Effects of population size on virus evolution: a baculovirus perspective

Mark P. Zwart

Promotor: Prof. dr. J. M. Vlak
Persoonlijk Hoogleraar bij de Leerstoelgroep Virologie

Co-promotoren:
Dr. W. van der Werf
Universitair hoofddocent, Leerstoelgroep Gewas- en Onkruidecologie
Dr. J.A.G.M. de Visser
Universitair Hoofddocent, Leerstoelgroep Erfelijkheidsleer

Promotiecommissie:
Prof. dr. C.J.F. ter Braak (Wageningen Universiteit)
Prof. dr. H.C.J. Godfray (University of Oxford, United Kingdom)
Dr. habil. J.A. Jehle (DRL RheinPfalz, Germany)
Prof. dr. S.F. Elena (CSIC-Public University Valencia, Spain)

Dit onderzoek is uitgevoerd binnen de onderzoekschool 'Production Ecology and Resource Conservation'.

Effects of population size on virus evolution: a baculovirus perspective

Mark P. Zwart

Proefschrift
ter verkrijging van de graad van doctor
op gezag van de rector magnificus
van Wageningen Universiteit
Prof. Dr. M. J. Kropff
in het openbaar te verdedigen
op dinsdag 30 september 2008
des namiddags te vier uur in de Aula

Zwart, M.P. (2008)

Effects of population size on virus evolution: a baculovirus perspective
Thesis Wageningen University – with references – with summary in Dutch
ISBN 978-90-8504-992-0

Subject headings: baculovirus, insects, evolution, fitness, modeling

Voor Marieke, Jade en Joy

‘What kind of promise or purpose can we possibly discern in the depths of a universe upon which we cannot yet train our sights in any settled way? The transience and expected death of the cosmos defy our attempts to state clearly what the “point” of it all might be. And the long tortuous epic of life’s evolutionary struggle and suffering only adds to our disquietude. We need to remain fully aware of all the travail and messiness, and not just the astonishing complexity and creativity, in life’s coming into being. […] The notion of cosmic purpose need not be forced to coincide simplistically with the stiff and lifeless idea of divine “intelligent design”. There is wonderfully intricate patterning in nature, of course, but there is much disorder and suffering as well. By anyone’s reckoning, this universe is not a perfectly ordered one, and all instances or order eventually dissolve into the torrent of entropy. […] What we *can* say, though, is that the universe, at the very minimum, has already given rise to instances of beauty, experience, enjoyment, personality and love. We can read these, of course, as accidental outcomes of a purposeless process, with no significance and no inherent connection to the whole scheme of things. On the other hand, recognizing the possibility that the universe is still barely emerging from the cosmic dawn, we may take them as promissory symbols of the ultimate depth into which all things are being drawn.’

John F. Haught, Deeper than Darwin

Preface

A Ph.D. thesis could be compared to a pilgrimage that requires ascetic devotion. The journey began with me tramping about rather aimlessly in the bogs: it took me about a year to realize that I didn't even know what questions I was really asking. Slowly the pace progressed to a purposeful and confident uphill stride: a clear direction was found and it was energetically pursued. And the journey has ended with nothing less than an outright sprint. One might expect it to be a lonely undertaking, but like many wanderers in foreign places, I have come across and banded with fellow sojourners who happened upon the same road. Looking back over the terrain traversed during the past years, I can say that there are many canyons and mountain ridges I would never have dared to negotiate without these fellow travelers!

I owe a great deal to a lot of people who have helped me out in all sorts of ways: help with analyses and experiments, important lessons about science and the art of living, and simply making the past years a great time! I feel privileged, as so many of you took an interest in me first as a person, and then as a colleague.

Just Vlak: thanks for giving me the opportunity to do this work under your supervision. Your commitment to your Ph.D.s makes them feel at ease, while you take every opportunity to challenge them to go further and deeper. It has been a pleasure to work with you and I have enjoyed our interaction. Thanks for all your input, and there is always yet another lesson to be learned from you.

Wopke van der Werf: the extent of the your input into this project in terms of enthusiasm, great ideas, sharp criticism, encouragement, time and effort was immeasurable, and without your wit and courage to tackle new problems we would never have gotten to where we are now. I also thank you for being a friend and the life lessons learned.

Arjan de Visser: thanks for constant support, optimism and always finding new, refreshing perspectives on problems. I must also thank you for being flexible: I admit I may have visited more social events than work discussions organized by the Genetics group.

Monique van Oers: your input was invaluable to this project! Thanks for your expertise on baculoviruses and for always been willing to consider new problems – even if they were not at all related to your many areas of expertise. Thanks also for the 'gezelligheid', and that the door to your office was always open. I am greatly indebted to you!

Lia Hemerik: thanks for all your ideas, help, insight, patience, humor and enthusiasm. I also have to thank you and Wopke for the most memorable moments of the Ph.D. work, which comprise the following two elements: (1) the three of us sitting in your office having extremely useful and enjoyable interaction (2) at some point someone – usually Wopke – raising the point that we *still* didn't understand Ridout *et al.* (1993) very well!

Jenny Cory: I am greatly indebted to you, most of all for putting an end to my aimless tramping about in the bogs and actually starting to go somewhere. It was enjoyable being around you – especially the visit to Vancouver. Thanks for reading *and* responding in depth to my endless email 'bulletins', it was very much appreciated. I hope to have the opportunity to work with you again!

Rolf Hoekstra: thanks for your support. I feel particularly indebted to you for your enthusiasm and positive outlook in taking new directions. Felix Bianchi: I hope you realize that you have been important in my own evolution. Thanks for all the help and encouragement! Judy Myers: Thanks for your hospitality, letting us stay at Saturna, and the stimulating discussion. Mart de Jong: thanks for your input at various times. It is a pleasure to now be working with you, and thanks for letting me wrap up some serious loose ends while at your group. Kelli Hoover: it was serious fun to have you over in Wageningen. Thanks for your enthusiasm and input!

Rob Goldbach: thanks for your consistent stream of encouragement and 'outside the box' ideas. It has been a pleasure to work in your laboratory. Els Roode: I have very much enjoyed working with you. Always talking about our families has been a lot of fun. Thanks for all the help! Jan van Lent: Whether behind the EM of in the coffee room, there is never a dull moment when you are around. Thanks for your invaluable help! Never stop instilling wisdom and realism into tomorrow's virologists (i.e. nieuwe afleveringen van 'De wereld volgens Jan van

Lent'). Hanke Bloksma: I am greatly indebted to you for your help and company. Marcel Westenberg: thanks for your help in starting things up. You are always a positive influence, both inside and outside of the lab. Gorben Pijlman: thanks for encouraging me to complete the work on baculovirus deletion mutants. That was first rate advice! Thea van Bemmelen & Wout Rozenboom: thanks for all your help with all sorts of practical problems and the good company. Dick Peters: you are truly an inspiration to many young folks, including myself. And then a word of thanks to the M.Sc. students who contributed to this project, Eloy Erro, Ioannis Papageorgiou and Sandra Bokkers: it was a pleasure to work with you, and thanks for your tangible input.

My fellow Ph.D.s, post docs, technicians, staff & M.Sc. students at the Virology lab: thanks for making these years very enjoyable. It has been a lot of fun just hanging out, although many of you have also helped me out in the lab or given invaluable ideas. Hans, Dave, Salva, Agata, Christina, Hendrik, Richard, Gang, Jeroen, Daniela, Cristiano, Marcel Prins, Simone, Esther, Lucas, Rozemarijn, David, Martin, Liljana, Magda, Fang, Juiliette, Agah, Henriek, Nina, and Angela.

I would like to thank the following people for stimulating conversations: Charles Godfray, Ian Smith, Santiago Elena, Diana Cox-Foster, Johannes Jehle and David Theilmann. I am greatly indebted to Frans van Aggelen, Andre Gidding, Leo Koopman for rearing *S. exigua*, and Tim Carty, Bob Possee and Adolfo Hernandez for providing insect eggs. Thanks to Maria Forlenza and Beja de Vries for their invaluable help with setting up the real-time PCR. I am indebted to Drs. Mukawa and Goto for sharing their real-time PCR data. A big thank you to Henk Brandsma and Gon van Laar for advice and help with thesis formatting.

A word of thanks to my parents, Jan and Roberta, for their encouragement. I would also like to thank my family for their support throughout this often very busy time. A special thanks to Peter Verboom and Lia Verboom for reading through sections of my thesis and giving me feedback. And to my long lost friends: although my Ph.D. work is finally finished, I'm afraid that with the kids it will still be a few years until I really have more time…

Finally, I would like to thank Marieke, Jade and our newcomer Joy for their love and support. Doing research is great, but without you it would mean nothing.

Contents

Chapter 1

General Introduction

1. Viruses: how much do we really know?

The viruses are a diverse group of obligate intra-cellular micro-parasites. They are known to infect eukaryotic, prokaryotic and archaeal host organisms – in many instances causing disease. They have been the subject of considerable scientific scrutiny, in part because of these pathogenic attributes. As with all scientific endeavours, there are more new questions being raised than answers being found. It may be stated that there are real gaps in our understanding of viruses and their complex behaviour in hosts and the environment over time and space. 'Gaps' here denote areas that remain untouched despite their relevance, the availability of research approaches and the necessary technologies.

What then are these alleged gaps in the field of virology? I propose that important areas include (1) The systems biology of viral infection: although molecular biology has elucidated many details of the infection process, an overall perspective of how the virus-host complex behaves as a whole and disease results is lacking. (2) The role of virus genetic diversity: for example, Vignuzzi *et al.* (2006) recently demonstrated that complementation between poliovirus variants is required for pathogenesis, confirming the concept that an RNA virus operates as a quasi-species (e.g. Domingo *et al.*, 1985). Although this is important work, what mechanism results in the complementation observed by Vignuzzi *et al*. remains unclear and why this is all necessary to cause disease is hence enigmatic. (3) The role of population genetics of viruses in disease and disease transmission: although considerable research has been done on experimental evolution of viruses in cell culture – experimental evolution studies with Vesicular Stomatitis Virus (VSV) being a notable example (Holland *et al.*, 1991; see also Moya *et al.*, 2000) - there is a paucity of studies performed with actual multi-cellular host organisms. As an example, the genetic bottleneck that exists in initial infection of a host organism (Sacristan *et al*., 2003; Ali *et al*., 2006) and its relatedness to viral fitness (de la Iglesia and Elena, 2007) have only started to be systematically investigated *in vivo* very recently.

2. The baculoviruses

Some 700 insect species have been described to contain baculoviruses (Miller, 1997), but most have been poorly investigated. Baculoviruses owe their name to the rodded shape of their nucleocapsids in virions. These viruses have large (80-180 kbp), circular

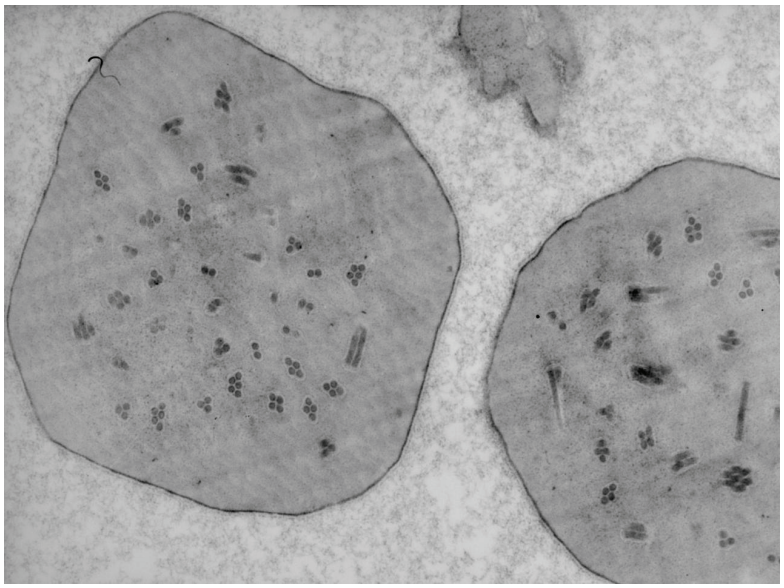


Figure 1: Occlusion bodies (OBs) of *Autographa californica* MNPV (J. van Lent).

double-stranded DNA genomes with 100-200 open reading frames (ORFs) (Theilmann *et al*., 2005). Baculoviruses are obligate host killers: their hosts must succumb to infection if they are to achieve horizontal transmission (the spread to other host individuals via new infections, Ebert and Weisser, 1997).

Baculoviruses have been extensively studied, more so than any other invertebrate virus. One of the reasons for this is that this virus family is "most beneficial from an anthropocentric viewpoint" (Miller, 1997, pg. 1). Baculoviruses have been used extensively for biological control of pest insects. Some notable examples are: the control of the velvet bean caterpillar (*Anticarsia gemmatalis*) in Brazil, the codling moth (*Cydia pomonella*) in Europe, the pine sawfly (*Gylpinia herciniae*) and gypsy moth (*Lymantria dispar*) in North America and the cotton bollworm (*Helicoverpa armigera*) in China, which is achieved in part through the use – to varying degrees - of baculoviruses (see Moscardi, 1999). Baculoviruses are frequently used as vectors for the expression of copious amounts of heterologous proteins in insect cells (King and Possee, 1990), for a variety of purposes including vaccines and diagnostics for animal and human health (van Oers, 2006). The baculovirus-insect cell system offers a relatively quick and economic means for expressing complex eukaryotic proteins that require post translational modifications (i.e. glycosylation – the addition of carbohydrate side chains necessary for correct protein folding and/or functionality). Moreover, baculovirus have shown some potential for gene therapy vectors (Hoffman, 1995).

The horizontal transmission stage (i.e. the virus morphtype which spreads the infection from host to host) of baculoviruses is the occlusion body (OB, Figure 1). It contains

either a single or multiple virions – referred to specifically as occlusion derived virus (ODV). ODV are embedded in a proteinacious matrix which protects them from environmental factors and against post-mortem decay in the host (Theilmann *et al.*, 2005). The number of ODV per OB was previously used to distinguish the genera Granulovirus (GV, single virion per OB) and Nucleopolyhedrovirus (NPV, multiple virions per OB). Recently, a new taxonomy of baculoviruses has been proposed at the genus level (Jehle et al., 2006), which has been adopted by the International Committee on Taxonomy of Viruses in 2008. Phylogenetic analysis has been confirmed that, for baculoviruses infecting Lepidoptera, the OB morphology-based taxonomical division between NPVs and GVs is a reflection of evolutionary history (Herniou *et al*., 2001), and these genera are therefore named Alphabaculovirus and Betabaculovirus respectively. Hymenopteran and Dipteran NPVs are phylogenetically distinct from the lepidopteran NPVs (and GVs), and can best be seen as separate lineages (Herniou *et al*., 2004). Hence, they are now recognized as separate genera (Gammabaculovirus and Deltabaculovirus). Alphabaculoviruses can have either single (single nucleopolyhedrovirus, abbreviated with SNPV) or multiple nucleocapsids (MNPV) per ODV. Interestingly, this morphology-based division does not reflect the evolutionary history of the NPVs. The Lepidopteran members of the genus NPV can be divided – based on phylogenetics - into Group I and II, both of which contain SNPVs and MNPVs (Herniou *et al*., 2004).

The canonical baculovirus life-history is different from that of many other viruses as there are two highly divergent virion morphotypes involved, ODV and so called budded virus (BV). In describing the baculovirus infection process[1], the terminology used to refer to steps in this process will also be defined (Figure 2). The OB (polyhedra, granula) containing ODV can persist in the environment for years under the proper conditions. For example, studies have shown that OBs can remain infectious in the soil for years (Thomas *et al*., 1972). OBs must first be ingested by a host insect larva, a process that will be referred to in this thesis as ingestion (Figure 2, step 1). They then fall apart as a result of the alkaline pH of the larval midgut juice. The embedded ODV are released in the gut lumen. We refer to this process as OB or polyhedra dissolution (Figure 2, step 2). Liberated ODV can then enter midgut epithelial cells (Figure 2, step 3; Federici, 1997). They bypass the peritrophic

[1] What then do we understand as infection? Infection involves at least two components. First, the host must be colonized by the pathogen. In other words, the pathogen must physically gain access to host tissues and produce viable progeny therein. Second, what differentiates mere colonisation of a host from infection is the significant adverse effects on host fitness in the latter case: the actual diseasing of the host.

membrane – a barrier protecting midgut cells - as the latter is in some cases made permeable by viral enhancins (GV) (Figure 2, step 3a and 3b) and then enter the midgut cell (Figure 2, step 3c). A key host defence mechanism is the so-called 'sloughing' of the midgut epithelium: the constant renewal of the outermost layer of midgut cells (Englehard and Volkman, 1995). It is therefore advantageous for the viral invasion process if the entry into midgut cells and the production of viral progeny in the form of BV proceed quickly before sloughing occurs.

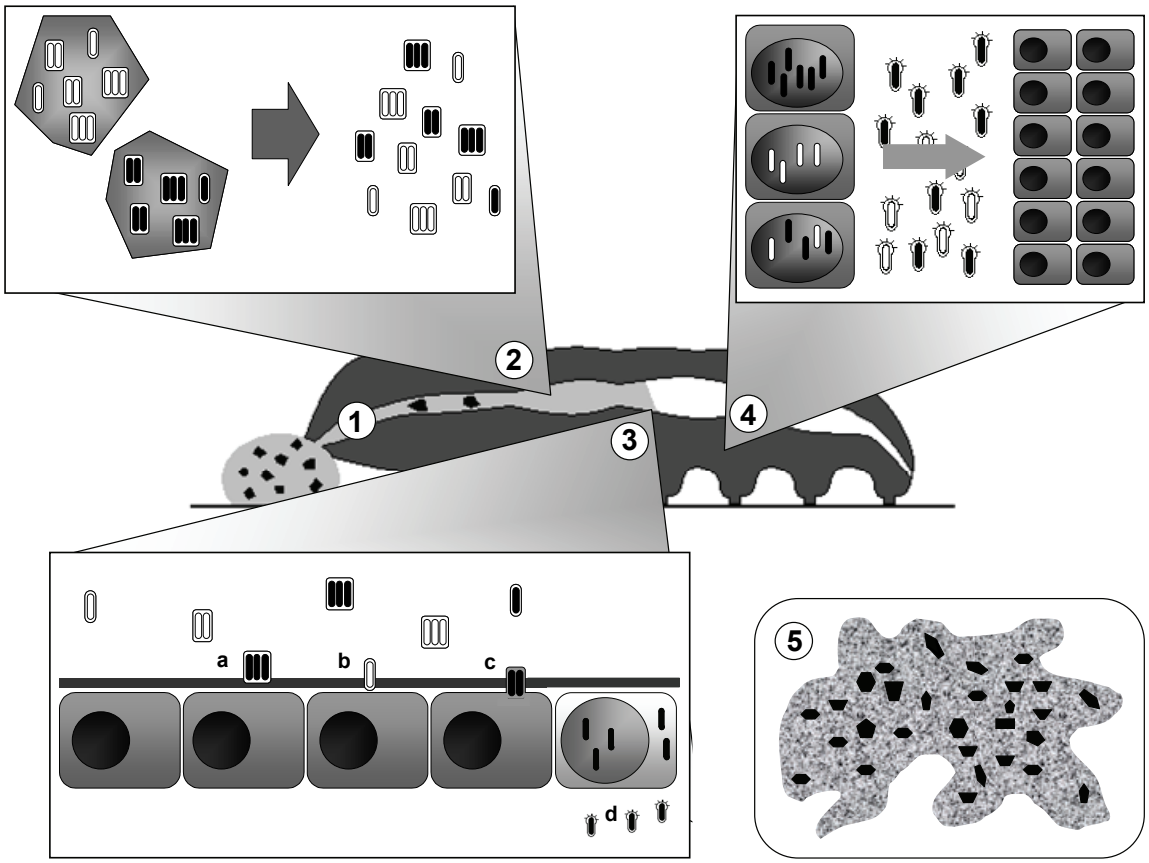


Figure 2: The NPV infection process in a typical laboratory setup: (1) Ingestion: the uptake of polyhedra through droplet feeding. Larvae are fed an aqueous suspension of polyhedra. (2) Polyhedra degradation: In the alkaline mid-gut environment, the protein matrix of the polyhedra dissolves, liberating the ODV. Note that black and white shading of nucleocapsids represents two different viral genotypes. (3) Invasion of the host: The free-swimming ODV can degrade the peritrophic membrane, making it possible to pass (a). Upon passing the membrane (b), ODV can fuse with the underlying mid-gut cells, eventually releasing their nucleocapsids into the cell (c). Multiple nucleocapsids can be present in a single ODV. Invaded midgut cells produce BV to spread the infection within the host (d). Viral replication in most host tissues follows (4): cells can be infected by a single viral genotype or a combination of both genotypes. Ultimately the host dies and liquefies, releasing newly formed occlusion bodies (5).

There is evidence that MNPVs achieve quick passage through midgut cells by 'shunting' of nucleocapsids through these cells - without the need for time-consuming viral replication (Washburn *et al*., 1999, ibid. 2003). To achieve this shunting, one or more nucleocapsids – from an MNPV ODV containing multiple nucleocapsids - enter the nucleus and commence gene expression necessary for the production of new viruses (e.g. *gp64*). The remaining nucleocapsids pass through the cytoplasm and are then simply repackaged as BV. The whole process, from entering a midgut cell to the production of BV which access the haemolymph and/or tracheal system (Figure 2, step 3d), we refer to as invasion of the host.

BV is a single enveloped nucleocapsid devoid of any occlusion that spreads the infection within the host (i.e. from cell to cell). BV egress from the midgut cells (Figure 2, step 3d), somehow pass the basal lamina and travel via the haemolymph and/or tracheal system to infect other cells types (Engelhard *et al*., 1994; Flipsen *et al.*, 1995, Federici, 1997). The subsequent process of viral amplification we refer to as replication (Figure 2, step 4). Eventually most host tissues become infected and the host dies (Figure 2, step 5). Late in the process of cellular infection, OBs are formed in the nucleus. Many baculoviruses code for lytic enzymes, chitinase and cathepsin, that cause liquefaction of the host body thereby releasing the OBs produced. The yield of OBs can be very large, depending on the species and life stage of the host insect. Yields of up to 10^{10} OBs per cadaver have been reported for Alphabaculoviruses (Sun *et al*., 2005).

It is important to note that although baculoviruses are obligate host killers, sub-lethal and latent infections of insect hosts are also common. In other words, at a relatively low dose a host may survive despite being successfully invaded by a virus (sub-lethal infection; Sait *et al*., 1994; Goulson and Cory, 1995). Such an infection can be passed on vertically (parent to progeny) without becoming apparent (latent infection; Burden *et al*., 2002). The virus may persist for many passages in the insect population, until it is triggered by 'stressing' of the host to cause disease. For example, latent infections can be triggered by the infection of the host by another baculovirus or environmental stresses (Jurkovičová, 1979, Hughes *et al.*, 1993, Cooper *et al.*, 2003). Many Lepidopteran insect cultures and natural populations have been reported to harbour latent baculoviruses (see Il'inykh and Ul'yanov, 2005), suggesting that latent infections are ubiquitous and important for baculovirus persistence (Burden *et al*., 2003). The possibility of latent infections is not only important for a complete understanding of baculovirus ecology; it may require due consideration when performing and interpreting laboratory experiments.

3. Questions arising

Even in such a well studied virus family such as the baculoviruses, there are many questions that remain to be answered. Some examples of interesting questions with respect to the biology of baculoviruses are: (1) The nature of the host receptor used by the fusion protein (F or GP64) of BV for attachment and entering of cells is unknown. For ODV, the mechanism of cellular entry has altogether not been elucidated, although it is known that four proteins called *per os* infectivity factors (PIF-1, PIF-2, PIF-3 and P74) are required for oral infectivity (Slack and Arif, 2007). A putative receptor has been assigned but poorly studied so far (Horton and Burand, 1993) (2) The genetic basis for single and multiple nucleocapsid ODV phenotypes and their functional significance is unknown. (3) In contrast to some assertions in the literature, there is no empirical basis for knowing if the multiple nucleocapsids in a single ODV from an MNPV can represent multiple virus genotypes. (4) The viral state in an insect larva persistently infected with a baculovirus is unknown. Similarly, the mechanisms by which latent baculoviruses respond to 'stress' or other baculoviruses invading an insect host are unknown. (5) And finally, what determines baculovirus host specificity? Although all of these questions are extremely interesting, they are not the subject of this thesis. These areas are however relevant and linked – to varying extents – to the questions being asking here and therefore worthy of mention.

The three gaps in virology as a whole mentioned previously in this chapter (systems biology, the role of genetic diversity and population genetics) are also highly relevant to our understanding of the baculoviruses. (1) A great deal is known about the molecular and cellular biology of baculoviruses, the course of infection and host pathology, and about the dynamics of the infection process in insect cells and insect larvae (Miller *et al.*, 1997). However, a systems perspective on the infection process (e.g. interactomics; the interaction between viral transcription, translation and the host machinery in time and space) is missing: (i) the knowledge at different levels has - for a large part - not been integrated and (ii) there have consequently not been any empirical tests of a model resulting from this synthesis. (2) What is the function of genotypic variation in a baculovirus population and why is it often stably maintained over time and space? The genetic diversity in natural baculovirus isolates has in many instances been described (e.g. Cory *et al*., 2005, Munoz *et* al., 1998, Smith and Crook, 1988, Lee and Miller, 1978), but the functional significance of this diversity remains elusive, while this is pivotal to our understanding of baculovirus diversity and evolution. A notable exception is the role of certain virus deletion mutants defective in

autonomous replication, whose functional significance has been partially elucidated (Lopez-Ferber *et al*., 2003, Simon *et al*., 2006). (3) Similarly, little work has been done on the population genetics of baculoviruses. Godfray *et al.* (1997) put forward a model for baculovirus population genetics - specifically for considering the fate of viruses incapable of autonomous horizontal transmission. This model was used as a basis for estimating that the multiplicity at which insect cells are infected by viruses – during the final round of infection *in vivo* - is high (multiplicity of infection ≈ 4.3; Bull *et al*., 2001, 2003). Despite these advances, there are still a plethora of questions concerning the population genetics of baculoviruses – and viruses in general – that remain unanswered.

This thesis is mainly concerned with the population genetics of the baculoviruses. One of the key questions that emerge is: how many virus individuals typically initiate disease in a host? This is a key issue, for a number of reasons: (1) It calls for the description of the fundamental mode of action of a virus: do virus entities (infectious particles) cooperate, coexist or constrain each other in the processes of infection and pathogenesis? (2) The number of viruses initiating infection will be of central importance for understanding the population genetics of a virus. For example, the roles that genetic drift and natural selection will play will be largely dependent on it. Genetic drift is a change in allele frequency due to stochastic processes – typically the random sampling of a small number of individuals (i.e. a bottleneck). Genetic drift will have a larger effect on allele frequencies as fewer individuals are sampled. Hence the number of individuals initiating infection will determine – in part – the role that genetic drift will play in virus evolution. Whether virus genotypes will mainly compete with each other only at the 'between-hosts' level – or also at the 'within-hosts' level - will also depend in part on the number of individuals typically infecting a host.

Despite a number of studies which do bear on this issue, this is a gap in our understanding of viruses in general, and their population genetics specifically. There are data published in many studies which can be used to argue that the number of virions invading a host and causing disease is small. These arguments can be based on (1) patterns of genetic drift, (e.g. Sacristan *et al*., 2003; Ali *et al*., 2006; Smith and Crook, 1988) (2) the number of foci of initial invasion of the host (e.g. Smith *et al.*, 2008), (3) dose-response data (e.g. Ridout *et al.*, 1993; see also Bianchi *et al.*, 2000) and (4) low median lethal dose values obtained from injection of virions quantified by biological activity assays in cell culture (Spieker *et al.*, 1996)[2].

[2] Note that some of these studies do not draw any conclusions of the number of virions initiating disease.

However, many of these arguments are not watertight. If foci of initial invasion data (2) are used as an argument, then one must consider whether multiple individuals could not have passed through one focus. Arguments from dose-response data (3) have until present required the estimation of a key parameter from the same data: the variation in susceptibility between hosts. Making an argument based on existing median lethal dose data (4) requires using the biological activity of virions (i.e. plack forming units, PFU) as a proxy for the actual number of virions, which in many cases leads to predictions that can underestimate the latter and are clearly wrong[3]. Moreover, there is not a quantitative understanding of this issue. For example, by examining patterns of genetic drift (1) it has been argued that a small number of virions is initiating invasion. But how small is small? What is the number of virions initiating disease? Is one enough? How does this number change if the host is exposed to different viral doses? These questions have simply not been answered satisfactorily.

