OXFORD BROOKES UNIVERSITY

Maintaining the balance: persistent baculovirus infection in insect cells

Raquel Baptista Arinto Garcia

A thesis submitted in partial fulfilment of the requirements of Oxford Brookes University for the award of Doctor of Philosophy

March 2019

Abstract

Baculoviruses such as Autographa californica nucleopolyhedrovirus (AcMNPV) are insect-specific viruses that usually kill the host. However, many species harbour persistent, non-lethal covert baculovirus infections although these have rarely been observed *in vitro*. Hence, little is known about the mechanisms for establishment or maintenance of persistent infections. Re-infection with a homologous baculovirus (known as superinfection) has been shown to activate the persistent infection to an overt form. *In vitro* studies have shown that infected cells can be resistant to superinfection although the mechanisms involved are largely unknown.

In 2011, a serendipitous event led to the creation of a persistently infected *Trichoplusia ni* cell line (Hi5) with an AcMNPV *p10*-deletion mutant (AcUW1.*lacZ*). From this culture, a clonal cell line (C20) was established that has been maintained to this day. This thesis aimed to use C20 to further our knowledge of the fine balance between host and virus (designated AcC20) that promotes the establishment and maintenance of a persistent infection. The cell line also enabled study of the superinfection exclusion phenomenon.

C20 cells demonstrated remarkable viability and diversity of cell size/shape compared to the parental cells. The persistent AcC20 virus was maintained at low levels but budded virus (BV) titres varied considerably within and between different cultures. Presence of BV and the major membrane protein GP64 confirmed C20 harboured a persistent rather than latent infection. Analysis indicated that only a small proportion (<7%) of cells were producing BV whilst superinfection studies demonstrated that the majority of cells were resistant to a secondary infection. Microscopy studies indicated that superinfection may be blocked as early as the cell binding stage although virus uptake was also affected.

Whole genome analysis of AcC20 showed four major deletions or insertions with other mutations in 34 coding regions affecting eight essential genes and a homologous region. Transcriptome analysis showed that all virus genes were expressed in C20 cells although the regulated phases of gene expression observed in a typical lytic infection were absent. Host cell transcriptome analysis also revealed differences between the cell response to a lytic and a persistent infection. In particular, endocytic pit formation and host defence pathways appeared to be compromised. The former may help explain resistance to superinfection and the latter may help maintain the persistent infection.

Attempts to create further persistent cell lines proved difficult. However, using viruses with a mutation in *lef2*, in which BV production is severely reduced, overcame this problem. These persistent infections were also used to evaluate the possibility of continuous recombinant protein production by expressing *dsred* and *urokinase* under strong promoters. DsRed but not urokinase was produced over several cell passages.

This thesis demonstrated establishment and maintenance of persistent infections in cell culture. Further research is needed to understand this process and why these cells are resistant to superinfection.

Presentations

The work described in this thesis has been presented at international conferences and seminars in the form of posters and oral presentations.

Posters

Arinto-Garcia, R., Irons, S., Hughes, L., Hawes, C., King, L., Possee, R. D. (2017). Characterization of a persistent baculovirus infection established in an insect cell line. 50th International Congress of the Society of Invertebrate Pathology, San Diego, USA. **Microbiology Society travel award winner.**

Arinto-Garcia, R., Irons, S., Hughes, L., Hawes, C., King, L., Possee, R. D. (2018). Characterization of a persistent baculovirus infection established in an insect cell line. Symposium. Faculty of Health and Life Sciences Postgraduate Research Student Symposium, Oxford Brookes University, Oxford, UK.

Arinto-Garcia, R., Irons, S., Hughes, L., Hawes, C., King, L., Possee, R. D. (2018). Characterization of a persistent baculovirus infection established in an insect cell line. Oxford Brookes University Graduate College Research Student Exhibition, Oxford, UK.

Oral Presentations

Arinto-Garcia, R., Irons, S., Hawes, C., King, L., Possee, R. D. (2017). Characterization of a persistent baculovirus infection established in an insect cell line. Faculty of Health and Life Sciences Postgraduate Research Student Symposium, Oxford Brookes University, Oxford, UK.

Arinto-Garcia, R., Irons, S., Hughes, L., Leite, D., Hawes, C., King, L., Possee, R. D. (2018) Characterization of a persistent baculovirus infection established in an insect cell line. Department of Biological and Medical Science Seminar Series, Oxford Brookes University, Oxford, UK

Arinto-Garcia, R., Bannach, C., R., Bao, K., Fei, Z., Blissard, G. W., Hawes, C., King, L., Possee, R. D. (2018) Characterization of a persistent baculovirus infection established in an insect cell line. South Wilts Virus Group Meeting, Oxford Brookes University, Oxford, UK.

Arinto-Garcia, R., Bannach, C., Leite, D., Hawes, C., King, L., Possee, R. D. (2018). Whole genome analysis of a baculovirus isolated from a persistently infected insect cell line. 51th International Congress of the Society of Invertebrate Pathology, Gold Coast, Australia.

Acknowledgements

I left this section until very last because I could not believe I was actually finishing my PhD, but I am! I take with me not only new knowledge, but also life skills I would not have gained without embracing this challenge. I learnt that the bigger picture is made up of many small pieces; my experiments taught me to be resilient and to keep persisting, but also how to celebrate the small successes. But certainly, I could not have taken this journey alone.

Firstly I want to thank my supervisors, Robert Possee, Linda King and Chris Hawes, for offering me the chance to do this PhD, for all their patience and thoughtful advice, for the practical wisdom and virology lessons that brought me up to speed for this project (Bob), for offering new experimental points of view and careful corrections (Linda) and for teaching me valuable tricks to improve my imaging techniques (Chris).

I want to thank the amazing team I was lucky to be a part of, the Insect Virus Research Group and OET: Mine, Adam, Leo, Riyadh, Carina, Vicky, Sarah, Olga, Cassandra, Dani, Laura, Alice, Durva, Alex, Becky and Leonie. Thanks for never being afraid of being asked silly questions, for always being ready to "think with me" and to explain the "whys". I have learnt and enjoyed so much more by having you around!

My friends also deserve a huge thank you! "Old" friends from Lisbon, who even from a distance were always present and looking forward to seeing me: Ana Sofia, Sofia, Madalena MV, Madalena A, Duarte, Zé, Joana S, Joana L, LCS girls and all my childhood friends. Thank you for encouraging me to go on this adventure and for making me believe I could do it. "New" friends for always being there exactly when I needed. Lia and Bianca, who accompanied me on this PhD journey, for all the steak and gin or camomile evenings spent discussing hard-core science and cheap gossip. Aleks, thanks for keeping me assured that both the PhD and winter would eventually come to an end, and that summer would arrive, even when all I could see was clouds. Finally, to Girlguiding for encouraging girls to believe in themselves, to the game evening group for the dinner-time strategic challenges (great minds also play games!), and to all the other inspiring friends I have in Oxford.

Thank you to my family, to whom I am incredibly grateful. To my mum, who never turns off the phone for us and who is always demanding life-updates; to my dad and my brother, the quiet motivators who were always there for me; and to my grandparents, who had very little idea what I was really doing but always believed I could do it. To all my lovely family, I want to say a big thank you for encouraging me, for all of your kind words and support, and for the regular visits.

Finally, I want to thank Oxford Brookes University for the opportunity to carry out this PhD and for the generous financial support my project received.

- Obrigada -

Table of content

Figures

Tables

List of Abbreviations

a.a. Amino acids
AcMNPV Autographa californica multiple nucleopolyhedrovirus
ACS Acyl-CoA synthetase
ATCC American type culture collection
BCPI 5-bromo-4-chloro-3-indolyl-phosphate
BEVS Baculovirus expression vector system
BSA Bovine serum albumin
BV Budded virus
CLSM Confocal laser scanning microscopy
C_t Cycle threshold
DAPI 4',6-diamidino-2-phenylindole, fluorescent stain that binds the DNA
DGAT Diacylglycerol acyltransferase
DGE Differential gene expression
DMF Dimethylformamide
DNase I Deoxyribonuclease I
DPBS Dulbecco's PBS
d p.i. Days post-infection
dsRNA Double stranded RNA
erm ezrin radixin moesin
EGFP Enhanced green fluorescent protein
F-actin Filamentous actin
FBS Fetal bovine serum
FHV Flock House virus
G-actin Globular actin
GPAT Glycerol-3-phosphate
GP64 Glycoprotein 64 (*gp64* refers to the respective gene)
GV Granulovirus
Hi5 cells High five cell line: commercial form of *Trichoplusia ni* BTI-TN-5B1-4 cells, originated from the ovarian cells of the cabbage looper
h p.i. Hours post-infection
h p.s. Hours post-seeding
hr *Homologous region*
HRP Horse radish peroxidase
iap Inhibitor of apoptosis
ie1 AcMNPV immediate early gene 1
IMS Industrial methylated spirit
IVRG Insect virus research group at Oxford Brookes University
kbp Kilobase pair
lacZ β-galactosidase gene (from *E. coli*)
MOI Multiplicity of infection
miRNA MicroRNA
min. p.i. Minutes post-infection
mRNA Messenger ribonucleic acid

NBT	Nitro blue tetrazolium chloride
NP40	Nonidet™ P 40
NPC	Nuclear pore complex
NSF	N-ethylmaleimide sensitive fusion protein
NTC	No template control
OB	Occlusion body
ODV	Occlusion derived virus
ORF1629	Open-reading frame 1629
OXSR	Oxidative stress response gene
P10	AcMNPV P10 protein (*p10* refers to the respective gene)
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PI	Proteinase inhibitor
PIF	*Per os* infectivity factor
PIPES	Piperazine-N,N'-bis(2-ethanesulfonic acid) used as buffer
Polh	Polyhedrin protein (*polh* refers to the respective gene)
pfu	Plaque-forming unit
qPCR	Quantitative polymerase chain reaction
RFU	Relative fluorescence units
RNAi	RNA interference
RT	Room temperature
RT-PCR	Reverse transcription polymerase chain reaction
SCD	Stearoyl-CoA desaturase
SD	Standard deviation
SDS-PAGE	Sodium dodecyl sulphate polyacryamide gel electrophoresis
SEM	Scanning electron microscopy
SE	Standard error of the mean
SeMNPV	Spodoptera exigua multiple nucleopolyhedrovirus
Sf9	*Spodoptera frugiperda* 9, a clonal cell line derived from Sf21 cells
Sf21	Cell line derived from *Spodoptera frugiperda* ovarian cells
SfMNPV	Spodoptera frugiperda multiple nucleopolyhedrovirus
siRNA	Small interfering RNA
SNARE	Soluble N-ethylmaleimide-sensitive factor receptors
SNPV	Single nucleopolyhedrovirus
ssDNA	Single stranded DNA
T. ni	*Trichoplusia ni*
***T. ni* 368**	*Trichoplusia ni* 368 cell line
TREX	Transcriptional export complex
VLP	Virus-like particles
VP39	AcMNPV viral protein 39
WGS	Whole genome sequencing
WT	Wild-type

Chapter 1

Introduction

1.1 The Baculoviridae

Baculoviruses are insect-specific viruses that often lead to the death of the larval host. They have a double-stranded, circular, supercoiled DNA with genomes varying from 80 to over 180 kilobase pair (kbp), encoding between 90 and 180 genes (Rohrmann, 2013). The genome is enclosed within rod-shaped nucleocapsids that are packaged into two forms over the bi-phasic life cycle: budded virions (BVs) and occlusion bodies (OBs) (Summers and Volkman, 1976; Volkman *et al.*, 1976; Volkman and Summers, 1977; Danquah *et al.*, 2012). BV envelopes from this family derive from the plasma membrane and contain an envelope fusion glycoprotein, such as GP64 or F protein, at one end of the virion (reviewed by King *et al.*, 2012). Based on the predicted amino acid alignment of 29 baculovirus core genes common to all sequenced isolates, the family is further subdivided into four genera: *Alphabaculovirus*, *Betabaculovirus*, *Deltabaculovirus* and *Gammabaculovirus* (Jehle *et al.*, 2006). The first two genera encompass viruses isolated uniquely from the insect order Lepidoptera (King *et al.*, 2012). Currently 40 species of *Alphabaculovirus* are recognized and their OB contain either single (S) or multiple (M) nucleocapsids within a single viral envelope embedded in a crystalline protein matrix. Rod-shaped nucleocapsids are 30-60nm×250-300 nm in size and contain a single dsDNA molecule (King *et al.*, 2012). Autographa californica nucleopolyhedrovirus (AcMNPV) is the type member of this genus.

1.2 Baculoviruses and host interactions

The majority of baculovirus infections culminate with the death of the host (Washburn *et al.*, 2003b) since the physiological changes caused in the host's short life span can hinder their co-existence as persistent or latent infections (Slack and Arif, 2007). However, insects can activate a range of virus defence mechanisms to protect themselves from infection or to minimize its impact, therefore possibly allowing persistent infections to be established. These include the release of antimicrobial peptides, phagocytosis, cell apoptosis and sloughing of infected midgut cells among others (Washburn *et al.*, 2003a; Narayanan, 2004; Cheng *et al.*, 2006; Clem, 2015; Williams *et al.*, 2017).

Nevertheless, baculoviruses have created mechanisms to bypass these defences, such as *p35* apoptosis suppressor (in AcMNPV) and *inhibitor of apoptosis* (*iap*) (Clem, 1997; Ikeda *et al.*, 2013). Apoptosis is triggered by the expression of viral early genes. However, P35, synthesized during the early and late phases of infection, localizes in the cytoplasm where it binds to effector caspases and prevents apoptosis (Prikhod'ko and Miller, 1999; reviewed by Ikeda *et al.*, 2013).

Furthermore, *p35* and early genes essential for viral DNA replication are involved in triggering host cell translational arrest that occurs after AcMNPV infection (Schultz and Friesen, 2009; Ikeda *et al.*, 2013). During the late and very late phases of infection host transcription is suppressed in favour of AcMNPV transcription (Braunagel *et al.*, 1998). Surprisingly baculovirus genes with pro-apoptotic effects were also identified, possibly as a mechanism to supress the larvae immune response and lead to higher production of OB. SeMNPV-encoded protein tyrosine phosphatase 2 (PTP2) was shown to induce mild apoptosis in *Spodoptera frugiperda* 21 (Sf21 cells Han *et al.*, 2018a).

The baculovirus infection cycle finishes with cell lysis, liquification of the insect and consequent release of OB. This promotes the horizontal transmission of the virus to neighbouring insects. Furthermore, it has been shown that the baculovirus infection is able to manipulate host behaviour by inducing hyperactivity or climbing behaviours as a strategy to enhance dispersal and transmission to new insects (Vasconcelos *et al.*, 1996; Goulson, 1997; Hoover *et al.*, 2011; reviewed by van Houte *et al.*, 2013). Infected larvae climb to the top of the trees just before dying and releasing OB, a light dependent behaviour known as tree-top disease (Han *et al.*, 2018b). The *egt* encoding ecdysteroid UDP-glucosyl transferase (EGT) was the first identified viral gene related to the induction of this behaviour in Lymantria dispar multiple nucleopolyhedrovirus (LdMNPV)-infected insects (Hoover *et al.*, 2011). The same link was observed in SeMNPV-infected insects, where *egt* was thought to facilitate tree-top disease by prolonging the larval time to death (Han *et al.*, 2015). Even though AcMNPV also induces this disease in *T. ni* and *S. exigua* larvae, AcMNPV *egt* does not play a similar role (van Houte *et al.*, 2014; Ros *et al.*, 2015). However, AcMNPV protein tyrosine phosphatase (*ptp*) was shown to induce hyperactive behaviour in *S. exigua* larvae, via an independent mechanism from the tree-top disease (van Houte *et al.*, 2012, 2014).

Baculoviruses may also exist in the population of insects in a covert form (usually in a persistent infection, reviewed by Williams *et al.*, 2017). Persistent infection is defined as the co-existence of both virus and host in a synergy that ensures the virus is transmitted at a low fitness cost for the host cells. A low expression of most of the viral genes and the production of viral particles distinguishes it from a latent infection. This type of infection is non-lethal and can occur through all the stages of the insect's life (Goic and Saleh, 2012; Goic *et al.*, 2013; Jakubowska *et al.*, 2014; Lieberman *et al.*, 2016). Although only sporadically observed *in vitro*, it is commonly detected in natural populations of insects infected by the four genera of Baculoviridae but has been mainly studied in NPV infecting lepidoptera (*Alphabaculovirus* genus) (Williams *et al.*, 2017). This ability of baculovirus to remain in the environment has been utilized in the development of pest control plans using NPV as biopesticides (Cory and Bishop, 1997) (see 1.6).

The maintenance of a persistent infection enables vertical transmission to progeny insects. It has been suggested that insect pathogens adopt a mixed-mode of transmission as a survival strategy (Williams *et al.*, 2017). This way, the virus has an unique evolutionary advantage by being able to follow two transmission pathways according to changes in the expected host survival, therefore optimizing the fitness obtained from each host (Fig. 1.1). This will permit the virus to persist under a wider range of ecological conditions and at a higher prevalence than by adopting a strict single-mode transmission (Ebert, 2013; Williams *et al.*, 2017).

Covert infections facilitate virus persistence in the environment when opportunities for horizontal transmission are limited, as during periods of low insect population densities, and enable virus dispersion over long distances by host migration (Vilaplana *et al.*, 2010; Fig. 1.1). In this state, the virus remains fully competent in the host cell and can be triggered to produce a lethal infection at any time (Burden *et al.*, 2006). For example, if host survival appears to be threatened, the viral opportunities for survival are in favour of horizontal transmission, so the infection can be activated to a lytic mode with production of OB (Williams *et al.*, 2017); Fig. 1.1). Possible threats to the insect population that have been reported as potential activators of overt disease include environmental and physiological stressors such as overcrowding, temperature and humidity changes, ingestion of toxic compounds, parasitism or changes in nutrient availability (Hughes *et al.*, 1997; Burden *et al.*, 2006; Murillo *et al.*, 2011; reviewed by

Williams *et al.*, 2017). This way, the virus can be re-activated to produce disease upon increase of the host density, above a threshold that allows horizontal transmission, killing the host and releasing progeny by horizontal transmission (Cooper *et al.*, 2003). This often triggers outbreaks in dense lepidopteran populations in natural habitats, hence rapidly reducing the host population to below the threshold density (Myers and Cory, 2016). Studies indicate that in this way, baculovirus infections are able to regulate the population density of their hosts (Anderson and May, 1980).

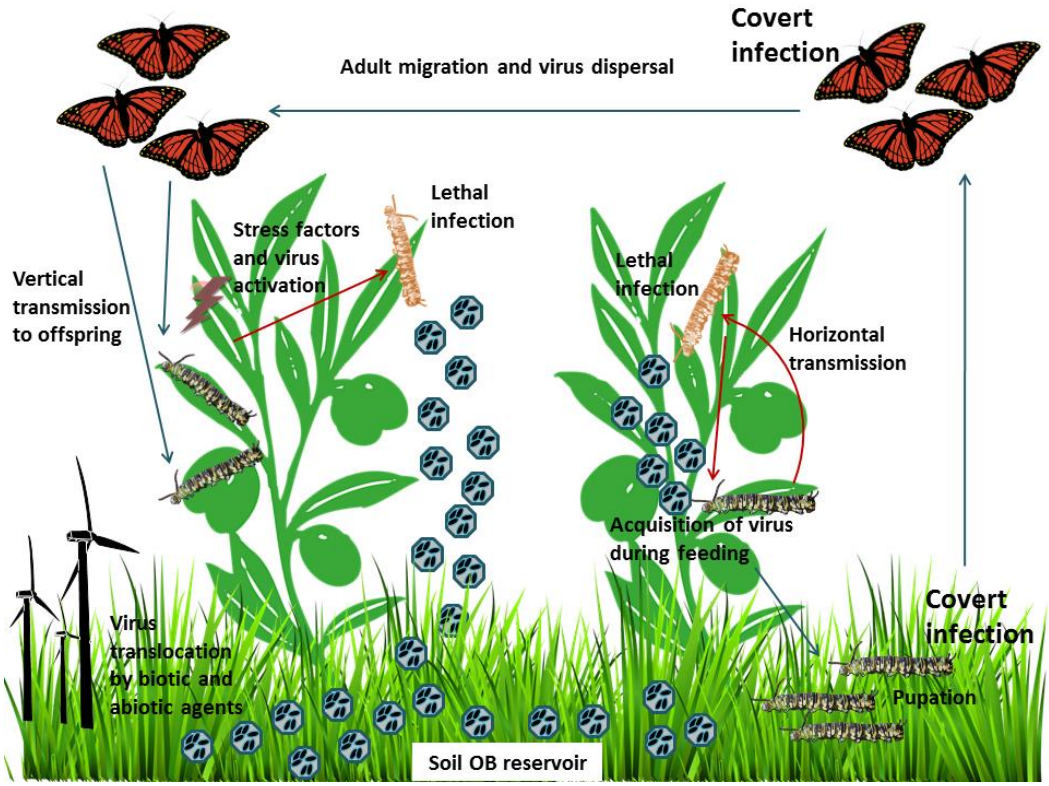


Figure 1.1 Baculovirus transmission routes and forms of infection
Baculoviruses can persist in the insect population in a covert form. However, stress factors can reactivate the virus to a productive infection releasing OB into the environment, which accumulate in the soil and are translocated by biotic and abiotic agents. When OB are ingested by larvae, another cycle of productive infection can be triggered (horizontal transmission cycle) or on the contrary the insects can survive the infection and continue to develop until the adult phase with a covert infection. This allows the virus to migrate and to be transmitted vertically to their offspring. Figure based on Williams *et al.*, 2017.

Finally, a challenge by a second pathogen, known as superinfection (see 1.7) has been shown by several studies to be a major trigger of a lytic infection (Burden *et al.*, 2003, 2006). This occurs often in nature and regularly involves inoculation with a heterologous virus, commonly isolated from a different host than the one under study (Hughes *et al.*, 1993; Cooper *et al.*, 2003). In this case, infection by a second pathogen constitutes a serious threat to host survival, hence, the horizontal transmission strategy is activated. Furthermore, the virus ability to persist indicates it was able to evade but also to achieve

a balance to some extent with the host immune system (Oldstone, 2006). This balance might be altered by the infection of a second pathogen which disrupts the immune system and possibly enables the first virus to reactivate to a lytic infection.

Research indicates that the viral strategy for maintaining covert infections might include differential infection of host cell types, modulation of gene expression and avoidance of host immune system clearance (Kane and Golovkina, 2010). Furthermore, the molecular basis for the maintenance of the covert infections might reside in the regulation of host and virus interactions through the action of microRNAs (miRNA) – small, non-coding, hairpins RNAs that are produced by both to regulate transcription and translation (reviewed by Asgari, 2015). An example is the miRNA *bantam*, which plays an important role in insect growth and in baculovirus-insect interaction (Shi *et al.*, 2016). Feeding a *bantam* inhibitor to larvae of *S. litura* and *S. exigua* resulted in abnormal larval growth and a decreased pupation rate (Shi *et al.*, 2016). Moreover, when the inhibitor was administrated together with AcMNPV infection in the first host, larvae died 3.5 days sooner than the control (Shi *et al.*, 2016). Host RNAi genes have also been implicated in preventing the re-activation of non-baculovirus persistent/latent infections by Glossina pallidipes salivary gland hypertrophy virus (GpSGHV) in *Glossina* (Meki *et al.*, 2018). Bmnpv-miR-3, micro-RNA encoded by Bombyx mori nucleopolyhedrovirus (BmNPV) was shown to be expressed during early stages of infection *in vivo*, and to regulate expression of the DNA binding protein P6.9 and other late genes (Singh *et al.*, 2014). These examples show the importance of microRNA as a host defence mechanism and their role in virus-host interactions (Singh *et al.*, 2014; Shi *et al.*, 2016).

1.3 Baculovirus replication-cycle in lytic infection

The AcMNPV OBs are transmitted between insects via the oral route and initiate the replication cycle by infecting larval midgut cells (reviewed by Rohrmann, 2013). The OBs dissolve in the alkaline environment of the larval midgut and release the occlusion derived virus (ODV) from the polyhedrin matrix (Williams *et al.*, 2017). The ODV fuse to the membrane of the midgut cells initiating the primary infection (Federici and Hice, 1997; Fig. 1.2]. Actin polymerization enables the efficient transport of the nucleocapsid towards the nucleus where it enters through the nuclear pores (Granados and Lawler, 1981). Here, the DNA is replicated and new nucleocapsids assembled (Volkman, 1988;

Danquah *et al.*, 2012; Fig. 1.2). The new nucleocapsids move through the cytoplasm and bud from GP64-enriched zones on the plasma membrane, which are responsible for initiating the secondary or systemic infection in the organism (Herniou *et al.*, 2003; Washburn *et al.*, 2003b; Fig. 1.2]. Viral progeny are first observed around 8 hours post-infection (h p.i.) and the complete replication cycle is completed by 24h p.i. (Granados and Lawler, 1981).

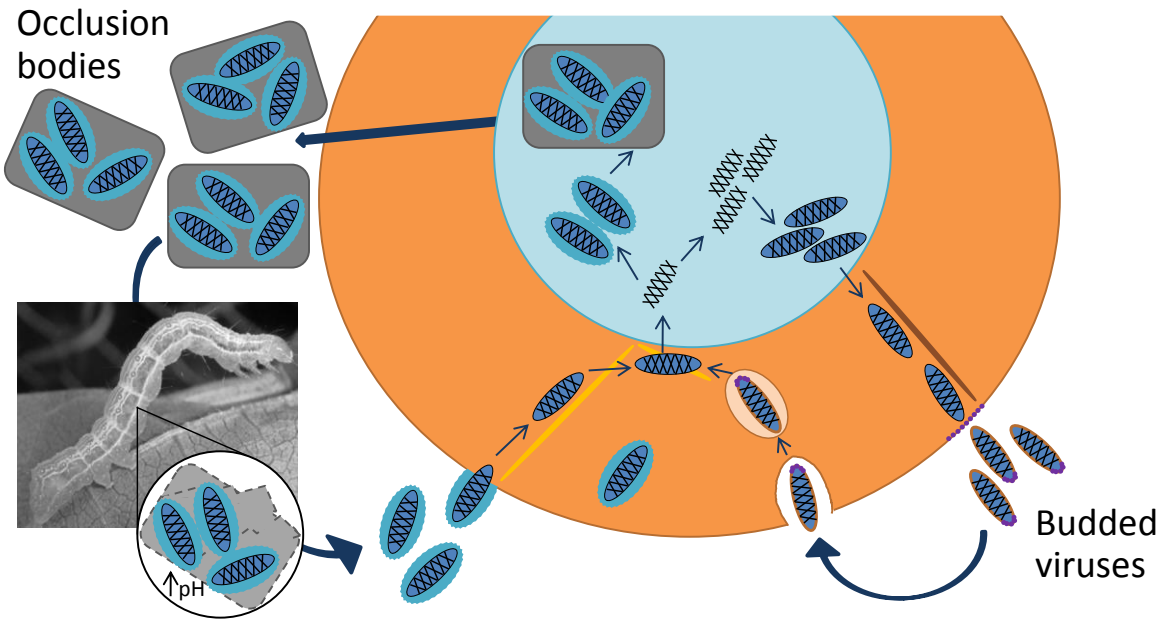


Figure 1.2 Baculovirus replication cycle
The budded viruses (BV) enter the cells via the endocytic pathway and are transported to the nucleus. Here gene expression is initiated and in the late phase new BV are generated and transported to the cell membrane to bud out in GP64 rich regions. In the very late phase, newly produced virions are enclosed into a polyhedrin matrix (occlusion body) and are released into the environment following cell lysis.

1.3.1 Entry of AcMNPV into cells

After ODV release from the OB, the envelope interacts with the microvilli of columnar cells (Granados and Lawler, 1981). A specific protein receptor on the ODV membrane is responsible for fusion to either the host cell brush border membrane vesicles (Granados and Lawler, 1981; Horton and Burand, 1993) or, alternatively, directly to the membrane through a non-endocytic pathway. The latter route has been shown to be less efficient but occurs over a wide pH range; although highly favoured for alkaline conditions, as found in the insect midgut (Horton and Burand, 1993). For BV entry into the cell, clathrin-dependent endocytosis is assumed to be the major pathway (Volkman and Goldsmith, 1985; Long *et al.*, 2006; Dong *et al.*, 2010). The GP64 glycoprotein receptor is responsible for the fusion of the virus envelope to the endosome membrane during

endocytosis (Blissard and Wenz, 1992). This process is pH dependent requiring pH5 or lower. Hence, the fusogenicity of GP64 is activated by the low pH of the acidified late endosomes (Blissard and Wenz, 1992; Leikina *et al.*, 1992; van Loo *et al.*, 2001) .

1.3.2 The role of cytoskeleton in virus infection

Actin and microtubules are major cytoskeletal components in all eukaryotic cells. BV enters cells by endocytosis (see 1.3.1) and its release into the cytoplasm induces the polymerization of thick actin cables (Charlton and Volkman, 1991). Globular actin (G-actin) polymerization into filamentous actin (F-actin) is regulated by the cellular Arp2/3 complex (Pollard, 2007), stimulated by the viral P78/83 nucleocapsid protein in infection (Ohkawa *et al.*, 2010). This is crucial for efficient translocation of nucleocapsids throughout the cytoplasm toward the nucleus and through the nuclear pores (Charlton and Volkman, 1993; Ohkawa *et al.*, 2010). However, these cables are transient structures (Charlton and Volkman, 1993). Actin is also critical for baculovirus progeny production (Volkman *et al.*, 1992; Gandhi *et al.*, 2012). Expression of *ie1*, *pe38*, *ac004*, *ac152*, *he65* and *ac102* early genes is necessary for nuclear localization of G-actin; ac102 is a key player in baculovirus manipulation of actin (Ohkawa *et al.*, 2002; Gandhi *et al.*, 2012). Nuclear F-actin polymerization is similarly P78/83 (*orf1629*)- and Arp2/3-dependent (Lanier and Volkman, 1998; Goley *et al.*, 2006). The lack of this event disrupts BV production (Volkman and Kasman, 2000). Viral DNA synthesis occurs at a normal rate but genomes are not packaged since viral capsids are malformed, and the virogenic stroma loses its shape (Volkman, 2007). Studies inactivating F-actin with cytochalasin D or latrunculin A demonstrated its essential role for nucleocapsid morphogenesis appears to be a characteristic common among viruses in this genus, with lepidopteran hosts (Ohkawa and Volkman, 1999; Volkman and Kasman, 2000). F-actin also acts as a regulator of chromatin remodelling by interacting with P6.9 (Volkman, 2007). Finally, F-actin has shown to be essential for viral transit to nuclear periphery and disruption of nuclear envelope integrity that enables nuclear egress to the cytoplasm (Ohkawa and Welch, 2018).

For BV production, nascent nucleocapsids need to travel to the cytoplasmic membrane where they will bud from the cell. Microtubules are highly dynamic polymers of heterodimers of α- and β-tubulin. Co-localization and interaction of EXON0 (encoded by

orf141) with β-tubulin is essential for the efficient transport of progeny nucleocapsids from the nucleus to the cytoplasm (Volkman and Zaal, 1990; Fang *et al.*, 2007, 2009a). In the absence of EXON0 levels of BV production are extremely reduced, hampering infection to the neighbouring cells (Dai *et al.*, 2004). However, actin-based motility was recently shown to be essential for movement to the plasma membrane (Ohkawa and Welch, 2018).

1.4 Baculovirus gene expression

Gene expression is divided into four tightly regulated phases: immediate early, early, late and very late (Fig. 1.3). This initiates following virus nucleocapsid uncoating in the nucleus and release of DNA. Four genes have been described as part of the immediate early phase, *ie0*, *ie1*, *ie2* and *pe38* that encode for transcriptional regulators for early viral genes. These coordinate the activation of the transcriptional cascade. Part of the IE1 population is translated from a larger spliced mRNA and it is called IE0 (Chisholm and Henner, 1988). This is the only major splicing event described for baculoviruses (Rohrmann, 2013). Although IE1 and IE0 can function independently, they are not equivalent and to achieve a wild-type (WT) infection both are required in the right ratio (Stewart *et al.*, 2005). The ability of IE1 to transactivate transient transcription is stimulated by sequence specific DNA binding to AcMNPV *homologous regions* (*hrs*) that work as origins of DNA replication in the absence of insect cell factors (Guarino *et al.*, 1986; Pearson *et al.*, 1992; Choi and Guarino, 1995; Rodems *et al.*, 1997). In AcMNPV, six *hrs* have been described and have been shown to comprise repeated sequences (Pearson *et al.*, 1992). These act as enhancers that *cis* activate the transcription of regulatory genes and are interspersed throughout the AcMNPV genome, at least 3kbp from *ie1* (Guarino *et al.*, 1986). The *ie1* is also transcribed through the late phase (Kovacs *et al.*, 1991) and it is also involved in the negative regulation of immediate early gene expression (Carson *et al.*, 1991; Leisy *et al.*, 1997). The *ie2* encodes for the transcriptional regulator IE2, which increases IE1-mediated transactivation of the early gene *pp31* (also called *39k*) (Carson *et al.*, 1988, 1991; Yoo and Guarino, 1994). The *pp31* encodes a phosphorylated DNA binding protein that locates to the nucleus associating with the virogenic stroma. Although not essential for viral replication, its absence

resulted in lower viral titres and reduced transcription from early and late genes (*p6.9*, *orf97*, *orf60*, *orf98*, *orf102* and *chitinase*) (Yamagishi *et al.*, 2007).

The early phases of gene expression are mediated by the host cell RNA polymerase II while later genes are transcribed by the virus encoded α-amanitin resistant RNA polymerase (Morris *et al.*, 1994; van Munster *et al.*, 2006). The viral RNA polymerase recognises unique promoter motifs (TAAG) from which transcription of late and very late genes initiates (reviewed by Rohrmann, 2013). This enzyme is composed of four protein subunits, Lef-8, Lef-4, Lef-9, and p47 (Guarino *et al.*, 1998)

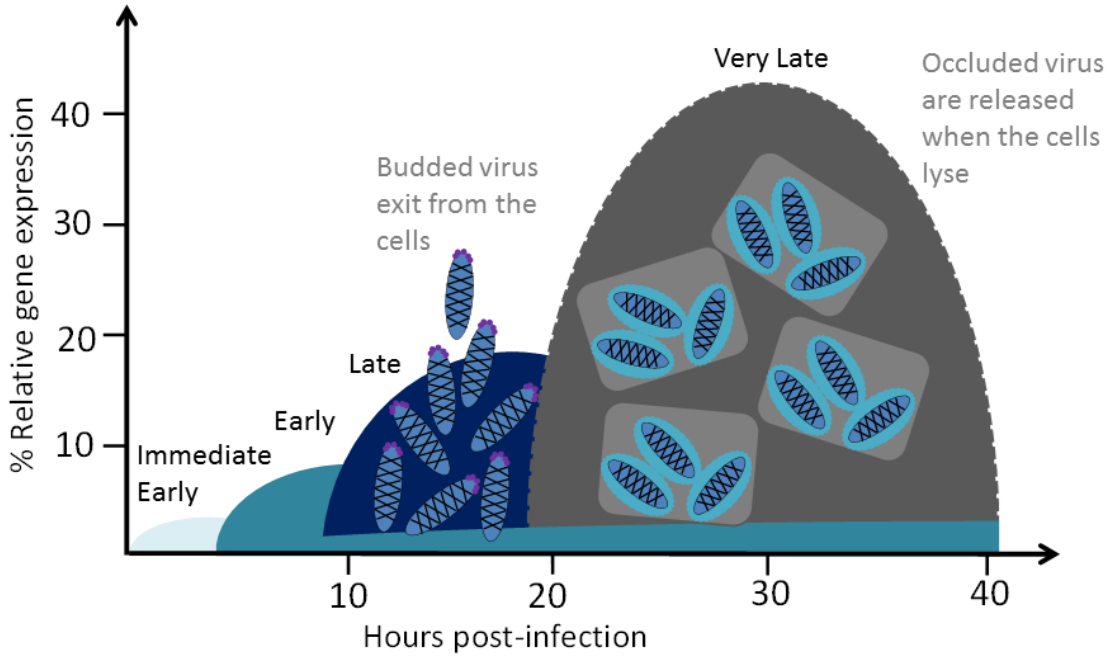


Figure 1.3: AcMNPV gene expression model
The accepted model for relative gene expression over time in a lytic infection defines four tightly regulated infection phases, immediate early, early, late and very late. The establishment of the infection occurs during the early phases when viral genes are transcribed. In the late phases of infection occurs the preparation of the new infectious particles either BV (late phase) or ODV (very late phase). The very late phase culminates with the release of the polyhedra enclosing ODV upon cell lysis.

The start of DNA replication can occur as early as 6h p.i. and up to 12h p.i. and marks the progression from early to late stages of infection (Blissard, 1996). DNA replication initiates from the six *hrs* although it has been suggested these might not be the only points of origin (Pearson *et al.*, 1992; Habib and Hasnain, 2000). Nine genes were identified as being involved in DNA replication, six of them essential (*helicase*, *dna polymerase*, *ie1*, *lef1*, *lef2*, and *lef3*) and three genes with a stimulatory effect (*p35*, *ie2*, and *pe38*) (Kool *et al.*, 1994).

Assembly of nucleocapsids takes place in the virogenic stroma inside the nucleus, in the late phase of infection (Fig. 1.2). In this phase, some early genes are suppressed, whereas others contribute to regulate the expression of late genes (Todd *et al.*, 1995, 1996; van Munster *et al.*, 2006). This phase is focused on the production of BV and therefore capsid protein genes are highly expressed. Mass spectrometry techniques have shown GP64, VP39 and P6.9 are the three most abundant proteins in the AcMNPV BV (Wang *et al.*, 2010b). VP39 is the major capsid protein and it is transcribed maximally at 12 to 24h p.i. It co-localizes with ME53 and GP64 at the cell plasma membrane (de Jong *et al.*, 2011). P6.9 is a protamine-like protein that it is phosphorylated immediately upon synthesis, it localizes to the nuclear matrix and is dephosphorylated when binding non-sequence-specifically to DNA (Wilson *et al.*, 1987; Dong *et al.*, 2010; Rohrmann, 2013). Although not essential for viral DNA replication, it is essential for infectious virus production (Wang *et al.*, 2010a). Its association with DNA contributes to transcription regulation (Peng *et al.*, 2012). The presence of DNA free of proteins can facilitate the activation of late promoters but once all DNA is densely packaged with P6.9, the transcription of late genes is inhibited (Balhorn, 2007; Peng *et al.*, 2012; Rohrmann, 2013).

The transition to the very late phase (from around 18h p.i.) is marked by the hyper expression of *p10* and *polyhedrin* (*polh*) viral genes. At 24h p.i., these generate the most abundant viral mRNA (Chen *et al.*, 2013). In this phase most of the recently produced virions become occluded and no longer bud out of the membrane. Polyhedrin is a very robust protein that offers protection to the ODV while in the environment (Duncan *et al.*, 1983; Ji *et al.*, 2010). The role of P10 in infection is not completely clear but it has been suggested that it forms a cage-like structure around nuclei that aids in nuclear lysis (Graves, 2016). The fact that these proteins are produced in large quantities (Fig. 1.3) but that replication *in vitro* can occur without them, contributes to the increasing interest in exploring this phase of infection for recombinant protein production using the baculovirus expression vector system (BEVS) (Hitchman *et al.*, 2009).

1.5 Baculovirus applications

Historically, baculoviruses have been extensively used as biopesticides while more recently, applications as an expression system have been extended and used for the

generation of virus-like particles (VLPs) to produce vaccines (reviewed by Haase *et al.*, 2015; Kost and Kemp, 2016). Chemical pesticide resistance and the pressure to reduce pesticide residues due to the impact on the environment and health has limited the options to control some globally important crop pests (Whalon *et al.*, 2008; Haase *et al.*, 2015). Baculoviruses, as natural pathogens of some of these pests, have been used as biocontrol agents for over 60 years (Cory and Bishop, 1997). Baculoviruses can be used in a range of conditions from forests to greenhouses (Cory and Bishop, 1997; Moscardi, 1999). However, only lepidopteran-specific NPV from *Alphabaculovirus* spp. and granuloviruses from *betabaculovirus* spp. have been commercially developed for this purpose (reviewed by Lacey *et al.*, 2015). Although the production process is time-consuming and labour-intensive, the robustness of the OB enables easy production and storage, consequently justifying its extensive use (Cory and Bishop, 1997; Lacey *et al.*, 2015). These biopesticides are highly pathogenic and adapted to their host and therefore have a narrow and specific host range (Cory and Bishop, 1997; Ohkawa and Volkman, 1999; Haase *et al.*, 2015). Therefore they are safe to be integrated in wide control plans as they do not infect other insects in the same environment like beneficial invertebrates (Cory and Bishop, 1997; Lacey *et al.*, 2015). However, some limitations are still to be overcome. A practical *in vitro* mass production system has yet to be optimized. The slow host killing rate and the very high specificity are still restricting the market use of baculoviruses as biopesticides (Lacey *et al.*, 2015).

The use of BEVS is an application of increasing relevance. It is commonly used for the synthesis of recombinant proteins for structural studies but also for the development of vaccines (reviewed by Kost *et al.*, 2005; Kost and Kemp, 2016). The system was first demonstrated in 1983 to express human beta interferon in insect cells (Smith *et al.*, 1983). The heterologous genes are inserted under the control of highly expressed very late promoters, including *polh* and *p10* of AcMNPV and Bombyx mori nucleopolyhedrovirus (BmNPV). The system has the ability to produce very high yields of foreign protein in insect cells. The expression of foreign genes in insect cells using BEVS has many advantages compared to prokaryotic systems, as it enables: a) post-transcriptional modifications, like folding, oligomerization, glycosylation, phosphorylation and many others, similar or identical to the ones produced by mammalian cells (reviewed by Assenberg *et al.*, 2013; Liu *et al.*, 2013); b) expressing

multiple proteins simultaneously in a single infection to obtain multimeric processed proteins. This way this system can be used for the generation of VLPs like vaccine-subunits (Kemp *et al.*, 2012; Wang and Roden, 2013; Kost and Kemp, 2016).

Since 2007, three recombinant vaccines produced in BEVS have been approved for use in humans and many others for veterinary use (reviewed by van Oers *et al.*, 2015). Cervarix® is prophylactic against human papillomavirus (HPV)-16/18, the two most common high-risk HPV types, and is approved both in the European Union and USA. In this system, the major surface protein L1 is produced and self-assembles into a VLP in Hi5 (*T. ni*) cells (Schiller and Lowy, 2001). This success was followed by the approval of Flublok® against seasonal influenza virus. In this case, BEVS replaced the traditional and long process of vaccine generation in chicken eggs (Cox *et al.*, 2008). The influenza glycoprotein HA is produced under the *polh* promoter in *Spodoptera frugiperda* 9 (Sf9) cells. Finally, therapeutic vaccine Sipuleucel-T (Dendreon®) has been approved in the European Union and USA for treatment of prostate cancer (Kantoff *et al.*, 2010; Cheever and Higano, 2011).

Even though baculoviruses are host specific, GP64 enables a broad entry tropism (Kataoka *et al.*, 2012). They are capable of entering mammalian cells via transduction and expressing foreign genes under the control of mammalian promoters without virus replication occurring (Hofmann *et al.*, 1995; Boyce and Bucher, 1996; Miyamura *et al.*, 1997; Tani *et al.*, 2003). Therefore, BEVS have also been used for gene delivery into mammalian cells with the advantage of low toxicity in comparison to mammalian virus-derived vectors. Examples of its applications in gene therapy include cancer therapy for suppression of tumour development, toxin vector for elimination of malignant glioma cells in the brain, herpes simplex and glioblastoma treatment (reviewed by Ono *et al.*, 2018).

Currently BEVS production is based on a productive infection with release of the desired recombinant proteins after cell lysis. This process of infecting, harvesting cells and purifying the protein is repeated in cycles and it is therefore time-consuming and labour-intensive. An option to facilitate this process would be the production of recombinant proteins continuously using a persistently infected cell line, which has yet to be explored.

1.6 Covert infections in insect cell lines

1.6.1 RNA genome viruses

Covert infection can assume the form of a persistent or a latent infection, both characterized by the co-existence of virus and host. However, while in a persistent infection viral genes are expressed and virus particles are released, albeit possibly at a low level, in a latent infection the viral genome is retained without being expressed until the virus is reactivated (Speck and Ganem, 2010; Lieberman, 2016). Due to the reduced or undetectable signs of infection, these conditions were usually missed in insects and cell cultures. Nonetheless, genetic techniques have revealed the presence of latent infections in commercial cell lines (Li *et al.*, 2007; Merten, 2007; Ma *et al.*, 2014).

A homologue of Flock House virus (FHV), an *Alphanodavirus*, has been identified as a latent infection in Hi5 cells (Li *et al.*, 2007; Merten, 2007). The nodaviridae family comprise small, non-enveloped viruses containing two molecules of a single-stranded, positive-sense RNA genome, co-packaged in spherical virus particles (Ball and Johnson, 1998). The family is divided into two genera, *Alphanodavirus* and *Betanodavirus*, based on their host, insects or fishes respectively (Ball and Johnson, 1998; Yong *et al.*, 2017). Although the infection of this FHV *Alphanodavirus* in *T. ni* cells is latent, the production of virus particles can be reactivated when the cells are challenged with baculovirus (Li *et al.*, 2007). However some *T. ni*-derived cell lines have been reported as nodavirus-free (Zhang and Thiem, 2010).

On the other hand, using degenerate PCR assays and next-generation sequencing, a novel RNA virus was identified in Sf9 cells from the Mononegavirales order (Ma *et al.*, 2014). This was called *Sf-rhabovirus* (SfRV) and phylogenetic analysis showed it has branched with Taastrup virus (Ma *et al.*, 2014). Sequences from SfRV were also found in the Sf21 parental cell line. However, it was shown that some earlier American type culture collection (ATCC) lots of Sf9 cells were free of this infection (Hashimoto *et al.*, 2017).

1.6.2 Baculoviruses

Sequencing amplified regions of *polh* and *ie1* genes indicated that many insects in natural populations harbour persistent virus infections and multiple colonization of the same host are possible (Burden *et al.*, 2003, 2006; Jakubowska *et al.*, 2014). However, persistent baculovirus infections have only rarely been observed in laboratory cultures (Hughes *et al.*, 1993; Lee *et al.*, 1998; Weng *et al.*, 2009).

In the early 1980's two independent studies described the establishment of a persistent baculovirus infection in Sf21 cells infected with S. frugiperda nucleopolyhedrovirus (SfMNPV) (McIntosh and Ignoffo, 1981; Crawford and Sheehan, 1983). An initial characterisation of this persistent infection described the cells as morphologically different, able to grow to a higher density but with a longer doubling time than the parental cell line. The persistent virus produced a BV titre 100x lower than the one produced by SfMNPV in Sf21 cells but it was still infectious; polyhedra production was severely reduced (McIntosh and Ignoffo, 1981).

Crawford and Sheehan (1983) described the process of establishing a persistent infection in Sf21 cells in three distinct phases. The early phase was identified by high levels of cell infection and death, which changed into a second phase where cell infection levels decreased. Finally, in the third phase, less than one percent of the cells were infected in any culture passage (Crawford and Sheehan, 1983).

Other persistent infections in insect cells cultures were described as being present at a very low level (Hughes *et al.*, 1993, 1997). A characteristic persistent infection was also thought to have been achieved more recently with Spodoptera exigua MNPV (SeMNPV), however, the isolation of a clonal cell line suggested that it was more likely a latent infection since the production of viral progeny was not observed. In the clonal cell line, cells harboured incomplete SeMNPV genomes so only a small portion of SeMNPV-specific DNA and transcripts were present in the cells (Weng *et al.*, 2009).

Deletion or mutation of the anti-apoptotic *p35* in AcMNPV causes infected Sf9 cells to undergo apoptosis (Clem *et al.*, 1991; Kelly *et al.*, 2006). However, it was described to also allow the cloning of survival cells harbouring a persistent infection, shown by detecting viral early gene expression (Lee *et al.*, 1998). Continuous release of viral

progeny was observed during passaging of cells. This artificially established persistent infection was shown to be reversed as soon as *p35* was reintroduced by stable transfection or insertion of the Cydia pomonella granulosis virus *iap* gene (Lee *et al.*, 1998).

Several studies focusing on host-virus interactions *in vitro* have highlighted the role of miRNAs and suggested they could be one of the factors in favour of persistent infection maintenance. By producing their own miRNA, baculovirus can target host and viral genes to facilitate their replication and escape the early immediate immune response (Hussain and Asgari, 2014; Singh *et al.*, 2014). It was demonstrated that upon inhibition of the host miRNA *bantam* in Sf9 cells, expression of *lef8*, *gp41* and *p10* viral genes increased. This generated an increase in viral DNA replication, although virus yields were not affected (Shi et al., 2016). Equally, miRNAs are likely capable of autoregulating replication and extending the temporal window for DNA replication, maximizing BV production while maintaining uncompromised host survival (Singh *et al.*, 2014). An example is AcMNPV miRNA-1, that was shown to suppress *ac94*, facilitating the switch from BV to ODV production (Zhu *et al.*, 2013). This is helpful considering that AcMNPV has a major requirement for the host machinery throughout its replication cycle.

1.7 Superinfection exclusion

Infecting the host with a second virus is one of the mechanisms suggested for re-activation of covert infections (Burden *et al.*, 2003, 2006). However, in some instances, an established virus infection has the capacity of interfering and preventing an infection by a second identical or closely related virus (Weng *et al.*, 2009; Beperet *et al.*, 2014). This phenomenon is called superinfection exclusion and, in lytic AcMNPV infections, it coincides with the end of the early phase and release of the first BV in the late phase of infection (12 to 15 h p.i.) (Beperet *et al.*, 2014). After this period, subsequent infection by another AcMNPV or a closely related virus (e.g. SfMNPV) is not possible (Beperet *et al.*, 2014). Similarly, this phenomenon was studied with SeMNPV, a group II MNPV, which was carried in a cell line in a latent-like infection. However, in this case, resistance to homologous but not heterologous superinfection was found (AcMNPV) (Weng *et al.*, 2009). Very little is known about the superinfection exclusion phenomenon and it is still

unknown at what stage the infection is blocked for the second virus or the mechanism involved.

In vivo, superinfection was indicated as a mechanism of re-activation of a persistent infection to an overt form. Nonetheless, very few studies have focused on this problem *in vitro*. In one of the very first descriptions of the phenomenon, the persistently infected cell line seemed to be resistant to both homologous and heterologous baculovirus (McIntosh and Ignoffo, 1981). The same was demonstrated with a *p35* deletion viral persistent infection, where the cells were resistant to superinfection with both the wild-type and *p35* mutant viruses (Lee *et al.*, 1998). Neither of these studies mentioned that re-activation might have occurred.

The complexity of the superinfection exclusion mechanism has been examined for several years in different virus families and has been described for human, animal and plant viruses (Adams and Brown, 1985; Lecoq, 1991; Karpf *et al.*, 1997). A wide range of mechanisms have been indicated including reduced binding of virus to cells, competition for cell receptors or intracellular host factors (Adams and Brown, 1985; Nasar *et al.*, 2015), entry or uncoating blockage of the superinfection virus (Singh *et al.*, 1997), trans-acting proteases and other specific proteins interactions (Karpf *et al.*, 1997; Folimonova, 2012). Furthermore, it has been suggested that the exclusion mechanisms might be differently regulated at the whole-organism or cellular level (Bergua *et al.*, 2014).

1.8 Serendipitous establishment of a cell line with a persistent baculovirus infection

The establishment of a cell line, named C20, that is persistently infected with baculovirus occurred five years before the start of this thesis (R. Possee, personal communication). Cultures of Hi5 cells were infected with AcUW1.*lacZ*, to study virus-infected cell lysis. In this virus, the first 152 nucleotides of the *p10* coding region were deleted and replaced with *lacZ* sequences (Weyer *et al.*, 1990). In the absence of P10, nuclei fail to lyse after the virus replication cycle is complete (Graves, 2016).

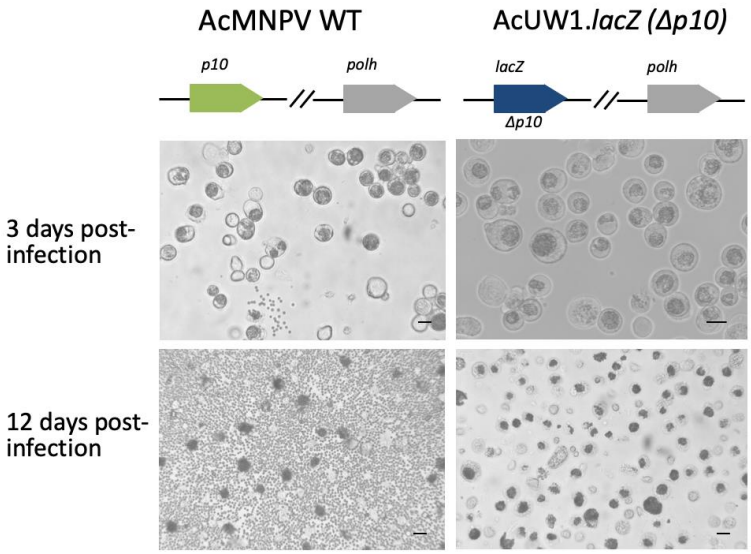


Figure 1.4: Bright field microscopy of AcMNPV- and AcUW1.*lacZ*-infected Hi5 cells.
Cells were infected with AcMNPV and AcUW1.*lacZ*. A schematic representation of the *p10* locus for the two viruses is shown. Pictures were taken using a bright field microscope at 3 days and 12 days post-infection. Scale bars 20μm.

Figure 1.4 shows a culture of Hi5 cells infected with AcMNPV in comparison to AcUW1.*lacZ* at three and 12 days post-infection (d p.i.). At 3d p.i, virus-infected cells appear similar, with polyhedra accumulating in the expanded nuclei ready to be released when the cells die. After 12 days, in AcMNPV infection, the cells had died and nuclear and cell membrane lysed, releasing polyhedra into the medium. However, in AcUW1.*lacZ* infection, although most, if not all, the cells have died and lost their cytoplasm, the nuclei remained intact, and contained large numbers of polyhedra that were not released into the cell medium (Fig. 1.4).

Instead, when generating the persistent virus infection described in this chapter, a few days after infection of Hi5 cells with AcUW1.*lacZ*, it was noticed that there remained clusters of live cells apparently uninfected. These clusters of cells were left to amplify and eventually were passaged when the monolayer became confluent. Figure 1.5 shows passage one and five of this early persistently infected culture. By passage 42, as the cells had been surviving with the virus for a long time, clones were isolated by serial dilutions. Several isolated clones were generated at this stage and all were shown to contain virus (R. Possee, personal communication). Clone 20 was selected and used to start the C20 cell line (Fig. 1.5), which has been grown continuously for eight years now and had reached passage 170 at the time this thesis was submitted.

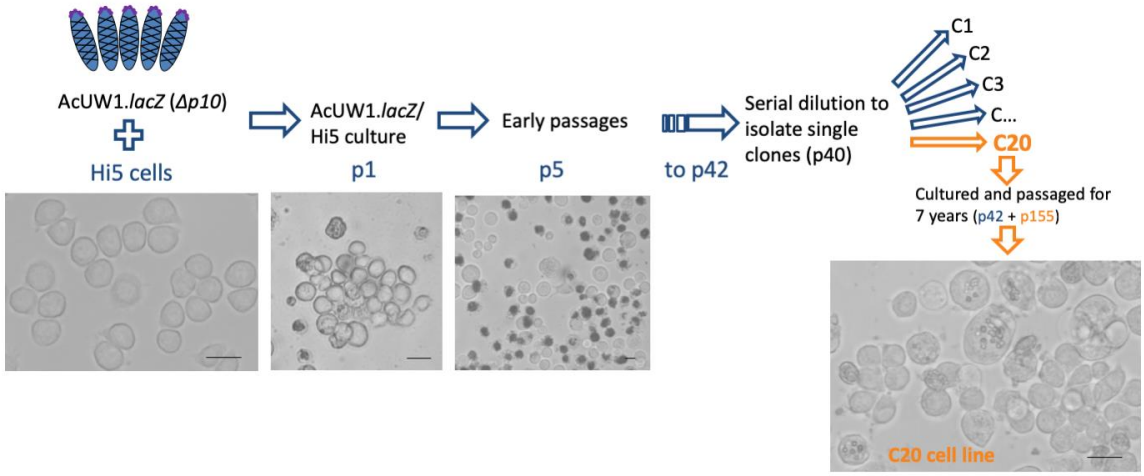


Figure 1.5: Generation of the C20 cell line
Hi5 cells were infected with AcUW1.*lacZ* and since clusters of live cells remained after a few days, they were left to grow until confluent. Subcultures were created up to passage 42, when serial dilution was performed in a multi-well plate to isolate single clones. Of the generated clones, C20 was chosen to be cultured further. Early passages of AcUW1.*lacZ* infected-Hi5 cells (p1 and p5) and passage 155 of C20 cell line is shown. Scale bars 20μm.

1.9 Project aims

Although persistent baculovirus infection is common in insects, its covert nature hinders its characterization. Consequently, very little is known about the mechanisms behind the delicate balance between virus and host.

The present study focused on a persistent infected cell line (C20), generated in our lab eight years ago and maintained in culture since then. It constitutes a valuable tool to study different aspects of persistent infection, including cellular characteristics, virus maintenance and genetic evolution. Furthermore, it provides an opportunity to study the phenomenon of superinfection exclusion. Molecular biology techniques will be combined with basic virology methodologies, microscopy studies and genomic and transcriptomic analysis tools to:

- Compare the growth characteristics and morphology of the C20 cells with its parental cell line - **Chapter 3**;
- Confirm the presence of a persistent virus infection in C20 and its extension under different conditions - **Chapter 3**;
- Determine if superinfection of C20 cells is possible, given that they have survived an initial virus challenge - **Chapter 4**;

- Evaluate if genomic changes occurred in the persistent AcC20 virus and whether any changes may have had an impact on the establishment or maintenance of the persistent infection - **Chapter 5**;
- Explore the virus and host transcriptome - **Chapter 5**;
- Establish a reliable system for establishing persistent baculovirus infection in insect cells and explore its potential for recombinant protein production - **Chapter 6**;

Chapter 2

Materials and Methods

2.1 Materials

2.1.1 Buffers and solutions

General chemicals were purchased from the following companies unless stated otherwise: BDH Chemicals Ltd, Fermentas, Invitrogen™, Life Technologies, New England Biolabs® Inc. (NEB®), Promega, Sigma-Aldrich, Thermo Fisher Scientific or VWR. Reagent percentages are w/v unless otherwise stated. Cell culture media was supplied by Gibco, PAN-biotec or Oxford Expression Technologies Ltd (OET). Internally prepared solutions (table 2.1) were sterilized either by autoclaving for 15 minutes (15 pounds per square inch) or by filtration (0.2µm). Enzymes and appropriate buffers were supplied by NEB® or Promega and manufactures recommendations followed for reaction optimization.

Table 2.1: Buffers and solutions

Application	Name	Composition
General buffers	Phosphate-buffered saline (PBS, 1x)	From tablets – 10mM sodium phosphate, 2.68mM KCl, 140mM NaCl
	TE buffer (10x)	100mM Tris-HCl, 1mM EDTA pH 8.0
PCR	TBE (5x)	445mM Tris base, 445mM boric acid, 10mM EDTA pH 8.0 (final pH ~8.3)
	Loading buffer (6x)	Gel Loading Dye, Purple (6X), no SDS (NEB®) - 3.3mM Tris-HCl (pH 8.0), 2.5% (v/v) ficoll-400, 10mM EDTA, 0.02% dye 1 and 0.0008% dye 2.
Protein gels	SDS running buffer (5X)	124.6mM Tris base, 0.8% SDS, 1.25M glycine (final pH ~8.3)
	PBS-Tween	PBS, 0.1% (v/v) Tween
	PBS-T milk powder (PBS-TMP)	PBS, 0.1% (v/v) Tween, 5% milk powder
	Coomassie blue staining	20% (v/v) industrial methylated spirit (IMS), 10% (v/v) acetic acid, 0.1% coomassie blue G-250
	Coomassie de-staining buffer	20% (v/v) IMS, 10% (v/v) acetic acid
	Alkaline Phospatase	100mM Tris-HCl pH 9.2, 100mM NaCl, 5mM $MgCl_2$
	SDS loading buffer (6x)	250mM Tris-HCl pH 6.8, 10% SDS, 50% (v/v) glycerol, 0.05-0.5% bromphenol blue, 25% (v/v) β-mercaptoethanol (added fresh)
	Towbin buffer	25mM Tris base pH 8.3, 192mM glycine, 20% (v/v) methanol
Cloning and preparation of competent cells	Super optimal broth (SOB) medium	2% tryptone, 0.5% yeast extract, 8.56mM NaCl, 2.5mM KCl, 10mM Mg_2Cl pH 7.0
	SOB with catabolite repression (SOC) medium	SOB medium, 20 mM glucose
	Inoue transformation buffer	10mM piperazine-N,N'-bis(2-ethanesulfonic acid) (PIPES) pH 6.7, 55mM $MnCl_2 \bullet 4H_2O$, 15mM $CaCl \bullet 2H_2O$, 250mM KCl
Confocal microscopy/ Flow cytometry	Blocking solution (PBS-BSA)	PBS, 1%/3% bovine serum albumin (BSA)
	Permeabilizing solution	PBS, 1% bovine serum albumin, 0.1% v/v Triton X-100
	DPBS-BSA	Dulbecco's PBS 1X (DPBS, Invitrogen™), 1% BSA
Fluorescence assays	Cell lysis buffer	PBS, 1% Nonidet™ P 40 (NP40)

2.1.2 Plasmids and bacterial strains

Plasmids used in this thesis are described in Table 2.2. For generation, maintenance or amplification of the plasmids, DH5α or NM522 disabled strains of *E. coli* strains were used (Invitrogen™). Bacterial stocks were stored at -80°C.

Table 2.2: Plasmids

Plasmid	Description	Reference	Chapter
pOET1	Transfer vector encoding for AcMNPV LEF-2, ORF603 and *polh* promoter and *polh* 3'UTR region	OET	4
pOET3	Transfer vector encoding for AcMNPV LEF-2, ORF603 and *p6.9* promoter replacing *polh* promoter and orf.	OET	4
pOET7	Transfer vector encoding for AcMNPV LEF-2, ORF603 and *ie1* promoter replacing *polh* promoter and orf.	OET	4
pEGFP-N1	Transfer vector encoding *egfp*.	Clontech, USA	4
pIE1-eGFP	Transfer vector encoding for AcMNPV LEF-2, ORF603 and *egfp* driven by *ie1* promoter replacing *polh* promoter and orf.	This thesis	4
p39K-eGFP	Transfer vector encoding for AcMNPV LEF-2, ORF603 and *egfp* driven by *39k* promoter replacing *polh* promoter and orf.	This thesis	4
pP6.9-eGFP	Transfer vector encoding for AcMNPV LEF-2, ORF603 and *egfp* driven by *p6.9* promoter replacing *polh* promoter and orf.	This thesis	4
pPOLH-eGFP	Transfer vector encoding for AcMNPV LEF-2, ORF603 and *egfp replacing polh* orf driven by *polh* promoter.	This thesis	4
pDsRed-Monomer-N1	Transfer vector encoding for the monomeric version of DsRed.	Promega	6
pAcUK13	Transfer vector encoding for the urokinase.	(King *et al.*, 1991)	6
pP6.9-lef2C84A-DsRed	Transfer vector with C84A mutation in *lef2* gene and ORF603 but with *polh* orf and promoter replaced by *dsred* under *p6.9* promoter.	This thesis	6
pP6.9-lef2C84A-UK	Transfer vector with C84A mutation in *lef2* gene and ORF603 but with *polh* orf and promoter replaced by *urokinase* under *p6.9* promoter.	This thesis	6
pPOLH-lef2C84A-DsRed	Transfer vector with C84A mutation in *lef2* gene, ORF603 and *polh* promoter but with *polh* orf replaced by *dsred* driven by *polh* promoter.	This thesis	6
pPOLH-lef2C84A-UK	Transfer vector with C84A mutation in *lef2* gene, ORF603 and *polh* promoter but with *polh* orf replaced by *urokinase* driven by *polh* promoter.	This thesis	6
pBS.SK+CB1$^{LEF\text{-}2,C84A}$	Transfer vector encoding for AcMNPV LEF-2^{C84A}, ORF603 and polyhedrin	(Bannach, 2018)	6

According to the bacterial vector selection marker, antibiotics were added to growth media in the following working concentrations: ampicillin 100µg/ml, chloramphenicol 15µg/ml and kanamycin 30µg/ml.

2.1.3 Insect cell lines

During this project the following cell lines were used for different purposes: *Spodoptera frugiperda* IPLB-SF-21 (Sf21) cells isolated from the fall army worm (Vaughn *et al.*, 1977), Sf9 cells (Thermo Fisher Scientific), a clonal cell line derived from Sf21 cells, *Trichoplusia ni* HighFive™ (BTI-TN-5B1-4; Hi5) cells (Wickham *et al.*, 1992; Wickham and Nemerow, 1993; Granados *et al.*, 1994) and TN-368 (Hink, 1970) derived from the cabbage looper. The C20 cell line, this project focus, was established from Hi5 cells in 2011, as detailed in section 1.8. The Sf9, Hi5 and C20 cells were maintained in ESF921 medium (OET), and Sf21 were grown in TC100 (Gibco) medium supplemented with 10% (v/v) foetal bovine serum [FBS; Labtech, UK; (King and Possee, 1992)]. Cells were maintained shaking cultures (Sf9 and Hi5; 135rpm), monolayer cultures (C20) and spinner cultures (Sf21; 80-90rpm). Occasionally cells were adapted to different growth conditions as experiments required. All cell lines were grown at 28°C.

2.1.4 Viruses

Table 2.3 summarizes all viruses used or generated in this project.

Table 2.3: Viruses

Virus	Description	Reference	Chapter
AcMNPV C6	Plaque isolate of the WT prototype type I *Alphabaculoviruses*	(Possee, 1986; Possee *et al.*, 1991)	4
AcUW1.*lacZ*	Recombinant AcMNPV with *lacZ* under *p10* promoter, replacing *p10* coding region.	(Weyer *et al.*, 1990)	3,4,5
AcEGFP-VP39	Recombinant AcMNPV containing EGFP fused to VP39 capsid gene.	(Danquah *et al.*, 2012)	4
AcRP23.*lacZ*	Recombinant AcMNPV lacking the polh coding region, *lacZ* was inserted under *polh* promoter	(Possee and Howard, 1987)	4
AcIE1-EGFP	Recombinant AcMNPV based on BacPAK6 (OET). The *polh* orf and promoter were replaced by *egfp* driven by *ie1* promoter.	This thesis	4
Ac39K-EGFP	Recombinant AcMNPV based on BacPAK6 (OET). The *polh* orf and promoter were replaced by *egfp* driven by *39k* promoter.	This thesis	4
AcP6.9-EGFP	Recombinant AcMNPV based on BacPAK6 (OET). The *polh* orf and promoter were replaced by *egfp* driven by *p6.9* promoter.	This thesis	4
AcPolh-EGFP	Recombinant AcMNPV based on BacPAK6 (OET). The *polh* orf was replaced by *egfp* driven by *polh* promoter.	This thesis	4
AcP6.9-lef2C84A-DsRed	Recombinant AcMNPV with C84A mutation in *lef2* gene and *polh* orf and promoter replaced by *dsred* under *p6.9* promoter.	This thesis	6
AcPolh-lef2C84A-UK	Recombinant AcMNPV with C84A mutation in *lef2* gene and *polh* orf replaced by *urokinase* driven by *polh* promoter.	This thesis	6

2.1.6 Primers and probes

Primers and probes (DNA Oligos; Table 2.4) were obtained from Eurofins MWG Operon (Germany) except the *actin* probe (Eurogentec, Belgium). All DNA oligos were dissolved to 100 µM stock concentrations [in dH_2O for primers or 1x TE buffer pH 8 (Eurogentec) for probes] and stored at -20°C. A working primer stock (10µM) was maintained at -20°C.

Table 2.4: Primers and probes

DNA oligo name	Sequence (5'→3')	Source	Chapter
RDP1104	TTAATACGACTCACTATAGGGGAAAACAGTCGTCGCTG TCAC	R. D. Possee	3
RDP1105	TTAATACGACTCACTATAGGGGAGTA CATGAAGTTGCGCGTG	R. D. Possee	3
Noda-Fw1	GTTTTCGAAACAAATAAAACAGAAAAGCG	This thesis	3
Noda-Rv1	GCTCGGTGACATAGTAATCAACATCAA	This thesis	3
qPCR gp64 F2	CTCGGTGCTCGACTTTGC	(Bannach, 2018)	3
qPCR gp64 F^{OET}	CGGCGTGAGTATGATTCTCAAA	BaculoQUANT™, OET	3,4,5
qPCR gp64 R^{OET}	ATGAGCAGACACGCAGCTTTT	BaculoQUANT™, OET	3,4,5
qPCR probe $gp64^{OET}$	5´FAM-TTTGGCGCGTGGTGAACGTAGACTTTT-3´TAM	BaculoQUANT™, OET	3,4,5
qPCR gp64 R4	GACGACACGGACGAGTG	(Bannach, 2018)	3
qPCR probe gp64	5´FAM-TTTGGCGCGTGGTGAACGTAGACTTTT-3´BHQ1	(Bannach, 2018)	3
qPCR actin 26	CTCACGCTCAAGTACCCCATC	This thesis	3
qPCR actin 27	TCATTGTAGAAGGTGTGGTGCC	This thesis	3
qPCR actin probe	5´TAM-AGCACGGCATCGTCACCAACTGG-3´BHQ2	This thesis	3
EcoRV-39K	TCATTGATATCTAGGCGGGTGTGATTCGGAC	This thesis	4
BamHI-39K	CTATAGGATCCGTTTGCTTCTTGTAAACCTTTGAAACAAC	This thesis	4
egt-ac16 Fw2	GCGTACTTGCGGCTGTC	This thesis	5
egt-ac16 Rv2	ACCACTAGCAATCTCGTCC	This thesis	5
p95 Fw3	ACCCGGGTAGAATAGGCGTTTTTGT	This thesis	5
p94 Fw	CAATCCACTCCACAGCGT	This thesis	5
P94 Rv	CACAATCGCAACAAGTCATAC	This thesis	5
p95 rv BamHI 2	ATAGGATCCGGAATGATGTCATTTGTTTTCGAC	This thesis	5
egt-ac16 Fw5	GCAAACGAGTTTCTTTACATGTTAATCA	This thesis	5
egt-ac16 Rv5	AATGCCGCAGTTTCTTTTTATGTACA	This thesis	5
p95 fw4	GCCCGAGCCAGAACCA	This thesis	5
p95 rv4	TTGGATTGCGTGTTCAGGTT	This thesis	5
egt-C20_Fw1	CGGTGAAATTCCTTTTGTGTGAT	This thesis	5
egt-C20_Rv1	GCGAGAAAAACCGTGTGC	This thesis	5
egt-C20 Fw2	AATGGAATGAGATGGAATGGAATG	This thesis	5
egt-C20_Rv2	CAAGATCCAGGACCACAACC	This thesis	5
LacZ_Fw1	ACCAGAAACAGCACCTCG	This thesis	5
LacZ_Rv1	CCCAAGTTCAGTCTGCCG	This thesis	5
RDP213	TGAGACGCACAAACTAATATCAC	R. D. Possee	4,6
RDP214B	ATACGTACAACAATTGTCTGTAAATCAAC	R. D. Possee	4,6
RDP737	GTCTGTACACGAACGCG	R. D. Possee	5
RDP738	GTCTGGCAAGAATCACAAC	R. D. Possee	5
RDP887	ATGGCGAATGCATCGTATAACGTGTG	R. D. Possee	6
RDP1275	TCCGAGTTTGTCGACAAGG	R. D. Possee	5
RDP1279	CCAGTTCGGTTATGAGCCG	R. D. Possee	5
Uk-Fw-HindIII	ATAGAAGCTTAATATGAGAGCCCTGCTGG	This thesis	7
Uk-Rv-EagI	TATTCGGCCGTGTCGGGTTTAACATTACG	This thesis	7

2.1.5 Antibodies

Primary and secondary antibodies used in this study were diluted in PBS-Tween with 5% milk or PBS-BSA (Table 2.1) to the appropriate concentration (Table 2.4).