4. Aim and Scope of this Thesis

This thesis explores the population genetics of one baculovirus species: *Autographa californica* multicapsid nucleopolyhedrovirus (AcMNPV). The focus is on two issues: (1) understanding the most basic virus-host and virus-virus interactions that lead to invasion and disease of the host, specifically the number of virus individuals invading a host insect and (2) the interactions that can occur between virus genotypes during the process of diseasing the host (e.g. competition or co-operation). The aim of the thesis is therefore to put forward and test a framework for making basic predictions for how the genotype composition of a baculovirus population will change over the infection of a host. Specifically, the goal is to better understand how genetic drift and, to a more limited extent, selection will operate on a baculovirus population. This question is particularly relevant with the introduction of genetically modified baculoviruses in the field.

Chapter 2 introduces much of the technology developed in order to research the population genetics of baculoviruses. Two technologies are central: (1) The ability to

[3] For example, Spieker *et al.* found a LD_{50} of <.01 PFU. It must take at least one physical virion to initiate disease! Note that the authors draw no conclusions on the number of founders of infection.

generate clonal virus populations that are 'tagged'. These populations are similar to a wild-type virus, but carry DNA sequences which allow them to be specifically recognized. This is achieved by the use of so called 'bacmids': full-length virus genomes containing an insert that (i) allows them to propagate as a plasmid in *Escherichia coli* and (ii) facilitates the generation of recombinant viruses (Luckow *et al*., 1993). Insect larvae were transfected with bacmid DNA in order to generate occlusion bodies. (2) The ability to measure the amount of each specific tagged virus in a mixed virus population consisting of the two tagged viruses. In order to do this, a quantitative real-time PCR (qPCR) based assay was developed and this method was experimentally validated.

Chapter 3 describes the investigation of the fate of baculovirus deletion mutants, which were generated in cell-culture, when they are re-introduced into larval insect hosts. This mimics the situation where viral genotypes, that are not infectious by themselves, can exist as parasites depending on an intact virus, which has the full genomic content. We show that these deletion mutants are quickly purged from the baculovirus population, and provide evidence that this probably arises as a consequence of their inability to replicate on their own. In other words, their persistence is dependent on co-replication with a helper (wild-type) virus.

Chapter 4 is a pivotal chapter in this thesis. Here we test whether the independent action hypothesis (IAH) can describe baculovirus invasion and diseasing of insect larvae. IAH was formulated by Druett (1953) and states that each pathogen individual has a chance of infecting and killing the host that is independent of other pathogen individuals. We show that a model that assumes a chance to infect and disease the host for each virion, independently from all other virions, allows for predictions on the number of pathogen individuals which have entered a host, based on the level of mortality experienced by hosts during a pathogen challenge. We find that IAH describes the infection process for AcMNPV if the host insects are of a susceptible species and early in their larval development. One of the implications of IAH is that a single virion can cause host disease and, in this pathosystem, death.

Chapter 5 takes a more in-depth look at the infection process in final instar larvae. In chapter 4 we found that the infection process in developmentally late (fifth instar) larvae could not be described by the independent action hypothesis as incorporated in the model. In Chapter 5 we first further consider the infection process in fifth instar

larvae experimentally, and then formulate a series of models which can describe this process by invoking additional phenomena. Here we find that a model which is congruent with IAH bests predicts the data; all that is required is the assumption there is variation in susceptibility between hosts.

Chapter 6 explores how a fast-killing baculovirus - an AcMNPV-derived mutant lacking an ecdysteroid UDP-glycosyl transferase gene (vEGTDEL) - behaves in co-infection of larval hosts with a wild type virus. Specifically, we consider the within-host fitness of vEGTDEL relative to wild type virus and find that it is reduced: the virus is lost from the population over passaging in insect larvae.

Chapter 7 is a general discussion of the research presented in this thesis, with an emphasis on connecting all the foregoing research chapters and providing a unified conceptual outlook on baculovirus population genetics and its relation to disease.

Acknowledgements:

I would like to thank Just and Ellen Vlak, Wopke van der Werf, and Arjan de Visser for their comments.

Chapter 2

A quantitative real-time PCR-based assay to determine baculovirus genotype frequencies

---

Mark P. Zwart, Monique M. van Oers, Jenny S. Cory, Jan W.M. van Lent, Wopke van der Werf, and Just M. Vlak

This chapter has been published as 'Development of a quantitative real-time PCR to determine genotype frequencies for studies in baculovirus population biology' (2008), *Journal of Virological Methods* **148**: 146-154

Abstract

Two bacmid-derived *Autographa californica* multicapsid nucleopolyhedrovirus genotypes - that differ only in a short tag sequence for differential PCR recognition - were generated. By electron microscopy, these genotypes were found to have identical polyhedra morphology. Mixtures of quantified polyhedra were made and used to validate a SYBR Green I-based quantitative real-time PCR (qPCR) to determine genotype frequencies in mixed genotype populations. The PCR could accurately quantify genotype ratios over a range of 8 orders of magnitude. Only a small correction of the genotype ratio was necessary to obtain a valid result. Low levels of aspecific background (a fluorescent signal when the template corresponding with the primer set used is not present) were measured in these validation experiments and in a typical laboratory setup. A small fitness difference between the genotypes generated was observed in a median lethal dose bioassay. The bacmid-derived virus genotypes generated and the qPCR assay are valuable tools for studying the population biology of baculoviruses.

1. Introduction

The Alphabaculoviruses (formerly known as Nucleopolyhedroviruses or NPVs) are one of four genera in the *Baculoviridae* family of arthropod viruses (Jehle *et al.*, 2006). One of their most characteristic features is the horizontal transmission vehicle of this genus, the polyhedron: a large proteinacious body in which a large number of virions are embedded (Funk *et al.*, 1997). The distinct biology and complex population structure of NPVs makes the study of their population dynamics over space and time an extremely interesting but challenging venture. From an applied perspective, elucidating NPV population biology is crucial to understanding the biological control potential of these viruses in the field and in addressing bio-safety concerns pertaining to the use of e.g. fast-acting recombinant NPVs. Natural isolates of NPVs typically contain numerous genotypes (Cory *et al.*, 2005, Knell and Summers, 1981, Lee and Miller, 1978). The evolutionary significance of this diversity – specifically if and how it is related to NPV life history traits – is unclear. A coherent theoretical framework of NPV population genetics has been proposed (Godfray *et al.*, 1997) and validated experimentally to a limited extent (Bull *et al.*, 2001, 2003). However, the current understanding of NPV population dynamics is fragmentary. Although some aspects have been better studied, such as the occurrence and role of deletion mutants defective in autonomous infection (Lopez-Ferber *et al.*, 2003, Simon *et al.*, 2004, 2006), the

premises for formulating a basic model have not been validated adequately.

In assaying genotype frequencies in a given NPV population and interpreting the significance of the results, it is necessary to have an understanding of the biology and population structure of NPVs. NPVs are obligate host-killing dsDNA viruses of arthropods with relatively large genomes (80-180 kilo bp) encoding 100-200 predicted ORFs (Theilmann *et al.*, 2005). Most known NPVs are infectious to lepidopteran insects and have a narrow host range (Theilmann *et al.*, 2005). NPVs are transmitted horizontally by polyhedra, which are able to persist outside of the host body for variable periods of time. Polyhedra are large (1-5 μm), proteinacious occlusion bodies containing numerous virions (> 100), referred to as Occlusion Derived Virus (ODV; Funk *et al.*, 1997). For the multiple nucleocapsid NPVs (MNPVs), each ODV typically contains multiple nucleocapsids. Each nucleocapsid contains a single copy of the viral genomic DNA (Theilmann *et al.*, 2005).

Polyhedra are ingested by the larval stage of the host insect and broken down in the alkaline midgut environment, releasing the ODV which in turn infect midgut epithelial cells. After initial infection, a second viral phenotype, budded virus (BV) is formed, which contains a single nucleocapsid within an envelope originating from the cell membrane. This form of the virus spreads the infection from cell to cell within the host, often transported through the haemolymph. Most of the host tissues are eventually infected and, except for the midgut, produce polyhedra (Federici, 1997). Because polyhedra are stable, easily purified, readily countable (e.g. Sun *et al.*, 2005 and the major unit for host-to-host transmission, a method for determining viral genotype frequencies in polyhedra was developed. With other purification and DNA isolation procedures, the method could easily be adapted to determine genotype frequencies in BV or intra-cellular viral DNA.

Studies of NPV population biology to date have deliberately used different genotypes with significant fitness differences, such as an inability to form polyhedra or the absence of a gene required for mid-gut infection (Bull *et al.*, 2001, 2003, Lopez-Ferber *et al.*, 2003, Simon *et al.*, 2006). Such experiments can lead to an understanding of the competitive processes between NPV genotypes and why certain genotypes persist in populations. To understand NPV population biology further, experiments with viral genotypes - preferably with no fitness differences - incorporating molecular markers allowing for quantification of genotype frequencies in mixed populations would be of great value. This would make it possible to consider the role that genetic drift has on NPV populations, especially at low virus doses, and to ascertain whether infection through the mid-gut acts as a genetic bottleneck. To this end, two genotypes differing in only a short stretch of non-coding sequence (100b.p.) were generated using bacmid

technology (Luckow *et al.*, 1993). Bacmids are full-length, biologically active viral genomes with an insertion containing: (1) a bacterial replicon and selection marker, allowing for propagation as a plasmid in *Escherichia coli* (2) transposition sites for insertion of an expression cassette.

Previous studies of NPV population biology have relied on methods such as restriction fragment length polymorphism (RFLP) analysis (Erlandson *et al.*, 2007, Cory *et al.*, 2005, Arends *et al.*, 2005, Smith and Crook, 1988), Southern blotting (Bull *et al.*, 2001, 2003) or a semi-quantitative competitive PCR (Lopez-Ferber *et al.*, 2003). However, quantitative real-time PCR (qPCR) not only has the potential to be more accurate, but also allows the development of a higher throughput assay. qPCR is therefore becoming the technique of choice for both relative and absolute quantification of levels of template DNA (Mackay *et al.*, 2002). For baculoviruses, quantitative PCR methods have been used to titrate BV levels (Hitchman *et al*., 2007, Lo *et al.,* 2004), but only semi-quantitative PCR methods have been used to determine genotype frequencies to date (Lopez-Ferber *et al.*, 2003). A qPCR-based assay allows for more accurate quantification of genotype ratios over a much larger range of ratios than existing methods. In order to quantify virus genotype ratios, specific PCR primers for amplifying the short inserts in the two bacmid-derived genotypes were used for qPCR reactions on viral DNA purified from polyhedra.

In order to further understand NPV population dynamics and population genetics, two tools were developed and validated: 1. two virus genotypes with a common genetic background differing only in a small, molecular marker, and 2. a simple yet accurate assay to determine the frequencies of these two virus genotypes in a mixed population. To this end, the following steps were taken: 1. The morphological similarity of the polyhedra of the two bacmid-derived genotypes was assessed by electron microscopy. Identical morphology would allow for the use of quantified polyhedra to validate the qPCR assay. 2. The infectivity of polyhedra of the two genotypes was determined by means of a bioassay, as having genotypes with equal fitness would be beneficial. 3. In order to validate the qPCR assay, its ability to determine genotype ratios of DNA isolated from mixtures of quantified polyhedra was tested. 4. As a final test, the qPCR assay was validated in an AcMNPV laboratory infection assay. The results demonstrate that a qPCR assay can be used conveniently to accurately quantify genotype ratios in mixed genotype populations over a large range of ratios.

2. Methods

2.1 Bacmid generation

The *Autographa californica* MNPV (AcMNPV [NC_001623], the archetype MNPV) bacmid was constructed by Luckow *et al.* (1993), with the bacterial insert located in lieu of the *polyhedri*n gene. In order to restore expression of polyhedrin, the corresponding ORF was excised from the pGEM-T/AcPolh vector (constructed by G.P. Pijlman in a similar manner to the PGEM-T/Sepol vector (see Pijlman *et al.*, 2002)) using *Eco*RI and *Pst*I, and cloned downstream of the *polyhedrin* promoter in pFastBac-Dual (Invitrogen), rendering the pFastBac-DUAL/Polh vector construct. Into this pFastBac-DUAL/Polh vector, a template for differential recognition with qPCR was inserted in the cloning site downstream of the *p10* promoter. By inserting a recognition sequence of either type A or type B, two bacmid "genotypes" were generated (bPolhA and bPolhB, respectively), which differed only in this recognition sequence (see Figure 1). In order to generate these recognition sequences, a small (60 bp) bacterial template was amplified and designed primer sequences (type A or type B) were introduced by tags on the PCR primers (see Figure 1). It was chosen to amplify a template in *E. coli* (strain DH10β) genomic DNA (*E. coli* K12 sequence [NC 000913]: 27,749 – 27,806). To generate the type A sequence, PCR was performed with the forward primer 5'- CTCGAGTCCGACGAGTTCCCTAAGGCTGGGCAAT CCTCGGCGATC-3' and reverse primer 5'-CTCGAGCACACTTGGTTGGTAGCCG TTCCGCGGTGTGAACG GC-3'. For generation of the type B sequence, the forward primer 5'-GTTTGTCAAACCCGAACTGCTGGCATTGCGGCAATCCTCGGCGA TC-3' and the reverse primer 5'-TCGGTAGTTTAAGTGTCTTGCGCATTAAGCCG CGGTGTGAACGGC-3' were used. Both PCR products were cloned into pGEM-T easy (Promega). The insert was excised from the pGEM-T easy vector with *Nco*I and *Nsi*I and cloned downstream of the p10 promoter in pFastBac-Dual/Polh, giving the pFastBac-DUAL/Polh/A and pFastBac-DUAL/Polh/B plasmids. The p10 promoter upstream of the A or B recognition sequence was then excised using *Sma*I and *BstZ*191, so as to prevent the production of transcripts of the recognition sequences inserted. With the resulting plasmids, the pMON14272 parent bacmid and helper plasmid pMON712 (Luckow *et al.*, 1993), recombinant bacmids bPolhA and bPolhB were generated. Bacmid DNA was isolated and the recombinant bacmid verified by PCR. M13F and M13R primers (Invitrogen) were used to detect 'empty' parent bacmid; M13F and the reverse primer 5'-AGCCACCTACTCCCAACATC-3', hybridizing in the gentamicin marker, were used to detect bacmids with transposed inserts.

2.2 Virus preparation

bPolhA and bPolhB bacmid DNA was used to transfect insect larvae instead of cultured insect cells, so as to prevent the generation of defective interfering (DI) viruses (Kool *et al.*, 1991). It has been demonstrated that DI viruses are generated rapidly in insect cells (Pijlman *et al.*, 2001).

To isolate high purity infectious bacmid DNA, a 200 ml LB liquid culture was grown overnight. The cells were collected and re-suspended in 1 ml STE Buffer (0.1 M NaCl, 10 mM Tris-HCl [pH = 8.0], 1 mM EDTA [pH = 8.0]). Lysozyme (in 10 mM Tris-HCl [pH = 8.0]) was then added to a concentration of 0.83 mg/ml and the sample was boiled for 30 s. After 1 minute on ice, the sample was centrifuged at 15,000 g for 15 min in a table-top micro-centrifuge. RNase A was added to a concentration of 160

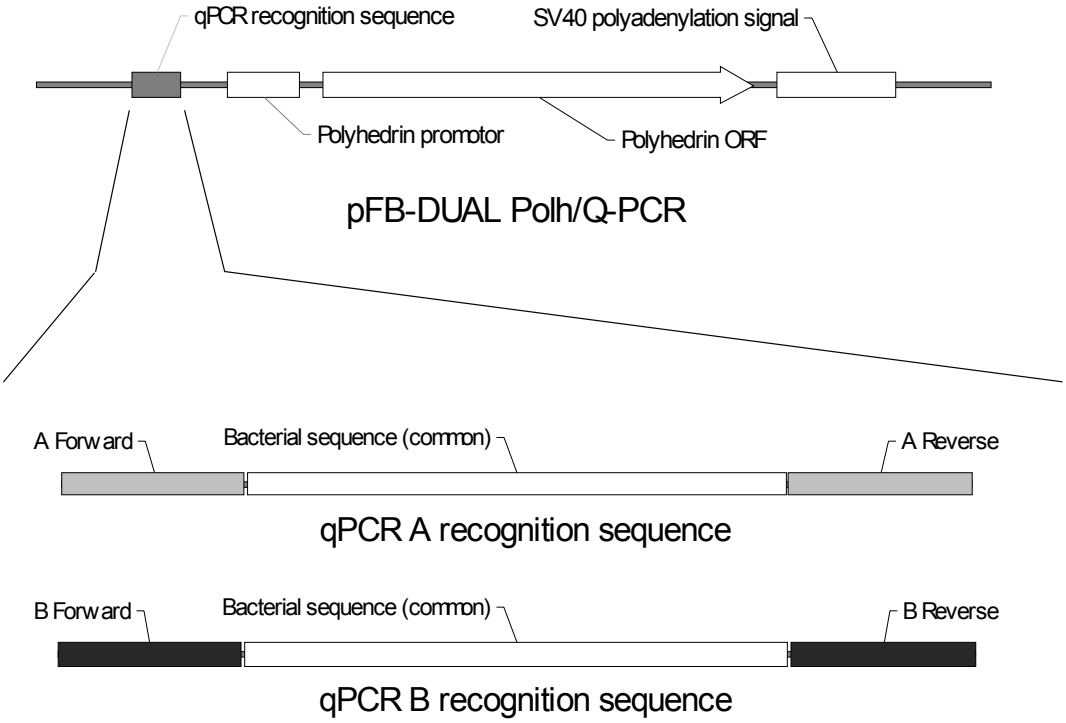


Figure 1: pFastBac-DUAL vectors generated for transposition into bacmids. Polyhedrin was inserted downstream of the polyhedrin promoter. Subsequently, either a type A (qPCR A) or type B (qPCR B) recognition sequence was inserted. These sequences only differ at the ends where the PCR primers annealed. The intermittent bacterial (*E. coli*) sequence was identical for both recognition sequences. Finally, the P10 promotor (originally between the qPCR recognition sequence and the Polyhedrin promoter) was removed.

µg/ml and the sample was incubated at 37° C for 30 min. Subsequently, Proteinase K was then added to a final concentration of 320 µg/ml, followed by 2 h incubation at 55° C. DNA isolation proceeded according to King and Possee's protocol (1992; pg 121-124; from step 2) with minor modifications. In brief, this procedure involves ultra-centrifugation over a cesium chloride gradient followed by dialysis against TE buffer (10 mM Tris-HCl [pH = 8.0], 1 mM EDTA [pH = 8.0]). The final DNA solution was then concentrated to a volume of approximately 100 µl with a speed-vac. This typically yielded a DNA concentration of 80 ng/µl.

In order to reconstitute the virus, 100 µl of bacmid DNA was mixed with 40 µl sterile milli-Q water and 60 µl Cellfectin (Invitrogen). After 15 minutes of incubation with occasional shaking, approximately 10 µl of the mixture was injected into late 4^{th} instar *Trichoplusia ni* larvae as described by Hajos *et al.* (1998). Insects that died of polyhedrosis were macerated, and polyhedra were purified by filtration through a double layer of cheese-cloth (grade #90, 'butter muslin') and three centrifugation steps (2500 g for 15 min) for washing. The pellet from the final step was re-suspended in 50% (v/v) glycerol and stored at -20° C until use. The polyhedra obtained were used to infect 30 *T. ni* L4 larvae with a high dose (>LD_{99}) in order to amplify the viruses, and purified as described above. Polyhedra were quantified by counting in an improved Neubauer haemocytometer (16 counts for each genotype).

2.3 Electron microscopy

Polyhedra of the vPolhA and vPolhB isolates, and of the authentic AcMNPV isolate (Vail *et al.*, 1971, also amplified in *T. ni*), were analyzed by electron microscopy (EM). A pellet of purified polyhedra was fixed and dehydrated, followed by standard embedding, thin sectioning and staining for EM analysis as described by van Lent *et al.* (1990). By means of electron microscopy of thin sections of polyhedra, the number of nucleocapsids per ODV was counted (>130 ODV analyzed per isolate). The number of ODV per polyhedron cross-section was also counted (> 70 Polyhedra analyzed per isolate).

2.4 Infection of insect larvae

Spodoptera exigua and *T. ni* (Lepidoptera: Noctuidae) were reared at 27° C with a 16:8 photoperiod. *S. exigua* were fed an artificial diet based on whole wheat germ and cornmeal (Smits *et al.*, 1986). *T. ni* were fed a whole wheat germ-based artificial diet. For infection of *S. exigua* L3 larvae with polyhedra, insects that molted within an 8 h time span were subsequently starved for another 16 h until use. *T. ni* was used for

transfection of bacmid DNA, as this procedure was not successful in *S. exigua* larvae – probably due to lower susceptibility to AcMNPV. *S. exigua* was used for subsequent bioassays and experiments, as AcMNPV-like sequences could not be detected in the colony by a nested PCR - whereas these could be detected in *T. ni* (data not shown).

A bioassay was performed to determine the infectivity of vPolhA and vPolhB. *S. exigua* L3 were infected using the droplet feeding method (Hughes and Wood, 1981; Bianchi *et al.*, 2000) and subsequently reared individually in 12 well plates on artificial diet. Larvae were challenged with either vPolhA or vPolhB, with 10-fold dilutions ranging from 10^6 to 10^3 polyhedra/ml. Twenty-four larvae were used per concentration, and 3 replicates of the bioassay were performed in total. The median lethal concentration (LC_{50}) was calculated using probit analysis (SPSS 12.0). A pairwise T-test was used to compare the LC_{50} values for the two viruses.

To determine how the qPCR assay performed in a typical laboratory setting, larvae were challenged with a suspension of 10^6 polyhedra/ml (approx. 10 x LC_{50} dose) as described above. Larvae were challenged with either vPolhA, vPolhB, or a 1:1 mixture of polyhedra of both viruses. Dead larvae were collected and stored individually at -20° C.

2.5 Purification of DNA from polyhedra

Larvae were macerated individually in 500 µl milli-Q water and filtered through a double layer of cheese-cloth. The samples were subsequently centrifuged for 30 sec at 100 g in a table-top centrifuge to clear any remaining large debris. The polyhedra were then pelleted by 15 min centrifugation at 2500 g and resuspended in 200 µl milli-Q and stored at -20° C.

One-hundred µl of polyhedra were used for all DNA extractions. The concentration of polyhedra from *S. exigua* larvae was not determined, but was low as larvae died in the third instar. For the quantified polyhedra mixtures, a concentration of 10^8 polyhedra/ml was used. The polyhedra were first incubated for 10 min at 37° C in 0.1 M Na_2CO_3. Tris-HCl (pH = 8.0) was added to a concentration of 30 mM. Proteinase K was then added to a final concentration of 540 µg/ml and the samples were incubated at 55° C for 2 h. A phenol:chloroform:isoamyl alcohol (25:24:1) and a chloroform:isoamyl alcohol (24:1) extraction were subsequently performed. The DNA was precipitated in isopropanol and the pellet washed with 70% ethanol. The DNA pellet was dissolved in 25 µl milli-Q water and analyzed on a 1% agarose gel. The concentration of genomic viral DNA after dissolution in water was estimated to be 5 ng/µl for the polyhedra mixtures. For the samples from *S. exigua* larvae, genomic viral DNA could only be

observed in a few samples, and the concentration was therefore estimated to be an order of magnitude lower than in the polyhedra mixtures samples. A 1:500 dilution of the DNA sample was used as the template for qPCR.

2.6 qPCR

The forward primer 5'-TCCGACGAGTTCCCTAAGGCTG-3' and reverse primer 5'-CACACTTGGTTGGTAGCCGTTC-3' were used for the detection of vPolhA. The forward primer 5'- GTTTGTCAAACCCGAACTGCTGGCATTGC-3' and reverse primer 5'-TCGGTAGTTTAAGTGTCTTGCGCATTAAGC-3' were used for the detection of vPolhB. All reactions were performed using Qiagen Quantitect SYBR Green Master Mix in a total reaction volume of 15 µl. Seven and a half µl Master Mix was used, forward and reverse primers were added to give 300 pm final concentrations, and 5 µl of template DNA was added. Non-template controls (NTCs) and standard curves (5×10^6 - 5×10^3 template copies – in the form of plasmids - per reaction) were performed *in duplo*; samples were analyzed *in triplo*. Purified plasmid DNA of the pFastBac-DUAL/Polh/A and pFastBac-DUAL/Polh/B constructs was used for standard curves. The DNA was quantified with a Nanodrop (Nanodrop Technologies) and verified by gel electrophoresis, and the number of plasmid copies were calculated based on the DNA concentration.

The PCR reactions were performed in a RotorGene 2000 thermal cycler (Corbett Research) and fluorescence was recorded on the FAM channel with a gain setting of 8. The program used for both primers was: 15 min at 94° C; 40 cycles of 94° C for 15 sec, 60° C for 30 sec, and then 72° C for 30 sec (acquisition of fluorescence signal); followed by a 'melting curve' (60° – 99°).

Data were analyzed using Rotorgene 6.0 (RG6) software (Corbett Research). Quantification was performed using the comparative analysis method in this software. Unlike analysis with standard curves, comparative analysis requires only a reference sample. For each sample to be analyzed, a take-off value (similar to Ct-value in standard curve analyses) is calculated and the second derivative of fluorescence over cycles is used to determine the efficiency of amplification. These values are then used to calculate the concentration of template DNA in a sample as compared to the reference samples. The 5×10^6 template copies standard curve samples were used as a reference. A 10% increase in total fluorescence (a setting referred to as 'NTC Threshold' in RG6) was required for a sample to be considered positive. Likewise, a reaction efficiency ('Reaction Efficiency Threshold' in RG6) of 10% was also required for a sample to be considered positive. Estimated mean copy numbers ($mean_a$,

$mean_b$) were used to calculate the ratio of vPolhA:vPolhB (R) for a particular sample:

1. $$R = \frac{mean_a}{mean_b}$$

The standard error of the ratio ($s.e._R$) was determined as follows:

2. $$s.e._R \approx R \cdot ([\frac{s.e._a}{mean_a}] + [\frac{s.e._b}{mean_b}])$$

$S.e._a$ and $s.e._b$ are the standard error of the mean corresponding to $mean_a$ and $mean_b$, respectively.

In order to validate this PCR assay, mixtures of quantified polyhedra of vPolhA and vPolhB were made (10,000:1 till 1:10,000, with 10 fold intervals). Polyhedra were purified and DNA extracted as described above and the DNA in the mixtures quantified by PCR. The quantification was performed *in duplo*.

The bacmid inserts – which include a type A or type B recognition sequence in the recombinant bacmids – are rather large (12 kbp) and this may result in gradual loss over passage, as is the case *in vitro* (Pijlman *et al.*, 2001). A stability check was therefore performed. The forward primer 5'- TCGGAATCCCTTGAGCAGCCTG-3' and reverse primer 5'- TTGCCGATGGTTGGTTCACACC-3' were used to amplify the immediately early gene 1 (*ie-1*, essential for virus replication) in a quantitative PCR. The reaction conditions were otherwise identical to those for vPolhA/vPolhB as described above. The ratio of recognition sequence (type A or type B) to *ie-1* was determined in DNA from the *T. ni* larva transfected with bacmid DNA (passage 1) and in two *S. exigua* larvae infected with a high dose of vPolhA or vPolhB (passage 3). A T-test was used to assess whether the ratios obtained for passage 1 and 3 were identical (performed separately for each passage 3 larva).

3. Results

3.1 Number of ODV per polyhedron cross-section for bacmid derived viruses

In order to use mixtures of quantified polyhedra to validate the qPCR assay developed here, it first had to be determined whether the polyhedra morphology of the two

bacmid derived viruses was identical. Moreover, it would be desirable that the morphology of bacmid derived viruses was similar to that of a wild-type virus. In order to determine whether there were any differences between the viruses in polyhedra morphology, the number of ODV per polyhedron cross-section was estimated. vPolhA, vPolhB and the 'Vail' isolate (Vail *et al.*, 1971) had similar numbers of ODV per polyhedron cross-section (AcMNPV-'Vail' (mean ± standard error of the mean) = 5.91 ± 0.52, vPolhA = 6.46 ± 0.57, vPolhB = 6.41 ± 0.57; Kruskal-Wallis Test, Chi-squared = 0.215, d.f. = 2, Significance = 0.898).

3.2 Number of nucleocapsids per ODV for bacmid derived viruses

ODV of the two bacmid-derived viruses generated (vPolhA and vPolhB) and of the 'Vail' AcMNPV isolate were examined (Figure 2). There was no significant difference between these three viruses in the number of nucleocapsids per ODV (Kruskal-Wallis Test, Chi-squared = 3.709, d.f. = 4, Significance = 0.447). This is therefore not a factor which influences quantification of genotype frequencies when comparing mixtures of single genotype polyhedra. Since the morphology of the two bacmid derived viruses is identical, making a mixture with a given ratio of polyhedra will result in the same ratio of ODV, viral nucleocapsids, and copies of viral genomic DNA.

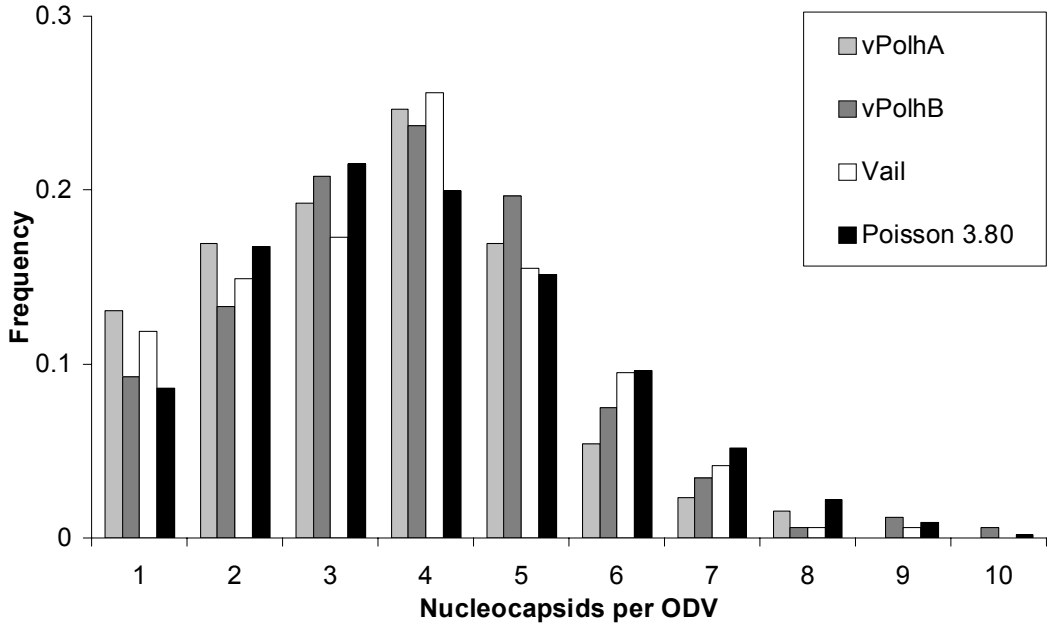


Figure 2: The distribution of the number of nucleocapsids per ODV. On the x-axis is the number of nucleocapsids per ODV, on the y-axis the frequency of this number of nucleocapsids per ODV was counted. The observed frequencies are shown for vPolhA, vPolhB and the Vail AcMNPV isolate. A truncated Poisson distribution with a mean of 3.80 is shown for comparison.

The data on nucleocapsids per ODV of all viruses were combined and found to follow a Poisson distribution (One-sample Kolmogorov-Smirnoff test, Poisson $\lambda = 3.71$, N = 688, Z = 1.319, Significance = 0.062). Since an ODV must contain at least one nucleocapsid, a truncated Poisson distribution – one that includes only non-zero integers – is a more appropriate description of this distribution. Given the λ (which for a true Poisson distribution describes the mean and variance) of a truncated Poisson distribution, Olkin *et al.* (1994) state that the mean, μ_χ, is:

3. $$\mu_x = \frac{\lambda \cdot e^{\lambda}}{e^{\lambda} - 1}$$

Using this formula a mean of 3.80 nucleocapsids per ODV was calculated.

3.3 Infectivity of polyhedra of vPolhA and vPolhB

The infectivity of the polyhedra of vPolhA and vPolhB was determined by means of a droplet feeding bioassay in *S. exigua* L3 larvae as described by Bianchi *et al.* (2000). The LC_{50} of vPolhB was higher than that of vPolhA, despite the common genetic bacmid background (mean log LC_{50} (polyhedra/ml), ± standard error of the mean; vPolhA = 4.52 ± 0.08, vPolhB = 4.96 ± 0.11). The difference in LC_{50} between the polyhedra of the two viruses is significant (pairwise T-test; t = 5.297, d.f. = 2, Significance (2-tailed) = 0.034).

3.4 Quantification of vPolhA:vPolhB virus mixtures by qPCR

It was established above that polyhedra morphology of the bacmid derived viruses was identical. It is therefore feasible to make mixtures of quantified polyhedra, extract DNA, perform qPCR, and use the resulting data to evaluate the performance of the qPCR assay (Figure 3). The qPCR assay could quantify virus ratios in a range from 10,000:1 to 1:10,000 ($r^2 = 0.9986$, for a linear regression analysis on both replicates). When analyzed independently, the two replicates gave highly similar lines of best fit (replicate 1: $y = 0.858x + 0.198$, replicate 2: $y = 0.820x + 0.143$), another indication that the assay gives consistent and reproducible results. When the log polyhedra mixture input versus log qPCR output is plotted (see Figure 3), the m-value (slope) for the line of best fit deviates from 1, and the b-parameter (intercept) deviates from zero. An assay with a valid outcome - despite being unadjusted - would have a regression with a slope of 1 and have a y-intercept of zero. Thus, this indicates that a correction

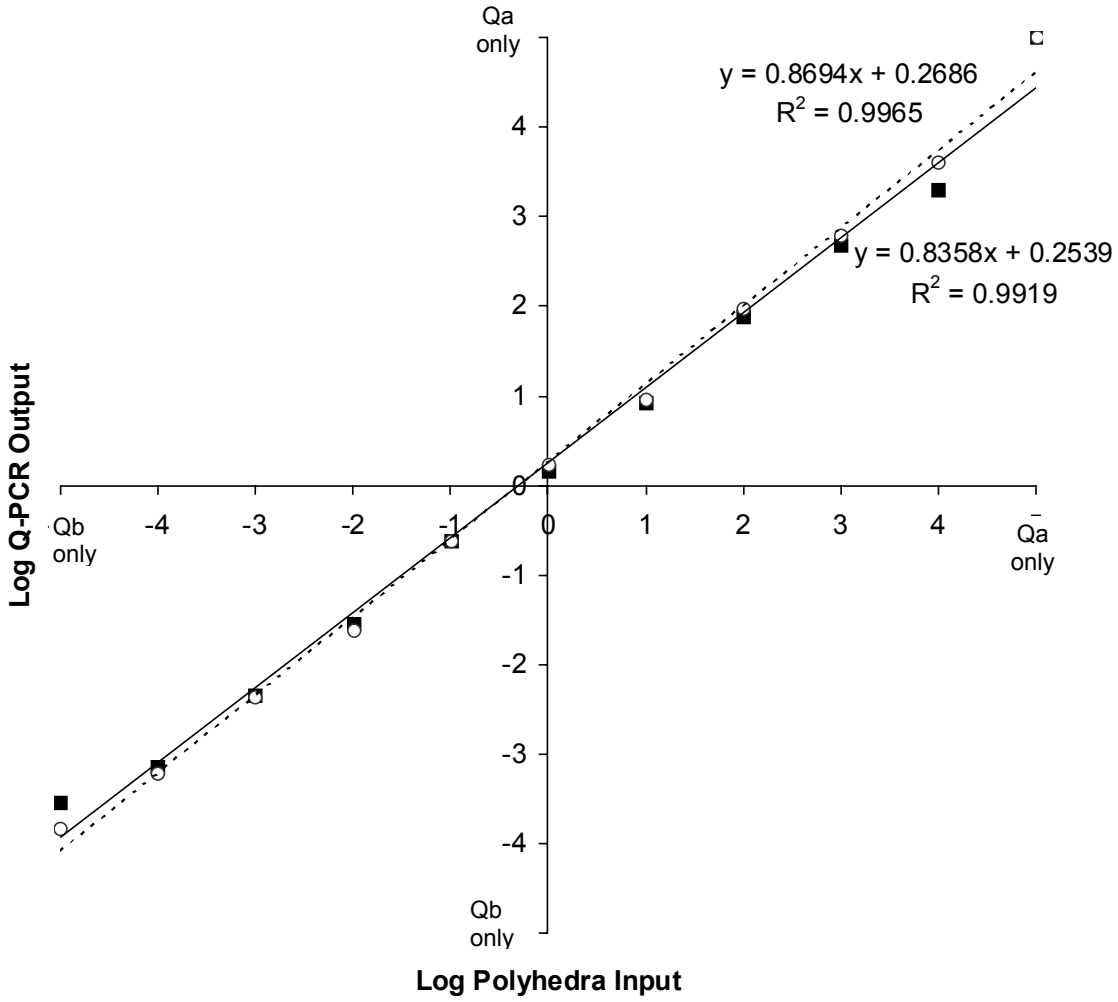


Figure 3: Graphical representation of the qPCR validation experiment. Polyhedra mixtures containing known ratios of vPolA and vPolB, as input, are indicated on the x-axis (Log Polyhedra Input). The qPCR output ratio is plotted on the y-axis (Log Q-PCR Output). Log indicates a $\log_{10}$ transformation. The black squares indicate the first replicate; the white circles indicate the second replicate. Note that the axis extends to infinity (Qa only) or -infinity (Qb only) beyond Log ± 4.

of the qPCR output ratio obtained is necessary for a valid result. This correction is obtained from the regression line; the measured qPCR ratio obtained is the y-value, and actual genotype ratio measured the corresponding x-value. Moreover, as the slope of the regression line is not one, the validation must be based on a range of ratios (e.g. it is not enough to determine the output with only a 1:1 mixture). A thorough validation of this type of qPCR assay, and in this case the need for a correction, is therefore essential.

3.5 qPCR background levels in a typical laboratory set-up

The level of background signal in a typical laboratory setup was investigated. Background means the vPolhA:VPolhB ratio when it is thought that only vPolhA or vPolhB is present. The source of the incorrect signal could be a lack of assay specificity, or contamination. NTCs were however always negative. This indicates that if background was caused by contamination, this must have occurred prior to the PCR procedure. Viral genotype ratios were determined for *S. exigua* larvae infected with a single genotype (see Figure 4). The results indicate that there are typically very low levels of background, allowing for quantification of NPV genotype ratios over a large range. No background was detected for larvae infected with pure vPolhA. For vPolhB, 1:11,000 (vPolhA:vPolhB) background was measured in 1 larva, whereas all others were negative for vPolhA.

The number of gene copies of the recognition sequence was compared to that of the essential gene *ie-1* by quantitative PCR. The constant ratio between recognition and *ie-1* sequences observed for passages 1 (transfected *T. ni)* to 3 (infected *S. exigua*) confirmed that the bacmids containing the recognition sequences were stable throughout the experiments performed (data not shown).

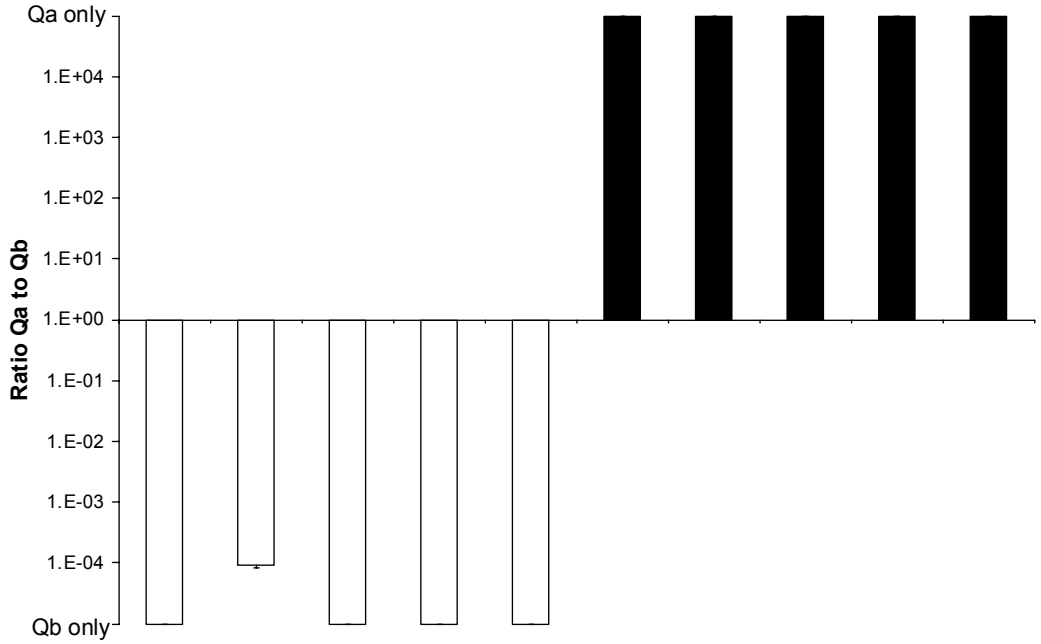


Figure 4: Log ratio of Qa:Qb obtained by qPCR for individual *S. exigua* larvae. Larvae were infected with either only vPolhB (white bars) or only vPolhA (black bars). The error bars denote the standard error of the ratio. Note that the axis extends to infinity (Qa only) or - infinity (Qb only) beyond Log ± 4.

3.6 Infection of insects with a high dose of a 1:1 mixture of vPolhA and vPolhB

In order for these bacmid-derived viruses and the qPCR assay to be useful tools for experiments in population biology, the fitness differences between the viruses must not lead to a rapid exclusion of one genotype by selection (e.g. vPolhA displacing vPolhB in a single cycle of infection). Therefore, genotype ratios were measured in larvae infected with a high dose (10^6 polyhedra/ml) of a 1:1 mixture of polyhedra (Figure 5). All larvae challenged with this dose of virus died. Given the miniscule background levels observed in the single genotype control larvae, it can be said with certainty that both genotypes were present in all larvae. A good deal of between-host variation in genotype frequencies was noted. As expected from the bioassay data (infectivity vPolhA > vPolhB), the mean of the log transformed data shifted towards the vPolhA genotype - as compared to the mixture used to initiate infection - although the difference was not significant (One-sample T-test, T = 1.801, d.f. = 9, Significance (2-tailed) = 0.105). It can be concluded that the difference in fitness between the two bacmid derived viruses – although present – does not form an impediment to carrying

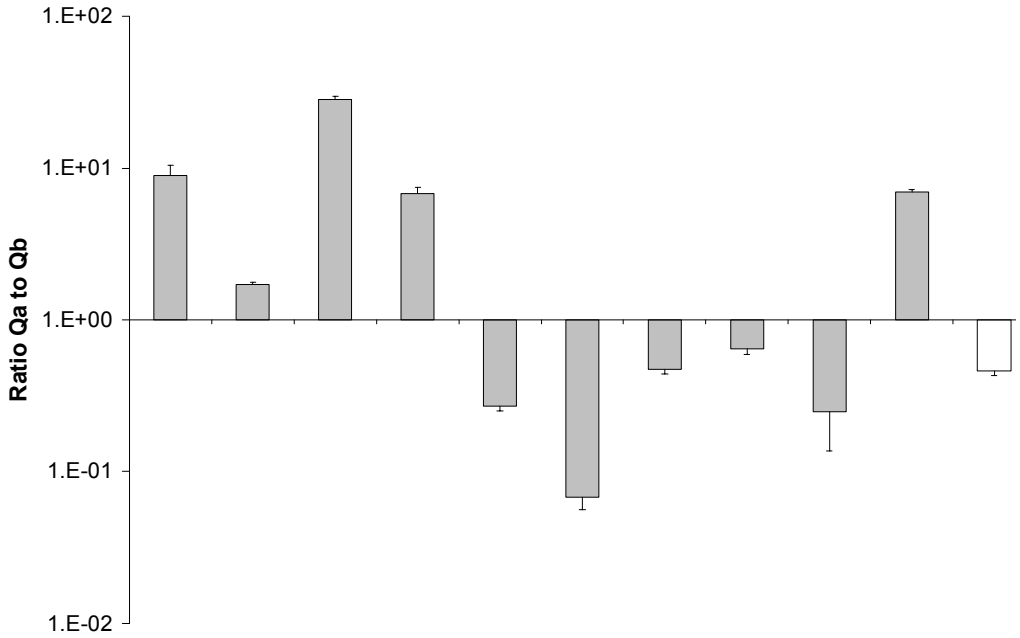


Figure 5: Co-infection of larvae with vPolhA and vPolhB. Log ratio of Qa:Qb obtained by qPCR for 10 *S. exigua* larvae infected with a 1:1 ratio of vPolhA and vPolhB polyhedra (gray bars). The white bar represents a direct qPCR on the input polyhedra mix (1:1) used in this experiment. The error bars denote the standard error of the ratio. Note that as the scale is logarithmic, values greater than 1 indicate that more vPolhA than vPolhB is present.

out studies of co-infection dynamics. This difference should, however, most likely be taken into account in analyzing the results of such studies.

4. Discussion

Two bacmid derived viruses designed to study NPV population genetics (vPolhA and vPolhB) – particularly genetic drift in mixed infections - were generated. These viruses had identical polyhedra morphology, in terms of number of nucleocapsids per ODV and number of ODV per polyhedron cross-section. Interestingly, it was noted that the number of nucleocapsids per ODV appears to be a truncated Poisson distribution (Figure 2). To model and understand NPV population genetics, this may be an important observation.

The identical polyhedra morphology allowed us to make mixtures of these polyhedra in order to validate a qPCR assay for determining the ratio of these viruses in a mixed genotype population. The results of this validation experiment indicated that the qPCR assay could quantify virus ratios over a range of eight orders of magnitude of genotype ratios (10,000:1-1:10,000). However, the experiment also demonstrated that a correction of the ratios generated was necessary to ensure the validity of the results. Analyzing the data by a comparative analysis ($r^2 = 0.9986$, see Figure 3) was superior to analysis by the standard curve method ($r^2 = 0.9781$, standard curve analysis of data not shown).

Polyhedra can be accurately counted and the morphology of polyhedra for the two genotypes used was shown to be very similar if not identical (Fig. 2); these unique conditions allow for a rigorous validation of the qPCR assay described in this paper. The method for quantification developed here therefore has important advantages over existing methods used for determining baculovirus genotype ratios, not only in terms of accuracy, but also in terms of the range of ratios that can be quantified and high-throughput. Although other validated real-time PCR-based assays for determining viral genotype frequencies have been described (e.g. Carrasco *et al.,* 2007), this is the first report concerning a qPCR-based assay for determining virus genotype ratios that has been subjected to validation by qPCR on mixtures of quantified viruses. Many assays for determining virus titers – without discrimination of different genotypes - have been validated thoroughly (e.g. Hitchman *et al.*, 2007). The method described in this report allows for the extension of qPCR-based methods into a multiple genotype situation to quantify genotype ratios.

It was chosen to develop a qPCR-based assay for determining virus ratios using SYBR

Green I fluorogenic dye. There appears to be no real consensus on whether DNA binding fluorophores (e.g. SYBR Green I) or 5' nuclease oligoprobes (e.g. Taqman probes) are superior for quantification, although some studies have made a direct comparison. SYBR Green I shows an either equal (Ponchel *et al.*, 2003, Maeda *et al.*, 2003) or superior (Schmittgen *et al.*, 2000) relative performance in quantification at intermediate DNA concentration ranges, whereas in some studies Taqman probes are claimed to have a higher specificity (Yin *et al.*, 2001, Ravva and Stanker, 2005). Studies in NPV population biology are often associated with the analyses of large numbers of samples. Therefore, the detection tools and application should be rapid, robust and economic. SYBR Green I, was therefore chosen and this appears to have worked quite well.

The qPCR assay described here exhibited low levels of background, specifically with the vPolhA primer set. Melting curve analysis did not reveal the presence of primer dimers, and non-template controls were always negative. Given that these background levels were extremely low - 1:11,000 was the highest background measured - it does not present a major problem in the quantification of virus genotype ratios to intermediate ratio ranges (1:1,000-1,000:1). However, beyond this range, genuine mixed genotype populations (vPolhA and vPolhB present) cannot be distinguished with absolute certainty from single genotype populations (vPolhA or vPolhB present) due to level of background seen in control samples. This limitation of this assay should always been given due consideration when the results are analyzed. This problem is, however, a corollary of high sensitivity of PCR-based assays; alternative techniques such as RFLP cannot even detect such extreme ratios.

The two bacmid derived viruses differed in their respective median lethal concentration (LC_{50}). It is unclear what causes this fitness difference as the genetic background (bacmid) of vPolhA and vPolhB is identical. One possibility could be the difference in the 100 bp recognition sequences. Although it cannot be excluded, this is highly unlikely since these sequences have been inserted in a non-coding locus, with no known promoter sequences (except for the polyhedrin promoter and ORF, which are cloned immediately downstream of these sequences). Transcription analysis of this region may give clues as to whether this is the case. Another possibility is that a deleterious mutation has occurred somewhere else in the viral genomic DNA of one of the bacmids during the construction and infection process. Given the large genome size (134 kbp), sequencing for genetic differences is not an attractive option. Others have also reported that generating marked genotypes with no fitness differences can be capricious (e.g. Sacristan, 2003).

Despite the fact that they differ slightly in fitness, vPolhA and vPolhB are useful tools

for exploring NPV population biology. Infecting larvae with a high dose (10 x LC_{50}) resulted in all larvae being infected with both genotypes (Figure 5). Although there was considerable between-host variation in the genotype frequencies, both viruses were still present in the population. A significant shift towards the vPolhA genotype in one infection cycle was not found. Hence, the fitness difference between the viruses is not such that the vPolhA genotype immediately displaces the vPolhB genotype. This means that these viruses can indeed be used to study co-infection dynamics, although the fitness differences must be taken into account (corrected for) when analyzing the data.

The between-host variation in the genotype frequencies observed in the small co-infection experiment is a form of genetic drift (see Figure 5). Given that the dose administered induced 100% mortality, this result is perhaps surprising. It is possibly an indication that, even at large virus doses resulting in high mortality, only a small number of virions are infection founders. One of the first issues to be addressed using the viruses and qPCR assay developed here is to determine the number of founders in an infection. The bacmid-derived viruses and sensitive qPCR provide the tools to conduct these exciting experiments.

The generation of two recombinant viruses and a SYBR Green I-based real-time PCR assay for population biology studies of NPVs are described here. Genotype ratios can be determined accurately in mixed populations, although the assay does have limitations in measuring extreme ratios (e.g. < 1:1,000). Viruses with purposely inserted non-coding recognition sequences were used in this study. SYBR Green I-based qPCR assays are in principle not limited to this approach. In other words, a qPCR assay could also be designed for two or more unmodified viruses – although there are then more constraints for primer design that could lead to reduced performance of primers. Recombinant viruses lend themselves best to this approach, as they will contain either exogenous sequences or 'junctions' over deleted endogenous genes that can be used as unique template sequences. For studies of NPV population biology, this qPCR assay is accurate, valid and amenable to high-throughput.

Acknowledgements

The authors thank Hanke Bloksma for invaluable assistance with the EM work. Maria Forlenza and Beja de Vries (Cell Biology and Immunology Group, Wageningen University) gave advice on setting up the qPCR assay, and granted the use their thermal cycler. Ioannis Papageorgiou constructed the pFastBac-DUAL/Polh/A vector. *T. ni* eggs were provided by Tim Carty (Centre for Ecology and Hydrology, Oxford, UK). *S. exigua* eggs were provided by Adolfo Hernandez (Certis, USA).

Chapter 3

Low multiplicity of infection *in vivo* results in purifying selection against baculovirus deletion mutants

---

Mark P. Zwart, Eloy Erro, Monique M. van Oers, J. Arjan G.M. de Visser, and Just M. Vlak

This chapter has been published as 'Low multiplicity of infection *in vivo* results in purifying selection against baculovirus deletion mutants' (2008), *Journal of General Virology* **89**: 1220-1224.

Abstract

The *in vivo* fate of *Autographa californica* multicapsid nucleopolyhedrovirus deletion mutants originating from serial passage in cell culture was investigated by passaging a population enriched in these mutants in insect larvae. The infectivity of polyhedra and ODV content per polyhedron were restored within two passages *in vivo*. The frequency of occurrence of deletion mutants was determined by real-time PCR. The frequency of the non-homologous region origin (non-HR *ori*) of DNA replication was reduced to wild-type levels within 2 passages. The frequency of the *polyhedrin* gene did not increase and remained below wild-type levels. Low multiplicity of infection (MOI) during the initial infection in insect larvae, causing strong purifying selection for autonomously replicating viruses, could explain these observations. The same virus population used *in vivo* was also passaged once at different MOIs in cell culture. A similar effect (i.e. lower non-HR *ori* frequency) was observed at low MOIs only, indicating that MOI was the key selective condition.

1. Introduction

The passage of baculoviruses in cultured insect cells leads to the rapid accumulation of deletion mutants, some of which have been demonstrated to have defective interfering properties (Pijlman *et al,*, 2001; Kool *et al.*, 1991; Wickham *et al.*, 1991). Populations with a high frequency of these deletion mutants can completely lose infectivity for insect larvae (Heldens *et al.*, 1996). The presence of such deletion mutants is therefore likely to be deleterious to the fitness of a virus population *in vivo*, although deletion mutants are found in some natural baculovirus isolates which can increase fitness (Lopez-Ferber *et al.*, 2003). Dai *et al.* (2000) found that alternate passaging in insect cells and larvae resulted in sustained infectivity for insect larvae, suggesting that purifying selection (Li, 1997) occurs in larvae. In other words, there may be stabilizing selection for a particular trait, in this case the ability to replicate autonomously. Consequently, individuals which do not posses this trait are removed from the population. However, the fate and dynamics of these cell culture-derived deletion mutants after reintroduction into insects have not been systematically addressed. Therefore, a clonal baculovirus was first serially passaged in insect cells to enrich for deletion mutants, and the resulting population was reintroduced to insect larvae. The effects of both *in vitro* and subsequent *in vivo* passaging were investigated by determining the virulence and occlusion derived virus (ODV) content of polyhedra, and the frequency of occurrence of deletion mutants by quantitative real-time PCR (qPCR).

2. Results

In order to generate a clonal population of *Autographa californica* multicapsid nucleopolyhedrovirus (AcMNPV), bacmid technology (Luckow *et al.*, 1993) was employed. A bacmid with restored polyhedrin expression was generated using the pFastBac-DUAL/Polh construct (Zwart *et al.*, 2008). This bacmid was used to transfect *Spodoptera frugiperda* (Sf-AE-21) cells (Vaughn *et al.*, 1977) according to Pijlman *et al.* (2001). The virus resulting from transfection of Sf21 cells was then passaged serially with minimal dilution (1:4) in Sf21 cells for 20 passages (final population: $P_{Cell}20$). RFLP analysis revealed that viruses with large genomic deletions appeared in the population starting at $P_{Cell}10$ as shown previously (Kool *et al.*, 1991).

Viruses of passages $P_{Cell}2$, $P_{Cell}5$, $P_{Cell}10$ and $P_{Cell}20$ were subsequently amplified in Sf21 cells. After 96 h, the cells where detached by agitation and sedimented by centrifugation (5 min at 2500 g). The supernatant was removed and stored at 4° C, and subsequently used as budded virus preparation (BV- $P_{Cell}20$). The pellet was resuspended in 5 ml 1% Triton X-100, and the cells were lysed by sonification. The resulting suspension was layered onto a 5 ml 30% sucrose (in 0.2% Triton X-100) cushion, and polyhedra were pelleted by centrifugation (30 min at 15,000 g at 4° C). The pellet was resuspended in 10 ml milli-Q water, pelleted (5 min centrifugation at 2500 g), and finally resuspended in 500 µl milli-Q water and stored at 4° C.

The median lethal concentration (LC_{50}) was determined by a droplet feeding bioassay (Hughes and Wood, 1981) for polyhedra of $P_{Cell}2$, $P_{Cell}5$, $P_{Cell}10$ and $P_{Cell}20$. *S. exigua* larvae were reared as described by Smits *et al.* (1986). Twenty-four newly molted *S. exigua* L4 larvae were challenged per concentration, and 24 larvae were taken as healthy controls. Ten-fold dilutions, in sterile distilled water, of polyhedra (10^6 - 10^2 polyhedra/ml) were used and the bioassay was performed *in triplo*. Probit analysis was employed to determine the LC_{50}. Statistical tests were performed with SPSS 12.0 unless specified otherwise. Linear regression demonstrated that there was a significant increase in LC_{50} over passages (Figure 1A; $F_{1,10} = 6.937$, $P = 0.025$). Unlike *S. exigua* MNPV (Heldens *et al.*, 1996), passaged populations of AcMNPV retain infectivity – albeit reduced - through passage 20. It was thus possible to use polyhedra of the $P_{Cell}20$ population to reinfect insect larvae.