Table 2.5: Antibodies used in this thesis

Antibody	Source species	Working dilution	Supplier
Anti-GP64 AcV5-tag	Mouse	1:1000 (WB)	Sigma-Aldrich
Anti-GFP	Rabbit	1:3000 (WB)	Abcam
Anti-P10	Rabbit	1:500 (WB)	GeneScript Biotech. Corporation
Anti-Polh	Guinea pig	1:500 (WB)	Made in house (N. Groome)
Anti-mouse IgG conjugated to horseradish peroxidase (HRP)	Goat	1:50000 (WB)	Jackson ImmunoResearch
Anti-rabbit IgG conjugated to HRP	Goat	1:50000 (WB)	Sigma-Aldrich
Anti guinea pig Immunoglobulin G (IgG, Alkaline Phosphatase)	Goat	1:10000 (WB)	Sigma-Aldrich
Alexa Fluor 488 anti-mouse	Goat	1:1000 (IF)	Life techologies
Anti-α-Tubulin	Mouse	1:750 (IF)	Sigma-Aldrich
Alexa Fluor 594 Phalloidin conjugated	-	1:200 (IF)	Invitrogen™

2.2 Virus protocols

2.2.1 Virus amplification

Virus stocks were amplified in Sf9 cells as described before (King and Possee, 1992). Cells were incubated until majority showed signs of infection (4 or 5 days) and removed from the culture by centrifugation (4000 rpm; 15min; 4°C; TYJS 4.2 rotor, J6-MI Beckman centrifuge). After this clarification, the culture medium containing the BV was supplemented with 2% (v/v) FBS and stored at -80°C. Working stocks were stored at 4°C up to one month.

2.2.2. Titration of virus stocks

Virus titration was performed following the baculoQUANT™ All-in-one qPCR method (OET; 7500 Real Time PCR system, Applied Biosystems™) or by plaque assay using Sf21 cells as previously described (Brown and Faulkner, 1978; King and Possee, 1992;

Hitchman *et al.*, 2007). Plaque assays were incubated for four days before staining except for AcC20 virus isolated directly from the C20 cells, incubated for 5 days.

2.2.3 Generation of recombinant viruses by co-transfection of insect cells

Recombinant viruses were generated by co-transfecting Sf21 cells with the bacterial transfer vector containing the gene of interest and BacPAK6 (OET, Ltd), following supplier's protocol. After incubating for 5 days, the medium containing BV was clarified by centrifugation and used for plaque-purification.

2.2.4 Plaque purification of recombinant BacPAK6 virus

To separate the recombinant virus from the parental virus, a plaque assay was performed using neat and 10^{-1}, 10^{-2} and 10^{-3} dilutions (see 2.2.2). After standard incubation, the cultures were stained with 1ml of neutral red at 0.125mg/ml and 8µl of a 5% (v/v) X-gal solution in dimethylformamide (DMF). After allowing the culture to de-stain, isolated colourless plaques were picked with a Pasteur pipet and mixed extensively in 500µl of TC100 with 10%FBS to release the BV into the media. Preparations were used successively in plaque assays until only colourless plaques were observed in a single well (2-3 plaque assays). The BV suspension from the last plaque assay, passage 0, was used to amplify passage 1 as described in 2.2.1.

2.3 Bacteria protocols

2.3.1 Preparation of ultra-competent cells by the Inoue method

DH5α or NM522 *E. coli* cells were plated in Lysogeny broth agar (LBA) Petri dishes and grown overnight at 37°C. A single colony was picked and grown in 5ml SOB media. The next morning this culture was used to inoculate 250ml of SOB in a shaking culture (200rpm) until an 0.55 optical density was reached and competent cells prepared as described (Green *et al.*, 2012). Aliquots of 300µl were snap-frozen and stored at -80°C until needed.

2.3.2 Transformation of competent *E .coli* cells

Competent cells (see 2.3.1), were thawed on ice. Per reaction, 50µl of *E. coli* competent cells were transformed with 2-5µl of the ligation mix (see 2.4.4) as described (Green *et al.*, 2012). A non-digested vector (1-10ng) was treated in the same way to test cells efficiency.

2.3.3 Evaluation of plasmid constructs and DNA purification

From the LBA Petri dishes with transformed cells (see 2.3.2) incubated overnight, 10 single colonies were picked and diluted in 10µl of 1x TE. Each sample was then tested by PCR with appropriate primers to screen the presence of the insert (see 2.4.1). The positive colonies were grown overnight (12-15h) in shaking cultures (200rpm) of LB supplemented with the appropriate antibiotic. The plasmid DNA was purified using QIAprep Spin Miniprep kit, following the supplier's protocol (Qiagen, Netherlands). A bacterial stock in LB with 25% (v/v) glycerol was stored at -80°C. Alternatively to the PCR screening, 6 colonies were grown in liquid media as described above. The plasmid DNA was extracted and then digested with appropriate restriction enzymes to test the presence of the right insert. Independently of the screening method, the DNA from the constructs presenting the right insert was sequenced by the Sanger method (see section 2.4.6).

2.4 Molecular biology protocols

2.4.1 Polymerase chain reaction (PCR)

Table 2.4. lists the primers used in this project. Reactions were performed with *Taq* or Q5 polymerase depending on the final aim of the experiment. Q5 Proofreading activity makes it more reliable. It was used when amplifying fragments for cloning or sequencing to identify possible SNPs in specific genes. Manufactures recommendations were followed and reactions were performed with 5ng of template samples/positive control DNA and a "no template" control (NTC, template replaced for dH_2O), at described conditions (Table 2.6). Annealing temperatures for primer pairs were calculated using Tm Calculator v1.9.13 (NEB, 2018) and optimized with replicas of the positive and negative control amplified at the recommended temperature and +/-3°C. PCR

amplifications were carried out in a SureCycler 8800 thermocycler (Agilent technologies, USA) or a ProFlex PCR system (Thermo Fisher Scientific, USA).

After amplification, 5µl of each sample were mixed with 1µl of loading buffer (6x; NEB®) and analysed in an agarose gel (see 2.4.3). For sequencing the products were then purified with the QIAquick PCR purification kit, following the manufacture's protocol (Qiagen).

Table 2.6: Thermocycling conditions for a PCR with *Taq* or Q5 polymerase

***Taq* polymerase**			
Step		**Temp.**	**Time**
Initial Denaturation		95°C	1min
30 cycles	Denaturation	95°C	20s
	Annealing	50-65°C	40s
	Extension	68°C	1min/kb
Final extension		68°C	5min
Hold		4°C	

Q5 polymerase			
Step		**Temp.**	**Time**
Initial Denaturation		98°C	1min
30 cycles	Denaturation	98°C	10s
	Annealing	55-68°C	20s
	Extension	72°C	30s/kb
Final extension		72°C	2min
Hold		4°C	

2.4.2. Restriction enzyme digests

In general, 1.5-3µg of plasmid DNA or 400-600ng of DNA from PCR products were digested using enzymes, buffers and protocols provided by NEB®. The reactions were incubated for 2h-4h and heat inactivated as described. From each reaction, 50ng were applied to an agarose gel for electrophoresis (see 2.4.1 and 2.4.3). For subsequent cloning steps, the vector DNA was dephosphorylated by addition of 2µl antarctic phosphatase (NEB®) and 1x of its own buffer in a 60 µl reaction. This was incubated at 37°C for 30min and heat inactivated at 80°C for 2 minutes. Samples were purified using the QIAquick PCR purification kit (Qiagen) and stored at -20°C for the ligation step.

2.4.3 Agarose gel electrophoresis

Agarose gels (1%, Thermo Fisher Scientific) were prepared as described (Green *et al.*, 2012) and safe view nuclei acid stain (NBS Biologicals) added as recommended. The solidified gel was submerged in 1x TBE buffer and the samples were loaded. DNA electrophorese were run horizontally at 80-120 volts depending on the tank size for 45min-1.5h. The DNA was visualized using a UV transilluminator (Bio Rad, ChemiDoc™ Imaging Systems).

For purification of a DNA fragment by electrophoresis, the restriction digestion or PCR reaction was mixed with 1x loading buffer and the DNA fragments separated as described above. Appropriate DNA fragments were visualized using a UV transilluminator, excised from the gel and purified using the QIAquick Gel Extraction Kit (Qiagen).

2.4.4 DNA Ligation

Ligation between two DNA fragments with complementary ends was performed using T4 DNA Ligase (Promega). Molar ratios of vector and insert were calculated using the online Promega Biomath Calculator (Promega, 2018); molar ratios of 1:1, 1:3, 3:1 were used. The reactions were carried out in a 10µl reaction with 30-40ng vector DNA, 1µl T4 Ligase buffer and 0.5µl of enzyme. Controls with and without T4 ligase were prepared to test: the restriction enzyme digest, by missing the insert; and the ligation reaction by using single digested vectors. Reactions were incubated according to supplier's instructions and used for transformation in *E. coli* competent cells (see 2.3.2).

2.4.5 DNA sequencing

All DNA sequencing was outsourced to Source Bioscience (U.K.), and the resultant sequencing data was analysed using SnapGene software (GSL Biotech, 2017).

2.4.6 Multiplex-qPCR assay for viral replication analysis

A qPCR was designed using *gp64* as the target gene and actin as the endogenous control. Primers and probe for *actin* were designed using OLIGO 7 software (Molecular Biology Insights; Table 2.4). For *gp64*, primers and probe designed by Carina Bannach were used [(Bannach, 2018); Table 2.4]. After optimization (see 3.2.6.2a), 2x Low rox mix, 400nm of each primer, 100nm probe and 0.5ng of DNA were mixed in a total volume of 20µl per reaction. Each reaction was performed in duplicate in a 96-well plate, except for the optimization steps and the calibrator sample, where triplicates were prepared. A DNA sample of Hi5 cells infected with AcUW1.*lacZ* at 0.1 MOI harvested at 24h p.i., was used as a calibrator for all the plates. The plate was centrifuged at low speed to remove air bubbles and the 7500 Real Time PCR system (Applied Biosystems™) was used to perform the reaction. The thermal profile consisted of 4.5min of initial

denaturation at 95°C, followed by 40 cycles of 10s denaturation at 95°C and 35s for annealing at 60°C. Fluorescence measurements were taken on the annealing step.

The C_t from the technical replicates was averaged and the ΔΔCt method was used to analyse the experiment. The result $2^{-\Delta\Delta Ct}$ represents the fold difference to the calibrator (Livak and Schmittgen, 2001; Applied Biosystems, 2008; Invitrogen, 2008). The comparison between samples is achieved by doing the ratio in between the samples $2^{-\Delta\Delta Ct}$.

$$\text{Fold difference} = 2^{-\Delta\Delta Ct}$$

$$\Delta Ct \text{ sample} - \Delta Ct \text{ calibrator} = \Delta\Delta Ct$$

$$Ct\ gp64^{sample} - Ct\ actin^{sample} = \Delta Ct \text{ sample}$$

$$Ct\ gp64^{calibrator} - Ct\ actin^{calibrator} = \Delta Ct \text{ calibrator}$$

2.4.7 Whole genome sequencing and bioinformatics analysis

2.4.7.1 Genome Assembly

AcC20 DNA-Seq was used to assemble a strain specific genome with SPAdes v3.11.1 (Bankevich *et al.*, 2012) using kmers 21,33,55,77 and AcUW1.*lacZ* genome as a trusted contig. Scaffolds generated were aligned against AcUW1.*lacZ* genome using BLASTN v2.6.0 (Altschul *et al.*, 1990). Four AcC20 scaffolds hits covered most of the AcUW1.*lacZ* genome. Regions of this genome not aligned to the AcC20 scaffolds were extracted and used as queries to identify other possible scaffolds. Overlapping AcC20 scaffolds recovered could be assembled into one, except for a gap at 12,478-13,211bp (AcUW1.*lacZ* genome coordinates). PCR and Sanger sequencing completed the genome assembly (egt-ac16 Fw5 and egt-ac16 Rv5; Table 2.4). The coverage of AcC20 DNA-Seq was assessed. Reads were mapped with BWA v0.7.17 (Li and Durbin, 2009) and BAM files were converted to bedgraph format with Bedtools v2.26.0 (Quinlan and Hall, 2010) and visualised in R v3.5.2 (R Core Team, 2014) with the ggplots2 package (Wickham, 2016). The *hr*s were mapped by identifying regions accumulating *Eco*RI restriction sites.

2.4.7.2 Transcriptome assembly and annotation

To identify transcripts in the AcC20 assembled genome, AcC20 RNA-Seq samples were used to assemble a *de novo* transcriptome (Trinity v2.2.0; Grabherr *et al.*, 2011). The

transcript sequences were translated with Transdecoder v3.0.0 (Haas *et al.*, 2013) and longest ORFs with homology to the Pfam (Eddy, 1998; Finn *et al.*, 2014) and the SwissProt databases (Altschul *et al.*, 1990; Bairoch and Apweiler, 2000) were retained. For annotation, all AcMNPV protein sequences downloaded from NCBI and the AcC20 genome were used to respectively BLASTP or TBLASTN v2.6.0 (Altschul *et al.*, 1990) query all the retained sequences. Hits to both were retained as initial C20 genes. These were further manually inspected in relationship to the AcMNPV previous described gene annotation (Chen *et al.*, 2013).

2.4.7.3 Variant analysis between AcUW1.*lacZ* and C20 DNA-seq

GATK v4.0.5.1 (DePristo *et al.*, 2011; Van der Auwera *et al.*, 2013) was used to identify variants between the AcUW1.*lacZ* and AcC20 strains. The AcC20 DNA-Seq was mapped to the AcUW1.*lacZ* genome with BWA mem (Li and Durbin, 2009) with–M –t 8 options. Duplicate reads were removed (PICARD tools), SAM files converted to BAM files and BAM/fasta files indexed with Samtools v1.7 (Li *et al.*, 2009). Due to the lack of known variants in AcMNPV four rounds of GATK base quality score recalibration were required before variants were called. A hard filter for SNPs (QD<2.0, MQ<40.0, FS>60.0, SOR>3.0, MQRankSum<-12.5, ReadPosRankSum<-8.0) and INDELs (QD<2.0, ReadPosRankSum<-20.0, FS>200.0, SOR>10.0) was used to retain possible variants.

2.4.7.4 Visualisation of data

Circos v0.69-6 was used for visualisation of the assembly scaffolds and variant calling analysis (Krzywinski *et al.*, 2009). The AcUW1.*lacZ* and AcC20 genomes were aligned using MAFFT v7.310 (Katoh and Standley, 2013) with --op 5 --ep 0 –maxiterate 100 parameters. A custom python script was used to adjust the position of variant calling locations, genes and *hrs*, based on the gaps created from the MAFFT alignments.

2.4.8 Reverse transcription-PCR

Reverse transcription-PCR (RT-PCR) was used to detect the presence of contaminant FHV genome in insect cell lines. The total RNA extracted from the cells using the RNeasy mini kit (Qiagen) was treated with deoxyribonuclease I (DNase I) according to recommendations (NEB®). It was then amplified with primers Noda-Fw1 and Noda-Rv1

(Table 2.4) using SuperScript III one-step RT-PCR system with Platinum™ Taq DNA Polymerase (Invitrogen™) and following manufactures instructions.

2.4.9 RNA sequencing

Superinfected and non-infected C20 cells, infected and mock Hi5 cells were prepared in replicate samples at the same time to be harvested at 0, 2, 8, 12, 18, 24, 48h p.i. Total RNA was isolated using TRIzol® (Invitrogen™) following the supplier´s recommendations. Strand-specific library generation was performed as previously described (Zhong *et al.*, 2011; Chen *et al.*, 2013).

2.5 Protein analysis

2.5.1 Production of virus-infected cells

Hi5 and C20 cells were seeded at $0.5x10^6$cells/well in a 6-well plate and infected at 0.1-5 MOI or mock-infected as described (King and Possee, 1992). Infected cells were incubated until the desired harvesting time.

At each time point, the media was removed and the cells re-suspended in 1ml of 1x TE with 1x proteinase inhibitor (PI; Sigma-Aldrich). The cells were concentrated by centrifugation for 3min at 9000rpm at 4°C in a microfuge. Cell pellet was re-suspended in 160µl of 1x TE containing 1x PI and 40µl of SDS loading buffer was added. The samples were boiled for 5min and stored at -20°C until the protein analysis was performed.

2.5.2 Polyacryamide gel electrophoresis

Protein samples were analysed using sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) as described (Green *et al.*, 2012). According to target protein sizes, 10, 12 or 15% gels were prepared in house (Green *et al.*, 2012). Before loading on the gel the samples were boiled for 5min and mixed thoroughly. Samples were then electrophoresed using the mini-PROTEAN vertical gel electrophoresis system (Bio-Rad) at 160-180V. Once the appropriate separation has been achieved, the electrophoresis was stopped and the proteins were either stained with Coomassie blue or transferred to a membrane for western blot (WB) analysis.

2.5.2.1 Coomassie staining of proteins

Following electrophoresis, the gels were submersed in Coomassie blue staining (see Table 2.1) and heated in a microwave for 10-20s avoiding boiling. The stain was washed off with dH_2O after 15min and the gel submerged in Coomassie de-staining buffer. After heating for 10-20s, they were allowed to de-stain for 30min. Coomassie de-staining buffer was replaced and the same procedure was followed. A picture was taken (Bio Rad ChemiDoc™ Imaging System) when a clear background was observed (2h to overnight).

2.5.2.2 Western Blots for protein analysis

Separated proteins in the gel were transferred to a nitrocellulose membrane as described (Green *et al.*, 2012) and following supplier's instructions for the Trans-Blot Turbo blotting system (BioRad). After transfer, the membrane was incubated in 5% milk (Marvel) in PBS-Tween for 30 min, followed by the required dilution of the primary antibody in the same buffer for 1h at RT (Table 2.5). The membrane was washed six times for 5min in 5% milk before the appropriate secondary antibody was added for 1h. All incubations were carried out at RT using a horizontally rotating platform. The membrane was washed as before and developed. For HRP secondary antibodies Clarity western ECL blotting substrate kit was used following the supplier's instructions (Bio Rad) with a 5min incubation and revealed using a Bio Rad, ChemiDoc™ Imaging System. Alternatively, the alkaline phosphatase method was followed to detect polyhedrin protein. Per membrane, 5ml of alkaline phosphatase were mixed with 16.5µl 5-bromo-4-chloro-3-indolyl-phosphate (BCPI) and 33 µl nitro blue tetrazolium chloride (NBTI). The membrane was incubated until bands were observed and imaged with the same system. For comparing proteins in between cell lines, membranes were processed and imaged simultaneously to ensure the exact same exposure conditions were performed.

2.6 Flow cytometry

Flow cytometry was used to identify the percentage of infected cells producing a target protein as a sign of infection. For non-fluorescent proteins (GP64), immunostaining was used. The cells were seeded and infected in 12 well-plates and at the desired time point fixed with 4% formaldehyde in ice cold PBS, followed by an additional fixation step with 70% ethanol. After blocking with PBS-BSA 3%, 200µl of 1:1000 mouse α-gp64 primary

antibody was added. As a secondary antibody 200µl of Alexa Fluor 488 anti-mouse 1:1000 was used. The antibodies were left for 45min each. Finally, 200 µl of DAPI 1:5000 was added to stain the nucleus. Washes in appropriate media were performed in between each step. The final sample was re-suspended in 1X DPBS (Invitrogen™).

For identification of cells producing a fluorescent protein (DsRed), the cells were grown as described above and at the desired time point, they were stained with 10µl of Hoechst 33342 solution in their previous media [2 drops of NucBlue® Live ReadyProbes® Reagent (Invitrogen) per 50µl of ESF921] and incubated for 1h. The cells were then harvested, washed in DPBS-BSA 1% and resuspended in 200µl of DPBS.

NovoCyte Flow Cytometer with NovoExpress 1.3.0 software (ACEA Biosciences, USA) was used to perform the experiment using the appropriate laser and detector for each fluorochrome (Table 2.7). From the total cases, the cells were selected by the presence of blue fluorescence (DAPI/Hoechst 33342) followed by size and single cell selection. Finally, non-infected cell samples were used as a negative control to establish the positive/negative threshold for Alexa Fluor 488 or DsRed fluorescence. These gates and were applied to the remaining samples to be tested.

Table 2.7: Flow cytometry laser and detector settings for fluorochromes

Fluorochrome	Laser	Detector
Alexa Fluor 488	488nm	530
DsRed	488nm	615/620nm
Dapi	405nm	445nm
Hoechst 33342	405nm	445nm

2.7 Microscopy techniques

2.7.1 Equipment and lasers

Bright field and fluorescence microscopy was performed using a Zeiss Axiovert 135 and a ProgRes®MF camera (Jenoptik, Germany), while for confocal microscopy a LSM 880 laser scanning microscope (Zeiss) was used. For scanning electron microscopy, samples were imaged using a Zeiss Merlin compact field emission gun SEM. Samples were captured with technical help from Dr Louise Hughes (Oxford Brookes Bioimaging Unit).

All confocal images were acquired with a transmitted light channel combined with a multi-track setup with appropriate filters for the different dyes or fluorescent proteins to prevent signal cross-over. An argon laser with an excitation of 488nm was used to capture images with EGFP or Alexa Fluor 488, a diode laser with an excitation of 405nm, was used to capture DAPI and a HeNe laser was used exciting at 561nm for cresyl violet and Alexa Fluor 594.

Images were acquired using an oil immersion objective; Plan-Apochromat 63X (1.4 numerical aperture) or 100x objective optionally combined with Airyscan mode. Zen 2 blue edition (Zeiss) was used for post-acquisition image processing.

2.7.2 Confocal laser scanning microscopy

For confocal microscopy, cells were infected with either AcEGFP-VP39 to track virus uptake or AcUW1.*lacZ* for immunostaining of the cytoskeleton. For live cell imaging, $0.5x10^6$ C20 or Hi5 cells were seeded in glass bottom dishes (Miltenyi Biotec) in triplicate and incubated overnight at 28°C. The cell medium was then replaced by 1µM cresyl violet in 1ml of PBS and incubated for 5min at 28°C. The cells were washed twice with PBS and one dish at a time infected with 200 MOI of AcEGFP-VP39 directly on the microscope platform and imaged as “time zero”. A time-series experiment comprising 95-102 frames over 8min was performed with a 100x objective and repeated at 30min, 60min, 90min and 120min for each of the infected cell line samples and non-infected controls.

For imaging fixed cells infected with the same virus, $2.2x10^5$cells/well were seeded in 12-well plates with No 0 coverslips (Scientific Laboratory Supplies Ltd., UK) and incubated overnight. Cells were infected at 500 MOI and the fixing process started 20 and 40min post-infection. The cells were fixed in 4% v/v formaldehyde (diluted from a 16% v/v formaldehyde solution, Thermo Fisher Scientific) in PBS buffer for 10min and washed 4 times for 5min in PBS buffer. The preparations were mounted in slides with Vectashield mounting media containing DAPI to mark the DNA (Vector Laboratories). The coverslip was sealed against the slide with nail vanish and the samples were stored at 4°C protected from light for imaging within 2 weeks.

For immunostaining the cells were seeded the same way as for imaging of fixed cells and, after overnight incubation at 28°C they were infected with 500 MOI of AcUW1.*lacZ*. The fixing and staining process was started 96h p.i. by permeabilizing the cells with 1ml of permeabilizing solution (Table 2.1) for 10min. The cells were then washed twice with 2ml of PBS, followed by two washes with the blocking solution (PBS-BSA). The cells were immunostained with 200µl of primary antibody for α-tubulin (1:750) to detect microtubules and Alexa Fluor 594 Phalloidin conjugated (1:200) to detect F-actin diluted in the blocking solution and incubated at RT for 1h. Unbound antibody was removed by washing the preparations three times with PBS-BSA. The secondary antibody Alexa Fluor 488 anti-mouse (1:1000) was then added to the cells, diluted in the blocking solution and incubated for 45min. Following another set of three washes, 1µg/ml of DAPI in 200µl was added for 5min (Sigma-Aldrich). Samples were washed and mounted with Vectashield mounting media without DAPI (Vector Laboratories). The samples were sealed and stored as described before.

2.7.3 Scanning electron microscopy

For scanning electron microscopy (SEM), $2.2x10^5$cells/well were seeded in 12-well plates with round No 0 coverslips (Scientific Laboratory Supplies Ltd., UK) in the bottom and incubated overnight. Two samples from each cell line were infected with AcEGFP-VP39 at 500 MOI and the fixing process started at 20min p.i. Infected cells and controls were fixed in 2% (v/v) paraformaldehyde and 2.5% (v/v) glutaraldehyde in 0.1M sodium cacodylate buffer for 30min and washed four times for 5min in 0.1M of the same buffer. Secondary fixation was performed with 1% osmium with 0.2M sodium cacodylate buffer for 30min and the preparations washed as before. Dehydration steps were performed in a sequential series of increasing ethanol concentration for 10min each (10%, 30%, 50%, 70% and 90%) followed by three washes in 100% ethanol and one in 100% dried ethanol. The coverslips were mounted into stubs and sputter-coated with 16.4nm of gold particles using the Agar auto sputter coater with thickness monitor control.

For imaging of BV, drops of 50µl AcEGFP-VP39 containing approximately $5.45x10^6$pfu were placed in coverslips and allowed to settle for 15min. To this initial inoculum, 50µl of 8% (v/v) formaldehyde in PBS was added and fixed for 10min. The inoculum was

carefully removed by pipetting and the protocol followed as before for fixing infected cells for SEM. The sample was sputter-coated with 12.1nm of gold particles.

2.8 Fluorescence and colorimetric assays

Fluorescent and colorimetric assays were used to analyse gene expression or production of recombinant proteins. To evaluate gene expression by fluorescence assays, cells seeded at $3.2x10^4$cells/well in 200µl in black-wall 96-multiwell plates. Virus inoculum at the desired MOI was diluted in 20µl ESF921 and left incubating with the cell monolayers for 45min. Mock-infected cells were inoculated at the same time. The virus inoculum was removed and replaced by 200µl ESF921. At the desired time p.i. the media was removed and the cells carefully washed, leaving them in PBS before the florescence reading. To evaluate production of recombinant protein the infected and mock cells were grown in 24 deep-well plates in a shaking platform at 1100rpm. At the desired time p.i. the cells were concentrated by centrifugation at 4000rpm and washed in PBS. After a second concentration step the cells were resuspended in PBS with 1% Nonidet™ P40 (NP40) to lyse the cell membrane and release cytoplasmic proteins. Finally, the samples were concentrated at 10,000rpm and the supernatant was collected and dispensed in black-wall 96 multiwell plates. To measure background fluorescence, PBS with 1% NP40 was dispensed in replicate wells. Fluorescence was read in nine equidistant points per well at 480nm/515nm for EGFP and 556nm/586nm for DsRed respectively for excitation and emission using the SpectraMax iD3 Microplate reader (Molecular devices, USA).

Colorimetric assays were used to test the production of urokinase. Infected and control cells were grown in shaking cultures as described above and concentrated by centrifugation at the desired time points p.i. Since urokinase is secreted, 100µl reactions were prepared with 70µl of the cell media to evaluate its production using the uPA Activity Assay Kit and following instructions (Merck, Germany). The plates were read using the LT-4500 Microplate Absorbance Reader (Labtech, France).

2.9 Statistical analysis

Statistical analysis was performed using SPSS (IBM Corp., 2017). Shapiro-Wilk test was used to test for normality. In accordance, a t-test for independent samples or a Mann

Whitney test for single comparisons was used for parametric or non-parametric datasets respectively. The result was considered significant if $p<0.05$. Graphs were developed using GraphPad Prism 7 (GraphPad Software, USA) or SPSS and errors bars represent standard deviation (SD) when possible or standard error of the mean (SE) as described in the figure legends. These were calculated from datasets including three independent experiments with three biological replicas each, unless otherwise stated.

Chapter 3

Characterisation of an insect cell line harbouring a persistent baculovirus infection

3.1 Introduction

Persistent baculovirus infections are rarely observed *in vitro* and only a few cases have been described (McIntosh and Ignoffo, 1981; Crawford and Sheehan, 1983; Hughes *et al.*, 1993; Lee *et al.*, 1998; Weng *et al.*, 2009). Therefore very little is known about how the host cell adapts to enable it to harbour the persistent infectious virus, or how the virus may change to enable it to persist, or the relative contributions of each. From the host cell perspective, a previous study has suggested that changes in cell morphology as well as growth patterns may play a role (McIntosh and Ignoffo, 1981). The same study reported the persistent virus had an infectious titre 100 times lower than the one produced by a similar lytic infection. In addition, polyhedra production was severely reduced (McIntosh and Ignoffo, 1981). Each persistently infected cell line consists of a unique balance between the host cell and the infecting virus. Hence, although previous studies might enable comparison, extrapolations cannot be made to the current study.

The C20 cell line was established by sub-cloning from Hi5 cells persistently infected with AcUW1.*lacZ* (R.D. Possee, personal communication). This was achieved eight years ago (in 2011), giving ample time for both the cell line and virus to evolve and thus differ from the parental cell and virus. The first step in this study was the characterisation of the C20 cell line growth pattern. A careful analysis of the virus infection pattern within and between different cultures was also carried out. Persistent and latent infections are often confused, hence this detailed characterisation aimed to distinguish between the two.

It is well known that the cytoskeleton undergoes major changes after the cell is infected by a baculovirus. Actin filaments have a key role for efficient retrograde transport of nucleocapsids toward the nucleus (Charlton and Volkman, 1993; Ohkawa *et al.*, 2010). G-actin is recruited to the nucleus where it is polymerized into F-actin, which is essential for BV morphogenesis (Volkman *et al.*, 1992; Charlton and Volkman, 1993; Lanier and Volkman, 1998; Ohkawa and Volkman, 1999; Volkman and Kasman, 2000; Ohkawa *et al.*, 2002, 2010; Goley *et al.*, 2006; Volkman, 2007; Gandhi *et al.*, 2012). On the other hand, microtubules, particularly β-tubulin are essential for the anterograde transport of nucleocapsids to the plasma membrane so they can bud out of the cell (Dai *et al.*, 2004; Fang *et al.*, 2007, 2009a; Danquah *et al.*, 2012). However, all these baculovirus studies

were based on a lytic infection cycle. Hence it was important to evaluate the impact of a persistent/latent baculovirus infection on the cytoskeleton host.

Finally, as part of this initial characterisation it was important to ascertain whether C20 were latently infected with nodavirus (FHV). Previous studies have shown that some stocks of Hi5 cells harbour a latent FHV infection (Li *et al.*, 2007; Merten, 2007) although some others (e.g. MSU-TnT4 cells) are clean (Zhang and Thiem, 2010). Even though when present, the infection is in a latent state, the possible re-activation induced by baculovirus is an important factor to consider when studying the finely constructed balance between host cell and AcC20 persistent virus in the C20 cells. In addition to testing C20 cells, other insect cell lines in regular use in the laboratory (Hi5 and TN368 and Sf9) were also tested for the presence of nodavirus.

3.2 Results

3.2.1 Morphological characterisation of C20 cells

Insect clonal cell lines such as Sf9 or Hi5 cells usually have a very well defined appearance. They are round, normally with a regular size and so called “shiny” surface.

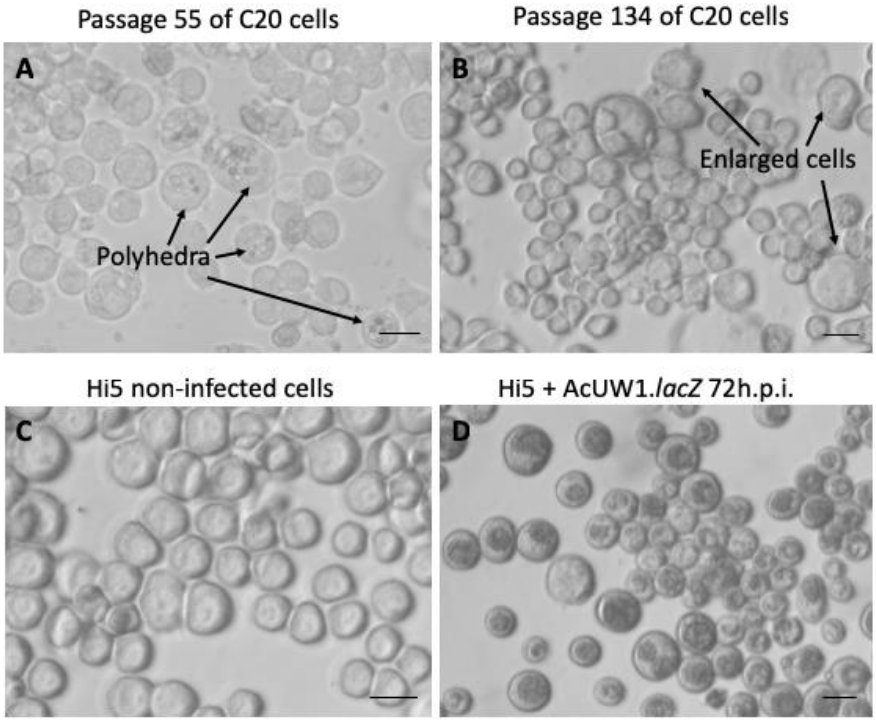


Figure 3.1: Bright field microscopy examination of non-infected and virus-infected cells.
Representative images of A) passage 55 and B) passage 134 of C20 cells, C) Hi5 cells non-infected and D) Hi5 infected with 5 MOI of AcUW1.*lacZ* at 72h p.i. Scale bars 20μm.

When infected with AcMNPV, cytopathic effects include increased cell size, odd shaped cells, presence of vacuoles and vesicles and ultimately the formation of polyhedra. As part of the very late infection phase, the nucleus enlarges to accommodate the newly produced polyhedra and, therefore, virus-infected cells are bigger than non-infected controls. Examples of healthy Hi5 cells and those infected with AcUW1.*lacZ* at 72h p.i are presented in Figure 3.1C and D respectively. The C20 cells show signs of virus infection that are more evident in some passages than others. Examples of enlarged cells, odd shaped cells and cells producing polyhedra are shown (Fig. 3.1 A and B).

Using a bright field microscope (Echo Revolve, VWR), C20 cell sizes were measured and compared with Hi5 cells. Cells were stained with trypan blue to exclude non-viable cells. The boxplots in Figure 3.2 represent a population of Hi5 and C20 cells grown in suspension cultures or monolayers. C20 cells had a much higher size variability than Hi5 cells as observed by the box size and inner fences in both shaking and monolayer cultures. There was a significant increase in size for both Hi5 and C20 cells when these were seeded in monolayer instead of shaking cultures ($p_{Mann\text{-}Whitney}<0.001$). This is possibly due to the attachment to the growth surface that is easily observed under a bright field microscope. However, this increase in size was not as accentuated for C20 cells as for Hi5 cells. In C20 cells, the median of size in monolayer and shaking cultures was 22.6μm and 20.3μm respectively, while for Hi5 the medians were 20.1μm and 16.7μm for monolayer and shaking cultures, respectively. Most importantly this analysis showed that C20 cells, independently of the culture method, were larger than the parental Hi5 cells ($p_{Mann\text{-}Whitney}<0.001$).

Finally, the number of C20 cells presenting polyhedra was counted. Only 5 out of 668 analysed cells showed polyhedra production. For this reason, statistical analysis was not performed in relation to the culture method.

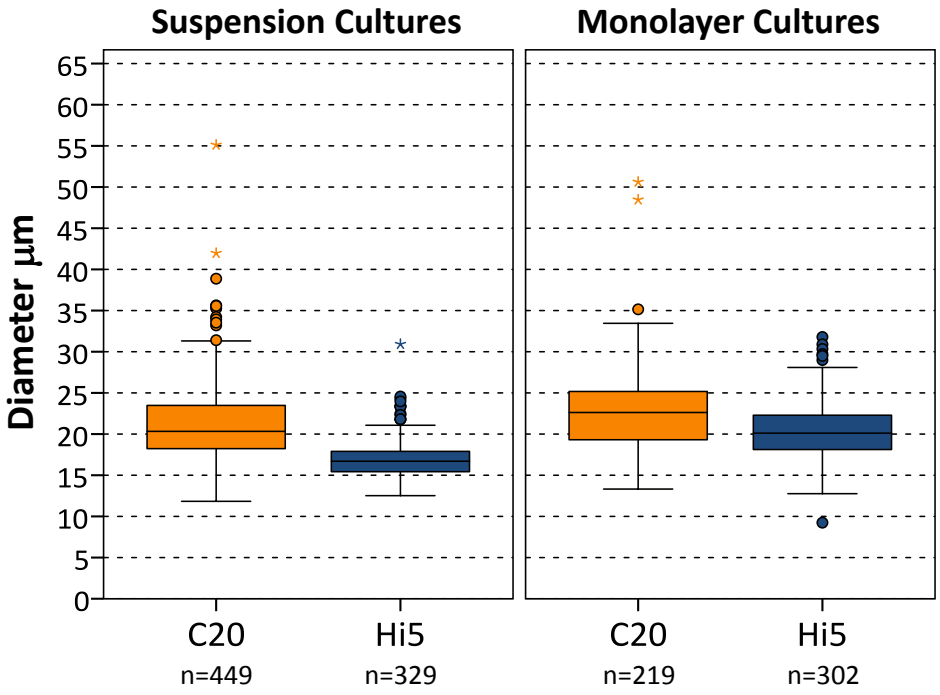


Figure 3.2: Cell diameters
Box plots representing diameters (µm) under a bright field microscope of C20 and Hi5 cells growing in suspension or monolayer cultures. Sample numbers (n) are shown. Boxplots including quartiles and inner fences were generated according to Tukey method. Cases not falling in the inner fences were marked as outliers (circles) or extreme outliers for values more than three times the height of the boxes (asterisks).

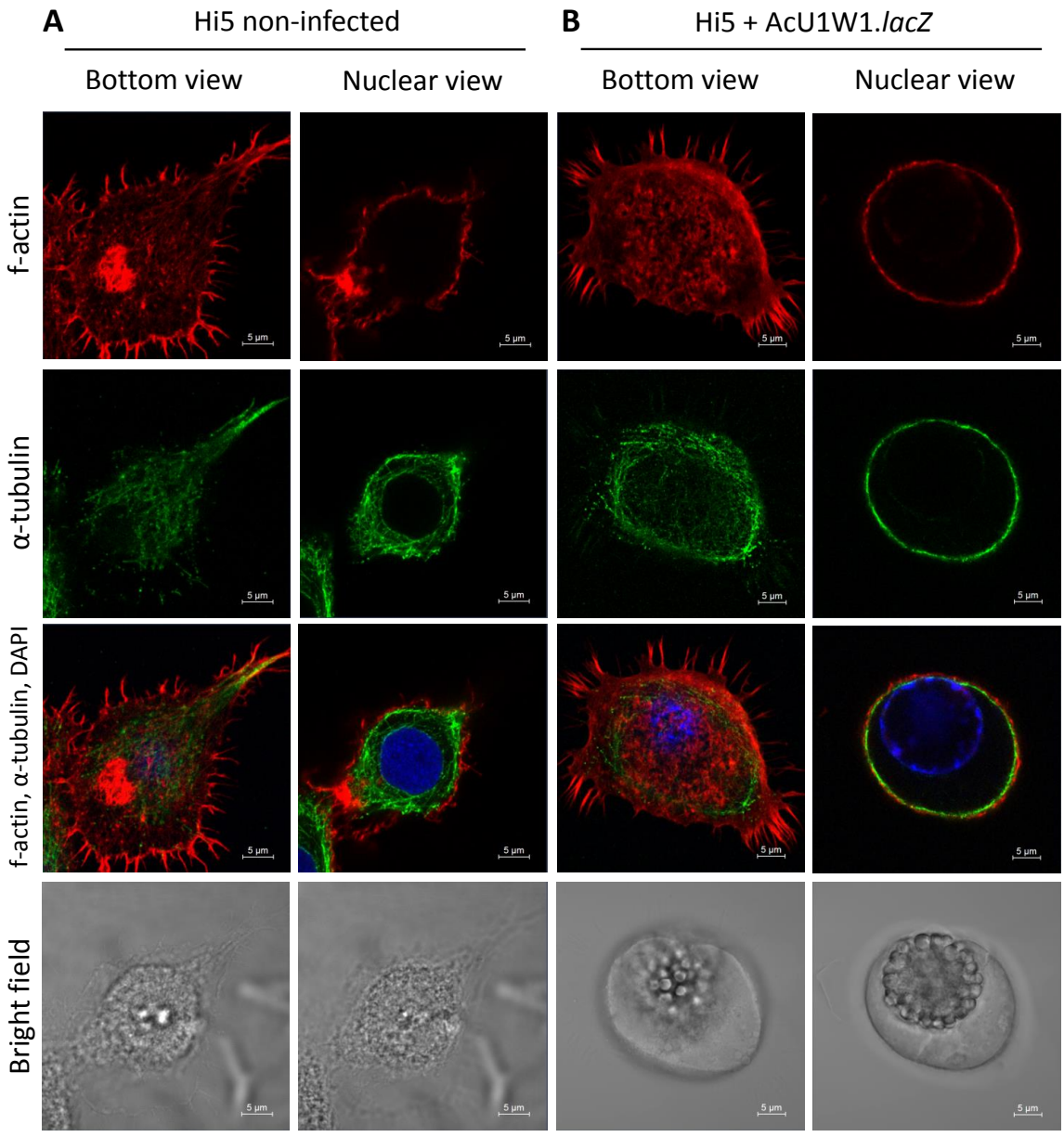
3.2.2 Analysis of the C20 cell cytoskeleton by Confocal Laser Scanning Microscopy

To further explore the morphology of C20 cells, the cytoskeleton was stained and observed with Confocal Laser Scanning Microscopy (CLSM). C20 cells, Hi5 cells non-infected and infected with AcUW1.*lacZ* were prepared for immunostaining as described in section 2.7.2. Alexa Fluor 594 Phalloidin conjugated was used to detect F-actin, anti-α-tubulin followed by anti-mouse Alexa Fluor 488 to detect microtubules and DAPI labelled the nucleus.

Figure 3.3A shows an example of a non-infected Hi5 cell. It is clear the attachment that these cells form, acquiring a flat shape against the coverslip where they were grown. At this focus plane, actin projections were observed using the bright field and red channel images. At the nuclear view, the complexity of tubulin structure was clear. On the other hand, when the cell was virus-infected (Fig. 3.3.B) this complex structure underwent major changes. The cells detached from the surface and acquired a more rounded shape, which was accompanied by shorter actin projections. The tubulin structure changed as it no longer created a net around the nucleus to support the cell but instead it mainly accumulated in a ring beneath the plasma membrane. The nucleus enlarged to

accommodate polyhedra that contain the new occluded virus ready to be released when the cell dies. This is clearly observed in the bright field and merged image (Fig. 3.3B).

The C20 cells demonstrated an array of structures (Figure 3C, D). For the majority of the cells, the cytoskeleton resembled the structure of its parental cell line, although the projections present were usually not as elongated as in Hi5 cells (Fig. 3.3C). Furthermore, some cells showed a more rounded up shape and shorter projections, which were clear signs of virus infection (Fig. 3.3D bottom left cell). Polyhedra were rarely observed and the actin and tubulin rings around the cytoplasm, as shown in Figure 3.3B, were not observed in C20 cells. Scanning electron microscopy images of virus-infected and non-infected Hi5 presented in chapter 4 show the exterior of the cell and corroborate the shape observed here for both Hi5 and C20 cells (see 4.2.5).



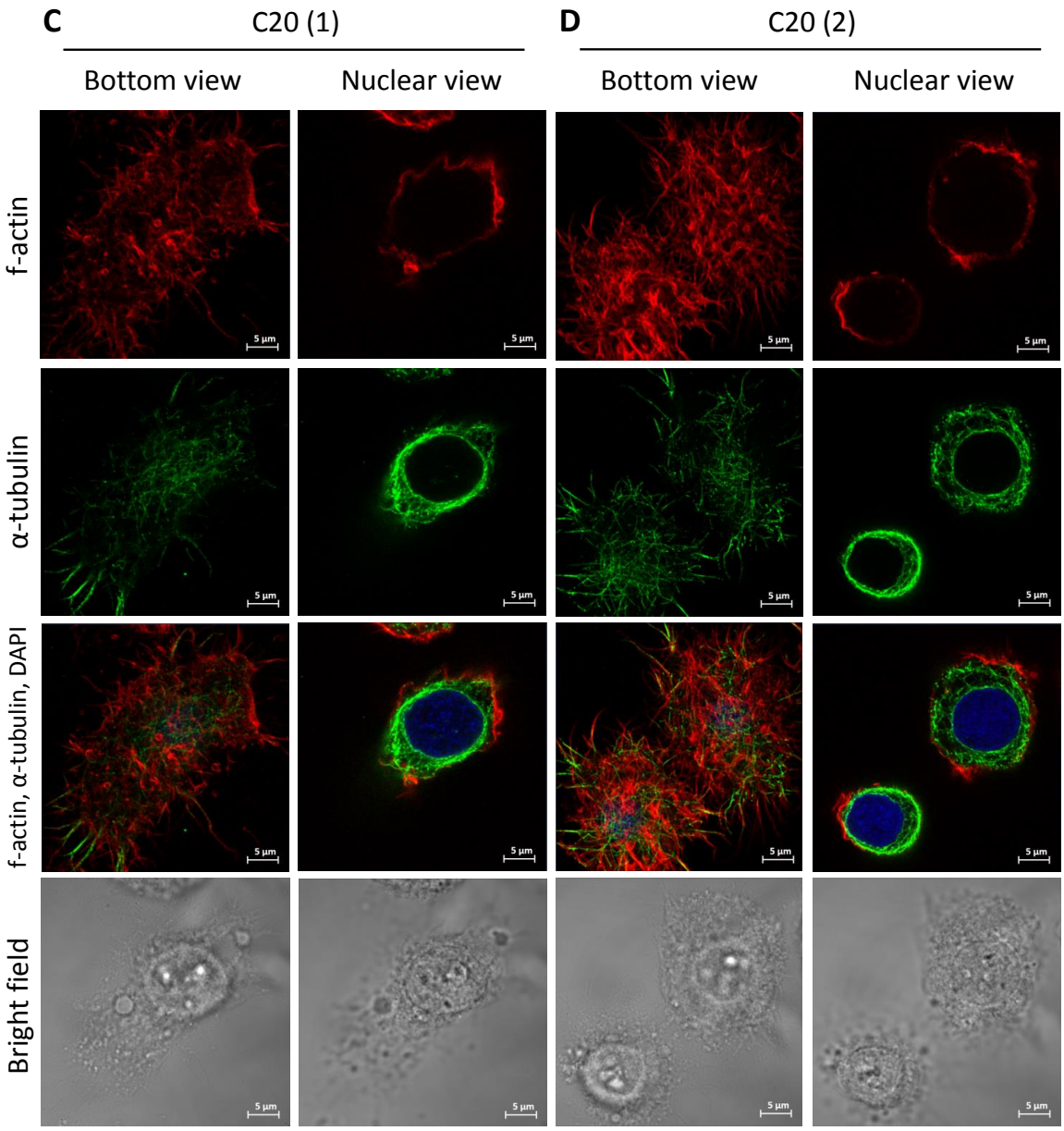


Figure 3.3: F-actin and α-tubulin structures in C20 and Hi5 cells
Hi5 cells A) non-infected or B) infected with AcUW1.*lacZ* and C-D) C20 cells were fixed with formaldehyde (at 96h p.i. in the case of virus-infected Hi5) and stained with: anti-α-tubulin and anti-mouse Alexa Fluor 488 for visualizing microtubules (green); Alexa Fluor 594 Phalloidin conjugated to detect F-actin (red); and DAPI to stain the nucleus by binding to double-stranded DNA (blue). A bottom view and nuclear view from the same cell are shown. For each panel, a merged image of the three fluorescent channels is present on the third row, with the last row showing the bright-field view. Representative images of replicate samples are shown. Scale bars are 5µm.

3.2.3 Assessing the growth kinetics and viability of the C20 cell line

C20 cells and Hi5 cells were established in suspension cultures at two different initial densities, $0.5x10^6$cells/ml and $1x10^6$cells/ml, and incubated at 28ºC. The experiment was repeated three times with two replicas of C20 cells from each initial seeding density. Aliquots of 0.2-0.5ml were taken daily to enable the cells to be observed and counted

with the light microscope until there were no viable cells in each flask. For each initial density, the three most diverse C20 growth patterns have been included in Figure 3.4, which shows mean viable cells (A, C) and percentage of viable cells (B, D).

The Hi5 cells rapidly initiated exponential growth, particularly when seeded at the higher density. These cultures reached their peak after 48h of incubation whereas those seeded at the lower density took 24h longer (Fig. 3.4C, A respectively). The viabilities of both cultures of Hi5 cells dropped dramatically after they had reached their peaks and no viable cells were observed from day four or six onwards, respectively, for high and low initial seeding densities (Fig. 3.4D, B). The culture of these cells was discontinued shortly after.

In contrast, C20 cells took longer for their growth to become exponential but after the number of cells peaked, between days five and eleven, the viability did not decline precipitously as in the Hi5 cell cultures. A careful observation of the high seeding density graphs (Fig. 3.4C, D) reflects the variability that was characteristic of different cultures of C20 cells. Three flasks seeded at the same density and incubated under the same conditions produced very different growth curves. The viability initially peaked for all the represented flasks at day five (Fig. 3.4C). After this, one of the flasks initiated a rapid decline and there were no viable cells observed after 8 days. On the other hand, the other two cultures survived up to 28 and 42 days respectively. For several periods, in between the second and fifth weeks, the number of live cells continuously increased even after a temporary decline was noted (Fig. 3.4C). This demonstrated they were still able to replicate even after weeks of culture. In these two cases, by the sixth day, more than 70% of cells were still viable (Fig. 3.4D). The viability dropped considerably after that but retained a fluctuating pattern until all the cells died.

All the C20 cell tested cultures were maintained for the period of the experiment without the addition of fresh culture media. The number and viability of these cells gradually reduced. It was of interest, therefore, to determine if the C20 cells that had been growing in an environment with declining nutrients and accumulating waste products, were able to divide and recover the viability of the culture if diluted into fresh medium (i.e. through passaging the cells). Two of the C20 replica cultures that still had live cells remaining 28 days after seeding, were diluted to $0.5x10^6$ cells/ml in fresh media

and incubated in suspension cultures. These batches are not represented in Figure 3.4 since almost all the culture was needed for passaging. At this point their viabilities were 9.5% and 17.3%. These cultures were observed with and without trypan blue staining on the day they were subcultured (Fig. 3.5A, B) and after three passages (14 days; Fig. 3.5C, D). Surprisingly, these cultures rapidly recovered their initial viability as observed by the proportion of cells taking up trypan blue before and after passaging (Fig. 3.5B and D, respectively). Unfortunately, no quantitative data was recorded from this experiment.

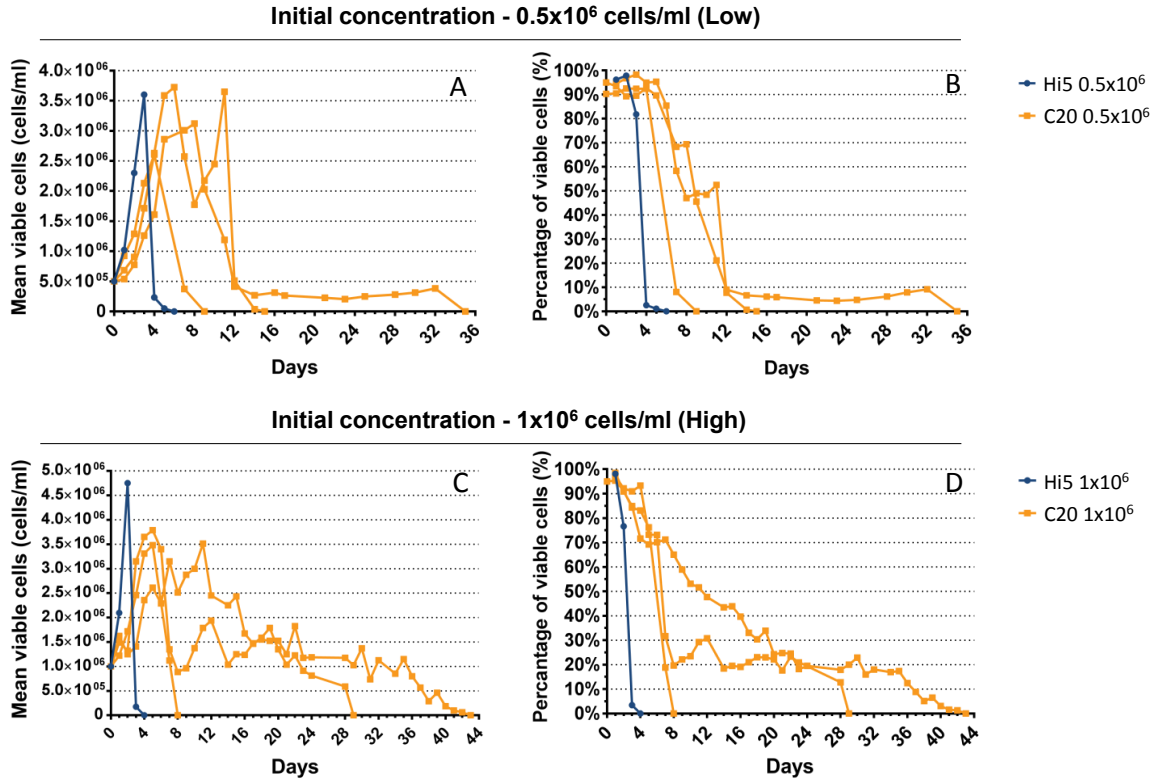


Figure 3.4: C20 and Hi5 growth curves
C20 and Hi5 cells were established in suspension cultures using two different seeding densities: low (0.5x10^{6}cells/ml; A, B) and high (1x10^{6}cells/ml;C, D). Cultures were incubated at 28°C until there were no viable cells remaining. For each initial seeding density, three C20 cultures are shown that represent the wide range of growth patterns observed. Daily samples were taken to quantify the number of total and viable cells per millilitre (counts were done in triplicate). A, C) Mean number of viable cells per millilitre and B, D) percentage of viable cells.

This study demonstrated that the C20 cell line can remain viable, albeit at a low level, for over a month without new nutrients being added and surviving in the accumulation of cellular toxic waste. Even after a long period of stress, the cultures were still functional and able to recover viability when passaged. This suggests that the C20 cells are much more resilient than the parental Hi5 cells.

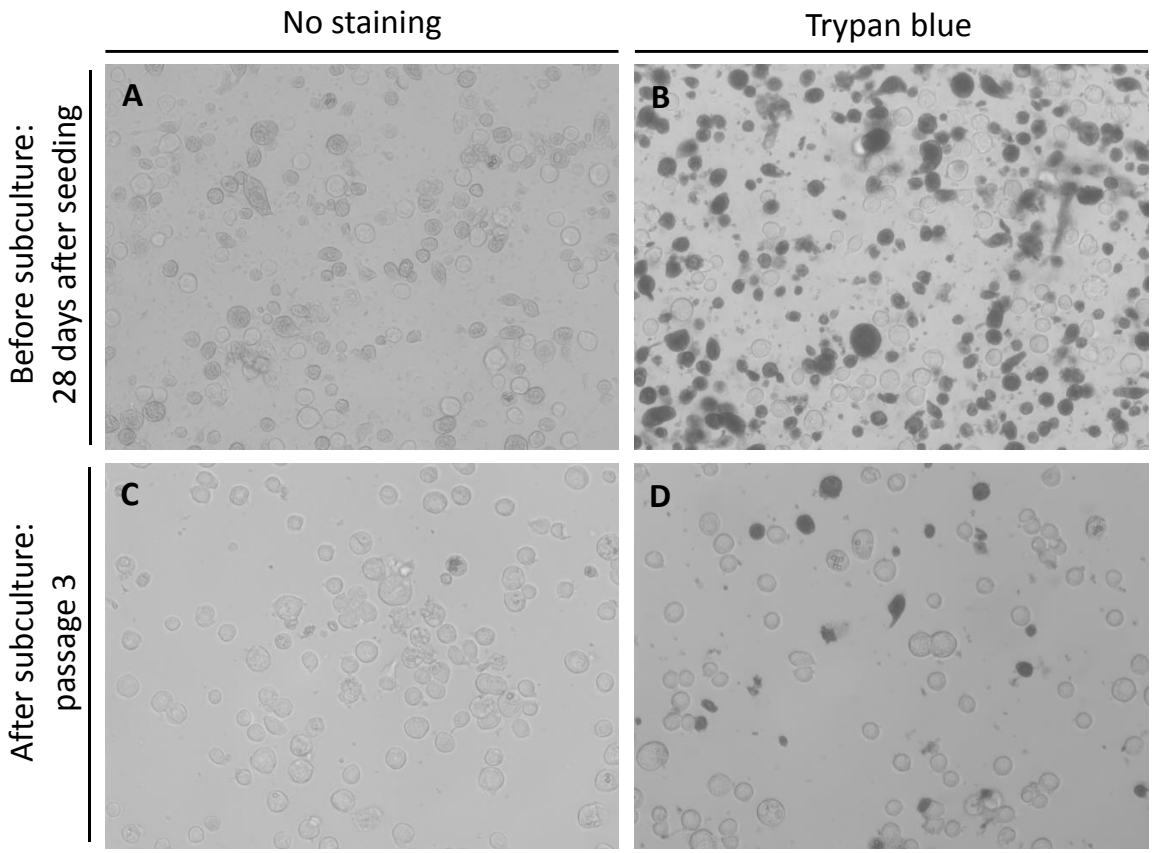


Figure 3.5: Long term viability of C20 cells in suspension culture.
C20 cells were passaged 28 days after being seeded in suspension/shaking cultures. Panels A and B show the cells on the day they were passaged. Panels C and D show the cells after three passages (14 days), when the viability was recovered. Panels B and D, cells were stained with trypan blue. Representative images are shown.

3.2.4 The presence of nodavirus in C20 cells

Previous studies have shown the presence of a latent nodavirus in Hi5 cells (Li *et al.*, 2007; Merten, 2007). Since C20 cells were created from Hi5 cells, it was important to verify if the nodavirus was still maintained in the presence of the baculovirus. As nodavirus contains two linear segments of positive-sense RNA, an experiment based on RT-PCR was designed to detect its genome (Fig. 3.6).

Total RNA was extracted from Hi5, C20, Sf9 and TN368 cells, and also from AcUW1.*lacZ*-infected Hi5 cells (RNeasy mini kit; see 2.4.8). Sf9 cells were used as a negative control since this cell line was shown not to be contaminated before (Li *et al.*, 2007).

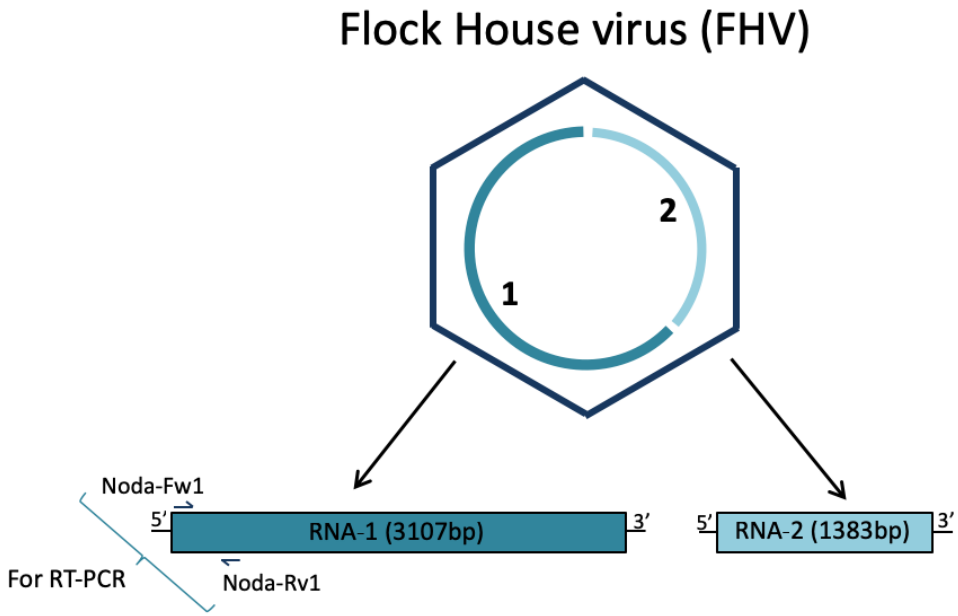


Figure 3.6: Schematic representation of FHV genome
Flock House virus genome comprises two single-stranded, positive-sense RNA molecules. Primer pair Noda-Fw1 and Noda-Rv1 (Table 2.4) was designed to amplify 544bp from the 5' end of the RNA segment 1 by RT-PCR.

Initial electrophoretic analysis of the RNA samples showed the presence of DNA in all of the samples tested (example for Hi5 and C20 in Fig. 3.7). Therefore, prior to the RT step, all samples were treated with DNase to eliminate any cellular or viral DNA. This generated cleaner preparations for RT-PCR therefore minimizing any non-specific reactions between primers and DNA (Fig. 3.7).

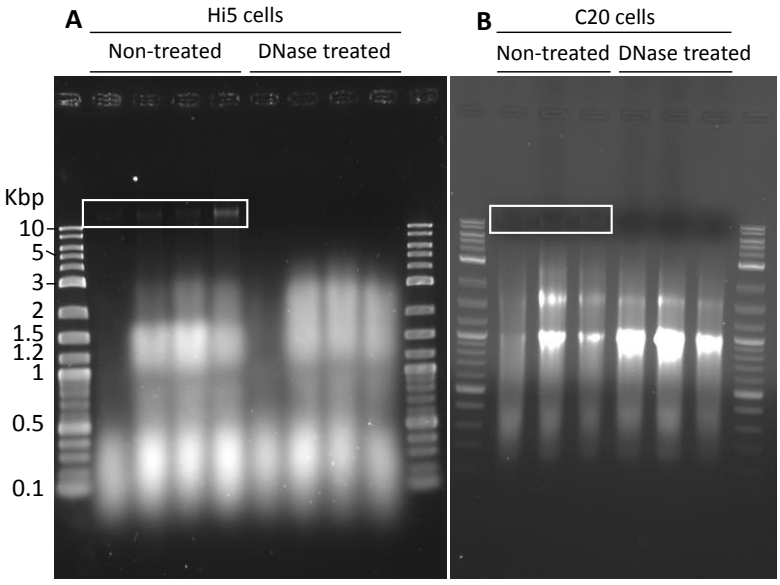


Figure 3.7: Treatment of intracellular RNA samples with deoxyribonuclease I.
Total RNA was extracted from replicate samples of A) Hi5 or B) C20 cells and analysed before or after treatment with DNase I. First and last lanes of each gel contain 2-log ladder. White boxes highlight the DNA bands presented in each gel.

A primer pair (Noda-Fw1 and Noda-Rv1; Table 2.4) was designed to amplify 544b from 5' end of the nodavirus RNA segment 1 as a double-strand DNA copy using SuperScript III one-step RT-PCR system (see 2.4.8). The results showed that both the C20 and Hi5 cell lines but not TN368 or Sf9, were infected with FHV (Fig. 3.8A, B). Futher analysis at the nature of the RNA would be needed to define if it represents the FHV genome or transcribed genes and therefore a covert or an overt infection.

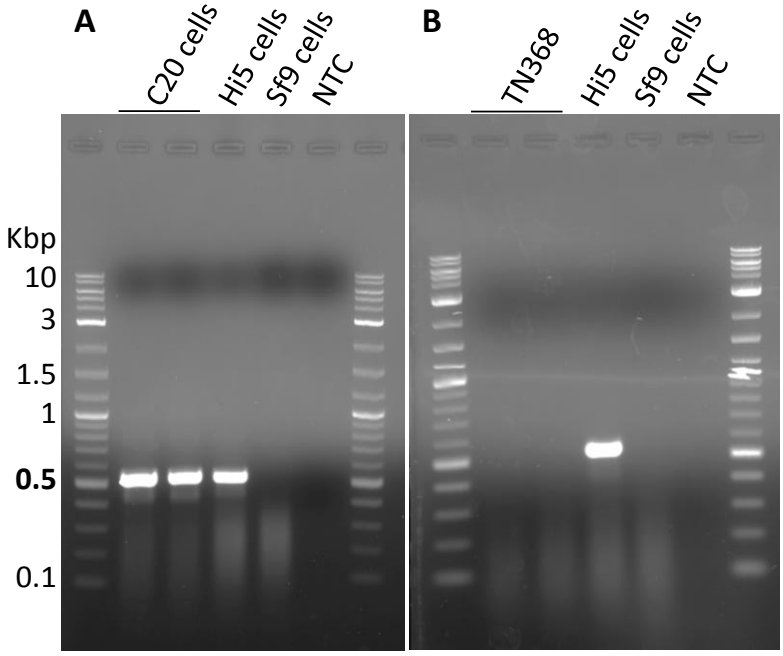


Figure 3.8: Detection of nodavirus in insect cells.
DNase I pre-treated RNA samples from C20 (A), Hi5, Sf9 (A,B) and TN368 cells (B) were tested by RT-PCR for the presence of nodavirus RNA using primers Noda-Fw1 and Noda-Rv1. NTC, non-template control. First and last lanes of each gel contain 2-log ladder.

3.2.5 Budded virus production by C20 cells over passages

To determine that C20 cells were persistently infected, medium form a number of passages was titrated to assess virus infectivity. The cell media were clarified using centrifugation to remove any remaining cells and then titrated for infectious virus using plaque assays (Fig. 3.9; see 2.2.2). The AcC20 infectious titres varied from $1.6x10^7$ to $7.8x10^3$pfu/ml, which although fluctuating between passages demonstrated an overall downward trend. This confirmed that infectious BV were produced in every C20 cell passage.

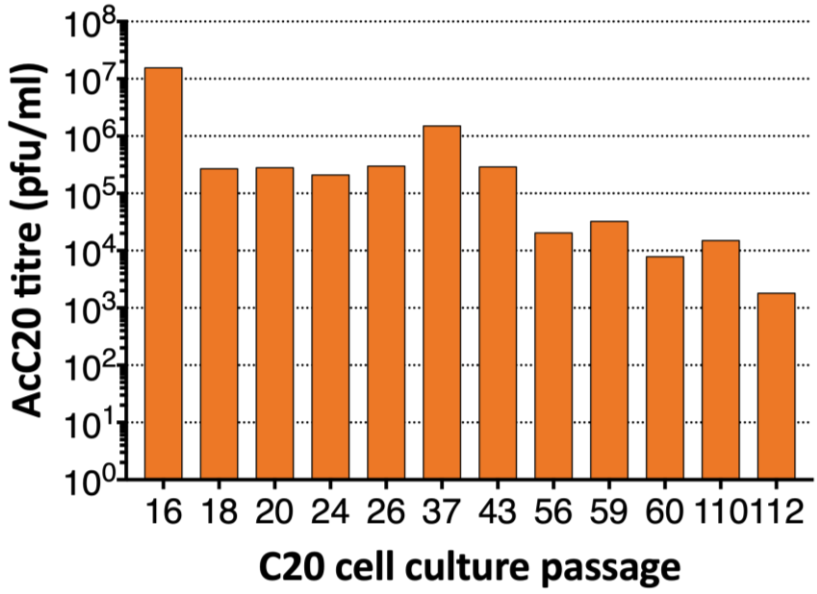


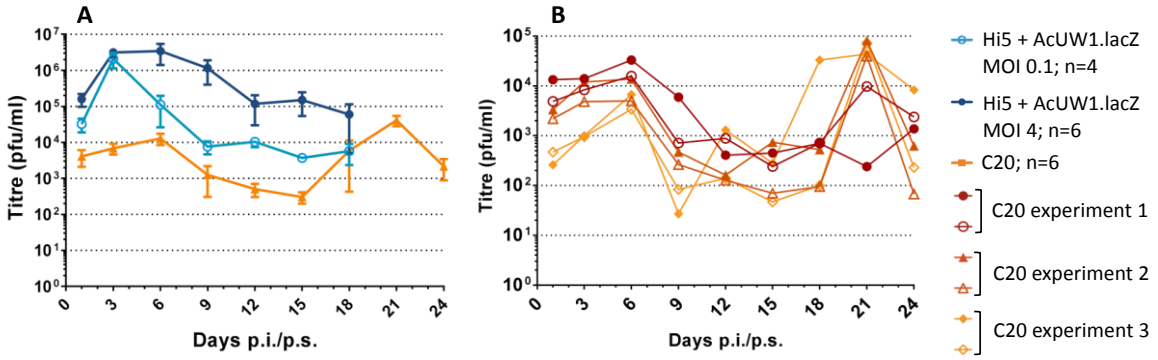
Figure 3.9: Infectious virus production in successive C20 cell passages
At different passages of the C20 cells, the cell media was clarified using centrifugation and used to titrate the virus by plaque assay. Titres in pfu/ml.

3.2.6 Analysis of virus replication in C20 cells over time

Figure 3.9 suggested considerable variation in infectious BV production by C20 cells in successive subcultures. To understand how AcC20 is generated within the duration of a single passage, replica monolayer cultures were established and then two flasks were harvested every three days up to 24 days post-seeding (p.s.) to collect both cells and medium by centrifugation. Control cells comprised Hi5 infected with AcUW1.*lacZ* at 4 MOI, which were collected every three days up to 18d p.i. Three independent experiments were performed.

3.2.6.1 AcC20 BV production over time in single cultures

Infectious virus in media samples from C20 or AcUW1.*lacZ*-infected Hi5 cultures was determined by plaque assay (Fig. 3.10a). In Hi5 infected cells, the titre followed a normal course of infection. Viral titre production increased over the first 3 days but then decreased as the cells started dying of the infection. This decrease was not very accentuated since accumulated virus in the media was measured.



Cumulative titre in pfu/ml of C20 cells (n=6) and Hi5 infected at 0.1 (n=4) or 4 MOI (n=6). A) Average of the biological replicas from three independent experiments (1,2 and 3) over 24 days, shown separately in B for C20 cells. Error bars represent standard error of the mean (SeM).

For C20 cells, the titre fluctuated over time as demonstrated in the independent flasks from each experiment (Fig. 3.10B). Although the patterns are different over time, in all of the flasks the titre has increased on the first few days and peaked 6 days after the cells were seeded, which represented an increase of ~10x from the initial titre. This was followed by a decrease in viral titre reaching a minimum around 12-15 days after seeding. In five out of six cultures the titre peaked again after 21 days. This late peak gave the highest titre overall (4 to 8.1x10^4pfu/ml) in four of the cultures. This indicates that over time the production of AcC20 is not uniform in a single flask.

3.2.6.2 Analysis of relative intracellular AcC20 DNA over time in single cultures

a) Establishing a multiplex-qPCR assay for relative intracellular viral DNA analysis

A qPCR was designed and optimized to examine virus DNA replication in C20 over the course of the growth of cells in a single culture passage. To enable more accurate results TaqMan™ probe chemistry was chosen. Since the probes can have fluorescent dyes with different emission spectra attached, this technology enables amplification of endogenous controls and genes of interest in the same reaction tube (multiplex qPCR), thus minimizing technical errors and false negatives. It is also highly specific since only targeted amplification products are detected and there is a very high signal-to-noise ratio. It minimizes reagent costs and starting sample by requiring only one reaction per sample.

For this particular experiment, *gp64* was chosen to measure viral replication since GP64 glycoprotein is reported as the limiting factor for production of BV (Oomens and

Blissard, 1999). As an housekeping gene in lepidopteran insects, *actin* was used as the endogenous control gene (Teng *et al.*, 2012).

The *actin* primer pair and probe were designed paying close attention to the following details to enable a highly efficient reaction: a) the secondary structure by avoiding hairpins, self-dimers and hetero-dimers tested using OligoAnalyser 3.1 (Integrated DNA Techologies, 2018); b) keep the proximity between the forward primer and probe; c) keeping melting temperature relatively high (50-65°C); d) ensuring a GC of 50-60% and low number of nucleotide repeats. For gp64, primers and probe (Table 2.4) were designed by Carina Bannach (Bannach, 2018). *Actin* and *gp64* primer pairs and probes were optimized in singleplex and multiplex reactions in non-limiting conditions. To develop a highly sensitive qPCR, the efficiency of the reaction was tested at different temperatures to ensure: a) a linear standard curve [Coefficient of determination R^2>0.98, measured from the Person's correlation coefficient (r) (Bio-Rad, 2006)]; b) an efficiency between 90-105% for each reaction; c) the efficiency for endogenous and target gene are identical when multiplexed (difference not greater than 5% and a slope <|0.1| of the ΔC_t across dilutions; d) a difference between the singleplex and multiplex reaction's efficiency in each gene is not greater than 5%, proving the sensitivity of the amplification is not compromised by the presence of another reaction occurring with the same reagents.

A standard PCR showed the designed primer pair correctly amplifying a short segment of *actin* (data not shown). Therefore the qPCR was optimized as follows.

Using a 1:5 dilution factor, standard curves for *actin* and *gp64* reactions in multiplex were performed with 0.004ng to 12.5ng of DNA per reaction. The efficiency at different temperatures was calculated according to the formula $E_{(\%)} = (10^{-1\text{-slope}}–1) \times 100$, in independent experiments (Fig. 3.11). At 59°C the highest efficiency was achieved for each gene (94.16% for *gp64*, 97.07% for *actin*). At 60°C the closest efficiency between the two genes was achieved (92.09% for gp64, 93.31% for actin). At 61°C although it was as close, the lowest dilution was not detected, therefore 59°C and 60°C were the chosen temperatures to continue the optimization. Both these temperatures generated a difference not greater than 5% in between the reactions efficiencies.

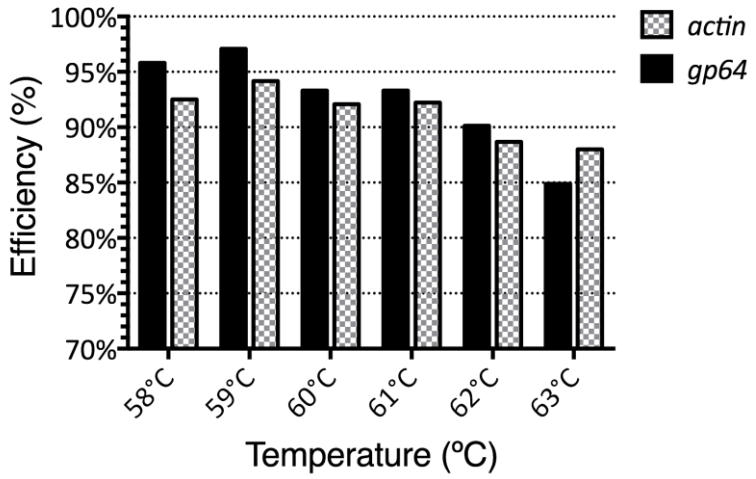


Figure 3.11: Efficiency test for qPCR.
Efficiency for *actin* and *gp64* reactions tested in multiplex reactions at 58°C to 63°C. Calibrator sample DNA diluted 1:5 from 12.5ng to 0.004ng. Technical triplicates were averaged before generating the standard curve. From the curve the slope and efficiency for each reaction were calculated. E(%) = $(10^{-1/slope}-1) \times 100$.

The next step consisted of confirming if the efficiency for endogenous and target genes were identical by calculating the relative efficiency (from the ΔCt). By plotting it, the absolute slope obtained was 0.05 and 0.02 for 59°C and 60°C respectively, therefore they are both <|0.1| (data not shown). Since 60°C had the lowest absolute slope and because it is good practise to work with the highest temperature that fits the requirements to enable more specificity, this temperature was chosen. The last test was comparing the multiplex reactions with the equivalent singleplex ones to see if the sensitivity of either of the reactions was compromised in the multiplex. The efficiency for the singleplex reactions were 92.08% and 98.06%, respectively, for *gp64* and *actin*. While for *gp64* the efficiency was very similar (92.09% when multiplexed), for *actin* the difference was bigger (93.31% when multiplexed). However, it was still within the 5% limit so this reaction was considered optimized and was used for the relative quantification of *gp64* and viral replication analysis.

b) Relative viral DNA replication analysis by multiplex-qPCR

From the concentrated cells, one replica per cell line and experiment was used to extract DNA. From the second replica, the volume of cells was split into two. While one half was used to extract DNA, the second one was prepared for protein analysis by Western Blot (Results 3.2.6.3). Therefore, DNA from each time point, cell line and experiment, was

available and used for relative quantification of *gp64* by qPCR, using *actin* Hi5 as the endogenous control.

Figure 3.12 shows relative *gp64* quantification by qPCR. If the replication cycle is completed as expected, an increase of *gp64* copy number correlates to an increase of BV since these are directly proportional; each virus genome has only one copy of each gene. The $2^{-\Delta\Delta Ct}$ method means the number of *gp64* copies has been standardized against the number of cells present. The largest increase in relative quantity of *gp64* occurs from day one to three in Hi5 cells infected with virus at either high or low MOI (Fig. 3.12A). Given that after 72h p.i. Hi5 cells are dying of the infection, this clarifies why the rate of relative BV production decreases. On the other hand, AcC20 relative intracellular DNA fluctuates over time without a specific pattern. When analysing Figure 3.12B, representing the relative quantity of *gp64* in the different biological replicas of C20, these do not fluctuate as much as when the BV titre was analysed per se (Fig. 3.10B). The difference between any two time points of the same culture rarely reached 10-fold (3.12B). Nonetheless this experiment showed there was a balance established between the C20 cells and AcC20 persistent virus since the fluctuation in the relative quantity of *gp64* was not as accentuated as in experiments that only dealt with the titre or number of cells independently (Fig. 3.10A, 3.10B and 3.4).

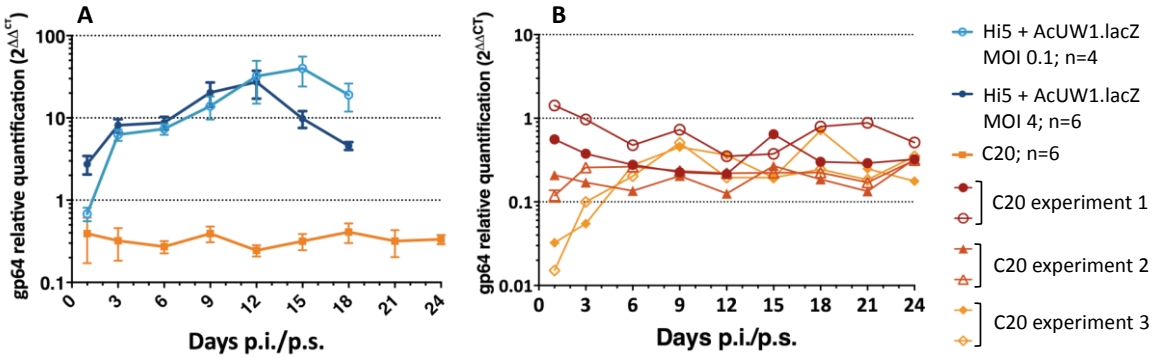
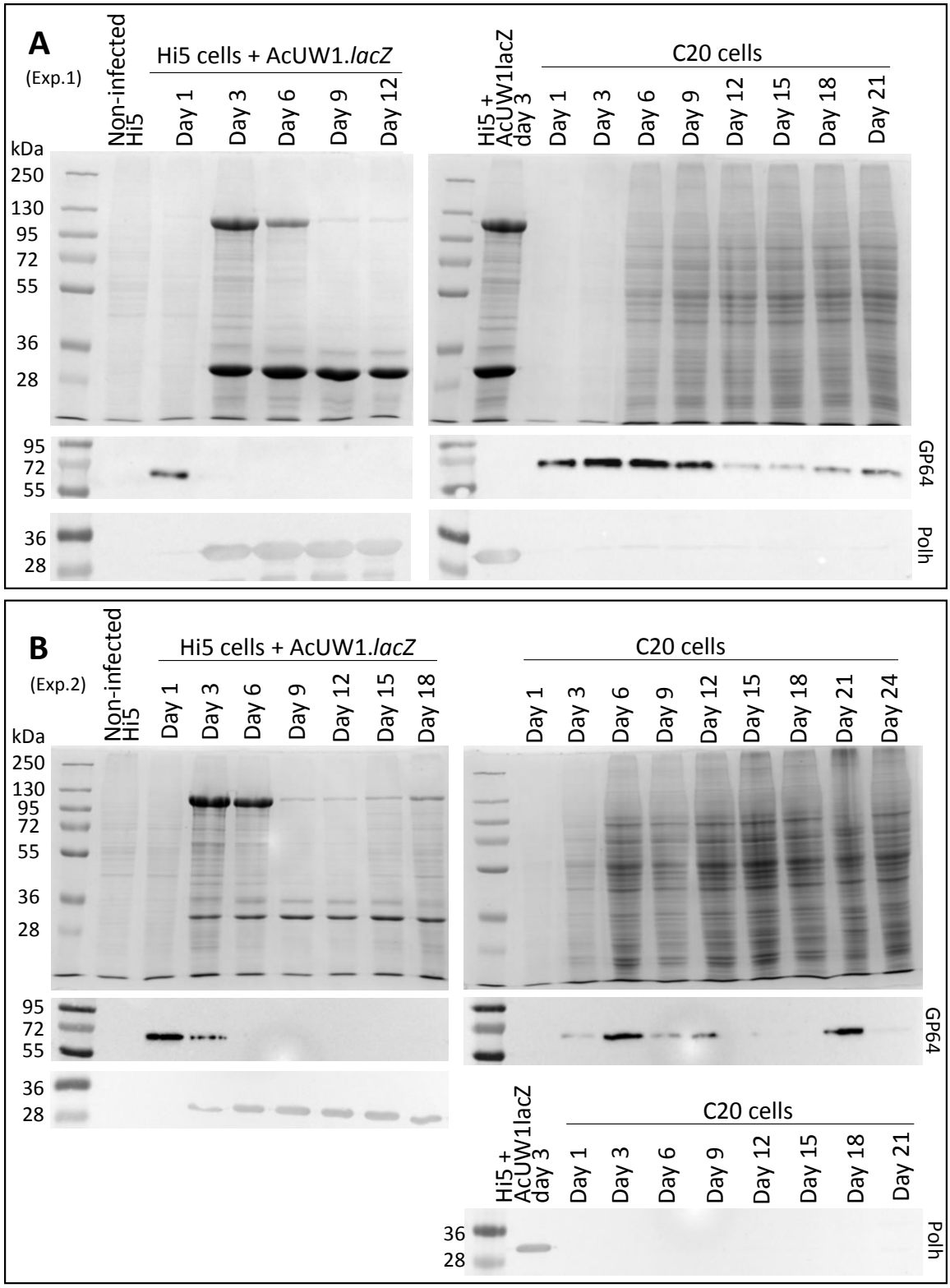


Figure 3.12: Relative cumulative intracellular AcC20 DNA over 24 days
Relative cumulative intracellular *gp64* quantification ($2^{-\Delta\Delta Ct}$) of C20 cells (n=6) and Hi5 infected at 0.1 (n=4) or 4 MOI (n=6). A - Average of the biological replicas from three independent experiments (1,2 and 3) over 24 days, shown separately in B for C20 cells. Error bars represent standard error of the mean (SeM).

3.2.6.3 Viral protein production over time in single cultures

Figure 3.13 shows the protein analysis of C20 and AcUW1.*lacZ*-infected Hi5 cells. The synthesis of these proteins was similar in Hi5 infected with a 0.1 or 4 MOI. The latter was chosen as the representative for the figure since most of the studies presented in this

project were done using a high MOI. Coomassie stained gels were used to observe the total protein and as a loading control for the western blots. Anti-GP64 and anti-Polh immunoblotting were performed at the same time. While Hi5 infection with AcUW1.*lacZ* served uniquely as a positive control, it was possible to observe the normal pattern of a lytic infection.



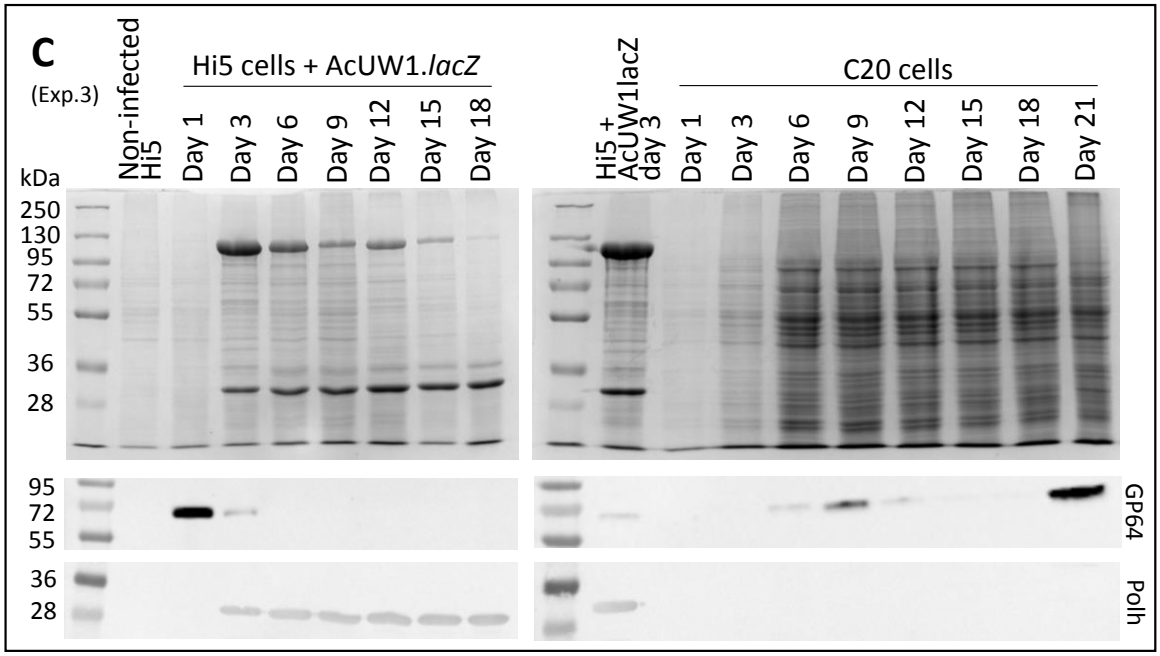


Figure 3.13: Time course analysis of cellular proteins from C20 cells and virus-infected Hi5 cells
Hi5 and C20 replica monolayers where seeded at the same time and Hi5 cells were infected at 4 MOI. At 24h p.i. and then every three days 2 replicas were harvested. Panels A, B and C represent independent experiments with one replica per time point and cell line of each experiment. Proteins were separated by a 10% SDS-PAGE and analysed by Coomassie staining. Simultaneously anti-GP64, anti-Polh immunoblotting were performed after protein transfer to a nitrocellulose membrane. Loading quantities were as follow: Coomassie staining, 5µl per sample; GP64 immunoblotting 10µl per sample. Polh immunoblotting 10µl for C20 cells, 5µl for Hi5 infected cells samples.

Two main proteins are particularly overexpressed from 72h p.i. onwards in the AcUW1.*lacZ* infection: Polh (~29 kDa (Ji *et al.*, 2010)) and β-galactosidase (116.482 kDa). However while Polh was strongly present from day 3 onwards, β-galactosidase was particularly evident at day 3 and 6, and from day 9 there was a decrease of this protein on the stained gel. This pattern was similar in the three independent experiments (left gel of Fig. 3.13A, B and C). Western blots for Polh confirmed it was expressed from day 3 onwards. GP64 immunoblotting revealed this protein was strongly present at 24h p.i. and there is still some accumulated produced protein after 3 days but none of the three experiments showed any GP64 production after that [GP64 ~59kDa (Hohmann and Faulkner, 1983)].

In C20 cells a very different pattern was observed. It was noticeable from the Coomassie stained gels the absence of β-galactosidase. This was expected since C20 cells treated with X-gal failed to produce a blue colour (R. D. Possee, personal communication). Therefore the *lacZ* gene present in C20 cells might have been lost or compromised (discussed in Chapter 5). Polyhedrin was not observed accumulating along the infection

at either time point. This was confirmed by immunoblotting for this protein, which indicated an absence of Polh in any of the 3 experiments, even when these were developed for extended periods. This generated a non-specific line around the size of Polh (Fig. 3.13A, B, C, Polh panels for C20 cells), which also happened in Hi5 cells infected with AcuW1.*lacZ* (day 1, experiment A). Surprisingly, GP64 was present throughout the 21 days of the C20 culture although it accumulated differently in the three experiments (Fig. 3.13, GP64 panels for C20 cells).

These experiments (section 3.2.6) demonstrated that the AcC20 virus was present at all times in C20 cells, reflecting the presence of a persistent infection. The three protocols (3.6.1, 3.6.2 and 3.6.3) were applied on samples coming from the same cultures so it is possible to see that over time in culture, there are new AcC20 genome copies being produced, new BV, and viral proteins being expressed.

3.2.7 Quantifying the proportion of C20 cells producing infecting virus

From the previous results it is clear that C20 cells produce infectious virus, at every passage. Individual cultures of cells also exhibited fluctuations in levels of virus throughout the duration of their propagation and viral proteins were also present. However, the titre was always lower than in a lytic infection, which indicates either virus was not produced at a high level or that not all the cells were productive. It was therefore important to quantify what was the proportion of C20 cells producing infectious virus. This was tested using a modified version of a plaque assay (3.2.7.1) and a flow cytometry protocol for GP64 (3.2.7.2).

3.2.7.1 Proportion of C20 producing infecting virus using a modified plaque assay

A method was devised to determine the proportion of C20 cells that were producing infectious virus. In principle, it consisted of mixing either C20 or AcUW1.*lacZ*-infected Hi5 cells with reporter Sf21 cells to detect virus produced by their companion cells in a plaque assay format.

C20 monolayers were washed 6 times before the cells were harvested and counted. Different quantities (between 500 and 10000) of C20 cells were mixed with $0.4x10^6$ Sf21 cells and allowed to attach to cell culture dishes for 1h. At the same time Hi5 cells were infected with 5 MOI of AcUW1.*lacZ* as a positive control and incubated for 4h. This

longer incubation time was to take into account that some nucleocapsids transit the cell and bud out of the membrane without undergoing replication (Washburn *et al.*, 2003b; Chen *et al.*, 2013; Rohrmann, 2013). Therefore an exhaustive wash after this step would help to bring the background virus in the media closer to zero.

Different quantities of virus-infected Hi5 cells (10-500) were mixed with Sf21 cells as was done with C20 cells and incubated for 1h. From the stock of harvested cells used to make the dilutions, an aliquot was centrifuged for media clarification and filtered before used to perform a control plaque assay. This determined the residual virus in cell culture media after washing. After this incubation for both cell lines and control media with Sf21, the media/inoculum was removed and the standard plaque assay procedure was followed (see 2.2.2). The experiment was incubated for five days for both cell lines and controls and repeated 3 times.

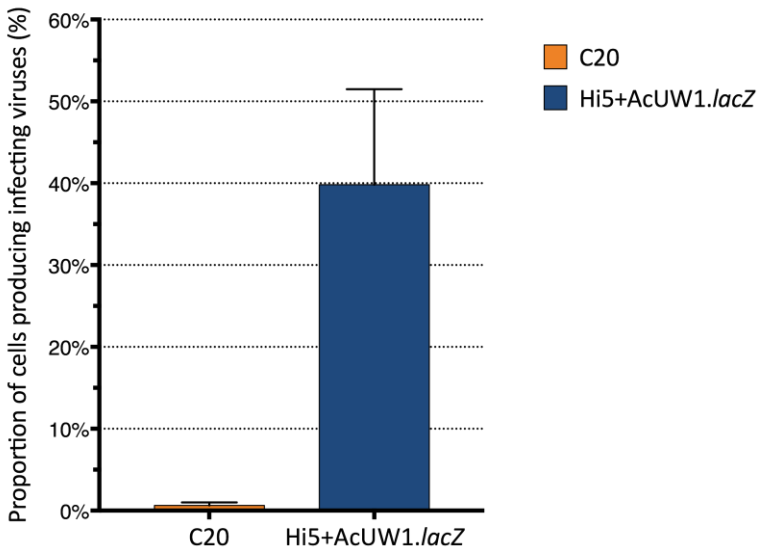


Figure 3.14: Proportion of C20 cells producing infecting virus
A modified version of a plaque assay was performed by mixing C20 and infected-Hi5 cells (5 MOI) with Sf21 cells. The background virus presented in the media for each cell line was accounted for and deducted from the final results. The percentage of cells producing infecting virus was calculated for both cell lines. Graph represents average and SD (error bars) of three independent experiments.

The plaques in each of the wells were counted and from the initial quantity of cells used it was possible to calculate how many were producing infectious virus. To these, the background virus was discounted as calculated from the control plaque assay on the media for each cell line. The percentage of C20 cells producing virus was very low, 0.64% on average (Fig. 3.14). However, although 5 MOI was chosen for Hi5 cells to ensure all

the cells were infected, only 39.81% of cells were apparently producing infectious virus according to this experiment. Assuming all Hi5 cells should be infected when using 5 MOI, 39.8% is also the efficiency of the experiment. Therefore one can possibly assume that 1.61% of the C20 cells are effectively producing virus. Nonetheless, this shows the experiment was not sensitive enough to have an accurate quantification of the proportion of cells producing infecting virus.

3.2.7.2 Proportion of C20 producing infecting virus by immunofluorescence flow cytometry

A flow cytometry protocol was designed in order to distinguish virus-infected from non-infected cells (see 2.6). In order to do this, C20 cells as well as Hi5 cells non-infected or infected with AcUW1.*lacZ* (5 MOI) were seeded and fixed 18h after. Cellular GP64 was then stained by immunofluorescence (Alexa Fluor 488) and the nucleus stained with DAPI. To distinguish cells from other events in the sample, the cases were gated by the presence of blue fluorescence (DAPI) and then gated by size. The non-infected cells were used to define the threshold between negative or positive cells. The presence of fluorescence reflected the proportion of cells expressing GP64. When this criterion was applied to Hi5 cells infected with AcUW1.*lacZ*, 78.6% were fluorescent, therefore, almost 80% of the cells were infected after 18h. On the other hand, only 7% of C20 cells demonstrated the presence of GP64 by this method (Fig. 3.15).

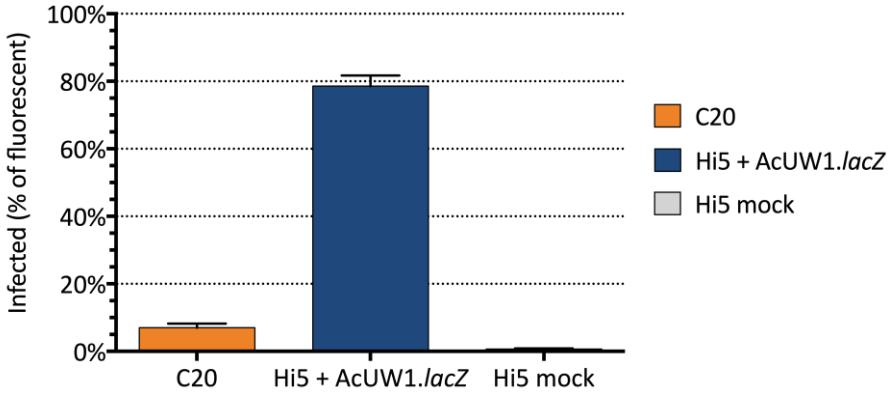


Figure 3.15: Immunofluorescence flow cytometry in C20 and AcUW1.*lacZ*-infected Hi5 cells.
C20 cells (passage 156) were fixed and GP64 immunolabeled with Alexa Fluor 488. As controls, Hi5 cells non-infected and infected with AcUW1.*lacZ* at 5 MOI at 18h p.i. were treated at the same time. All cells were marked with DAPI. Flow cytometry events were gated by blue fluorescence (DAPI), size and single cells. Analysis of green fluorescence (Alexa Fluor 488) in non-infected Hi5 cells defined the negative/positive threshold for the presence of GP64. Error bars represent SD in between 3 biological replicas.

3.3 Discussion

In this chapter a basic characterisation of the C20 cell line was undertaken. This cell line was generated by infection of Hi5 cells with AcUW1.*lacZ* in 2011. This generated a persistent infection which was cloned to isolate a clonal cell line, the C20 cells. Therefore throughout the chapter several comparisons were made with the parental Hi5 cell line non-infected or infected with AcUW1.*lacZ*.

It was noticeable that C20 cell morphology differed from the parental cell line. Enlarged or irregular-shaped cells were often observed and occasionally cells produced polyhedra, a clear sign of infection. The generally less healthy appearance of the cells and the presence of vacuoles also suggested they were infected with virus. Experimental work confirmed C20 cells are in general larger than Hi5 cells and present a higher variability of sizes (Fig. 3.2). This differs from the persistent infection established by McIntosh et. *al* in 1981 where the generated persistently infected Sf21 cell line presented rounder and smaller cells than the original cells (McIntosh and Ignoffo, 1981). Insect cells usually attach strongly to the growth surface, creating extended protuberances. This makes the cells look larger when grown in a monolayer and it is extensively observed in Hi5 cells. Our data suggested the diameter of the C20 cells significantly increased when they were grown in a monolayer, a consequence of the attachment to the surface that was however, not as accentuated as in Hi5 cells (Fig. 3.2). This was also observed using confocal microscopy where in C20 cells, shorter projections of F-actin and α-tubulin or their absence was observed (Fig. 3.5; data not statistically analysed). Furthermore, a big variability of sizes was present and the cells appeared more spherical and loosely attached to the growth surface in comparison to the parental Hi5 cell line.

The viability of C20 cells was consistently slightly lower than that of Hi5 cells. This might also have affected the results for the observation of the cytoskeleton. Some of the cells did not look as healthy as others with noticeable rounding up, a normal process observed before the cells die. In future work, non-viable cells would need to be marked for confocal microscopy studies, so they are not taken into account in the interpretation. Nonetheless, the variability present in these experiments was the first indication that we might be facing a mixed population.

Surprising results were observed for the positive viral infection control used for confocal microscopy. Hi5 cells infected with AcUW1.*lacZ* by 96h p.i. showed a ring of F-actin and microtubules around the cytoplasm. Previous studies have shown F-actin is essential for nucleocapsid morphogenesis (Ohkawa and Volkman, 1999; Volkman and Kasman, 2000) and acts as a regulator of chromatin remodelling by interacting with P6.9 (Volkman, 2007). At 24h p.i. actin localizes in the nucleus and in the late phase of infection F-actin filaments together with VP80 penetrate and link the virogenic stroma with the nuclear periphery (Marek *et al.*, 2011). In the current study only the very late phase of infection was observed (96h p.i.) therefore the observations from the late phase might not apply. However, being F-actin filament necessary for the maintenance of the virogenic stroma, it was surprising that at 96h p.i in a nuclear view no F-actin was observed around or inside the nucleus. Instead it formed a ring around the cytoplasm that was not described before to my knowledge. However, actin was recently suggested to be essential for virus movement to the plasma membrane (Ohkawa and Welch, 2018). From the bright field images, a structure was observed around the very late phase enlarged nucleus. P10 forms a cage-like structure around cell nucleus that aids in nuclear lysis (Graves, 2016); however AcUW1.*lacZ* lacks *p10* and has *lacZ* instead. Therefore this unknown structure might be in a very late phase of infection preventing F-actin and microtubules to reach the cell nucleus.

By looking at the growth kinetics and viability of C20 cells it was noticeable that they are much more resilient then Hi5 cells. They can survive for up to 42 days without being subcultured in a nutrient-exhausted medium that probably accumulates toxic waste. While it is important to note that not all the tested batches survived that long, the fact that most of them survived for over a month is an achievement not documented before to our knowledge for any persistent infected cell line. The C20 cell line is likely a result of a delicate balance between an infecting virus and the host cells, including its defence mechanisms. Therefore a great deal of variability is naturally inherent to the cell line. Furthermore, a distinct growth pattern is observed from the parental cells. The growth curves demonstrated that in C20 cells and contrary to Hi5 cells, once the maximum number or viable cells is reached, due to high cell density and media exhaustion, there are several cycles of decline and recover of viability over time, before the cells die. We hypothesize that the AcC20 might be playing a role in this growth pattern since it could

be affecting some of the cell defence mechanisms including apoptosis, enabling them to survive for longer. In agreement with the first persistent infection described in baculovirus, C20 cells have a longer doubling time then the parental cell line although they did not seem to be able to grow to a higher density as it was described before (McIntosh and Ignoffo, 1981).

Section 3.2.4 showed that C20 cells along with their parental cells present a latent infection by FHV. However, TN368 and Sf9 cells, also used in our group in the same physical spaces, are free of nodavirus infection. This will be an important variable to consider throughout the thesis, since it might be influencing the balance between the C20 cells and the AcC20 persistent infecting virus as well as C20 cell behaviour when superinfected (Chapter 4). It would be interesting to understand if this latent infection has a role in the establishment of a persistent infection by, for example, determining if a *T. ni* cell line lacking nodavirus can develop this type of infection with a baculovirus.

After this close characterisation of the C20 cells the attention was switched to the AcC20, the virus persistently infecting these cells. It was observed that the virus was present in all of passages of the culture (up to 112) that were tested. This is probably the longest persistent infection described *in vitro* reported to date. Nevertheless, Figure 3.9 shows a downwards tendency for the virus titre over time.

A careful study of the AcC20 in culture over a period of 24 days was performed combining measurements of BV cumulative production, relative intracellular DNA quantification by qPCR and protein production at regular time points. Analysis of the cumulative virus titre showed virus replication occurs in C20 cells although not at a steady rate, since the cumulative titre fluctuated over the course of the experiment in all the replicates. Also, AcC20 never reaches very high titres as is observed in a lytic infection of Hi5 cells ($1.2x10^7$ pfu/ml in this experiment at 4 MOI). It is important though to remember that C20 cells keep producing infectious virus even if not passaged regularly.

Relative intracellular DNA quantification showed once more that the production of BV varies over time in C20 cells. However results for Hi5 infected cells produced a different pattern then when analysing the virus titre alone and the variation in the relative BV production by C20 did not look as accentuated as when the BV production was analysed

alone (fig. 3.10B). This is a consequence from the $2^{-\Delta\Delta Ct}$ method used which provides relative quantification of *gp64* standardized against the number of cells present, reflected by the *actin* presence. Therefore the variation of number of cells and BV production are both considered to the final result. From Figure 3.4 it was clear that the cells suffer cycles of more and less intense growing. Therefore a peak in the BV production alone as was observed at 21 days for most of the replicas (Fig. 3.10B) might be a consequence of an increase of the number of cells that all together generate more BV, and hence will not be outstanding in a relative quantification. For the same reason, the less accentuated drop in relative BV production observed in Hi5 cells might be caused by the increased mortality of Hi5 cells at latter time points and therefore the proportion of relative *gp64*/*actin* does not drop as much as the BV production alone. This will also justify the lower variability in the relative quantification of BV production by C20 compared to the variability of the titre over the same period. These results suggest that an increase in the BV production is intrinsically connected to an increase in the number of C20 viable cells and vice-versa although further work needs to be done to confirm this. For these reasons and because the detection of gene replication occurs before the detection of BV by plaque assay comparisons need to be made carefully and the fluctuation patterns observed in Figures 3.10 and 3.12 cannot be compared.

The same experiment analysed the protein profile and accumulated production of GP64 and polyhedra over time. This showed viral proteins are being produced in C20 cells over time although the protein profile does not seem to change considerably and is similar to non-infected Hi5 cells. Polyhedrin was not detected by immunobloting in C20 cells, which was expected by the lack of observed polyhedra in the cells when microscopy techniques are applied (Fig. 3.2). On the other hand, GP64 was present in the three independent experiments performed and confirmed the fluctuation previously observed in the other two methods. The results for virus-infected Hi5 cells corroborated literature results since there was a clear shift observed from GP64 to polyhedra production characteristic from the shift from late to very late phase of infection (Rohrmann, 2013). The *lacZ* seems to have been lost as reflected by the lack of β-galactosidase in the stained gels from C20 samples over time. This result was not surprisingly since when challenging C20 cells with X-gal the media does not achieve the characteristic blue colour (R. D. Possee personal communication).

These three experiments from the long-term culture samples confirm AcC20 is constantly present and being produced in the C20 cell line, demonstrating we are facing a persistent infection. However, considering the longevity of the culture and the low titres observed it was hypothesised that currently not all the cells were infected at all time points. Two tests were developed to test the proportion of infected cells at each time point. There were based on the production of infectious BV (plaque assay-based test) and the presence of the GP64 protein in the cells (immunofluorescence flow cytometry). The control cells for the first test proved this was not sensitive enough. Nevertheless, flow cytometry revealed 7% of C20 cells were infected at passage 156. This result was supported by the positive control since almost 80% of 18h p.i. Hi5 cells showed the presence of GP64. This supports the morphologic observations of C20 cells. Using different microscopy techniques it was observed that although there was clear signs of infection, some cells resembled Hi5 cells in size, shape and morphology. Even though this result seems very low, less than 10% of C20 cells are infected, this analysis was based on the late phase of infection, production of GP64. It is possible that the infections are not complete and not all the cells reach this phase although they are infected. Moreover, due to time limitations this test was only performed once. Although it included biological replicas, it would be important to repeat the experiment independently.

In the current chapter a relative quantification qPCR method for BV was developed to determine relative viral replication. The method fulfilled all the criteria to be able to be used as such although it was observed in the end of the optimization process that the reaction for *actin* has a higher efficiency when used in singleplex then when multiplexed with *gp64*. This probably means the *gp64* reaction is limiting *actin* reaction; so limiting the primers for *gp64* would slow down *gp64* reaction and enable the actin reaction to have an increased efficiency when multiplexed. Nevertheless, this method was tested and can be used in the future for similar tests.

Persistent infections are distinguished from latent infections (e.g. Herpes Simplex Virus-HSV) by the expression of viral genes and production of progeny virus even if at low levels. On the other hand latency is defined as a long-term infection where the viral genome is retained in the host but no expression occurs. Therefore no infectious viruses are produced. It is however capable of reactivation to repeat the infection cycle (Speck

and Ganem, 2010; Lieberman, 2016). Since RNA genomes would rapidly be degraded by the cell, retrovirus reverse transcribe its genome into DNA and integrate it into the host cell genome, staying dormant until reactivation. Previous studies mentioned the presence of a latent FHV infection in Hi5 cell stocks (Li *et al.*, 2007; Merten, 2007). However, nodaviruses are uncapable of integrating their genome into the host genome. The genome needs to be replicated by its own polymerase to be maintained. In Hi5 cells, no signs of infection are present but gene expression needs to occur to some extent for polymerase to become available. Therefore the FHV infection in Hi5 cells might be a persistent infection rather than latent since some viral expression occurs. However, some authors consider latent infection can include gene expression from a limited number of genes to maintain latency (Chao *et al.*, 1998; Kane and Golovkina, 2010).

Although important data was generated in this chapter, a slightly different design of the experiments could have answered more questions. The experimental analysis of BV production (Fig. 3.10A, B) measured cumulative titre. Measuring the BV production every day would have provided another perspective on the AcC20 replication in the cells. In order to do that, an exhaustive wash of the cells and media replacement would have to be done every day. The developed qPCR method analysed *gp64* relative quantity, reflecting the relative number of AcC20 copies to the endogenous cell control and standardized to a calibrator sample. However, to better understand the mechanism occurring in a persistent infection, especially in a long term culture it would have been interest to optimize the experiment from the gene expression point, using cDNA as the starting sample.

The variability in C20 cultures presented in this chapter is going to be a factor throughout this thesis and it is important to keep in mind when analysing the results presented. Although the experiments were designed to enable as much reproducibility as possible to enhance comparison, and this was ensured in all the controls, in many of the experiments, the inherent variability of the C20 cells is reflected and cannot be interpreted as a fault of the technique or experiment but rather an important feature of the C20 cells.

C20 cells were generated from a single cell, however, this thesis focuses on passages 50 to 160 of the cell line. Every cell line suffers slight changes once every few passages. This

is probably exacerbated in C20 cells by the presence of AcC20, causing cells and virus to evolve together in a balance for continuous survival. In the present chapter, concerns that we might be facing a mixed population arose. This raises important questions that need to be acknowledged in this thesis: are all C20 cells infected at every time point? It might be that besides the persistently infected cells, some cells actually die of the infection while some others are free of virus and can be infected by the newly generated virus at any point. One of the very early persistent infections described, mentioned that in a later phase not all the cells were infected in any culture passage and less than one percent of the cells were infected. However we need to consider the screening tests available at the time were more limited than those of the present day so this percentage might have been higher (Crawford and Sheehan, 1983).

The fact that only 7% of C20 cells may be producing infectious virus demands a future experiment to re-clone the cells by limit dilution. If the frequency of virus-infected cells is really so low, it should be a comparatively simple matter to isolate those free of infection. In the initial cloning process that generated C20, all other clones were found to produce virus. Over 160 passages later, it might not be surprising if some C20 cells have either lost the virus completely, or harbour only part of the original virus genome. Such a phenomenon was observed in *S. exigua* cells harbouring a persistent virus genome (Weng *et al.*, 2009).

3.4 Conclusion

This characterisation of the C20 cell line has showed that the cells are in general larger than the parental cell line although more diverse. They have a slower growth pattern and are more resilient then the parental Hi5 cell line as they can survive for a long time without being passaged and can still recover full viability if split when the viability is as low as 10%. Even under proper culturing conditions, viability of each culture is usually lower then Hi5 cells, a first prove of the balance occurring with the persistent infectious AcC20 virus.

Besides the persistent infection, C20 cells, as well as its parental cell line, present a latent infection by FHV from the nodavirus family. This is an important factor to be

considered in future data interpretation. Sf9 and TN368 proved to be free from this infection.

AcC20 virus was present in every tested passage of the C20 cells and its production into BV varies with time in an independent culture. This balance might be related with the growth pattern presented by C20 cells. Viral protein production also fluctuates with time although GP64 was always present.

Together the data from this chapter shows C20 cells are persistently infected. Although AcUW1.*lacZ* was the initial virus used to generate this cell line, eight years of continuous adaptation of cells and virus have passed, so the virus currently infecting C20 cells will be referred as AcC20.

Chapter 4

C20 cells are partially resistant to superinfection

4.1 Introduction

The data from Chapter 3 confirmed C20 cells harbour a persistent infection. Although a low level of BV was shown to occur at all times, a lytic infection was not triggered. Why the viruses produced do not infect neighbouring cells and generate an overt infection is unknown. It was hypothesised that the cells might already be infected even though not all cells show signs of infection. Nevertheless, analysis of the presence of GP64 in the cells suggested that not all the cells may be infected at all times or, at least, they do not reach the late phase of infection. Another possible explanation could be that some or all C20 cells are resistant to infection by the endogenous virus. In that case, it would be interesting to clarify if this resistance is only for AcC20 virus or if the C20 cells are resistant to infection by a different AcMNPV. This phenomenon is called superinfection exclusion and it is described as the capacity of an established virus infection to interfere or prevent the infection of a second virus (Tscherne *et al.*, 2007). In the examples where the cells are resistance to superinfection, infection of the same cell by more than one virus is not possible. Several studies have suggested this phenomenon happens frequently with baculoviruses, as infected insect cells appear to have the ability to interfere with subsequent infection by a closely related virus, becoming resistant to superinfection (Weng *et al.*, 2009; Beperet *et al.*, 2014). In overt AcMNPV infections, exclusion seems to correlate with the release of the first BV in the late phase of infection. However, it is not clear how or why superinfection exclusion happens. Similarly, although very little is known about this process in persistently infected cells, studies have suggested they are refractory to secondary infection by both homologous and heterologous baculoviruses (McIntosh and Ignoffo, 1981; Lee *et al.*, 1998) .

The current chapter examined the superinfection phenomenon in an AcMNPV persistent infected cell line aiming to clarify if their resistance to superinfection is the key for their long time survival in culture. It was intended to expand the knowledge of the mechanisms underpinning superinfection exclusion. To follow homologous superinfection in C20 cells, a virus with EGFP fused to AcMNPV major capsid protein VP39 (AcEGFP-VP39), was used in confocal microscopy studies (previously AcEGFP-VP39NatP, (Danquah *et al.*, 2012)).

In an unsuccessful virus superinfection it is not known at which stage the process is halted. Although there are unlimited hypotheses, several important cell regulated steps will be studied. After recognition of the GP64 receptor, baculovirus BV enter the cell through endocytosis by fusion of the virus envelope with the endosome membrane (Volkman and Goldsmith, 1985; Blissard and Wenz, 1992; Long *et al.*, 2006; Dong *et al.*, 2010). Thus endocytosis might be a barrier for superinfection. Once the viral DNA is uncoated in the nucleus, virus gene expression starts in a tightly regulated process through immediate early, early, late and very late phases of infection (Rohrmann, 2013). To determine if the blockage to superinfection occurred in one of these phases, representative gene promoters active in each of them (linked to a reporter gene) were used to monitor gene expression by the superinfecting virus. These comprised: *ie1* as a representative of the immediate early genes responsible for transactivation of the early genes (Rohrmann, 2013); *39k* (or *pp31*) contains both early and late promoters and encodes a phosphorylated DNA binding protein that locates to the nucleus associating with the virogenic stroma (Yamagishi *et al.*, 2007) and *p.6.9,* responsible for condensation of viral DNA, are both examples of late genes (Wilson *et al.*, 1987; Thiem and Miller, 1989; Dong *et al.*, 2010; Rohrmann, 2013); finally, *p10* and *polh* are hyper-expressed in the very late phase of infection and are responsible for aiding nuclear lysis (Graves, 2016) and forming the polyhedra structural protection (Duncan *et al.*, 1983; Ji *et al.*, 2010), respectively.

In order to study the different phases of infection and their possible role in the superinfection resistance, several approaches were considered based on confocal and electron microscopy techniques, gene expression and traditional adsorption and uptake essays. This will help us clarify if the C20 cells are resistant to a secondary infection and if they are, in which phase the blockage occurs.

4.2 Results

4.2.1 Analysis of superinfection in C20 cells

To test if C20 cells could be superinfected, a recombinant virus (AcEGFP-VP39; (Danquah *et al.*, 2012) with EGFP fused to VP39, the major AcMNPV capsid protein, was used to infect C20 cells and Hi5 cells as a positive control. Tagging of VP39 with EGFP allows BV

to be tracked by confocal microscopy with their accumulation generating a strong EGFP signal, for example, in the nucleus where nucleocapsids are being produced (Danquah *et al.*, 2012).

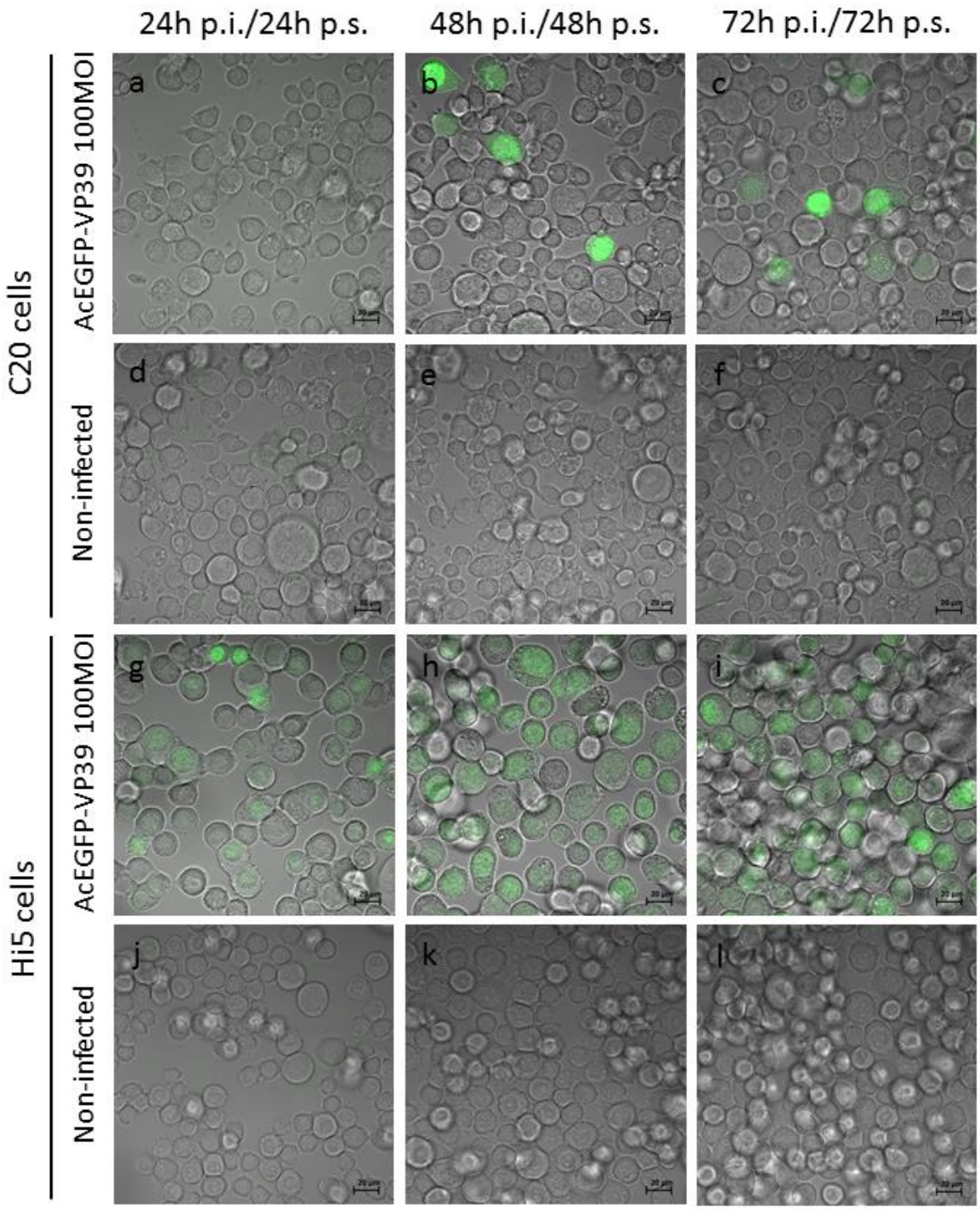


Figure 4.1: Live cell imaging for superinfected C20 and infected Hi5 cells.
C20 (a-c) and Hi5 (g-i) cells were infected with AcEGFP-VP39 at 100 MOI and observed on the confocal microscope at 24h p.i., 48h p.i. and 72h p.i. Negative controls were imaged at the same time points, C20 (d-f) and Hi5 (j-l). Scale bars 20μm. Representative images are showed.

Briefly, both cell lines were seeded in glass bottom dishes in triplicate. One dish from each cell line was mock-infected, while the other two were infected with AcEGFP-VP39 at 100 MOI. Cells were incubated at 28°C and imaged at 24, 48 and 72h p.i. (Fig. 4.1) using confocal microscopy combining transmitted light with an EGFP channel (See 2.7).

Mock-infected controls shown no fluorescence above background (Fig. 4.1d-f, j-l). In the Hi5 cells, a strong EGFP signal was observed in infected cells at 24h p.i. An increase in the proportion of infected cells was observed from 24h p.i. to 48h p.i. and then to 72h p.i., although at this late time point, the accumulation of floating cells made it harder to understand the infection pattern (Fig. 4.1 g-i).

In C20 cells by comparison, at 24h p.i. there were no signs of infection (Fig. 4.1a). Nonetheless, at 48h p.i. and 72h p.i., a small proportion of cells showed a strong EGFP fluorescence in the nucleus, suggesting that a small number of cells had been superinfected (Fig. 4.1b, c). This was a marked contrast to the Hi5 cells, in which all, or almost all cells were infected (Fig. 4.1g-i)).

4.2.2 Gene expression in different phases of superinfection

4.2.2.1 Construction of recombinant viruses expressing *egfp* under different promoters

Four virus constructs were created with *egfp* driven by promoters from the different phases of infection: AcIE1-EGFP, Ac39K-EGFP, AcP6.9-EGFP and AcPolh-EGFP. In order to generate these, plasmid vectors pOET1, pOET3 and pOET7, with *polh*, *p6.9* and *ie1* promoters, respectively, before a multi-cloning site, were used to develop constructs with *egfp* under different promoters. To obtain a vector with the *39k* promoter, this was amplified from AcMNPV WT DNA by PCR with primers *Eco*RV-39K and *Bam*HI-39K (See 2.4.1; Table 2.4). The PCR product was inserted by restriction cloning into pOET1 to derive pOET1-39k (see 2.4.2 and 2.4.4). The integrity of the constructs was confirmed by PCR screening, enzyme digestion and Sanger sequencing of the successfully transformed colonies DNA (See 2.3.2, 2.3.3 and 2.4.5). The *egfp* was retrieved from pEGFP-N1 and inserted into the four vectors containing different promoters: pOET7, pOET1-39k, pOET3 and pOET1 (Table 2.3, Fig. 4.2). Following ligation and transformation the integrity of the generated constructs, pIE1-eGFP, p39K-eGFP, pP6.9-eGFP and pPOLH-eGFP was confirmed as described above.

Recombinant viruses were prepared by co-transfection of Sf21 cells with the relevant vector and BacPAK6 (see 2.2.3) to derive AcIE1-eGFP, Ac39K-eGFP, AcP6.9-eGFP and AcPolh-eGFP (Fig. 4.2). Viruses were plaque-purified and amplified to 25ml P1 stocks in Sf9 cells as described in section 2.2.1 and 2.2.4.

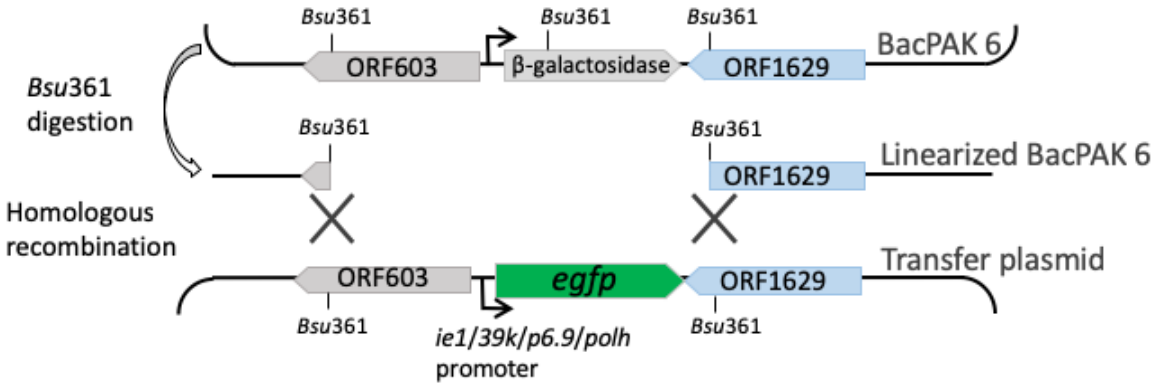


Figure 4.2: Schematic representation of *egfp* recombinant virus generation
Transfer plasmids containing *egfp* under *ie1*, *39k*, *p6.9* and *polh* promotors were used for co-transfection with BacPAK 6. Homologous recombination originated AcIE1-egfp, Ac39K-egfp, AcP6.9-egfp, AcPolh-egfp that were purified by plaque assays to obtain passage 0.

4.2.2.2 EGFP production in superinfected C20 cells

As a preliminary test for superinfection in C20 cells and to further check the recombinant viruses constructed in a lytic infection, protein extracts from infected cells in between 8 and 72h p.i. were analysed. A western blot for EGFP (predicted molecular weight: 26.9 kDa) and the control SDS-PAGE's to observe sample loading were performed (Fig. 4.3A). The Coomassie stained gels showed that in Hi5 cells the protein profile had changed from 36h p.i. onwards. In contrast, in C20 cells, although the general protein load increased with time, due to cell replication, the profile did not change. WB revealed the production of EGFP in Hi5 infected cells from 36h p.i. when under the *39k* or *p6.9* promoter, at 48h p.i. when under *ie1* and at 72h p.i. when under *polh* promoter. A non-specific band at 36kDa was observed with the anti-GFP antibody in all samples. Furthermore, the constructs were tested for the production of P10. *p10* is present in the four recombinant viruses under its own promoter so protein was expected to be present at similar quantities in between the four viruses and increase over time in a lytic infection. Non-superinfected C20 cell extracts are not expected to present P10 since the persistent infecting virus is *p10* $^{-}$. Figure 4.3B shows the WB and stained gels for P10 for Hi5 and C20 cells. In Hi5 cells, the infection developed as normal for the four viruses with P10 accumulating from 24h p.i. and increasing up to 72h p.i. However, sample

crossover in between wells was observed since there is signs of P10 on the non-infected control. In C20 cells, as predicted from the previous section results (see 4.2.1), low levels of P10 were produced. These experiments confirmed the recombinant virus are able to express EGFP and P10 production confirmed the viruses are infectious.

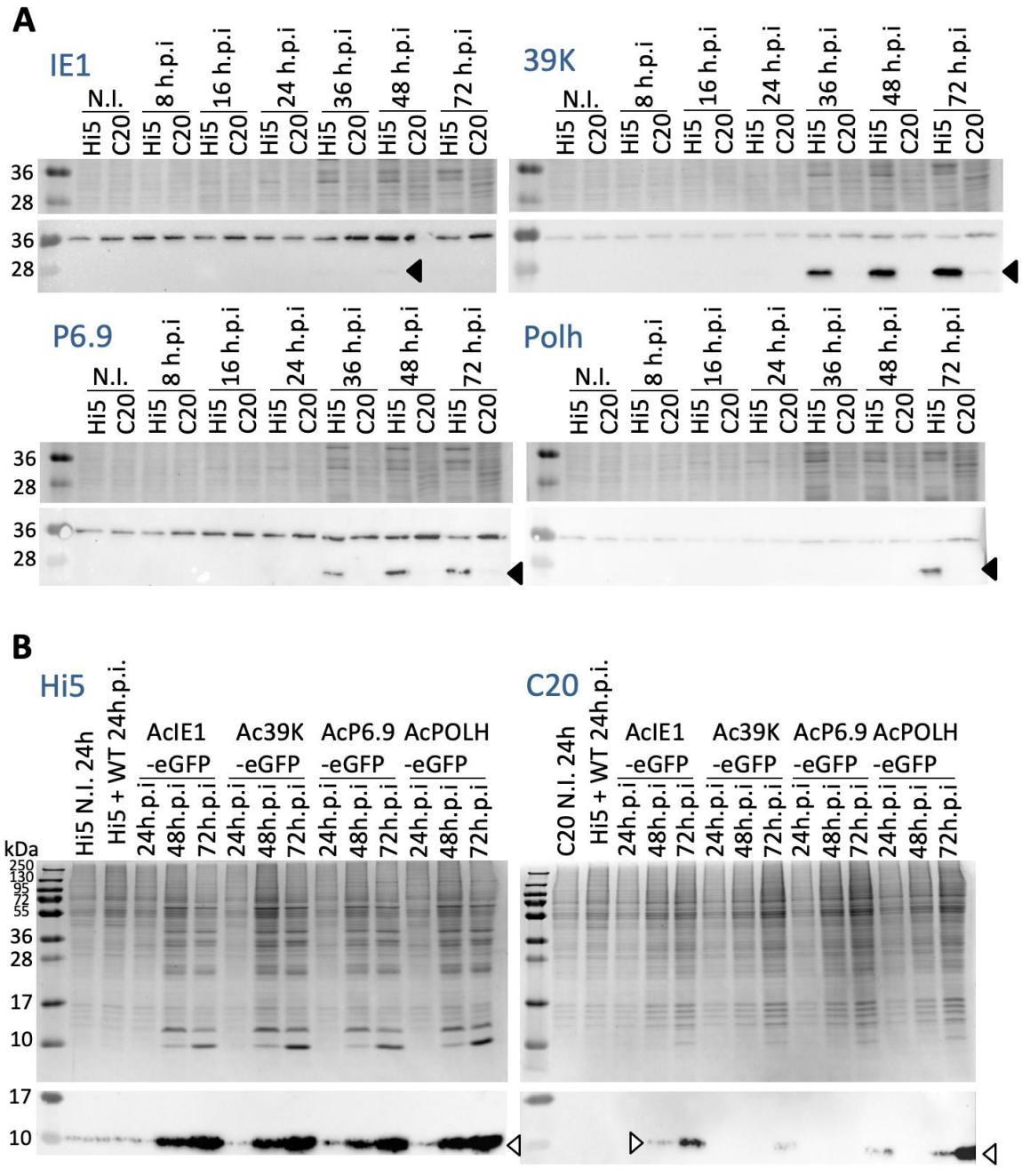


Figure 4.3: Protein analysis for recombinant EGFP viruses
Hi5 and C20 cells were infected with AcIE1-EGFP, Ac39K-EGFP, AcP6.9-EGFP and AcPolh-EGFP and infections harvested at 24, 48 and 72h p.i. Proteins were analysed by immunoblotting for A) eGFP and B) P10 and SDS-page revealed by Coomassie staining as a loading control. Page ruler marker was included in the first lane of each gel. All gels were composed of 12% polyacrylamide. 5µl and 10µl were loaded per sample, for Coomassie staining and P10 immunoblotting respectively. Exposure times for WB: A) 203.2s, 35.5s, 249.9s and 132.5s for IE1, 39K, P6.9 and POLH respectively; B) 183.6s, 218.5s for Hi5 and C20 respectively. Arrowheads highlight EGFP (black) or P10 (white).

The next test quantified EGFP production by viruses utilizing the different gene promoters in C20 superinfected cells and used lytic infections in Hi5 cells as a control. A single batch of C20 and Hi5 cells were used to seed all the multiwell plates, one plate per time points p.i. Staggered infections in both cell lines were created at 5 MOI in replicate for the four recombinant viruses, AcIE1-EGFP, Ac39K-EGFP, AcP6.9-EGFP and AcPolh-EGFP and a mock-infected control to be tested. At the desired time point, the plate was prepared for fluorescent analysis (see 2.8).

Figure 4.4A combines the results of the three independent experiments performed. As observed by WB, EGFP production from *p6.9* and *polh* promoters was not as high as expected from late promoters (Fig. 4.3). P6.9 is important for viral DNA packaging and nucleocapsid assembly while polyhedrin is the protein forming a matrix to occlude BV into ODV (Rohrmann, 2013). Therefore, during late and very late phase of infection respectively these need to be highly expressed. A high level of expression was expected from these genes in comparison to the previous two genes (*ie1* and *39k*). Expression of *ie1* was weak even in Hi5 cells due to the expected low expression from this promoter. For *39k*, the expression increased over time, as predicted, being significantly higher than in *ie1*. The next step consisted on testing expression in Sf9 cells in comparison with Hi5 cells to see if the problem with *p6.9* and *polh* expression could be linked with the cell line being used. Figure 4.4B shows the results when 3 independent experiments were performed for these two cell lines. Expression levels for *p6.9* and *polh* were still below what is expected although differences were seen in between cell lines (Fig. 4.4B).

At this stage, use of the *39k* promoter was chosen for further analysis to represent early and late gene expression since this gene contains both promoters. At all time points and for both cell lines, expression was always significantly higher than the background in PBS (for graphs in Figure 4.4, background signal was discounted from the fluorescence in each sample). This shows that even at low levels, C20 cells were able to express this early gene from a superinfecting virus. When comparing expression of *39k* in C20 cells with the parental Hi5 cell line, there was a significant difference in between them at each time point. These results suggest superinfection in C20 is possible although it occurs at a low level. Moreover, since this major difference for the parental cell line was observed in such an early phase, the blockage to superinfection is probably before the early phase of infection.

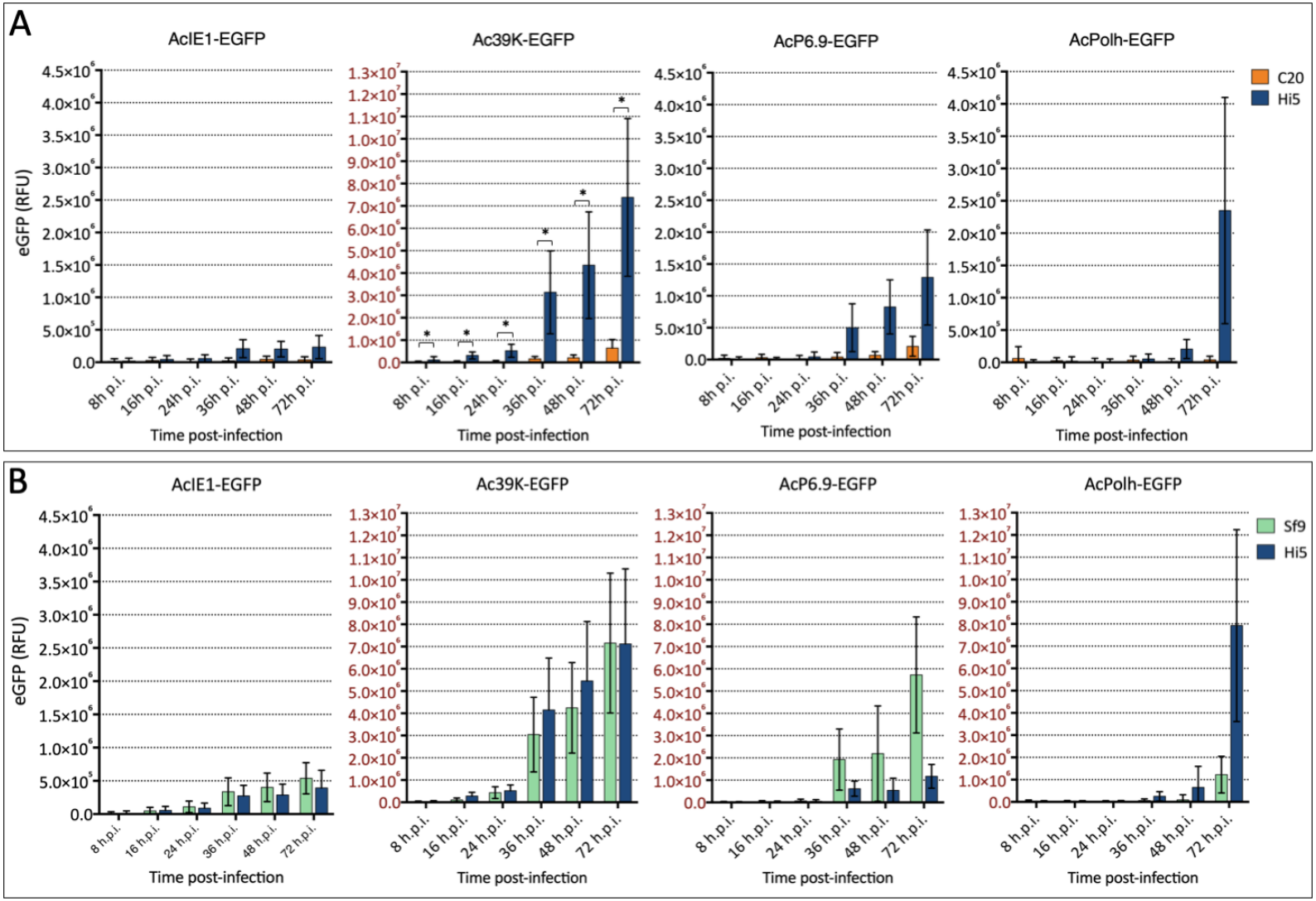


Figure 4.4: Expression of EGFP under different viral promoters
A) Expression of *egfp* in C20 cells superinfected and Hi5 cells infected with AcIE1-EGFP, Ac39K-EGFP, AcP6.9-EGFP and AcPolh-EGFP. B) Expression of *egfp* in Sf9 and Hi5 cells infected with the same recombinant viruses. Each panel incudes results from three independent experiments with four replicas each. Error bars represent SD. RFU, Relative fluorescence units.

From the four virus infections in Hi5 cells, the *39k* promoter was the one producing the strongest fluorescence from EGFP. Figure 4.5 gives a visual result of what was just described. C20 and Hi5 cells were infected with Ac39K-EGFP and pictures taken with bright field and fluorescent microscopy using a GFP filter. Expression on the positive control parental cells was strong from 24h p.i. and yet increases over time. In C20 cells superinfected with the same virus, fluorescence was not observed at 24h p.i. but the signs of infection were observed at 48h p.i. and 72h p.i. when the number or cells expressing EGFP under *39k* seems to have increased. This indicates that C20 cells are able to be superinfected but the infection seems to take longer to develop and only a small amount of cells are able to be superinfected.

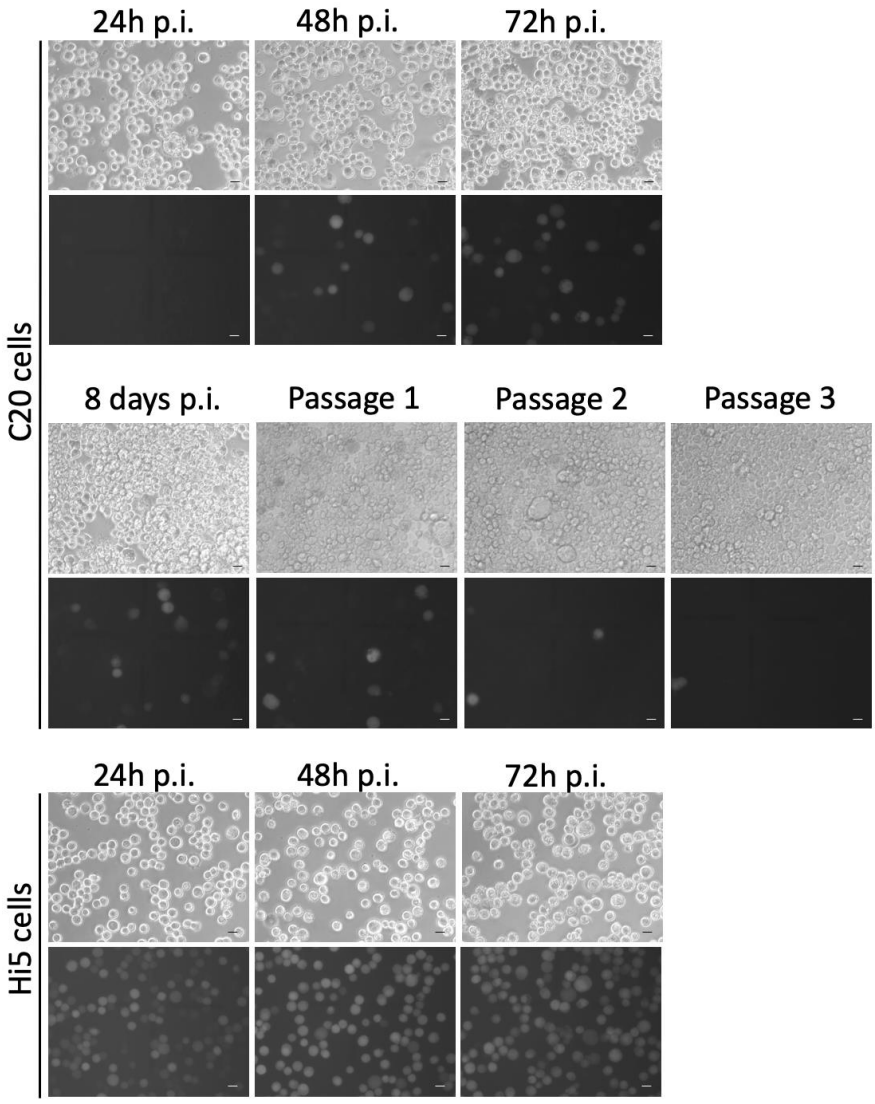


Figure 4.5: egfp expression under 39k AcMNPV promoter
C20 cells were superinfected with Ac39K-EGFP and imaged at 24h p.i., 48h p.i., 72h p.i., 8d p.i. and up to passage 3. Hi5 cells infected with Ac39K-EGFP and imaged at 24h p.i., 48h p.i. and 72h p.i. are shown as a positive control. Bright field and green fluorescent channel for each field are shown. Images are representative. Scale bars 20μm.