Polyhedra of $P_{Cell}2$, $P_{Cell}10$ and $P_{Cell}20$ were also analyzed by electron microscopy (EM) as described by Zwart *et al.* (2008). The number of ODV per polyhedron cross-section was counted, for minimally 30 polyhedra per sample. There was a significant

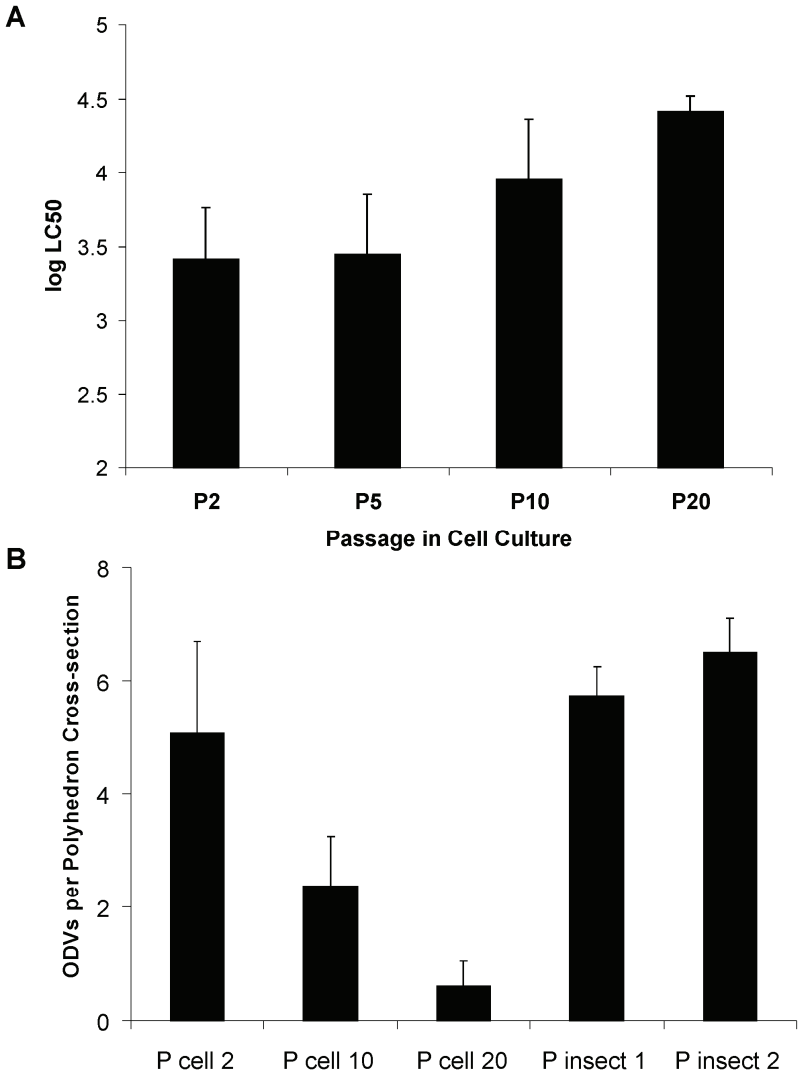


Figure 1: A. The $\log_{10}$ LC_{50} of the virus over passages in cell culture, with standard error of the mean. B. The mean number of ODV per polyhedron cross-section for passaging in Sf21 cells (P cell 2, 10, 20) and subsequent passaging of $P_{cell}20$ in *S. exigua* larvae at an LC_{99} (P insect 1, 2). Error bars indicate the standard error of the mean.

drop in ODV per cross-section over passages (Figure 1B; Jonckheere-Terpstra Test, N = 151, Standardized JT-statistic = -7.237, $P < 0.001$). By $P_{Cell}20$, most polyhedra no longer contained ODV.

DNA purification from polyhedra and qPCR were performed as described by Zwart *et al.* (2008) to determine the frequency of occurrence of deletion mutants. Primers used for amplification of the immediate early gene *ie-1* (Kovacs *et al.*, 1992) (Figure 2A) were described previously (Zwart *et al.*, 2008). *Ie-1*, an early transcriptional regulator, is essential for replication and was expected to be relatively stable over passaging and therefore a suitable reference locus. Primers for the non-homologous region origin (non-HR *ori*) of DNA replication (forward 5'-CCGAGACATACCACAAAGCCG-3', reverse 5'-GCACATAAACGACGCAGAATACAT-3') were also used, as this locus

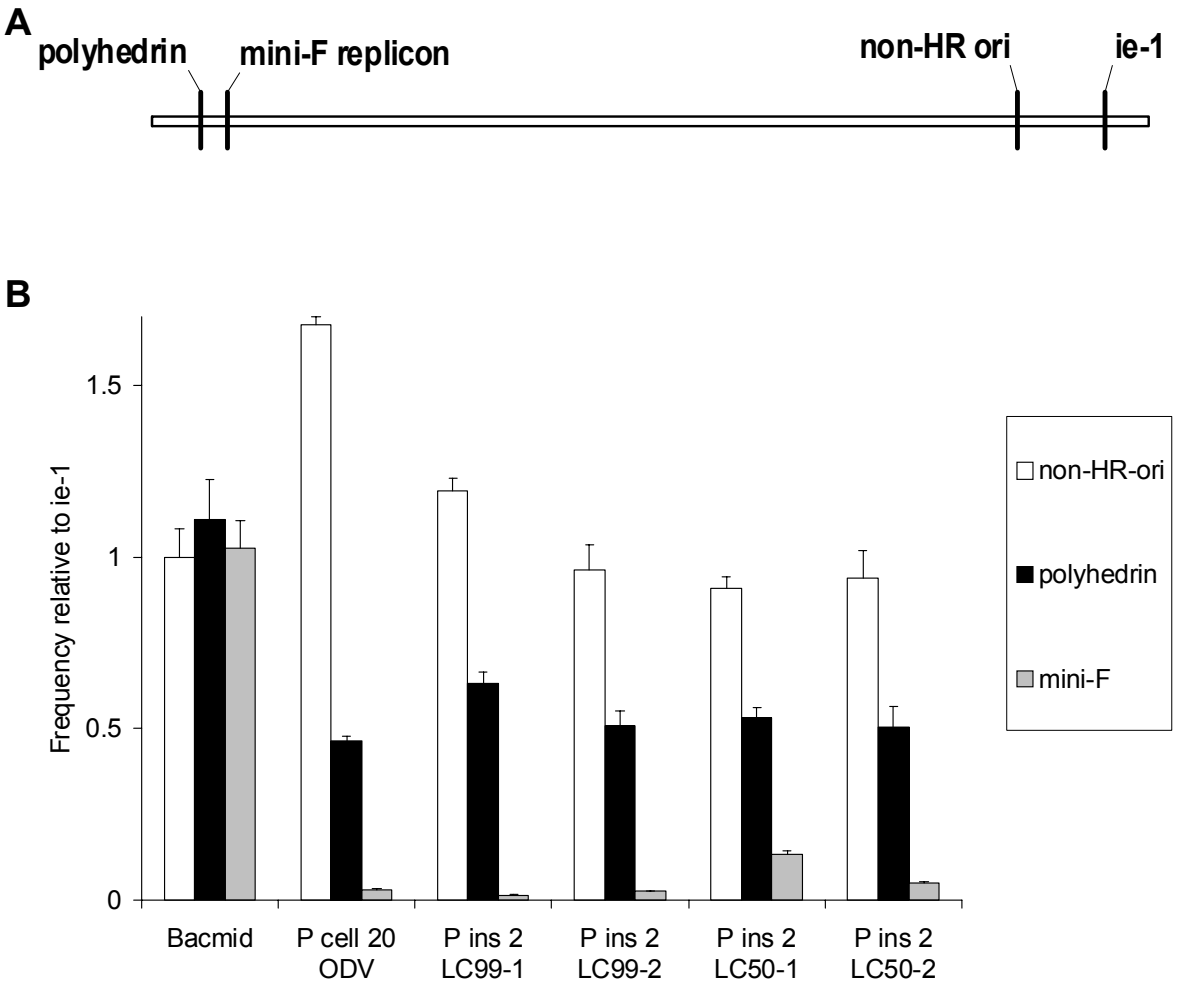


Figure 2: A. The location of the loci used for qPCR in the bacmid. B. Loci frequencies, relative to the *ie-1* locus (y-axis), in the original bacmid DNA, and DNA isolated from ODV of $P_{cell}20$ and the $P_{cell}20$ population after passaging in insects at different doses (P ins 2). Error bars indicate the standard error of the mean.

has been found to become enriched during passaging (Pijlman *et al.,* 2002). Primers for *polyhedrin* (Smith *et al.*, 1983) (forward 5'-GGGTGGGCAGCAACAACGAGTA-3', reverse 5'-CCGATGTAAACGATGGGCTTGTAG-3') were included to test for the presence of the gene. No selection for this gene was expected *in vitro* as BV was passaged, while *in vivo* there will be weak selection for polyhedrin (Bull *et al.*, 2001). Primers for the mini-F replicon in the bacmid insert (Luckow *et al.*, 1993) (forward 5'-CTTTACGACGGCGACTCCCATC-3', reverse 5'-GCTTACTGAGGACGCACTGG ATG-3') were also used, as this locus is not expected to be under selection in either insect cells or larvae. Ratios of the individual loci to *ie-1* and their corresponding standard errors were calculated as described by Zwart *et al.* (2008), rendering a locus frequency relative to *ie-1* (referred to as frequency).

qPCR analysis demonstrated that the frequency of the non-HR *ori* had increased to

1.68 by $P_{Cell}20$ (Figure 2B). The frequency of polyhedrin decreased to 0.47. The mini-F replicon decreased in frequency to 0.03, probably also due in part to instability of a bacmid insert (Pijlman *et al.*, 2003). The $P_{Cell}20$ virus population thus has an unbalanced gene content, as confirmed by qPCR, and has polyhedra with fewer ODV and lower infectivity.

Amplified polyhedra from the $P_{Cell}20$ where used to perform two passages in *S. exigua* L4 larva (final population: $P_{Insect}2$). To this aim insects were challenged with an LC_{50} (2×10^4 polyhedra/ml, final population: $P_{Insect}2\text{-}LC_{50}$) and LC_{99} (6×10^7 polyhedra/ml, final population: $P_{Insect}2\text{-}LC_{99}$), as calculated from the bioassay data for polyhedra from $P_{Cell}20$. Two replicates were performed for each concentration, with polyhedra being pooled from 10 larvae per replicate. Polyhedra of $P_{Insect}1\text{-}LC_{99}$ and $P_{Insect}2\text{-}LC_{99}$ (replicate 1) were analyzed by EM as described above. The LC_{50} of $P_{Insect}2\text{-}LC_{99}$ (replicate 1) was also determined by bioassay as described above. For comparative purposes, $P_{Cell}2$ was also included in the bio-assay. DNA isolation and qPCR - as described above - were performed for all four $P_{Insect}2$ populations.

After 2 passages in insects at high dose ($P_{Insect}2\text{-}LC_{99}$, replicate 1), the infectivity of polyhedra was restored to control values (mean log LC_{50} ±SE: $P_{Cell}2$ = 3.754 ±0.064, $P_{insect}2\text{-}LC_{99}$ = 3.963 ±0.086; Pairwise *t*-test: $t_2 = -2.392$, $P = 0.139$). The ODV content of polyhedra from this population ($P_{Insect}2\text{-}LC_{99}$, replicate 1) was also restored to wild-type (Figure 1B). The observed number of virions per polyhedron cross-section was not significantly different from that in polyhedra from larvae transfected with the original bacmid (mean ±SE: $P_{Insect}1\text{-}LC_{99}$ = 5.74 ±0.58, Bacmid *in vivo* = 7.60 ±1.17; Mann-Whitney *U*-test, $U = 1568.5$, N = 123, $P = 0.168$).

To analyze the qPCR data on the $P_{Insect}2$ populations (Figure 2), a two-way ANOVA (Genstat 7.0) was performed to test if there was an effect of dose (LC_{50} or LC_{99}) on the frequency of the three loci. Both dose ($F_{1,8} = 0.87$, $P = 0.397$) and the dose*locus interaction ($F_{2,8} = 0.77$, $P = 0.554$) did not have a significant effect, whereas locus did ($F_{2,8} = 72.71$, $P < 0.001$). Since there is not a significant effect of dose on the frequency of the three loci, locus frequencies for the four populations were compared to that in the $P_{cell}20$ by means of a one-sample *t*-test. For the non-HR *ori* there was a significant decrease in frequency ($t_3 = -10.491$, $P = 0.002$). Moreover, the mean $P_{Insect}2$ frequency of the non-HR *ori* was not significantly different from the wild-type frequency ($t_3 = 0.006$, $P = 0.996$). The mean increase in *polyhedrin* frequency (0.08) was not significant ($t_3 = 2.717$, $P = 0.073$). The mini-F replicon also did not significantly increase in frequency ($t_3 = .957$, $P = 0.409$).

The 3000-fold difference in dose between the LC_{50} and LC_{99} treatments did not have a significant effect on changes in locus frequencies. If the number of initially infecting

viruses is proportional to the dose, then this could be an indication that the number of initially infecting viruses was very small, even at the high dose. In other words, despite the higher number of viruses initiating infection at high dose, the initial *in vivo* MOI was still so low it had effectively not changed. This is another indication that the number of founders of infection is small, as has previously been suggested (Smith and Crook, 1988).

Deletion mutants missing *ie-1*, or with multiple copies of the non-HR *ori*, were rapidly purged from the populations during passaging *in vivo* (Figure 2). During continuous bioreactor experiments with baculoviruses, cyclic fluctuations in virus titers similar to those described by Von Magnus (1954) have been observed (van Lier *et al.*, 1992). The explanation proposed for this phenomenon is that following a strong drop in virus titers due to defective interference, viruses capable of autonomous replication are enriched at low multiplicities of infection (MOI, de Gooijer *et al.*, 1992). The fact that infection of larvae is initiated by a small number of viruses will similarly lead to low MOIs during initial infection. Thus, the purifying selection observed *in vivo* may come about mainly because of these low MOIs. This is not to suggest that there is no selection for the presence of, for example, *polyhedrin* during passaging in insect larvae, but rather that the predominant selective condition resulting in the purifying selection observed is the ability to autonomously replicate.

In order to test whether this explanation has merit for the rescue effect observed by passaging in larvae, the titer of BV-P_{Cell}20 was determined by an end point dilution assay and one round of infection of Sf21 cells was performed at MOIs 10, 1, 0.1, and 0.01. BV was purified, DNA isolated and qPCR performed as described above. The experiment was performed *in duplo*. Linear regression analysis was used to see if locus frequencies changed as the MOI decreased. For the non-HR *ori* there was a significant decrease in frequency ($F_{1,6}$ = 21.937, P = 0.0034), while for *polyhedrin* ($F_{1,6}$ = 0.117, P = 0.7445) and the mini-F replicon ($F_{1,6}$ = 0.1750, P = 0.6903), no change in frequency was observed (Figure 3A). It was therefore possible to mimic *in vitro* the purifying selection observed *in vivo* simply by altering the MOI. This is in agreement with results obtained by van Lier (1996) in bioreactor systems: it was empirically shown that low MOIs resulted in the slower accumulation of defective viruses. The BV-P_{Cell}20 sample and a number of time points during passaging were analyzed by qPCR, to confirm that the ancestral virus (BV-P_{Cell}0) had the same profile as the bacmid and to determine when deletion mutants could first be detected (Figure 3B). Changes in locus frequencies were only visible around P12, in agreement with RFLP data (not shown).

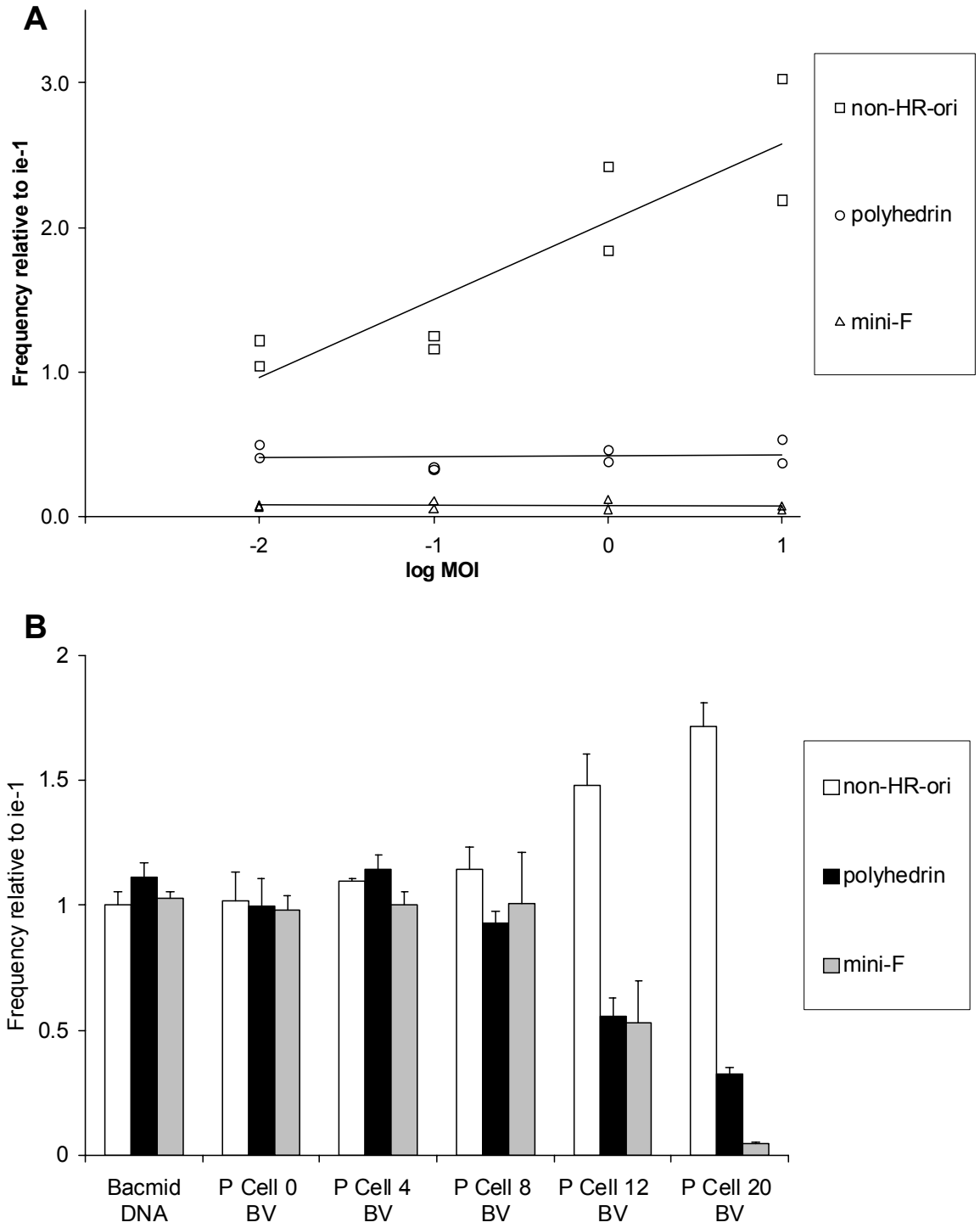


Figure 3: A. Loci frequencies, relative to the *ie-1* locus (y-axis), after a single passage in Sf21 cells (log_{10} MOI on x-axis) are shown. Note that this analysis was performed on DNA obtained from BV collected after a single passage of $P_{cell}20$ BV at different MOIs. B. Loci frequencies relative to the *ie-1* locus are given, with standard errors of the mean. The bacmid DNA, BV of the ancestral virus after transfection ($BV_{cell}0$), and BV of passages 4, 8, 12 and 20 ($BV_{cell}4$ - $BV_{cell}20$) were analyzed.

3. Discussion

In the experiments described here, *polyhedrin* did not increase significantly in frequency by $P_{Insect}2$. Bull *et al.* (2001) found that a *polyhedrin* negative virus persisted, in co-infection with a *polyhedrin* positive virus, for six serial transfers in insect larvae. The *polyhedrin* negative virus needs to co-infect a cell with a *polyhedrin* positive virus in order to produce polyhedra and be carried over to the next round of infection. Since most cells in the insect will be infected during the final round of infection, Bull *et al.* reason that the rate at which a *polyhedrin* negative virus is lost from a population depends directly on MOI in the final found of cellular infection. So presumably there was not strong selection against the *polyhedrin* negative virus because of a high *in vivo* MOI at the end of infection. ODV of the $P_{Cell}20$ population has an observed *polyhedrin* frequency (f_{virus}) of 0.463. If the frequency of *polyhedrin* is conserved during infection, than in the final wave of infection, the frequency that cells will be infected by at least one BV containing *polyhedrin* in its genome (f_{cell}) is:

1. $$f_{cell} = 1 - (1 - f_{virus})^m$$

Here is m is the *in vivo* MOI, estimated at 4.3 BV per cell (Bull *et al.*, 2001, 2003). This results in a f_{cell} of 93%. Since most cells infected in the final wave of infection have been infected by a BV containing the *polyhedrin* gene, it can be understood that selection is weak.

Contrary to *polyhedrin* locus frequency, ODV content of polyhedra was rescued to wild type levels within a single passage ($P_{Insect}2$-LC_{99}, replicate 1). One explanation for this discrepancy is that the presence of deletion mutants, for example those missing *ie-1*, is somehow interfering with the generation of polyhedra. The relative frequency of *polyhedrin* to non-HR *ori* - rather than *ie*-1 - gives an indication of the frequency of viruses capable of generating polyhedra to all viruses in the population, not only those containing *ie-1*. The frequency of *polyhedrin* as compared to the non-HR *ori* rises from 0.28 ($P_{Cell}2$) to 0.54 (mean $P_{Insect}2$) during passaging *in vivo*. Another compatible explanation for this observation is that the purifying selection observed results in a critical increase in frequency of another gene essential for generation of normal polyhedra (e.g. *25k*).

During baculovirus infection of insect larvae, the *in vivo* MOI changes drastically. Since the number of viruses initiating infection appears to be small (e.g. Smith and

Crook, 1988), initially the MOI is extremely low. Towards the end of infection the MOI is rather high (Bull *et al.*, 2001, 2003). These different conditions will result in different selection pressures on the virus population during the course of infection. Genes involved in replication (*ie-1*) or transmission (*polyhedrin*) will also be under different selection pressures. There will be a very strong selection on genes involved in replication early in larval infection. Later in infection, selection is somewhat relaxed as MOIs become high - if the presence of a single gene copy is sufficient for cellular infection to normally proceed. Conversely, there will not be selection for genes involved in transmission early in infection. Later in infection there will be moderate selection for these genes, but not of the strong purifying type seen for replication early in infection. The fate of baculovirus deletion mutants derived from tissue culture and reintroduced into insect larvae illustrates a virological principle that may apply to many other virological systems.

Acknowledgements:

Thanks to Jan van Lent and Hanke Bloksma for assistance with EM, and Gorben Pijlman for invaluable discussion. This project was partially supported by EU grant LSHB-CT-2006-037541. MZ was supported by a strategic Ph.D. grant from the C.T. de Wit Graduate School for “Production Ecology and Resource Conservation”.

Chapter 4

An experimental test of the independent action hypothesis in virus-insect pathosystems

Mark P. Zwart, Lia Hemerik, Jenny S. Cory, J. Arjan G.M. de Visser, Felix J.J.A. Bianchi, Monique M. van Oers, Just M. Vlak, Rolf F. Hoekstra, and Wopke van der Werf

A modified version of this chapter has been submitted for publication.

Abstract

The 'independent action hypothesis' (IAH) - postulated in 1952 - states that the action of single pathogen individual can lead to host death and that pathogen individuals act independently. IAH was never rigorously tested or demonstrated due to lack of adequate experimental approaches and of a theoretical framework for making predictions. In this paper we (1) develop a parsimonious probabilistic framework for testing IAH, and (2) demonstrate that, in two out of six virus-insect pathosystems tested, IAH gives pertinent predictions for the number of invading pathogens. We first show that IAH inextricably links host survivorship to the number of invading pathogen entities, and then develop a model to predict the frequency of single and dual genotype infections when a host is challenged with a mixture of two genotypes. Model predictions were tested using near genetically marked, near identical baculovirus genotypes and insect larvae from three host species, both early and late in larval development. IAH predictions were not significantly different from experimental observations in early instar larvae of susceptible host species (*Trichoplusia ni* and *Spodoptera exigua*). Predictions were not in agreement with the data for more resistant late instar larvae of susceptible host species, and larvae of a less susceptible host species (*Mamestra brassicae*). The model therefore gives good estimates of the number of invaders in pathosystems characterized by high host susceptibility only, implying that in these systems a single virus individual can invade the host and cause disease.

1. Introduction

In the vast majority of pathosystems it is unknown how many pathogen individuals have invaded a host and caused disease. Estimates of this number are a prerequisite for making predictions on the likelihood that diseased hosts were invaded by multiple genotypes, genetic drift (Chao, 1990), the evolution of pathogen genotypes (Cooper *et al.*, 2003), and on interactions between pathogen genotypes within infected hosts. These interactions include competition (Turner and Chao, 1999), complementation (Vignuzzi *et al.*, 2006) and recombination (Gibbs *et al*., 2001). For viruses there is experimental evidence that the number of individuals initiating an infection can be small. This conclusion is supported by analysing patterns of genetic drift (Smith and Crook; 1988; Ali *et al.*; 2006) or attempting to directly count the number of initially

invaded cells (Smith *et al.*, 2008). The notion that the number of virions initiating disease can be small is also supported by median lethal dose (LD_{50}) values obtained from challenge experiments, where the biological activity of the inoculum has been determined in cell culture (e.g. the number of plack forming units (PFU); e.g. Spieker *et al.*, 1996)[1]. There is, however, no framework for making predictions on the number of virus individuals initiating an infection. This is a significant gap in our knowledge, given the importance of the number of individuals initiating infection for pathogen population dynamics and evolution.

The independent action hypothesis (IAH) was originally formulated to explain dose-response data for mammals exposed to anthrax spores by inhalation, as IAH leads to a distinct dose-response relationship (Druett, 1952). IAH states that each pathogen individual has a non-zero probability of invading the host and causing disease, and that conspecific pathogen entities do not affect each other during this process (Druett, 1952). IAH therefore concerns the entire sequence of pathogen related events leading to host disease: from initial pathogen invasion of the host until the establishment of host disease. A rigorous test of IAH has not been reported in any system. What is understood by host disease varies between pathosystems. Since the pathogen used in our experimental system - a baculovirus - must kill its host in order to achieve horizontal transmission, we define host disease as pathogen-induced host death (Federici, 1997). We subsequently refer to pathogen-induced host death simply as death, since in our experimental system these are the only deaths we have to consider; the insect larvae used are free of other pathogens and attrition in our experimental setup is negligible.

IAH offers an important initial null hypothesis of a pathogen's mode of action and perhaps has wide applicability. It is inherently a limited hypothesis; if it is rejected, alternative hypotheses could take into account complications, such as interactions between individuals of different genotypes, which could conceivably occur at all levels leading ultimately to host disease. Complementation between genotypes, for example, has been previously shown to be important for understanding how viruses cause disease (e.g. Vignuzzi *et al.*, 2006). What IAH postulates – and what we here attempt to test – is on a more fundamental level: that conspecific pathogen individuals act independently of one another in invading the host and causing disease. In other words, if two identical pathogen individuals challenge a host, the probability that each

[1] Note that since the LD50 obtained is based on biological activity, it does not reflect the actual number of virions initiating disease E.g. Spieker *et al.* found a LD_{50} of <.01 PFU. This indicates that this is not a measure of the actual number of virions invading, since invasion must entail the entry of at least one virion.

pathogen individual brings about host disease can be (1) bigger: there is complementation between individuals (as opposed to different genotypes), (2) smaller: there is antagonism between individuals, or (3) unchanged: there is no added benefit or cost to the number of pathogen individuals present, as each acts on its own. The last case corresponds to IAH, since the probability that disease is caused by the cumulative action of all individuals challenging a host follows from the number of individuals present, each having a fixed, independent chance of causing disease (assuming no differences in pathogen dose and host susceptibility).

In this paper we show that the number of pathogen individuals that has invaded a diseased host can be predicted under IAH, and this extends the utility of the theory beyond making predictions on dose-response relationships. The prediction for the number of individuals invading the host follows from the proportion of hosts that have not been diseased during a pathogen challenge experiment (i.e. host survival in our pathosystem). Hence, if the predictions made by IAH can be corroborated for a particular pathosystem, IAH is a powerful tool. Knowing whether or not IAH renders useful estimates may be of great importance to understanding the population genetics of a particular pathogen: under IAH there will be a high frequency of single genotype infection when host mortality is moderate to low (< 50%). Note that this prediction does assume a homogenous host population exposed to the same pathogen dose, and as such its application beyond laboratory settings is circumspect. But it is not just the method for estimating the number of founders which may be relevant. Considering the possibility that host disease initiated by a single or small number of individuals can occur on a regular basis - if hosts typically encounter low doses - and its many ramifications, may be of great importance. Given the association between host survival and the number of pathogens invading hosts which become diseased, we devised a probabilistic framework for the expected frequency of dual genotype invasions when hosts are challenged with a pathogen population containing two genotypes. This allows for a test of the pertinence of IAH predictions.

To test whether the IAH model gives pertinent estimates of the number of pathogens invading a host, we used a virus-insect system: *Autographa californica* multicapsid nucleopolyhedrovirus (AcMNPV, an Alphabaculovirus) genotypes and their larval insect hosts. Note that in this system we understand virions to be 'pathogen individuals'. This system was chosen because of the relative ease with which marked conspecific genotypes can be generated (Luckow *et al.*, 1993) and the availability of different hosts and developmental stages with different susceptibility to AcMNPV (Payne, 1986; Possee *et al.*, 1993). We previously described baculovirus genotypes that are suitable for testing IAH (Zwart *et al.*, 2008). These genotypes - from here on

named A and B - are launched from bacterial artificial chromosomes encoding the entire AcMNPV genome (Luckow *et al.*, 1993), and differ only in short (100 bp, > 0.1% of the AcMNPV genome) marker sequences in a non-coding region. These marker sequences allow for detecting genotype presence by means of a sensitive quantitative real-time PCR (qPCR) based assay (Zwart *et al.*, 2008). Here we demonstrate that the IAH model renders pertinent host survival-based estimates of the frequency of dual genotype invasions in two out of six pathosystems investigated.

2. Materials and Methods

2.1 The frequency of dual genotype infection under IAH

Under IAH pathogens act in a strictly independent manner (Druett, 1952). The mean number of invading pathogens (λ) for a homogenous population or for any given genotype in a mixed population will therefore be:

(1) $$\lambda = n \cdot p$$

Here n is the number of pathogens the host is challenged with (e.g. the number of pathogens ingested by the host, as opposed to the number that actually invades the host) and p the probability of invasion for each pathogen individual[2]. Typically, p will be very small, requiring a large n to produce even a low numbers of invaders. If the probability of invasion for each pathogen individual is independent of that of the others, it follows that the number of invading pathogens per host will follow a Poisson distribution with mean λ (Olkin *et al.*, 1994). Under IAH, the probability that a pathogen individual which has invaded the host will subsequently induce full-blown systemic infection and ultimately death is one. In other words, even if only one single successful invasion event occurs then the host will die. The probability of host survival

[2] It was originally suggested that for a pathogen organism invasion consisted of reaching a propitious site and proliferation (Druett, 1952). NPVs have two different viral forms that are genetically identical: occlusion–derived virus (ODV, in polyhedra) for spreading the infection from host to host and budded virus (BV) for spreading the infection within the host from cell to cell (Federici, 1997). In applying IAH to a baculovirus we propose that invasion requires (1) an ODV to access a midgut cell and (2) production of BV which have access to other interior host tissues. In other words, the host must be colonized by the virus.

after a pathogen challenge (*S*, i.e. fraction surviving) is then the zero-term of the distribution of the number of invading pathogens:

(2) $$S = e^{-\lambda}$$

IAH thus inextricably links the number of invading pathogens to host survival. Importantly, this association is independent of exact n and p values, which cannot always be experimentally determined.

Technically it would be difficult, if not impossible, to empirically determine the number of invading pathogen entities in a host after a challenge. However, if the host is challenged with a pathogen population consisting of multiple genotypes, then the distribution of genotype frequencies in individual dead hosts results from genetic drift and will depend on the number of invading pathogens. Assume a population entirely comprised of pathogen individuals of genotype A and B with invasion chances p_{A} and p_{B}, respectively. If the challenge dose is n_{A} for genotype A and n_{B} for genotype B, then under IAH $\lambda_{\mathrm{A}} = p_{\mathrm{A}} n_{\mathrm{A}}$ and $\lambda_{\mathrm{B}} = p_{\mathrm{B}} n_{\mathrm{B}}$. The probability distribution of the number of pathogens invading the host under IAH is Poisson with mean $\lambda_{\mathrm{A}} + \lambda_{\mathrm{B}}$:

(3) $$P(K = k) = \frac{(\lambda_{\mathrm{A}} + \lambda_{\mathrm{B}})^k}{k!} e^{-\lambda_{\mathrm{A}} - \lambda_{\mathrm{B}}}$$

Here *k* is the number of invading pathogen entities. The probability that only genotype A will be present in a host, $P(\mathrm{A} \cap \overline{\mathrm{B}})$ is:

(4) $$P(\mathrm{A} \cap \overline{\mathrm{B}}) = \sum_{i=1}^{k} P(\mathrm{A} \cap \overline{\mathrm{B}} \mid K = i) P(K = i) = \sum_{i=1}^{\infty} \left(\frac{\lambda_{\mathrm{A}}}{\lambda_{\mathrm{A}} + \lambda_{\mathrm{B}}} \right)^i \frac{(\lambda_{\mathrm{A}} + \lambda_{\mathrm{B}})^i}{i!} e^{-\lambda_{\mathrm{A}} - \lambda_{\mathrm{B}}}$$

This can be simplified to give:

(5) $$P(\mathrm{A} \cap \overline{\mathrm{B}}) = (1 - e^{-\lambda_{\mathrm{A}}}) e^{-\lambda_{\mathrm{B}}}$$

Likewise, $P(\overline{\mathrm{A}} \cap \mathrm{B}) = e^{-\lambda_{\mathrm{A}}} \left(1 - e^{-\lambda_{\mathrm{B}}}\right)$, and the probability that both genotypes are present, $P(\mathrm{A} \cap \mathrm{B})$, is:

(6) $$P(\mathrm{A} \cap \mathrm{B}) = P(\mathrm{A})P(\mathrm{B}) = \left(1 - e^{-\lambda_{\mathrm{A}}}\right)\left(1 - e^{-\lambda_{\mathrm{B}}}\right)$$

The expected proportion of dual genotype invaded cadavers is obtained by dividing $P(\mathrm{A} \cap \mathrm{B})$ by the probability of infection $\left(1 - P(\bar{\mathrm{A}} \cap \bar{\mathrm{B}})\right)$:

(7) $$P'(\mathrm{A} \cap \mathrm{B}) = \frac{P(\mathrm{A} \cap \mathrm{B})}{1 - P(\bar{\mathrm{A}} \cap \bar{\mathrm{B}})} = \frac{\left(1 - e^{-\lambda_{\mathrm{A}}}\right)\left(1 - e^{-\lambda_{\mathrm{B}}}\right)}{1 - e^{-\lambda_{\mathrm{A}} - \lambda_{\mathrm{B}}}}$$

A Venn diagram (Fig. 1) illustrates the proportions of challenged hosts that contain no pathogens, one pathogen genotype (A or B) or both pathogens of both genotypes, explaining in a non-mathematical guise the rationale of above formulas. Given (1) host survival, (2) the frequencies of genotypes in the original pathogen population, and (3) the invasion probabilities of the pathogen genotypes, IAH permits predictions of the expected frequency of single and dual genotype invaded hosts (Figure 2). This allows for an unambiguous test of IAH, by comparing the experimentally observed frequency of dual genotype invasion with the model prediction.

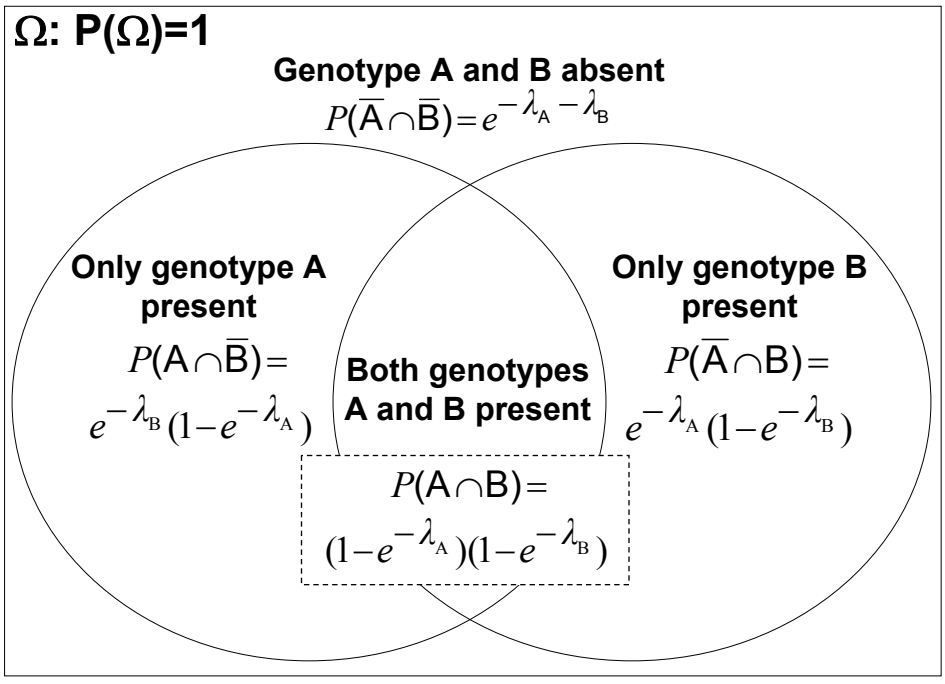


Figure 1: The IAH-based model for the frequency of dual genotype invasions; it is summarized in a Venn diagram, where Ω is the set of all possible outcomes.

(b) Challenge experiments

All challenge experiments and qPCR analyses were basically performed as previously described (Zwart *et al*., 2008). Briefly, developmentally synchronous larvae were starved overnight and subsequently allowed to drink from small droplets of a polyhedra suspension with a given concentration (minimally 24 larvae per dose). The larvae were then reared separately on semi-synthetic diet and upon death collected and individually stored. DNA was isolated from polyhedra present in the cadavers of individual larvae. A separate qPCR reaction was then performed to determine the presence of genotypes A and B in the DNA sample.

We assume there is not variation in ingested dose since this does not impact the dose-response relationship very strongly (Ridout *et al.*, 1993). Hence polyhedra concentration in the inoculum can be used as a proxy for dose. For determining the dose-response relationship for *S. exigua* L3, different ranges of polyhedra concentrations were taken for different replicates (4 replicates in total). The total number of replicates taken was (log concentration:total number of replicates) 2:1, 3:3:, 4:4, 5:4, 6:3 and 7:1. At least 24 larvae were taken per dose per replicate, and 24 larvae were also taken as non-virus controls. The best fitting p value (1.12×10^{-5}) was found by non-linear regression (mortality = 1-exp(-p· concentration), SPSS 15.0).

The invasion probabilities for virions of genotypes A and B were likewise estimated by non-linear regression on dose-response data obtained individually for each genotype in *Spodoptera exigua* L3 (Zwart *et al.*, 2008), giving a ratio of the invasion probabilities of the two genotypes (A:B = 3.17). The exact ratio of the two genotypes in the polyhedra mixture used for the experiments had also been determined by qPCR (A:B = 0.46; Zwart *et al.*, 2008). The expected ratio $\lambda_A : \lambda_B$ under the IAH model is therefore 1.46.

L3 and L5 larvae of all 3 species and 2 instars were exposed to different doses. If there were less than the required number of cadavers for qPCR analysis, challenge experiments were repeated. Survival was then averaged over the different replicates performed. For each instar (L3 or L5) of each species, 24 larvae were individually analyzed by qPCR. In *S. exigua* and *T. ni*, the fifth instar is the final larval instar, whereas for *M. brassicae* it is the penultimate larval instar. A polyhedra concentration resulting in survival level S in the range 0.14-0.34 was used for *S. exigua* and *T. ni*, and $S \approx 0.7$ for *M. brassicae*. A lower survival rate could not be reached in *M. brassicae* L5 regardless of dose.

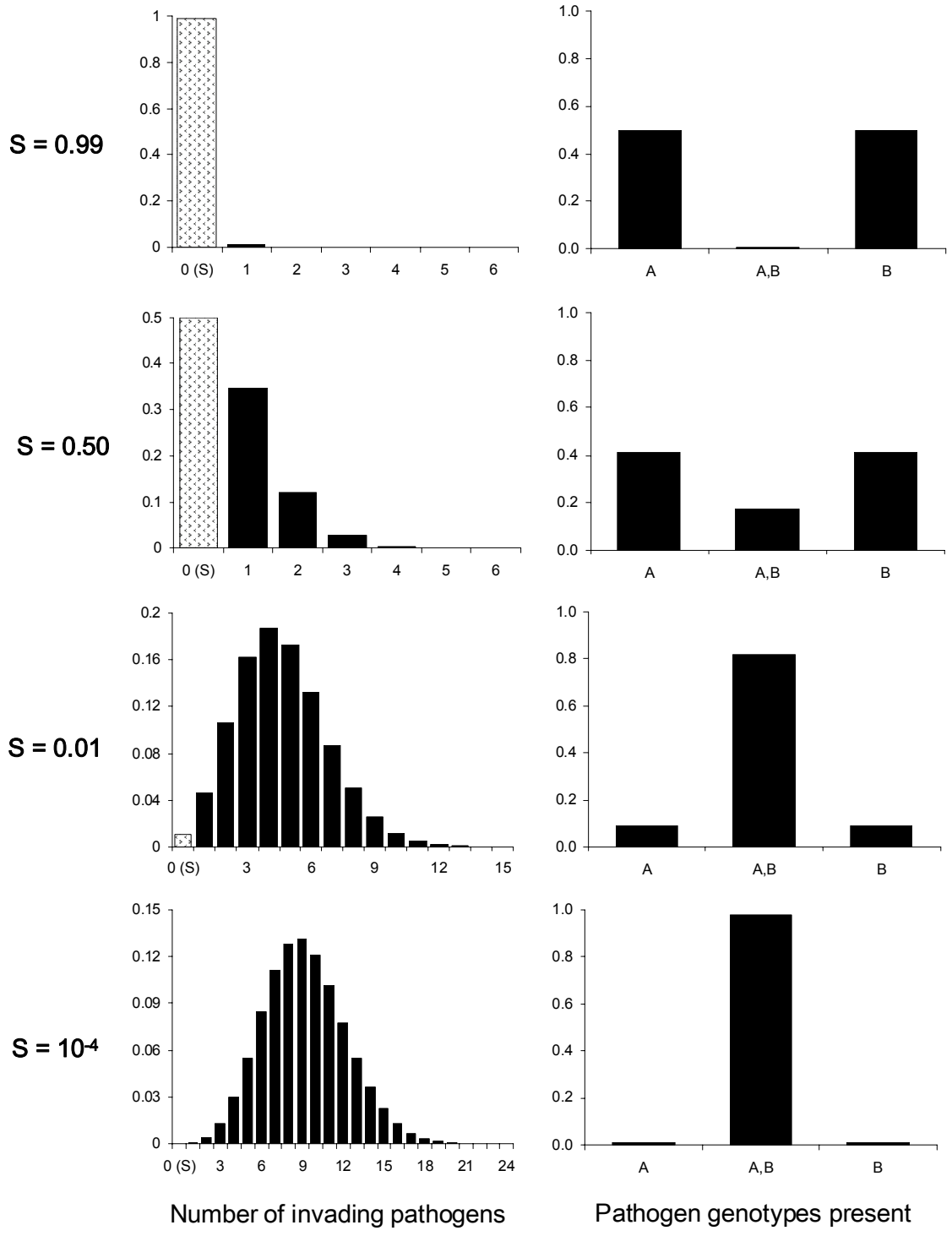


Figure 2: Model predictions under IAH on the frequency distribution of invading pathogens (left panels) and the frequency of dual genotype presence in diseased hosts (right panels). The corresponding frequency of host survival (*S*) is given by the grey bar on the far left. For all graphs, frequency is on the y-axis. Note that for the number of invading pathogens, the scales on both axes are not the same for different levels of host survival. The frequency that zero pathogens invade a host is also host survival (*S*). A starting population with a 1:1 ratio of pathogen genotypes A and B, and an equal invasion chance ($p_A = p_B$) between these pathogen genotypes was assumed.

(c) Binomial test for comparing experimental data to model predictions:

In order to test whether the qPCR data conformed to IAH, the frequency of dual genotype invasions observed in experiments was compared to that predicted by the model, using an exact one-tailed binomial test (SPSS 12.0). Observed host survival was used to calculate λ (equation 2), which was partitioned into λ_A and λ_B in proportion to the ratio of A and B in the inoculum, and the relative virulence of A and B (Zwart *et al*., 2008). The expected frequency of dual genotype presence in dead hosts, $P'(A \cap B)$, is predicted by the model (equation 7). This is the probability of a success in the binomial test. The number of host cadavers with both genotypes present (as determined by qPCR) was the number of successes observed, whereas the total number of host cadavers analyzed was the number of observations. A significance level of 0.05 was used.

3. Results

3.1 Dose – response relationship upon challenge of *Spodoptera exigua* L3 larvae with AcMNPV

We first considered whether the dose-response relationship observed is compatible with IAH, as IAH should lead to a distinct dose-response curve (Druett, 1952). A 1:1 mix of our AcMNPV genotypes A and B was used to challenge *Spodoptera exigua* third instar (L3) larvae, a permissive lepidopteran host species (Bianchi *et al.*, 2000), by droplet feeding (Zwart *et al.*, 2008; Hughes and Wood, 1981). The observed dose-response relationship does not only depend on whether IAH holds, but also on the between-hosts variation in ingested dose (Ridout *et al*., 1993) and susceptibility (Ridout *et al*., 1993; Ben-Ami *et al*., 2008). Discriminating between departure from IAH and the effects of variation in host susceptibility will therefore be problematic (Ben-Ami *et al*., 2008), hindering a rigorous test of IAH by means of dose-response data. Our dose-response data are, however, clearly compatible with IAH (Figure 3), given that between-hosts and between-cohort variation in dose and susceptibility would lead to a more shallow dose-response curve (Ridout *et al*.; 1993; Ben-Ami *et al*., 2008). Including an estimate of between-hosts variation would therefore further improve the fit for *S. exigua* L3, which is already quite good.

3.2 The frequency of dual genotype invasion in *Spodoptera exigua* L3 larvae

The frequency of dual genotype invasion was then determined for three different doses in *S. exigua* L3 (Figure 4, i-iii). The frequencies of dual genotype invasion in the experimental data were not significantly different from model predictions of $P'(A \cap B)$. We therefore conclude that IAH gives a pertinent estimate of the number of invading virions, as well as the dose-response relationship, for AcMNPV challenge of *S. exigua* L3 larvae. This is the first report of a thorough experimental test of IAH in any pathosystem. We did however consistently see a somewhat higher frequency of dual genotype invasion than predicted by the model. This may simply be a chance effect, or it may be that factors which this simple model does not take into account – e.g. variation between hosts in susceptibility – do have a minor effect of dual genotype invasion frequency in this setup.

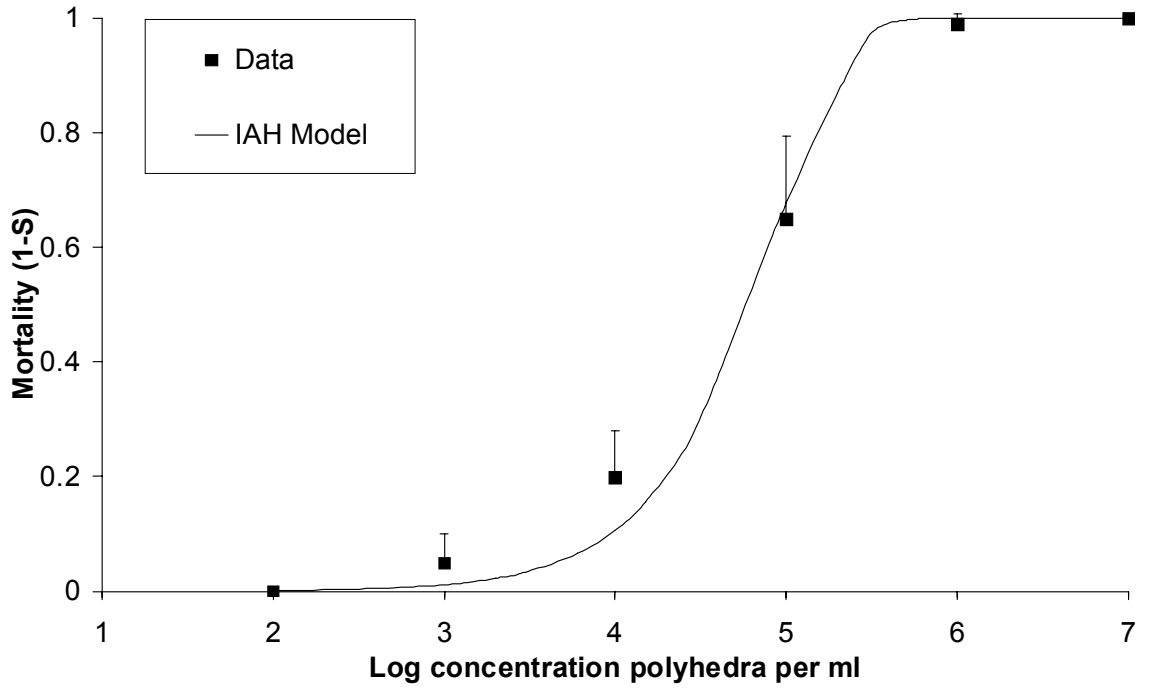


Figure 3: Dose – response data for AcMNPV in *Spodoptera exigua* L3 larvae. On the x-axis is the polyhedra concentration, and on the y-axis mortality. Bioassay mortality data (squares), with standard errors, and the fitted dose-response relationship under IAH (line) are displayed. See methods section for the number of replicates per dose. Note that for simplicity concentration was taken as dose, rather than an estimate of the number polyhedra or virions ingested.

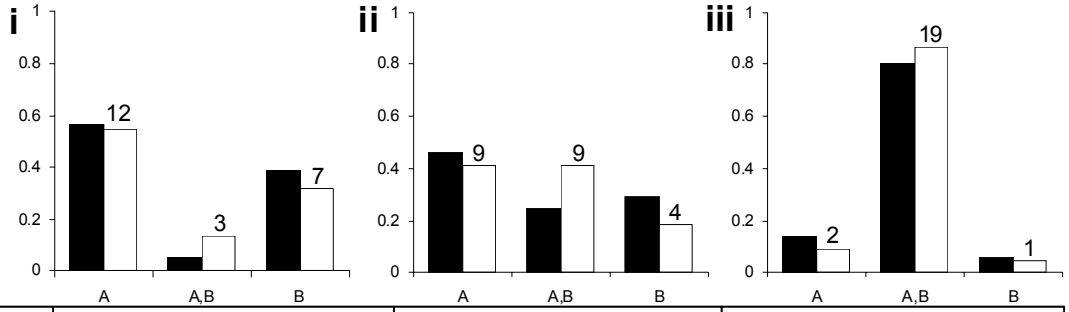
3.3 The frequency of dual genotype invasion in L3 or L5 larvae of three host species

To test the generality of the IAH model for AcMNPV, we determined if IAH also pertains to the AcMNPV challenge of other host species and in different larval instars of the same host species. We used L3 and L5 larvae of *S. exigua*, *Trichoplusia ni* (another permissive host) and *Mamestra brassicae* (a semi-permissive host). Here, we found that host compatibility and larval instar modulate the applicability of IAH. L3 larvae of *S. exigua* and *T. ni* showed a frequency of dual genotype invasion that was not different from what is predicted with IAH (Figure 4, iv-v). The frequency of dual genotype invasions was significantly greater than predicted by the IAH-based model in L3 larvae of the semi-permissive *M. brassicae* (Figure 4, vi) and in all of the challenge experiments with L5 larvae (Figure 4, vii-ix). For the sake of later discussion, we also provide dose-response data for all five pathosystems – those other than *S. exigua* L3 (Figure 5).

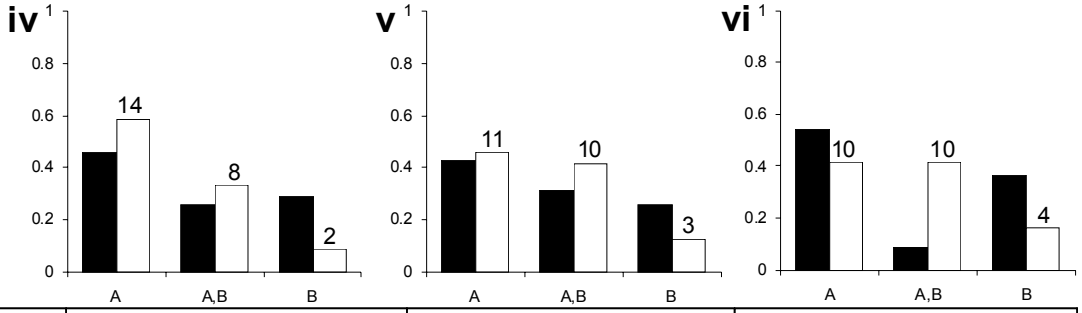
4. Discussion

We have formulated a model of pathogen infection based on IAH which gives satisfactory estimates of the frequency of dual genotype invasion for the baculovirus AcMNPV, but only for early instar larvae of a permissive host (Figure 4). A recent report has shown that complementation between RNA virus genotypes is important for pathogenesis (Vignuzzi *et al*., 2006). Our findings illustrate that the invasion and diseasing processes of a DNA virus do not necessarily involve complementation

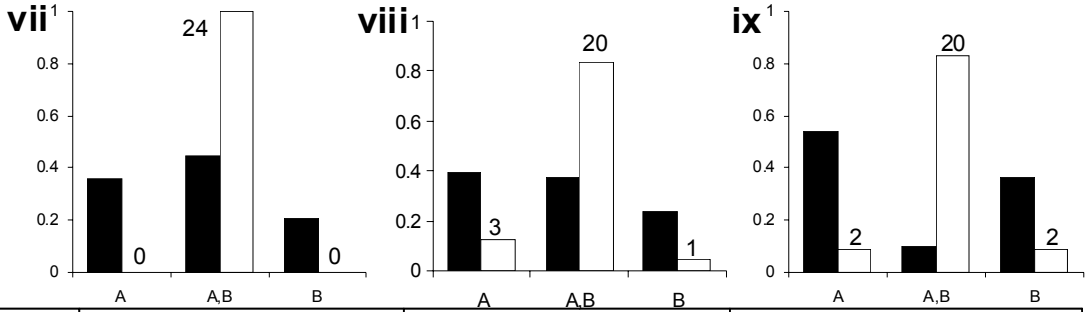
Figure 4 (right): Model predictions and experimental data for the frequency of dual genotype invasions. The presence of genotypes A and B was determined at different doses in *S. exigua* L3 (i-iii, 22 larvae per dose), and in different host species in L3 (iv-vi) and L5 (vii-ix) larvae (24 larvae per treatment). Model predictions are depicted with black bars. Experimental data are depicted with white bars, and the number of observations noted above each column. *S* is survival and $f(A \cap B)$ is the relative frequency of dual genotype invasions in the experimental data. $P'(A \cap B)$ is the frequency of dual genotype invasions predicted under IAH. Binom. *P*-value is the observed significance of a binomial test comparing $f(A \cap B)$ to $P'(A \cap B)$. An asterisk marks significant departure from IAH. Panels i-iii demonstrate that IAH holds in the third larval instar of *S. exigua*, irrespective of dose or mortality. Panels iv-vi show that IAH holds in L3 of *S. exigua* and *T. ni*, but not in L3 of semi-permissive *M. brassicae*. Panels vii-ix show significant departure from IAH in challenge experiments with fifth larval instars in all three species.