To ascertain if the established persistent infection could develop into a second persistent infection or lytic infection by provoking cell lysis, superinfected C20 cells were passaged until no more fluorescent cells were observed (passage four). From passage one to three it is clear that the quantity of superinfected cells decreases, probably a sign that this secondary infection is not able to spread from cell to cell or not at a high enough speed for the infection to spread (Fig. 4.5).

4.2.3 Early observation of superinfection in C20 cells using live cell imaging

Previous results in the current chapter suggested although superinfection in C20 cells is possible in a small proportion of the cells, there might be a blockage even before early gene expression. Therefore, it was decided to test by confocal microscopy earlier steps from the infection cycle, the virus binding and uptake.

To observe virus binding, $0.5x10^6$cells (C20 and Hi5) were seeded per glass bottom coverslip and incubated for 24h. To increase the changes of observing interaction in between virus and cell wall of C20 cells, AcEGFP-VP39 at 200 MOI was used and the virus added directly on the microscope stage, enabling observation of the very early phase of infection. The cells were examined at 2s p.i. (referred as 0min p.i. from now), 30, 60, 90 and 120min p.i. using time series mode to enable the visualization of any virus movement and adsorption to the cells. This approach enabled observation of the viruses moving around the cells and often adsorbing and being taken up by the Hi5 cells. Viruses were observed interacting with the membrane from 0m p.i.

It has been suggested that AcMNPV uses the endosomal system to enter the cell. Thus, cresyl violet was used to stain the lysosomes and enable co-localization with the EGFP marked viruses that are taken up by the cell. The BV were observed interacting with Hi5 cell membranes from time 0. Even though most viruses were observed outside of the membrane, from 30min p.i. virus particles were recognized inside Hi5 cells, and its number increased with the time (Fig. 4.6; Supplementary CD S4.2). In C20 cells, viruses were observed in the media and on rare occasions attached to the membrane. No viruses were observed inside the cells (Fig 4.6; Supplementary CD S4.1). Even though the lysosomal marker helped visualizing lysosomes, very little co-localization was detected.

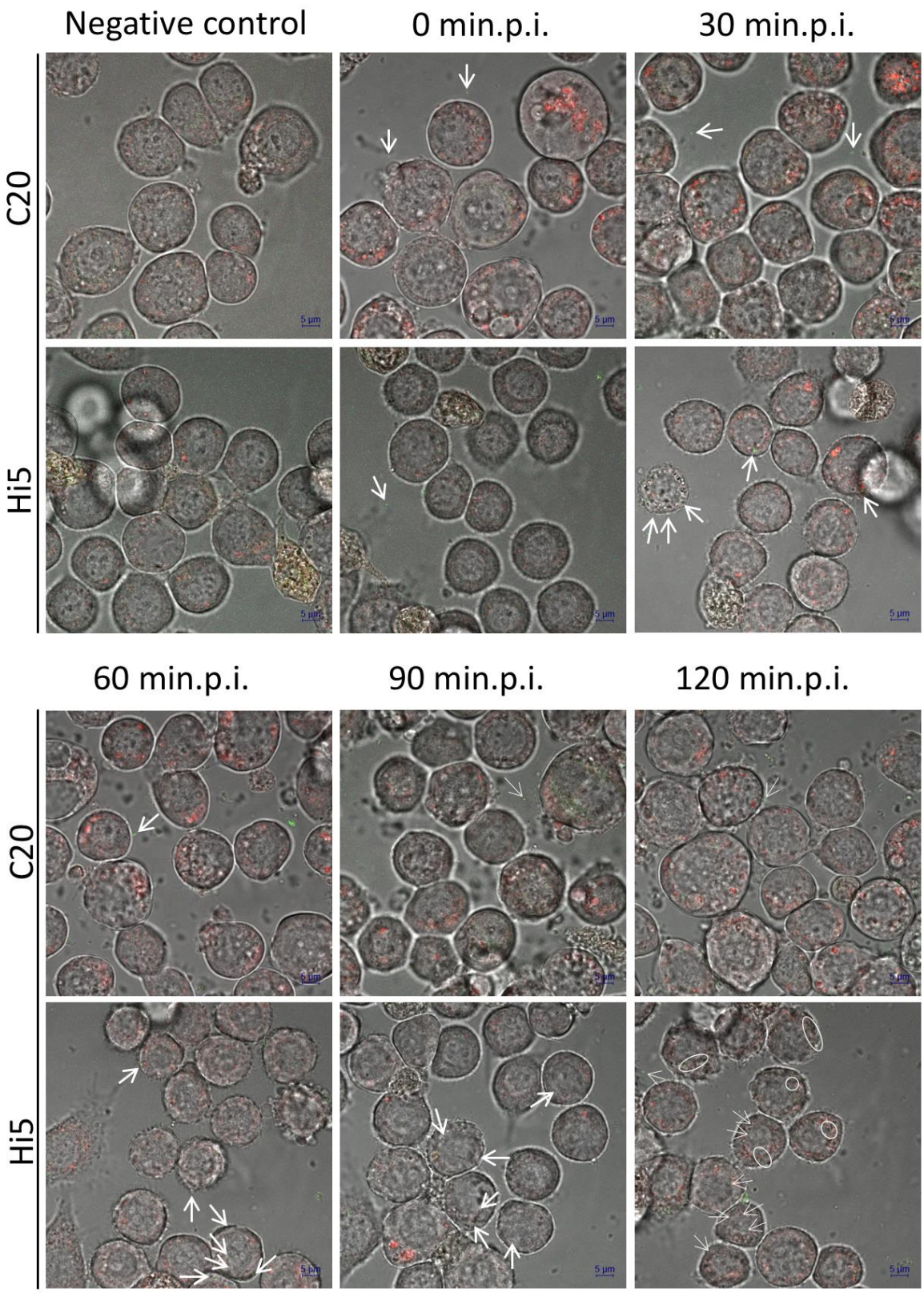


Figure 4.6: Live cell confocal microscopy images of uptake in superinfection and lytic infection.
C20 cells superinfected and Hi5 cells infected with AcEGFP-VP39 were observed at 0m.p.i., 30m.p.i, 60m.p.i., 90m.p.i., 120m.p.i. and incubated at 28°C in between time points (viruses marked in green). Lysosomes were marked with Cresyl violet (red). White arrows point to virus particles. Images are of representative cells. Videos from 60min p.i. are available for both cell lines in supplementary CD (Video S4.1 and S4.2).

Although this approach achieved good results, a method was needed that enabled direct quantification and statistical analysis. With live cell imaging the cells are in different optical planes and the viruses are constantly moving, which prevents accurate counting of the total number of viruses inside a cell or on its membrane. Furthermore, the confocal microscope being used had an Airyscan unit, which would enable improvement in the fluorescent imaging but was not fast enough to follow virus movement while keeping well resolved images necessary to observe particles as small as viruses. Hence, the next step was to optimize a protocol for fixing infected cells and visualize uptake.

4.2.4 Early observation of superinfection in C20 cells with fixed cells

The previous experiments suggested uptake was occurring within the first hour. Therefore, 20 and 40min p.i. were the time points chosen for optimization for fixed cell infections. The C20 and Hi5 cells were grown on coverslips inside multi-well plates and infected at 500 MOI with the AcEGFP-VP39 (to account for the viruses that could be washed off before the fixing process). The cells were then fixed with 4% formaldehyde (in PIPES buffer) at 20 and 40min and prepared with vectashield mounting containing DAPI, to stain cellular nuclei (which replaced the lysosomal staining). CLSM with Airyscan enabled the observation of these preparations with high resolution. This detector is composed of several detector elements to generate a final image with increased spatial resolution (Huff, 2015). Z-stacks were performed to analyse virus presence in the whole cell. In Hi5 cells, virus uptake was observed at 20min p.i. in some of the imaged cells (Fig. 4.7i,j). At 40min p.i., the number of cells with virus interacting with the membrane increased (Fig. 4.1k,l). At this time point some viruses were also observed inside the cell and crossing the nuclear membrane (Fig. 4.7l). In C20 cells a lower number of virus particles were observed attached to the membrane and only very rarely were they observed inside the cell (Fig. 4.7c-f).

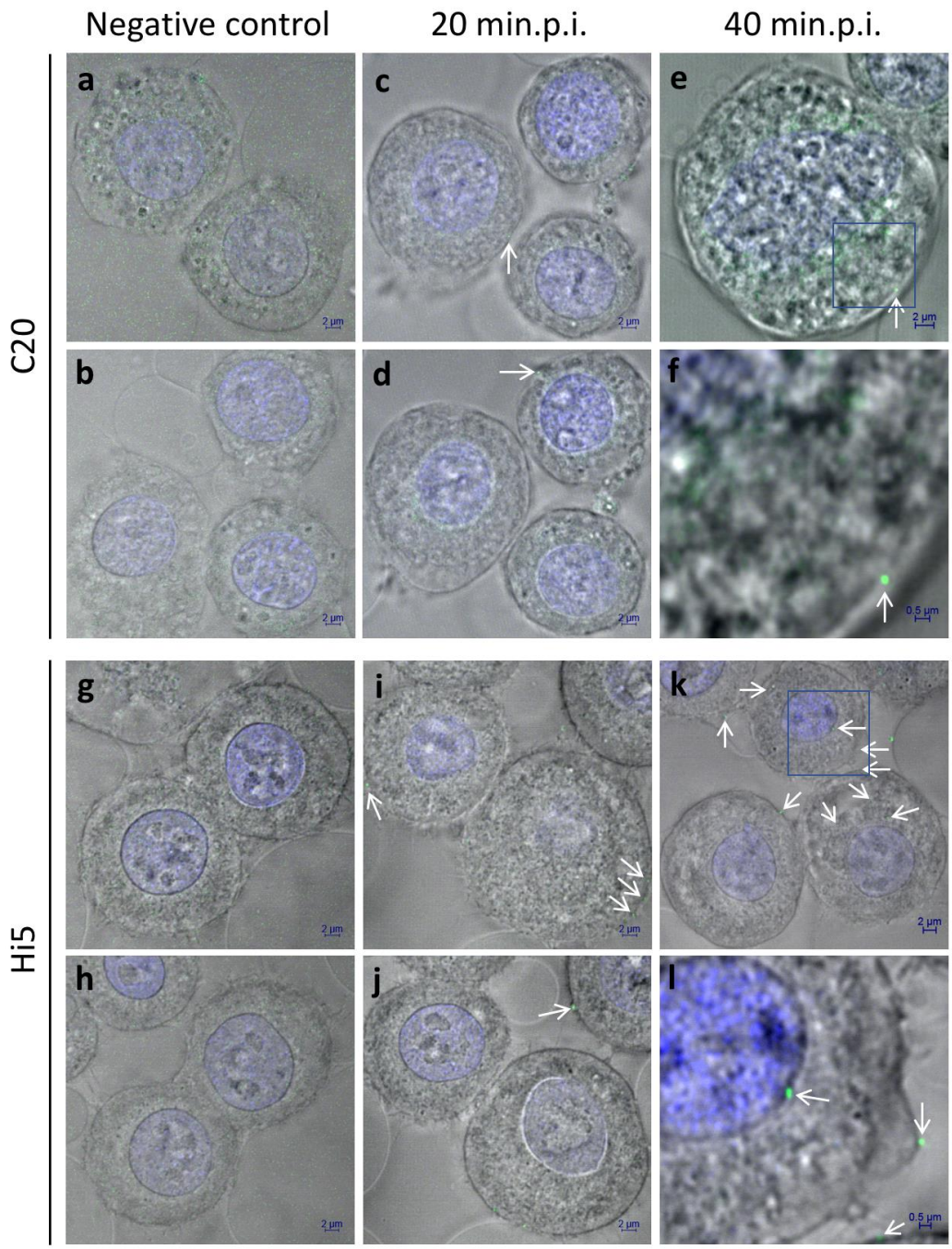


Figure 4.7: Confocal microscopy images of uptake in superinfection and lytic infection using fixed cells. C20 cells superinfected (c-f) and Hi5 cells infected (i-l) with AcEGFP-VP39 were fixed at 20m.p.i. and 40m.p.i. (viruses marked in green). C20 (a,b) and Hi5 (g,h) negative controls are shown. Preparations were mounted with Vectashield with DAPI to mark the DNA (blue). Panels c-d, and i-j represent different focal planes of the same cells. Panels f and l represent magnifications of the blue squares marked in images e and k respectively. White arrows point to observed viruses. Images are of representative cells.

A total of 129, 123, 107 and 108 cells for Hi5 fixed at 20min p.i., Hi5 fixed at 40min p.i., C20 fixed at 20min p.i. and C20 fixed at 40min p.i. respectively, were observed in this experiment. Each whole cell was examined and the number of viruses observed on the membrane and inside each cell counted. In both cell lines the number of virus particles

observed inside the cells was lower that observed on the membrane. Statistical analysis confirmed the early observations (Fig. 4.8A). There was a significant increase in the number of virus particles observed at the membrane or inside Hi5 cells from 20 to 40min p.i. ($p_{Mann\text{-}Whitney}<0.05$). In C20 cells, at both time points and virus location, the number of viruses was significantly lower when compared to Hi5 cells ($p_{Mann\text{-}Whitney}<0.05$). However, there was a significant increase in the number of viruses binding to the membrane of C20 with time, which was not reflected in a similar increase in number of viruses inside the cells. Figure 4.8B shows the histograms for virus particles at the cell membranes or inside the cells for both cell lines. A clear difference is observed in the distribution of the frequencies in between cell lines.

The frequencies of virus particles inside C20 cells compared to Hi5 cells suggests an important blockage to the superinfection occurs before or at the uptake step. When analysing the presence of viruses in the cytoplasm, only 2.8% (n=3, out of 107) and 5.6% (n=6, out of 108) of C20 cells at 20min p.i. and 40min p.i., respectively showed virus particles in the cytoplasm. In Hi5 cells, 45.7% and 89.4% of cells had virus particles present in the cytoplasm 20 and 40min p.i. respectively. In C20 cells, no more than two virus particles were observed in the cytoplasm of any cell while Hi5 cells were recorded with a maximum 6 and 30 virus per cell in the cytoplasm, respectively for 20 and 40min p.i. (Fig. 4.8B).

Examining the distribution of BV binding to the cell membrane, indicates the resistance to superinfection can occur even earlier by a blockage in the adsorption step. At 40min p.i., in C20 cells a maximum of 10 virus particles were observed adsorbing to the cell membrane, in the whole cell. In contrast, in Hi5 cells, up to 73 BV were observed binding to the same cell. At this time point, only 3.3% (n=4, out of 123) of Hi5 cells did not have BV adsorbing to its membrane while 47.2% (n=51, out of 108) of C20 cells were in this situation.

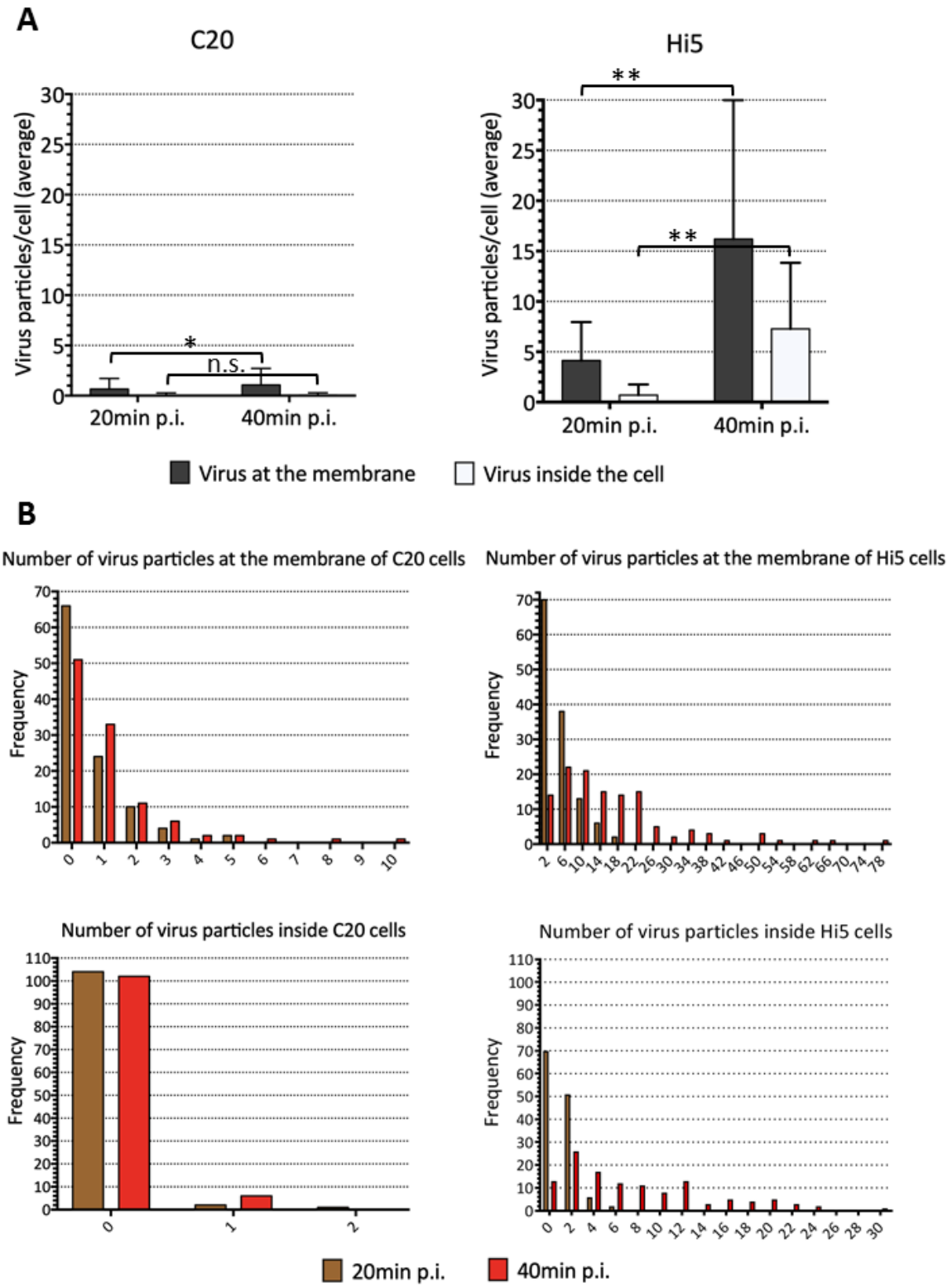


Figure 4.8: Quantitative results from confocal microscopy of uptake in fixed cells.
A) Viral particles at the membrane or inside C20-superinfected or Hi5 infected cells were counted in both cell lines and time points (cell numbers: 129 Hi5 at 20min p.i., 123 Hi5 at 40min p.i., 107 C20 at 20min p.i. and 108 C20 at 40min p.i.). (*) $p<0.05$, (**) $p<0.001$ and (n.s.) non-significant using Mann-Whitney test. Error bars represent SE. B) Histograms for frequency distribution of virus particles at the membrane or inside C20 and Hi5 cells For C20 cells absolute values are represented in the x-axis. For Hi5 cells the x-axis represents the centre of bins (bin width: at the membrane - 4; inside the cells - 2)

4.2.5 Observation of virus binding to cells using scanning electron microscopy

Scanning electron microscopy (SEM) was performed to further study the secondary infection in the early stages. Hi5 and C20 cells were infected with AcEGFP-VP39 at high MOI and fixed as described (see 2.7.3)

At 20min p.i, no virus particles were found on the C20 cell surface (Fig. 4.9). One virus particle was observed on the surface of the coverslip but not in contact with the cells. In contrast, several viruses were found on the surface of Hi5 cells, adsorbing to the cell, demonstrating that this result was not affected by the multiple washing steps involved in the protocol. However, with this technique we cannot ensure this BV is from the AcEGFP-VP39 recombinant virus and not from the AcC20 naturally maintained in the persistent infected C20 cell line instead.

Differences in morphology discussed in chapter 3 were also observed between C20 and the parental Hi5 cell line. Hi5 cells presented extended protuberances a consequence of the attachment created with the growth surface (Fig. 4.9v-viiii). Examples of cells going through division were also observed since some cells were re-gaining the round shape characteristic of this step (data not shown). In contrast, in C20 cells, although long membrane protuberances were observed they were never as flat and accentuated as in Hi5 cells suggesting a weaker attachment to the surface (Fig. 4.9Ai-iv).

The viruses present were 280-320nm long and 60-70nm wide. To confirm that these structures were actually virus, a preparation of BV alone was made. Approximately $5.45x10^6$pfu were set in a coverslip and fixed for SEM as described in section 2.7.3. The structures observed presented the same size and shape as those recorded previously (Fig. 4.9B).

Both CLSM and electron microscopy indicated there was a partial blockage to secondary infection occurring on the adsorption step, therefore this step was explored further.

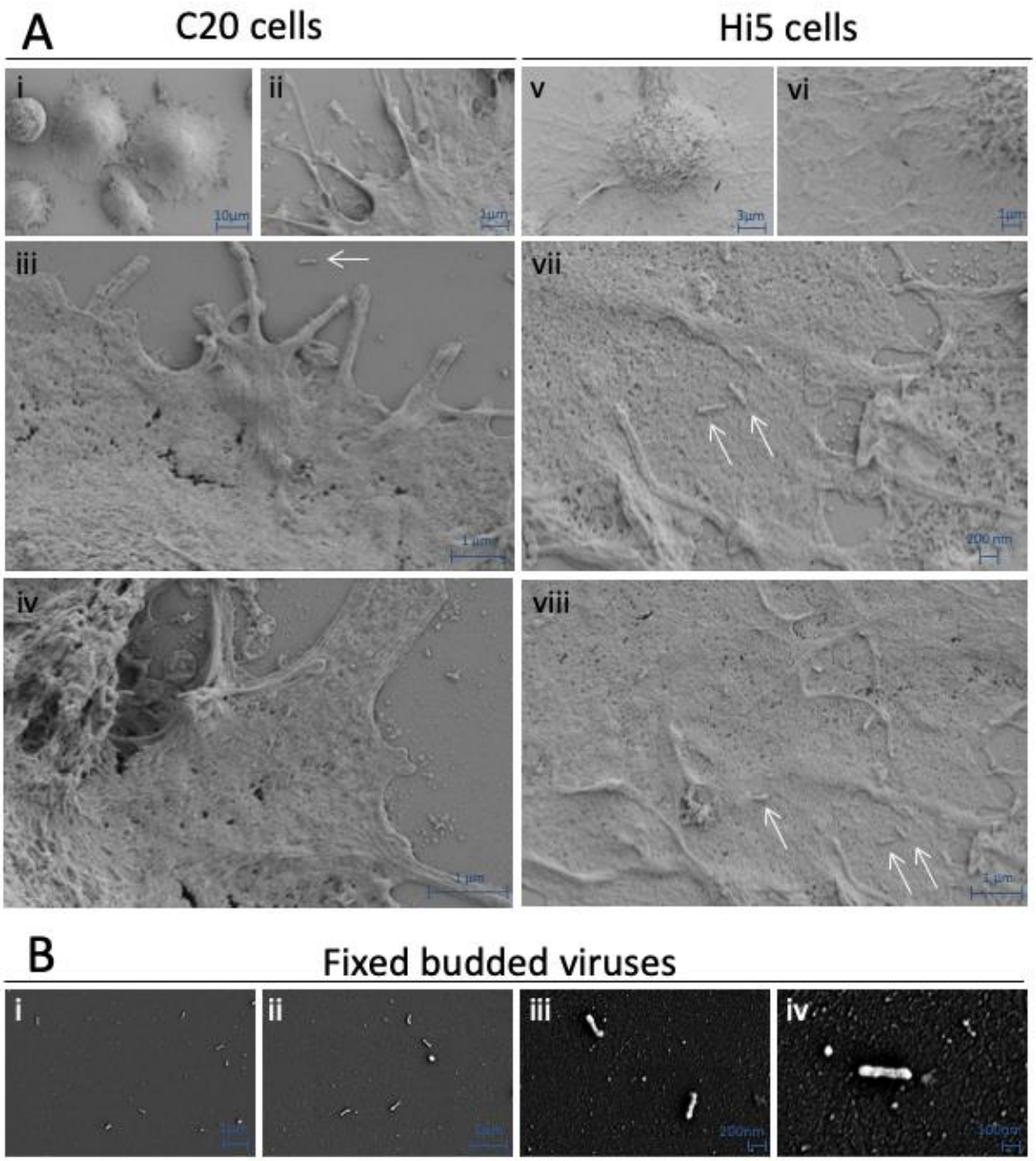


Figure 4.9: Imaging of C20 superinfected, Hi5 infected cells and BV using SEM.
A) C20 superinfected and Hi5 infected cells with AcEGFP-VP39 recombinant virus (500 MOI) and fixed 20min post-infection. B) Fixed budded viruses. After fixing, samples were subjected to critical point drying in CO_2 and gold coated with 15nm of gold particles. Blue arrows point to some of the structures identified as budded viruses.

4.2.6 Pull down assay to measure virus adsorption to cells

Minimal superinfection occurs in C20 cells and the experiments conducted so far suggest a blockage to secondary infection occurring at the binding step. Therefore, an experiment was designed to quantify the viruses adsorbing to the C20 membrane. Virus uptake is thought to occur in the first hour (Brown and Faulkner, 1978; King and Possee, 1992), therefore, cells were infected in suspension with virus and incubated for one hour

at 900rpm. It was considered that after this period virus adsorption would have occurred. The cells were then separated from the media by centrifugation and the media used in a plaque assay in Sf21 cells (see 2.2.2).

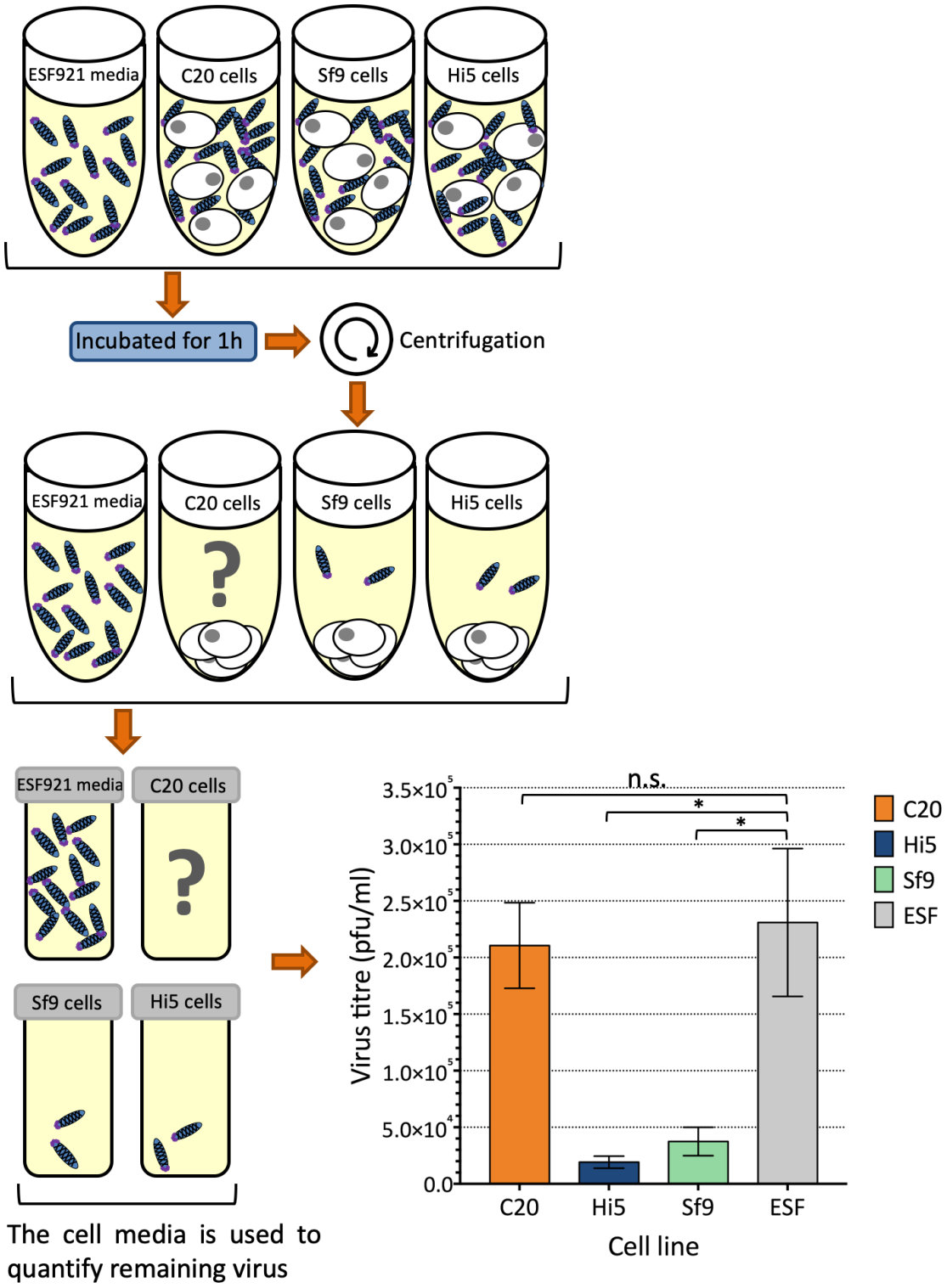


Figure 4.10: Pull down essay to evaluate adsorption and uptake
Schematic representation of the pull down assay. Briefly, C20, Sf9 and Hi5 cells were mixed with 0.1 MOI of AcUW1.*lacZ* as well as a control without cells (ESF media). After 1h incubation the cells were span down and the media used to perform a plaque assay. The graph shows the virus titres in the cell media. (*) $p<0.001$ and (n.s.) non-significant using Mann-Whitney test. Error bars represent SD.

A drop in the number of viruses in the media would be consistent with adsorption or uptake to the cell. Sf9 and Hi5 cell lines where used as positive controls and ESF921 insect medium alone as a negative. C20 cells and controls were seeded at $1x10^6$cells/ml in deep round bottom 12 well plates non-cell treated. Three replicas of Sf9, Hi5 and C20 cells and ESF921 in equal volumes were inoculated with 0.1 MOI of AcUW1.*lacZ*. At the same time three wells containing C20 cells were maintained as a control to assess how high the inherent virus titre was and make sure this factor was not interfering with the experimental results. The plates were incubated for 4 days or 5 days in the case of the C20 negative control, given the fact that C20 virus plaques always need a longer incubation time in order to be clearly counted after staining.

As shown in Figure 4.10, in both Sf9 and Hi5 cells there was a drop in the number of viruses present in the media compared to the negative control (ESF; $p_{Mann\text{-}Whitney}<0.001$ for both). The control for C20 cells revealed that the titre of the virus present ranged from $2.4x10^3$ to $4.7x10^4$ pfu/ml, values higher than expected. This value was subtracted from the C20 titre obtained on the AcUW1.*lacZ*-inoculated C20 cells. The results show there was no adsorption taking place on C20 cells since there was no significant difference to the ESF921 negative control.

4.2.7 Virus titre fluctuation in superinfected C20 cells

The previous sections showed C20 are susceptible to superinfection although: a) only a small proportion of the cells seems to be able to be secondarily infected; b) the infection seems to progress slower and c) the blockage to superinfection seems to occur as early as the adsorbtion step of infection. Also, the superinfection does not seem to be able to progress into a lytic infection or persistent infection since over passages the secondary virus was lost (section 4.2.2.2). However, nothing was known about the progression of the virus titre during a secondary infection event in C20 cells.

In order to evaluate the BV replication after super infection , replicas of C20 and Hi5 cells were seeded in 35mm dishes and infected in duplicate with AcRP23.*lacZ* or with insect cell media as a negative control. The C20 cells were washed 6 times with ESF921 before the experiment was started to remove endogeneous virus. AcRP23.*lacZ* has *lacZ* inserted under the *polh* promoter, which distinguishes it from residual AcC20. Cells were

incubated for 45min with the virus and treated with PBS at pH3 for 10min[1] to eliminate non-attached particles or those adhering to the cell surface that could interfere with the results of the early time points. Buffer at pH11 was then added and the cells washed with pH7 PBS and cell media before subsequent incubation. Samples of the cell media were collected at 1h p.i. and 3h p.i. to establish a baseline and then every 24h for 6 days. After removal of residual cells by centrifugation, the cell media was used to perform plaque assays. These were them stained with neutral red and X-gal to identify plaques generated by AcRP23.*lacZ*.

For C20, a superinfection was created with AcEGFP-VP39 to control the possible reactivation of the persistent virus AcUW1.*lacZ*. However, a previous test challenging C20 with X-gal has revealed blue color is not generated anymore. From this viral control no blue plaques were observed after the plaque assays were stained (data not shown).

Virus titres for the super infecting AcRP23.*lacZ* and for the persistent infecting AcC20 were calculated. Figure 4.11A combines the results of the three independent experiments with biological replicas performed and showed individually in Figure 4.11B. The baseline time points (1h p.i. and 3h p.i.) were averaged and marked at 2h p.i for more accuracy, since at this time points very low titres were expected. Although the cells were treated with buffer at pH3 and washed before the first samples were taken, only in C20 non-superinfected was the infectious titre below detection. In Hi5 and C20 infected with AcRP23.*lacZ*, a baseline titre in between 5 and $1.2x10^3$ pfu/ml was observed. In the control Hi5 cells, a normal lytic exponential infection was developed although low titres (~10^6) were produced. In C20 non-infected cells a normal titre for AcC20 virus was achieved (~10^3) after 72h p.i. showing the cells are still able to produce

[1] The GP64 triggers fusion of the virus envelope with the endosome membrane during BV entry into the cell by endocytosis. The protein fusogenic activity depends on the low pH that causes conformational changes and enables merging of the membranes (Blissard *et al.*, 1992). Therefore it was hypothesised that a low pH environment would trigger GP64 conformational changes, possibly causing BV inactivation or membrane disruption. To check for virus inactivation in reaction to a low pH, PBS based solutions at different pH's were tested for different times in the present study. Plaque assay experiments shown that after 10min of treatment at pH3, 99.9% of virus were inactivated. Equally, when treating Hi5 and C20 cells for the same time at this pH the cell viability was maintained. Also, mixing equal volumes of solutions at pH3 and pH11 has shown to restore a neutral pH to the solution even though the buffer proprieties were lost at these extreme pHs (data not shown). The pH adjustment with HCl and and NaOH caused an increase on Na^+ and Cl^- ions besides H^+ and OH^- that should be acknowledge when analysing the solution effect on the virus particles.

new virus after treatment with a low pH buffer. For C20 super infected with AcRP23.*lacZ* the virus titre fluctuated along time altough titres were generally low, as the maximum reached in one replica was $3.4x10^3$pfu/ml, at 120h p.i. In this superinfection, distinguish patterns from the lytic infection produced in Hi5 cells were observed. Figure 4.11B shows the progression of infection in the individual biological replicas. While the initial titre after the acid treatment is different between replicas the fluctuation pattern from this baseline seems to be similar in between them. Titres in between 10 and $3.4x10^3$pfu/ml were achieved in the superinfection in between 2 and 6d p.i. (Fig. 4.11B).

This experiment has shown once more that superinfection is possible and occurs at a low level but the secondary virus is not able to create a productive infection.

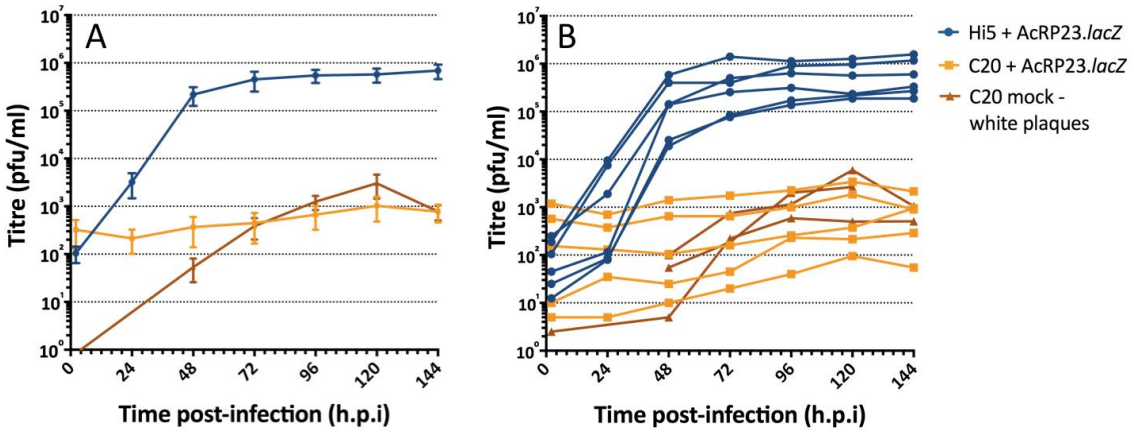


Figure 4.11: Virus titre fluctuation over time in superinfected cells.
C20 cells superinfected and Hi5 cells infected with AcRP23.*lacZ* and C20 non-infected titres are represented over a period of 6 days. Three independent experiments with two replicas each were performed. A) overall mean titre for all the experiments and replicas, showed separately in B. Errors bars represent SE.

4.3 Discussion

The fact that the C20 cells present a persistent infection for over seven years and although the produced virus is infectious, they do not succumb to a lethal infection presents an interesting finding. Therefore it was important to study if the cells were resistant to superinfection with an homologous virus that was not AcC20. Existing references that studied the superinfection exclusion in persistent infected cells affirmed the cells are resistant to both homologous and heterologous baculovirus secondary infection (McIntosh and Ignoffo, 1981; Lee *et al.*, 1998). However, a third study suggested cells were only partially resistant to superinfection, with both classes

(Crawford and Sheehan, 1983). The results shown AcMNPV-persistent infected cells were highly resistance, requiring 1000 times more virus paricles to initiate an infection while a SfMNPV-persistent cell clone was only partially resistant. Nonetheless these tests were carried out in clonal cell lines from the persistent infection, absent of signs of infection that were considered cured. The slow growth of this cured cells and resistance to superinfection suggests though (Crawford and Sheehan, 1983), in light with other works, that this cell lines were still persistently infected. Superinfection exclusion in a lytic infection was investigated in a variety of virus infecting humans and plants. Several mechanisms have been suggested to be related with the resistance to a secondary infection, involving changes in the plasma membrane proprieties or reduced acidity of early endosomes preventing virus fusion (Singh *et al.*, 1997), competing for host factors essential for viral replication (Adams and Brown, 1985; Nasar *et al.*, 2015) and protein interactions (Karpf *et al.*, 1997; Folimonova, 2012). Why the resistance happens in baculovirus was not explored in detail before. The process followed in this chapter was designed to clarify if C20 cells were resistant to superinfection by a homologous AcMNPV virus. As a partial blockage was suggested to occur, the research was focused on understanding the blockage to superinfection by focusing in the different steps of viral infection, from the physical barriers of the viral cell entrance to the gene expression.

The first sign that the blockage to superinfection exists but it might not be complete was drawn using confocal microscopy to observe superinfected cells with a virus containing EGFP-tagged VP39, over three days. From 24h p.i a strong fluorescent signal was observed in Hi5 cells nucleus, likely to be attributed to the many new BV being produced in the virogenic stroma in the late and very late phases of infection (Fig. 4.1). Moreover, superinfection seemed to be possible in C20 cells however the process is slower than in a first lytic infection in Hi5 cells. Only from 48h p.i. fluorescence was observed in C20 cell's nucleus. However, even after 96h p.i. the proportion of superinfected cells is very small compared to a lytic infection, showing either the existing blockage to superinfection is only partial or only some of the cells are able to be superinfected. In the case of the second hypothesis, this suggests we might be facing a mixed population, where not all the cells are persistently infected at the same time or not all the population is currently infected with the same virus. The cells and AcC20 virus in them

have been passaged over 100 times. Passaging a virus increases the chances of genomic changes to occur in a virus and it is why is usually not recommended to passage a virus for more than a few times in succession. An example is the few polyhedra phenotype (FP) that appears upon virus passaging (Potter *et al.*, 1976), discussed in detail on Chapter 5. Besides that, AcC20 has the additional factor of surviving in a persistent stage which might increase the selective pressure at all times. Therefore, it is valid to consider we might be facing a mixed virus population and therefore a persistently infected cell population that keeps diverging in every passage. This was suggested before in section 3.2.3 when a range of cell cytoskeleton arrangement was observed.

Following the first hypothesis, that the blockage might now be total, gene expression studies were designed to check where the superinfection was stopped. Unfortunately, only *39k* had robust enough results to draw conclusions. When C20 cells were superinfected with Ac39k-EGFP, the level of emitted fluorescence was significantly lower then when infecting Hi5 cells with the same virus. This happened at all time points tested, from 8h p.i. to 72h p.i. Using a fluorescent microscope to look at the same cells, it becomes clear that this drop of fluorescence is not caused by a lower emission level from all the C20 cells but rather from an emission that looks like the same level as in lytic infected cells, but only from a few cells.

Confocal microscopy experiments were designed to check if C20 cells were able to internalize a second virus. Using live cell imaging, it was possible to follow a virus binding, getting taken up by the cell and in some cases reaching the nucleus, but this process was not very practical for a quantification experiment since only one focal plane was observed at each time point (Supplementary CD videos S4.1 and S4.2). On the other hand, by fixing infected and superinfected cells at defined time points it was possible to observe full cells by using Z-stacks. This enabled a precise count of how many viruses are interacting or have entered the cell at 20 and 40m p.i. Results showed that although there was a significant increase in the number of viruses binding to the membrane of C20 with time, the same did not happen when accounting for the number of viruses inside the cells. At 40min p.i. only 5.6% of C20 cells showed up to two viruses in the cytoplasm comparing to 89.4% of Hi5 cells which could present up to 30 viruses in the same cell cytoplasm. The fact that Hi5 cells could present more than 10 viral particles inside the cytoplasm at 40min p.i is in agreement with a previous study that indicated

superinfection exclusion starts when the first BV's are released from the first infection, so the end of the early phase of infection around 12 to 15h p.i. (Beperet *et al.*, 2014). The confocal experiment here described suggests a secondary virus is able to bind to the C20 cell membrane, showed by the significant increase of virus adsorbing to the membrane in between the two tested time points. However, the amount of binding virus is significantly lower than in a lytic infection. At 40min p.i., almost 50% of C20 cells did not present any virus binding to its membrane, suggesting a possible difference in the membrane composition might reduce the ability of BV to bind as suggested for Semliki Forest Virus (Singh *et al.*, 1997). Over the same time frame, the percentage of viruses that were taken up by the cell did not increase. This indicates that even if the superinfection blockage is only partial, it might occur very early in infection, possible involving binding step. Nonetheless, only a few of the BV that are resistant to this blockage and bind to the cell, are internalized and able to cause a secondary infection. Nevertheless, taking into account the confocal observations that showed that when the superinfection occurs, the process is slower than in a lytic infection (section 4.2.1), later time points should be considered to enable drawing further conclusions considering the uptake step. Following these results, further assays were prepared to check the binding and adsorption step in infection.

C20 cells superinfected and infected Hi5 cells were observed with electron microscopy to observe the adsorption and binding of the viruses to the cells. In the C20 cells no viruses were found bound to the membrane at 20min p.i. However, a careful interpretation of this result is necessary since the technique is working at a greater magnification and resolution. This enables highly detailed imaging but the visualization of a smaller number of cells. CLSM retrieved an average 0.65 virus per C20 cell membrane across more than 100 cells. Given the amplification used on SEM to enable virus visualization, the probability of observing a virus by SEM was very low.

An assay was prepared to check these early and important steps of infection that enabled to consider a bigger quantity of cells (see 4.2.6). With a pull down essay, the virus titre left on the media after 1 hour of incubation with C20 was analysed and compared to what happens in a lytic infection. This showed the number of virus left in the media of C20 cells was no different from our negative control where no cells were added. These results together with the microscopy observations might indicate the main

blockage to superinfection occurs very early in infection involving the adsorption step. Nonetheless, superinfection in C20 cells seems to have a longer viral cycle then lytic infection has it was shown for other phases of infection (section 4.2.1), so the uptake could also be slower than in a lytic infection. Also, previous techniques have suggested only a small proportion of cells can be superinfected, so not all techniques might have enough sensitivity to detect such subtle differences.

Since previous techniques have suggested superinfection in C20 cells is possible, an experiment was performed to evaluate how the titre would fluctuate in C20 superinfected cells, and if a pattern was followed. It was shown that superinfection can be maintained for 6 days and the titre is always considerably lower than in a lytic infection, although different fluctuation patterns are followed by different replicas of the same experiment. This secondary infection was not transformed in a lytic infection over the course of the experiment, the same as observed in section 4.2.2.2 when the superinfected cells were observed up to 8 days. Nevertheless, it would be interesting to extend the experiment to clarify for how long a superinfection can be maintained on the same samples of C20 cells and corroborate the results from fluorescent gene expression experiment where it was shown that the superinfecting viral infection disappears with time and passaging (section 4.2.2.2).

4.4 Conclusion

Throughout the chapter, different techniques have shown that superinfection is possible in C20 cells and occurs at a slower rate than lytic infection. The main barrier to superinfection seems to be the adsorption step although the uptake is also compromised. This hypothesis needs further testing to ensure the sensitivity of the techniques used is not affecting the results. On the other hand the main conclusion drawn from this chapter is that only a small subpopulation of the cells can be superinfected. This was shown by using a virus with a fluorescent protein fused to the capsid protein to follow uptake and observe infection up to 96h p.i. with confocal microscopy and studies on expression of *39k*. This means we are most certainly facing a mixed population of cells when referring to the C20 cell line.

Chapter 5

Virus genome sequencing and transcriptome analysis of C20 cells

AcC20 DNA whole genome sequencing was outsourced to Source BioScience. Bioinformatics analysis was performed by D. Leite (Oxford Brookes University). I carried out the subsequent analyses and necessary confirmation of each particular gene.

For the transcriptome analysis, I performed the infection studies and RNA extractions. The RNA-seq libraries were generated by C. Bannach (Oxford Brookes University) in G. W. Blissard´s lab at BTI in the USA under a visiting scholarship program. Bioinformatics analyses were carried out by K. Bao (BTI). Downstream analyses were initially performed by Carina Bannach and I re-analysed and interpreted the data in light of the biological characterisation of C20 cells and AcC20 virus (Chapters 3 and 4).

5.1 Introduction

In previous chapters a detailed characterization of C20 cells and AcC20 was performed to understand the mechanisms behind the persistent infection. The C20 cells seemed to be more resilient than parental Hi5 cells and the virus replicated at a lower rate although it was still able to establish a lytic infection when isolated and amplified in Sf cells. This suggested the virus infection capability had not been compromised. The percentage of cells producing infectious virus was lower than 10%, and the cell line contained a subpopulation of cells permissive to a secondary infection.

From an evolutionary point of view, baculovirus survival depends on its adaptation to the host. This host, however, is constantly evolving and naturally selecting more resistance mechanisms to counter virus virulence. Virus dependence on the host machinery for survival and replication, enhances the need for it to evolve new strategies to overcome cell defences and adapt to new conditions. This process provides strong selective pressure for virus adaptation. This need for adaptation is exacerbated in the case of persistent infections as they need to co-exist.

The AcMNPV genome was first sequenced in 1994. Its 133.9 Kbp encodes at least 156 densely packed open reading frames (ORFs) (Ayres *et al.*, 1994). This initial achievement supported further genomic characterizations that enabled attribution of function to individual *orf*'s. A recent study has detected a conserved complex of genes particularly prone to evolution. This included both essential and non-essential genes responsible for late expression like helicase and *lef* genes; very late phase *polh*, *ec27*, vlf-1 and *chitinase* (Hill and Unckless, 2017).

The C20 cell line was first established eight years ago (in 2011). Over 160 passages enabled both virus and cell adaptation and genomic evolution. To study the AcC20 genome, whole genome sequencing (WGS) was carried out at passage 40 of the cells and virus. After alignment with the original virus (AcUW1.*lacZ*) used to establish the persistent infection, *orf*s with major revealed deletions were analysed in the early passages of the virus to understand their evolution. Moreover a *de novo* assembly of the AcC20 genome was generated.

Baculovirus expression is temporally regulated in four phases necessary for the propagation of the infection. Based on *orf* promoter motifs it was estimated that the AcMNPV genome is comprised of 34.6% early, 51.9% late and 13.5% early and late genes (Chen *et al.*, 2013). The majority of viral mRNA results from expression of late (e.g. *p6.9*) and very late (e.g. *polh*, *p10*) genes, which are expressed in different temporal profiles (Fig.1.2). The shutdown of host cell transcription is highlighted by the fact that late in infection 80% of the total cellular mRNA population is comprised of viral transcripts (Chen *et al.*, 2013).

Baculovirus infections are followed by morphological changes in host cells. These includes cytoskeleton rearrangement (as observed in Fig 3.3), modifications at cellular signalling levels, the cell cycle arrest (Braunagel *et al.*, 1998), increase of oxidative stress (Wang *et al.*, 2004; Rohrmann, 2013), activation of a stress response (Lyupina *et al.*, 2010) and other antiviral defence mechanisms together with the shutdown of host transcription (Braunagel *et al.*, 1998; Nobiron *et al.*, 2003). Nevertheless, the virus replication cycle depends on the host cell machinery (Blissard and Rohrmann, 1990; Possee *et al.*, 2010; Rohrmann, 2013). Superinfection-exclusion seemed to involve the uptake and binding steps (Chapter 4) suggesting related mechanisms might be differential regulated in C20 cell cells. To explore this and other cellular pathways in persistent infection and to better understand how viral and host gene expression differs in persistent infection and superinfection from a lytic infection, strand-specific RNA-sequencing was performed on RNA extracts from C20 and Hi5 cells, both non-challenged and challenged with AcMNPV WT.

5.2 Results

5.2.1 Whole genome analysis of AcC20

At passage 40 of the C20 cells, the culture volume was enlarged to 1L. From this culture, BV was purified and subsequently DNA was extracted (R. D. Possee, personal communication). This DNA represents a snapshot of the virus produced by C20 cells at this early passage, and was subjected to WGS. Since the C20 cell line originated from an infection of Hi5 cells with AcUW1.*lacZ*, this virus genome was used as a scaffold for paired-end mapping (see 2.4.7).

5.2.1.1 AcC20 major deletions and insertions

Three regions with major deletions in the AcC20 genome were predicted when compared to the parental AcUW1.*lacZ* (Fig. 5.1). These comprised: 469bp of the 3' end of egt, an intergenic region and 120bp of the 5' end of the adjacent orf16; an internal deletion of 378bp of *p95*; 4756bp spanning the 3' end of *lacZ* (2742bp), the remaining nucleotides at the 3´end of *p10* (128bp), the 3' of *p74* (1476bp) and 410bp of intergenic regions. This analysis also resulted in a peak in the read depth at the *hr5* (Fig. 5.1). This is thought to be an artefact caused by the similarity between the six *hr* sequences present in the AcMNPV genome (Pearson *et al.*, 1992). These regions comprise large AT-rich regions that have a high degree of similarity. The genome reads were 50bp long, which might have been insufficient to align *hrs*-specific sequences at their correct regions. Finally, visual analysis of the aligned genome against AcUW1.*lacZ* revealed an insertion of 287bp in *p94*. Primers were designed to read across each of the identified regions by Sanger sequencing amplifying DNA from passage 40 (p40; used for WGS) and a much later passage (p135).

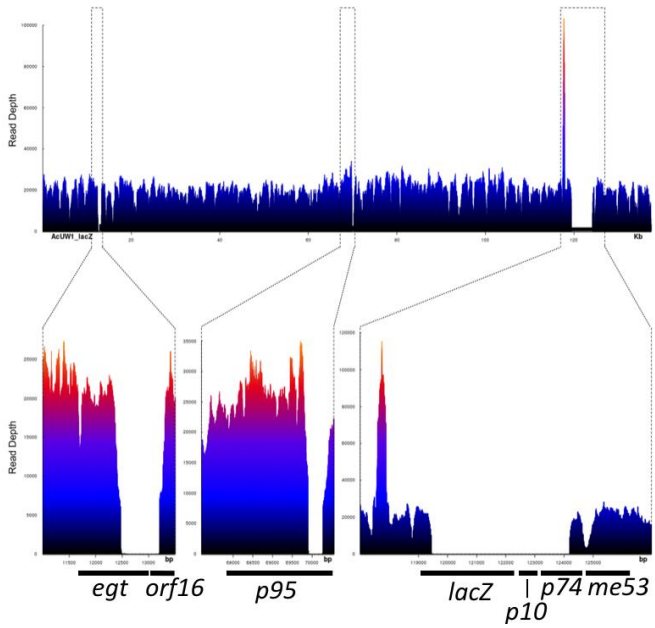


Figure 5.1: Magnification of deleted regions in AcC20 genome
Magnification of deleted regions and surrounding read depth. The three regions cover partially or totally: *egt* and *orf16*; *p95*; *lacZ*, *p10* and *p74*. The spike before lacZ represents *hr5*.

a) The *egt*/*orf16* region

The *egt* and *orf16* region was analysed using PCR with primers egt-ac16 Fw2 and egt-ac16 Rv2, which revealed an increase in size of the target rather than the deletion

identified in the whole virus genome analysis described above (data not shown). Therefore, primers egt-ac16 Fw5 and egt-ac16 Rv5 were designed that were closer to the deletion. This primer pair was predicted to generate fragments of 1001bp from the WT AcMNPV DNA and 267bp from AcC20 DNA. The predicted amplicon was confirmed for AcMNPV (Fig. 5.2A, band 1). However, AcC20 DNA from p40 and p135 both produced fragments of ~1700bp and ~370bp (Fig. 5.2A, bands 2 and 3). The apparent insertion in the virus genome that resulted in the band 2 product may have been partially lost in subsequent passages of the C20 cells after 40 as band 3 was produced to a higher concentration using AcC20 p135 DNA as a template in the PCR.

The DNA from bands 1-3 indicated in Figure 5.2A was isolated from the agarose gel. Sequencing of band 1 from AcMNPV *egt*/*orf16* matched that previously described for this region (Ayres *et al.*, 1994; Fig. 5.3A).

The results for band 2 confirmed an insertion but also identified a deletion in this region. The insertion was 1350bp long and occurred after position 12476bp relative to the AcMNPV genome (Fig. 5.3B). The AT-rich nature of the insert (67%) and the presence of long repeats of A and T including a poly A repeat of 12nt, required the use of several primers to obtain accurate sequence. The insertion just after a TTAA motif generated a duplication of the tetranucleotide after the insert. According to the nucleotide online database, both ends of the insert seem to relate to transposable elements from the *T. ni* host genome (NCBI, 1982; Altschul *et al.*, 1990). The 305bp 5' flanking region corresponds to the *T. ni* transposable element-derived protein 4-like while the last 164bp of the insertion are similar to *Trichoplusia ni* peptidyl-prolyl cis-trans isomerase G-like. Immediately after this insertion a 96bp fragment of *egt* remained before a deletion of 635bp encompassing 373bp of the 3' end of *egt*, 145bp of the intergenic region and 117bp of the 5' end of orf16. Around the deleted region further mutations included an extra T between the now fused egt and orf16 and an SNP from a cytosine (C) to an adenine (A) at the start of the remaining orf16 (equivalent to position 13211bp in AcUW1.*lacZ*).

Sequence analysis of band 3 confirmed that although lacking the insertion in band 2 describe above, it had an equivalent deletion of DNA and other nucleotide polymorphisms.

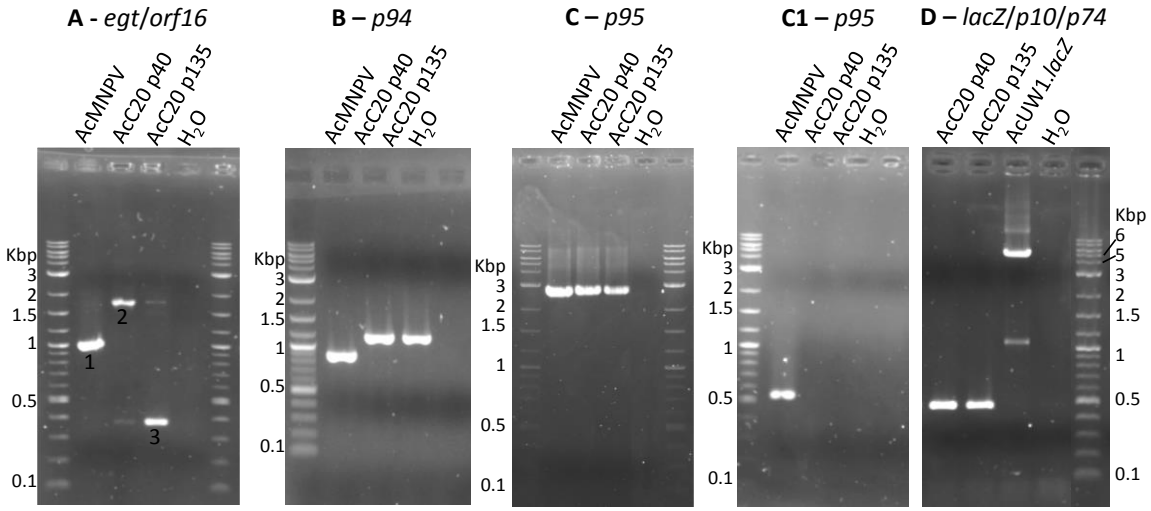


Figure 5.2: Analysis of AcC20 genome regions with major deletions and insertions.
AcC20 DNA from cell passages 40 and 135 were analysed using PCR: A) *egt*/*orf16* region, with primers egt-ac16 Fw5 and egt-ac16 Rv5; B) *p94*, with primers p94 Fw and p94 rv1; C) *p95*, with primers p95 Fw3, p95 Rv BamHI 2; C1) *p95* inner primers, p95 fw4 and p95 rv4; D) *lacZ*/*p10*/*p74*, with primers LacZ Fw1 and LacZ Rv1. AcMNPV WT or AcUW1.*lacZ* DNA was used as positive control, NTC (H_2O) control is present. Numbered bands were purified from the gel and analysed further using Sanger sequencing.

The predicted synthesised proteins for each of the sequenced AcC20 polymorphisms of the *egt* and *orf16* region are shown in Figure 5.3. Predicted translation of the longer polymorphism of this region, originated by a deletion, an insertion and smaller mutations, generated three distinct predicted proteins (Fig. 5.3B). The first corresponded to a 406 a.a. long EGT transcribed and translated from 1052bp of the *egt* and 166bp of the inserted nucleotides (predicted protein 1). From this same polymorphism, a short 76 a.a. protein, also seemed to be generated. Transcription started at the inserted fragment, included a truncated segment of *egt* and had a stop codon at *orf16* gene (predicted protein 2, Fig. 5.3B). Finally, a third transcribed predicted protein (3), starts from a fragment of *egt* (between the insertion and deletion) and follows into the turbulent intergenic zone and truncated *orf16* creating a shorter fusion of EGT and DA26 (predicted protein 3; 210a.a.). For the truncated version of the region without any further insertion (Fig. 5.3C), all the genomic changes occurred in frame. Therefore, a unique protein resulting from the fusion of EGT and DA26 was probably generated (predicted protein 4; 570 amino acids; a.a.). The original EGT protein was comprised of 507 a.a., while DA26 is a shorter protein with 226 a.a.

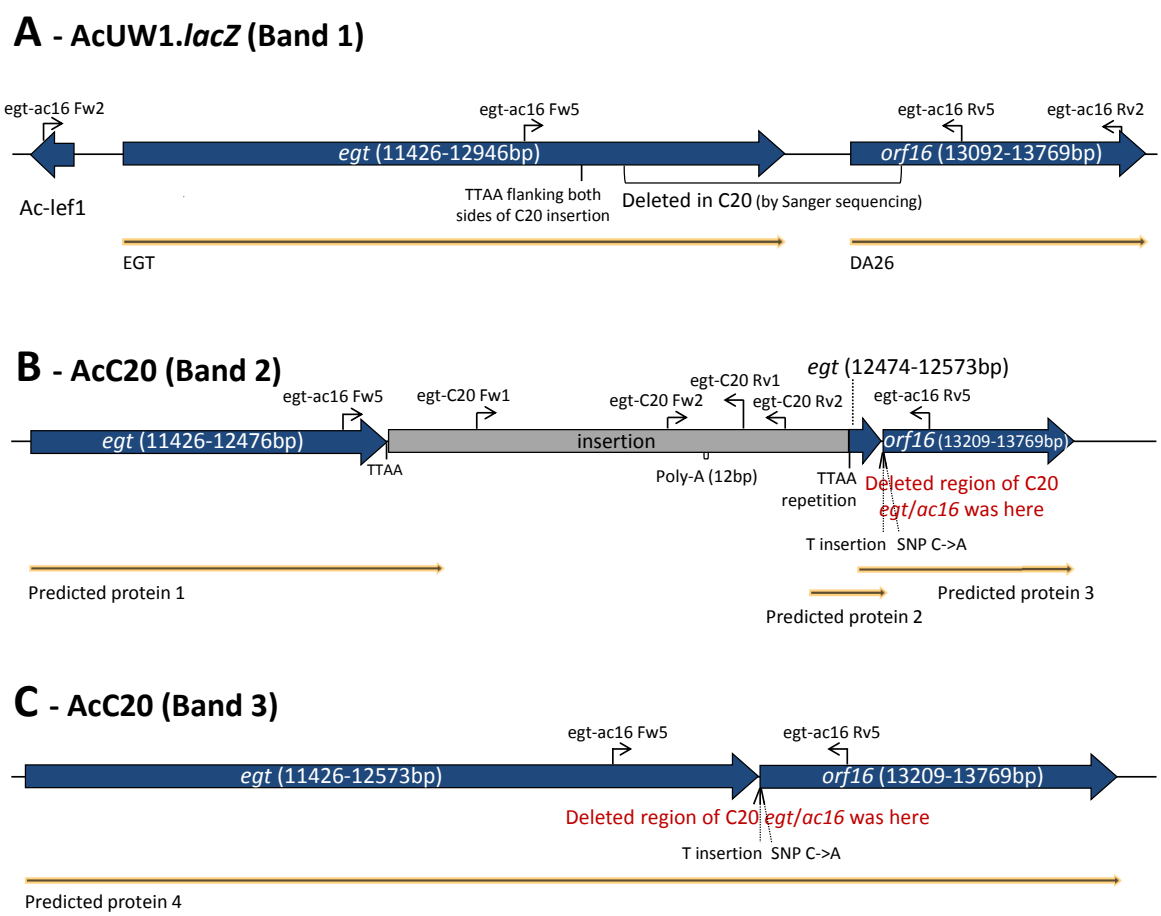


Figure 5.3: Genome organisation of the *egt*/*orf16* region in AcC20
The *egt*/*orf16* sequenced region of: A) AcUW1.*lacZ* (band 1, Fig. 5.2A); B) long fragment (band 2) amplified from AcC20 virus from passage 40 of C20 cells; C) short fragment (band 3) amplified from AcC20 virus from passage 135 of C20 cells; Specific mutations in AcC20 are marked. Primers represented by black arrows. Yellow outlined black arrow represents predicted proteins. Bp count origin was 12nt after start of AcUW1.*lacZ hr1*.

b) The *p94* gene (ODV-E25)

Amplification by PCR with p94 fw and p94 rv1 primers confirmed an insertion in AcC20 (Fig. 5.2B) and sequencing results agreed with its predicted size (287bp). This generated a stop codon 68nt after the beginning of the insertion. The predicted truncated ODV-E25 has ~60% of its original size. Moreover Sanger sequencing revealed nucleotides 915 and 916 of *p94* in AcMNPV are GG and not TA as previously described (Ayres *et al.*, 1994).

c) The *p95* gene

The *p95* was amplified by PCR with primers p95 Fw3 and p95 Rv *Bam*HI 2. With this primer pair, an amplicon with 2725bp and 2347bp was expected for WT AcMNPV and AcC20 respectively, according to whole genome analysis. While the first amplicon was present and confirmed the experiment design was correct, the second questioned if there was a deletion in AcC20 *p95* (Fig. 5.2C). Primers p95 Fw4 and p95 Rv4 were

designed to bind inside the predicted deleted fragment. As shown in Figure 5.2C1 no amplification was observed from these inner primers in the AcC20 DNA suggesting the absence of the target sequences. Sanger sequencing results confirmed the predicted 378bp deletion of virus sequences, although they revealed an insertion of 444bp that replaces the deleted region (Supplementary data Sequence 2). Although the majority of the insertion did not match with any known relevant sequence, 96bp were a partial repetition from *orf71* (142bp from the 5' end). A stop codon was created shortly after the beginning of the insertion, hence, a truncated P95 protein was predicted to be generated.

d) The *lacZ*/*p10*/*p74* coding region

Primers lacZ Fw1 and Rv1 were designed to check the region covering *lacZ*, the remaining 3' *p10* and *p74* sequences. By PCR and Sanger sequencing, predicted WGS sizes and sequences for the products from AcC20 p40 and AcUW1.*lacZ* positive control were observed (472bp and 5227bp respectively; Fig. 5.2D). The p135 had the same sequence as p40.

5.2.1.2 Evolution of AcC20

Given the major changes identified in the AcC20 genome at passage 40 by WGS and confirmed by Sanger sequencing, it was important to explore the possible relation between these and the establishment and maintenance of the persistent infection. Virus clones from previously stored BV samples from passages before (Hi5/AcUW1.*lacZ*) and after C20 cultures were established, were isolated through three rounds of plaque assay in Sf21 cells. Six independent virus clones were obtained both from Hi5/AcUW1.*lacZ* before the cells were cloned into the C20 cell line, and from the C20 cell line itself. These were then amplified in Hi5 cells and the DNA extracted from the concentrated cells. The DNA samples were tested by PCR with the previously described primers for *egt*/*orf16* region, *p94* and *p95* (Fig. 5.4).

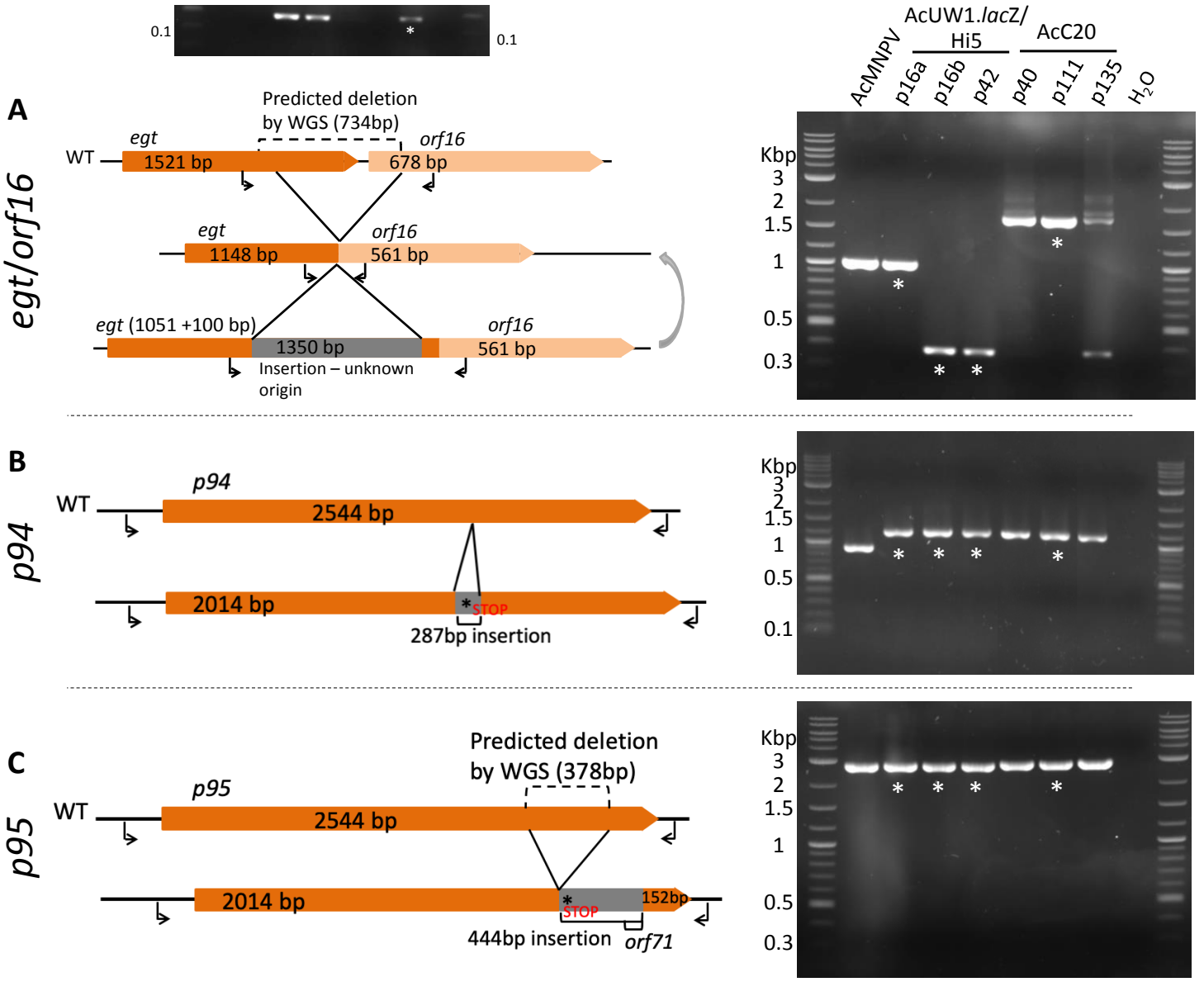


Figure 5.4: Evolution of *egt/orf16*, *p94* and *p95* regions over virus passaging
Schematic representations and electrophoretic analysis of A) *egt/orf16*; B) *p94*; and C) *p95* mutated regions over passaging. Passages of the virus in the initial Hi5 and AcuW1.*lacZ* culture, before the culture was cloned to originate the C20 cell line and passages of the AcC20 virus are presented. All bands marked with * represent single clones (plaque picks). P16a and b represent two clones from the same virus passage. NTC – H_2O.