Panel	i	ii	iii
Host	*S. exigua* L3	*S. exigua* L3	*S. exigua* L3
Polyhedra/ml	10^4	10^5	10^6
S	0.80	0.35	0.01
f (A n B) - data	0.136	0.409	0.864
P'(A n B) -model	0.054	0.248	0.799
Binom. *P*-value	0.113	0.071	0.328



Panel	iv	v	vi
Host	*S. exigua* L3	*T.ni* L3	*M. brassicae* L3
Polyhedra/ml	10^5	10^5	10^7
S	0.34	0.26	0.69
f (A n B) - data	0.333	0.417	0.417
P'(A n B) -model	0.254	0.314	0.089
Binom. *P*-value	0.248	0.192	< 0.001*



Panel	vii	viii	ix
Host	*S. exigua* L5	*T.ni* L5	*M. brassicae* L5
Polyhedra/ml	10^7	10^5	10^9
S	0.14	0.20	0.67
f (A n B) - data	1.000	0.833	0.833
P'(A n B) -model	0.441	0.370	0.096
Binom. *P*-value	< 0.001*	< 0.001*	< 0.001*

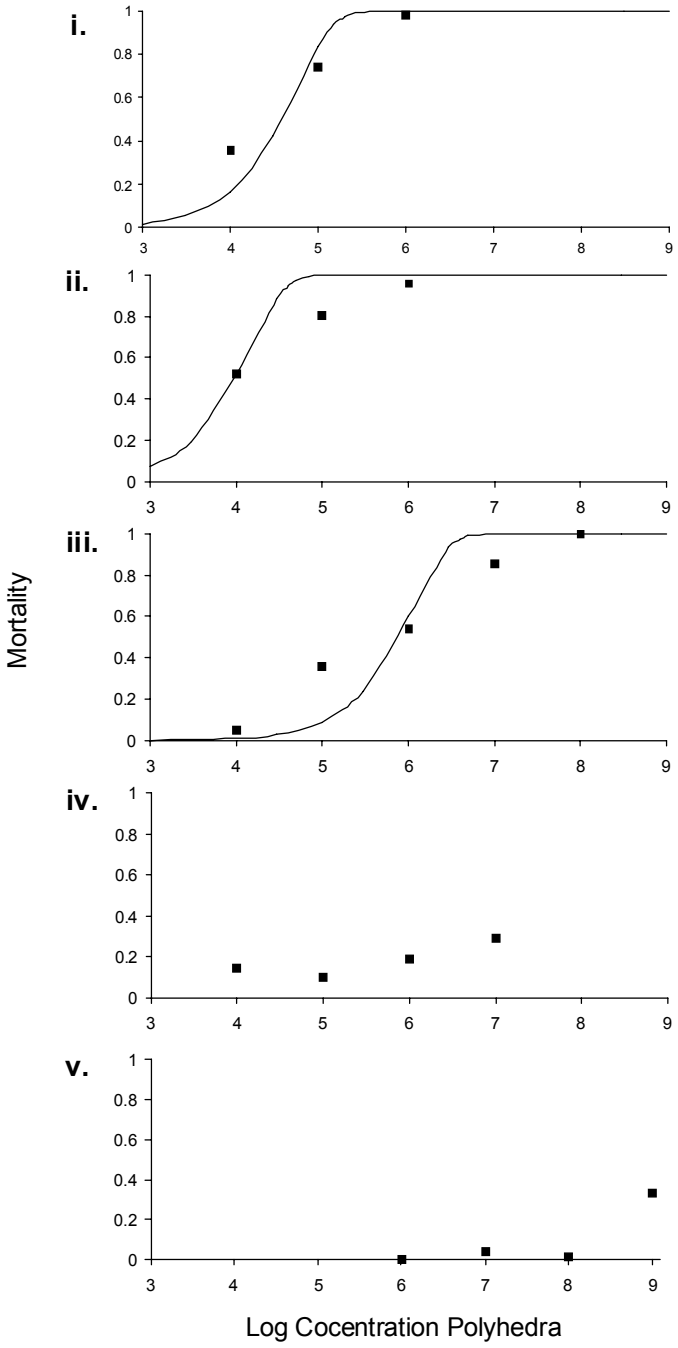


Figure 5: Dose-response relationships in semi-permissive larvae and L5. Single replicates of dose-response bio-assay data of *T. ni* L3 (i) and L5 (ii), *S. exigua* L5 (iii), and *M. brassicae* L3 (iv) and L5 (v). On the x-axes are the concentrations of polyhedra in the innoculum, and on the y-axes mortality. Experimental data are denoted by black squares, and the IAH model for dose-response is given by the line. This is done only for comparative purposes, and only in those instances were the model could be fit using non-linear regression (see Methods). The data suggest that in these other instances, dose response curves are shallower. Whether the *T. ni* L5 data are really shallower cannot be ascertained from these data. Note that only *T. ni* L3 dual genotype invasion frequency data conform to the IAH model.

between conspecific individuals. The two findings are of course not contradictory, as complementation is thought to occur between genotypes and IAH concerns the interactions between conspecific individuals. For example, IAH may very well apply to a clonal RNA virus population, although this conjecture may be hypothetical given the speed with which diversity is generated in RNA virus populations (e.g. Cuevas *et al.*, 2005).

Results obtained in L3 and L5s of three different host species demonstrate unequivocally that the IAH model with a fixed dose and invasion probability does not apply in L5s of any of the tested species, nor in the L3 of the semi-permissive *M. brassicae* (Figure 4). IAH predictions were not significantly different from observed genotype frequencies in L3 larvae of *S. exigua* and *T. ni.* The methodology therefore appears to be suitable for determining situations in which the IAH model must be rejected. On the other hand, we have formulated a basic model of the infection process, one that does not include, for example, variation in ingested dose and differences in host susceptibility. It may be that an IAH-based model which includes these factors may describe the infection process in those instances that the model formulated here fails to do so. Such a test must be performed before IAH itself – and not just the model we have formulated here - can be categorically rejected for the baculovirus infection process in L5 and semi-permissive species.

In all cases in which the IAH model was rejected, shallow dose-response curves were observed (Figure 5). Shallow dose-response curves have been reported earlier for late instar larvae (Bianchi *et al.*, 2000, 2002). Our data suggest that there may be agreement between the two methodologies for determining whether IAH holds: considering dose-response data as originally proposed (Druett, 1952), and determining the presence of pathogen genotypes in host cadavers as proposed here. One advantage of our method is that it may work without an estimate of the variation in host susceptibility. This is required for a test of IAH using the dose-response relationship (i.e. any variation in host susceptibility will lead to a smoother dose-response relationship). A combination of the two methods is a reasonable test for the applicability of IAH in a particular system.

The failure of the IAH model to predict the frequency of dual genotype invasions in four out of six pathosystems investigated raises the issue of what the main difference between these systems is. If there is really dependent action in these instances – which remains to be tested - larval resistance mechanisms to baculoviruses may be the reason for this. Larval resistance tends to increase as development progresses, both within and between instars (Engelhard and Volkman, 1995; Hoover *et al.*, 2000; Washburn *et al.*, 1998). This age-dependent resistance may help explain our findings in L5. But what

exactly is the reason that the IAH model fails in these instances? The entire disease process can be divided up into at least two steps to help identify potential reasons: (1) invasion of the host and (2) establishing disease at the systemic/organismal level. If these resistance mechanisms act mostly on the former, this does not necessarily lead to the failure of IAH, as it only becomes harder for a virion to get into the host (host mortality at a given dose will be lower, but the response will still conform with IAH). On the other hand, if a resistance mechanism acts on the latter, this may lead to dependent action. For example, if a minimum or threshold number of virions must be reached before the host dies, the frequency of dual genotype invaded hosts at a given level of host mortality will be higher. Such a threshold could arise if there were a host resistance mechanism that could be overcome by pathogen swamping. However, if there were a minimum or threshold number of virions necessary to cause host death, a steep dose response curve would be expected (Ridout *et al*., 1993). In our experiments, rejection of IAH appears to be accompanied by shallower dose response curves (Figure 5). Although resistance mechanisms which might be responsible for the failure of the IAH model would be expected to act at the systemic level, the observed combination of high prevalence of dual genotype invasions and a shallow dose response curve cannot be attained by this simple extension of the IAH model. Other models will therefore have to be considered and tested.

Alphabaculoviruses are transmitted as polyhedra, i.e. proteinacious occlusion bodies containing many virus particles (virions). The infectious unit is the virion (Federici, 1997). IAH was originally formulated to describe a bacterial pathogen that also had a horizontal transmission stage which contains multiple pathogen infectious units (Druett, 1952). Likewise, the fact that MNPVs have multiple nucleocapsids per virion does not affect model predictions in this particular instance. The polyhedra used in the challenge experiments were separately derived from bacmids; hence the virions they contain have only nucleocapsids of a single genotype (Zwart *et al*., 2008). Despite the intricacies of the baculovirus transmission stage, the frequency of dual genotype invasions in this particular setup can therefore be described with a general model.

It has been previously suggested that IAH may apply to some pathogens, based on distributions of pathogen genotype frequencies in individual hosts when low host mortality is observed (Smith and Crook, 1998; Moxon and Murphy, 1978). These conclusions, however, were tentative; no analysis was performed, perhaps for lack of a theoretical framework. For the bacterium *Haemophilus influenza,* predominantly single genotype infections were observed at low doses. However, the observed dose – systemic response relationship is in contradiction to IAH (Moxon and Murphy, 1978). For *Artogeia rapae* Granulovirus (Baculoviridae), results were found which are in

agreement with IAH (Smith and Crook, 1998). However, multiple virus doses were not included, the technique used for detecting genotypes was not sensitive enough, and only a small sample size was analyzed.

Given the general nature of the model proposed here, we expect that it may be applied to many different types of pathogens. Moreover, the experimental technique of challenging hosts with a mixture of pathogen genotypes and subsequently determining genotype frequencies is available in many pathosystems (Carrasco *et al*., 2007; Lenhoff *et al.*, 1998, Marks *et al.*, 2005), and techniques for tracking genotypes of human viruses are also readily available (Giannini *et al.*, 1999). Since independent action can be seen as a limiting case, we hope the present report will lead to testing of this model in other pathosystems, and demonstrating which extensions are necessary and sufficient to explain deviations from this simplest of hypotheses. Knowing whether IAH applies to a particular pathosystem may be indispensable to understanding pathogen evolution; the demonstration of conditional IAH in virus-insect systems and the availability of our methodology render testing IAH in other pathosystems feasible.

Acknowledgements:

We thank Frans van Aggelen, Andre Gidding and Leo Koopman for providing *S. exigua* and *M. brassicae* eggs. We thank Mart de Jong, Cajo ter Braak, Kelli Hoover, Jaap Molenaar, Paul Struik, and Gorben Pijlman for comments on various aspects of the manuscript. Mark Zwart was supported by a grant from the C.T. de Wit Graduate School for “Production Ecology and Resource Conservation”.

Chapter 5

A probabilistic model of baculovirus infection of late instar host larvae

---

Mark P. Zwart, Lia Hemerik, Just M. Vlak, and Wopke van der Werf

This chapter is to be submitted for publication.

Abstract

We previously demonstrated that a model based on the independent action hypothesis (IAH) did not describe the infection process of *Autographa californica* multicapsid nucleopolyhedrovirus (AcMNPV) in fifth instar larvae (L5) of permissive hosts, whereas it did hold for L3 larvae. For L5 larvae, more virions appeared to be invading the host than predicted by IAH, and dose-response curves were shallower. To better understand the infection process in this case, we challenged *Spodoptera exigua* L5 larvae with multiple doses of polyhedra of two isogenic bacmid-derived AcMNPV genotypes, differing only in a 100 bp quantitative real-time PCR (qPCR) recognition sequence. Host survival was recorded and genotype frequencies in larval cadavers were determined by qPCR. A 'shallow' dose-response relationship was found. At a dose eliciting low host mortality (30%), there were predominantly single genotype infections. At doses eliciting higher mortality (50%-100%), there were almost exclusively dual genotype infections. We postulated four probabilistic models – all based on the IAH model - to describe these data. The best fitting model made a single additional assumption: that there was variation between hosts in susceptibility. Variation in host susceptibility results in a higher frequency of dual genotype invasions than predicted by an IAH-based model that does not allow for variability in the susceptibility of the host population. Since there are many pathosystems in which a combination of IAH and uniform host susceptibility may not correctly describe the infection process, the models described here may be widely applicable.

1. Introduction

The independent action hypothesis (IAH) describes a pathogen's mode of action by asserting that each pathogen individual has the ability to invade and disease the host, and that pathogen individuals act in an independent manner (Druett, 1952). The hypothesis results in a singularly shaped dose-response curve (Druett, 1952, Meynell and Stocker, 1957). We previously showed that the number of infecting pathogens can be predicted with IAH, based on the level of host survival in a pathogen challenge. This model correctly predicts the infection process for *Autographa californica* multicapsid nucleopolyhedrovirus (AcMNPV) in permissive, third instar insect larvae (Chapter 4). This was done by considering the frequency of dual genotype pathogen invasion, when a population of hosts is challenged with a pathogen population consisting of two similar genotypes. The IAH model did not, however, predict the

outcome of the infection process when host larvae were in their fifth instar and/or of a semi-permissive host species: (1) the dose-response curves were less steep than the prediction of the IAH model and (2) the frequency of dual genotype invasion was higher – suggesting that more virions were entering the hosts than predicted. More polyhedra are typically required to reach equal levels of mortality in L5 larvae, compared to preceding larval instars (e.g. Bianchi *et al.*, 2000). Note that invertebrates, such as lepidopteran insect larvae, show both age-dependent resistance as well as semi-permissiveness for baculoviruses. Our data, however, suggest that the infection process may be fundamentally different in L5 larvae.

The most obvious mechanism which would describe the low frequency of single genotype invasions is that multiple pathogen invasion events are necessary in order for the host to become diseased. However, if there is a threshold for the number of invading virions required to disease the host, this has implications for dose response. Ridout *et al.* (1993) have already argued that a 'multiple-hit' model will lead to steeper dose-response curves than a single hit model. Here, we look at the AcMNPV process of infection of *S. exigua* L5 larvae in detail. Our aim is to identify mechanisms that can explain why independent action appears to fail for late instar larvae, and formulate a parsimonious model that does correctly describe the infection process in this case, reconciling the frequency of dual genotype invasions with a shallow dose-response.

The IAH model had previously not been rigorously tested for viruses, or in any pathosystem (Chapter 4). As such, there are no reports on reasons why this model is not applicable in specific instances. Understanding the infection process when the IAH model fails is highly relevant, for the following reasons: (1) there could be deviation from independent action in many pathosystems. For example, we found that IAH predictions were adequate in 2 out of 6 pathosystems investigated – those pathosystems in which there was no independent action were characterized by higher levels of host resistance (Chapter 4). This suggests that for a broad understanding of virus and general pathogen infection dynamics, these situations must be given due consideration. (2) In order to better understand those pathogens that are most relevant from societal perspective – namely those infecting humans – models of disease in a host characterized by complex host defense and resistance mechanisms must be developed.

To better understand the baculovirus infection process in late instar larvae, we first characterized AcMNPV infection of *Spodoptera exigua* L5 in more detail: the dose response curve and genotype distributions in infected hosts - at multiple doses - were determined. For these experiments we used polyhedra of the previously described vPolhA and vPolhB bacmid-derived AcMNPV genotypes (Zwart *et al.*, 2008). We

then postulated four stochastic models - all extensions of the IAH model we previously presented (Chapter 4) – and tested which best described the outcome of these experiments.

2. Methods

2.1 Bioassays and quantitative real-time PCR

A droplet feeding bioassay was performed with newly molted 5th instar *S. exigua* larvae (0-8 hours), starved for 16 h, as described previously (Chapter 4). The larvae were challenged with a 1:1 polyhedra mixture of vPolhA and vPolhB bacmid-derived genotypes (Zwart *et al*., 2008). Twenty-four larvae were taken as a non-virus control. The same number was infected per dose, with 10-fold incrementing doses ranging from 10^3–10^9 polyhedra/ml. As the number of larvae in a synchronized cohort was limited, the entire range was not taken in a single replicate, but a subset of ranges taken in each replicate. Mortality was recorded daily and dead larvae collected and stored individually at -20° C. Subsequent polyhedra purification, DNA isolation and qPCR were performed as previously described (Chapter 4).

2.2 Proposed probabilistic models of the infection process

We first give a conceptual overview of the infection process, followed by an overview of the models we are proposing. We view the baculovirus infection process in general as being composed of 5 main steps (Figure 1): (Step 1) Polyhedra are ingested by a host insect larva. In the case of the setup we are trying to model here, this occurs by drinking a suspension of polyhedra. (Step 2) Subsequently, polyhedra fall apart as a result of the alkaline pH in the larval midgut, a process we refer to as polyhedra dissolution. (Step 3) Liberated occlusion derived virus (ODV) can bypass the peritrophic membrane and enter an epithelial midgut cell. Subsequently, budded virus (BV) are secreted from the midgut cells into the interior tissues of the host (trachea, hemocytes). This entire process we refer to as invasion of the host (Steps 1-3). (Step 4) The virus can then be further amplified in the host, a process we refer to as replication (Step 5). If the virus is amplified sufficiently, the host insect dies. Sufficiently amplified virus can be measured by qPCR.

To describe the infection process of a baculovirus in L5 larvae, we propose to focus on two key steps in the infection process: invasion and inducing host disease. First, the host must be invaded by the pathogen (Step 3 in Figure 1). Second, a full-blown

systemic infection and pathogenesis can occur, and in our model system, host death (Step 5 in Figure 1). All the four models proposed consider only these processes. We ignore the steps of ingestion of polyhedra and the degradation of polyhedra in the midgut, as (1) L5 larvae are typically highly resistant and must therefore ingest large numbers of polyhedra before any mortality is observed (2) the variation in ingested dose has been reported to have only minor effects on the dose-response relationship (Ridout *et al.*, 1993). We also ignore the viral replication phase as a process, and simply focus on the outcome of this phase: does the host succumb to infection or survive? Since we are attempting to predict the frequency of dual genotype infections at present, and not the actual ratios of genotypes in dual genotype infected larvae, the variation introduced by this step can be ignored.

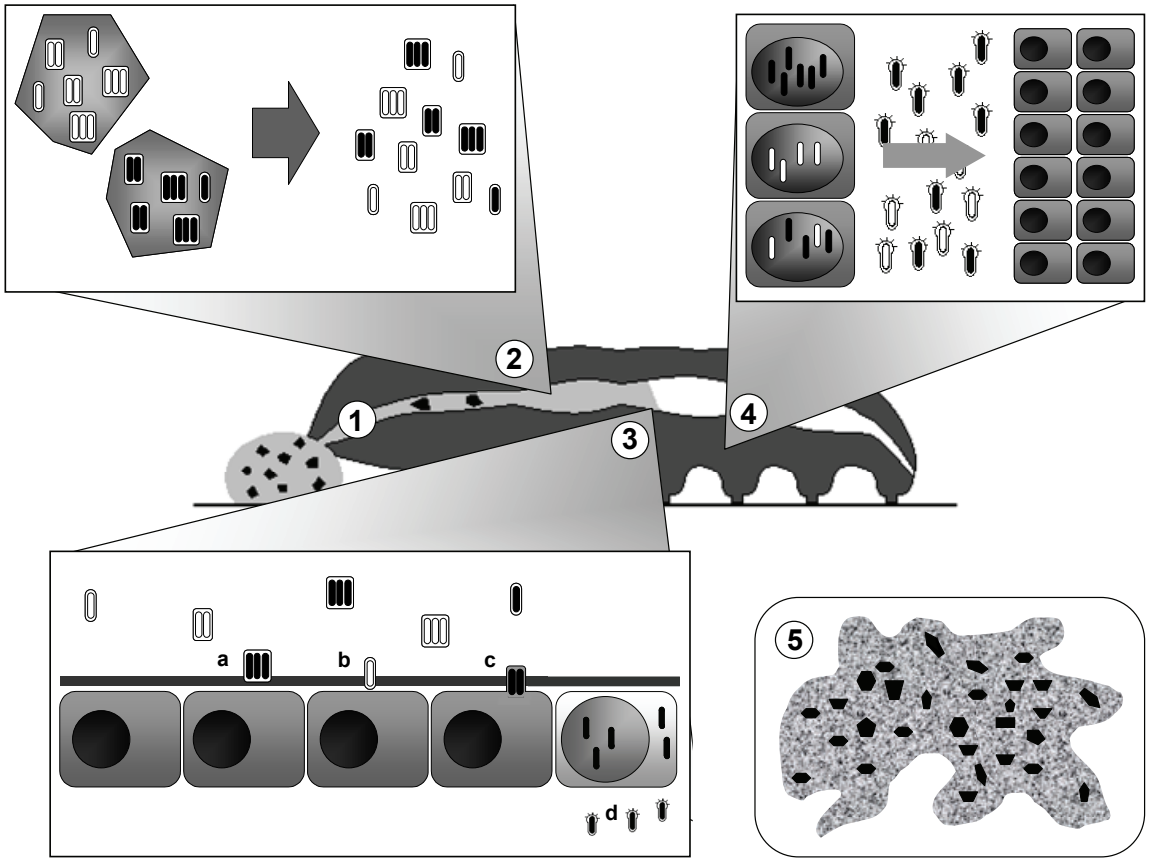


Figure 1: the AcMNPV infection process in our laboratory setup: (1) ingestion of polyhedra, (2) polyhedra dissolution, (3) invasion; this step consists of breaking down (a) and bypassing the peritrophic membrane (b), entering a midgut cell (c), and producing infectious progeny in the form of budded virus (d), (4) replication and (5) host death, upon which the larval cadaver liquefies and polyhedra are released.

We here consider four probabilistic models of the infection process. The IAH model previously described (Chapter 4) is here considered as a null-model. This model assumes that each pathogen individual challenging the host has a non-zero probability of invading and causing host disease, and that pathogen individuals act independently during the entire process (Druett, 1952). For all four models the following holds: given a virus population composed of n_A and n_B individuals of genotypes A and B, with invasion probabilities p_A and p_B, the mean number of invaders for each genotype is $\lambda_A = p_A n_A$ and $\lambda_B = p_B n_B$, λ_A and λ_B are Poisson distributed (Chapter 4) and $\lambda = \lambda_A + \lambda_B$. An overview of the four models postulated is given in Figure 2. Details of how model equations were derived are given in the Appendix.

2.2.1 Model 1: Invaded pathogens have a fixed probability of causing disease

The postulate that not all invading pathogens will cause disease is relevant because baculoviruses are known to cause sub-lethal infections (Sait *et al.*, 1993; Myers *et al.*, 2000). In the framework of model 1, sub-lethal infections would be considered as invaded virions that do not cause host disease and perhaps remain quiescent in the

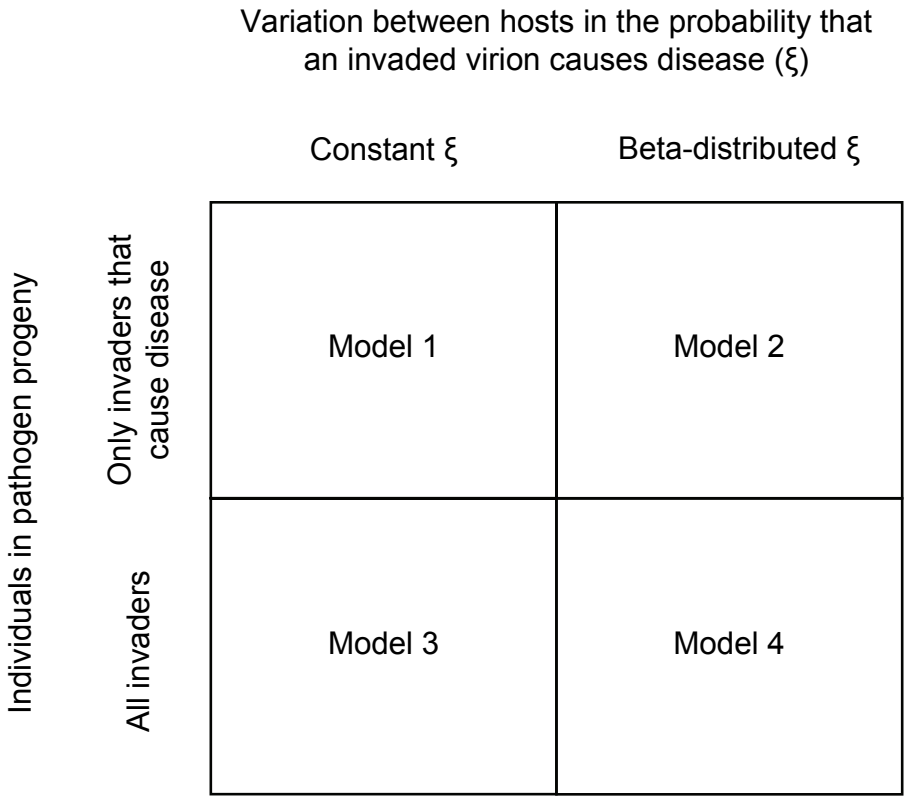


Figure 2: An overview of the four models postulated.

host. Sub-lethal infection may indicate that invasion does not necessarily lead to death (as assumed by all four models presented here).

The first model (1) assumes independent action in host invasion and (2) relaxes the assumption that the invasion of the host by a pathogen will inevitably lead to host disease. Instead, we assume that each invaded pathogen has a fixed probability ξ of causing disease. This probability is constant over the population of hosts. We also assume that only those pathogens which have caused diseased are represented in the pathogen progeny: these pathogens have been established in the host. In this model, it is possible to invade but not cause disease. These individuals are not represented – or alternatively not measurable – in the pathogen population which emerges from a diseased host. There is also independent action in the process of host disease.

It can be appreciated that this model is similar to the IAH model: the definition of host invasion, as proposed for model 1, allows pathogens to be removed after having invaded, but pathogen individuals still operate independently and the chance of bringing about host death is the invasion chance (p) multiplied by ξ. This model is useful for illustrating that the IAH model does not strictly define the number of steps involved in true 'invasion' of the host: pathogens must simply act independently at all levels of the process. For model calibration (2.4), the product $\xi \cdot p$ must be estimated.

2.2.2 Model 2: Invaded pathogens have a beta-distributed probability of causing disease

It has previously been suggested that differences in host susceptibility may be important in understanding dose response. For example, Ridout *et al.* (1993) and Bianchi *et al.* (2002) recognized this for baculovirus infection of late instar larvae. Another study investigates whether a model assuming departure from IAH (pathogen antagonism) or differences in host susceptibility better describes experimental dose-response data in another pathosystem, and finds that it is the latter (Ben-Ami *et al.*, 2008). The effects of this assumption (variation in host susceptibility) on the rate of dual genotype presence – or even co-infections in general – have not been given consideration to our knowledge.

Model 2 is an extension of model 1, and adds the assumption that the probability ξ of causing disease varies over hosts. ξ is now the realization of a stochastic variable σ; for each host one realization is valid and the stochastic variable σ follows a beta distribution over hosts. The beta-distribution was chosen because of its versatility (e.g. Olkin *et al.*, 1994) and because there is a precedent for its use in describing the variation in host susceptibility (Ridout *et al.*, 1993). As in model 1, ξ and p are very

closely linked and cannot be identified separately. Since ξ follows directly from the two parameters chosen for the beta distribution (α, β, see Appendix), p was set to a convenient value for calculating the beta probability density function, and α and β could then be estimated.

2.2.3 Model 3: Invaded pathogens have a fixed probability of causing disease and all invaded virions are represented in the pathogen progeny

Baculoviruses are known to latently infect insect larvae (Jurkovičová, 1979; Hughes *et al.*, 1993). Latent infections that become activated and cause host death – presumably coming about because host defenses have been weakened by biotic or abiotic stress or disease – would suggest that once an invader has broken the host defense, other invaded pathogens could be represented in the viral progeny subsequently generated. This postulate is incorporated by this model.

Model 3 assumes – like models 1 and 2 – that each invading virion has a probability ξ of bringing about host disease. As in model 1, we again assume there is no variation in susceptibility in the host population. However, unlike models 1 and 2, we here assume that all pathogens which have invaded the host will be represented in the pathogen progeny. In other words, if any one invading pathogen manages to disease the host, than all other pathogens which have invaded (but had not been successful in inducing host disease) will nevertheless be represented in the pathogen progeny. To illustrate the principle consider the following example: pathogen individuals X and Y invade a host. X manages to cause disease in the host, while Y is not successful in causing disease. Where in models 1 and 2, Y would not be represented in the viral progeny, in model 3 (and model 4) Y is also represented in the viral progeny. Note that for this model, different combinations of ξ and p values will render different results; although the dose-response relationship is determined by their product, the prevalence of dual genotype infections at a given level of mortality will be determined by the exact p and ξ values.

2.2.4 Model 4: Invaded pathogens have a beta-distributed probability of causing disease and all invaded virions are represented in the pathogen progeny

This model is a further extension of model 3: all pathogens which have invaded the host will be represented in the pathogen progeny. However, as in model 2, the probability that an invaded pathogen will cause host disease (ξ) is again the realization of a stochastic variable σ, that follows a beta distribution over hosts. In order to fit this model, three parameters must be found: p, α and β.

2.3 Calibration of models

Calibration of the models was conducted in MatLab R2006b, using the least squares optimization function lsqnonlin. This function uses a large-scale algorithm for searching the parameter space. The algorithm is a subspace trust region method and is based on the interior-reflective Newton method (Coleman & Yi, 1994; 1996). Inputs to lsqnonlin are a function with the model equations, initial values for the parameters and their bounds, and the data. As data we used the absolute frequencies of death at seven doses, observed in the experiment, and the absolute frequencies of single genotype (A alone, and B alone) and mixed genotype infection at four doses. The number of degrees of freedom was calculated as: df = 7 + 4 x (3-1) – n, where df is the number of degrees of freedom for model estimation, and n is the number of parameters. The error sum of squares was calculated by summing the squared residuals, and residuals were defined as the unweighted difference between observed absolute frequency and model-predicted absolute frequency. Lsqnonlin estimates the Jacobian matrix of sensitivities of parameter values to data values, which allows estimation of standard errors of the parameters and their correlations.

In the function defining the mathematical model for the data, calculation of population level probabilities was achieved by numerical integration of model equations as given in Appendix 1 over the probability distribution of σ characterizing the variability in host susceptibility. Integrals were calculated numerically using a recursive adaptive higher order integration method, as implemented in the function QUADL in MatLab R2006b. QUADL is based on Lobatto quadrature (Gander & Gautschi, 2000). All results of numerical integrations over the beta PDF were verified by stochastic simulations of the infection process, based on a sufficiently large numbers of random draws to average out the variability, and results of parameter optimizations with lsqnonlin were corroborated by grid searches in MatLab and R.

3. Results

3.1 The dose-response relationship in *S. exigua* L5 larvae

The dose response curve for AcMNPV infection of *S. exigua* L5 was first determined (Figure 3). Previous reports suggested that dose response curves in L5 larvae of this insect were shallow (Bianchi *et al*., 2000, 2002). Here we found a dose-response curve that was much shallower than that previously observed for *S. exigua* L3, were the dose

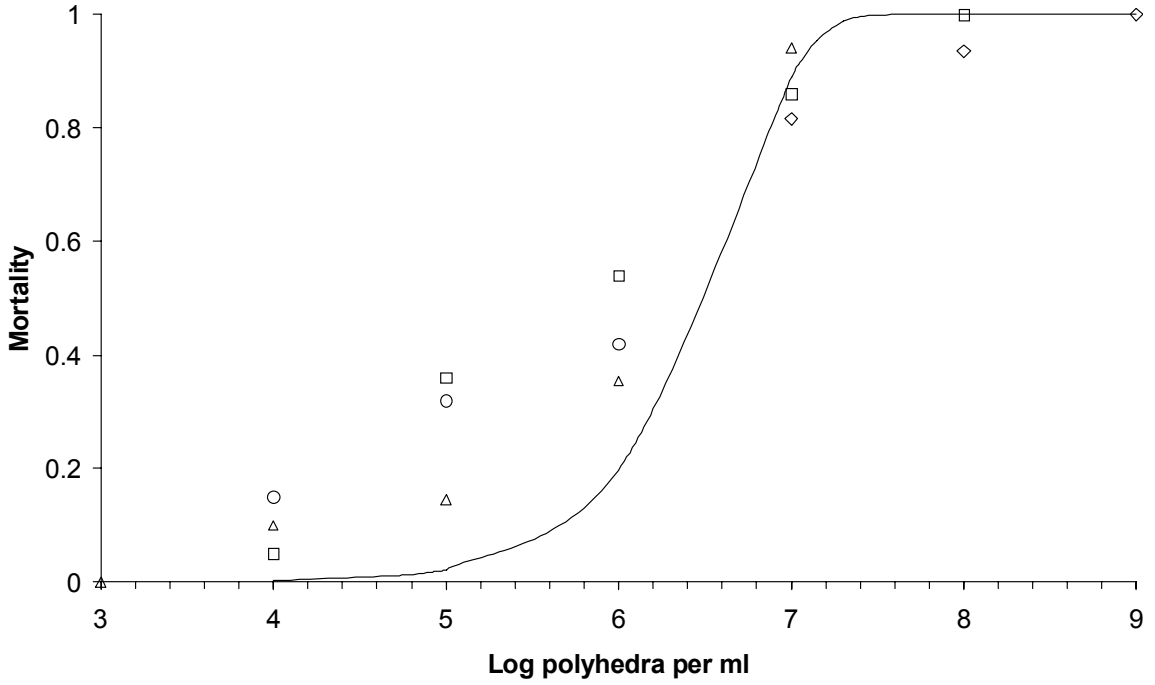
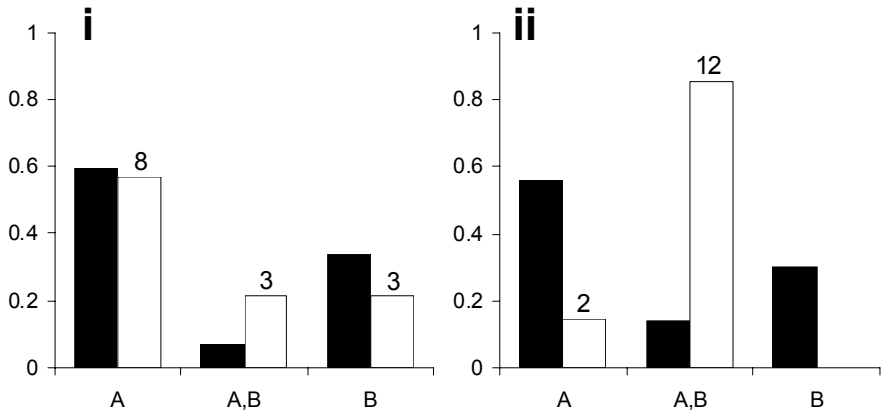


Figure 3: The dose-response curve for AcMNPV infection of *S. exigua* L5. On the x-axis the $\log_{10}$ of the polyhedra dose per ml droplet fed. The proportion of hosts dying is on the y-axis. The different symbols represent different replicates. Note that different replicates covered different ranges in dose. The line represents the fitted IAH-based dose-response curve (P- value (2.205×10^{-7}) found by non-linear regression (mortality = 1-exp(-p· concentration), SPSS 12.0).

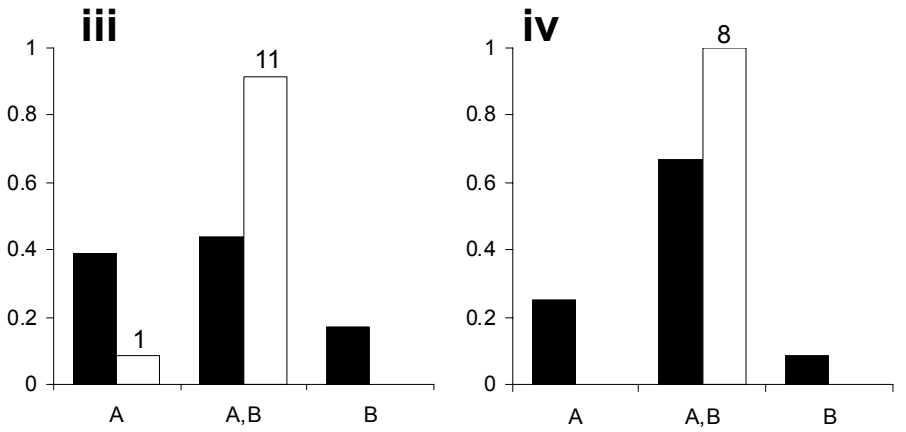
response was similar to that predicted by the IAH model (Chapter 4). There is a clear discrepancy between the data observed here and the prediction from an IAH-based model based on a fixed probability of infection (Chapter 4; Ridout *et al.,* 1993; Ben-Ami *et al.*, 2008).

3.2 Genotype frequencies at different doses determined by qPCR

The frequency of genotypes A and B was determined at different doses by qPCR. The data are presented categorically, as single or dual genotype infected hosts (Figure 4). These data were compared to the predictions of IAH by considering the frequency of dual genotype presence in cadavers, $P'(A \cap B)$, as described previously (Figure 4; see Chapter 4). The frequency of dual genotype presence contradicts IAH at all doses except the lowest, 10^5 polyhedra/ml, where there are numerous single genotype infections ($\approx 80\%$). We found that at all higher doses ($\geq 10^6$ polyhedra/ml),



Panel	i	ii
Polyhedra/ml	10^5	10^6
Number of larvae	14	14
S	0.72	0.55
f(A and B)	0.214	0.857
P'(A and B)	0.076	0.138
Significance level	0.085	>0.001*



Panel	iii	iv
Polyhedra/ml	10^7	10^8
Number of larvae	12	8
S	0.13	0.03
f(A and B)	0.917	1
P'(A and B)	0.440	0.667
Significance level	0.001*	0.039*

Figure 4: The frequency of dual genotype presence in S. exigua L5 larvae at different levels of mortality. *S* is mean observed mortality, *f*(A and B) the frequency of dual genotype presence in experiments, and *P*'(A and B) the model prediction for the frequency of dual genotype presence. Significance level is obtained by a one-sided binomial test (SPSS 12.0). The number of observations is included above each measured frequency.

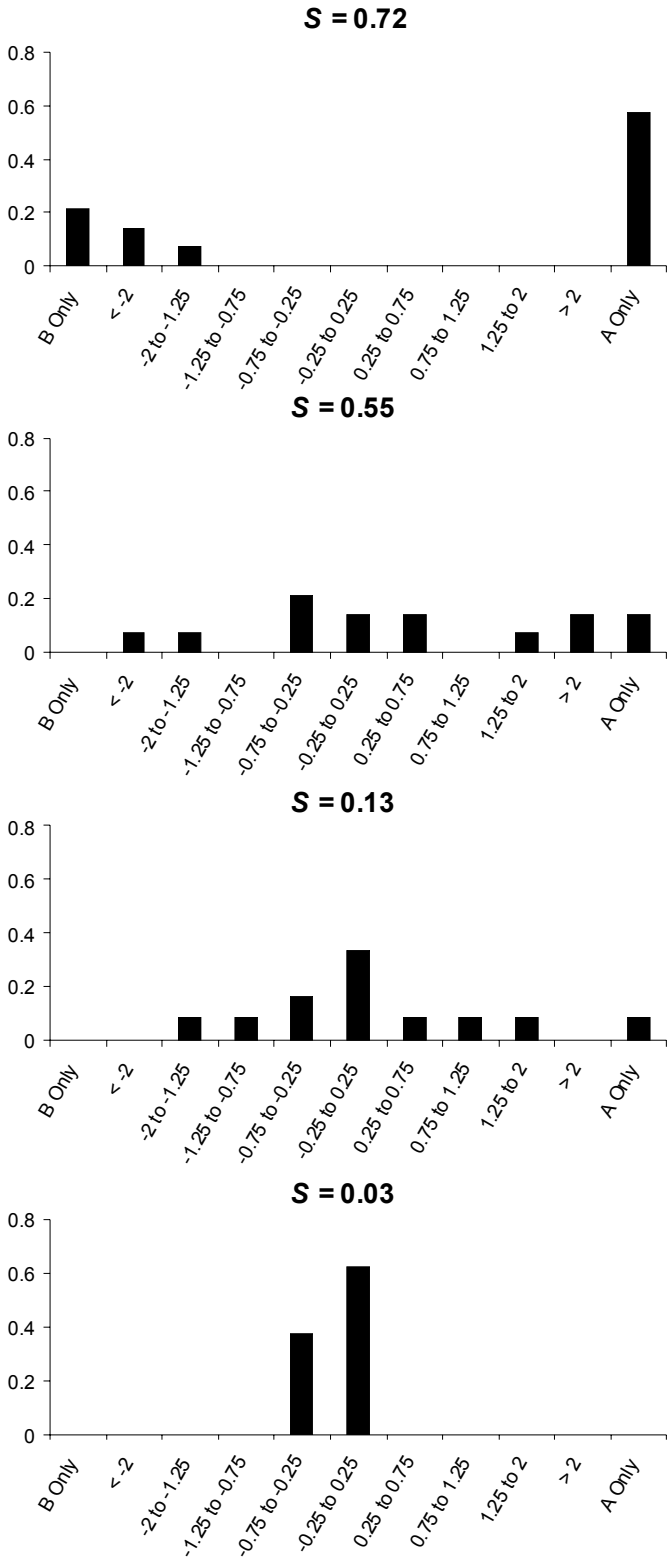


Figure 5: Genotype ratios, as measured by qPCR, in *S. exigua* L5 larvae at different levels of mortality. *S* is mean observed mortality. On the x-axis is the log of the genotype ratio (A:B), and on the y-axis the frequency. 8,12,14 and 14 larvae were analyzed (*S* = 0.03 to *S* = 0.72).

dual genotype infections predominate. These data are therefore in agreement with our previous findings for *S. exigua* L5, where we rejected an IAH-based model (Chapter 4). The genotype ratio data (Figure 5) show that as dose and mortality increase, the variation in genotype ratio appears to decrease. One would expect this trend since more virions will invade the host as dose is increased.

3.3 Calibration of models

The four postulated models were each calibrated (Table 1) and the solutions were plotted with the data (Figures 6 & 7). Models 1 (SSQ = 538) and 3 (SSQ = 384) gave the poorest fit; the dose response curves are much steeper than the data, due to the absence of variability in host susceptibility. Model 3, incorporating the idea that all invaded virions contribute to the measured genotype presence in cadavers, gives a

Table 1: Model calibration results. For the different models, different numbers of parameters needed to be estimated. SSQ is the sum of squares (see Methods 2.3). Note that 'gene pool in the cadaver' refers to those pathogen genotypes which can be measured

Model	Hypotheses	SSQ	Parameter	Estimate ± S.E.
1*	- Fixed ξ - Only disease causing virions enter the gene pool in the cadaver	537.8	$p \cdot \xi$	$1.89 \times 10^{-6} \pm 1.8 \times 10^{-8}$
2**	- ξ is a realization of a beta distribution - Only disease causing virions enter the gene pool in the cadaver	190.6	α	0.286 ± 0.0046
			β	2.78 ± 0.102
3	- Fixed ξ - All invaded virions enter the gene pool in the cadaver	383.6	p	$1.31 \times 10^{-5} \pm 6.3 \times 10^{-7}$
			ξ	0.137 ± 0.0067
4	- ξ is a realization of a beta distribution - All invaded virions enter the gene pool in the cadaver	213.9	p	$3.55 \times 10^{-5} \pm 0.168 \times 10^{-5}$
			α	0.310 ± 0.0066
			β	1.38 ± 0.093

* Only the product of p and ξ, and not the actual separate parameters, can be estimated. See Methods section for motivation.

** $\xi = \alpha/(\alpha+\beta)$, and p and ξ cannot be separate estimated. Therefore, p was arbitrarily set to 10^{-4} and only estimates of α and β were made.

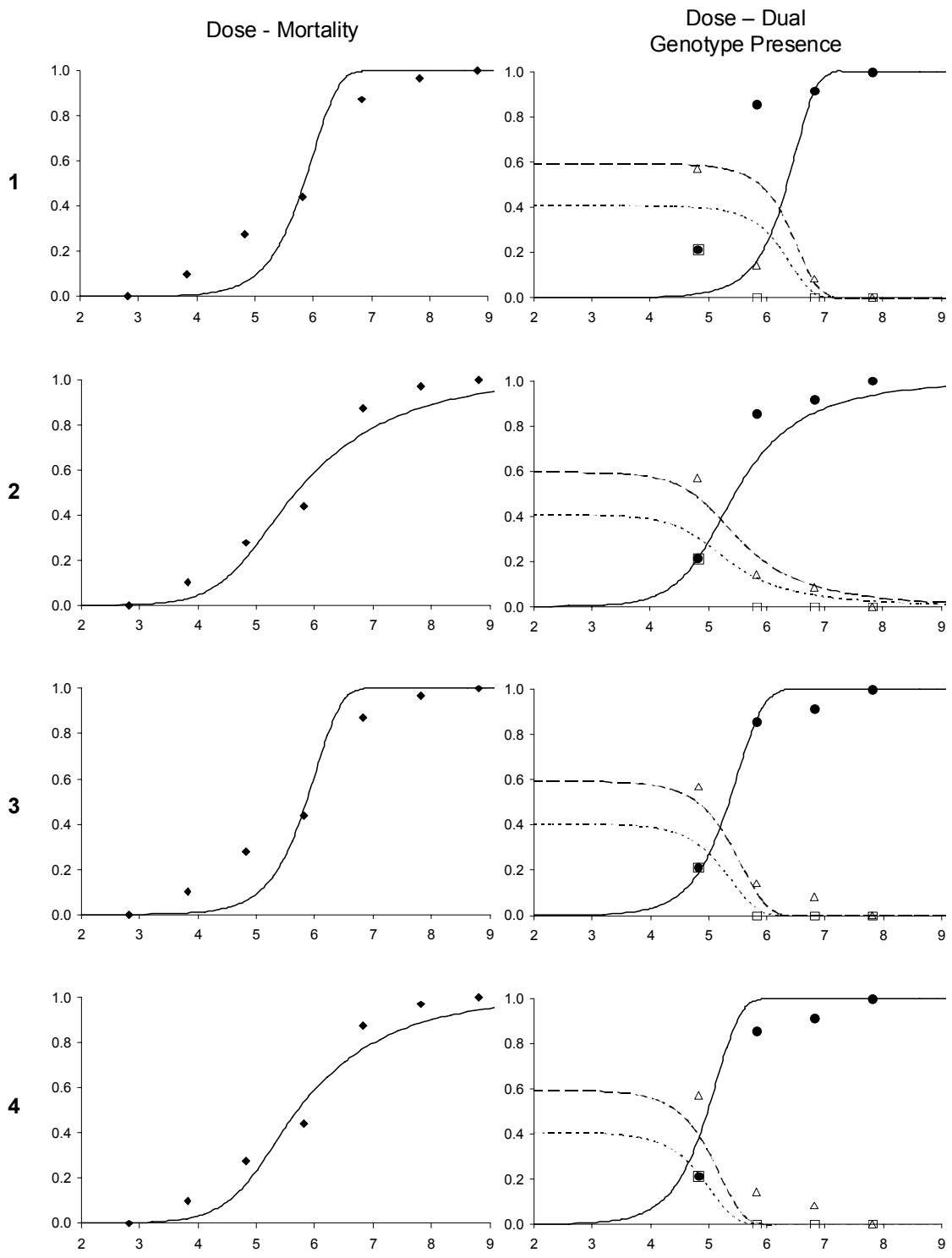


Figure 6: Fitted models 1-4 compared with experimental data. The fitted model is indicated by the numbers on the far left. For all panels, the log of the dose is on the x-axis and frequency on the y-axis. The dose-mortality relationship is in the left hand panels: diamonds are the experimental data and the line the model prediction. The relationship between dose and dual genotype presence is in the right panels. The symbols denote experimental data, and lines model predictions: black circles and the solid line denote the presence of genotypes A and B, triangles and the dotted line genotype A only, and squares and the fine dotted line genotypes B only, respectively. Model 2 gives the best fit.

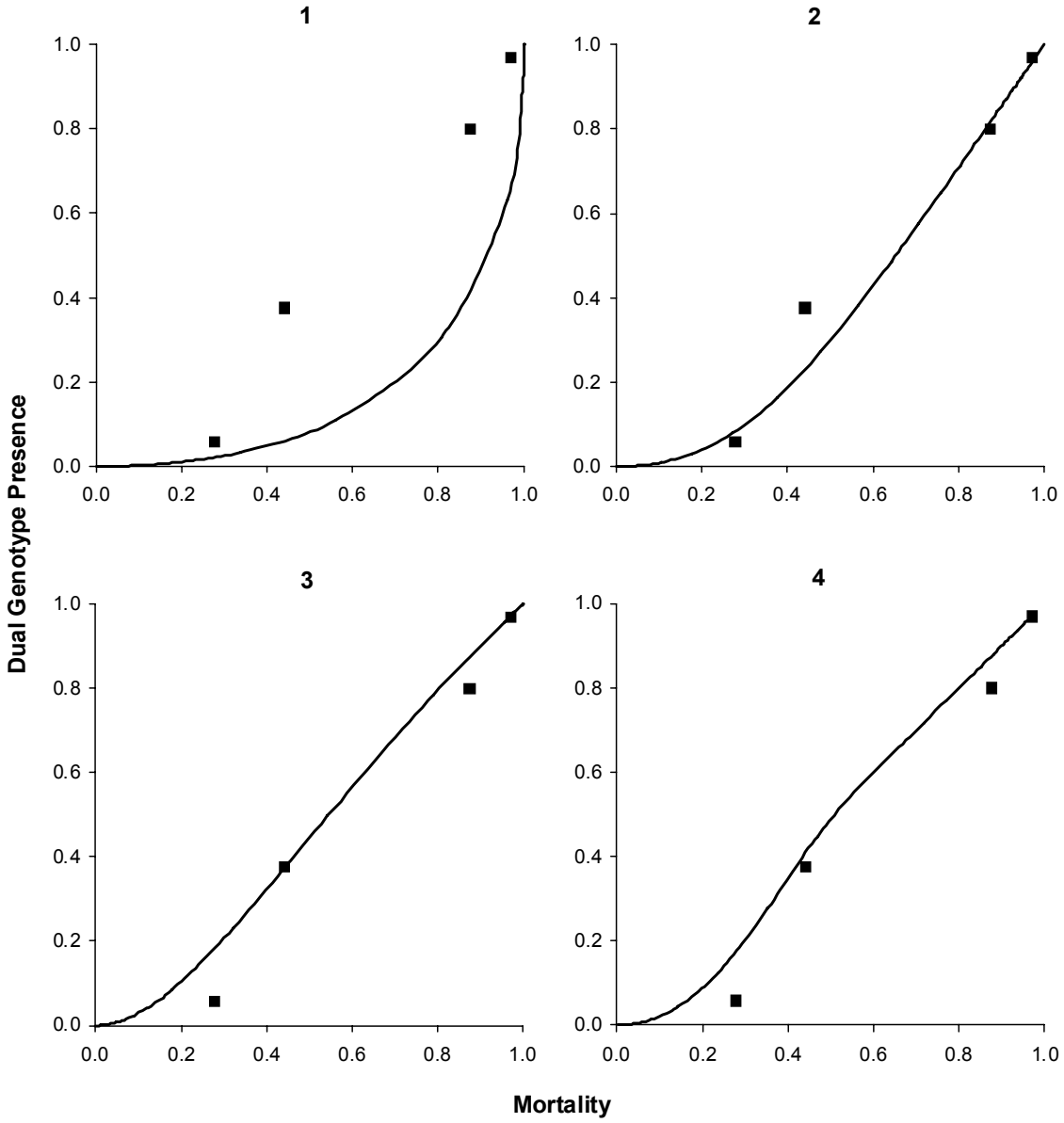


Figure 7: Fitted models 1-4 compared with experimental data. The relationship between mortality (x-axis) and the frequency of dual genotype presence (y-axis) is plotted. The squares represent the experimental data (mortality vs. dual genotype presence) and the line the model prediction.

good description of the frequency of dual genotype invasion, whereas model 1 does not. Models 2 and 4 both gave the best fits: dose response curves are much closer to the data, although the model mortality responses and frequencies of dual genotype presence approach 1 slower than the data. Model 2 (SSQ = 191) gives a somewhat lower sum of squares than model 4 (SSQ = 214). Whether the difference in model fit is statistically significant is moot; since model 2 requires only one assumption on top of

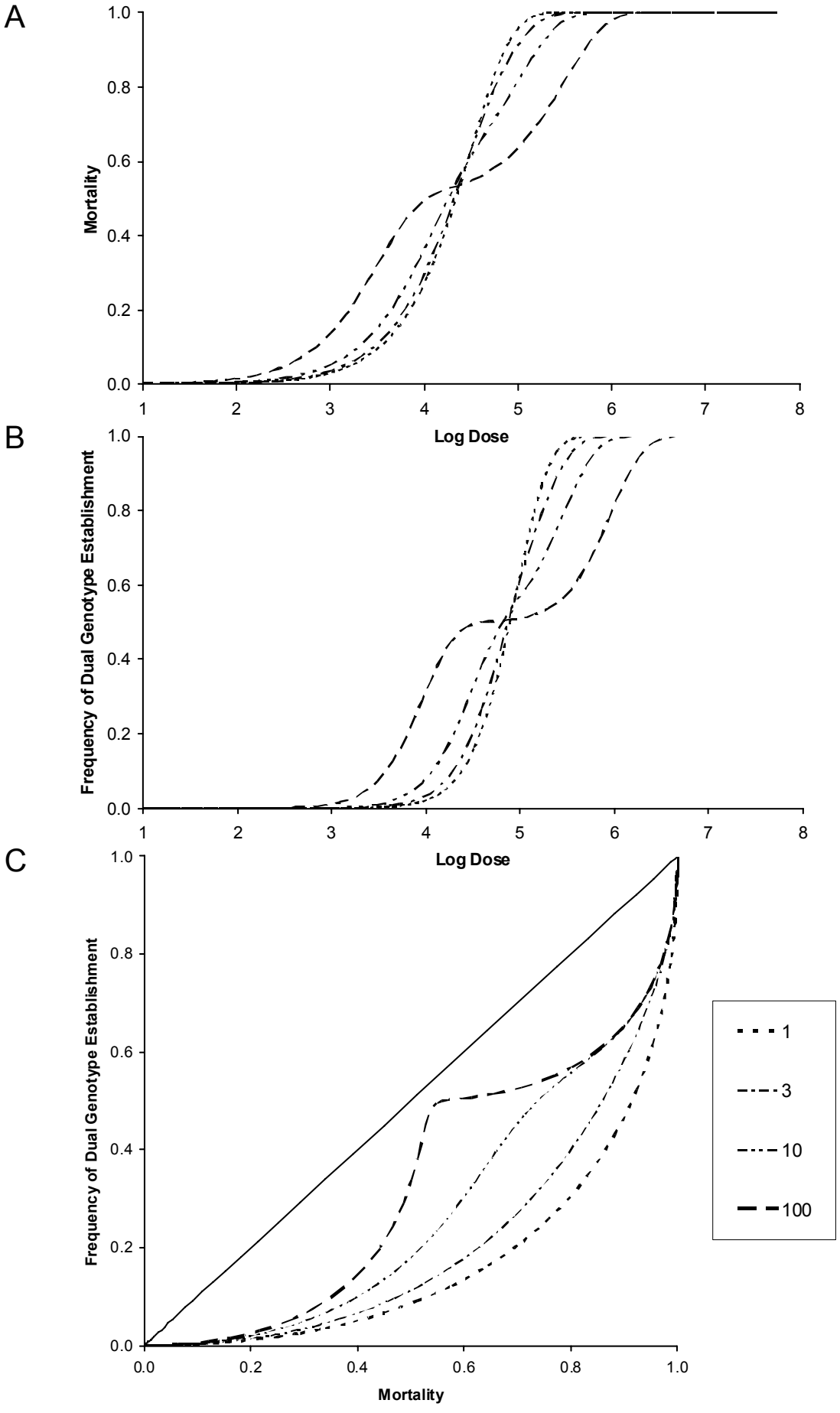
the IAH based model (there is variation in ξ between hosts) - whereas model 4 requires two assumptions – model 2 is the best of the four models for describing these data.

4. Discussion

Previous work (Chapter 4) on whether the IAH model can describe dose-response and the frequency of dual genotype presence in the baculovirus – insect larvae pathosystem raised an intriguing question: for those instances were the IAH model was rejected, how can shallow dose-response relationships and a high frequency of dual genotype presence be explained? The question is interesting in its own right: having a model which correctly predicts the number of invading baculovirus virions is important for understanding baculovirus evolution and ecology. But it is particularly relevant in a more general context to consider these sorts of models; the original IAH model is bound to fail in many other pathosystems, and the number of invading pathogens is an important issue. How does the pathogen infection process operate when IAH appears to fail?

Here we considered a number of models that could potentially describe the pathogen infection process when the IAH model fails. All of these extended models build on the IAH model (Zwart *et al.*, 2008; Druett, 1952). The different extensions made to this model are all arguably plausible given the current understanding of baculovirus biology (see methods for a motivation of each model).