In the *egt* and *orf16* region of the first passage tested (p16) of the initial non-cloned culture, two clones (p16a and p16b) had different polymorphisms for this region (Fig. 5.4A). While p16a presented the same amplified fragment as the WT AcMNPV DNA, p16b had a deletion. This is the same as described in section 5.2.1.1. and it is present up to passage 42, the one used to isolate the C20 cell line. For AcC20, information is only available from passage 40, the same used to extract DNA for WGS. As observed before, a mixed population exists. From passage 111 the two isolated clones tested presented a higher molecular band that included the previous described insertion (only one clone shown, Fig. 5.4). This insertion seems to have been partially lost over time since the later passage 135 described above presents a higher proportion of the smaller band compared to passage 40 (Figs. 5.2, 5.3 and 5.4). It is clear though that in this region a deletion has occurred before an insertion, which has been lost after passaging of the virus in the C20 culture (Fig. 5.4A).

For *p94*, as observed in the agarose gel and confirmed by Sanger sequencing, all passages tested, from either the early AcUW1.*lacZ* culture in Hi5 or cloned C20 cells, had the insertion described (Fig 5.4B).

Regarding the *p95*, since the occurred insertion is the same size as the region of DNA that was deleted, very little information is available from the agarose gel after PCR. However the sequencing results revealed that the virus passages originated from the initial Hi5/AcUW1.*lacZ* culture had a sequence similar to the WT AcMNPV. Therefore both the insertion and deletion occurred after the cloning of the C20 cell line. However, for this gene it was not possible to conclude if either the deletion or insertion occurred first as all the AcC20 passages tested had both (Fig. 5.4C).

5.2.2 *De novo* assembly for AcC20

The generation of short reads (50bp) from WGS and the repetitive nature of the *hr* regions, has hampered the generation of an independent *de novo* assembly. Hence, a reference based *de novo* assembly is presented, where AcUW1.*lacZ* was used as a reference genome (see 2.4.7). This approach generated a single scaffold with a gap around *egt*/*orf16* region, completed with the Sanger sequencing results presented in section 5.2.1.1. The final assembly is shown in Figure 5.5, which incorporates the read

coverage, CG content, previous discussed major deletions and insertions, other smaller insertions, deletions and SNP.

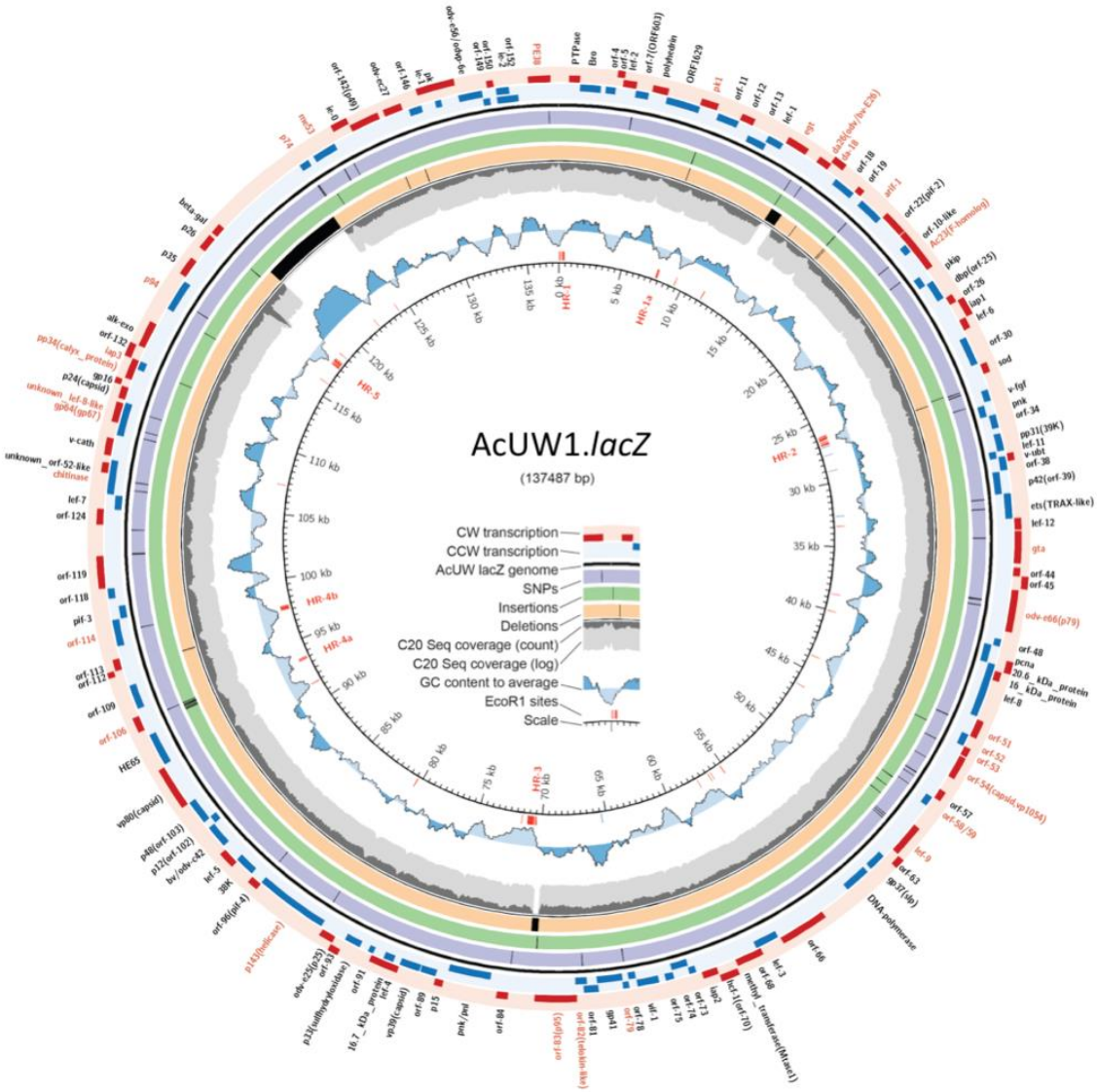


Figure 5.5: AcC20 genome aligned against AcUW1.*lacZ*
Black outside ring represents the AcUW1.*lacZ* scaffold genome (137487bp) against which the AcC20 genome was aligned. AcC20 genes are represented in red and blue according to the direction of transcription, clockwise and anti-clockwise respectively. SNPs, insertions and deletions in the AcC20 genome are marked as black lines in the purple, green and orange ring respectively. The grey ring represents the sequencing coverage for every region and GC content in relation to average is graphically represented in blue. Homologous regions (HR). BWA was used for paired-end mapping to the AcUW1.*lacZ* genome.

Table 5.1 summarizes all the mutated *orf*'s in AcC20 as identified by WGS. Eight of the genes presenting non-synonymous SNP or major INDELS encoded essential proteins for AcMNPV BV replication: PK1, DA26, ARIF-1, AC53, Lef-9, helicase, GP64 and AC83. The *hr2*, thought to be a transcriptional enhancer and origin of replication (Guarino *et al.*, 1986), was also affected in AcC20.

Table 5.1: Variants in AcC20 in comparison to AcUW1.*lacZ* and effects on CDS

Genes	Protein function	Essential *in vitro*	Mutation(s) observed in AcC20	Results in the translated protein
pk1 (*ac10*)	PK-1 is a serine/threonine kinase. Phosphorylates viral or cellular proteins essential for nucleocapsid assembly (Liang *et al.*, 2013).	Yes	1 nt inserted and 1 nt deletion	Frame shift from insertion is corrected by the deletion, 5 a.a. affected
egt	Encodes an ecdysteroid UDP-glucosyl transferase (EGT) that catalyzes the transfer of glucose from UDP-glucose to ecdysteroids (insect molting hormones). Thus allowing the virus to interfere with normal insect development (O'Reilly and Miller, 1989, 1990).	No	373 nt deleted at the 3' end. 1350 nt inserted in one polymorphism.	Possible merge with truncated DA26 or generation of a truncated EGT with 54a.a. in the C-terminal transcribed form the insertion (see 5.2.1.1a).
orf16	Encodes for DA26 which binds to and regulates levels of IE0 and IE1. A domain within the acidic transcriptional activation domain is responsible to bind to the proteins (Nie *et al.*, 2009).	Yes	635 nucleotides deleted covering 3' end of *egt* and 5' end of *da26*.	Possible merge with EGT or generation of a protein with transcription starting from the 3' end of EGT (see 5.2.1.1a).
da18	Not described.	Not known	1 nt deleted	Frame shift extends the 5' end 45 a.a.
arif-1 (*ac20* and *ac21*)	ARIF-1 is a component of the actin rearrangement-inducing complex therefore locating at the plasma membrane (Dreschers *et al.*, 2001).	Yes	1 non-synonymous SNP, 3 insertions of 1nt, 2 deletion of 1nt.	5' extension of peptide. Note: region had low coverage
ac23	F gene homolog (AC23) that cannot compensate for the loss of *gp64*. It is found in BV and represents an important viral pathogenicity factor in larvae infected with AcMNPV (Lung *et al.*, 2003).	No	1 non-synonymous and 1 synonymous SNP	No change predicted.
hr2	Implicated both as transcriptional enhancers and origins of DNA replication (Guarino *et al.*, 1986).	Yes	Many SNP upstream of Ac-fgf fibroblast growth factor. 290bp deletion.	Not transcribed.
gta	Not described.	Not known	2 non-synonymous SNP	No change predicted.
p79 (odv-e66)	Core gene encoding for ODV-E66, a specific envelope protein. It is not essential for virus replication but plays an important role in oral infectivity (Xiang *et al.*, 2011).	No	1 non-synonymous SNP, 1 synonymous SNP	No change predicted.
ac47 - *ets* (TRAX-like)	Not described.	Not known	2 synonymous SNP	-

(Table 5.1 cont.) Genes	Protein function	Essential *in vitro*	Mutation(s) observed in AcC20	Results in the translated protein
orf51 (*ac51*)	*ac51* is a late gene required for efficient nuclear egress of AcMNPV nucleocapsids and is essential for virulence *in vivo* (Qiu *et al.*, 2018).	No	1 non-synonymous SNP	No change predicted.
orf52	Not described.	Not known	1 insertion	Frame shift extends 5' sequence 70a.a.
orf53	Although the role of the encoded AC53 protein is not entirely clear, it is known to be involved in nucleocapsid assembly and essential for BV production (Liu *et al.*, 2008).	Yes	1 non-synonymous SNP	No change predicted.
orf54	Encodes for VP1054, a minor capsid protein and plays a crucial role in the transport of capsid proteins like VP39 (the major capsid protein) to the nucleocapsid assembly site (Guan *et al.*, 2016).	Yes	1 non-synonymous SNP	No change predicted.
orf57	Not described.	Not known	1 synonymous SNP	-
orf58/*orf59*	Although without a clear function, the encoded proteins associate with AcMNPV BV and ODV (Braunagel *et al.*, 2003; Wang *et al.*, 2010b)	No	1 insertion, 1 non-synonymous SNP	Results in fusion of orf58-59
fp25k	Function unknown. Locus prone to mutation. A spontaneous insertion creates the characterized few polyhedra phenotype (Cary *et al.*, 1989; Wang *et al.*, 1989; Kelly *et al.*, 2006).	No	1nt insertion	Results in truncation of 70% of the C-terminal.
lef-9	Encodes for Lef-9, one of the four equimolar protein subunits of the AcMNPV DNA-dependent RNA polymerase. This allows specific transcription from late and very late promoters (Guarino *et al.*, 1998).	Yes	3 non-synonymous SNP	No change predicted.
orf79	Although the role is not clear, it was suggested that AC79 is required for efficient BV production (Wu and Passarelli, 2012).	No	1 non-synonymous SNP	No change predicted.
orf82	Late gene that encodes ORF82, required for BV production (Gauthier *et al.*, 2012).	No	1 non-synonymous SNP	No change predicted.
p95 (*ac83*)	PIF protein Required for assembly of the PIF complex as well as ODV and BV nucleocapsid assembly (Javed *et al.*, 2017)	Yes	Deletion of 378bp and insertion of 444bp	Results in truncation of the P95 to 683a.a. instead of 847a.a. Last 12a.a are a variant since they are from the inserted region (see 5.2.1.1c).
orf84	Not described.	Not known	1 non-synonymous SNP	No change predicted.
p94	Core gene that encodes ODV-E25. Protein associates with BV and ODV (Braunagel *et al.*, 2003; Chen *et al.*, 2012)	No	1nt insertion, 287bp insertion, 1 non-synonymous SNP and 1 synonymous SNP	Results in truncation of the C-terminal. The protein is 533a.a. instead of 803a.a. (22a.a. are transcribed from the insertion; see 5.2.1.1b)

(Table 5.1 cont.) Genes	Protein function	Essential *in vitro*	Mutation(s) observed in AcC20	Results in the translated protein
p143 (Helicase)	Encodes the helicase, essential for DNA replication (Kool *et al.*, 1994).	Yes	1 non-synonymous SNP	No change predicted.
orf-96	Encodes *per os* infectivity factor 4 (PIF-4) so it is required for oral infection (Fang *et al.*, 2009b)	No	1 non-synonymous SNP	No change predicted.
orf106/*orf107*	Not described. Probably these two genes are a single *orf* (Harrison and Bonning, 2003).	Not known	9 insertions 1 SNP	Insertions in *orf106* result in frame shift which deletes stop codon. So 5' sequence extends to the end of *orf107*.
orf112	Not described.	Not known	1 non-synonymous SNP	Frame shift extends 5' sequence to the end of *orf113*.
orf114	ODV and BV associated protein (Braunagel *et al.*, 2003; Wang *et al.*, 2010b)	No	1 non-synonymous SNP	No change predicted.
orf120	Not described	Not known	1 synonymous SNP, 1 non-synonymous SNP	No change predicted.
pk-2	Present both early and late during virus infection but it does not appear to have a significant influence on virus replication. (Li and Miller, 1995)	No	1 non-synonymous SNP	No change predicted.
chitinase	Together with AC127 (cathepsin), it is responsible for the liquefaction of insects late in infection. Gene product is localized to the cytoplasm and it associates with BV (Hawtin *et al.*, 1995; Wang *et al.*, 2010b).	No	3 non-synonymous SNP	No change predicted.
gp64	Encodes GP64, a low pH activated envelope fusion protein required for BV to enter the cells (Monsma *et al.*, 1996).	Yes	3 non-synonymous SNP	No change predicted.
ac131	The pp34 encoded protein is an integral component of the calyx/polyhedron envelope protein (Zuidema *et al.*, 1989; van Lent *et al.*, 1990)	No	Insertion	Frame shift extends 5' sequence 69a.a.
ac138	Encodes for P74 *per os* infectivity factor. It is required for oral infection but dispensable for infection of cultured cells. (Kuzio *et al.*, 1989)	No	Deletion of 1476nt in the 3' end (out of 1938nt)	Deletion removes 76% of the 3' coding region (see 5.2.1.1d).
me53	ME53 protein associates with BVs and ODV and is required for efficient budded-virus production (de Jong *et al.*, 2009).	No	1nt insertion 2 non-synonymous SNP	Insertion adds stop codon 3a.a. down
ie0	IE0 supports early gene and very late virus replication (Huijskens *et al.*, 2004; Stewart *et al.*, 2005)	No	1 synonymous SNP	-
pe38	Activates DNA replication (Kool *et al.*, 1994)	No	1 non-synonymous SNP	No change predicted.

For the essential genes, the variants in the protein were checked using the conserved domain database by NBCI. For GP64 and helicase, although the mutations were in a conserved region, they still aligned with proteins from the same family in baculovirus. For PK1, DA26 and ARIF-1, the affected amino-acids were not in a region of high conservation. In the case of Lef-9 and AC53, although the proteins belong to very conserved families, the variants were in regions that the AcMNPV WT has variants itself. Finally, P95 truncation affected the C-terminal of the protein, which did not seem to be a conserved dominion.

5.2.3 Few polyhedra phenotype and the production of apoptotic vesicles

In Hi5 cells freshly infected with AcUW1.*lacZ*, the parental virus used to generate the original persistent infection, generous amounts of polyhedra are produced in their nuclei during the very late phase of virus replication (Fig. 5.6Ai). However, in various passages of the C20 cell line it was observed that polyhedra were only present in a small proportion of the cells. Furthermore, these cells produced less polyhedrin protein than in a normal infection (Fig. 3.1).

To analyse if the AcC20 was responsible for this phenotype in the C20 cells, BV in the media were separated from the cells by centrifugation. This BV sample was then amplified in Sf9 cells and used to infect Hi5 cells (Fig. 5.6Aii). Finally, a virus with a previously described mutation in the *fp25k* gene, AcFP25 [previously AcdefrTp35^{r} (Kelly *et al.*, 2006); Fig. 5.6Aiii], was used to infect Hi5 cells. The AcFP25 virus has an insertion of an adenine in a run of seven of these nucleotides generating a truncation of the coding region shortly after, resulting in the few polyhedra phenotype (FP) (Kelly *et al.*, 2006, 2008). When comparing the three panels in Figure 5.6A, it is clear that although AcC20 is capable of producing polyhedra, the quantity per cell is similar to that produced by an FP25 mutant virus. Therefore this gene in AcC20 was analysed by PCR (primers RDP737 and RDP738; table 2.4) and Sanger sequencing at passage 40 of C20 cells. The results revealed that AcC20 had an adenine insertion resulting in a frame shift and therefore generated the truncated protein previously described (Fig. 5.6A1; Kelly *et al.*, 2006). This was also confirmed by WGS (Table 5.1). This mutation was also present in later passages of AcC20 (p111, p135) and in the two early stocks of AcUW1.*lacZ*/Hi5 (p16, p42).

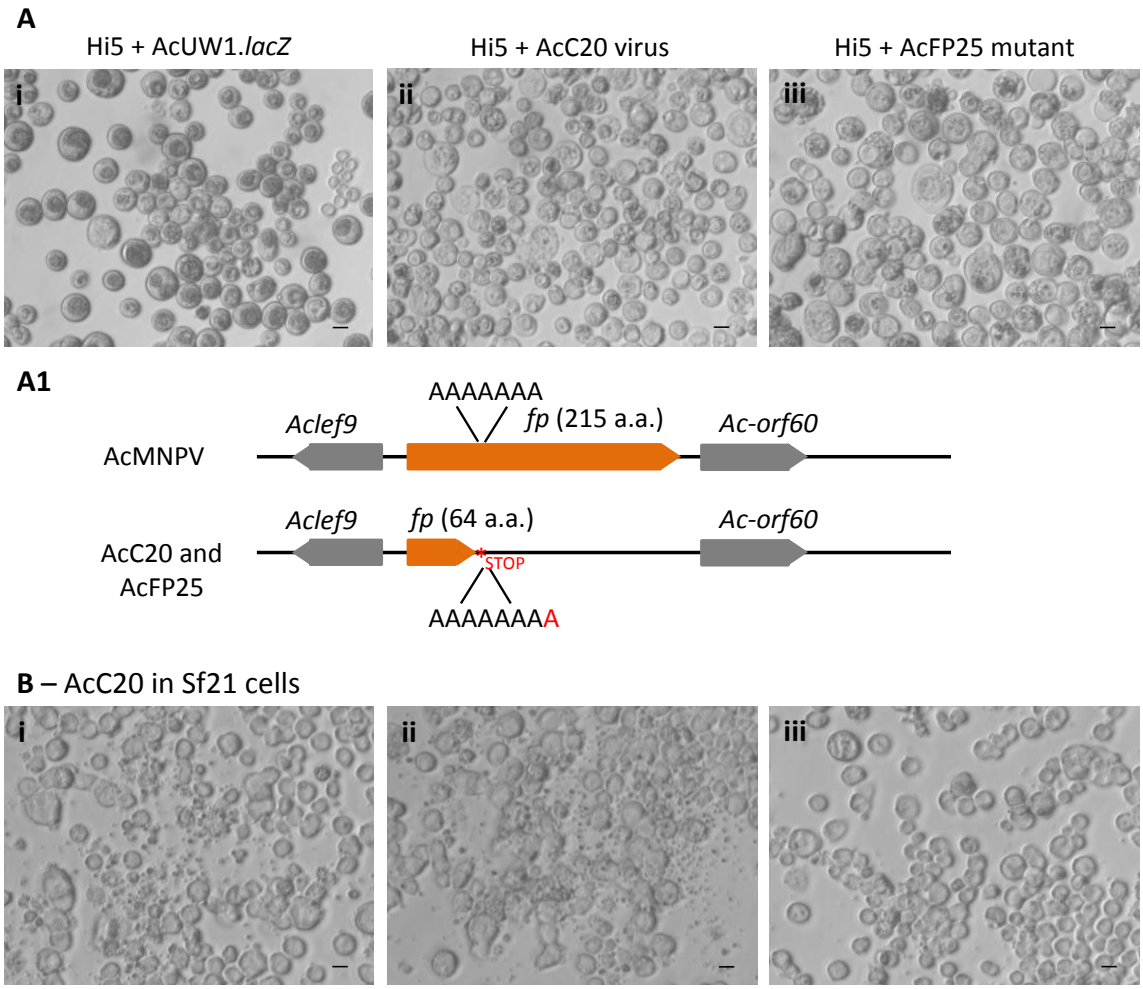


Figure 5.6: Imaging of lytic and persistently infected cells
A) comparison of nuclear polyhedra in Hi5 cells infected with: i) AcUW1.*lacZ*; ii) AcC2O virus previously amplified in Sf9 cells; iii) AcFP mutant. A1) genomic diagram of *fp25k* and neighbour genes in AcMNPV WT and AcC20. Bi, ii, iii) examples Sf21 cells infected with AcC20 to check for the presence of apoptotic vesicles. Images are representative and were obtained using phase contrast microscopy. Scale bars 20µm.

Plaque assays in Sf21 cells were used frequently to titrate the AcC20 virus and to isolate single clones. A close observation of some of the plaques generated revealed the presence of apoptotic vesicles (Fig. 5.6Bi-iii). The phenomenon, described as plasma-membrane blebbing most often results from a defect in the anti-apoptotic *p35* (Clem *et al.*, 1991; Kelly *et al.*, 2006). This was observed in a deletion mutant lacking the *p35* basal promoter region and the first 391 bases of the *p35* coding region (Kelly *et al.*, 2006). The *p35* from passage 40 of AcC20 virus and control AcMNPV WT DNA was amplified with primers RDP1275 and RDP1279 (Table 2.4). Since reverse primer aligned to *hr5*, adjacent to *p35*, several products were generated, so the correct bands were purified from an agarose gel prior to Sanger sequencing. Also, clonal DNA from passages of the Hi/AcUW1.*lacZ* early culture (p16a, p16b, p42) and subsequent AcC20 passages (p111, p135) were tested at the same time (data not shown). The sequencing results revealed a

possible deletion of 477bp in the 3' end of the *p35* gene. On the other hand, no mutations were observed for this gene using WGS (See 5.2.2).

5.2.4 RNA sequencing and data analysis

Total RNA was extracted from mock or AcMNPV-infected Hi5 cells and non-challenged or superinfected C20 cells at precise time points using the TRIzol method (see 2.4.9). All RNA samples were analysed by electrophoresis to check for degradation and DNA contamination (Fig. 5.7).

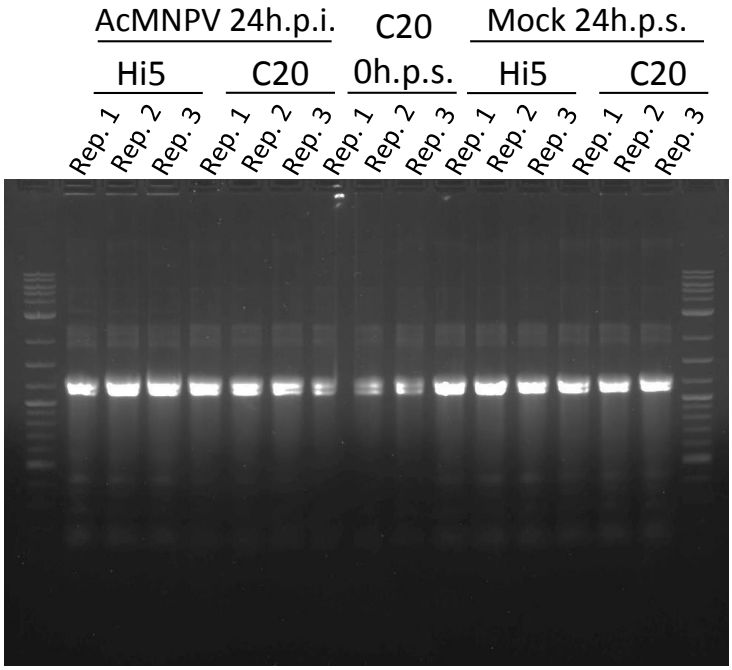


Figure 5.7: Example of electrophoresis for extracted RNA samples.
Sample purity and presence of contaminants was analysed by electrophoresis, 5µl of each RNA sample was checked in a 1% agarose gel.

Strand-specific RNA-seq libraries using 4µg of total RNA were generated as described before [work undertaken by C. Bannach; (Zhong *et al.*, 2011; Chen *et al.*, 2013)]. Purified and pooled libraries quality was assessed by agarose gel electrophoresis (data not shown). The mean size of DNA fragments in each of the pooled libraries was between 200-300bp. Illumina® RNA-seq was outsourced to Epigenomic core using the Illumina® HiSeq4000 platform (Fig. 5.8C; Weill Cornell Medical School, NYC, US). Per library at least 13 million filtered reads were generated except for the library "C20 18h p.i. replicate 2". Since this library only had 100 raw reads, the sample was removed from downstream analysis.

The reads were mapped against the tetra- and *Alphanodavirus* genome as a quality control and to remove those sequences. No mapped reads were obtained for the first virus. Nevertheless, since both Hi5 and C20 cells have a nodavirus latent infection (see 3.2.4.), several hundred per library were mapped to alphanodovirus for all samples. These reads were filtered off.

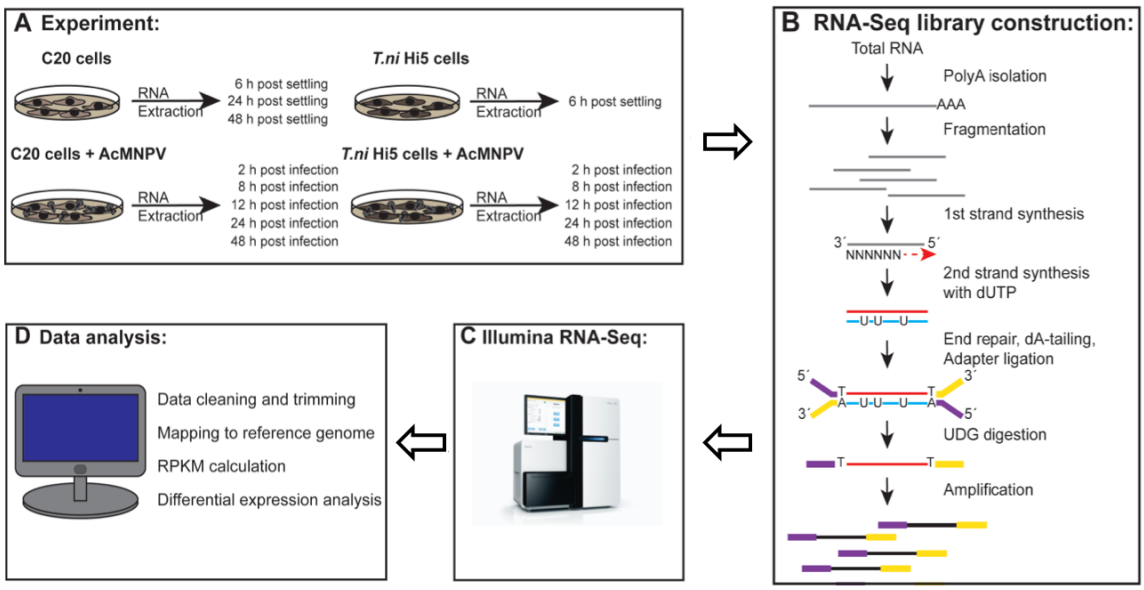


Figure 5.8: C20 RNA-seq experiment summary.
A) Total RNA from C20 cells and Hi5 cells or AcMNPV-challenged C20 and Hi5 cells was extracted at the required time points. B) Strand-specific libraries for RNA-seq were generated as previously described (Zhong et al., 2011). C) Sequencing was outsourced (Illumina®, Weill Cornell Medical School). D) Raw reads were cleaned and sub sequentially trimmed and mapped to the AcMNPV genome. Reads per kilo base per million mapped reads (RPKM) values were calculated (K. Bao, BTI).

The percentage of reads aligning to the AcMNPV genome in virus-infected Hi5 cells increased over time. Even though at 2h p.i only 0.3% of cleaned reads were virus-specific, by 48h p.i. this value reached 95%. On the other hand, for non-challenged C20 cells ~2% were mapped to the viral genome at any time point. This indicates that although virus transcription was detected at all times in C20 cells, transcription occurred at a low level (Fig. 5.9). In the superinfected C20 cells, the number of reads mapping to the AcMNPV genome increased from 1% at 12h p.i. to 4.2% at 48h p.i. This suggests that C20 cells can be superinfected although the secondary virus is expressed at a low level.

After this preliminary analysis, the reads per kilobase of transcript per million (RPKM) reflecting AcMNPV gene expression were calculated (K. Bao). Due to transcripts overlapping, resulting from overlapping *orfs*', RPKM values were adjusted as previously

described (Chen *et al.*, 2013). When the upstream viral gene was expressed very strongly this led to negative RPKM values. Since gene expression cannot be negative those values were considered as 0 for further analysis of viral gene expression. For *orf*s that share 3´transcriptional co-termination sites with adjacent genes and/or their mRNAs overlap with adjacent genes to a large extent, expression of their transcripts could not be determined accurately (Chen *et al.*, 2013). This occurred for *orf44*, *orf45*, *etm*, *orf56*, *lef-10*, *orf78*, *orf81*, *orf113* and *alk-exo*, therefore analysis of transcription in these genes was not carried forward.

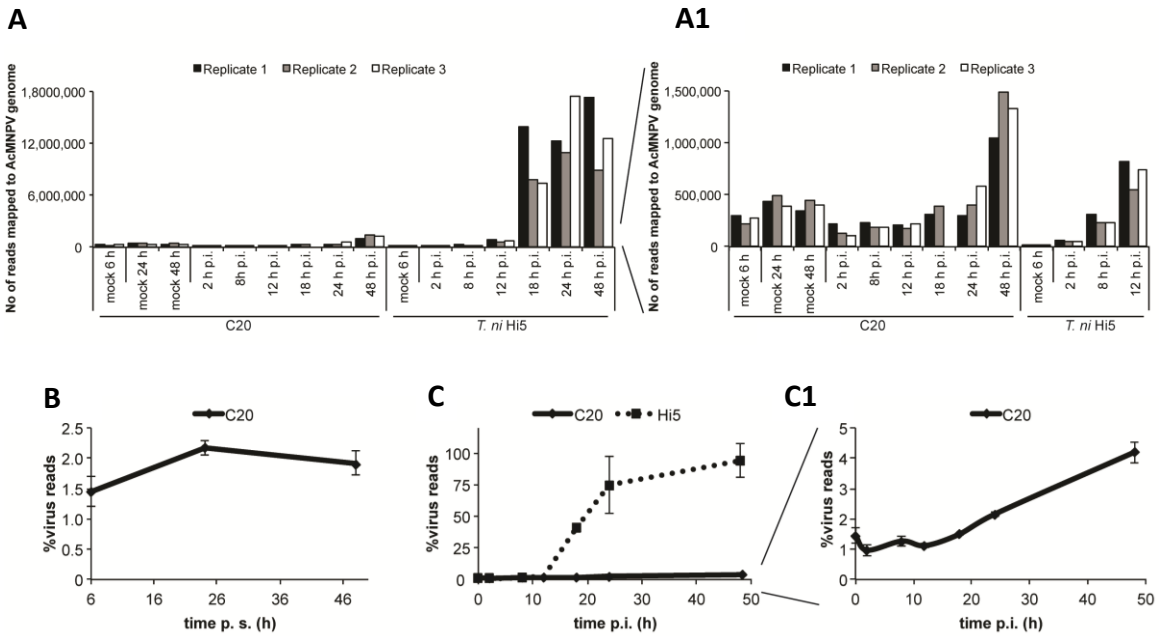


Figure 5.9: RNA-Sequencing statistics overview [adapted (Bannach, 2018)].
A) Total number of Illumina® sequencing reads for each replicate that was mapped to the AcMNPV genome. B, C) Average percentage of reads (as % of cleaned reads) in B) unchallenged C20 or C) AcMNPV infected C20 and Hi5 cells that were mapped to the viral genome. A1 and C1 represent magnifications of A and C respectively.

5.2.4.1 Virus transcription in C20 cells

The RNA sequencing analysis revealed that all virus genes were expressed in C20 cells. In superinfected C20 cells, *p10* expression was used as a marker to follow AcMNPV WT infection. The parental AcUW1.*lacZ* virus used to establish the persistent infection had most of the *p10* coding region replaced by *lacZ* (see 1.5; (Weyer *et al.*, 1990)) although the 5´leader sequence, 132nt at the 3´end and the 3´UTR were retained (Weyer *et al.*, 1990). This justifies the presence of a few reads mapped to *p10* in unchallenged C20 samples. However, in C20 superinfected cells, *p10* expression increased dramatically (160-fold at 48h p.i. compared to unchallenged C20) demonstrating superinfection had

occurred. Furthermore, the expression pattern was similar to what was observed in a lytic infection in Hi5 cells with a lag phase up to 18h and higher expression levels occurring very late in infection, at or after 48h p.i. since expression did not peak up to 48h p.i. (data not shown).

Figure 5.10 illustrates the temporal expression of viral genes in each of the cell lines and infections, by showing the six most abundant expressed genes during the time course. In a lytic infection at 2h p.i., immediate early genes *pe38*, *ie2* and *ie1* were the highest expressed genes, accounting for 38% of the total mRNA present. The expression of late gene *p6.9* at this time point was, however, surprisingly high (18%), along with expression of *me53* and *polh* (7% each). The *gp64* only accounts for the sixth highest expressed genes at 8h p.i., where it represents 4% of total mRNA. From 12h p.i. *p6.9* is highlighted, representing 14%, 44%, 52% and 35% of all viral transcripts at 12h p.i., 18h p.i., 24h p.i. and 48h p.i. respectively. The very late gene *polh* represents 40% of all viral transcripts at 48h p.i. although the expression for these transcripts does not peak up to this last time point. At the same time point *p10* only represented 2% of all viral mRNA present (Fig. 5.10).

In the non-challenged C20 cell less variation was observed between the three time points. Overall only eight genes featured in the graph covering the six highest expressed at each time point: *p6.9*, *polh*, *orf76*, *orf58*, *ac-bro*, *ctx*, *v-ubi*, and *odv-ec27*. The late genes *p6.9* and *orf76* were the most abundant genes reaching between 32-38% and 7-12% respectively during the experiment. Each of the remaining genes accounted for less then 4% of the mRNAs present at all time points tested.

The superinfection of C20 cells introduced some variability to the gene expression. Here, 12 genes accounted for the six highest expressed throughout the experiment time: *p6.9*, *polh*, *orf76*, *orf75*, *orf58*, *orf74*, *p94*, *ac-bro*, *ctx*, *v-ubi*, *odv-ec27* and *orf82*. The *p6.9* and *orf76* genes were still the highest transcribed genes from 8h p.i. to 48 h p.i. accounting for 17-32% and 6-11% of total mRNA respectively. Expression of *polh* only reached the top six genes at 48h p.i. and accounted for 5% of total mRNA.

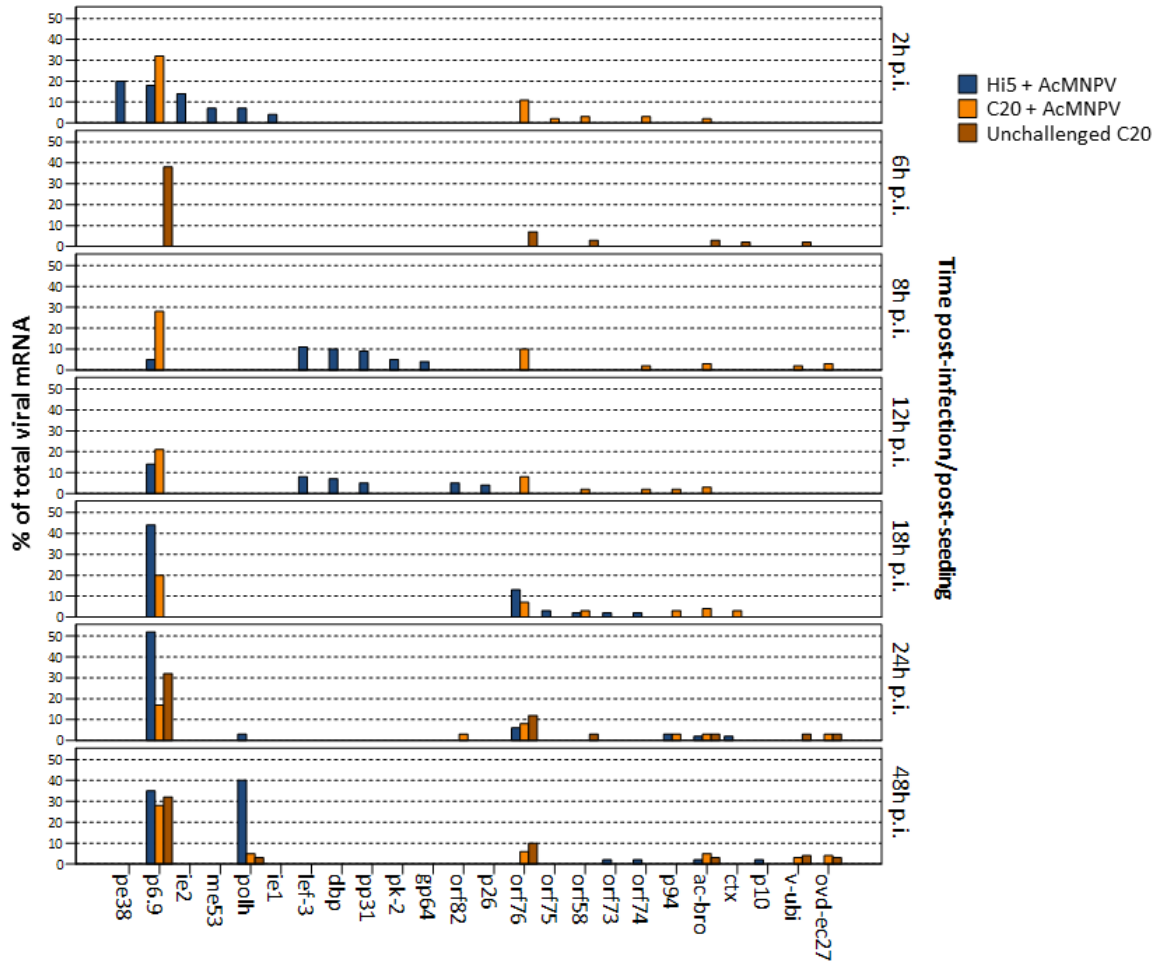
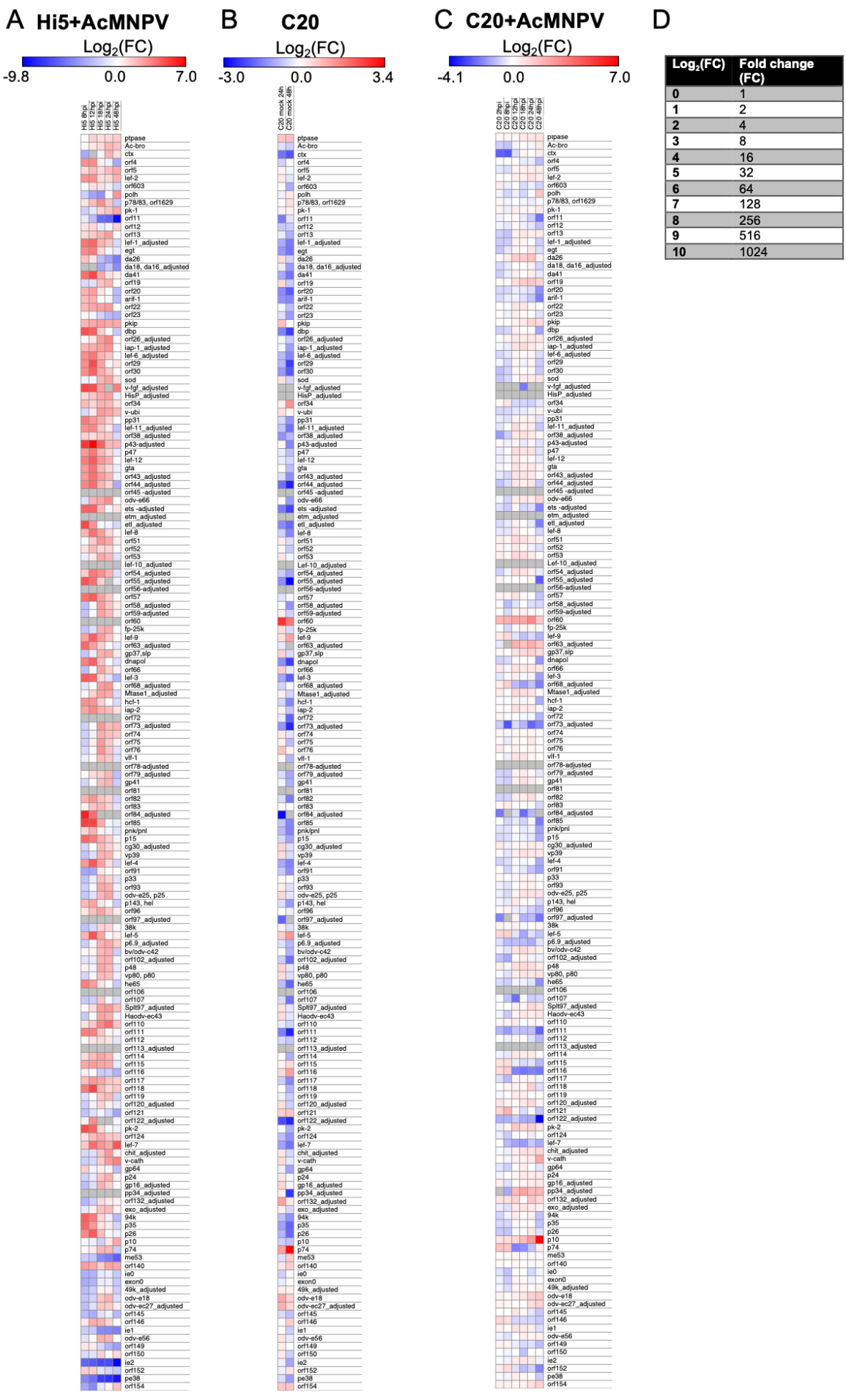


Figure 5.10: Temporal expression of viral genes in AcMNPV-infected Hi5 cells, superinfected and unchallenged C20 cells
Percentage of total mRNA for the six most abundant genes at each time point in AcMNPV-superinfected C20, infected Hi5 and unchallenged C20 cells.

5.2.4.2 Viral transcription patterns in C20 cells

a) Differential gene expression analysis

According to the results in the previous section, gene expression did not vary greatly over time either in superinfected or non-challenged C20 (Fig. 5.10). To investigate the results further, differential gene expression (DGE) analysis was performed. Firstly, the data from the lytic infection was used to observe the landscape of viral gene expression. To do this, the $\log_2$ difference of normalized gene expression for each gene and time point of virus-infected cells was compared to AcMNPV-infected Hi5 cells at 2h p.i (Fig. 5.11A). Gene expression for immediate early genes peaked at 2h p.i (e.g. *ie1*), early genes peaked between 8-12h p.i (e.g. *lef-3*), late genes between 18-24h p.i. (e.g. *p6.9*) and very late genes at 48h p.i. (e.g. *polh*).



Log_2(FC)	Fold change (FC)
0	1
1	2
2	4
3	8
4	16
5	32
6	64
7	128
8	256
9	516
10	1024

Figure 5.11: Differential gene expression analysis of superinfected C20, infected Hi5 and unchallenged C20 cells.
Normalised viral gene expression (RPKM values) of A) AcMNPV-infected Hi5 cells at the indicated time points p.i. were compared to AcMNPV-infected Hi5 cells at 2h p.i. or B) C20 at 24 or 48h p.s. were compared to normalised viral gene expression of C20 cells 6h p.s. or C) C20 6h p.s. was compared to normalised viral gene expression of superinfected C20 cells at the indicated time points. A-D) fold change (FC) of viral gene expression was presented on a $\log_2$ scale as heat maps. Grey indicated no detected gene expression.

Similarly, in unchallenged C20 cells, normalized gene expression at 6h p.s. was used as a baseline to compare with gene expression at 24 and 48h p.s. From the heat map in Figure 5.11B it is clear that although gene expression changed over time, only a few genes were up regulated. When the fold change was calculated and statistically analysed, little difference in gene expression was observed at 24h p.i. since 93.2% of the viral genes were not down or up regulated. At 48h p.i. 17.7% and 4.1% of the genes were significantly down or up regulated respectively ($p_{\text{student's t-test}}<0.05$). The highest observed fold change occurred for *p74* which at 48h p.i. was 10.2 times more highly expressed than at 6h p.i.

Finally, using the $\log_2$ difference, the DGS of superinfected cells was calculated against unchallenged C20 cells at 6h p.i. Here the pattern of gene expression obtained differed from that observed in a lytic infection (Fig 5.11C). In comparison, immediate early genes peaked at 2h p.i. in AcMNPV infected Hi5 and around 12h p.i. for superinfected C20 cells (*ie0*, *ie1*, *ie2* and *pe38*). Even though *p6.9* was the highest expressed gene in both unchallenged and superinfected cells, *p6.9* expression decreased after superinfection (Figs. 5.10 and 5.11). On the other hand, *polh* exemplifies how some genes seem to have followed the same transcription pattern as in a productive infection, although the fold change was not as accentuated (Fig. 5.11A and C).

b) Cluster analysis in infection phases of viral mRNA abundance over time

For a wider and more comprehensive understanding of transcription phases in superinfection, cluster analysis of viral mRNA abundance over time was carried out.

Genes were clustered into immediate early, early, late and very late based on the promoter sequences: genes containing a TAAG promoter motif were considered as late; other genes were considered early, as described (Fig. 5.12A; Chen *et al.*, 2013). If both motifs were present, the allocated phase corresponded to the time point of maximum expression. Finally, immediate early (*ie0*, *ie1*, *ie2, pe38*) and very late genes (*p10*, *polh*)

were classified based on the literature (Guarino *et al.*, 1986; Van Der Wilk *et al.*, 1987; Chisholm and Henner, 1988; Krappa and Knebel-Mörsdorf, 1991; Ayres *et al.*, 1994; Yoo and Guarino, 1994; Rohrmann, 2013). To obtain the transcription profiles for phases over time, at each time point, normalized gene expression values for all genes belonging to that phase were added up. This analysis enclosed a total of 147 genes distributed by immediate early (n=4), early (n=61), late (n=80) and very late (n=2) (Fig. 5.12).

In accordance with Figure 5.9, in the productive infection viral gene expression increased over time (Fig. 5.12A). As observed in Figure 5.12B, expression profiles of each phase of infection had the structure of a Gaussian distribution in a productive infection. Immediate early and early genes reached a maximum of transcription at 2h p.i. and 8-12h p.i. respectively (Fig. 5.12B). Late gene expression reached a maximum between 18-24h p.i. but surprisingly a second smaller peak was observed at 2h p.i. (Fig. 5.12B). Very late gene expression had a lag phase up to 24h p.i., from where expression started to increase without peaking up to the last sampled time point (Fig 5.12B).

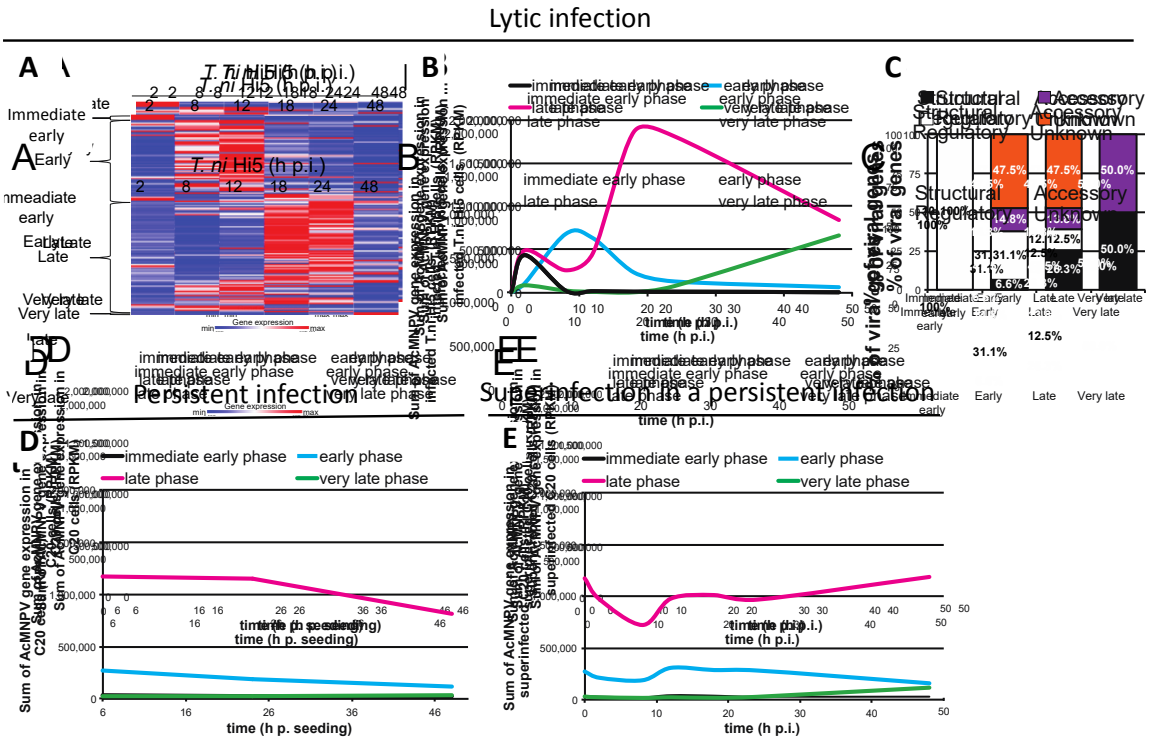


Figure 5.12: Viral gene expression in a lytic or persistent infections and superinfection in a persistent infection.

A) Schematics of normalised gene expression (RPKM) of AcMNPV-infected Hi5 cells. Normalised gene expression from immediate early (n=2), early (n=61), late (n=80) or very late genes (n=2) at indicated time points p.i. or p.s. was added up to obtain viral expression profiles of B) AcMNPV-infected Hi5 cells, D) unchallenged C20 cells or E) AcMNPV-superinfected C20 cells. C) Percentage of viral genes of AcMNPV-infected Hi5 cells corresponding to their function. Heat map was visualised using Morpheus Software.

According to their known function, genes were classified into regulatory, structural or accessory and the percentage for each phase was calculated (Fig. 5.12C). Of the genes with a known function from the early phase of infection, the majority encode for regulatory proteins (31.1%) while approximately 25% of late genes encode structural proteins (Fig. 5.12C).

In unchallenged or superinfected C20 cells, although gene expression changed over time, there were no obvious peaks for any of the phases of infection (Fig. 5.12D, E) as were evident in AcMNPV-infected Hi5 cells (Fig. 5.12B). Late genes were the most highly expressed, however the very late phase only included *polh* since AcC20 lacked *p10*. On the other hand, when superinfecting C20 cells, it was clear that the transcription profile of a persistent infection differed from a lytic infection (Fig. 5.12B and E) as the typical temporal regulated phases of infection were not observed (Fig. 5.12E). For all phases of infection, gene expression decreased until 8h p.i. and then fluctuated over time. For the very late phase, gene expression increased slightly overtime without peaking, while for the other phases transcription did not increase above baseline (Fig. 5.12E; C20 0h p.i. corresponds to C20 mock 6h p.s.). During the course of the superinfection, immediate early genes were the least expressed followed by very late genes, early genes and finally late genes as the highest expressed (Fig. 5.12E). This order matched what happened in non-challenged cells. In this case, no transition in between phases of infection occurred (Fig. 5.12E). Nevertheless, the data showed that superinfection was possible in C20 cells.

Virus gene expression in unchallenged C20 cells was classified into 3 profile groups based on cluster analysis by the hierarchical complete linkage method (Fig. 5.13B; C. Bannach; Morpheus Software). For genes clustered in group 1, transcription decreased over time. More than half of the genes here classified had an early promoter motif and 28.1% had a regulatory function. (Fig. 5.13C and D). For group 2 and 3, the majority of genes contained a late promoter motif, 62.1% and 80% respectively (Fig 5.13A and B). From the genes with a known function, the majority of genes clustered in these two groups encoded for structural proteins (27.6% and 32.0% for group 2 and 3 respectively; Fig. 5.13C, D).

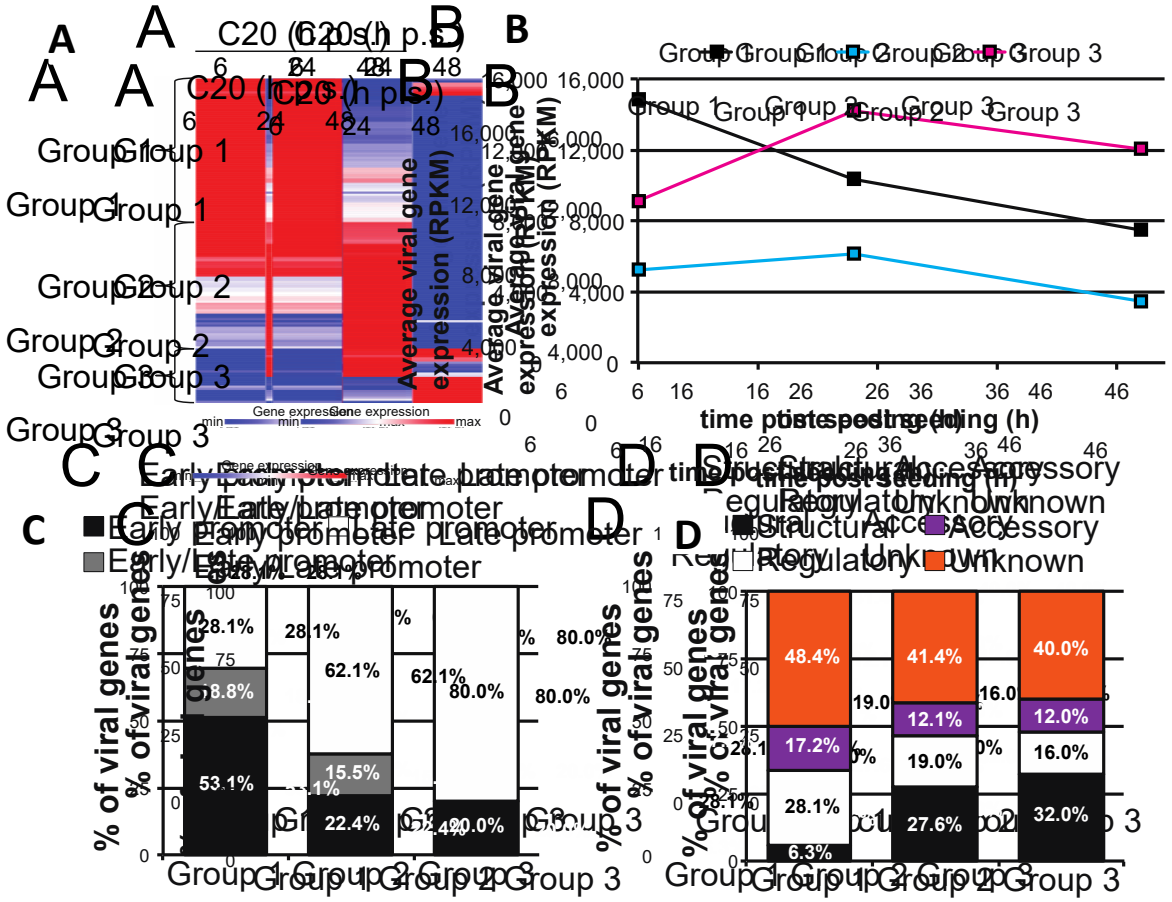


Figure 5.13: Cluster analysis of AcMNPV gene expression pattern in unchallenged C20 cells.
A) Schematics of normalised gene expression (RPKM) in C20 cells at 6, 24 and 48h p.s. Based on their gene expression pattern, viral gene expression was clustered by complete-linkage clustering into three groups: Group 1 (n=64), Group 2 (n=58) and Group 3 (n=25). B) Average gene expression pattern for each group as normalised gene expression levels. C-D) Percentage of viral genes from group 1-3 corresponding to C) function and D) promoter motifs. Heat map was visualised using Morpheus Software.

The same clustering method was applied to the superinfected C20 cells. Four clusters were identified, group A, B, C and D (Morpheus Software; Fig. 5.14B). Group B, C and D genes had an expression pattern similar to that seen in a lytic infection with genes from immediate early, early and very late phase respectively. However maximum expression was achieved later than in a normal virus infection, at 8h p.i., 12-18h p.i. and 48h p.i., respectively for groups B, C and D. On the other hand, transcription of group A genes decreased over the time in this superinfection (Fig. 5.14B). Even though the profiles resembled a lytic infection, distinct genes clustered in these groups (Fig. 5.14C). In group B, where expression peaked at 8h p.i., the majority of the genes clustered had late promoters (83.3%) and most of the genes with known function encoded for structural proteins (27.8%) (Fig. 5.14C and D). For group C, although the expression profile mimicked an early phase of a lytic infection, 46.8% of the encoded proteins had a late

promoter while 33.9% had an early promoter. However, most of the genes encoded regulatory proteins (40.3%; Fig. 5.14D). Group A was unexpected since this expression pattern is not observed in a productive infection. Genes clustered here had mainly early promoter motifs (71.1%) although the majority of them encode proteins for which function is still unknown (63.1%; Fig. 5.14D).

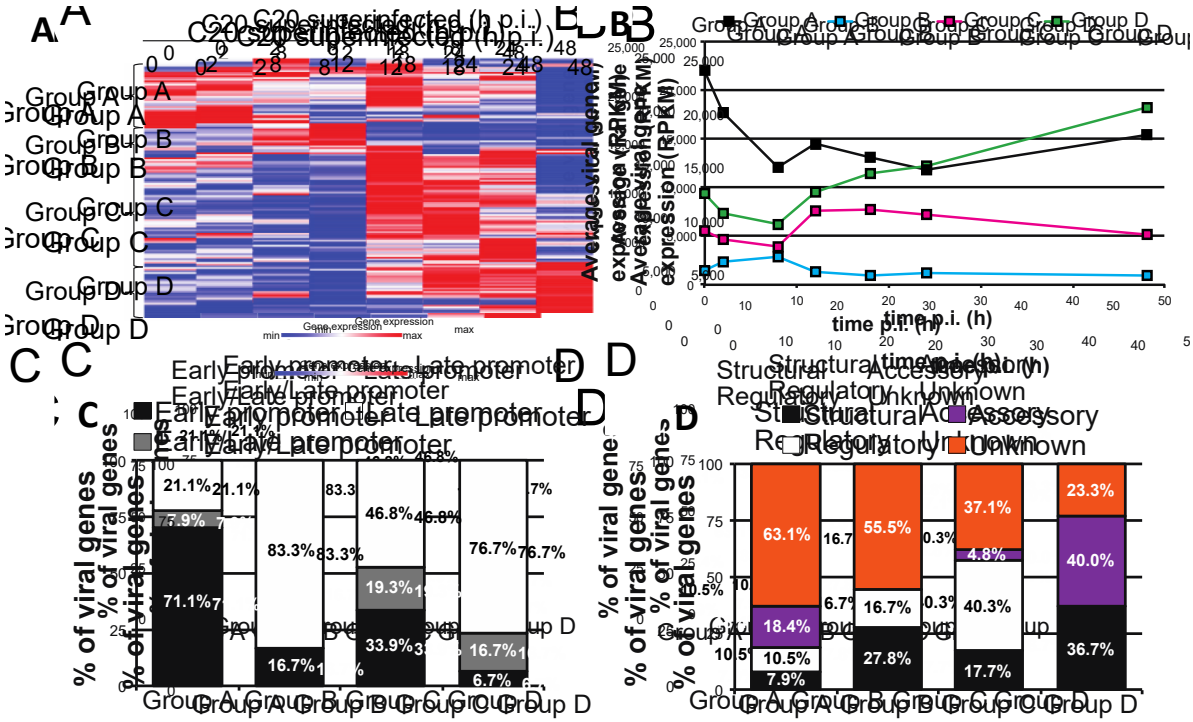


Figure 5.14: Cluster analysis of virus gene expression pattern in AcMNPV-superinfected C20 cells.
A) Schematics of normalised gene expression (RPKM) of superinfected C20 cells from 0 to 48 h p.i. Based on their gene expression pattern, viral gene expression was clustered by complete-linkage clustering into three groups: Group A (n=38), Group B (n=18), Group C (n=61) and Group D (n=30). B) Average gene expression pattern for each group as normalised gene expression levels. C-D) Percentage of viral genes from group A-D corresponding to C) function and D) promoter motifs. Heat map was visualised using Morpheus Software.

5.2.5 Host transcriptome analysis in mock and superinfected C20 cells

Transcription of genes in C20 cells, was compared to the transcriptome of uninfected Hi5 cells or when responding to an AcMNPV-lytic infection. Previously filtered reads from single-stranded RNA-seq were mapped to the Hi5 genome to calculate the RPKM values. It was first observed that in a lytic AcMNPV infection the number of reads mapped to the host genome decreased over time, which corresponded with an increase in reads mapped to the virus genome (Fig. 5.15A). In contrast, the number and the proportion of reads mapped to the host genome in C20 cells remained constant over

time (>95% cleaned reads; Fig. 16A and B). When these cells were superinfected with AcMNPV the proportion of reads mapping to the host genome slightly decreased to around 90% (Fig. 5.15C).

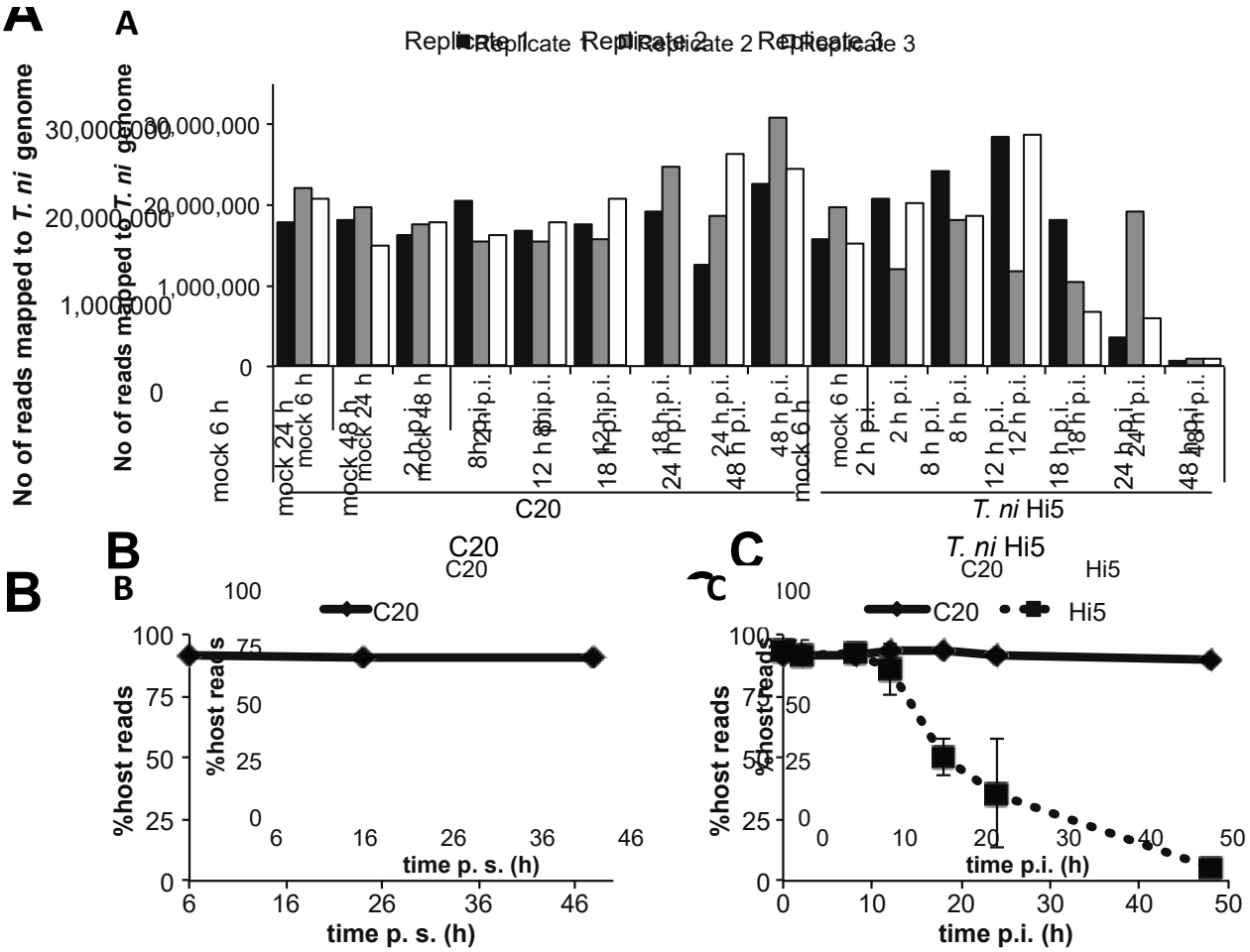


Figure 5.15 RNA-Sequencing statistics overview [adapted (Bannach, 2018)].
Total RNA of unchallenged C20 cells and AcMNPV-superinfected C20 or AcMNPV-infected Hi5 cells was extracted at the indicated time points p. s. or p. i. Strand-specific libraries were generated and deep sequencing was performed using Illumina® deep sequencing platform. Total number of Illumina® RNA sequencing reads for each replicate that was mapped to the A) *T. ni* genome. Average percentage of reads that were mapped to the host genome in B) unchallenged C20 or C) AcMNPV-superinfected C20 and AcMNPV-infected Hi5 cells.

5.2.5.1 Host transcriptome in persistent infection

To better investigate the response of the host C20 cells to the persistent virus, the expression of host genes at each time point post-seeding was compared to non-infected and AcMPNV-infected Hi5 cells as DGE (Table 5.2). In order to do this, gene expression was divided into up and down regulation categories, which in turn were classified into high (FC≥10 or ≤0.1), medium (FC<10-5 or >0.1-0.2) or low response (FC<5-2 or FC>0.2-0.5; Table 5.2). To identify significant changes in either up or down regulation for a gene, only RPKM≥5 were included (with $p<0.05$; (Chen *et al.*, 2014). A RPKM value of

approximately three, corresponds to one transcript in mammalian cells (Mortazavi *et al.*, 2008). Host genes with RPKM values lower than five were considered either as repressed (RPKM<5 in control=*T. ni* Hi5 mock but RPKM≥5 in C20 sample, with p<0.05) or as induced (RPKM<5 in control=*T. ni* Hi5 mock but RPKM≥5 in C20 sample, with p<0.05).

These analyses indicated that in unchallenged C20 cells only a small percentage of the Hi5 genes (n=14,372) were up or down regulated, 1.0% and 2.4% respectively, in C20 6h p.s. compared to non-infected Hi5 cells (Table 5.2). Regarding induced and repressed genes, 1.9% of the host genes were induced and 2.8% were repressed in unchallenged C20 cells (C20 6h p.s.; Table 5.2). This illustrates that the majority of host transcription is not regulated differently in response to a persistent infection. However, in the present study only 4.8% of transcription was down regulated at 48h p.i. in a productive infection (Table 5.2). Furthermore, no increase in down regulated or repressed genes was observed over the infection as shown in previous studies (Chen *et al.*, 2014). This might have been caused by the high standard deviations in Hi5 samples from 18-48h p.i. and therefore the host transcriptome response could not be considered significant.

Table 5.2: Pairwise DGE analysis of host genes compared to Hi5 mock cells *.

		Fold Change											
		Up regulated			Down regulated								
Sample	Induced	≥10	<10-5	<5-2	>0.2-0.5	>0.1-0.2	≤0.1	Repressed	Total regulated	%Up*	%Down*	%Regulated	%Not regulated
C20 6h	271	7	15	126	287	41	11	403	1161	2.92	5.16	8.08	91.92
C20 24h	384	4	15	160	908	47	11	564	2093	3.92	10.65	14.56	85.44
C20 48h	429	5	22	261	638	29	8	471	1863	9.69	7.97	17.66	82.34
Hi5 2h p.i.	0	0	0	3	2	1	0	14	20	0.02	0.12	0.14	99.86
Hi5 8h p.i.	259	0	1	103	234	3	0	254	600	2.53	3.42	5.94	94.06
Hi5 12h p.i.	45	1	1	15	69	7	0	65	203	0.43	0.98	1.41	98.59
Hi5 18h p.i.	176	1	1	158	341	12	0	181	870	2.34	3.72	6.05	93.95
Hi5 24h p.i.	85	0	3	72	254	4	0	74	492	1.11	2.31	3.42	96.58
Hi5 48h p.i.	232	11	37	31	631	18	1	42	1003	2.16	4.81	6.98	93.02

n=14,372 Hi5 genes. *Up and down regulated gene expression include induced and repressed genes, respectively

According to literature, a summary of the pathways involved or affected by virus infection is illustrated in Figure 5.16. Genes from specific pathways were identified by C. Bannach in collaboration with A. Shrestha and Z. Fei (Garry W. Blissard´s lab, BTI USA)

mainly based on Chen *et al*., 2013. Gene expression in C20 cells was compared to uninfected Hi5 cells (Supplementary CD Table S1).

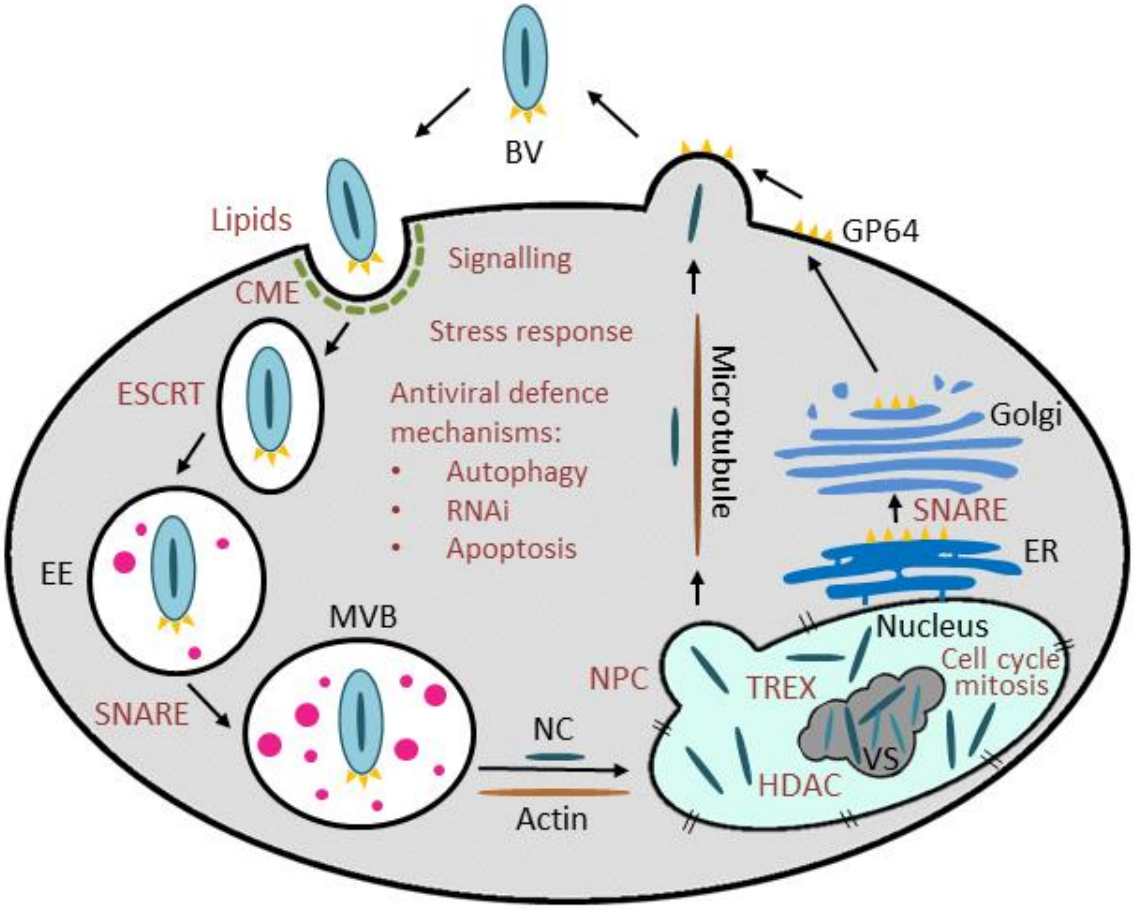


Figure 5.16: Summary of host pathways involved in AcMNPV infection.
AcMNPV enters the cell via clathrin-mediated endocytosis (CME) and traffics through the endosomal compartments using the endosomal sorting complex required for transport (ESCRT) machinery and the soluble N-ethylmaleimide-sensitive factor receptors (SNARE). Following early endosomal (EE) acidification this matures into a multivesicular body (MVB). Here, GP64 glycoprotein receptors enable the fusion of the nucleocapsid envelope to the endosome membrane. After their release, nucleocapsids traffic towards the nucleus using actin filaments. They then enter the nucleus through the nuclear pore complexes (NPCs), where virus uncoating takes place. Virus replication occurs in the virogenic stroma (VS) and may be dependent on histone deacetylases (HDAC) and the transcriptional export complex (TREX). The assembled nucleocapsids either bud out into the cytoplasm or exit the nucleus via the NPCs. The GP64 glycoproteins traffic to the plasma membrane via the endomembrane system through the endoplasmic reticulum (ER) and the Golgi apparatus, possibly in a SNARE protein dependent manner. Nucleocapsids then traffic towards the plasma membrane using microtubules and bud out of the cell, obtaining an envelope from the GP64-rich plasma membrane.

Regarding the pathways related with virus binding to the membrane and uptake, genes encoding for proteins required for the endocytic pit formation and some of the related key proteins were down regulated (Supplementary CD Table S5.1). In addition, *acyl-desaturase* involved in the fatty acid biosynthesis was one of the most highly induced

genes, while enzymes involved in the synthesis of sphingolipids were repressed in C20 cells.

Transcription of genes encoding for various transmembrane channels were between the most highly induced (*glutamate-gated chloride channel*, *trehalose transporter* and *neural acetylcholine receptor*) and expression of *caveolin-3*, gene required to form membrane invaginations [reviewed by (Cheng and Nichols, 2016)], was repressed. One of the most significant down regulated host genes in C20 was *ezrin radixin moesin* (*erm*) homolog 1, probably involved in membrane-associated signalling pathways and as a link between the actin cytoskeleton and transmembrane receptors (Batchelor *et al.*, 2004). An important part of the cell cytoskeleton, α-tubulin, was one of the most highly induced genes (Supplementary CD Table S5.1).

Transcription of *nup155*, a nuclear pore complex (NPC) protein (Radu, 1993), was reduced in C20 cells. The same happened with HDAC1 and N-ethylmaleimide sensitive fusion protein (NSF) gene expression. Inhibition of histone deacetylases (HDAC) leads to reduction of late gene expression (Peng *et al.*, 2007) while NSF is part of the soluble N-ethylmaleimide-sensitive factor receptors (SNARE) controlling fusion of cell membranes (reviewed by Han *et al.*, 2017) and involved in intracellular trafficking processes (Guo *et al.*, 2017a). Reduced expression in C20 cells when compared to Hi5 cells was also observed for *dead box protein 39A/39B* encoding for a putative RNA helicase and an essential splicing factor (Fleckner *et al.*, 1997). This is part of the transcriptional export complex (TREX) (Supplementary CD Table S3) that has a vital role in the gene expression pathway and prevents DNA damage (Heath *et al.*, 2016).

Concerning the cell cycle progression, transcription of the *structural maintenance of chromosome 3* (*smc3*), part of the chromatid cohesion complex, was reduced in C20 cells. Expression of key proteins for mitotic progression was also reduced including *aurora kinase*, *surviving and cell division cycle protein 20* (CDC20; Supplementary CD Table S3). Signalling pathways were differentially expressed as shown by genes involved in Mitogen-activated protein kinase (e.g. MAPK; *ras85D*, *MAPK4* and *p39 MAPK*), JAK/STAT [e.g. *ubiquitin-conjugated enzyme 13* (*UBC13*), *gram-negative binding protein* (*GNBP*) and SKPA)] and PI3K signalling cascades (*e.g. pi3k_68D*) (Supplementary CD Table S3).

Stress response was naturally affected in the persistent infected cells. Expression of genes from the heat shock protein family (hsp) was reduced in C20 cells, including *hsp21.4*, *hsp60*, *hsp70* and small heat shock protein (*sHsp*). The *hsp20.8* was the exception since its transcription was induced in C20 cells. Surprisingly, *glutathione s-transferase* (*gst*), a detoxification enzyme but also a cell signalling modulator in response to oxidative stress [reviewed by (Sheehan *et al.*, 2001), was one of the highest upregulated genes in C20 cells. Expression of *catalase* was reduced in C20 challenged cells, while superoxide dismutase was up regulated compared to Hi5 cells (Supplementary CD Table S3).

Antiviral response includes important defence mechanisms against virus infection such as autophagy, RNA interference (RNAi) pathways and induction of apoptosis. In respect to autophagy, a much conserved lysosomal degradative pathway, genes involved in vesicle induction (*atg101, atg17*), vesicle elongation (*atg4B, atg12*) and signalling (*vps20, ras*) were repressed in C20 cells. Furthermore, a decreased expression of *atg8*, *atg3* and *tsc* was observed when compared to Hi5 cells. The key components of insect cell small interfering RNA (siRNA) responses to long dsRNA, *dicer-2* and *ago-2*, (Chen *et al.*, 2014) had their expression down regulated in C20 cells. While dicer-2 processes dsRNA into short RNAs, ago-2 interacts with siRNA to form the RNA induced silencing complex (RISC) (Sashital and Doudna, 2010; Chen *et al.*, 2014). The *death-associated inhibitor of apoptosis* (*pdcd5*) and *caspase-1* were also down regulated in C20 cells. While PDCD5 (*pdcd5* encoded) activates other apoptotic related transcription factors (Li *et al.*, 2016), *caspase-1* activates the programmed cell death by lysis.

5.2.5.2 Host transcriptome in superinfected C20 cells

The previous chapters established that C20 cells are partially resistant to superinfection. Viral-cell invasion and the establishment of an infection are dependent on specific cellular pathways. Hence, these were carefully analysed to understand if the presence of a persistent infection could be compromising essential cellular pathways for virus infection. Gene expression in superinfected C20 cells was compared to unchallenged and AcMNPV-infected Hi5 cells. Samples from unchallenged C20 cells 6h p.i. were considered the time 0 for the superinfection analysis. Transcription profiles for each pathway are

plotted in the supplementary Figures S 5.1-5.18 while data analysis is shown in Supplementary CD Table S5.1.

In the AcMNPV-productive infection, genes encoding for proteins involved in endocytosis were highly expressed between 2-8h p.i. Even though these genes were expressed in superinfected C20 cells, their expression was delayed and peaked generally after 12h p.i. (Fig. S5.1). Furthermore, as observed for the non-challenged C20 cells, genes required for endocytic pits formation were repressed.

The regulation of lipid synthesis seemed to be affected. Expression of *scd*, encoding Stearoyl-CoA desaturase (SCD), a pivotal enzyme for monosaturated fatty acids synthesis, increased modestly until 18-24h p.i. in AcMNPV-infected Hi5, following by a decline. In contrast, *scd* transcription in C20 superinfected cells decreased until 8h p.i. and subsequently increased back to its basal level (Fig. S5.2). Furthermore, expression of genes encoding for proteins required for glycerine synthesis decreased in virus infected Hi5 over time, while it remained constant in challenged C20 cells, such as: diacylglycerol acyltransferase (DGAT), glycerol-3-phosphate (GPAT) and acyl-CoA synthetase (ACS) (Fig. S5.2).

Components of the endosomal sorting complex required for transport (ESCRT; reviewed by Hurley, 2015) play a role in efficient entry and egress of BV (Li and Blissard, 2012; Yue et al., 2018). Hence, in AcMNPV-infected Hi5 cells gene expression peaked between 2-12h p.i. However in superinfected C20 cells transcription of some of these genes did not reach a maximum as expression levels remained low during the time course (Fig. S5.3).

SNARE proteins Syx7, NSF and SNAP29 have shown a different pattern of gene expression in C20 superinfected cells compared to a productive infection (Fig. S5.4). Different expression patterns were also observed for genes encoding components of NPC, essential for AcMNPV transport into the nucleus (Au and Panté, 2012; Au *et al.*, 2013), such as *nup155* compared to infected Hi5 cells. Nup155 interacts with GLE1 involved in the mRNA transport through the NCP (Rayala *et al.*, 2004) and its expression showed a similar trend (Fig. S5.5). The increase in AcMNPV-infected Hi5 histone deacetylase HDCA1 expression was not observed in superinfected C20 (Fig. S5.6). Transcription of the *suppressor of transcriptional defects of hpr1 delta by overexpression complex* (*thoc2*), the largest subunit of transcription export pathway complex required

for mRNA biogenesis (Peña *et al.*, 2012), was activated in AcMNPV/infected Hi5 cells peaking at 2h p.i. In C20 cells, gene expression for *thoc2* was constant over time (Fig. S5.7).

Regarding the cell cycle progression and mitosis regulation, retinoblastoma protein (Rb) is an important part of the G1/S phase checkpoint (Weinberg, 1995). In infected Hi5 cells, *rb* transcription increased, reaching a maximum at 8h p.i. while it stayed constant during C20 superinfection (Fig S5.8). Genes involved in signal pathways seem to have a differential expression as evaluated by the transcription of *ras85D*, *p39 MAPK*, *DOMELESS*, *HEMIPTEROUS*, *SKPA, pi3k_68D* and *nuclear kappa-light-chain enhancer of activated B-cells* (*NF8kB19*) (Fig. S5.10-5.13). A substantially altered expression was observed for genes involved in stress response. Examples are the *hsp70* and *sHsp* for which gene expression remained constant in C20 AcMNPV-superinfected cells while it increased or declined respectively in a productive infection (Fig. S5.14). Furthermore, the maximum gene expression for the *oxidative stress response gene* (OXSR) was reached earlier in the superinfected cell line in comparison to the lytic infection (12h p.i. and 24h p.i. respectively; Fig. S15). In regard to the antiviral response, expression of RNAi genes was in general down regulated in the AcMNPV-Hi5 infection while it remained constant in C20 cells (Fig. S5.17). Transcription of *pdcd5* was strongly activated upon AcMNPV infection of Hi5 cells, while it only slightly increased in C20 AcMNPV-superinfected cells (Fig S5.18). Cytochrome C involved in programmed cell death [reviewed by (Jiang and Wang, 2004)] was increased upon infection of Hi5 cells, having a peak of expression at 12h p.i. However, its expression did not change in superinfected C20 cells (Fig S5.18).

5.3 Discussion

The present chapter examined the genome of a long term persistent virus infection in Hi5 cells as well as the virus and host transcriptome by RNA-sequencing. However, a gap of 10 months and over 35 passages occurred in between the genomic and transcriptomic analysis. Therefore, caution must be exercised in correlating the two datasets. In addition, the superinfection phenomenon described in Chapter 4 was reconsidered in light of these molecular techniques.

Early characterization of WGS data revealed three major deletions in the AcC20 genome, in independent regions: *egt* and *orf16* region; *p95* gene; *lacZ*, *p10* and *p74* region. Upon confirmation by PCR, large insertions were also detected in the first two regions. A close observation of the *de novo* assembly obtained later and AcUW1.*lacZ* alignments revealed a further insertion of 287bp in the *p94* coding region (ODV-E25). Due to the initial approach undertaken, using AcUW1.*lacZ* genome as a scaffold for assembly, gaps were created in these regions (Fig. 5.1). Therefore, the insert sequences would not have been mapped to the virus DNA, so were excluded. Furthermore, when reference based *de novo assembly* was firstly attempted the *egt*/*orf16* region still showed a gap since the size of the *egt* insertion (1350nt) hindered assembly without a scaffold. The assembly on this region was completed using primer walking. The insert comprised partial *T. ni* transposable sequences and generated a duplication of the AcMNPV *egt* TTAA as previously described in *egt* but also for TFP3 and IFP2 lepidopteran transposon insertions in the AcMNPV *fp25k* (Cary *et al.*, 1989; Wang *et al.*, 1989; O'Reilly *et al.*, 1990; Wang and Fraser, 1993). These transposable elements usually have no significant protein-coding potential and incorporated imperfect sequences repeats (Cary *et al.*, 1989; Wang *et al.*, 1989) as was the case in the *egt* described insertion. The insertion of transposable *T. ni* sequences into the AcMNPV genome suggests viral infections have a role as a vector for horizontal DNA transfer between eukaryote cells therefore contributing to species evolution (Gilbert *et al.*, 2014; Drezen *et al.*, 2017). Moreover, it was suggested that baculovirus DA26 protein could be involved in determining tissue tropism and virulence in host insects (Katsuma *et al.*, 2012).

In AcMNPV, EGT has a role for infection *in vivo* (O'Reilly and Miller, 1989, 1990) and ODV-E25 (product of *p94*) although it associates with both BV and ODV is not essential *in vitro* (Braunagel *et al.*, 2003; Chen *et al.*, 2012). However, AC83 and DA26, encoded by *p95* and *orf16* respectively, are essential for viral replication. The DA26 binds to and regulates the levels of IE0 and IE1 (Nie *et al.*, 2009) while AC83, besides its role as a PIF, is essential for BV and ODV nucleocapsid assembly (Javed *et al.*, 2017). For *p95* the deletion and insertion occurred after the cells were cloned to derive C20, sometime in between passage 1 and 40. For *egt* and *ac16* however, the deletion must have happened as early as passage 16 of the initial Hi5 AcUW1.*lacZ*-infected culture, or even earlier since this passage already presented both the WT genotype and the deleted one. The

observed insertion occurred after the cloning of the cell line but seems to be gradually lost over time. Finally, for *p94* the insertion happened at or before passage 16 of the initial AcUW1.*lacZ*-infected Hi5 culture and it was retained at the last passage tested (C20 p135). These mutations generated a truncated transcript for EGT, ODV-E25 and P95 proteins which was confirmed when the transcriptome data was assembled. Although the RNA alignment was not conclusive for DA26, its transcription seems to initiate within the *egt* coding region due to the deletion of the 3' *orf16* coding region and 5' *egt* coding region. The inconclusive results for this region might be caused by the presence of a mixed population and would have been overcome by performing single cell RNA-seq. Given the AC83 and DA26 essential roles in virus replication, it was hypothesised that these modified proteins are still functional although further proteomic studies would be needed to support this conclusion.

The AcC20 de novo assembly revealed variants to the original AcUW1.*lacZ* affecting thirty nine *orf*'s (see Table 5.1). The majority are either non-essential for the AcMNPV replication cycle or are only essential in *vivo* (e.g. *orf51*, *orf96*, *p79*, *ac138*). Nonetheless, eight genes encode for the essential proteins PK-1, DA26, ARIF-1, AC53, Lef-9, helicase, GP64 and AC83 while the last region covers *hr2*, probably a transcriptional enhancer and an origin of DNA replication (Guarino *et al.*, 1986; Pearson *et al.*, 1992). A basic alignment against the protein database revealed the variants were not affected in conserved regions for most of the proteins. Exploring the impact of these proteins *in vitro* would be achieved by performing Western blots. This will clarify the consequence of these truncations and amino-acid changes for the structure and function of the protein. However, minor variants detected could just reflect errors in the original AcMNPV sequence, from which AcUW1.*lacZ* was based at, since the sequencing was performed before the existence of automated sequencing methods (Ayres *et al.*, 1994).

As described and shown in Chapter 3 (Fig. 3.1), C20 cells rarely produced polyhedra but when these were observed in the nucleus the number present was lower than what was expected from AcUW1.*lacZ* (Fig 5.6). This phenotype was described in 1973 and it often occurs upon passaging the same virus several times (Hink and Vail, 1973; Potter *et al.*, 1976), particularly in cell lines derived from *T. ni*. It was later attributed to the truncation of the FP25 protein, due to the insertion of an extra A in a sequence of seven (Beames and Summers, 1988, 1989; Kelly *et al.*, 2006). However, the same or similar FP

phenotypes can occur due to the insertion of transposable elements from host *Sf* or *T. ni* genomes (Fraser *et al.*, 1983; Beames and Summers, 1988; Cary *et al.*, 1989; Wang *et al.*, 1989) as discussed above in the *egt* genotype. The FP25 was suggested to associate with several proteins, ODV-E26, ODV-E66 and GP67, ODV-E25 and P39, and together play an important role in the targeting and intracellular transport of viral proteins during infection (Braunagel *et al.*, 1999). The insertion of an extra A was confirmed in the AcC20 virus and it was shown to be present from passage 16 or earlier of the original pre-cloned culture of Hi5 AcUW1.*lacZ*-infected cells. Interestingly, this mutation usually leads to an increase in BV production, therefore a higher titre which leads to higher cell mortality (Potter *et al.*, 1976; Kelly *et al.*, 2006). However, in the present study a long term persistent infection is described presenting the FP phenotype while the cells produce a low virus titre (see Chapter 3). One can speculate that the C20 cell is a result of several dynamic subpopulations. Therefore at a defined time point, an infected subpopulation might be producing high quantities of BV but still be overall diluted and not enough to infect all the population and kill the cells. On the other hand, we might have a cell population harbouring an incomplete genome as described before (Weng *et al.*, 2009), which transcription is enough to drive resistance to superinfection.

From the plaque assays of AcC20 virus performed in Sf21 cells a particular phenotype was observed. When using a bright field microscope, while some plaques were clear, some others were surrounded by apoptotic vesicles. Infection of Sf21 cells with AcMNPV causes a transient state of plasma-membrane blebbing, however the observation of this in a large scale is unusual. This phenomenon was previously described and occurs when the apoptotic response is not blocked resulting into plasma membrane blebbing (Clem *et al.*, 1991; Kelly *et al.*, 2006). It was attributed to a deletion in the 5' coding region of the anti-apoptotic *p35.* (Kelly *et al.*, 2006). Similar to what happens in apoptosis, this results in the disintegration of cells into apoptotic bodies (Kelly *et al.*, 2006). A previous study reported the establishment of a persistent baculovirus infection in *Sf* cells with a mutated or deleted *p35* (Lee *et al.*, 1998). However, this persistent infection was blocked by stable transfection of the same gene in the host genome (Lee *et al.*, 1998). In the current study, plasma-membrane blebbing occurred when AcC20 was used to infect *Sf* cells, however, in this case a possible deletion in the 3' end of the *p35* gene was found. Nonetheless, RNA sequencing showed anti-apoptotic *p35* gene was expressed in

C20. It would be important to test though if this corresponds to a fully competent protein.

Subsequently, RNA sequencing was performed in C20 cells to explore the dynamics between the host and the persistent infecting virus transcriptomes. Using the same tools, gene expression in superinfection was also analysed and compared to a lytic infection. An initial analysis of the read number and distribution between host and virus gene expression showed only ~2% of the C20 transcriptome corresponds to viral genes in contrast to a productive infection where it can reach up to 95%. Although transcription occurred at a low level, all viral genes were expressed in C20 cells. Three time points 24h apart from each other were tested, which showed that only a few genes were significantly up or down regulated in between the time points. This corroborates the dynamic behaviour described when analysing AcC20 replication by several other techniques (Chapter 3).

The AcUW1.*lacZ* virus initially used to establish the persistent infection had *p10* deleted, therefore the presence of *p10* transcripts during the time course of the superinfection in C20 illustrated this cell line can be superinfected. The proportion of total reads mapped to the viral genome was itself higher than in unchallenged C20 cells and increased during the time course, confirming the previous conclusion. Moreover, the exponential increase of *p10* expression up to 48h p.i., corroborated the accumulation of P10 observed in western blots, for superinfected C20 cells but not in the same unchallenged cells (see 4.2.2.2 and Fig. 4.3).

Even though all viral genes were expressed in C20 cells, interestingly the genes expressed most highly at the three time points all had late promoters, belonging to the late or vey late phase of infection: *p6.9*, *polh*, *orf76*, *orf58*, *ac-bro*, *ctx*, *v-ubi*, and *odv-ec27*. Considering the infection is already established, it could be hypothesised that different cells have different cycles of virus infection at all times and therefore this result will just corroborate that the late genes when expressed have a strong transcription. Therefore, the most abundant genes at any time would be from these phases. For both unchallenged and superinfected C20 cells the highest expressed genes at any time point were *p6.9* and *orf76*. The P6.9 is responsible for the dense packing of the DNA and inactivation of late genes (Wilson *et al.*, 1987; Dong *et al.*, 2010), while AC76 is involved

in intranuclear microvesicle formation and in its absence the OB formed lack ODV (Hu *et al.*, 2010). In virus-infected Hi5 cells these genes were most highly expressed at 24h and 18h p.i. respectively. It is unknown why these specific transcripts are so high, accounting together for 42-45% and 25-43% of the total viral mRNA in unchallenged or superinfected C20 respectively. The remaining highly expressed genes in unchallenged C20 cells are all late genes except *polh*. Although *polh* followed the same upwards tendency observed in a lytic infection, its expression only accounted for 5% of total mRNA at 48h p.i. Genes *orf58* and *ctx* are not essential for viral replication although *orf58* encoded protein associates with AcMNPV BV and ODV (Braunagel *et al.*, 2003; Wang *et al.*, 2010b). The ODV-EC27 seems to be involved in cell cycle regulation during virus infection while it associates with both forms of the virus (Braunagel *et al.*, 1996; Belyavskyi *et al.*, 1998). Finally *v-ubi* encoded protein localizes at the inner envelope surface and might have a role in protecting viral proteins from degradation by the host (Guarino *et al.*, 1995; Haas *et al.*, 1996) while *c-bro* is a possible DNA binding protein (Zemskov *et al.*, 2000).

In superinfected C20, four new genes reached the top six most expressed at different time points [*orf75* (2h p.i.), *orf74* (2,12,18h p.i.), *p94* (12,18,24h p.i.) and *orf72* (24h p.i.)]. Of these, *orf74* and *p94* encode for proteins that associate with BV and ODV (Braunagel *et al.*, 2003; Chen *et al.*, 2012), *orf75* is required for nuclear egress of nucleocapsids and intracellular vesicle formation (Guo *et al.*, 2017b; Shi *et al.*, 2017) while *orf72* function is unknown. Although more diversity was observed after superinfection, transcription did not resemble the pattern of a productive infection and the typical temporal expression in AcMNPV infection was not present, hence, expression of early genes never overcome late gene transcripts (Fig. 5.12). From an increasing order of gene expression the phases were ordered as immediate early, very late, early and late. Differential gene expression suggested delayed expression for immediate early genes in superinfected C20 cells (12h p.i.) compared to AcMNPV infected Hi5 cells (2h p.i.) which might be one of the causes for slow rate of superinfection described in chapter 4.

In the present study, clustering gene expression in infection phases enabled experimental validation of the empirical model of baculovirus gene expression that describes the phases of infection (Fig. 1.3 and 5.12B). The maximum gene expression

was achieved at 2h p.i., 8-12h p.i. and 18-24h p.i. for immediate early, early and late genes respectively. An important discrepancy from the model was detected for the very late phase in which the expression maximum is probably reached after 48h p.i. and not before as predicted (Kelly and Lescott, 1981; Friesen and Miller, 1985). However, these predictions were made with *Sf* cells and used different cell culture media which might affect the transcription pattern. Also, up to 48h p.i. very late expression was not higher than late gene transcription. Taking the experiment up to 72h p.i. would be needed to clarify the very late expression pattern and its expression peak.

Finally RNA sequencing enabled the analysis of host gene expression in a persistent infection and in superinfection for the first time. Contrary to a productive infection, where the proportion of reads mapped to the host genome decreased as the infection progresses, in a persistent infection this remained constant. In this covert infection, host genes are transcribed at high levels (>95% of the total reads) and the majority of host transcription is not regulated differently besides the presence of a persistent infection. In superinfection this percentage decreases to 90% suggesting a subpopulation of C20 cells could be superinfected as previously suggested by AcMNPV transcriptome analysis and in chapter 4. Supporting previous studies that indicated host transcription is shut down in baculovirus lytic infection, a rapid decreased in host transcription of AcMNPV-infected Hi5 cells was observed (Braunagel *et al.*, 1998; Nobiron *et al.*, 2003).

In order to explore which main cellular pathways might be expressed differently and therefore allow the maintenance of the persistent infection these were analysed separately. Furthermore, the same pathways were explored in superinfection of the C20 persistent infection to understand if the partial blockage to superinfection observed in chapter 4, might be linked to repression of particular genes or pathways.

Host gene expression was compared to the parental Hi5 cell line infected or non-infected. Firstly explored and with the more informative results were pathways related to virus binding to the membrane and uptake. Genes required for the endocytic pit formation were repressed and related genes were down regulated, indicating a deficiency of endocytosis in the C20 cell line. On the other hand, surprisingly, transcripts encoding several transmembrane channels were the most highly induced.

Virus uptake is dependent on the membrane lipid composition (Chernomordik *et al.*, 1995; Tani *et al.*, 2001). Therefore the observed differential regulation of lipid synthesis in C20 cells can possibly affect membrane composition and therefore the ability of BV to bind, as it was suggested before (see 4.2.4). The reduced expression of *gle1* and *nup155* indicated some mechanisms for mRNA transport and AcMNPV nuclear import through the NPC might be affected (Radu, 1993; Rayala *et al.*, 2004). The observed decreased expression of *rb* and other mitotic proteins can be leading to slow cycle and mitosis progression and the cause of the slow growth of C20 cells (see chapter 3).

It was speculated that to enable the maintenance of the persistent infection in C20 cells, stress response and defense mechanisms could be down regulated. Baculovirus infection leads to oxidative stress in Sf9 and Hi5 AcMNPV-infected cells (Wang *et al.*, 2004; Rohrmann, 2013). The *hsp* genes are an important part of the cell stress response to aggressions such as virus infection (Lyupina *et al.*, 2010). Furthermore, manganese superoxide dismutase (MnSOD) expression was shown to be able to overcome the oxidative stress caused by the infection in Hi5 cells but not Sf9 (Wang *et al.*, 2004). Enzymes from the GST family encode genes important for detoxication and toxification mechanisms but are also involved in cell signaling modulation as a response to oxidative stress [reviewed by (Sheehan *et al.*, 2001; Nebert and Vasiliou, 2004)]. A productive AcMNPV infection in Sf9 cells has been demonstrated to increase expression of proteins from the hsp family as hsp/hsc70 (Lyupina *et al.*, 2010). In our persistent infection, expression from *hsp* family was in general reduced while *gst* and superoxide dismutase had a highly regulated expression. As a consequence of the persistent infection, a higher oxidative stress was hypothesised to be present in the cell and therefore the response mechanisms could be up regulated. However, this was not verified in the C20 cells as the response mechanisms to oxidative stress did not seem to be in general up regulated. A possible reason might be that the maintenance of the infection at a persistent level does not substantially increase the oxidative stress thus negating the need for a response.

Autophagy degradative pathway represents an important viral defense mechanism as it can target viral particles for degradation, promote its interaction with specialized receptors for innate immunity activation or facilitate anti-viral adaptive immunity activation (Viret *et al.*, 2018). A previous study suggested AcMNPV infection did not

inhibit autophagy but instead triggered apoptosis under starvation pressure (Wei *et al.*, 2012). However, genes involved in autophagy were repressed in C20 cells. Key components from cellular siRNA response to invading dsRNA (*dicer-2* and *ago-2*) contribute to an efficient antiviral response towards AcMNPV. In turn, viral *p35* was shown to suppress the siRNA pathway (Mehrabadi *et al.*, 2015; Karamipour *et al.*, 2018), besides its main role as an anti-apoptotic protein by inhibiting *caspase-1* (Prikhod'ko and Miller, 1999). As mentioned before, AcMNPV *p35* expression was not differentially regulated in C20 cells, however, cell activators of apoptotic transcription factors were down regulated including *caspase-1* and *pdcd5*, probably a result of the interaction with *p35* and other antiapoptotic factors (Prikhod'ko and Miller, 1999). The balance between anti-apoptotic viral genes and apoptosis inducers as part of the cell defense mechanisms, together with a repression of autophagy, might be responsible for the maintenance of the C20 persistent infection over the years.

The same tools enabled the analysis of the gene expression in response to superinfection of persistently infected cells. Previous results in this thesis showed that although a partial blockage to superinfection occurs, a small proportion of the cells can be secondarily infected. Microscopy and adsorption studies described in chapter 4 suggested the blockage could be in the very early steps of superinfection, involving the binding or absorption steps. However, a previous study suggested the exclusion mechanism might arise due to actin rearrangements during virus infection as the inhibition of actin polymerisation by cytochalasin D enabled infections with a second virus (Beperet *et al.*, 2014). As discussed before, genes encoding for the endocytic pit formation were repressed in non-challenged C20 cells. This suggests a lower capacity of C20 cells to take up new BV. The same result was observed when superinfecting these cells. Furthermore, genes encoding for proteins involved in endocytosis had a late expression compared to a productive infection. In a lytic infection, transcription peaked around 2-8h p.i. while it decreased thereafter, possibly enabling the formation of the endocytic vesicles at early times points. However, this later decrease in gene expression might account for the resistance observed to infection with a second virus 12-15h p.i. of the first one (Beperet *et al.*, 2014). The late expression of endocytic proteins in superinfection might be one of the reasons for the slow rate at which superinfection develops (see 4.2.4). Furthermore, delayed expression of this proteins might be

accounting for the apparent high resistance to superinfection since when cells were challenged the same absorption time as for a lytic infection was allowed (i.e. virus inoculum was left in contact with the cells for 45min only). Increased times could have increased the rate of superinfection. Other pathways showing differential expression for some of their genes included lipid synthesis, endosomal trafficking system through the ESCRT complex, SNARE proteins responsible for the fusion of transport vesicles with their target membranes, NPC complex transport and signalling. This suggests the regulation of multiple trafficking steps upon infection of C20 is altered when comparing to its parental cell line. Finally, regarding the defence mechanisms, transcription of *sHsp* and *RNAi* genes in general was down regulated in response to an AcMNPV-infection while it remained constant in C20-superinfected cells. For *pdcd5*, apoptosis activator, which was down regulated in unchallenged C20 cells but strongly activated in AcMNPV-infected Hi5 cells, expression only slightly increased upon superinfection. The same was observed for other proteins involved in programmed cell dead. In general, the defence mechanisms seem not to be activated or not as strongly as in a lytic infection which was expected considering they were found to be mainly down regulated in the unchallenged C20 cells.

5.4 Conclusion

Whole genome sequencing revealed nine essential *orf*s were affected in AcC20 compared to the parental AcUW1.*lacZ* virus. These included two major truncated proteins (AC83 and DA26) and the *hr2* region. Also, a defective *fp25k*, responsible for the low polyhedra production, was detected in both early and late passages of the virus. Although transcription occurred at a low level, all viral genes were expressed in C20 cells. The RNA sequencing studies also showed a low level of superinfection is possible and the infection probably progresses slower, although gene expression is considerably different form a productive infection. The majority of viral transcripts present in either non-challenged or superinfected C20 were from the late phase of infection. Transcription studies suggested endocytic pit formation and related key proteins as well as defence mechanisms including apoptosis and autophagy seem to be compromised. While the first set of genes might be the cause for large resistance to superinfection, the second can help to understand the maintenance of a persistent viral infection.

Further proteomic studies would be needed to understand the impact of the mutations in essential *orf*'s to the translated protein and associated proteins. It would also be important to clarify if the mutations in *hr2* is impacting viral replication. Repeating transcriptomic analysis at a different passage will be needed to corroborate the preliminary results generated in this chapter. Enlarging the sampling time points will enable a better study of the very late phase of expression in a productive infection. Furthermore, it will be important to understand how superinfection develops at later time points, since the results showed it might progress slower in general. Finally, although gene expression analysis of both virus and host hinted at several reasons for the maintenance of the persistent infection as well as the resistance to superinfection, including repression of genes for endocytic pit formation, differential lipid synthesis regulation and differential regulation of activators of apoptosis transcription factors, these would need to be analysed by other molecular methods to confirm the results presented here.

Chapter 6

Establishing cell lines persistently infected with recombinant baculoviruses

6.1 Introduction

The previous chapters focused on the C20 cell line, established after Hi5 cells were infected with AcUW1.*lac*Z. Shortly after this, several other similar cultures were created with baculoviruses expressing genes such as *urokinase*, influenza virus neuraminidase or fluorescent proteins (R.D. Possee, personal communication). Unfortunately, these cultures were discarded before detailed studies were conducted.

During the course of the present project several unsuccessful attempts were made to re-establish a persistent infection using AcUW1.*lacZ* virus and Hi5 cells. Although the cells seemed to survive for three to six weeks, and clusters of apparently healthy cells were observed, they all eventually succumbed to overt infection. Lack of success could be due to the passage number of the virus inoculum or to the different media currently used in our laboratory. However, even though different media were tested and a stored sample of the same passage number of AcUW1.*lacZ* was used to infect Hi5 cells, no persistent infection was generated.

Chapter 5 revealed a mutation in *fp25k* appeared at passage 16 or earlier in the Hi5/AcUW1.*lacZ* culture. This FP mutation is responsible for a decrease in polyhedra numbers but an increase of BV production (Kelly *et al.*, 2006) and appears often after virus passaging (Potter *et al.*, 1976). Therefore, the early appearance of the FP profile and consequent increase in BV may be a possible reason for the unexpected decrease in viability of the new cultures. It was hypothesised that using a virus with a low BV production may overcome this problem.

A recent study from our group suggested that a mutated virus having impaired production of BV is able to infect Hi5 cells in a persistent form. The virus used had two non-synonymous SNP in *lef2* affecting amino-acid 84 that changed from a cysteine to an alanine (Bannach, 2018). Therefore, in the present chapter, persistent infections were created by infecting Hi5 cells with virus containing this mutation but also harbouring recombinant genes under strong promoters.

Currently, although BEVS have been extensively used for the production of recombinant proteins and VLPs with important applications such as the production of vaccines (reviewed by Liu *et al.*, 2013; Kost and Kemp, 2016), the system is still based on a batch

production process. Recombinant proteins are generated by infecting insect cells with a high MOI of a recombinant virus expressing genes of interest under a late or very late AcMNPV promoter, usually *p10*, *polh*, or *p6.9*. The protein of interest is then purified from the total culture during the very late phase of infection. Large quantities of protein can be generated by creating batches of infected cells in a discontinuous process. The present chapter focuses on investigating the use of a persistent recombinant baculovirus infection for the continuous production of proteins, thus offering an alternative to the present BEVS. The use of a persistent culture would avoid the recurrent steps of scaling up the cell culture, infecting and waiting for the very late phase to be achieved before extracting the protein samples. This would enable the process to be expedited by extracting recombinant products in a continuous manner, analogous to stable mammalian cell lines expression systems (Zhu, 2012; Khan, 2013).

6.2 Results

6.2.1 Recombinant viruses with a mutated *lef2* to express genes of interest

Plasmid pBS.SK+CB1$^{LEF\text{-}2,C84A}$ was digested with *Eco*RV and *Bsp*EI and the fragment containing the mutated *lef2* was used to replace the homologous region in pOET1 and pOET3, producing pOET1-lef2C84A and pOET3-lef2C84A. The monomeric version of DsRed was inserted into these plasmids between *Hin*dIII and *Not*I after isolation from pDsRed-Monomer-N1. Urokinase was obtained from pAcUK13 (King *et al.*, 1991) by PCR with primers Uk-Fw-*HindIII* and Uk-Rv-*EagI* (table 2.4), digested with the appropriate enzymes and cloned into pOET1-lef2C84A and pOET3-lef2C84A. This process generated the plasmid transfer vectors pP6.9-lef2C84A-DsRed, pPOLH-lef2C84A-DsRed, p6.9-lef2C84A-UK and pPOLH-lef2C84A-UK.

The DNA from AcMNPV∆lef2 (Chambers, 2012) was linearized with *Sfr*I and co-transfected with the transfer vectors described above (see 2.2.3). Before the cell media was harvested for plaque purification, fluorescence was confirmed in the co-transfections with DsRed constructs using microscopy. Recombinant viruses were isolated from the cell culture medium using three rounds of plaque purification to derive: AcP6.9-lef2C84A-DsRed, AcPOLH-lef2C84A-DsRed, AcP6.9-lef2C84A-UK and

AcPOLH-lef2C84A-UK (see 2.2.4). During this stage, the slow replication of the virus was noticed since the plaques took longer than usually to develop (data not shown). The final plaque picks were used to amplify small virus stocks from which DNA was extracted and tested for the presence of the inserted genes by PCR with primers RDP887 and RDP214 (data not shown). Sanger sequencing was performed to confirm the mutation in *lef2*. These analyses showed that all AcPOLH-lef2C84A-DsRed isolates lacked the DsRed coding region, therefore this virus was omitted from further study. The other three viruses were amplified in Sf9 cells to working stocks (see 2.1.1).

6.2.2 Establishment of persistent infections harbouring *lef2* mutated virus

To establish persistent virus infections, $0.3x10^6$ Hi5 cells per well were seeded in 12-well plates and inoculated with 5, 15 or 45µl of a p0 stock of AcP6.9-lef2C84A-DsRed, AcP6.9-lef2C84A-UK or AcPOLH-lef2C84A-UK. Simultaneously, AcUW1.*lacZ* was used as a control while three wells were mock-infected. Over a period of seven to twelve days, most cells inoculated with virus displayed cytopathic effects consistent with virus infection resulting in their demise.

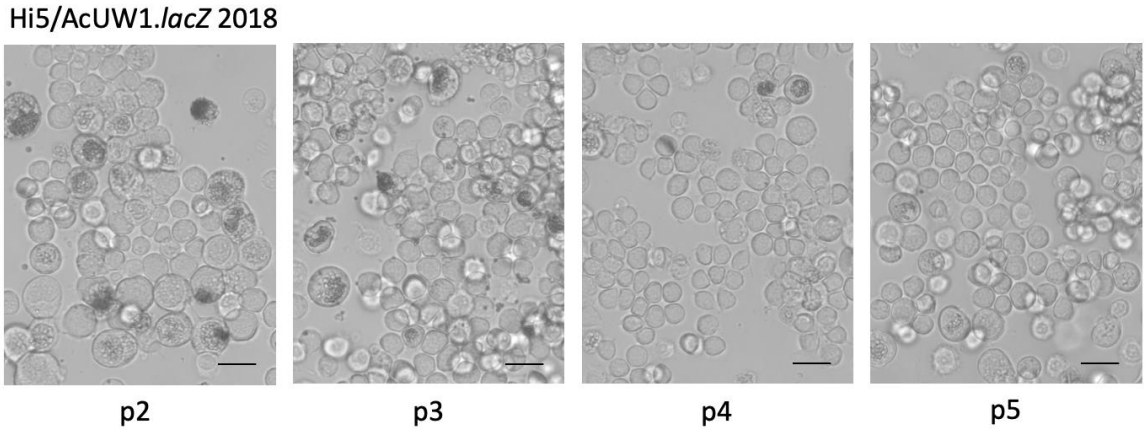


Figure 6.1: Bright field microscopy in sequential passages of a persistent infection
Persistent infected culture of Hi5 cells harbouring AcUW1.*lacZ* generated in 2018 at passage 2-5. Scale bar 20µm. Images are representative.

However, small clusters or colonies of healthy cells with shiny aspect, apparently uninfected, developed in all cultures from ten days, including those infected with AcUW1.*lacZ*. At passage two to five of the persistent infection harbouring AcuW1.*lacZ* established in 2018 (Hi5/AcUW1.*lacZ* 2018), some cells still showed evidence of

occlusion body formation even though a decrease was noted from passage 3 to 4 (Fig. 6.1).

For the four persistent virus infections created with Hi5 cells, once the cells had reached confluency (2-3 weeks) they were transferred into 6-well plates and designated passage 1. The original well usually retained a small number of cells. These were also cultured further by adding fresh medium. These wells served as a source of cells for several additional transfers to 6-well plates. It was noticeable that after the first passage in 6-well plates, cells frequently succumbed to virus infection, whereas those left in the original 12-well plate survived. In rare cases this would happen in some replicas of the same batch while others would maintain the persistent virus infection. Those passage 1 cells that did survive were transferred into larger flasks and from passage 2 onwards, were diluted 10 times when sub cultured. Thereafter, most of the batches survived. As passaging continued, the proportion of apparent healthy cells increased. Those cells showing cytopathic effects, including vacuoles and enlarged nucleus size, decreased; although these cells were present in every culture.

For the Hi5/AcUW1.*lacZ* 2018 culture, after passage 9 was reached (10 weeks), the cells succumbed to overt virus infection. At the time this thesis was written, AcP6.9-C84A-DsRed and AcPolh-C84A-UK infections had reached passage 15 and had been in culture for five months. For AcP6.9-C84A-UK persistent infection, two attempts were needed, therefore, by the time the project was finished the cultures had reached passage 7 and were three and a half months old. Different passages of these cultures are shown (Fig. 6.2). Fluorescent imaging of the Hi5/AcP6.9-lef2C84A-DsRed culture demonstrated a strong fluorescent signal. However, fluorescence was not present in every cell and usually corresponded to the enlarged cells (Fig. 6.2 b and b').

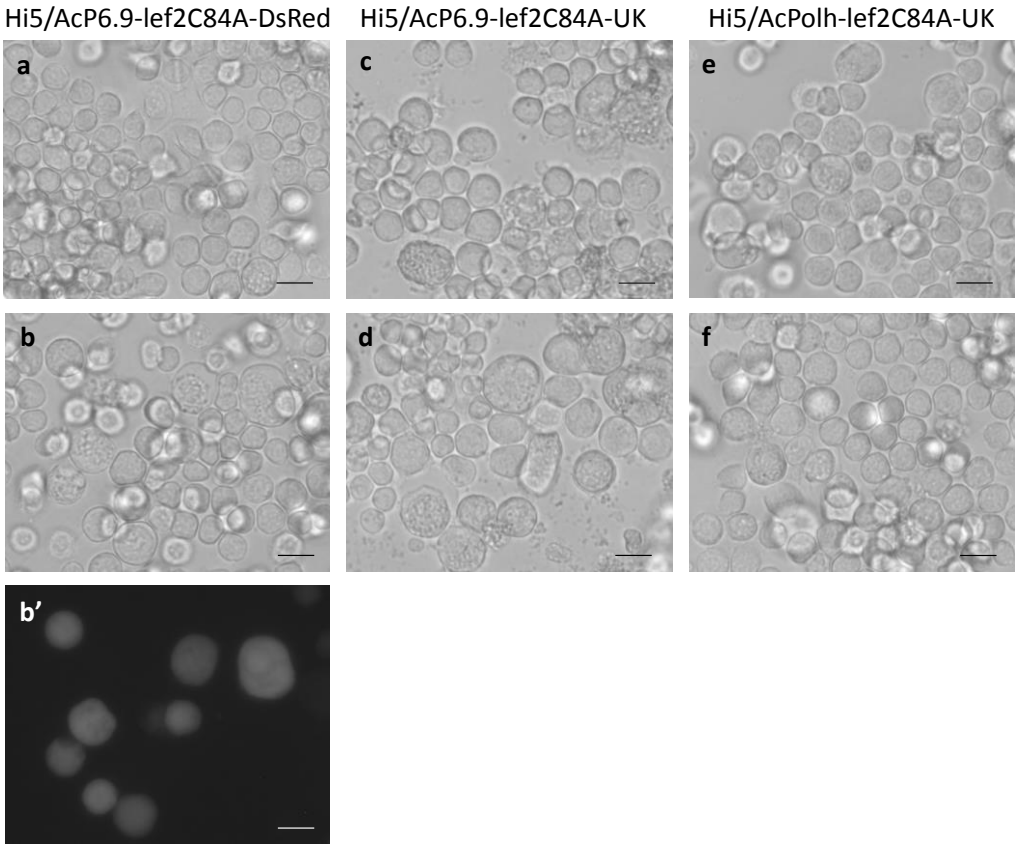


Figure 6.2: Bright field and fluorescence microscopy observation of persistent virus-infected cell cultures.
Bright field images of persistent virus-infected culture of Hi5 cells harbouring: AcP6.9-lef2C84A-DsRed at (a) passage 5 and (b) passage 8; AcP6.9-lef2C84A-UK at (c) passage 1 and (d) passage 2; and AcPolh-lef2C84A-UK at (e) passage 4 and (f) passage 6. Panel b' represents a fluorescent microscopy image of panel b (red channel). Scale bar 20μm. Images are representative.

6.2.3 Production of urokinase and DsRed by persistently infected Hi5 cell cultures

To test urokinase production, an activity assay kit was used as described in section 2.8. At 72h p.s. of passage 9 of the persistent Hi5/AcPolh-lef2C84A-UK culture, the production of urokinase was very low and, therefore with this small number of samples, no significant difference was observed compared with the mock samples even 24h after the beginning of the test ($p_{t\text{-}test}>0.05$; data not shown). Furthermore, urokinase expression was observed when Hi5 cells were infected with the same virus.

Production of DsRed in Hi5/AcP6.9-C84A-DsRed was quantified between passages 9 and 12. Florescence assays using a plate reader were carried out to analyze the total expression of DsRed in the persistently infected culture (Fig 6.3A), while flow cytometry studies enabled an estimation of the proportion of cells that were infected (Fig. 6.3B).

Overall production of DsRed in the Hi5/AcP6.9-C84A-DsRed culture was measured and compared to a lytic infection with the same virus and with a virus expressing the fluorescent gene under *polh* but carrying a WT *lef2* gene (see 2.8). The persistently infected cells produced fluorescence and the signal was significantly higher than in a productive infection with the same virus (Figure 6.3A). This surprising result was probably due to an accumulation of protein in the cells before they were passaged for the present experiment. Moreover, the signal produced by this virus in either persistent or productive infection was dramatically inferior (61 and 238-fold respectivelly) to the signal produced by a lytic infection with a WT *lef2* virus (AcDsRed, $p_{Mann\text{-}Whitney}<0.01$, Fig. 6.3A).

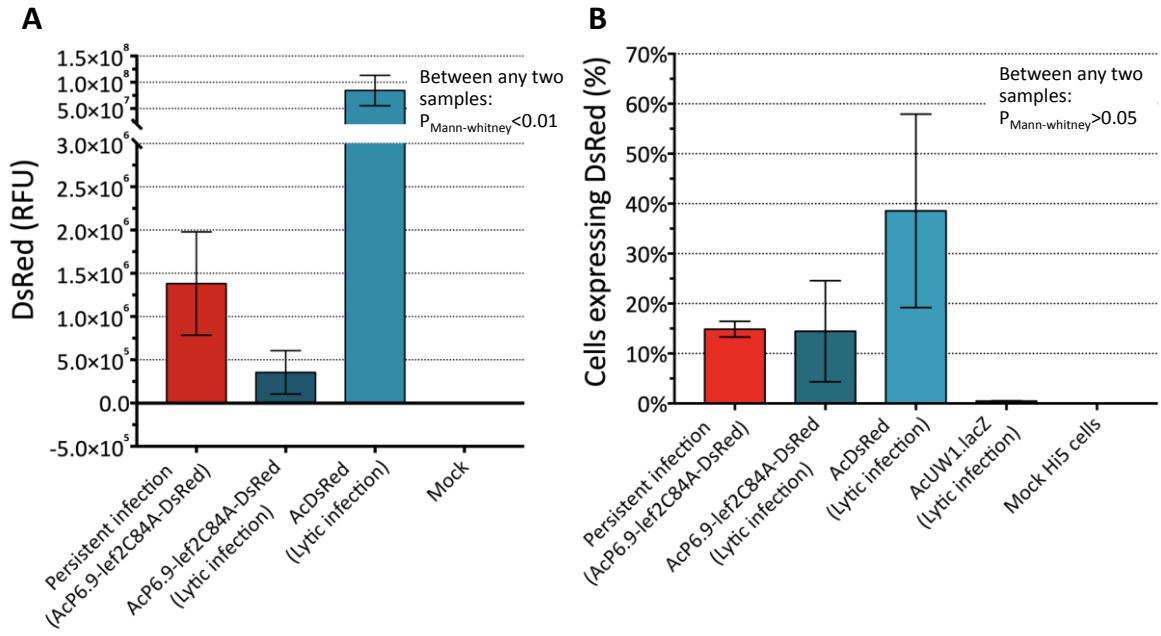


Figure 6.3: Production of DsRed in persistent and lytic infections.
Hi5/AcP6.9-lef2C84A-DsRed persistent infected cells and Hi5 cells freshly infected with 0.1 MOI of AcP6.9-lef2C84A-DsRed, AcDsRed and AcUW1.*lacZ* (B only) were seeded in shaking cultures (A) or monolayers (B) in multi-well plates and left incubating for 96h. Mock infected cells were grown at the same time. A) The cells were washed and treated with PBS with 1% NP40. The media were recovered and used to measure DsRed fluorescence signal in a plate reader (556nm/586nm for excitation and emission respectively). Average background fluorescent from six wells containing PBS with 1% NP40 only was discounted from all samples. B) Cells were harvested, washed and resuspended in DPBS for flow cytometry analysis. Scale bars represent SD from biological and technical replicas of A) the three independent experiments at passage 9, 10 and 11 or B) from biological replicas of a single experiment at passage 12. Mann-whitney test performed in between all samples for both graphs.

In agreement with previous imaging observations at passage 5 (Fig 6.2b and b'), flow cytometry reported that only 15% of persistently AcP6.9-lef2C84A-DsRed-infected cells produced the recombinant fluorescent protein (Fig. 6.3B) and that the strongest fluorescent signal was obtained from the largest cells in the population (data not

shown). This was similar to what happened in a lytic infection in Hi5 cells with the same virus (14.5%) and considerably lower than a productive infection with AcDsred (38.6%). In this last control, a high variability between samples was observed (27-61% of cells expressing DsRed). The fact that some of the positive control AcDsRed-infected Hi5 cells were not producing fluorescence was confirmed by fluorescence microscopy observation just before the experiment was carried out (data not shown). Therefore, owing to a high variability present in the positive controls (AcDsRed productive infections) and not enough replicates due to time limitations, no significant results were obtained from either comparison.

6.2.4 Genomic analysis of persistent viruses

Chapter 5 identified genomic regions that had suffered major deletions, insertions and INDELS in AcC20. To determine if they were linked to the establishment of a persistent infection, PCR was used to screen *egt*/*orf16*, *fp25k*, *p35*, *p94* and *p95* regions in virus DNAs from the new cell lines. For these regions, DNA up to passage seven of AcP6.9-lef2C84A-DsRed and AcPolh-lef2C84A-UK, up to passage six of Hi5/AcUW1.*lacZ* 2018 (*fp25k* only) and up to passage three of AcP6.9-lef2C84A-UK was tested. For *egt*/*orf16*, *p95* and *p94*, the last passage available of each was further analyzed by PCR and Sanger sequencing and all presented the WT version of the region (data not shown). For *p35*, several bands were present due to the design of the primer in a *hr* region, however, the products presented the same size as in AcMNPV WT.

Finally, *fp25k* presented an interesting pattern of products. The first tested passage for each persistent virus had a size resembling the WT (Fig. 6.4). However, in subsequent subcultures one or two higher weight molecular bands appeared in all four persistent viruses. These were first detected at passage one, two, three or four of AcP6.9-lef2C84A-Uk, AcPolh-lef2C84A-UK, AcP6.9-lef2C84A-DsRed and AcUW1.*lacZ* persistent infections in Hi5, respectively. The results suggest one or more insertions of around 200bp-400bp might have occurred, generating two or more distinct populations. For Hi5/AcUW1.*lacZ* 2018, the lowest molecular weight band from passage 5 was purified from the electrophoresis gel and analysed by Sanger sequencing. Results showed the sequence was still the same as described for AcMNPV WT (Ayres *et al.*, 1994).

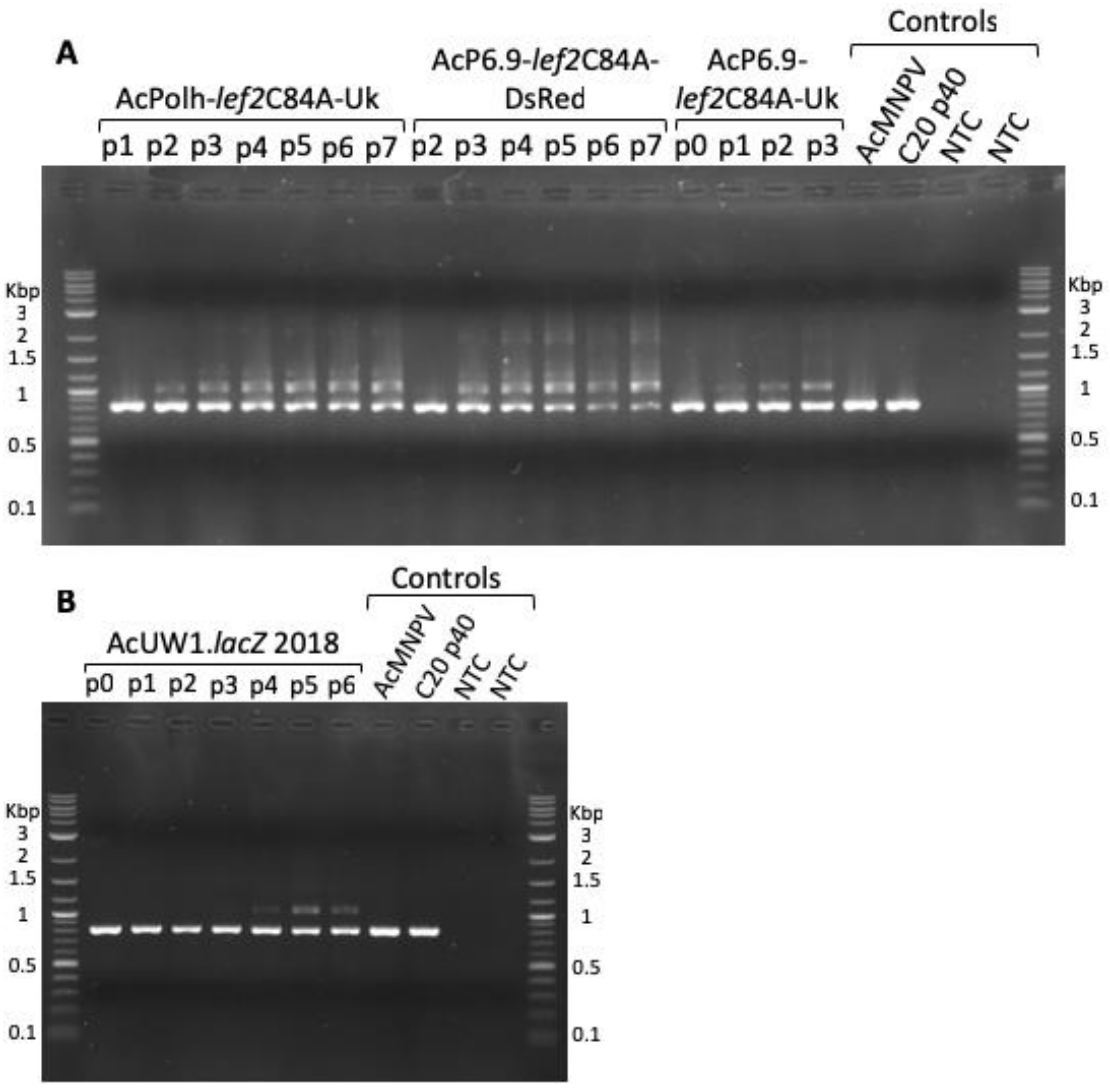


Figure 6.4: Evolution of *fp25k* region in persistent infections
Amplification of *fp25k* in passages of persistent infections: A) Hi5/AcPolh-lef2C84A-Uk passage 1-7, Hi5/AcP6.9-lef2C84A-DsRed passage 2-7 and Hi5/AcP6.9-lef2C84A-Uk passage 0-3. B) Hi5/AcUW1.*lacZ* passage 0-6. Reaction controls for both gels: AcMNPV WT, C20 p40 and NTC. First and last lanes of each gel contain 2-log ladder.

6.3 Discussion

This chapter aimed to produce persistent virus infections expressing genes of interest and analyse the possible application of this process for large scale protein production. The viruses constructed expressed *dsred* and *urokinase* under the *polh* or *p6.9* promoters. These viruses were *polh* negative and had a non-synonymous mutation in *lef2* that resulted in a virus with limited production of BV (Bannach, 2018). It was hypothesised that the impaired production of BV would give the cells more time to adapt to the virus infection, increasing the probability of generating a persistent infection.

It was noted that in the early passages minor changes in culture conditions such as the transfer to a new vessel or the total replacement of the media would affect the survival of the cells triggering a lytic stage in several batches of newly established persistent infections. This has been described before (Crawford and Sheehan, 1983) and was observed during the establishment of the Hi5/AcUW1.*lacZ* culture from which the C20 cells were cloned (R.D. Possee, personal communication) and might demonstrate the fine balance existing in a persistent virus infection. However, the persistent infections with the *lef2* mutation generated in this chapter survived up to at least seven passages and for at least three and a half months in culture. The Hi5/AcUW1.*lacZ* 2018 culture here generated was in culture for 10 weeks.

In comparison with the several unsuccessful attempts performed with AcUW1.*lacZ*, the three viruses used to establish persistent infections with the *lef2* mutation successfully developed into a persistent infection at the first or second attempts. These results demonstrated that reducing BV production in the early stages of infection creats more opportunity for the cells to adapt to the virus, thus increasing the likelihood of generating persistent infections.

Notably the last attempt in this project to establish a persistent infection with AcUW1.*lacZ* was the only one that worked. Over the project several attempts were performed with different MOIs. Although no clear change was noted in the protocol that originated this persistent infection in Hi5 cells, the cells were left for less time before sub culture. In the previous attempts the cells were left after reaching confluency hoping that only the healthy clusters survived. Therefore, when sub-cultured at this late stage (>4 weeks), high levels of BV were present in culture. On the contrary, in the present chapter, the cells were sub-cultured as soon as some started to deattach. This small change might support a higher survival rate since it kept down the level of BV in the media. Moreover, the experiment was performed independently twice with three quantities of infecting virus tested. For all the tested proportions of virus, separate persistent infections were generated and cultured independently, most surviving up to passage two, after which only two single cultures were further maintained. This suggests the method was more robust than previously.

Regarding production of the recombinant proteins, different results were observed for DsRed under *p6.9* promoter or urokinase under *polh*. The production of DsRed in the Hi5/AcP6.9-C84A-DsRed culture was obvious when observing the cells under the microscope as this protein emits in the visible spectrum (cells appear red). Measurement of DsRed fluorescent levels showed there was a significant amount of the recombinant protein being produced in this culture even though it was lower than that obtained in a lytic infection with AcDsRed (recombinant virus with DsRed under *polh* promoter but a WT *lef2*).

Flow cytometry confirmed that in the Hi5/AcP6.9-lef2C84A-DsRed culture only a proportion of the cells seemed to be producing the protein of interest (15%) and a relation between cell size and expression of DsRed might apply. This is not surprising considering virus-infected cells usually develop enlarged nuclei. The fact that less than 20% of the persistently infected cells were producing DsRed might demonstrate the presence of a mixed population in this early stage of a persistent infection or may reflect a cyclic nature undergone to maintain the host-virus balance. Furthermore, it can explain the low total fluorescence observed on the plate reader. However, no statistically significant results were obtained between the different virus infections by flow cytometry. The positive control for this experiment, a productive infection of AcDsRed in Hi5 cells, showed less than 40% of the cells were producing DsRed at 96h p.i; a very high variability in between samples was observed. This might highlight the need for further optimization of this experiment since at this late time point with *dsred* under the *p6.9* late promoter, a higher accumulation of the recombinant protein would be expected. Increasing the MOI might enable a higher protein expression and repeating the experiment could overcome the reproducibility problem, increasing experimental robustness. Nonetheless, together these results demonstrate a persistent infection was present and recombinant proteins were produced even if at a low level. Similarly, in C20 cells only a small proportion of the cells were shown to be infected (see section 3.2.7). However, these tests were done after the cells and virus were cultured for over 130 passages. Despite these results, at the time the C20 cell line was cloned (p42 of Hi5/AcUW1.*lacZ* culture), single cell dilution demonstrated every single cell was infected (R. Possee personal communication).

Unfortunately, tests for the production of urokinase revealed Hi5/AcPolh-lef2C84A-Uk persistently infected cells were not producing the enzyme. However, only one test was done and further optimizations could be necessary. Furthermore the lack of sufficient replicates hindered statistical analysis. The fact that all persistent viruses had a *lef2* mutation that hampered viral replication probably prevented higher levels of expression of DsRed and detectable levels of urokinase. Furthermore, these two genes were encoded under different promoters, from the late (*p6.9*) and very late (*polh*) phases of infection. In this case, higher expression levels were expected under the *polh* promoter since this gene is naturally highly expressed in the very late phase of infection (Chen *et al.*, 2013).

Problems of gene expression under *polh* were already experienced in chapter 4 (see 4.2.2). However, it was reported before that urokinase expression seems to be higher under *p6.9* rather than *polh* (Lawrie *et al.*, 1995). It was suggested that secreted proteins could be poorly expressed using *polh* promoter due to the severely compromised host cell secretory pathway at the very late stages of infection (Jarvis *et al.*, 1990; King and Possee, 1992). However, in the present study expression of urokinase was observed in the lytic infection established as a positive control with the same virus present in the persistent infection (AcPolh-lef2C84A-Uk), corroborating the previous studies showing expressiong of UK under this very late promoter is still possible (Lawrie *et al.*, 1995). Expression of DsRed and urokinase were tested at different times post-seeding (96 and 72h p.s., respectively), which might also have enabled a higher accumulation of recombinant protein in the first persistent infection. However, this would not explain the absence of protein production observed for urokinase since the cells were seeded at a high density ($1x10^6$cells/ml) and most of the enzyme-contained media was maintained.

At the genomic level, the persistent viruses were screened for the previously detected major changes in AcC20. An interesting result was achieved for *fp25k*, which indicated an insertion or duplication occurred in this gene that originated at least two subpopulations from very early in each culture (passage 1-4 depending on the persistent virus). Several studies have reported the insertion of host DNA transposons in the *fp25k* gene as the cause for the FP profile (Cary *et al.*, 1989; Wang *et al.*, 1989; Beames and Summers, 1990; Wang and Fraser, 1993). In the last generated Hi5/AcUW1.lacZ 2018 culture a decrease in the proportion of cells producing polyhedra was noted from

passage 3 to 4 by visual microscopy observation. The *fp25k* polymorphism was first detected in passage 4 of this culture, suggesting these events might be related as previously described.

However, in AcC20 this profile was actually generated due to the insertion of an A within a row of 7 that created a truncated FP protein (see section 5.2.3). This insertion had occurred before p16 of Hi5/AcUW1.*lacZ* persistently infected culture that preceded the C20 cell line. This genotype seems to often appear when viruses are passaged (Potter *et al.*, 1976). The fact that FP is not essential *in vitro* and the presence of several transposon target sites (TTAA) and a poly-A sequence, seem to increase the probability of mutations in this gene (Cary *et al.*, 1989; Wang *et al.*, 1989; Beames and Summers, 1990; Wang and Fraser, 1993; Kelly *et al.*, 2006). In the present chapter, the absence of the A insertion could be reported for the smaller molecular band obtained from the PCR of passage 5 of Hi5/AcUW1.*lacZ* 2018. However the remaining persistent infections, passages and PCR products were not analysed due to time restrictions and the proximity of the generated PCR products at each passage.

Although the nature of the higher molecular weight products was not clear, the presence in the four viruses suggests it was associated with the establishment of the persistency. Cloning of the cell population or sub cloning of *fp25k* into a plasmid vector would enable Sanger sequencing of these products to be performed and clarify its origin. Nevertheless, these results and previous documented work indicated that the use of persistently infected cultures for protein production needs to be considered cautiously. The cultures cannot be maintained for more than a few passages since it increases the risk of mutations. Stabilization of the *fp25k* gene by inserting SNPs that would break the runs of A's but maintained the a.a. coding region might prevent this mutation from occurring. Furthermore transposon target sites could also be stabilized as described before (Giri *et al.*, 2010). However, this concern might also be valid for other genes non-essential *in vitro* that could be susceptible to insertions. It would be interesting to determine if cells not expressing DsRed in the Hi5/AcP6.9-lefC84A-DsRed culture are still infected. This would be achieved by single cell analysis. This would also enable to determine if a relationship exists between the population expressing fluorescence and the presence of the FP25 or other FP phenotype.

6.4 Conclusion

In this chapter, persistent infections in Hi5 cells harbouring viruses expressing *dsred* and *urokinase* under late or very late promoters were generated. The viruses had a mutated *lef2* provoking a slow production of BV. The results suggested this might create a milder environment for the infected culture and promote the generation of a persistent infection. The generated cultures included cells with signs of infection while other clusters looked healthy as observed in C20 cells. Expression of DsRed from the Hi5/AcP6.9-C84A-DsRed covert infection was demonstrated but not all cells were expressing the recombinant protein. Why this occurs is not clear and would be important to clarify in future studies of persistent infections. Preliminary tests for the production of urokinase failed to demonstrate its production. Further analysis and optimization of culture conditions would be necessary. An insertion in the *fp25k* gene appeared in early passaging of the covert cultures and although this was not characterized, it was present in the three generated *lef2* mutant-persistent infections and Hi5/AcUW1.*lacZ* 2018. Further studies would be needed to clarify if this is related with the generation of persistent infections or with the particular virus mutation present.

The results demonstrate that although persistent infections can be used to purify recombinant proteins in a continuous way, further optimizations would be needed to obtain a higher expression of the desired product. Stabilization of the *fp25k* gene will reveal if by preventing the FP mutation to occur and therefore maintaining the quantity of produced BV, persistent infections could be established as easily as generated in this chapter with an impaired *lef2* virus. The use of persistent infections for production of recombinant proteins has been shown to be possible. However, this should only be used for a restricted number of passages to avoid possible genomic changes occurring in non-essential genes when sub cultured. Moreover, regular genomic tests should be performed to confirm the intact nature of the recombinant protein.

Chapter 7

Final discussion and future work

7.1 Introduction

After baculovirus entry and expression of the early genes, the host cell cycle is arrested and the virus utilises the host cell machinery to produce infectious particles (Braunagel *et al.*, 1998). Release of BV spreads the infection to neighbouring cells *in vivo* and *in vitro*. Later, virus particles become occluded, which is important for horizontal transmission. Baculoviruses, however, can also remain in a covert form, enabling maintenance of the virus in the host across generations. In this case, a balance between the host and virus is created that enables both to co-exist (Goic and Saleh, 2012; Jakubowska *et al.*, 2014). Although this is common in insects (Williams *et al.*, 2017) it is rarely achieved in culture, therefore very little is known about how this equilibrium is created and which interactions are important. The few previous studies focused on AcMNPV persistent infections *in vitro*, were largely carried out before the advance of the genomic techniques (McIntosh and Ignoffo, 1981; Crawford and Sheehan, 1983; Hughes *et al.*, 1993, 1997; Lee *et al.*, 1998). Therefore, a detailed analysis of the persistently-infected C20 cell line, combining morphological, viral, genomic and transcriptomic data, was performed in this thesis.

It has not been possible to demonstrate infection of insect cells with a second baculovirus (superinfection) after the first infection has been established *in vitro* (Weng *et al.*, 2009; Beperet *et al.*, 2014). Nonetheless in nature, many insects are colonized by more than one virus (Burden *et al.*, 2003, 2006; Jakubowska *et al.*, 2014). In AcMNPV, superinfection-exclusion seems to be regulated through gene expression and release of BV from the primary infection (Beperet *et al.*, 2014). However, what mechanisms prevent secondary virus infection are unknown. Furthermore, this interference phenomenon was not studied in a persistent infection so it was unclear if the usual low level infection present would interfere with a secondary infection the same way. This study included analysis of superinfection in the C20 cells to study any barriers to superinfection and explore the interactions occurring between the host and virus.

Baculovirus specificity for insect hosts has been studied primarily for the purposes of producing biopesticides for pest control (Haase *et al.*, 2015; Lacey *et al.*, 2015). Biopesticides from *Alphabaculovirus* and *Betabaculovirus* spp. have been used for over 60 years (Cory and Bishop, 1997; Lacey *et al.*, 2015). Their use in this way has been

facilitated by a virus biphasic replication cycle that culminates with the production of OBs, which are robust and stable in the environment. Over the last 30 years, the focus of applied baculovirus research has switched to the production of recombinant proteins, including more recently subunit vaccines (Cox *et al.*, 2008) or more complex VLPs (Schiller and Lowy, 2001; Kost and Kemp, 2016). Even though considerable progress has been made, recombinant protein production by BEVS is still a time-consuming technique. Proteins are largely made in the late to very late phase of infection, where cell lysis eventually occurs; thus proteins are made in a batch process. This thesis considered whether a continuous production method, based on persistently-infected cultures, could increase the efficacy of protein production. In addition, a more detailed understanding of the mechanisms involved in the establishment of a persistent infection may also enable us to generate more reliably cell lines containing persistent recombinant viruses expressing genes of interest.

7.2 Final discussion and future work

7.2.1 The persistent infection

The C20 cell line comprises cells generally larger than the parental stock with a greater variability of sizes. Increase in size, together with the presence of cytoplasmic vacuoles and nuclear polyhedra reflect signs of baculovirus infection. However, variation in size and lack of a uniform presence of cytopathic effects may indicate that not all cells are infected, as described in one of the earliest reported baculovirus persistent infections in culture (McIntosh and Ignoffo, 1981).

The most surprising result was the resilience presented by the C20 cells. These cells were able to grow in culture for over a month without replacement of media or nutrients, presumably whilst accumulating toxic waste products. When eventually the viability of the cells reduced considerably, the culture was able to be recovered by passaging into new media. This was demonstrated on a number of occasions, by passaging cells that had been in culture for between two and four weeks. No other reports in the literatures describe an insect cell culture that can survive for such a long period of time.

It was hypothesised that differences in host-virus regulation might be responsible for the maintenance of the persistent infection. However, transcriptome analyses revealed that the majority of host transcription was not differently regulated. The differences noted suggested mechanisms for mRNA transport through the NPC might be compromised and that cell cycle progression may be slower. Indeed, a slower growth pattern was observed for C20 cells compared to Hi5 cells (Fig. 3.4). Furthermore, cell activators of apoptotic transcription factors were shown to be down regulated and genes involved in autophagy were repressed, which might be directly related with the long survival pattern (above), as well as the maintenance of the persistent infection in cells. In order to understand the impact that regulation of these genes has on the maintenance of the persistent infection, it would be interesting to further explore how these genes are regulated in early passages of an infection of this kind.

Although the robustness of the C20 cells was impressive, these tests were carried out when the culture was well established (after passage 70). In comparison, the early Hi5/AcUW1.*lacZ* cultures preceding C20 were more sensitive to the stress of being "passaged" (R. Possee personal communication). An almost cyclical pattern was observed in these early cultures; the cells would appear healthy and suddenly, in the next passage, cytopathic effects would increase, probably a result of a higher viral load present, and a lytic infection would ensue leading to cell death. This happened in as many as 50% of the initial cultures. However, if the cells survived this critical phase, they would recover a healthy appearance. For this reason, several independent flasks of the cell line were kept at all times. Once the C20 clonal cell line had been established and, as the number of passages increased, these critical cycles happened less often. By the end of the thesis, at passage 160, it was common that C20 cells would have a less healthy appearance for a few passages, but the culture would survive and recover.

A similar pattern was observed upon the establishment of the *lef2*-C84A persistent infections, where up to passage 10, some replicate cultures died of re-activated lytic infection, while others survived (Chapter 6). These examples demonstrate the complex virus-cell interactions affecting BV production, and are probably influenced by many factors such as cell growth rate and subculture seeding density (Crawford and Sheehan, 1983). The observed resilience at higher passages may result from host and virus adaptation over time, creating a more robust balance, or may be a consequence of the

overall decreasing trend in BV titre observed in the C20 cells (Fig. 3.9), even though BV were present in every C20 passage.

The existence of a dynamic balance was also observed within single C20 cultures over time. An increase in BV production seemed to correlate to an increase in the number of viable cells, and *vice versa* (Fig. 3.10 and 3.12). Also, a dynamic up and down regulation of viral gene expression occurred although more and later time points would be needed to corroborate this hypothesis (tested at 0, 24 and 48h p.s.; Fig. 5.11).

It has been suggested that covert baculovirus infections are associated with low virulence (reviewed Williams *et al.*, 2017). In SeMNPV, genotypic variants from field-collected insects transmitted vertically in the population were less pathogenic and slower to kill than variants collected from the soil, which presumably were the consequence of an overt infection (Cabodevilla *et al.*, 2011). Similarly, OB from a SfMNPV persistent infection were determined to be three times less infectious for larvae than the ones produced in the parental cell line (McIntosh and Ignoffo, 1981). However, this was not the case in AcC20 in cell culture, since the BV isolated from C20 cells was still able to infect and kill Sf9 cells (R. Possee, personal communication).

One of the first studies describing the establishment of a persistent infection suggests there are three phases (Crawford and Sheehan, 1983). In the first phase, a high proportion of cells die from infection; in the second, the virus load decreases as well as the number of infected cells. The same process (described above) was observed upon the establishment of the C20 cell line and other persistent infections (Chapter 6); however, the virus present was not titrated at this stage. In phase three, less than 1% of cells were infected (Crawford and Sheehan, 1983). These cells were “cured” by cloning and selection of non-infected cells but the doubling time had increased and the cells became resistant to superinfection thereafter. This suggests these cells may have actually maintained the persistent infection or perhaps developed a latent infection; although at the time, molecular techniques were not as sensitive and may simply not have detected the infection (Crawford and Sheehan, 1983).

At later C20 passages, there were fewer visible signs of infection in C20 cells and a flow cytometry screen for C20 cells presenting GP64 revealed 7% of the cells were infected at passage 156 (this test was only performed once, so repetition is needed to enable

robustness). However, only cells with viruses replicating in the late phase of infection would be considered as infected by this method. Nonetheless, this might explain the diversity of C20 cell morphology observed and the fact that some cells resembled healthy Hi5. Host genes were still transcribed at high levels in these cells as 95% of the overall C20 transcriptome is from the insect genome. This may reflect that the infection was present at a low level in all the cells or indicate that only a subset of the cells was infected (although AcC20 gene expression occurs at normal levels in the population of infected cells this is diluted in the total population of cells). It is worth noting that during the cloning process to derive the C20 cell line, every clonal cell line established tested positive for virus (R. Possee personal communication). This raises the question whether AcC20 is being lost over time. A persistent SeMNPV infection has been shown to convert into a latent infection with cells harbouring and transcribing incomplete viral genomes (Weng *et al.*, 2009). However, in neither of these studies did the cell line survive for such a long time in culture before the effects were described (Crawford and Sheehan, 1983; Weng *et al.*, 2009).

As described above, the continuous co-evolution of virus and cells created variability over more than 160 passages. This could also explain the variable titres and production of GP64 within and between different cultures. Furthermore, genomic analysis demonstrated mutations had occurred over time. In some cases, more than one genotype was found in the same passage demonstrating the dynamic evolution of the virus. Mutations were detected in 38 genes and one homologous region (*hr2*).

A comparative genomic analysis of SeMNPV isolates from different backgrounds analysed virulence and host-virus interactions (Theze *et al.*, 2014). Mutations were found in 48 ORFs compared to the reference isolate involving virion proteins important for primary and secondary infection and for inhibition of apoptosis. This aimed to identify possible pathways linked to the tendency of generating a persistent or a lytic infection but concluded transmission phenotypes involve multiple molecular pathways (Theze *et al.*, 2014). In the present study, eight of the non-synonymous mutated genes encoded for essential proteins for BV replication: PK1, DA26, ARIF-1, AC53, Lef-9, helicase, GP64 and AC83. Truncated proteins for DA26 and AC83 and an extended ARIF-1 were predicted. These genes cover a range of pathways, from different phases of infection including regulating levels of immediate early proteins (DA26), assembly of

AcMNPV RNA polymerase (Lef-9), DNA replication (Helicase), inductors of actin rearrangement (ARIF-1) and essential genes for complete nucleocapsid assembly (PK1, ORF53, GP64, AC83). Remarkably, RNA sequencing revealed all AcMNPV genes were expressed in C20 cells, although transcription occurred at a low level. Late genes were the highest expressed with transcripts of *p6.9* and *orf76* the most abundant.

Mutations in the *fp25k* were identified in all the persistent infections described in this thesis. An extra A in a sequence of 7 causing the FP25 profile was detected in the C20 cell line from p16 of the Hi5/AcUW1.lacZ initial culture, the earlier tested passage. In the *lef2* mutant persistent infections, insertions appeared in the *fp25k* in very early cultures although a mixed population with the WT genotype exists (see 6.2.4 and 7.2.3). This demonstrates how vulnerable *fp25k* is to mutation but also how quickly mutations can occur and be maintained in a persistent virus.

Functional protein studies and 3D models would complete this study and reveal if the predicted proteins from mutated essential genes were still fully functional. Moreover it could indicate if any of these mutations played a role in the establishment and maintenance of the persistent infection. Although some comparison was possible with the earlier virus passages, the DNA was amplified in Hi5 cells rather than directly extracted from the cells as in passage 40 for WGS. In future persistent infections studies, to enable a more precise follow up of the genomic changes, the cultures should be scaled up regularly for DNA extraction. A comparison of the genomes at different phases of persistent infection, e.g., when cells are still vulnerable to an overt infection or after stabilization of the persistent infection, will help clarify which genomic changes are necessary for this type of infection to occur. However, several scenarios might be possible for the establishment of persistent infections, as probably involving one or more compromised genes, or different sets of them, could achieve the same phenotype.

7.2.2 Superinfection of a persistently infected cell line

In the present work, gene expression, microscopy, genomic and transcriptomic techniques demonstrated superinfection in a persistent infection is possible. Confocal microscopy studies showed fluorescence produced by a secondary infecting tagged-virus in C20 cells from 48h p.i. indicating the replication cycle of the superinfecting virus is slower than in a lytic infection. Furthermore, only a small percentage of the cells could

be secondarily infected, while others seemed to be resistant (Fig. 4.1). When analysing the early phases of infection almost half of the C20 cells observed seemed to be unable to adsorb new BV, while over 90% of the cells were unable to uptake virus (Fig. 4.8). Therefore, the primary blockage seems to occur as early as the binding step, even though the capacity for virus uptake also seems to be inhibited. This indicated the blockage to superinfection may involve several pathways instead of a single compromised step.

Transcriptome analysis of superinfection revealed the number of viral transcripts increased over 48h, although the typical temporal regulated phases of infection were not observed as gene expression did not vary greatly over time. It was highlighted that host genes encoding for proteins required for the endocytic pit formation were repressed. The *erm* homolog 1, probably involved in the linkage between the actin cytoskeleton and transmembrane receptors (Batchelor *et al.*, 2004), was down regulated. This might indicate a deficiency in C20 capacity for endocytosis and hence may be related to the reduced ability for virus uptake observed in C20 cells. Moreover, lipid synthesis in C20 cells appeared to be differentially regulated, which may affect membrane composition and therefore reduce the affinity for virus binding (Chernomordik *et al.*, 1995; Tani *et al.*, 2001).

Even though these results can not be directly extrapolated to the general resistance to superinfection observed with homologous and heterologous virus in AcMNPV overt infections, it is possible that the early phases of infection (binding and uptake) are the first ones to be compromised once host transcription starts being down regulated in a lytic infection. In a productive infection, resistance to a secondary infection was shown to start around 12-15h p.i., at the initiation of the late phase of infection. In this phase cell processes are highly essential for viral replication. Uptake mechanisms are not needed anymore, so it would be logical that the cell arrest shown to occur in the present study, and previously (Braunagel *et al.*, 1998), would start by shutting down non-necessary mechanisms for virus replication. However, further studies of superinfection in productive infections would be needed to corroborate this theory.

Superinfection studies suggested the current batches of C20 cells might actually be composed of a mix of one or more populations. The fact that when challenged, some

cells were able to accomplish gene expression up to the very late phase, while others seemed resistant, is intriguing. Furthermore, superinfection in C20 was self-contained as it did not seem to spread to the rest of the culture (Fig 4.5). This suggests BV from a secondary infection were not being released, while AcC20 virus particles were being produced in C20 cells, indicating the host machinery for BV production was not compromised. Furthermore, transcriptome studies showed genes from the late phase, necessary for BV production, were the highest expressed both in non-challenged or superinfected C20 cells.

An interesting follow up of these results would consist of trying to establish a second superinfection after C20 cells had survived the challenge with a first virus. If, in accordance with previous results, only a sub-population can be superinfected but the infection does not spread, this set of cells will likely succumb to the first superinfection. A second attempt of secondarily infecting the same C20 cell batch would reveal if it was only a specific sub-population that was susceptible, so none or fewer cells could get superinfected this time. On the other hand, if again a subpopulation of cells could be superinfected it meant that other underlying factors were determining the superinfection susceptibility. An example would be that the ability of each cell to be superinfected could dependent on the replication cycle phase of the primary persistent infection.

Both WGS and transcriptomic studies were undertaken using a sample of the whole population, hence, the presence of a complete genome and transcriptome does not ensure this is the case in every cell. Therefore, if C20 cells only harboured partial AcC20 genomes this would probably drive different levels of resistance to superinfection and explain the patterns of superinfection exclusion observed. A previous latent-like baculovirus infection was achieved where incomplete SeMNPV genomes and transcriptomes were present in the cells and were enough to interfere with a superinfection of an homologous virus (Weng *et al.*, 2009). This theory could be tested by cloning the linear AcC20 genome into a low copy number plasmid vector (bacmid; Luckow *et al.*, 1993) and characterize the genome population using this system (e.g. AcC20 genome sizes, presence of partial genomes, range of deleted mutants).

Studies on superinfection also raised the question whether the cells were resistant to its own virus or just to other AcMNPV virus. The maintenance of the C20 persistent infection itself, having infectious virus in culture, suggests some cells might be resistant to infection with AcC20. On the other hand, a cyclic pattern might be occurring where C20 cells are infected and dying but since virus titres are low (Fig. 3.9 and 3.10), the remaining cells keep replicating providing more cells to be infected. Two approaches could be undertaken to clarify these hypotheses: a) construction of a fluorescent tagged version of the current AcC20 virus to challenge C20 cells and detect if the proportion of superinfected cells is the same as observed before; b) re-cloning the C20 cells and characterize the cells either by single cell analysis or by examining early cultures obtained from each clone. Genomic and transcriptomic studies could be repeated in the new clones alone or when superinfected. The need for this second point was highlighted by the fact that less than 10% of C20 cells were shown to be infected in today's cultures, even though this could also be caused by the presence of partial genomes.

7.2.3 The generation of persistent infections *in vitro* and their applications

The last chapter aimed to develop a reliable system to create persistent infections and to test if these could be used for recombinant protein production, thus creating an improved continuous BEVS. Several unsuccessful trials were carried out to generate a new persistent infection with AcUW1.*lacZ*, although the very last attempt generated a persistent infection that was maintained up to passage 9. It was hypothesised that the generation of the FP25 profile in early passages of a persistent infection and consequent increase of BV in culture might be hindering this process. This was observed from the earliest tested passage of the infection that originated C20 cells. Therefore, in the last chapter persistent infections were created using viruses with impaired BV production, expressing recombinant proteins under late and very late genes. For two out of three viruses a persistent infection was successfully created at a first attempt, while the third one was only established after a second attempt. The three persistent infections reached between passage 7 and 15 at the time this thesis was written. This indicated a lower viral titre in the initial cultures may have increased the opportunity for the cells to adapt to the virus presence, hence facilitating the generation of persistent infections as

hypothesised. With continuing passaging, fewer batches have died of the infection probably indicating a more stable virus-host balance was being achieved.

Suitable tests for the production of each recombinant protein in two of the persistent infections (Hi5/AcP6.9-lef2C84A-DsRed and Hi5/AcPolh-lef2C84A-UK) revealed *dsred* was being expressed while *urokinase* was not. However, although the *dsred* expression tests were done repeatedly to enable reproducibility, due to time limitations, *urokinase* tests could not be repeated. Expression of *dsred* showed it was possible to use persistent infections in BEVS even though not all the cells seem to be infected at the same time. However, expression from the *lef2* mutant virus, both in persistent or productive infections was considerably lower than expression from a recombinant virus with a WT *lef2* (Fig. 6.3). The use of a mutated *lef2* virus implies a slower replication cycle since this is an essential gene and therefore other options might need to be considered for producing high levels of recombinant protein in BEVS.

A screen for the previously identified mutation regions in AcC20 genome revealed that a mutation in the *fp25k* gene occurred in the very early passages (passage 2, 3, 1 and 4 for AcPolh-*lef2*C84A-UK, AcP6.9-*lef2*C84A-DsRed, AcP6.9-*lef2*C84A-UK and AcUW1.*lacZ* respectively) although the genotype was different from that observed before. In this case a second longer amplicon was generated by PCR besides the predicted one. Sanger sequencing showed the inserted A detected in the AcC20 was not present in the lower band of AcUW1.*lacZ* passage 5, although the other cultures and passages were not tested. Nonetheless, this showed that mutations in the *fp25k* occur very early in infection as suggested in AcC20 and this might indeed be hindering the establishment of persistent infections. Moreover, the common AcMNPV FP25 mutation decreases the production of polyhedra. The *polh* promoter is usually the one chosen for expressing genes of interest since it is highly transcribed; therefore, the presence of the FP25 mutation will prevent high levels of recombinant protein production.

A previous study that stabilized the *fp25k* gene by removing the transposon target sites (TTAA) reported a delay on the appearance of the FP25 phenotype (Giri *et al.*, 2010). However the phenotype would still appear within a few subcultures. Furthermore, this did not stabilize the poly-A region, which is usually mutated upon virus passaging, as occurred in AcC20. Therefore, future work would include the establishment of persistent

infections not with impaired BV production but with a stabilized form of the *fp25k* gene, that will prevent BV production to suddenly increase and kill the culture. This could be achieved by inserting a point mutation in the run of As without disturbing the coding region, which could be combined with the stabilization of the transposon target sites. Further tests will be needed to test if this prevents the appearance of the FP phenotype. This would be followed by testing if the virus can be used to obtain recombinant protein production in a continuous system and compare the production level with a lytic infection with the same virus.

These results also highlight the importance of regular genetic analysis of the virus if using persistent infections for BEVS as other genes may be affected by point mutations that can affect directly or indirectly the desired protein.

All the persistent infected cell lines described in this thesis and previously established in our group (R. Possee, personal communication) were created in Hi5 cells. It is possible that the presence of a latent nodavirus infection in Hi5 cells (see 3.2.4) could be having a role. However, the generation of fresh persistently infected cultures was only achieved towards the end of the project so this possible relation was not yet studied. Published studies in persistent baculovirus infected cell lines were mainly achieved in *Spodoptera* spp. cells (McIntosh and Ignoffo, 1981; Crawford and Sheehan, 1983; Lee *et al.*, 1998; Weng *et al.*, 2009). Recently, these cells have been found to contain a persistent rhabdovirus infection (Ma *et al.*, 2014). As a reliable system to develop AcMNPV covert infections was thought to be developed in the present work, the next step would be to try the generation of covert infections in a nodavirus free *T. ni* cell line (Zhang and Thiem, 2010) with either *lef2* mutants or a *fp25k* stabilized recombinant virus. This will indicate whether the presence of a latent FHV infection is facilitating the generation of covert infections.

7.3 Concluding remarks

The persistently infected C20 cell line is characterized by a population of morphologically irregular cells with variable presence of cytopathic effects and high resilience in culture. They present both host and virus transcripts, although a low infectious virus titer is present. It was suggested that some of the cell defense genes might be down regulated

and eight essential virus genes were shown to have suffered mutations. However, a functional characterization of the proteins produced would be needed to understand its impact in the establishment or maintenance of this state of infection. After 8 years and over 160 passages it is possible that the C20 cells have diverged into two or more subpopulations with different susceptibilities to a secondary baculovirus infection. The apparent resistance to superinfection seen in the majority of cells seems to be related with a blockage in the binding and uptake mechanisms, possibly caused by a compromised endocytic pathway and changes in the membrane composition. However, transcriptomic data was retrieved from a single set of samples so repetition would be needed to confirm these preliminary results.

The creation of persistent virus infections seems to be facilitated by reducing BV output in the early stages of infection as demonstrated by the establishment of three infections of this kind with *lef2* mutant virus. In the future, stabilization of the *fp25k* gene and attempts to generate new persistent infections would be carried out. This would aim to understand if by stabilizing this gene, often responsible for generation of high virus titers in early passages, persistent infections can be generated and high levels of recombinant proteins can be produced in BEVS.

Literature

Adams, R.H. and Brown, D.T. 1985. BHK cells expressing Sindbis virus-induced homologous interference allow the translation of nonstructural genes of superinfecting virus. *Journal of Virology* 54: 351–357.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. 1990. Basic local alignment search tool. *Journal of Molecular Biology* 215: 403–410.

Anderson, R.M. and May, R.M. 1980. Infectious diseases and population cycles of forest insects. *Science* 210: 658–661.

Applied Biosystems 2008. Guide to performing relative quantitation of gene expression using real-time quantitative PCR.

Asgari, S. 2015. Regulatory role of cellular and viral microRNAs in insect–virus interactions. *Current Opinion in Insect Science* 8: 104–110.

Assenberg, R., Wan, P.T., Geisse, S. and Mayr, L.M. 2013. Advances in recombinant protein expression for use in pharmaceutical research. *Current Opinion in Structural Biology* 23: 393–402.

Au, S. and Panté, N. 2012. Nuclear transport of baculovirus: revealing the nuclear pore complex passage. *Journal of Structural Biology* 177: 90–98.

Au, S., Wu, W. and Panté, N. 2013. Baculovirus nuclear import: open, nuclear pore complex (NPC) sesame. *Viruses* 5: 1885–1900.

Ayres, M.D., Howard, S.C., Kuzio, J., Lopez-Ferber, M. and Possee, R.D. 1994. The complete DNA sequence of *Autographa californica* nuclear polyhedrosis virus. *Virology* 202: 586–605.

Bairoch, A. and Apweiler, R. 2000. The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Research* 28: 45–48.

Balhorn, R. 2007. The protamine family of sperm nuclear proteins. *Genome Biology* 8: 227.

Ball, L.A. and Johnson, K.L. 1998. Nodaviruses of Insects. In: Miller, L.K. and L.A. Ball (Eds.), The Insect Viruses, Springer US, Boston, MA, pp. 225–267.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A. and Pevzner, P.A. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology: A Journal of Computational Molecular Cell Biology* 19: 455–477.

Bannach, C. 2018. Regulation of late and very late gene expression in both lytic and persistent baculovirus infections. PhD. Thesis, Oxford Brookes University.

Batchelor, C.L., Woodward, A.M. and Crouch, D.H. 2004. Nuclear ERM (ezrin, radixin, moesin) proteins: regulation by cell density and nuclear import. *Experimental Cell Research* 296: 208–222.

Beames, B. and Summers, M.D. 1988. Comparisons of host cell DNA insertions and altered transcription at the site of insertions in few polyhedra baculovirus mutants. *Virology* 162: 206–220.

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

Beames, B. and Summers, M.D. 1990. Sequence comparison of cellular and viral copies of host cell DNA insertions found in *Autographa californica* nuclear polyhedrosis virus. *Virology* 174: 354–363.

Belyavskyi, M., Braunagel, S.C. and Summers, M.D. 1998. The structural protein ODV-EC27 of *Autographa californica* nucleopolyhedrovirus is a multifunctional viral cyclin. *Proceedings of the National Academy of Sciences of the United States of America* 95: 11205–11210.

Beperet, I., Irons, S.L., Simón, O., King, L.A., Williams, T., Possee, R.D., López-Ferber, M. and Caballero, P. 2014. Superinfection exclusion in alphabaculovirus infections is concomitant with actin reorganization. *Journal of Virology* 88: 3548–3556.

Bergua, M., Zwart, M.P., El-Mohtar, C., Shilts, T., Elena, S.F. and Folimonova, S.Y. 2014. A viral protein mediates superinfection exclusion at the whole-organism level but is not required for exclusion at the cellular level. *Journal of Virology* 88: 11327–11338.

Bio-Rad 2006. Real-Time PCR Applications Guide.

Blissard, G.W. 1996. Baculovirus-insect cell interactions. *Cytotechnology* 20: 73–93.

Blissard, G.W. and Rohrmann, G.F. 1990. Baculovirus diversity and molecular biology. *Annual Review of Entomology* 35: 127–155.

Blissard, G.W. and Wenz, J.R. 1992. Baculovirus gp64 envelope glycoprotein is sufficient to mediate pH-dependent membrane fusion. *Journal of Virology* 66: 6829–6835.

Boyce, F.M. and Bucher, N.L. 1996. Baculovirus-mediated gene transfer into mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America* 93: 2348–2352.

Braunagel, S.C., He, H., Ramamurthy, P. and Summers, M.D. 1996. Transcription, translation, and cellular localization of three *Autographa californica* nuclear polyhedrosis virus structural proteins: ODV-E18, ODV-E35, and ODV-EC27. *Virology* 222: 100–114.

Braunagel, S.C., Parr, R., Belyavskyi, M. and Summers, M.D. 1998. *Autographa californica* nucleopolyhedrovirus infection results in Sf9 cell cycle arrest at G2/M Phase. *Virology* 244: 195–211.

Braunagel, S.C., Russell, W.K., Rosas-Acosta, G., Russell, D.H. and Summers, M.D. 2003. Determination of the protein composition of the occlusion-derived virus of *Autographa californica* nucleopolyhedrovirus. *Proceedings of the National Academy of Sciences of the United States of America* 100: 9797–9802.

Braunagel, S.C., Burks, J.K., Rosas-Acosta, G., Harrison, R.L., Ma, H. and Summers, M.D. 1999. Mutations within the *Autographa californica* nucleopolyhedrovirus *FP25K* gene decrease the accumulation of ODV-E66 and alter its intranuclear transport. *Journal of Virology* 73: 8559–8570.

Brown, M. and Faulkner, P. 1978. Plaque assay of nuclear polyhedrosis viruses in cell culture. *Applied and Environmental Microbiology* 36: 31–35.

Burden, J.P., Possee, R.D., Sait, S.M., King, L.A. and Hails, R.S. 2006. Phenotypic and genotypic characterisation of persistent baculovirus infections in populations of the cabbage moth (Mamestra brassicae) within the British Isles. *Archives of Virology* 151: 635–649.

Burden, J.P., Nixon, C.P., Hodgkinson, A.E., Possee, R.D., Sait, S.M., King, L.A. and Hails, R.S. 2003. Covert infections as a mechanism for long-term persistence of baculoviruses. *Ecology Letters* 6: 524–531.

Cabodevilla, O., Ibañez, I., Simón, O., Murillo, R., Caballero, P. and Williams, T. 2011. Occlusion body pathogenicity, virulence and productivity traits vary with transmission strategy in a nucleopolyhedrovirus. *Biological Control* 56: 184–192.

Carson, D.D., Guarino, L.A. and Summers, M.D. 1988. Functional mapping of an AcNPV immediately early gene which augments expression of the IE-1 trans-activated 39K gene. *Virology* 162: 444–451.

Carson, D.D., Summers, M.D. and Guarino, L.A. 1991. Molecular analysis of a baculovirus regulatory gene. *Virology* 182: 279–286.

Cary, L.C., Goebel, M., Corsaro, B.G., Wang, H.-G., Rosen, E. and Fraser, M.J. 1989. Transposon mutagenesis of baculoviruses: analysis of *Trichoplusia ni* transposon IFP2 insertions within the FP-locus of nuclear polyhedrosis viruses. *Virology* 172: 156–169.

Chambers, A. 2012. The role of ODV structural proteins in baculovirus replication. PhD. Thesis, Oxford Brookes University.

Chao, Y.C., Lee, S.T., Chang, M.C., Chen, H.H., Chen, S.S., Wu, T.Y., Liu, F.H., Hsu, E.L. and Hou, R.F. 1998. A 2.9-kilobase noncoding nuclear RNA functions in the establishment of persistent Hz-1 viral infection. *Journal of Virology* 72: 2233–2245.

Charlton, C.A. and Volkman, L.E. 1991. Sequential rearrangement and nuclear polymerization of actin in baculovirus-infected *Spodoptera frugiperda* cells. *Journal of Virology* 65: 1219–1227.

Charlton, C.A. and Volkman, L.E. 1993. Penetration of *Autographa californica* nuclear polyhedrosis virus nucleocapsids into IPLB Sf 21 cells induces actin cable formation. *Virology* 197: 245–254.

Cheever, M.A. and Higano, C.S. 2011. PROVENGE (Sipuleucel-T) in prostate cancer: the first FDA-approved therapeutic cancer vaccine. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research* 17: 3520–3526.

Chen, L., Hu, X., Xiang, X., Yu, S., Yang, R. and Wu, X. 2012. *Autographa californica* multiple nucleopolyhedrovirus *odv-e25* (*ac94*) is required for budded virus infectivity and occlusion-derived virus formation. *Archives of Virology* 157: 617–625.

Chen, Y.-R., Zhong, S., Fei, Z., Hashimoto, Y., Xiang, J.Z., Zhang, S. and Blissard, G.W. 2013. The transcriptome of the baculovirus *Autographa californica* multiple nucleopolyhedrovirus in *Trichoplusia ni* cells. *Journal of Virology* 87: 6391–6405.

Chen, Y.-R., Zhong, S., Fei, Z., Gao, S., Zhang, S., Li, Z., Wang, P. and Blissard, G.W. 2014. Transcriptome responses of the host *Trichoplusia ni* to infection by the baculovirus *Autographa californica* multiple nucleopolyhedrovirus. *Journal of Virology* 88: 13781–13797.

Cheng, J.P.X. and Nichols, B.J. 2016. Caveolae: one function or many? *Trends in Cell Biology* 26: 177–189.

Cheng, T., Zhao, P., Liu, C., Xu, P., Gao, Z., Xia, Q. and Xiang, Z. 2006. Structures, regulatory regions, and inductive expression patterns of antimicrobial peptide genes in the silkworm Bombyx mori. *Genomics* 87: 356–365.

Chernomordik, L., Leikina, E., Cho, M.S. and Zimmerberg, J. 1995. Control of baculovirus gp64-induced syncytium formation by membrane lipid composition. *Journal of Virology* 69: 3049–3058.

Chisholm, G.E. and Henner, D.J. 1988. Multiple early transcripts and splicing of the *Autographa californica* nuclear polyhedrosis virus IE-1 gene. *Journal of Virology* 62: 3193–3200.

Choi, J. and Guarino, L.A. 1995. The baculovirus transactivator IE1 binds to viral enhancer elements in the absence of insect cell factors. *Journal of Virology* 69: 4548–4551.

Clem, R., Fechheimer, M. and Miller, L. 1991. Prevention of apoptosis by a baculovirus gene during infection of insect cells. *Science* 254: 1388–1390.

Clem, R.J. 1997. Regulation of Programmed Cell Death by Baculoviruses. In: Miller, L.K. (Ed.), *The Baculoviruses,* Springer US, Boston, MA, pp. 237–266.

Clem, R.J. 2015. Viral IAPs, then and now. *Seminars in Cell & Developmental Biology* 39: 72–79.

Cooper, D., Cory, J.S., Theilmann, D.A. and Myers, J.H. 2003. Nucleopolyhedroviruses of forest and western tent caterpillars: cross-infectivity and evidence for activation of latent virus in high-density field populations. *Ecological Entomology* 28: 41–50.

Cory, J.S. and Bishop, D.H.L. 1997. Use of baculoviruses as biological insecticides. *Molecular Biotechnology* 7: 303–313.

Cox, M.M.J., Patriarca, P.A. and Treanor, J. 2008. FluBlok, a recombinant hemagglutinin influenza vaccine. *Influenza and Other Respiratory Viruses* 2: 211–219.

Crawford, A.M. and Sheehan, C. 1983. Persistent baculovirus infections: *Spodoptera frugiperda* NPV and *Autographa californica* NPV in *Spodoptera frugiperda* cells. *Archives of Virology* 78: 65–79.

Dai, X., Stewart, T.M., Pathakamuri, J.A., Li, Q. and Theilmann, D.A. 2004. *Autographa californica* multiple nucleopolyhedrovirus *exon0* (*orf141*), which encodes a RING finger protein, is required for efficient production of budded virus. *Journal of Virology* 78: 9633–9644.

Danquah, J.O., Botchway, S., Jeshtadi, A. and King, L.A. 2012. Direct interaction of baculovirus capsid proteins VP39 and EXON0 with kinesin-1 in insect cells determined by fluorescence resonance energy transfer-fluorescence lifetime imaging microscopy. *Journal of Virology* 86: 844–853.

DePristo, M.A., Banks, E., Poplin, R., Garimella, K.V., Maguire, J.R., Hartl, C., Philippakis, A.A., del Angel, G., Rivas, M.A., Hanna, M., McKenna, A., Fennell, T.J., Kernytsky, A.M., Sivachenko, A.Y., Cibulskis, K., Gabriel, S.B., Altshuler, D. and Daly, M.J. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics* 43: 491–498.

Dong, S., Wang, M., Qiu, Z., Deng, F., Vlak, J.M., Hu, Z. and Wang, H. 2010. *Autographa californica* multicapsid nucleopolyhedrovirus efficiently infects Sf9 cells and transduces mammalian cells via direct fusion with the plasma membrane at low pH. *Journal of Virology* 84: 5351–5359.

Dreschers, S., Roncarati, R. and Knebel-Mörsdorf, D. 2001. Actin rearrangement-inducing factor of baculoviruses is tyrosine phosphorylated and colocalizes to F-actin at the plasma membrane. *Journal of Virology* 75: 3771–3778.

Drezen, J.-M., Josse, T., Bézier, A., Gauthier, J., Huguet, E. and Herniou, E.A. 2017. Impact of lateral transfers on the genomes of lepidoptera. *Genes* 8.

Duncan, R., Chung, K.L. and Faulkner, P. 1983. Analysis of a mutant of *Autographa california* nuclear polyhedrosis virus with a defect in the morphogenesis of the occlusion body macromolecular lattice. *The Journal of General Virology* 64 (Pt 7): 1531–1542.

Ebert, D. 2013. The epidemiology and evolution of symbionts with mixed-mode transmission. *Annual Review of Ecology, Evolution, and Systematics* 44: 623–643.

Eddy, S.R. 1998. Profile hidden Markov models. *Bioinformatics* 14: 755–763.

Fang, M., Dai, X. and Theilmann, D.A. 2007. *Autographa californica* multiple nucleopolyhedrovirus EXON0 (ORF141) is required for efficient egress of nucleocapsids from the nucleus. *Journal of Virology* 81: 9859–9869.

Fang, M., Nie, Y. and Theilmann, D.A. 2009a. AcMNPV EXON0 (AC141) which is required for the efficient egress of budded virus nucleocapsids interacts with β-tubulin. *Virology* 385: 496–504.

Fang, M., Nie, Y., Harris, S., Erlandson, M.A. and Theilmann, D.A. 2009b. *Autographa californica* multiple nucleopolyhedrovirus core gene ac96 encodes a *per os* infectivity factor (PIF-4). *Journal of Virology* 83: 12569–12578.

Federici, B.A. and Hice, R.H. 1997. Organization and molecular characterization of genes in the polyhedrin region of the Anagrapha falcifera multinucleocapsid NPV. *Archives of Virology* 142: 333–348.

Finn, R.D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Heger, A., Hetherington, K., Holm, L., Mistry, J., Sonnhammer, E.L.L., Tate, J. and Punta, M. 2014. Pfam: the protein families database. *Nucleic Acids Research* 42: D222–D230.

Fleckner, J., Zhang, M., Valcárcel, J. and Green, M.R. 1997. U2AF65 recruits a novel human DEAD box protein required for the U2 snRNP-branchpoint interaction. *Genes & Development* 11: 1864–1872.

Folimonova, S.Y. 2012. Superinfection exclusion is an active virus-controlled function that requires a specific viral protein. *Journal of Virology* 86: 5554–5561.

Fraser, M.J., Smith, G.E. and Summers, M.D. 1983. Acquisition of host cell DNA sequences by baculoviruses: relationship between host DNA insertions and FP mutants of *Autographa californica* and *Galleria mellonella* nuclear polyhedrosis viruses. *Journal of Virology* 47: 287–300.

Friesen, P.D. and Miller, L.K. 1985. Temporal regulation of baculovirus RNA: overlapping early and late transcripts. *Journal of Virology* 54: 392–400.

Gandhi, K.M., Ohkawa, T., Welch, M.D. and Volkman, L.E. 2012. Nuclear localization of actin requires AC102 in *Autographa californica* multiple nucleopolyhedrovirus-infected cells. *Journal of General Virology* 93: 1795–1803.

Gauthier, D., Thirunavukkarasu, K., Faris, B.L., Russell, D.L. and Weaver, R.F. 2012. Characterization of an *Autographa californica* multiple nucleopolyhedrovirus dual mutant: ORF82 is required for budded virus production, and a point mutation in LEF-8 alters late and abolishes very late transcription. *The Journal of General Virology* 93: 364–373.

Gilbert, C., Chateigner, A., Ernenwein, L., Barbe, V., Bézier, A., Herniou, E.A. and Cordaux, R. 2014. Population genomics supports baculoviruses as vectors of horizontal transfer of insect transposons. *Nature Communications* 5.

Giri, L., Li, H., Sandgren, D., Feiss, M.G., Roller, R., Bonning, B.C. and Murhammer, D.W. 2010. Removal of transposon target sites from the *Autographa californica* multiple nucleopolyhedrovirus *fp25k* gene delays, but does not prevent, accumulation of the few polyhedra phenotype. *The Journal of General Virology* 91: 3053–3064.

Goic, B. and Saleh, M.C. 2012. Living with the enemy: viral persistent infections from a friendly viewpoint. *Current Opinion in Microbiology* 15: 531–537.

Goic, B., Vodovar, N., Mondotte, J.A., Monot, C., Frangeul, L., Blanc, H., Gausson, V., Vera-Otarola, J., Cristofari, G. and Saleh, M.C. 2013. RNA-mediated interference and reverse transcription control the persistence of RNA viruses in the insect model Drosophila. *Nature Immunology* 14: 396–403.

Goley, E.D., Ohkawa, T., Mancuso, J., Woodruff, J.B., D'Alessio, J.A., Cande, W.Z., Volkman, L.E. and Welch, M.D. 2006. Dynamic nuclear actin assembly by Arp2/3 complex and a baculovirus WASP-like protein. *Science (New York, N.Y.)* 314: 464–467.

Goulson, D. 1997. Wipfelkrankheit: modification of host behaviour during baculoviral infection. *Oecologia* 109: 219–228.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N. and Regev, A. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* 29: 644–652.

Granados, R.R. and Lawler, K.A. 1981. *In vivo* pathway of *Autographa californica* baculovirus invasion and infection. *Virology* 108: 297–308.

Granados, R.R., Guoxun, L., Derksen, A.C.G. and McKenna, K.A. 1994. A new insect cell line from *Trichoplusia ni* (BTI-Tn-5B1-4) susceptible to *Trichoplusia ni* single enveloped nuclear polyhedrosis virus. *Journal of Invertebrate Pathology* 64: 260–266.

Graves, L. 2016. Taking a closer look: exploring the functional roles of P10 in baculovirus-infected cells. PhD. Thesis, Oxford Brookes University.

Green, M.R., Sambrook, J. and Sambrook, J. 2012. *Molecular Cloning: A Laboratory Manual*, 4th ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y, .

GSL Biotech 2017. *SnapGene Software*. GSL Biotech.

Guan, Z., Zhong, L., Li, C., Wu, W., Yuan, M. and Yang, K. 2016. The Autographa californica multiple nucleopolyhedrovirus *ac54* gene is crucial for localization of the major capsid protein VP39 at the site of nucleocapsid assembly. *Journal of Virology* 90: 4115–4126.

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

Guarino, L.A., Smith, G. and Dong, W. 1995. Ubiquitin is attached to membranes of baculovirus particles by a novel type of phospholipid anchor. *Cell* 80: 301–309.

Guarino, L.A., Xu, B., Jin, J. and Dong, W. 1998. A virus-encoded RNA polymerase purified from baculovirus-infected cells. *Journal of Virology* 72: 7985–7991.

Guo, Y., Yue, Q., Gao, J., Wang, Z., Chen, Y.-R., Blissard, G.W., Liu, T.-X. and Li, Z. 2017a. Roles of cellular NSF protein in entry and nuclear egress of budded virions of Autographa californica multiple nucleopolyhedrovirus. *Journal of Virology* 91.

Guo, Y.-J., Fu, S.-H. and Li, L.-L. 2017b. *Autographa californica* multiple nucleopolyhedrovirus *ac75* is required for egress of nucleocapsids from the nucleus and formation of *de novo* intranuclear membrane microvesicles. *PLOS ONE* 12: e0185630.

Haas, A.L., Katzung, D.J., Reback, P.M. and Guarino, L.A. 1996. Functional characterization of the ubiquitin variant encoded by the baculovirus Autographa californica. *Biochemistry* 35: 5385–5394.

Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M., MacManes, M.D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C.N., Henschel, R., LeDuc, R.D., Friedman, N. and Regev, A. 2013. *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols* 8: 1494–1512.

Haase, S., Sciocco-Cap, A. and Romanowski, V. 2015. Baculovirus insecticides in Latin America: historical overview, current status and future perspectives. *Viruses* 7: 2230–2267.

Habib, S. and Hasnain, S.E. 2000. Differential activity of two non-*hr* origins during replication of the baculovirus *Autographa californica* nuclear polyhedrosis virus genome. *Journal of Virology* 74: 5182–5189.

Han, J., Pluhackova, K. and Böckmann, R.A. 2017. The multifaceted role of SNARE proteins in membrane fusion. *Frontiers in Physiology* 8.

Han, Y., van Houte, S., van Oers, M. and Ros, V. 2018a. Baculovirus PTP2 functions as a pro-apoptotic protein. *Viruses* 10: 181.

Han, Y., van Houte, S., van Oers, M.M. and Ros, V.I.D. 2018b. Timely trigger of caterpillar zombie behaviour: temporal requirements for light in baculovirus-induced tree-top disease. *Parasitology* 145: 822–827.

Han, Y., van Houte, S., Drees, G., van Oers, M. and Ros, V. 2015. Parasitic manipulation of host behaviour: baculovirus SeMNPV EGT facilitates tree-top disease in *Spodoptera exigua* larvae by extending the time to death. *Insects* 6: 716–731.

Harrison, R.L. and Bonning, B.C. 2003. Comparative analysis of the genomes of *Rachiplusia ou* and *Autographa californica* multiple nucleopolyhedroviruses. *The Journal of General Virology* 84: 1827–1842.

Hashimoto, Y., Macri, D., Srivastava, I., McPherson, C., Felberbaum, R., Post, P. and Cox, M. 2017. Complete study demonstrating the absence of rhabdovirus in a distinct Sf9 cell line. *PloS One* 12: e0175633.

Hawtin, R.E., Arnold, K., Ayres, M.D., Zanotto, P.M., Howard, S.C., Gooday, G.W., Chappell, L.H., Kitts, P.A., King, L.A. and Possee, R.D. 1995. Identification and preliminary characterization of a chitinase gene in the *Autographa californica* nuclear polyhedrosis virus genome. *Virology* 212: 673–685.

Heath, C.G., Viphakone, N. and Wilson, S.A. 2016. The role of TREX in gene expression and disease. *Biochemical Journal* 473: 2911–2935.

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* 48: 211–234.

Hill, T. and Unckless, R.L. 2017. Baculovirus molecular evolution via gene turnover and recurrent positive selection of key genes. *Journal of Virology* 91.

Hink, W.F. 1970. Established insect cell line from the cabbage looper, *Trichoplusia ni*. *Nature* 226: 466–467.

Hink, W.F. and Vail, P.V. 1973. A plaque assay for titration of alfalfa looper nuclear polyhedrosis virus in a cabbage looper (TN-368) cell line. *Journal of Invertebrate Pathology* 22: 168–174.

Hitchman, R.B., Possee, R.D. and King, L.A. 2009. Baculovirus expression systems for recombinant protein production in insect cells. *Recent Patents on Biotechnology* 3: 46–54.

Hitchman, R.B., Siaterli, E.A., Nixon, C.P. and King, L.A. 2007. Quantitative real-time PCR for rapid and accurate titration of recombinant baculovirus particles. *Biotechnology and Bioengineering* 96: 810–814.

Hofmann, C., Sandig, V., Jennings, G., Rudolph, M., Schlag, P. and Strauss, M. 1995. Efficient gene transfer into human hepatocytes by baculovirus vectors. *Proceedings of the National Academy of Sciences of the United States of America* 92: 10099–10103.

Hohmann, A.W. and Faulkner, P. 1983. Monoclonal antibodies to baculovirus structural proteins: determination of specificities by western blot analysis. *Virology* 125: 432–444.

Hoover, K., Grove, M., Gardner, M., Hughes, D.P., McNeil, J. and Slavicek, J. 2011. A gene for an extended phenotype. *Science (New York, N.Y.)* 333: 1401.

Horton, H.M. and Burand, J.P. 1993. Saturable attachment sites for polyhedron-derived baculovirus on insect cells and evidence for entry via direct membrane fusion. *Journal of Virology* 67: 1860–1868.

van Houte, S., Ros, V.I.D. and van Oers, M.M. 2013. Walking with insects: molecular mechanisms behind parasitic manipulation of host behaviour. *Molecular Ecology* 22: 3458–3475.

van Houte, S., Ros, V.I.D. and van Oers, M.M. 2014. Hyperactivity and tree-top disease induced by the baculovirus AcMNPV in *Spodoptera exigua* larvae are governed by independent mechanisms. *Die Naturwissenschaften* 101: 347–350.

van Houte, S., Ros, V.I.D., Mastenbroek, T.G., Vendrig, N.J., Hoover, K., Spitzen, J. and van Oers, M.M. 2012. Protein tyrosine phosphatase-induced hyperactivity is a conserved strategy of a subset of baculoviruses to manipulate lepidopteran host behavior. *PloS One* 7: e46933.

Hu, Z., Yuan, M., Wu, W., Liu, C., Yang, K. and Pang, Y. 2010. *Autographa californica* multiple nucleopolyhedrovirus *ac76* is involved in intranuclear microvesicle formation. *Journal of Virology* 84: 7437–7447.

Huff, J. 2015. The Airyscan detector from ZEISS: confocal imaging with improved signal-to-noise ratio and super-resolution. *Nature Methods* 12.

Hughes, D.S., Possee, R.D. and King, L.A. 1993. Activation and detection of a latent baculovirus resembling Mamestra brassicae nuclear polyhedrosis virus in M. brassicae insects. *Virology* 194: 608–615.

Hughes, D.S., Possee, R.D. and King, L.A. 1997. Evidence for the presence of a low-level, persistent baculovirus infection of Mamestra brassicae insects. *The Journal of General Virology* 78 (Pt 7): 1801–1805.

Huijskens, I., Li, L., Willis, L.G. and Theilmann, D.A. 2004. Role of AcMNPV IE0 in baculovirus very late gene activation. *Virology* 323: 120–130.

Hurley, J.H. 2015. ESCRTs are everywhere. *The EMBO Journal* 34: 2398–2407.

Hussain, M. and Asgari, S. 2014. MicroRNAs as mediators of insect host–pathogen interactions and immunity. *Journal of Insect Physiology* 70: 151–158.

IBM Corp. 2017. *IBM SPSS Statistics for Windows*. Armonk, NY: IBM Corp.

Ikeda, M., Yamada, H., Hamajima, R. and Kobayashi, M. 2013. Baculovirus genes modulating intracellular innate antiviral immunity of lepidopteran insect cells. *Virology* 435: 1–13.

Integrated DNA Techologies 2018. OligoAnalyser 3.1. Available at https://www.idtdna.com/calc/analyzer.

Invitrogen 2008. Real Time PCR: from theory to practice.

Jakubowska, A.K., D'Angiolo, M., González-Martínez, R.M., Millán-Leiva, A., Carballo, A., Murillo, R., Caballero, P. and Herrero, S. 2014. Simultaneous occurrence of covert infections with small RNA viruses in the lepidopteran *Spodoptera exigua*. *Journal of Invertebrate Pathology* 121: 56–63.

Jarvis, D.L., Fleming, J.-A.G.W., Kovacs, G.R., Summers, M.D. and Guarino, L.A. 1990. Use of early baculovirus promoters for continuous expression and efficient processing of foreign gene products in stably transformed Lepidopteran cells. *Nature Biotechnology* 8: 950–955.

Javed, M.A., Biswas, S., Willis, L.G., Harris, S., Pritchard, C., van Oers, M.M., Donly, B.C., Erlandson, M.A., Hegedus, D.D. and Theilmann, D.A. 2017. Autographa californica multiple nucleopolyhedrovirus AC83 is a *per os* infectivity factor (PIF) protein required for occlusion-derived virus (ODV) and budded virus nucleocapsid assembly as well as assembly of the PIF complex in ODV envelopes. *Journal of Virology* 91.

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* 151: 1257–1266.

Ji, X., Sutton, G., Evans, G., Axford, D., Owen, R. and Stuart, D.I. 2010. How baculovirus polyhedra fit square pegs into round holes to robustly package viruses. *The EMBO Journal* 29: 505–514.

Jiang, X. and Wang, X. 2004. Cytochrome C-mediated apoptosis. *Annual Review of Biochemistry* 73: 87–106.

de Jong, J., Arif, B.M., Theilmann, D.A. and Krell, P.J. 2009. *Autographa californica* multiple nucleopolyhedrovirus *me53* (*ac140*) is a nonessential gene required for efficient budded-virus production. *Journal of Virology* 83: 7440–7448.

de Jong, J., Theilmann, D.A., Arif, B.M. and Krell, P.J. 2011. Immediate-early protein ME53 forms foci and colocalizes with GP64 and the major capsid protein VP39 at the cell membranes of *Autographa californica* multiple nucleopolyhedrovirus-infected cells. *Journal of Virology* 85: 9696–9707.

Kane, M. and Golovkina, T. 2010. Common threads in persistent viral infections. *Journal of Virology* 84: 4116–4123.

Kantoff, P.W., Higano, C.S., Shore, N.D., Berger, E.R., Small, E.J., Penson, D.F., Redfern, C.H., Ferrari, A.C., Dreicer, R., Sims, R.B., Xu, Y., Frohlich, M.W., Schellhammer, P.F. and IMPACT Study Investigators 2010. Sipuleucel-T immunotherapy for castration-resistant prostate cancer. *The New England Journal of Medicine* 363: 411–422.

Karamipour, N., Fathipour, Y., Talebi, A.A., Asgari, S. and Mehrabadi, M. 2018. Small interfering RNA pathway contributes to antiviral immunity in *Spodoptera frugiperda*

(Sf9) cells following Autographa californica multiple nucleopolyhedrovirus infection. *Insect Biochemistry and Molecular Biology* 101: 24–31.

Karpf, A.R., Lenches, E., Strauss, E.G., Strauss, J.H. and Brown, D.T. 1997. Superinfection exclusion of alphaviruses in three mosquito cell lines persistently infected with Sindbis virus. *Journal of Virology* 71: 7119–7123.

Kataoka, C., Kaname, Y., Taguwa, S., Abe, T., Fukuhara, T., Tani, H., Moriishi, K. and Matsuura, Y. 2012. Baculovirus GP64-mediated entry into mammalian cells. *Journal of Virology* 86: 2610–2620.

Katoh, K. and Standley, D.M. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* 30: 772–780.

Katsuma, S., Kobayashi, J., Koyano, Y., Matsuda-Imai, N., Kang, W. and Shimada, T. 2012. Baculovirus-encoded protein BV/ODV-E26 determines tissue tropism and virulence in Lepidopteran insects. *Journal of Virology* 86: 2545–2555.

Kelly, B.J., King, L.A., Possee, R.D. and Chapple, S.D.J. 2006. Dual mutations in the *Autographa californica* nucleopolyhedrovirus *FP-25* and *p35* genes result in plasma-membrane blebbing in *Trichoplusia ni* cells. *Journal of General Virology* 87: 531–536.

Kelly, B.J., Chapple, S.D.J., Allen, C., Pritchard, C., King, L.A. and Possee, R.D. 2008. Extended budded virus formation and induction of apoptosis by an AcMNPV *FP-25*/*p35* double mutant in *Trichoplusia ni* cells. *Virus Research* 133: 157–166.

Kelly, D.C. and Lescott, T. 1981. Baculovirus replication: Protein synthesis in *Spodoptera frugiperda* cells infected with Trichoplusia in nuclear polyhedrosis virus. *Microbiologica* 4: 35–57.

Kemp, T.J., Safaeian, M., Hildesheim, A., Pan, Y., Penrose, K.J., Porras, C., Schiller, J.T., Lowy, D.R., Herrero, R. and Pinto, L.A. 2012. Kinetic and HPV infection effects on cross-type neutralizing antibody and avidity responses induced by Cervarix(®). *Vaccine* 31: 165–170.

Khan, K.H. 2013. Gene Expression in Mammalian Cells and its Applications. *Advanced Pharmaceutical Bulletin; EISSN 2251-7308*.

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. *Academic Press, London*; Waltham, MA.

King, L.A. and Possee, R.D. 1992. The Baculovirus Expression System: A laboratory guide. Springer Netherlands, Dordrecht, .

King, L.A., Kaur, K., Mann, S.G., Lawrie, A.M., Steven, J. and Ogden, J.E. 1991. Secretion of single-chain urokinase-type plasminogen activator from insect cells. *Gene* 106: 151–157.

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

Kost, T.A. and Kemp, C.W. 2016. Fundamentals of baculovirus expression and applications. In: Vega, M.C. (Ed.), Advanced Technologies for Protein Complex Production and Characterization, *Springer International Publishing*, Cham, vol. 896. pp. 187–197.

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

Kovacs, G.R., Guarino, L.A., Graham, B.L. and Summers, M.D. 1991. Identification of spliced baculovirus RNAs expressed late in infection. *Virology* 185: 633–643.

Krappa, R. and Knebel-Mörsdorf, D. 1991. Identification of the very early transcribed baculovirus gene PE-38. *Journal of Virology* 65: 805–812.

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S.J. and Marra, M.A. 2009. Circos: an information aesthetic for comparative genomics. *Genome Research* 19: 1639–1645.

Kuzio, J., Jaques, R. and Faulkner, P. 1989. Identification of p74, a gene essential for virulence of baculovirus occlusion bodies. *Virology* 173: 759–763.

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

Lanier, L.M. and Volkman, L.E. 1998. Actin binding and nucleation by *Autographa california* M nucleopolyhedrovirus. *Virology* 243: 167–177.

Lawrie, A.M., King, L.A. and Ogden, J.E. 1995. High level synthesis and secretion of human urokinase using a late gene promoter of the *Autographa californica* nuclear polyhedrosis virus. *Journal of Biotechnology* 39: 1–8.

Lecoq, H. 1991. Control of zucchini yellow mosaic virus in squash by cross protection. *Plant Disease* 75: 208.

Lee, J.C., Chen, H.H. and Chao, Y.C. 1998. Persistent baculovirus infection results from deletion of the apoptotic suppressor gene *p35*. *Journal of Virology* 72: 9157–9165.

Leikina, E., Onaran, H.O. and Zimmerberg, J. 1992. Acidic pH induces fusion of cells infected with baculovirus to form syncytia. *FEBS Letters* 304: 221–224.

Leisy, D.J., Rasmussen, C., Owusu, E.O. and Rohrmann, G.F. 1997. A mechanism for negative gene regulation in *Autographa californica* multinucleocapsid nuclear polyhedrosis virus. *Journal of Virology* 71: 5088–5094.

van Lent, J.W., Groenen, J.T., Klinge-Roode, E.C., Rohrmann, G.F., Zuidema, D. and Vlak, J.M. 1990. Localization of the 34 kDa polyhedron envelope protein in *Spodoptera frugiperda* cells infected with Autographa californica nuclear polyhedrosis virus. *Archives of Virology* 111: 103–114.

Li, G., Ma, D. and Chen, Y. 2016. Cellular functions of programmed cell death 5. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* 1863: 572–580.

Li, H. and Durbin, R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics (Oxford, England)* 25: 1754–1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. and 1000 Genome Project Data Processing Subgroup 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics (Oxford, England)* 25: 2078–2079.

Li, T.C., Scotti, P.D., Miyamura, T. and Takeda, N. 2007. Latent infection of a new alphanodavirus in an insect cell line. *Journal of Virology* 81: 10890–10896.

Li, Y. and Miller, L.K. 1995. Expression and functional analysis of a baculovirus gene encoding a truncated protein kinase homolog. *Virology* 206: 314–323.

Liang, C., Li, M., Dai, X., Zhao, S., Hou, Y., Zhang, Y., Lan, D., Wang, Y. and Chen, X. 2013. *Autographa californica* multiple nucleopolyhedrovirus PK-1 is essential for nucleocapsid assembly. *Virology* 443: 349–357.

Lieberman, P.M. 2016. Epigenetics and genetics of viral latency. *Cell Host & Microbe* 19: 619–628.

Liu, C., Li, Z., Wu, W., Li, L., Yuan, M., Pan, L., Yang, K. and Pang, Y. 2008. *Autographa californica multiple nucleopolyhedrovirus ac53* plays a role in nucleocapsid assembly. *Virology* 382: 59–68.

Liu, F., Wu, X., Li, L., Liu, Z. and Wang, Z. 2013. Use of baculovirus expression system for generation of virus-like particles: successes and challenges. *Protein Expression and Purification* 90: 104–116.

Livak, K.J. and Schmittgen, T.D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2−ΔΔCT method. *Methods* 25: 402–408.

Long, G., Pan, X., Kormelink, R. and Vlak, J.M. 2006. Functional entry of baculovirus into insect and mammalian cells is dependent on clathrin-mediated endocytosis. *Journal of Virology* 80: 8830–8833.

van Loo, N.D., Fortunati, E., Ehlert, E., Rabelink, M., Grosveld, F. and Scholte, B.J. 2001. Baculovirus infection of nondividing mammalian cells: mechanisms of entry and nuclear transport of capsids. *Journal of Virology* 75: 961–970.

Luckow, V.A., Lee, S.C., Barry, G.F. and Olins, P.O. 1993. Efficient generation of infectious recombinant baculoviruses by site-specific transposon-mediated insertion of foreign

genes into a baculovirus genome propagated in Escherichia coli. *Journal of Virology* 67: 4566–4579.

Lung, O.Y., Cruz-Alvarez, M. and Blissard, G.W. 2003. Ac23, an envelope fusion protein homolog in the baculovirus *Autographa californica multicapsid* nucleopolyhedrovirus, is a viral pathogenicity factor. *Journal of Virology* 77: 328–339.

Lyupina, Y.V., Dmitrieva, S.B., Timokhova, A.V., Beljelarskaya, S.N., Zatsepina, O.G., Evgen'ev, M.B. and Mikhailov, V.S. 2010. An important role of the heat shock response in infected cells for replication of baculoviruses. *Virology* 406: 336–341.

Ma, H., Galvin, T.A., Glasner, D.R., Shaheduzzaman, S. and Khan, A.S. 2014. Identification of a novel rhabdovirus in *Spodoptera frugiperda* cell lines. *Journal of Virology* 88: 6576–6585.

Marek, M., Merten, O.-W., Galibert, L., Vlak, J.M. and van Oers, M.M. 2011. Baculovirus VP80 protein and the F-actin cytoskeleton interact and connect the viral replication factory with the nuclear periphery. *Journal of Virology* 85: 5350–5362.

McIntosh, A.H. and Ignoffo, C.M. 1981. Establishment of a persistent baculovirus infection in a lepidopteran cell line. *Journal of Invertebrate Pathology* 38: 395–403.

Mehrabadi, M., Hussain, M., Matindoost, L. and Asgari, S. 2015. The baculovirus antiapoptotic p35 protein functions as an inhibitor of the host RNA interference antiviral response. *Journal of Virology* 89: 8182–8192.

Meki, I.K., Kariithi, H.M., Parker, A.G., Vreysen, M.J.B., Ros, V.I.D., Vlak, J.M., van Oers, M.M. and Abd-Alla, A.M.M. 2018. RNA interference-based antiviral immune response against the salivary gland hypertrophy virus in Glossina pallidipes. *BMC Microbiology* 18: 170.

Merten, O.W. 2007. Attention with virus contaminated cell lines. *Cytotechnology* 55: 1–2.

Miyamura, T., Aizaki, H., Tani, H., Shoji, I., Matsuura, Y., Ishii, K., Chiba, T. and Saito, I. 1997. Efficient gene transfer into various mammalian cells, including non-hepatic cells, by baculovirus vectors. *Journal of General Virology* 78: 2657–2664.

Monsma, S.A., Oomens, A.G. and Blissard, G.W. 1996. The GP64 envelope fusion protein is an essential baculovirus protein required for cell-to-cell transmission of infection. *Journal of Virology* 70: 4607–4616.

Morpheus Software. Available at https://software.broadinstitute.org/morpheus.

Morris, T.D., Todd, J.W., Fisher, B. and Miller, L.K. 1994. Identification of lef-7: a baculovirus gene affecting late gene expression. *Virology* 200: 360–369.

Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L. and Wold, B. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods* 5: 621–628.

Moscardi, F. 1999. Assessment of the application of baculoviruses for control of Lepidoptera. *Annual Review of Entomology* 44: 257–289.

van Munster, M., Willis, L.G., Elias, M., Erlandson, M.A., Brousseau, R., Theilmann, D.A. and Masson, L. 2006. Analysis of the temporal expression of *Trichoplusia ni* single nucleopolyhedrovirus genes following transfection of BT1-Tn-5B1-4 cells. *Virology* 354: 154–166.

Murillo, R., Hussey, M.S. and Possee, R.D. 2011. Evidence for covert baculovirus infections in a *Spodoptera exigua* laboratory culture. *The Journal of General Virology* 92: 1061–1070.

Myers, J.H. and Cory, J.S. 2016. Ecology and evolution of pathogens in natural populations of Lepidoptera. *Evolutionary Applications* 9: 231–247.

Narayanan, K. 2004. Insect defence: its impact on microbial control of insect pests. *Current Science* 86.

Nasar, F., Erasmus, J.H., Haddow, A.D., Tesh, R.B. and Weaver, S.C. 2015. Eilat virus induces both homologous and heterologous interference. *Virology* 484: 51–58.

NCBI 1982. GeneBank. *Natl. Libr. Med. US Natl. Cent. Biotechnol. Inf.* Available at https://www.ncbi.nlm.nih.gov/nucleotide/

NEB 2018. Tm Calculator v1.9.13. Available at https://tmcalculator.neb.com/#!/main.

Nebert, D.W. and Vasiliou, V. 2004. Analysis of the glutathione S-transferase (GST) gene family. *Human Genomics* 1: 460–464.

Nie, Y., Fang, M. and Theilmann, D.A. 2009. AcMNPV AC16 (DA26, BV/ODV-E26) regulates the levels of IE0 and IE1 and binds to both proteins via a domain located within the acidic transcriptional activation domain. *Virology* 385: 484–495.

Nobiron, I., O'Reilly, D.R. and Olszewski, J.A. 2003. *Autographa californica* nucleopolyhedrovirus infection of *Spodoptera frugiperda* cells: a global analysis of host gene regulation during infection, using a differential display approach. *The Journal of General Virology* 84: 3029–3039.

van Oers, M.M., Pijlman, G.P. and Vlak, J.M. 2015. Thirty years of baculovirus-insect cell protein expression: from dark horse to mainstream technology. *Journal of General Virology* 96: 6–23.

Ohkawa, T. and Volkman, L.E. 1999. Nuclear F-Actin is required for AcMNPV nucleocapsid morphogenesis. *Virology* 264: 1–4.

Ohkawa, T. and Welch, M.D. 2018. Baculovirus actin-based motility drives nuclear envelope disruption and nuclear egress. *Current Biology* 28: 2153-2159.e4.

Ohkawa, T., Rowe, A.R. and Volkman, L.E. 2002. Identification of six *Autographa californica* multicapsid nucleopolyhedrovirus early genes that mediate nuclear localization of G-actin. *Journal of Virology* 76: 12281–12289.

Ohkawa, T., Volkman, L.E. and Welch, M.D. 2010. Actin-based motility drives baculovirus transit to the nucleus and cell surface. *The Journal of Cell Biology* 190: 187–195.

Oldstone, M.B.A. 2006. Viral persistence: parameters, mechanisms and future predictions. *Virology.* 344(1): 111-118.

Ono, C., Okamoto, T., Abe, T. and Matsuura, Y. 2018. Baculovirus as a tool for gene delivery and gene therapy. *Viruses* 10: 510.

Oomens, A.G.P. and Blissard, G.W. 1999. Requirement for GP64 to drive efficient budding of *Autographa californica* multicapsid nucleopolyhedrovirus. *Virology* 254: 297–314.

O'Reilly, D.R. and Miller, L.K. 1989. A baculovirus blocks insect molting by producing ecdysteroid UDP-glucosyl transferase. *Science (New York, N.Y.)* 245: 1110–1112.

O'Reilly, D.R. and Miller, L.K. 1990. Regulation of expression of a baculovirus ecdysteroid UDPglucosyltransferase gene. *Journal of Virology* 64: 1321–1328.

O'Reilly, D.R., Passarelli, A.L., Goldman, I.F. and Miller, L.K. 1990. Characterization of the DA26 gene in a hypervariable region of the Autographa californica nuclear polyhedrosis virus genome. *The Journal of General Virology* 71 (Pt 5): 1029–1037.

Pearson, M., Bjornson, R., Pearson, G. and Rohrmann, G. 1992. The Autographa californica baculovirus genome: evidence for multiple replication origins. *Science (New York, N.Y.)* 257: 1382–1384.

Peña, A., Gewartowski, K., Mroczek, S., Cuéllar, J., Szykowska, A., Prokop, A., Czarnocki-Cieciura, M., Piwowarski, J., Tous, C., Aguilera, A., Carrascosa, J.L., Valpuesta, J.M. and Dziembowski, A. 2012. Architecture and nucleic acids recognition mechanism of the THO complex, an mRNP assembly factor. *The EMBO Journal* 31: 1605–1616.

Peng, Y., Song, J., Lu, J. and Chen, X. 2007. The histone deacetylase inhibitor sodium butyrate inhibits baculovirus-mediated transgene expression in Sf9 cells. *Journal of Biotechnology* 131: 180–187.

Peng, Y., Li, K., Pei, R., Wu, C., Liang, C., Wang, Y. and Chen, X. 2012. The protamine-like DNA-binding protein P6.9 epigenetically up-regulates *Autographa californica* multiple nucleopolyhedrovirus gene transcription in the late infection phase. *Virologica Sinica* 27: 57–68.

Pollard, T.D. 2007. Regulation of actin filament assembly by Arp2/3 complex and formins. *Annual Review of Biophysics and Biomolecular Structure* 36: 451–477.

Possee, R.D. 1986. Cell-surface expression of influenza virus haemagglutinin in insect cells using a baculovirus vector. *Virus Research* 5: 43–59.

Possee, R.D. and Howard, S.C. 1987. Analysis of the polyhedrin gene promoter of the *Autographa californica* nuclear polyhedrosis virus. *Nucleic Acids Research* 15: 10233–10248.

Possee, R.D., Sun, T.P., Howard, S.C., Ayres, M.D., Hill-Perkins, M. and Gearing, K.L. 1991. Nucleotide sequence of the *Autographa californica* nuclear polyhedrosis 9.4 kbp EcoRI-I and -R (polyhedrin gene) region. *Virology* 185: 229–241.

Possee, R.D., Griffiths, C.M., Hitchman, R.B., Chambers, A., Murguia-Meca, F., Danquah, J., Jeshtadi, A. and King, L. 2010. Baculoviruses: biology, replication and exploitation. *Insect Virology,* Caister Academic Press, .

Potter, K.N., Faulkner, P. and MacKinnon, E.A. 1976. Strain selection during serial passage of *Trichoplusia* in nuclear polyhedrosis virus. *Journal of Virology* 18: 1040–1050.

Prikhod'ko, E.A. and Miller, L.K. 1999. The baculovirus PE38 protein augments apoptosis induced by transactivator IE1. *Journal of Virology* 73: 6691–6699.

Promega 2018. BioMath Calculators. BioMath Calc. - Ligations Molar Ratio Insert Vector Calc. Available at https://www.promega.co.uk/resources/tools/biomath/

Qiu, J., Tang, Z., Cai, Y., Wu, W., Yuan, M. and Yang, K. 2018. The Autographa californica multiple nucleopolyhedrovirus *ac51* gene is required for efficient nuclear egress of nucleocapsids and is essential for *in vivo* virulence. *Journal of Virology*.

Quinlan, A.R. and Hall, I.M. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics (Oxford, England)* 26: 841–842.

R Core Team 2014. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Radu, A. 1993. Nup155 is a novel nuclear pore complex protein that contains neither repetitive sequence motifs nor reacts with WGA. *The Journal of Cell Biology* 121: 1–9.

Rayala, H.J., Kendirgi, F., Barry, D.M., Majerus, P.W. and Wente, S.R. 2004. The mRNA export factor human Gle1 interacts with the nuclear pore complex protein Nup155. *Molecular & Cellular Proteomics: MCP* 3: 145–155.

Rodems, S.M., Pullen, S.S. and Friesen, P.D. 1997. DNA-dependent transregulation by IE1 of *Autographa californica* nuclear polyhedrosis virus: IE1 domains required for transactivation and DNA binding. *Journal of Virology* 71: 9270–9277.

Rohrmann, G.F. 2013. Baculovirus Molecular Biology, 3rd ed. National Center for Biotechnology Information (US), Bethesda (MD), .

Ros, V.I.D., van Houte, S., Hemerik, L. and van Oers, M.M. 2015. Baculovirus-induced tree-top disease: how extended is the role of *egt* as a gene for the extended phenotype? *Molecular Ecology* 24: 249–258.

Sashital, D.G. and Doudna, J.A. 2010. Structural insights into RNA interference. *Current Opinion in Structural Biology* 20: 90–97.

Schiller, J.T. and Lowy, D.R. 2001. Papillomavirus-like particle vaccines. *Journal of the National Cancer Institute. Monographs* 50–54.

Schultz, K.L.W. and Friesen, P.D. 2009. Baculovirus DNA replication-specific expression factors trigger apoptosis and shutoff of host protein synthesis during infection. *Journal of Virology* 83: 11123–11132.

Sheehan, D., Meade, G., Foley, V.M. and Dowd, C.A. 2001. Structure, function and evolution of glutathione transferases: implications for classification of non-mammalian members of an ancient enzyme superfamily. *The Biochemical Journal* 360: 1–16.

Shi, A., Hu, Z., Zuo, Y., Wang, Y., Wu, W., Yuan, M. and Yang, K. 2017. Autographa californica nucleopolyhedrovirus *ac75* is required for the nuclear egress of nucleocapsids and intranuclear microvesicle formation. *Journal of Virology* JVI.01509-17.

Shi, X., Ran, Z., Li, S., Yin, J. and Zhong, J. 2016. The effect of microRNA bantam on baculovirus AcMNPV infection *in vitro* and *in vivo*. *Viruses* 8: 136.

Singh, C.P., Singh, J. and Nagaraju, J. 2014. bmnpv-miR-3 facilitates BmNPV infection by modulating the expression of viral P6.9 and other late genes in *Bombyx mori*. *Insect Biochemistry and Molecular Biology* 49: 59–69.

Singh, I.R., Suomalainen, M., Varadarajan, S., Garoff, H. and Helenius, A. 1997. Multiple mechanisms for the inhibition of entry and uncoating of superinfecting Semliki Forest virus. *Virology* 231: 59–71.

Slack, J. and Arif, B.M. 2007. The baculoviruses occlusion-derived virus: virion structure and function. *Advances in Virus Research* 69: 99–165.

Smith, G.E., Summers, M.D. and Fraser, M.J. 1983. Production of human beta interferon in insect cells infected with a baculovirus expression vector. *Molecular and Cellular Biology* 3: 2156–2165.

Speck, S.H. and Ganem, D. 2010. Viral latency and its regulation: lessons from the gamma-herpesviruses. *Cell Host & Microbe* 8: 100–115.

Stewart, T.M., Huijskens, I., Willis, L.G. and Theilmann, D.A. 2005. The *Autographa californica* multiple nucleopolyhedrovirus *ie0-ie1* gene complex is essential for wild-type virus replication, but either IE0 or IE1 can support virus growth. *Journal of Virology* 79: 4619–4629.

Summers, M.D. and Volkman, L.E. 1976. Comparison of biophysical and morphological properties of occluded and extracellular nonoccluded baculovirus from *in vivo* and *in vitro* host systems. *Journal of Virology* 17: 962–972.

Tani, H., Nishijima, M., Ushijima, H., Miyamura, T. and Matsuura, Y. 2001. Characterization of cell-surface determinants important for baculovirus infection. *Virology* 279: 343–353.

Tani, H., Limn, C.K., Yap, C.C., Onishi, M., Nozaki, M., Nishimune, Y., Okahashi, N., Kitagawa, Y., Watanabe, R., Mochizuki, R., Moriishi, K. and Matsuura, Y. 2003. *In vitro* and *in vivo* gene delivery by recombinant baculoviruses. *Journal of Virology* 77: 9799–9808.

Teng, X., Zhang, Z., He, G., Yang, L. and Li, F. 2012. Validation of reference genes for quantitative expression analysis by real-time rt-PCR in four lepidopteran insects. *Journal of Insect Science (Online)* 12: 60.

Theze, J., Cabodevilla, O., Palma, L., Williams, T., Caballero, P. and Herniou, E.A. 2014. Genomic diversity in European *Spodoptera exigua* multiple nucleopolyhedrovirus isolates. *Journal of General Virology* 95: 2297–2309.

Thiem, S.M. and Miller, L.K. 1989. Identification, sequence, and transcriptional mapping of the major capsid protein gene of the baculovirus *Autographa californica* nuclear polyhedrosis virus. *Journal of Virology* 63: 2008–2018.

Todd, J.W., Passarelli, A.L. and Miller, L.K. 1995. Eighteen baculovirus genes, including *lef-11*, *p35*, *39K*, and *p47*, support late gene expression. *Journal of Virology* 69: 968–974.

Todd, J.W., Passarelli, A.L., Lu, A. and Miller, L.K. 1996. Factors regulating baculovirus late and very late gene expression in transient-expression assays. *Journal of Virology* 70: 2307–2317.

Tscherne, D.M., Evans, M.J., von Hahn, T., Jones, C.T., Stamataki, Z., McKeating, J.A., Lindenbach, B.D. and Rice, C.M. 2007. Superinfection exclusion in cells infected with hepatitis C virus. *Journal of Virology* 81: 3693–3703.

Van der Auwera, G.A., Carneiro, M.O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., Banks, E., Garimella, K.V., Altshuler, D., Gabriel, S. and DePristo, M.A. 2013. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Current Protocols in Bioinformatics* 43: 11.10.1-33.

Van Der Wilk, F., Van Lent, J.W.M. and Vlak, J.M. 1987. Immunogold detection of polyhedrin, p10 and virion antigens in Autographa californica nuclear polyhedrosis virus-infected *Spodoptera frugiperda* cells. *Journal of General Virology* 68: 2615–2623.

Vasconcelos, S.D., Cory, J.S., Wilson, K.R., Sait, S.M. and Hails, R.S. 1996. Modified behavior in baculovirus-infected Lepidopteran larvae and its impact on the spatial distribution of inoculum. *Biological Control* 7: 299–306.

Vaughn, J.L., Goodwin, R.H., Tompkins, G.J. and McCawley, P. 1977. The establishment of two cell lines from the insect *Spodoptera frugiperda* (Lepidoptera; Noctuidae). *In Vitro* 13: 213–217.

Vilaplana, L., Wilson, K., Redman, E.M. and Cory, J.S. 2010. Pathogen persistence in migratory insects: high levels of vertically-transmitted virus infection in field populations of the African armyworm. *Evolutionary Ecology* 24: 147–160.

Viret, C., Rozières, A. and Faure, M. 2018. Autophagy during early virus-host cell interactions. *Journal of Molecular Biology* 430: 1696–1713.

Volkman, L. 2007. Baculovirus infectivity and the actin cytoskeleton. *Current Drug Targets* 8: 1075–1083.

Volkman, L.E. 1988. *Autographa californica* MNPV nucleocapsid assembly: inhibition by cytochalasin D. *Virology* 163: 547–553.

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

Volkman, L.E. and Goldsmith, P.A. 1985. Mechanism of neutralization of budded *Autographa californica* nuclear polyhedrosis virus by a monoclonal antibody: inhibition of entry by adsorptive endocytosis. *Virology* 143: 185–195.

Volkman, L.E. and Zaal, K.J. 1990. *Autographa californica* M nuclear polyhedrosis virus: microtubules and replication. *Virology* 175: 292–302.

Volkman, L.E. and Kasman, L.M. 2000. Filamentous actin is required for lepidopteran nucleopolyhedrovirus progeny production. *Journal of General Virology* 81: 1881–1888.

Volkman, L.E., Summers, M.D. and Hsieh, C.H. 1976. Occluded and nonoccluded nuclear polyhedrosis virus grown in *Trichoplusia ni*: comparative neutralization comparative infectivity, and *in vitro* growth studies. *Journal of Virology* 19: 820–832.

Volkman, L.E., Talhouk, S.N., Oppenheimer, D.I. and Charlton, C.A. 1992. Nuclear F-actin: a functional component of baculovirus-infected lepidopteran cells? *Journal of Cell Science* 103: 15–22.

Wang, H.G. and Fraser, M.J. 1993. TTAA serves as the target site for TFP3 lepidopteran transposon insertions in both nuclear polyhedrosis virus and *Trichoplusia ni* genomes. *Insect Molecular Biology* 1: 109–116.

Wang, H.H., Fraser, M.J. and Cary, L.C. 1989. Transposon mutagenesis of baculoviruses: analysis of TFP 3 lepidopteran transposon insertions at the FP locus of nuclear polyhedrosis viruses. *Gene* 81: 97–108.

Wang, J.W. and Roden, R.B. 2013. Virus-like particles for the prevention of human papillomavirus-associated malignancies. *Expert Review of Vaccines* 12: 129–141.

Wang, M., Tuladhar, E., Shen, S., Wang, H., van Oers, M.M., Vlak, J.M. and Westenberg, M. 2010a. Specificity of baculovirus P6.9 basic DNA-binding proteins and critical role of the C terminus in virion formation. *Journal of Virology* 84: 8821–8828.

Wang, R., Deng, F., Hou, D., Zhao, Y., Guo, L., Wang, H. and Hu, Z. 2010b. Proteomics of the *Autographa californica* nucleopolyhedrovirus budded virions. *Journal of Virology* 84: 7233–7242.

Wang, Y., Oberley, L.W., Howe, D., Jarvis, D.L., Chauhan, G. and Murhammer, D.W. 2004. Effect of expression of manganese superoxide dismutase in baculovirus-infected insect cells. *Applied Biochemistry and Biotechnology* 119: 181–193.

Washburn, J.O., Trudeau, D., Wong, J.F. and Volkman, L.E. 2003a. Early pathogenesis of *Autographa californica multiple nucleopolyhedrovirus* and *Helicoverpa zea single nucleopolyhedrovirus* in *Heliothis virescens*: a comparison of the 'M' and 'S' strategies for establishing fatal infection. *The Journal of General Virology* 84: 343–351.

Washburn, J.O., Chan, E.Y., Volkman, L.E., Aumiller, J.J. and Jarvis, D.L. 2003b. Early synthesis of budded virus envelope fusion protein GP64 enhances *Autographa californica* multicapsid nucleopolyhedrovirus virulence in orally infected heliothis virescens. *Journal of Virology* 77: 280–290.

Wei, W., Gai, Z., Ai, H., Wu, W., Yang, Y., Peng, J., Hong, H., Li, Y. and Liu, K. 2012. Baculovirus infection triggers a shift from amino acid starvation-induced autophagy to apoptosis. *PloS One* 7: e37457.

Weinberg, R.A. 1995. The retinoblastoma protein and cell cycle control. *Cell* 81: 323–330.

Weng, Q., Yang, K., Xiao, W., Yuan, M., Zhang, W. and Pang, Y. 2009. Establishment of an insect cell clone that harbours a partial baculoviral genome and is resistant to homologous virus infection. *Journal of General Virology* 90: 2871–2876.

Weyer, U., Knight, S. and Possee, R.D. 1990. Analysis of very late gene expression by *Autographa californica* nuclear polyhedrosis virus and the further development of multiple expression vectors. *The Journal of General Virology* 71 (Pt 7): 1525–1534.

Whalon, M.E., Mota-Sanchez, D., Hollingworth, R. 2008. Global Pesticide Resistance in Arthropods. CABI, Wallingford, UK ; Cambridge, MA.Wickham, H. 2016. Ggplot2: Elegant graphics for data analysis, Second edition. Springer, Cham.

Wickham, T.J. and Nemerow, G.R. 1993. Optimization of growth methods and recombinant protein production in BTI-Tn-5B1-4 insect cells using the baculovirus expression system. *Biotechnology Progress* 9: 25–30.

Wickham, T.J., Davis, T., Granados, R.R., Shuler, M.L. and Wood, H.A. 1992. Screening of insect cell lines for the production of recombinant proteins and infectious virus in the baculovirus expression system. *Biotechnology Progress* 8: 391–396.

Williams, T., Virto, C., Murillo, R. and Caballero, P. 2017. Covert infection of insects by baculoviruses. *Frontiers in Microbiology* 8: 1337.

Wilson, M.E., Mainprize, T.H., Friesen, P.D. and Miller, L.K. 1987. Location, transcription, and sequence of a baculovirus gene encoding a small arginine-rich polypeptide. *Journal of Virology* 61: 661–666.

Wu, W. and Passarelli, A.L. 2012. The *Autographa californica* M nucleopolyhedrovirus *ac79* gene encodes an early gene product with structural similarities to UvrC and intron-encoded endonucleases that is required for efficient budded virus production. *Journal of Virology* 86: 5614–5625.

Xiang, X., Chen, L., Hu, X., Yu, S., Yang, R. and Wu, X. 2011. *Autographa californica* multiple nucleopolyhedrovirus *odv-e66* is an essential gene required for oral infectivity. *Virus Research* 158: 72–78.

Yamagishi, J., Burnett, E.D., Harwood, S.H. and Blissard, G.W. 2007. The AcMNPV *pp31* gene is not essential for productive AcMNPV replication or late gene transcription but appears to increase levels of most viral transcripts. *Virology* 365: 34–47.

Yong, C.Y., Yeap, S.K., Omar, A.R. and Tan, W.S. 2017. Advances in the study of nodavirus. *PeerJ* 5: e3841.

Yoo, S. and Guarino, L.A. 1994. The Autographa californica nuclear polyhedrosis virus *ie2* gene encodes a transcriptional regulator. *Virology* 202: 746–753.

Zemskov, E.A., Kang, W. and Maeda, S. 2000. Evidence for nucleic acid binding ability and nucleosome association of *Bombyx mori* nucleopolyhedrovirus BRO proteins. *Journal of Virology* 74: 6784–6789.

Zhang, F. and Thiem, S.M. 2010. The *Trichoplusia ni* cell line MSU-TnT4 does not harbor a latent nodavirus. *In Vitro Cellular & Developmental Biology. Animal* 46: 1–6.

Zhong, S., Joung, J.-G., Zheng, Y., Chen, Y., Liu, B., Shao, Y., Xiang, J.Z., Fei, Z. and Giovannoni, J.J. 2011. High-throughput illumina strand-specific RNA sequencing library preparation. *Cold Spring Harbor Protocols* 2011: 940–949.

Zhu, J. 2012. Mammalian cell protein expression for biopharmaceutical production. *Biotechnology Advances* 30: 1158–1170.

Zhu, M., Wang, J., Deng, R., Xiong, P., Liang, H. and Wang, X. 2013. A microRNA encoded by *Autographa californica* nucleopolyhedrovirus regulates expression of viral gene ODV-E25. *Journal of Virology* 87: 13029–13034.

Zuidema, D., Klinge-Roode, E.C., van Lent, J.W. and Vlak, J.M. 1989. Construction and analysis of an *Autographa californica* nuclear polyhedrosis virus mutant lacking the polyhedral envelope. *Virology* 173: 98–108.

Appendix

Sequence S5.1: *egt* AcC20 insertion sequence at passage 40.
Insert is 1350bp long and starts 1052bp after the beginning of the coding region. It is flanked by two TTAA boxes.
>CCTTTCGTGTGTTTTAGACGTGGATCCACGTTCGAAAAATCTTTACCCTCGTGTGCCTAAGCGT
GGATCCACGTCCAAAACTAACAAACTTTTTAAATGCGATTTTGTTCTCTATTCTTTGGACTTTTGG
GTAAATTTTTTTTTGCAACGAAAACGTACTGAATGAATCTGTCACGGGCGGCCATTAGACTGAAT
GGATTTCGTCATGGGAATGGGATCGACAAGTAAAAACAAGTGCATTCTGGTGACACGTGCGAC
CGGTGAAATTCCTTTTGTGTGATTCTCACGAAAATAAAcAACTGTAAGTATTCTTGTCTATTTTGC
TTTTGTTTATATTATGTAGAACATAAAACAATTGAAAATAAATGTATTATTAACTTGATTTGTGAT
ATTTCTACGATCAAATCTCTCGCGATGTTTTTTCTGTCAATGTTTTGATTTTCTTATCGATTTATATT
TGCATCTATCTTAAAAAAATGTCGTTATTATATATCATTATAAAGGAAATTTCATTAACTTTCGAA
TTGCATAAAGGTTTTAGTACATTTTTTATCAGAATATGAAAAATAATTGACATTTTCTAAGGTCAT
GCAAAATGCACGCACACATTCTCATAATAGGGCTTGCATTTACGTGACGCAGATGACGCAATCG
AATGGCGTCATGTTACGTCAGTCTTGATGTTTTTGTACAGTATGTTGTTTTGTTTGAATCCATTGA
TATATTGAACTATATATTATAATTGCAGTGTAAACTGTATATAGCGACAATGAAATAAATTGCGC
ATTTTATGGCGTTATTGACACTTGTTATTAAATAAATGGAATGAGATGGAATGGAATGAAACGA
GATGAAAAAAAAACAATATTATTTTACTATTTCAATATTTTTTACTAATAAACGATACTATACATG
CTCATGATATAAAAAAAAAAAATAACTGGTGTTGTTAATCGATTTCGCTCACGTAAAACTACTAG
CGTTTCATATGGTTTCGTCCAAGTTTTCGTAACTATTTTTCTGCACACGGTTTTTCTCGCGTATCTA
TCTTAACTTTATAATATTATAAAAGCGAAAGCAATAAATATCAGAAAGTATCAATAAGTCATTAG
TGGTTTTCTCTTCTTTTTAGGAATTCATCATGGTTGTGGTCCTGGATCTTGACCGGTTTTTGCATA
AACGAAATACAATCAAAAATAAACATCCTGTATCACCGGTCTATGTTATGGCGCGAAATTTAAAA
TGCTTACACACACATACAAAGGAACGCTCGCCGCTTCCTCGCGCCATTTCATTTCGGCCGTGCAT
ACGACGGTAATTATTACGAGCGTGGGTCAATCACACGAAAGG

Sequence S5.2: *p95* AcC20 insertion sequence at passage 40.
Insert is 444bp long and starts 2014bp after the beginning of the coding region.
>ATAATGTCTTGATAAGCAATCACCCAATATCAAATTAATCTGGCATCCGGCTGTAAAGACAAAT
ATAGGCCATGTGCGGCAATGTCTAATTTAAATCTCACGTCACCCCAGAGTGCAAAAAAGTAATA
AAAAAGAAAACGAATTTCATTTTACGCGCATGAAAGTCTATAAAAAATTAACGAGCCTTGTATA
AACATACAAGGTAGTTTTTTGGCAACAGATTCTTATAGCGTTTTCTTGTATCAATTCATTCAATTG
GCAAGTATGGCCCATTCCAATGTGATTCAGGAGAATCTTTATTTTTTCTTGTTAAAAATATTGTGA
CCGTTTGTGAGATAAAAGTTAATATAATATATATTTCTTTGTTTTCTTTCTTCGGTTTAATTTTCGA
TATATATGATGATGGACAATCTATATTTTTTGGATTGTTGCATATGTTGCATA

Figures S5.1 - S5.18 Host gene transcription profiles in AcMNPV-infected C20 and Hi5 cells.

Normalised gene expression (RPKM) in AcMNPV-infected Hi5 and C20 cells were plotted over the time of infection. Unchallenged C20 cells from 6h p.s. were considered as 0h p.i.;

Figure S5.1: Endocytosis-related genes

Figure S5.2: Genes involved in lipid synthesis

Figure S5.3: ESCRT-related genes

Abbreviations included in Fig. S5.1 to S5.18:

A - ACC: Acetyl CoA carboxylase, ACS: Acyl-CoA Synthetase, AIF: apoptosis inducing factor, Akt/PKB: protein kinase B, Alix: ALG-2 interacting protein X, ALYREF: ally of AML-I and LEF-1, AMPK: AMP-activated protein kinase, Ang2: another new gene 2, AP: adapter protein, APCC: anaphase promoting complex/ cyclosome, Apolp: apolipophorin, Are1: a region expressed 1, Arp: actin related protein, Arg/Ago: protein argonaute, α-SNAP: α-soluble NSF attachment protein, Atg: autophagy related gene, ATGL: adipose triglyceride lipase, Atm: ataxia telangiectasia protein, Atr: ataxia telangiectasia and Rad3-related protein;

B - Bet1: blocked early in transport, BIF: box- interacting factor, Bos1: Bet1 suppressor 1, Bub: budding uninhibited by benzimidazoles 1, BubR1. Bub related 1;

C - Cat: catalase, CDC: cell division cycle protein, CDH: CDC20 homologue, Cdk: Cycline dependent kinase, Chk: checkpoint kinase, CENP: centromere- associated protein, CHMP: Charged multivesicular body protein, CLHC: clathrin heavy chain, CLTA: clathrin light chain A, COPG1: coatomer protein subunit gamma 1, CPC: chromosomal passenger complex, CYP: cytochrome P450;

D - DDX: dead box protein, DHADPAT: dihydroxylacetone phosphate, DIAP: death-associated inhibitor of apoptosis, DIF: differential signalling, Dmkp: drosophila dual-specific phosphatase, DNM1L: dynamin 1-like protein, Dnr: defence repressor;

E - E2f: E2 transcription factor, EAP: ELL-associated protein, EPS: epidermal growth factor receptor; ESCRT: the endosomal sorting complexes required for transport;

F - FABP: fatty acid binding protein, FADD: Fas-associated protein with death domain, FAS: fatty acid synthase, FCHO: FCH domain only, FG: phenylalanine-glycine repeats;

G - GAK: cycline G associated protein, G: gap, GPAT: glycerol-3-phosphate phosphate acyltransferase, GPDHcyt: glycerol-3-phosphate DH (cyt), GDPHmit: glycerol-3-phophate DH (mit), GNBP: gram-negative binding protein, Gos: Golgi SNARE protein;

H - HCC8: hepatocellular carcinoma protein 8, HCCS1: hepatocellular carcinoma suppressor 1, HDAC: histone deacetylase, Hrs: hepatocyte growth factor regulated tyrosine kinase substrate, HSC: heat shock protein, HSL: hormone sensitive lipase, Hsp: heat shock protein;

I - IAP: inhibitor of apoptosis, ICAD: caspase-activated DNase, IKK: IkB kinase, INCEP: inner centromer protein, Ipr: isoproterenol, Ird: immune response-deficient;

J - JAK: Janus kinase;

L - LEPR: leptin receptor gene related protein, Lgl: lethal giant larvae, LIP5: LYST interacting protein 5, LPAAT: lysophosphaticacid phosphatase, Lp-R: lipophorin receptor, LPR: LDL receptor homologous gene, Lsd/PLIN: lipid storage droplet protein, LTp: lipid transfer particle;

M - M: mitosis, Mad: mitotic arrest deficient, MAPK: mitogen-activated protein kinase, MEKK: mitogen-activated protein kinase kinase kinase 1, Memb: mebrin, MGAT: monoacylglycerol acyltransferase, MGL: monoglyceride lipase, MIDAS: mitochondrial DNA absence sensitive factor, Mn: superoxide dismutase 2, MPK: mitogen-activated protein kinase, MTP: microsomal triacylglycerol transfer protein, Mvb12: Multivesicular body sorting factor 12, Myob: myobrevin;

N - NEMO: nuclear factor-kappa B essential modulator, NFAT: nuclear factor of activated T-cells, NFkB: nuclear factor kappa-light chain enhancer of activated B-cells, NLWASP: neuronal Wiskott-Aldrich Syndrome protein, NSF: N-ethylmaleimide sensitive fusion protein, Nup: nucleoporins, Nyv1: new yeast;

O - OXSR: oxidative stress response;

P - P70s6K: ribosomal protein S6 kinase, PAP: phosphatidic acid phosphatase, PDK: protein kinase D, PI3K: phosphoinositide-3-kinase, PLK: polo-like kinase, POLDPI3: polymerase delta-interacting protein 3, Pep12: carboxypeptidase Y deficient, PI3K: phosphatidylinositol-3-kinase, PP2A: protein phosphatase 2;

R - Rab: Ras-accociate protein, Ras: Rat sarcoma, Rb: retinoblastoma protein, RBM: RNA-binding protein, Rel: Rel-like domain containing proteins, Rod: rough deal, RZZ: Rod-ZW10, Zwilich complex;

S - S: synthesis, SAC: spindle assembly checkpoint, SARNP: SAP domain containing ribonucleoprotein, SCD: acyl-CoA delta-9 desaturase, Sec: secretory mutant protein, Sed5: suppressor of Erd2 deletion, sHSP: small heat shock protein, Sirt: Sirtuin, Sft1: suppressor of Sed5, Slt1: SNARE-like tail- anchored protein 1, Smc: structural maintenance of chromosome, Snc1/Snc2: suppressor of the null allele of CAP, SNAP: synaptosome- associated protein (size in kDa), SNX: sorting nexin, Sod: superoxide dismutase 1, SPN: serpin, SPO20: sporulation 20, Sro7/Sro77: suppressor of Rho3,

Sso1/Sso2: suppressor of Sec1, STAM: signal transducting adaptor protein, STAT: signal transducers and activators of transcription, Syb: synaptobrevin, Syx: syntaxin;

T - TGL: triglyceride lipase, Thoc: suppressor of transcriptional defects of hpr1 delta by overexpression complex, Tlg: T-snare affecting a late Golgi compartment, TRAF: tumour necrosis factor receptor-associated factor, Tom: tomsyn, Tor: target of rapamycin, TSC: tuberous sclerosis complex, Tsg101: Tumor susceptibility gene 101 protein;

U - UBC: ubiquitin-conjugating enzyme, Ufe1: unknown function essential, ULK: Unc-5 like autophagy activating kinase, Use1: unconventional SNARE in the ER, UVRAG: UV irradiation resistance-associated gene;

V - Vam: vacuolar morphogenesis, Vam6: vacuolar morphogenesis protein 6, VAMP: vesicle-associated membrane protein, viaf: IAP-associated factor; VPS: vacuolar protein sorting protein, Vtil: Vps10 interacting protein, γ-SNARE;

Y - Ykt6: YKL196c-encoding protein;

Z - ZC3H11A: human zinc finger CCCH domain- containing protein 11A, Zw10: zeste white 10;

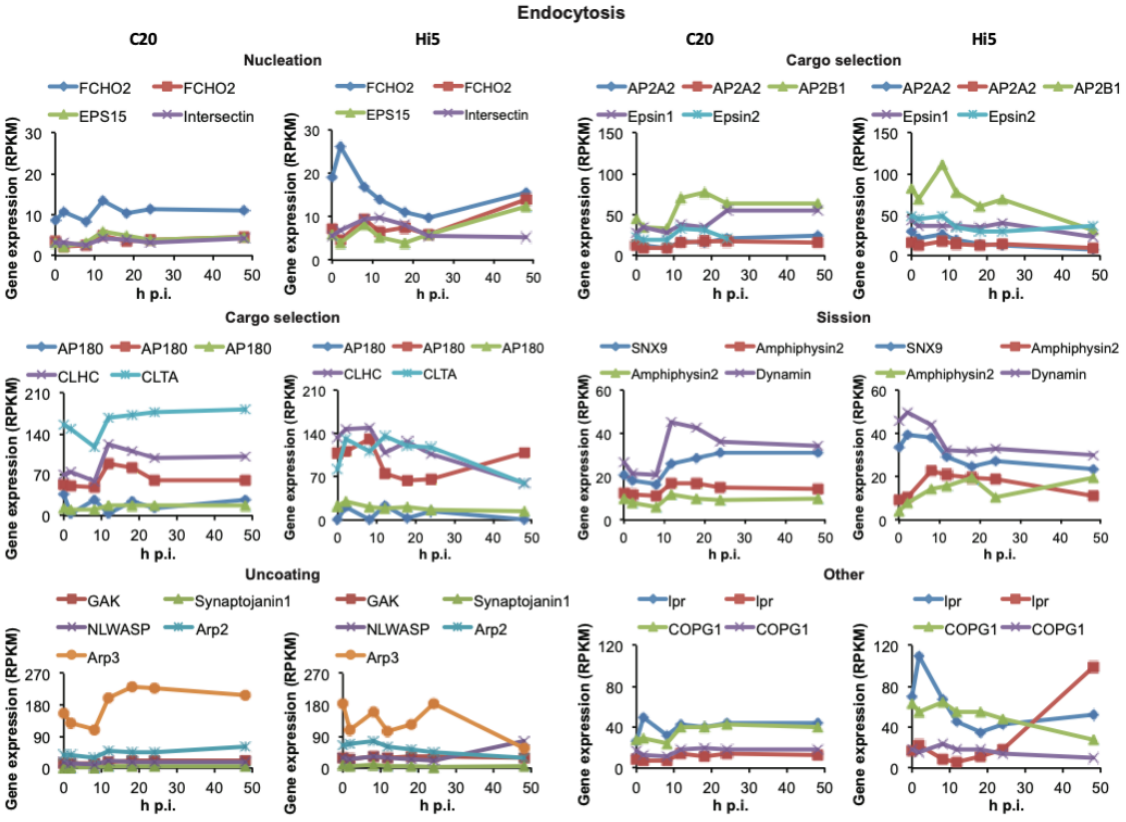


Figure S5.1: Endocytosis-related genes

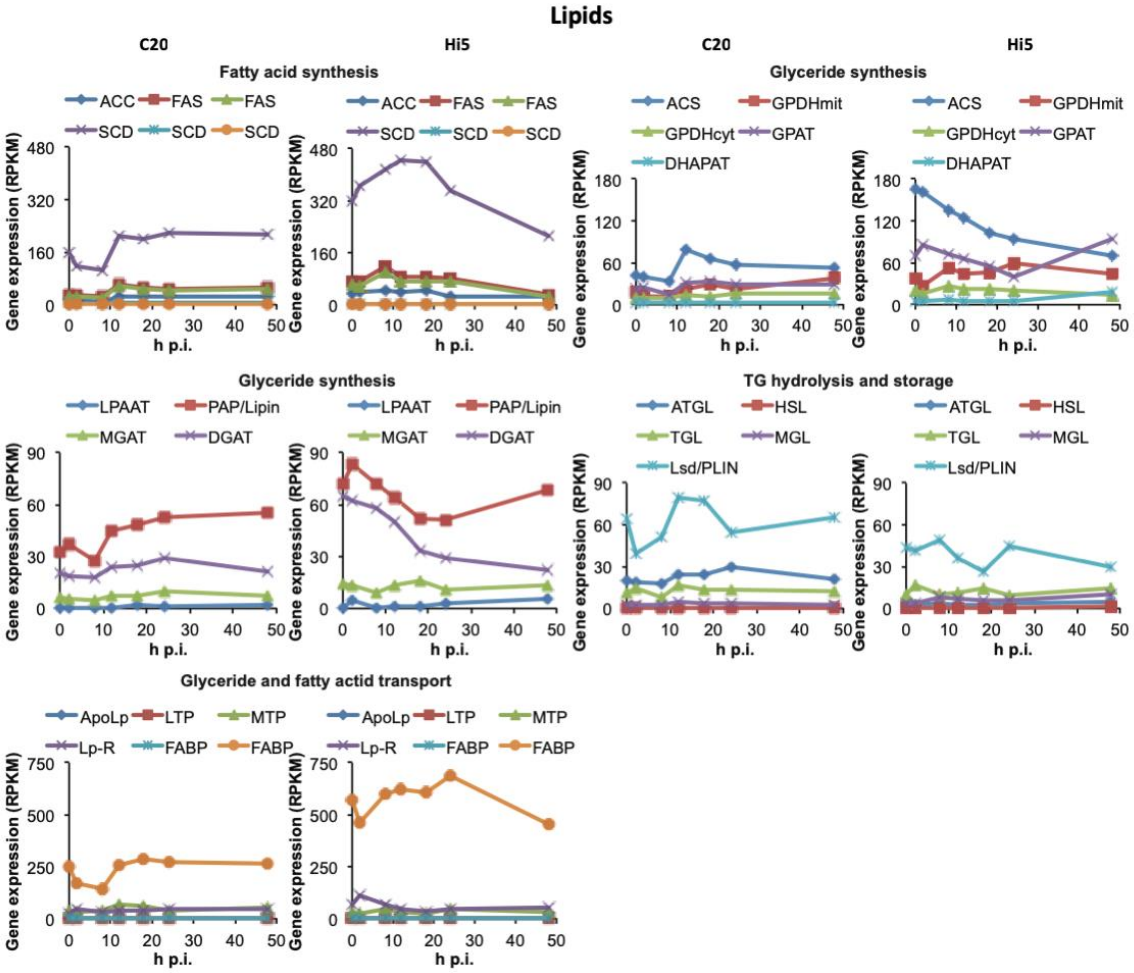


Figure S5.2: Genes involved in lipid synthesis

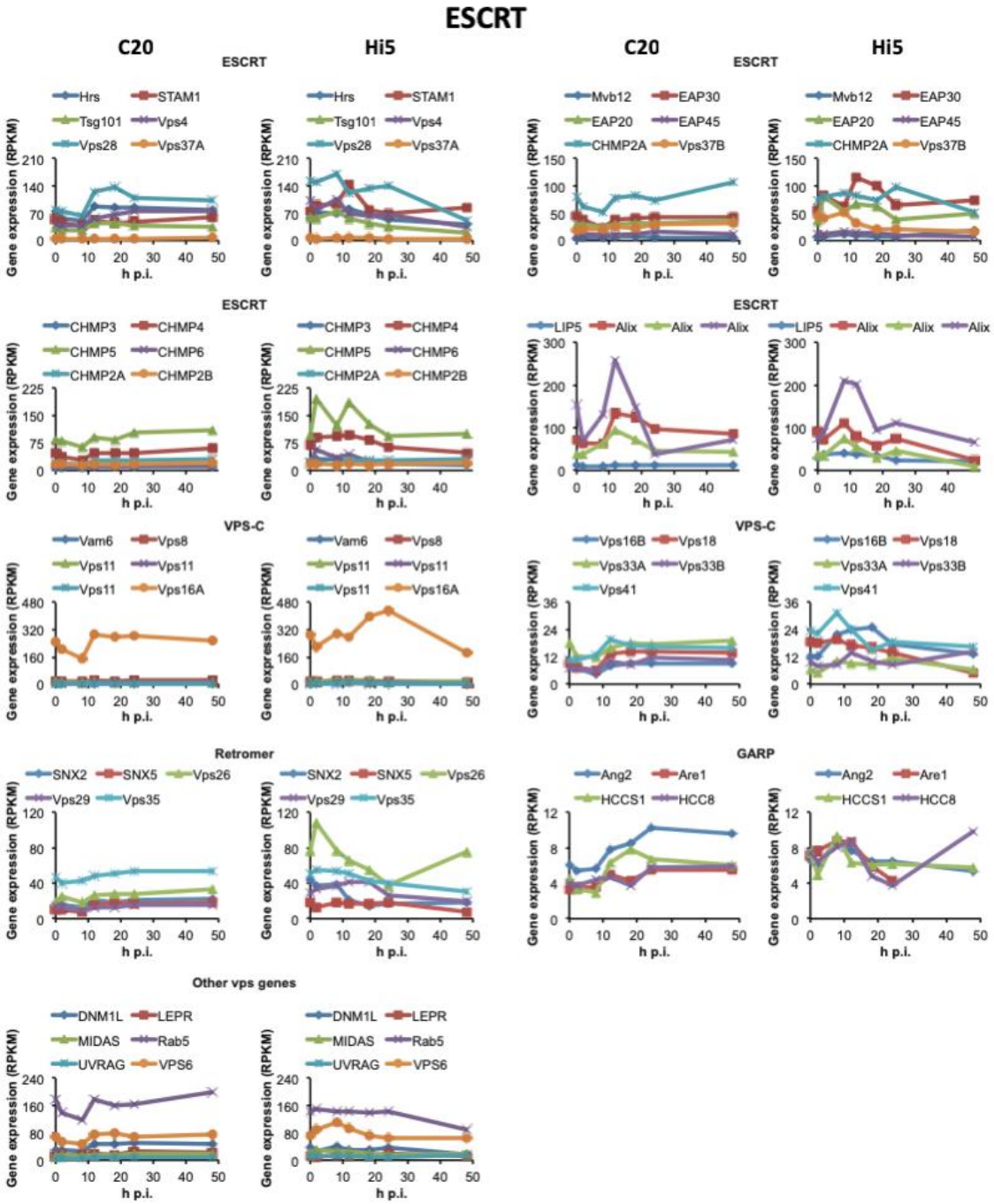


Figure S5.3: ESCRT-related genes

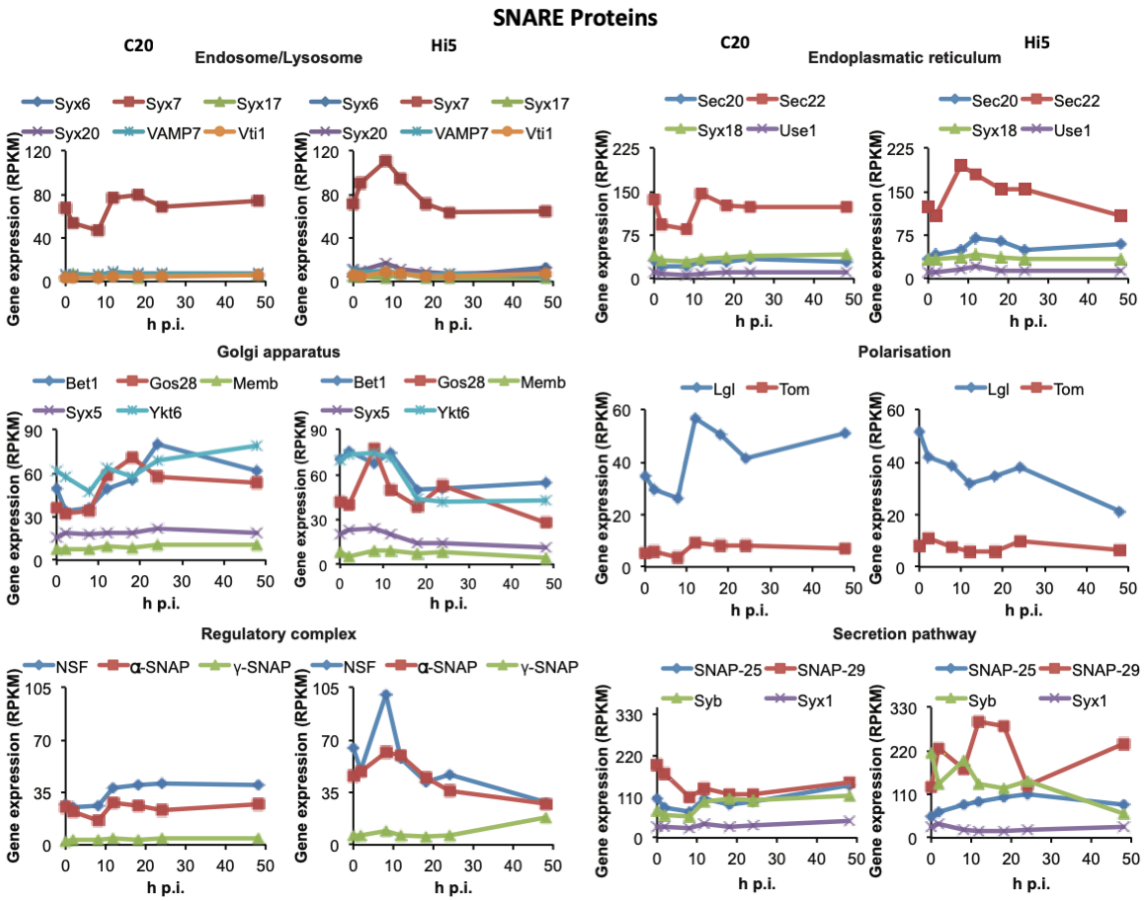


Figure S5.4: Genes encoding for SNARE proteins

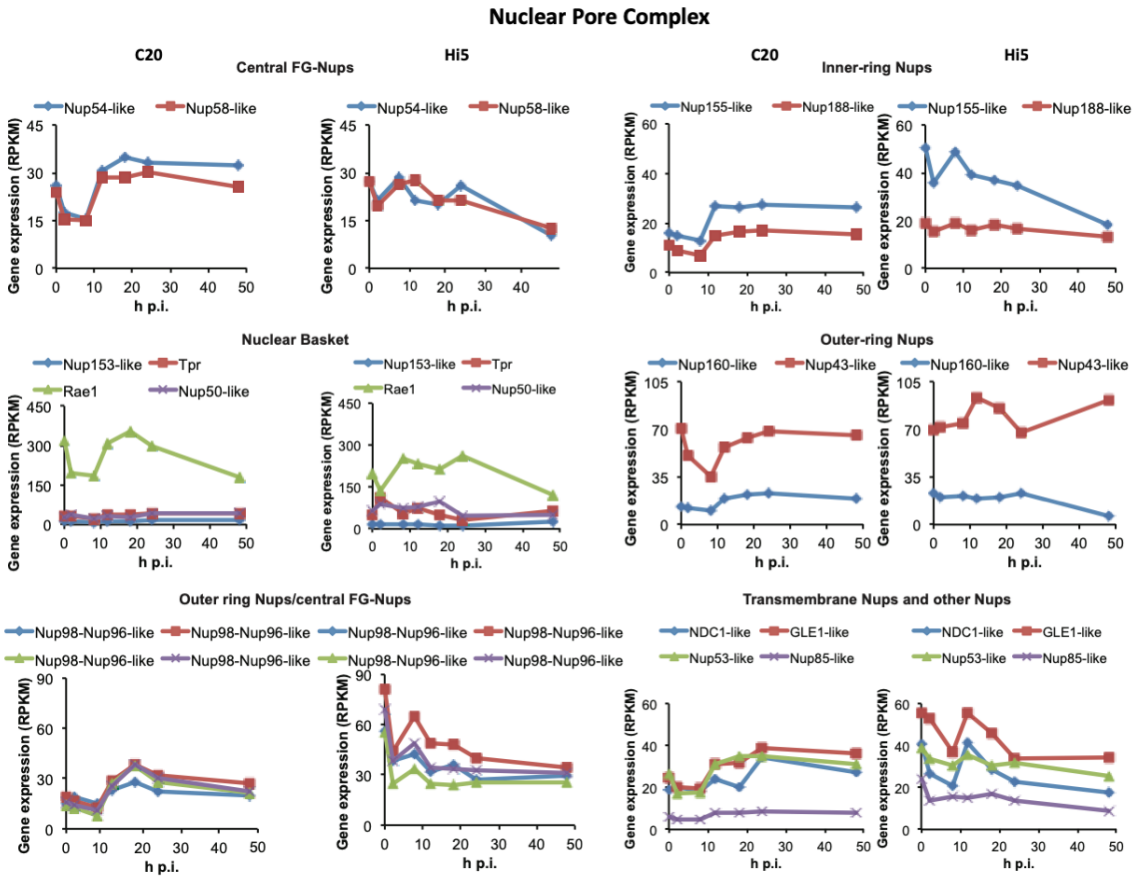


Figure S5.5: Genes encoding for the Nuclear Pore Complex

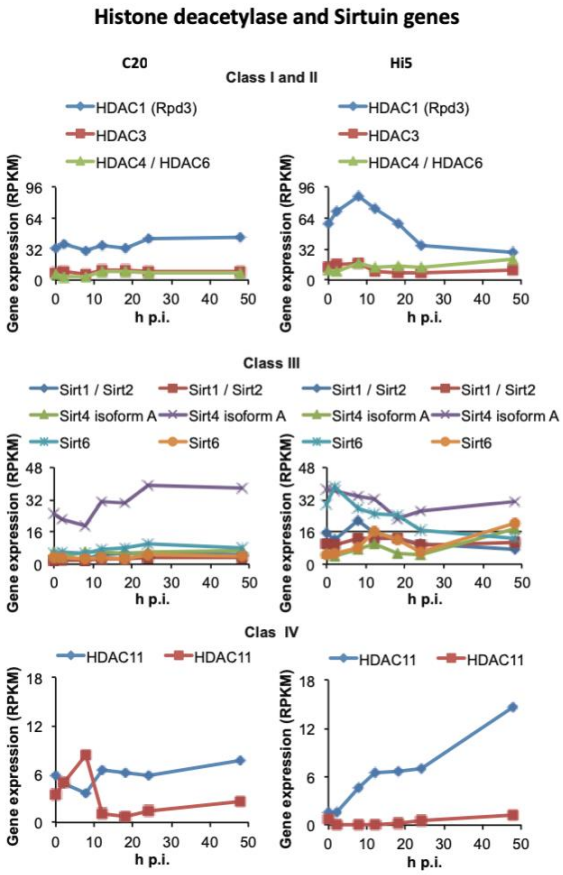


Figure S5.6: Histone deacetylase and Sirtuin genes

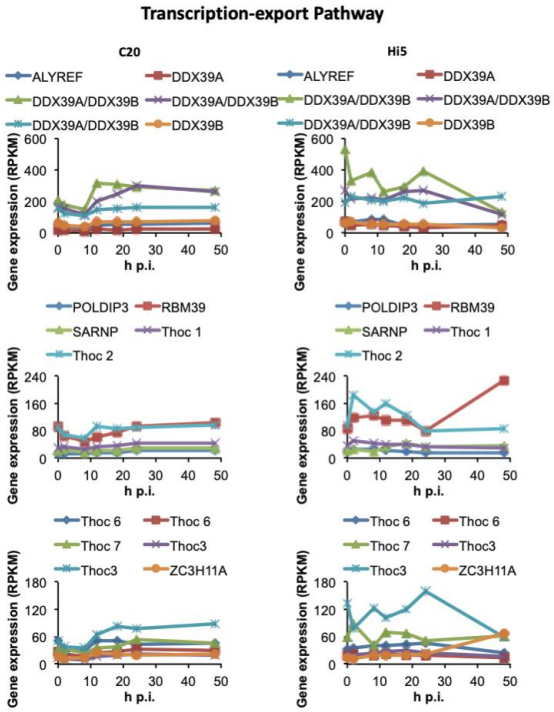


Figure S5.7: Genes encoding for the transcription export pathway

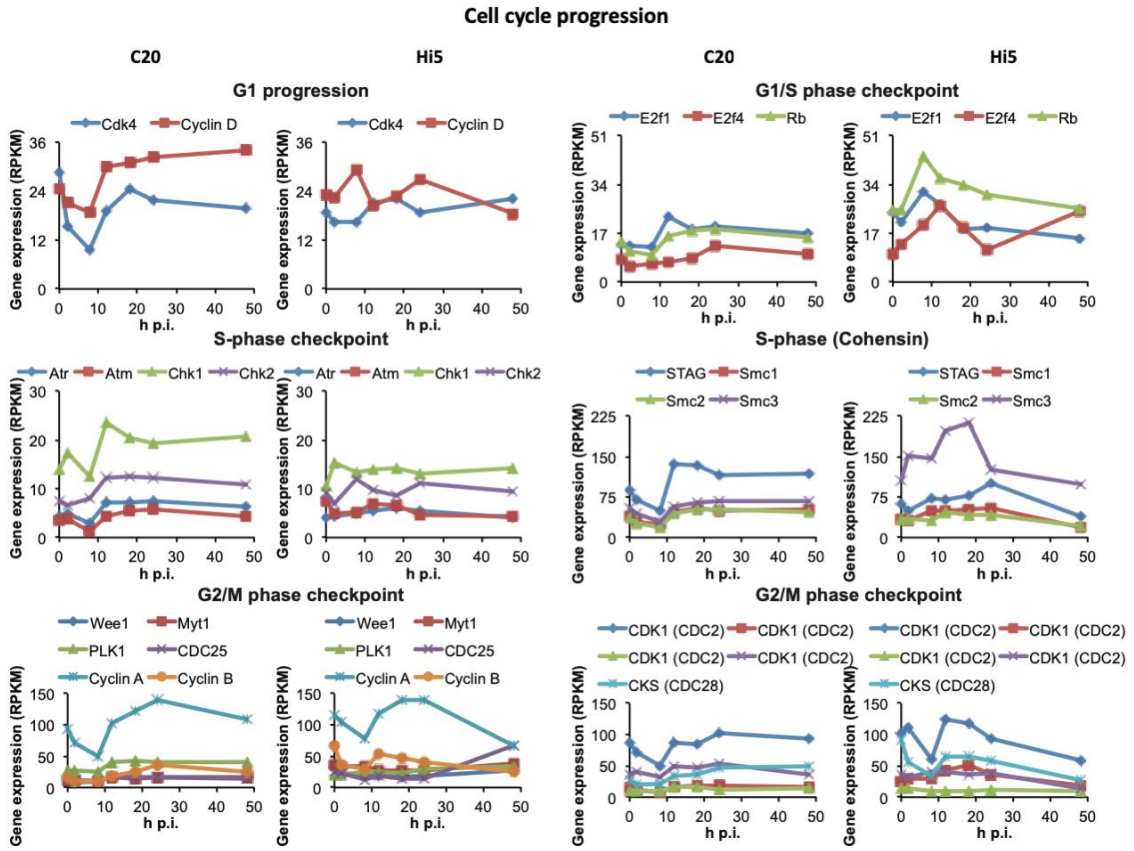


Figure S5.8: Cell cycle progression-related genes

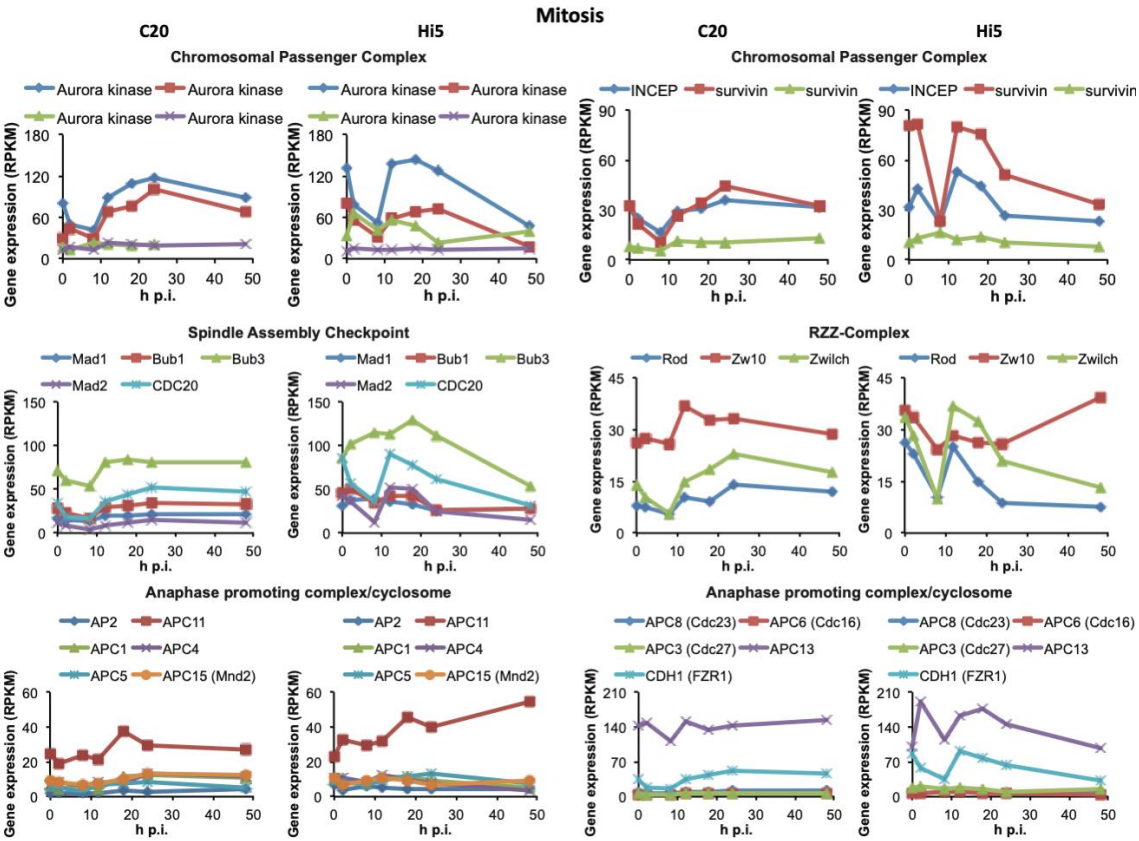


Figure S5.9: Mitosis-related genes

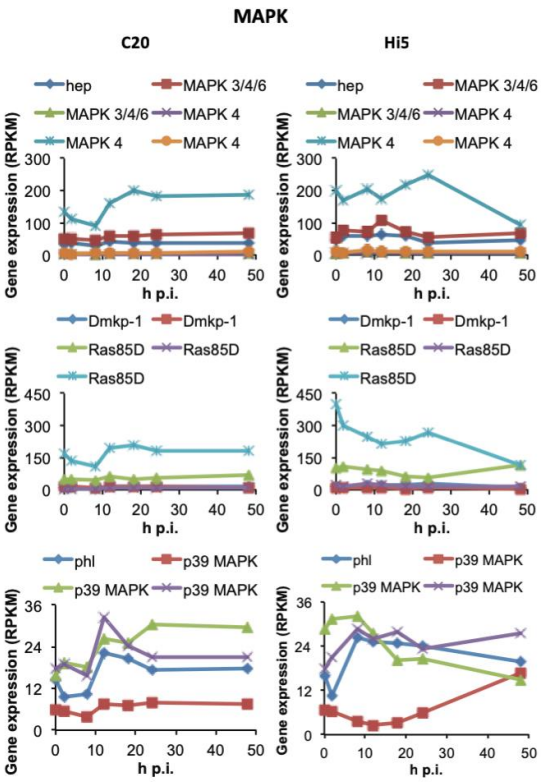


Figure S5.10: Genes encoding for proteins involved in MAPK- signalling

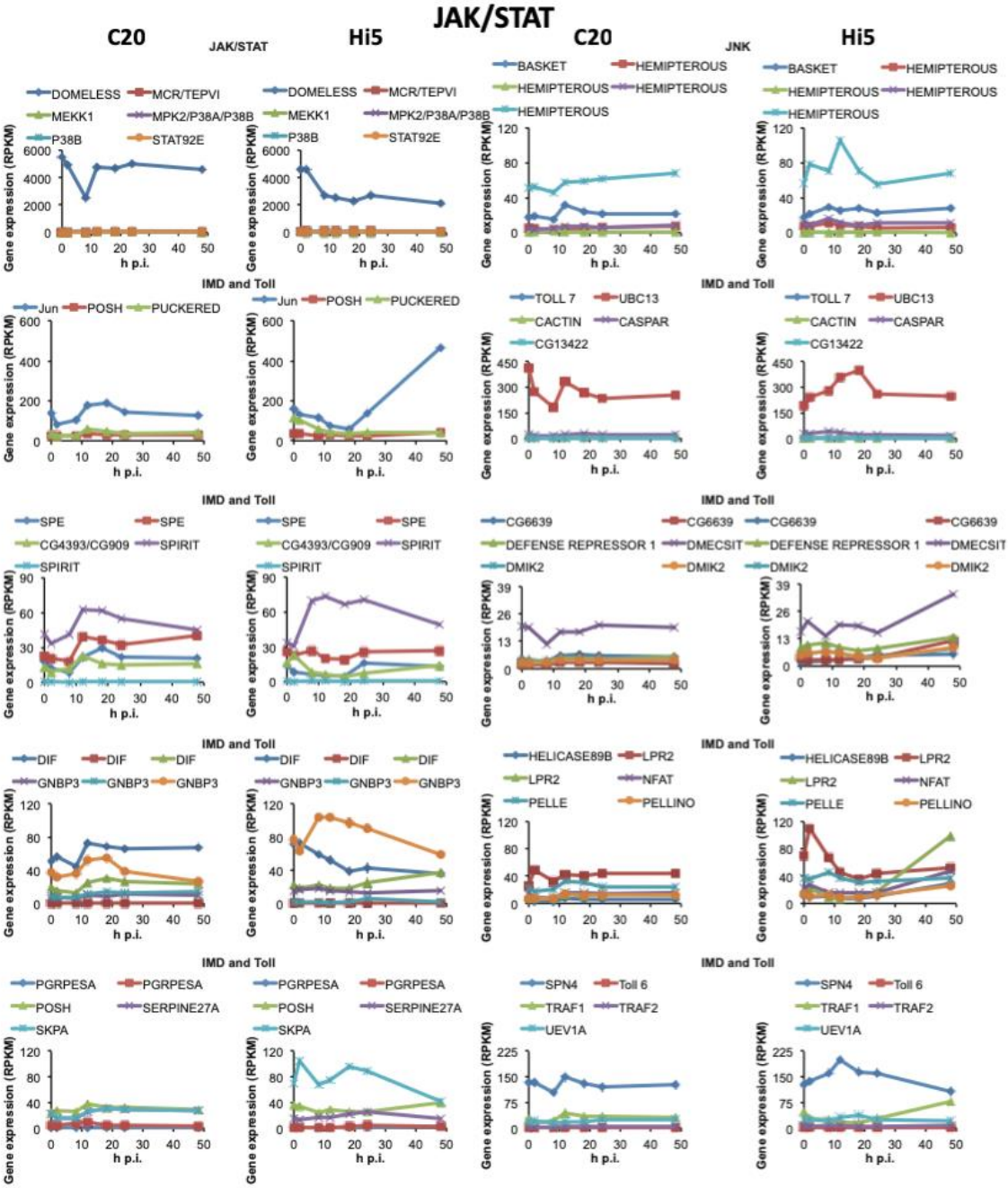


Figure S5.11: Genes encoding for proteins involved JAK/STAT signalling

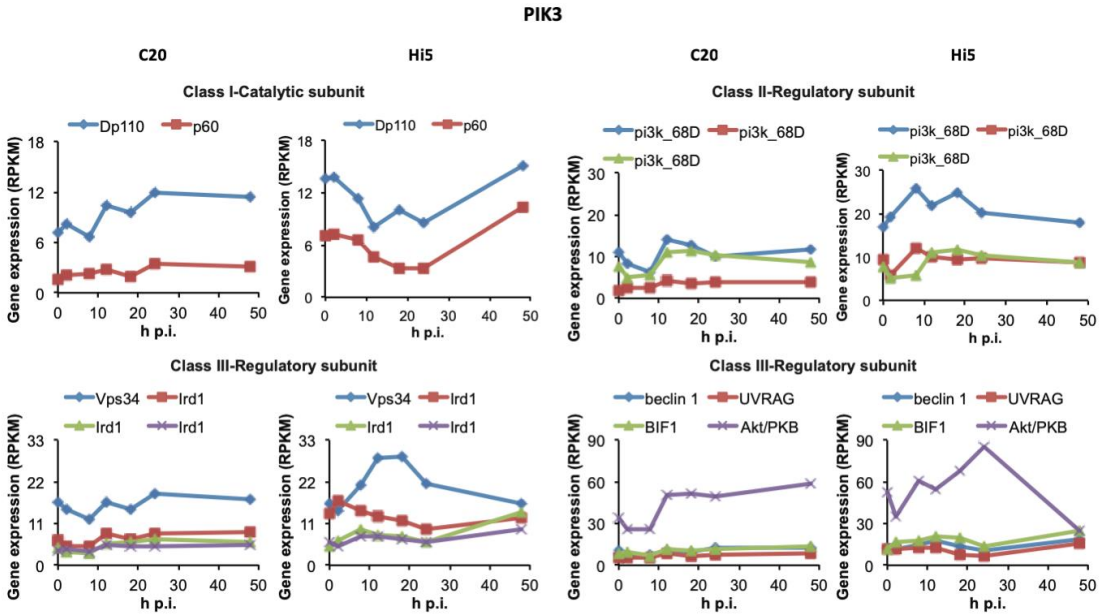


Figure S5.12: Genes encoding for proteins involved in PI3K signalling

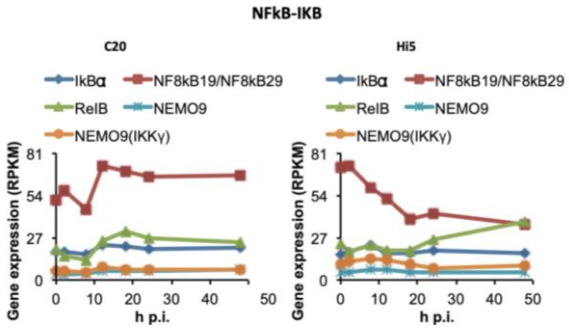


Figure S5.13: Genes encoding for proteins involved in NFkB-IKB signalling

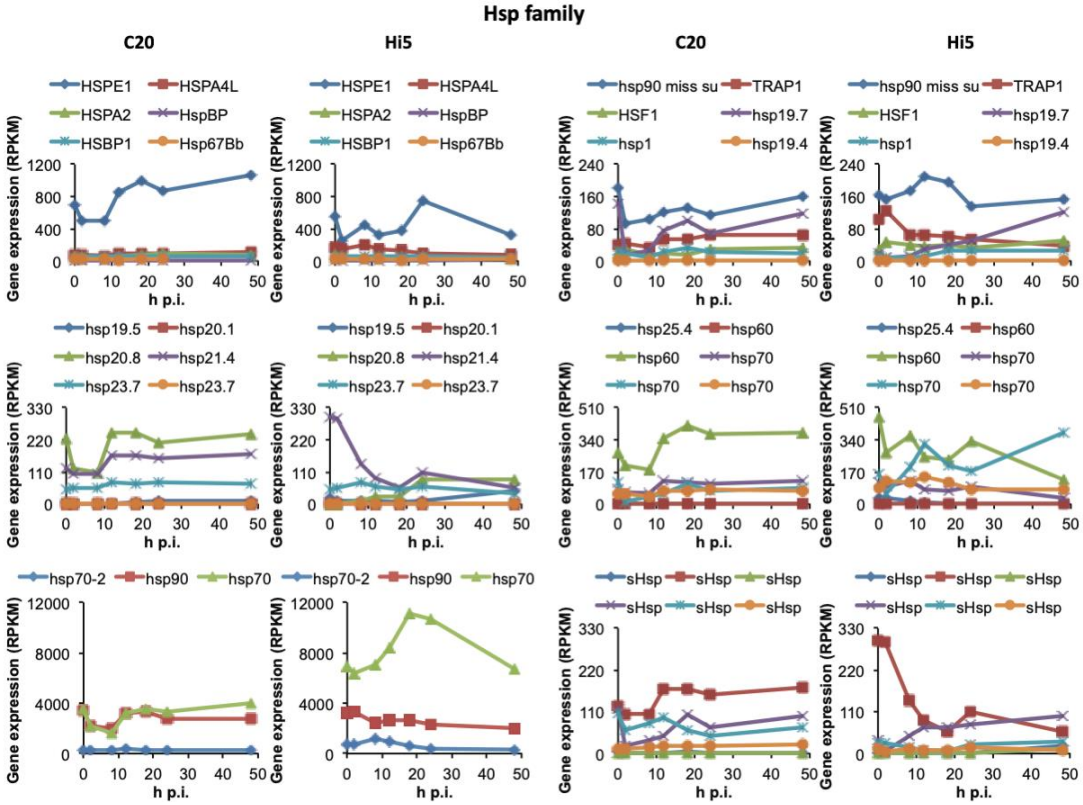


Figure S5.14: Hsp family-related genes

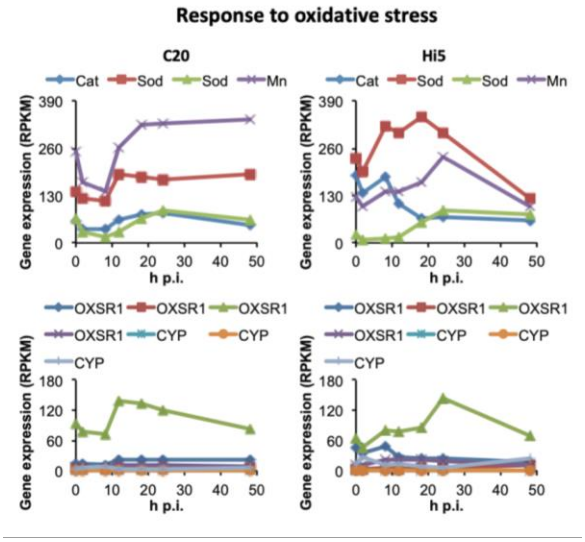


Figure S5.15: Response to oxidative stress-related genes

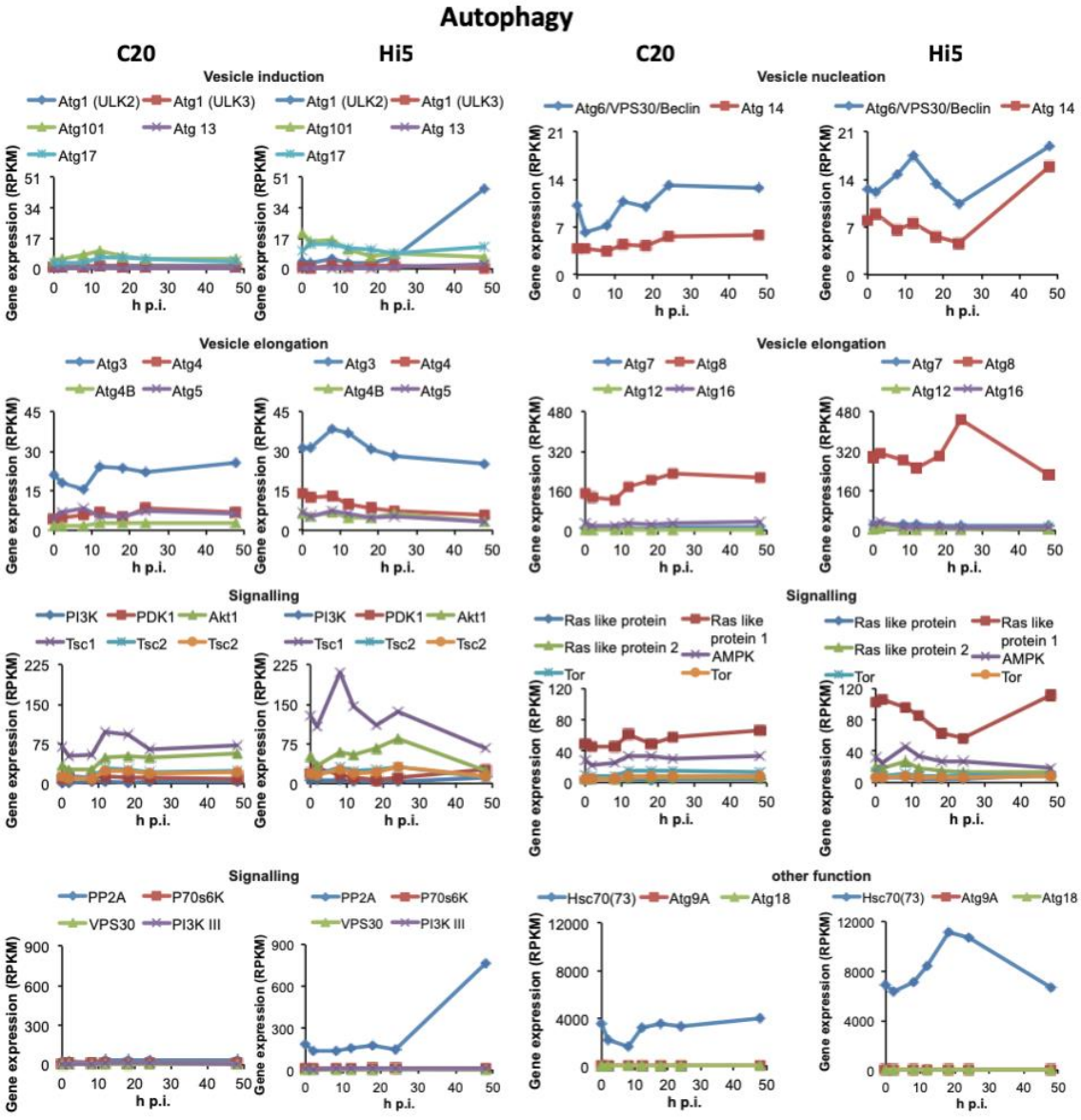


Figure S5.16: Autophagy-related genes

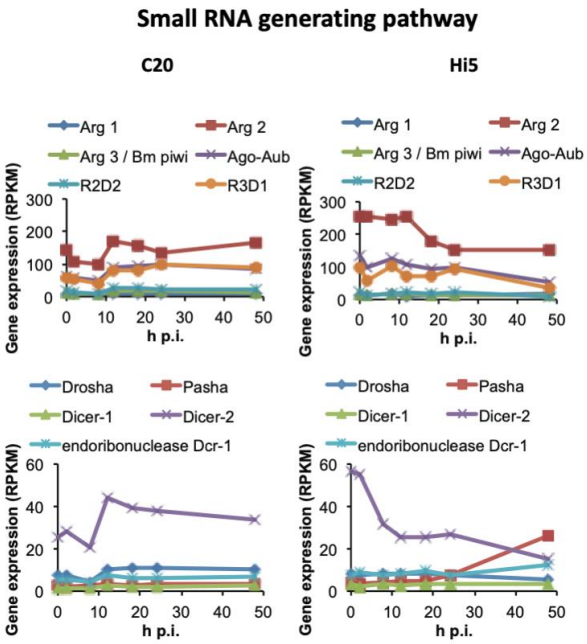


Figure S5.17: Genes encoding for RNAi-related proteins

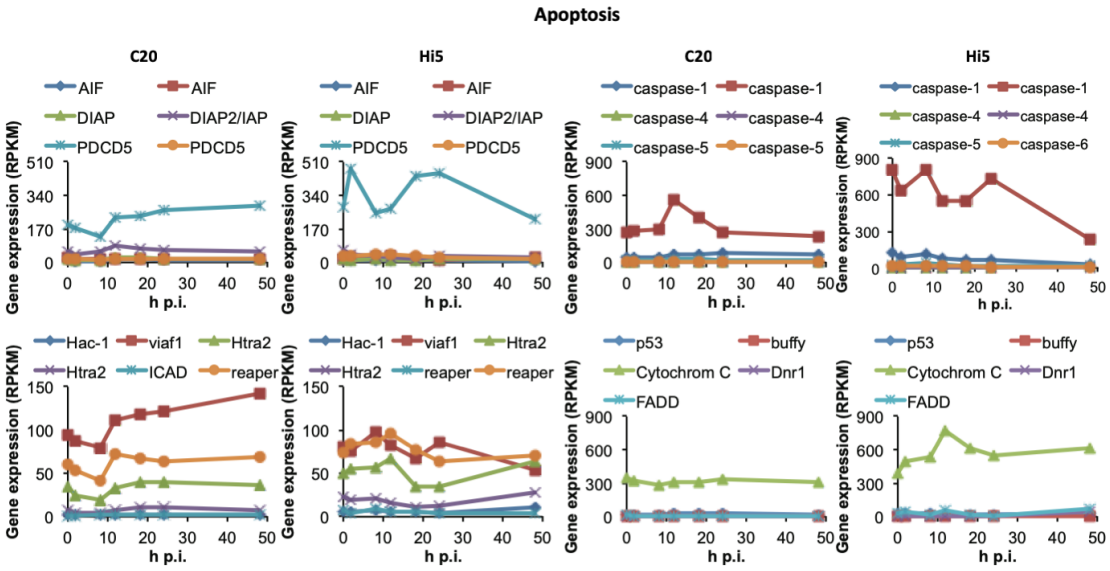


Figure S5.18: Transcription profiles of apoptosis-related genes

Supplementary data on CD:

Table S5.1: C20 host gene expression analysis.
S5.1a - Filtered reads from single-strand RNA-seq were mapped to the *T. ni* genome; S5.1b - Calculated *T. ni* RPKM values; S5.1c - Differential gene expression analysis of *T. ni* genes compared to *T. ni* Hi5 mock cells; S5.1d - *T. ni* functional groups associated with specific cellular pathways and processes: Endocytosis, Lipid synthesis, Autophagy, ESCRT, SNARE proteins, JAK-STAT, Cell cycle progression, Mitosis, Apoptosis, Heat Shock Proteins, Nuclear Pore Complex, Histone deacetylase pathway, Reactive oxygen species pathway, Small RNA generating pathway, Transcription-export pathway, MAP Kinases, NFkB-IKB, PI3 Kinases.

Video S4.1 and S4.2: Live cell confocal microscopy videos of uptake in superinfection and lytic infection.
C20 cells superinfected (S4.1) and Hi5 cells infected (S4.2) with AcEGFP-VP39 observed at 60m p.i. Viruses marked in green (EGFP) and lysosomes in red (Cresyl violet). Videos are of representative cells and incorporate 95-102 frames over 8min.