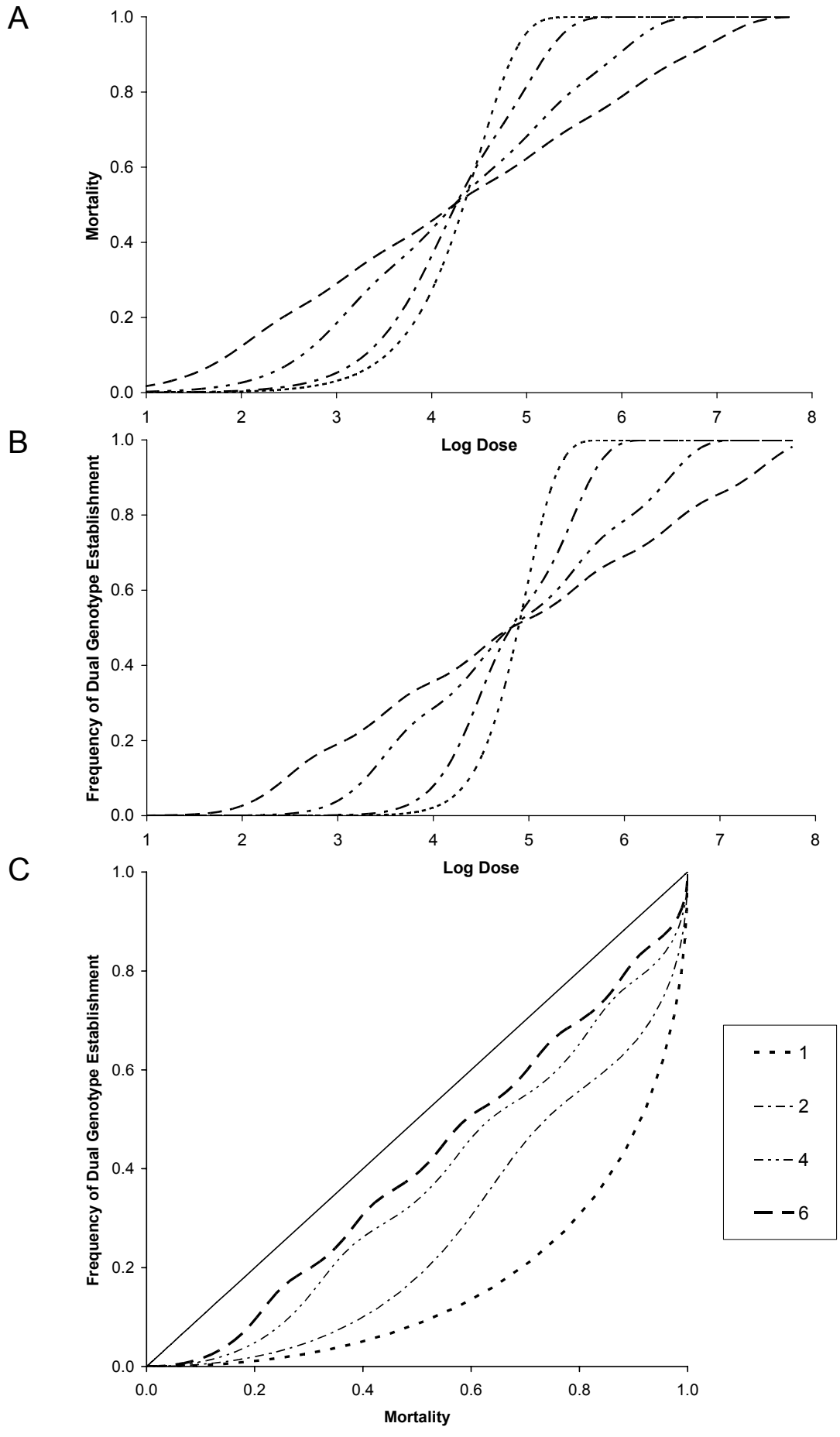
Figure 8 (Right): The effects of heterogeneity in host susceptibility. One half of the host population is assumed to be resistant, while the other half is susceptible. The ratio of pathogen invasion probability in the susceptible and resistant hosts is given in the legend, and applies to all three panels. (E.g. A value of '1' indicates the invasion probabilities are the same, so in this instance the host population is homogenous. '100' indicates the pathogen invasion chance is 100 times lower in the susceptible hosts than in resistant hosts.) A pathogen population composed of two genotypes in a 1:1 mixture, and no differences in invasion probability for these genotypes, were assumed. In Panel A, the dose response relationship is illustrated. On the x-axis is the log of dose, and on the y-axis mortality. Note that as the host population becomes more heterogeneous (i.e. the difference in pathogen invasion probability for the two halves of the host population becomes larger), the dose response relationship becomes smoother and bimodality more pronounced. In Panel B is the frequency of dual genotype establishment, which follows a similar trend. Panel C is the relationship between host mortality (x-axis) and the frequency of dual genotype establishment (y-axis). The solid line is a 1:1 relationship between mortality and the frequency of dual genotype establishment. For all heterogeneous host populations (i.e. 3, 10 and 100), the frequency of dual genotype establishment is higher than for a homogenous population (i.e. 1).


A
Mortality
Log Dose
B
Frequency of Dual Genotype Establishment
Log Dose
C
Frequency of Dual Genotype Establishment
Mortality
1
3
10
100

The mechanisms we considered were (1) differences in host susceptibility, (2) opportunism in the infection process: if one virion breaks the host systemic defense, than all invaded virions are represented in the viral progeny generated and detected by qPCR, and (3) a combination of these two mechanisms. We were surprised to find that the assumption of heterogeneity in host susceptibility was sufficient to describe the infection process (Figure 6). Interestingly, this mechanism has been previously suggested by Ridout *et al.* (1993) and Bianchi *et al.* (2002) for understanding dose-response, in the latter case also specifically for late instar larvae. We had not anticipated that the assumption of heterogeneity in host susceptibility would also help reconcile the IAH model with the observed higher frequency of dual genotype presence.

But why does host heterogeneity lead to a higher frequency of dual genotype presence? Consider a hypothetical situation similar to our experimental setup: hosts are challenged with a pathogen population composed of two genotypes. The host population is assumed to be heterogeneous with respect to pathogen susceptibility: half of the host population is susceptible and half is resistant. Using the IAH model, the effects of host heterogeneity on dose-response and the frequency of dual genotype establishment can be predicted. A simple way of doing this is first considering the susceptible and resistant host populations separately, and subsequently summing model predictions (e.g. mortality in susceptible hosts + mortality in resistant hosts): both relationships become bimodal (Figure 8A and B). If for these same instances we

Figure 9 (Right): The effects of heterogeneity in host susceptibility. Here we assume a population that is composed of 1, 2, 4 or 6 classes of individuals, varying in their susceptibility to a pathogen. The number of classes is given in the legend, and applies to all three panels. (E.g. A value of '1' indicates the invasion probabilities are the same for all hosts, so in this instance the host population is homogenous. '6' indicates that there are six host classes with different susceptibilities to the pathogen) The following invasion probabilities were assumed: 1 class, $10^{-4.5}$; 2 classes, 10^{-4}, 10^{-5}; 4 classes, 10^{-3}, 10^{-4}; 10^{-5}, 10^{-6}; 6 classes, 10^{-2}, 10^{-3}, 10^{-4}; 10^{-5}, 10^{-6}, 10^{-7}. The geometric mean invasion probability is $10^{-4.5}$ in all four cases. A pathogen population composed of two genotypes in a 1:1 mixture, and no differences in invasion probability for these genotypes, were assumed. In Panel A, the dose response relationship is illustrated. On the x-axis is the log of dose, and on the y-axis mortality. Note that as more host classes are introduced, the dose response relationship becomes shallower. In Panel B is the frequency of dual genotype establishment, which follows a similar trend. Panel C is the relationship between host mortality (x-axis) and the frequency of dual genotype establishment (y-axis). The solid line is a 1:1 relationship between mortality and the frequency of dual genotype establishment. As a more heterogeneity is introduced in the host population, the frequency of dual genotype establishment becomes higher and eventually approaches the 1:1 line (both pathogen genotypes are established in all hosts at any level of mortality).


A
Mortality
1.0
0.8
0.6
0.4
0.2
0.0
1
2
3
4
5
6
7
8
Log Dose
B
Frequency of Dual Genotype Establishment
Log Dose
C
Frequency of Dual Genotype Establishment
Mortality
0.0
0.2
0.4
0.6
0.8
1.0
1
2
4
6

plot mortality vs. the frequency of dual genotype establishment, it becomes evident that differences in pathogen susceptibility lead to a higher frequency of dual genotype presence at all levels of mortality (Figure 8C).

For a somewhat more realistic illustration, the host population is subdivided into a greater number of classes of susceptibility, covering a logarithmic series for the susceptibility parameter. The relationship between dose and mortality (Figure 9a), and dose and dual genotype establishment frequency (Figure 9b), become very smooth as heterogeneity in the host population increases. The relationship between mortality and dual genotype establishment (Figure 9c) approaches the line where both genotypes are established in all diseased hosts, as host heterogeneity is increased.

These figures also suggest a reason why the IAH model was convincingly rejected for intermediate mortality, but not at low and high mortality (Figure 4): differences in host susceptibility may lead to the greatest disparity between models with a constant or variable invasion chance at intermediate mortalities (e.g. Figure 6C). We expect that variation in host susceptibility will always lead to a higher frequency of dual genotype establishment – analogous to that it leads to a shallower dose-response relationship (Ridout *et al.*, 1993) – but a mathematical proof of this postulate must still be formulated.

We conclude that independent action may very well apply in those instances where we previously suggested an IAH-based model with a constant probability of host infection be rejected (Chapter 4): variation in host susceptibility leads to a higher frequency of dual genotype establishment. There is an interesting observation in our previous work that corroborates this conclusion: in those instances that the IAH model was not rejected, the frequency of dual genotype invasion was consistently higher- although not significantly so - than model predictions (Chapter 4). Since there will always be minor differences in host susceptibility, this would be expected under the framework we have adopted here. This raises the question of why there would be greater differences in host susceptibility in late instar larvae (L5) than in early instar larvae (L3). Two possible answers are: (1) Differences in intrastadial developmental resistance - an increase in host resistance within an instar (Hoover *et al.*, 2002, Grove *et al*., 2007) - arise quicker in L5 than in L3. Since for all our experiments larvae were selected in a fixed time window (12 h), this is a possible explanation. Variation may therefore be reduced when shorter windows (e.g. 2 h) are taken, and hence the original IAH-based model (Chapter 4; Druett, 1952) would then be applicable again. (2) Larval resistance to baculoviruses has a high phenotypic plasticity (Goulson and Cory, 1995; Reeson *et al*., 1998), and the variability in resistance therefore increases over development, even in a largely homogenous, artificial environment. Selection of larvae

based on whether melanization (an increase in skin pigmentation) has occurred or not, will probably lead to a more homogenous host population since melanization has been shown to be linked to resistance (Reeson *et al.*, 1998).

5. Appendix

Given that models 1, 2 and 3 are all special cases of model 4, we provide a detailed derivation of model 4 only. Model 1 and Model 2 are similar to the IAH model (Figure 10A), except that λ values must be multiplied by ξ, and model 2 accounts for variability in ξ in the host population (Figure 10B). Model 3 is analogous to Model 4 (Figure 10C), except that ξ is constant for model 3 and beta-distributed for Model 4.
Conditional upon a realization ξ that is (1) independent for each virion and (2) equal for all virions that have invaded a particular host, the number of invaders causing host disease - represented as the stochastic variable Ω - is Poisson distributed. A realization ω of this number given the fixed chance in a host of ξ and the total number n of invaded virions of genotypes A and B is:

$$\text{(A1)}\quad \Pr(\Omega = \omega \mid N = n \wedge \sigma = \xi) = \binom{n}{\omega} \xi^{\omega} (1-\xi)^{n-\omega}$$

However, we want the unconditional probability of the number of invaders causing disease. Therefore, we first include that the realization n comes from a Poisson distribution by using Bayes' rule and sum over all possible values for n:

$$\text{(A2)}\quad \Pr(\Omega = \omega \mid \sigma = \xi) = \sum_{n=0}^{\infty} \binom{n}{\omega} \xi^{\omega} (1-\xi)^{(n-\omega)} \left(\frac{\lambda^n}{n!} e^{-\lambda} \right)$$

It is easily shown that $\Pr(\Omega = \omega \mid \sigma = \xi)$ follows a Poisson distribution with mean $\xi \cdot \lambda$.
Above, ξ is a realization of the beta distribution:

$$\text{(A3)}\quad P(\sigma < \xi) = \int_0^{\xi} f(x)dx = \int_0^{\xi} \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)} dx$$

Here α and β are the two parameters that determine the shape of the beta distribution and $B(\alpha,\beta)$ is the beta function. The mean probability that an individual causes

systemic infection is $E(\sigma) = \alpha/(\alpha+\beta)$. From (A2) and (A3) we derive:

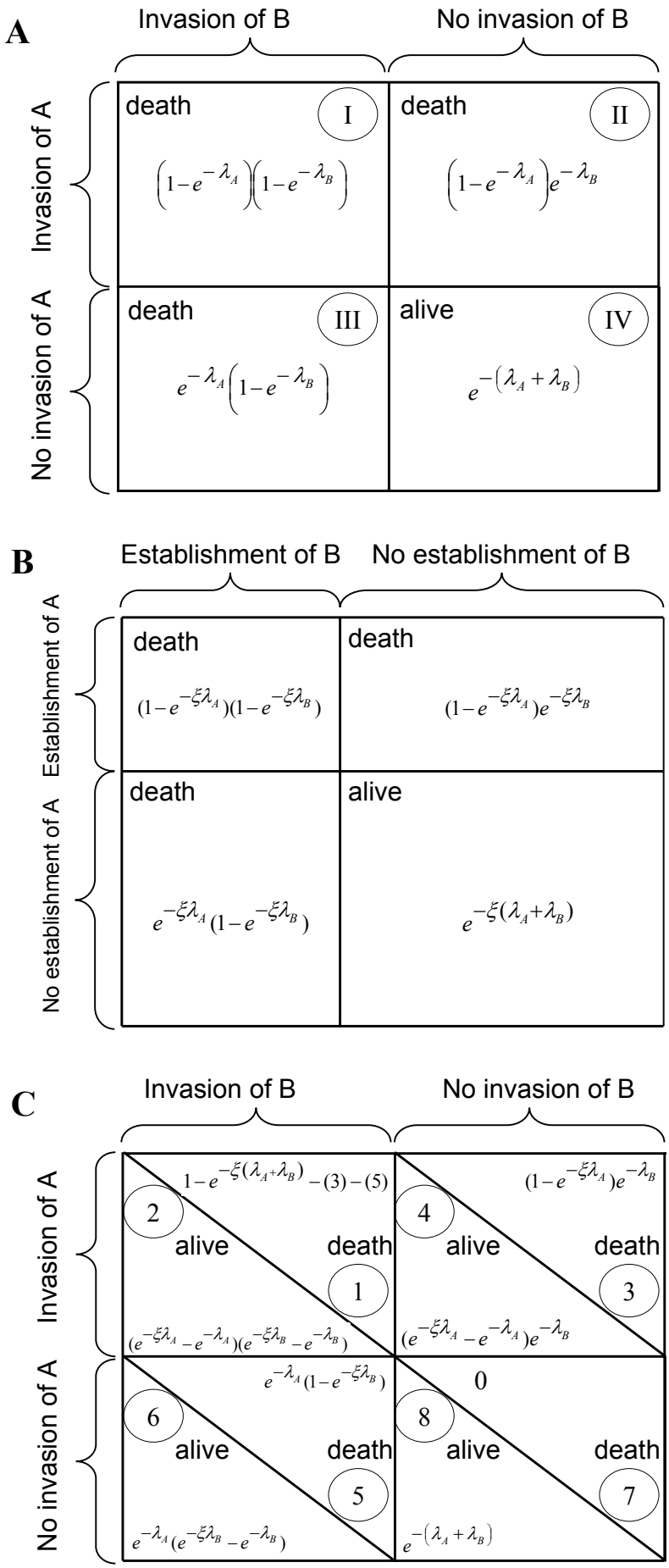
$$\text{(A4)}\quad P(\Omega=\omega) = \int_0^1 \frac{(\xi\lambda)^{\omega}}{\omega!} e^{-\xi\lambda} \frac{\xi^{\alpha-1}(1-\xi)^{\beta-1}}{B(\alpha,\beta)} d\xi$$

Hosts can survive if no virions managed to invade and if virions did invade but subsequently were not successful in causing host disease. Host mortality *(M)* therefore follows from subtracting the zero-term ($\Pr(\omega=0)$) of the Poisson distribution for the number of invaders successfully causing host disease from 1 (i.e. subtracting the proportion hosts surviving from the full host population at risk):

$$\text{(2)}\qquad M = 1 - P(\Omega=0) = 1 - \int_0^1 e^{-\xi\lambda} \frac{\xi^{\alpha-1}(1-\xi)^{\beta-1}}{B(\alpha,\beta)} d\xi$$

For clarification of the differences between the IAH model derived in Chapter 4 and the currently derived model 4 we refer again to Figure 10. When invasion inevitably leads to disease, and in our system death, IAH applies (Figure 10A), and the host survives only if neither A or B invade. Considering that here each invading virion has a chance of killing the host we arrive at the panels in figure 10C, where it is clear that if no background mortality is assumed panel 7 (Figure 10C) has a zero probability and panel 8 (Figure 10C) has equal probability as under IAH (panel IV, Figure 10A).

Figure 10 (Right): A diagrammatic representation of models presented in this paper. Above in A is a depiction of the null model: IAH with a fixed chance of invasion. If one pathogen invades, this will lead to disease. The host will only survive if both A and B do not invade and cause disease. In B, models 1 and 2 are illustrated. Invasion no longer automatically leads to disease, but there is a chance ξ that an invaded pathogen will cause disease, a process we refer to as establishment in the figure. For model 1, ξ is constant. This model is similar to the IAH-based model with a fixed invasion chance, but ξ values of less than one reduce the levels of host disease and co-establishment. For model 2, the probabilities apply to a single insect host individual, and they are conditional upon a realization ξ from the beta distribution. To obtain the probabilities for model 2 at the level of the population of the host, we have to multiply by the PDF of the beta distribution and integrate over the domain of ξ; from zero to one. C depicts models 3 and 4, were there is dependent action and pathogens which have invaded can benefit from other pathogens causing host disease. Given that we have assumed independent action in the process of invasion, the squares in the panels A and C (I through IV and 1+2 through 7+8) represent equal probabilities with respect to the distribution of invaded genotypes (e.g. III = 5+6). In models 3 and 4, only a proportion of the hosts invaded by any combination of genotypes dies. For model 3, ξ is constant. For model 4, the probabilities are again conditional upon a realization ξ from the beta distribution and are obtained as described for model 2.


A
Invasion of B
No invasion of B
Invasion of A
No invasion of A
death
I
$\left(1-e^{-\lambda_A}\right)\left(1-e^{-\lambda_B}\right)$
death
II
$\left(1-e^{-\lambda_A}\right)e^{-\lambda_B}$
death
III
$e^{-\lambda_A}\left(1-e^{-\lambda_B}\right)$
alive
IV
$e^{-\left(\lambda_A+\lambda_B\right)}$
B
Establishment of B
No establishment of B
Establishment of A
No establishment of A
death
$(1-e^{-\xi\lambda_A})(1-e^{-\xi\lambda_B})$
death
$(1-e^{-\xi\lambda_A})e^{-\xi\lambda_B}$
death
$e^{-\xi\lambda_A}(1-e^{-\xi\lambda_B})$
alive
$e^{-\xi(\lambda_A+\lambda_B)}$
C
Invasion of B
No invasion of B
Invasion of A
No invasion of A
$1-e^{-\xi(\lambda_A+\lambda_B)}-(3)-(5)$
2
alive
death
1
$(e^{-\xi\lambda_A}-e^{-\lambda_A})(e^{-\xi\lambda_B}-e^{-\lambda_B})$
$(1-e^{-\xi\lambda_A})e^{-\lambda_B}$
4
alive
death
3
$(e^{-\xi\lambda_A}-e^{-\lambda_A})e^{-\lambda_B}$
$e^{-\lambda_A}(1-e^{-\xi\lambda_B})$
6
alive
death
5
$e^{-\lambda_A}(e^{-\xi\lambda_B}-e^{-\lambda_B})$
0
8
alive
death
7
$e^{-\left(\lambda_A+\lambda_B\right)}$

Now we derive expressions for the panels (1)-(6) in figure 10C. For the full host population the probability for the number of invaders causing disease to be ω conditional upon the fixed chance ξ in a host is:

(A5) $$P(\Omega = \omega \mid \sigma = \xi) = e^{-\xi(\lambda_A+\lambda_B)} \frac{(\xi(\lambda_A + \lambda_B))^{\omega}}{\omega!}$$

Thus the total fraction that dies given ξ, i.e. the sum of the panels (1), (3), (5) and (7) (Figure 10C) is:

(A6) $$1 - P(\Omega = 0 \mid \sigma = \xi) = 1 - e^{-\xi(\lambda_A+\lambda_B)}.$$

For invasion with a number of *k* virions of genotype A given that no type B virions invaded and given the fixed chance ξ of killing the host for each virion is

(A7) $$P(\Omega = \omega \wedge A = k \mid B = 0 \wedge \sigma = \xi) = e^{-\lambda_A} \frac{{\lambda_A}^k}{k!} \binom{k}{\omega} \xi^{\omega} (1-\xi)^{k-\omega}$$

Using Bayes' rule and summing over all different realizations of *k* yields that $P(\Omega = \omega \mid B = 0 \wedge \sigma = \xi)$ follows a Poisson distribution with mean $\xi \cdot \lambda_A$. Therefore, the formula for the fraction dead in the "only invaded by genotype A" class as represented by panel (3) in figure 10C is:

(A8) $$\big(1 - P(\Omega = 0 \mid B = 0 \wedge \sigma = \xi)\big) P(B = 0) = (1 - e^{-\xi\lambda_A}) e^{-\lambda_B}$$

Because of the fact that genotype A and genotype B occur symmetrically in the formula panel (5) of figure 10C should contain $(1 - e^{-\xi\lambda_B})e^{-\lambda_A}$. The formula for panel (1) is now easily derived from (A6), (A8) and this last formula as:

(A9) $$1 - e^{-\xi(\lambda_A+\lambda_B)} - (1 - e^{-\xi\lambda_A})e^{-\lambda_B} - (1 - e^{-\xi\lambda_B})e^{-\lambda_A}.$$

Acknowledgements:

We thank Frans van Aggelen, Andre Gidding and Leo Koopman for providing *S. exigua* and *M. brassicae* eggs. Karel Keesman advised on the use of lsqnonlin.

Chapter 6

Mixed infections and the competitive fitness of faster-acting genetically modified viruses

---

Mark Zwart, Wopke van der Werf, Monique M. van Oers, Lia Hemerik,
Jan M.V. van Lent, J. Arjan G.M. de Visser, Just M. Vlak, and Jenny S. Cory

A slightly modified version of this chapter has been submitted for publication.

Abstract

Faster-acting recombinant baculoviruses have shown potential for improved suppression of insect pests, but their ecological impact on target and non-target hosts and naturally occurring pathogens needs to be assessed. Previous studies have focused on the fitness of the recombinant at the between-hosts level. However, the population structure of the transmission stages will be decided by within-host selection. Here we have experimentally quantified the within-host competitive fitness of a fast-acting recombinant *Autographa californica* multicapsid nucleopolyhedrovirus missing the endogenous *egt* gene (vEGTDEL), by means of direct competition in single and serial passage experiments with its parental virus. A quantitative real-time PCR assay was employed to determine the ratio of these two viruses in passaged mixtures. We found that vEGTDEL had reduced within-host fitness: per passage the ratio of wild type to vEGTDEL was on average enhanced by a factor of 1.6. There is also frequency-dependence: the lower the frequency of the Wt-virus, the stronger it is selected. Additionally, the virus ratio is a predictor of host death time and virus yield. Our results show that *egt* is an important determinant of within-host fitness and provide a framework for a more complete assessment of the ecological impact of recombinant baculovirus release.

1. Introduction

Insect pathogens offer viable alternatives to synthetic chemical pesticides for the suppression of insect pests. One particular pathogen, the bacterium *Bacillus thuringiensis* (or one or more of its toxins), is widely used in spray applications or by the incorporation of its toxins into a variety of crop plants (Chistou *et al.,* 2006). Representatives of other groups of entomopathogens, such as viruses and fungi, have a long history of use and the occasional widespread success (Moscardi, 1999; see also Scholte *et al.,* 2005), but in general they have not achieved their potential as insecticides. One reason for this lack of take-up is that they are perceived as slow acting, at least in comparison to synthetic chemicals and insect growth regulators. This has led to the development of faster-acting genetically modified pathogens, particularly in the insect baculoviruses (Incegolu et al. 2006).

Baculoviruses are highly pathogenic for certain insect species and can have a major impact on the population dynamics of their host (Cory and Myers 2003). It is therefore

important to be able to predict whether the release of genetically modified baculoviruses will displace native wild type strains and alter the dynamics of their host or hosts. To date, this issue has received some theoretical treatment (Dushoff and Dwyer 2001, Bonsall et al. 2005) and contained field experiments have compared the horizontal transmission of recombinant and natural viruses (Hails *et al.* 2002; Cory *et al.*, 1994). These limited studies focus on between host transmission: however, the interactions within the host can be equally important in parasite ecology and evolution and yet these have received little attention. Within-host selection will determine how many and what type of transmission stages are produced, and this will influence key traits, such as virulence, that are fundamental to the epizootiology of pathogens in natural populations (e.g. de Roode et al. 2005, Ben-Ami et al. 2008). Studies on multiple infection and within-host competition of parasites have tended to focus on how they shape the evolution of parasite virulence in relation to disease intervention strategies such as vaccination and the avoidance of drug resistance (e.g. de Roode et al. 2004). Within-host selection, however, is also of direct relevance to the risk assessment and environmental release of recombinant pathogens, such as those used for insect pest control.

The *Baculoviridae* are a large family of viruses of the arthropods, characterized by their rod shaped nucleocapsids and relatively large dsDNA genomes (Theilmann *et al*., 2005). Most known baculoviruses are infectious to Lepidopteran insect larvae, and cause lethal host infections. Due to their life-history traits, baculoviruses are considered obligate host killing parasites (Ebert and Weisser, 1997). The infection cycle is biphasic; after the initial infection, viral spread to different tissues within an individual host is mediated by single virions, known as budded virus (BV); BV has limited persistence outside of the host body. At the end of the infection cycle, the host insect dies and the cadaver releases what is typically a very large number of horizontal transmission stages, referred to as polyhedra for the Alphabaculoviruses (Jehle *et al.*, 2006; Federici, 1997). Polyhedra are proteinacious occlusions containing a large number of virions, which can remain infectious for long periods of time (Thomas, 1972; England, 1998). Thus horizontal transmission only occurs upon host death, and the horizontal transmission stage has prolonged persistence outside of the host. These characteristics are crucial to understanding baculovirus ecology and evolution.

A number of genetically modified baculoviruses with improved characteristics for pest control have been created (Inceoglu *et al.*, 2006). Many of these recombinant viruses kill the target pest more rapidly than the original wild type virus, leading to reduced crop damage in the field (Cory *et al*., 1994, Sun *et al.*, 2004). There are several reports in which the fitness components of wild-type viruses and their derived recombinants

have been experimentally determined in the laboratory. A faster speed of kill usually leads to reduced larval cadaver weight and virus yield (e.g. Burden et al. 2000, Cory *et al*., 2004), as the host insect has less time to develop and the virus in turn has less host tissue available to convert to transmission stages. Virus virulence, defined here as fatal infection, is not linked to speed of kill and thus pathogenicity is often unchanged in recombinant baculoviruses, compared to their parent wild type (e.g. Burden et al. 2000). This can vary with species, larval stage, bioassay design and the genes added and/or removed (e.g. Hernández-Crespo et al. 2001, Pineda et al 2003). However, no difference in pathogenicity compared to the parent wild type virus has been found in those few recombinants that have been tested under field conditions (Hernández-Crespo et al. 1999, Sun et al. 2004a). While understanding the effects of genetic modification on traits that affect efficacy and between-host transmission is crucial for predicting the environmental impact of recombinant insecticides, the population structure of the transmission stages will be decided by within-host selection. In particular, we need to ask, how will these fast-acting recombinant viruses fare when they are in direct competition with a wild-type virus in the same host?

Milks *et al.* (2001) found that a faster-acting recombinant baculovirus expressing *Androctonous australis* insect-selective toxin (AaIT) did not appear to have altered within-host fitness. However, in this study within-host fitness was evaluated only by considering passaged lines in which one virus went to fixation; what happened in replicates which remained co-infected over serial passage – the majority of cases - was not addressed. A few studies have focused on the co-infection dynamics of baculoviruses in general. Hamblin *et al.* (1990) showed that co-occlusion of different baculovirus genotypes in a single polyhedron can occur, using a polyhedrin negative virus (i.e. one that cannot produce the proteinacious matrix that the virions are embedded in). Baculovirus genotypes lacking the ability to autonomously generate polyhedra are maintained in serial transfer experiments, providing strong evidence for a high cellular multiplicity of infection (MOI) towards the end of infection *in vivo* (Bull *et al.* 2001, 2003). Wild type NPV populations, which are often mixtures of genotypes, can contain deletion mutants that have lost *per os* infectivity. Lopez-Ferber *et al.* (2003) demonstrated that a mixture of a naturally occurring *Spodoptera frugiperda* MNPV (SfMNPV) deletion mutant and a full-length SfMNPV genotype was more pathogenic than the full length virus alone. These deletion mutants reach an equilibrium frequency within a few transfers in serial passage experiments (SPEs) (Simon *et al.* 2006). Conversely, it was recently demonstrated that cell culture derived *Autographa californica* MNPV (AcMNPV) deletion mutants missing a gene essential to replication (*ie-1*) were quickly purged during *in vivo* passage, apparently due to low cellular MOI at the beginning of infection (Zwart *et al.*, 2008 [1]).

The general picture that emerges from these studies is that NPV genotypes that can achieve horizontal transmission are not quickly purged from populations, so long as they are not defective in autonomous replication. Given that fast-acting recombinant baculovirus genotypes can be fully capable of both autonomous replication and horizontal transmission, they may therefore very well co-exist in natural populations for long periods of time, despite having reduced performance with respect to certain fitness components (Dushoff and Dwyer 2001; Milks *et al.*, 2001). In fact, it could be suggested that fast-acting viruses may have enhanced within-host competitive fitness, as many of these viruses have endogenous genes deleted in order to increase the speed of kill (e.g. the ecdysteroid UDP-glucosyltransferase gene, *egt* (O'Reilly and Miller, 1989). This results in a marginally smaller genome and therefore the potential for faster DNA replication. There is therefore even a possibility, albeit remote, that certain fast-acting recombinant viruses could out-compete wild-type viruses at the within-host level. As natural baculovirus isolates are typically composed of numerous genotypes (Cory *et al*., 2005, Smith and Crook, 1988, Lee and Miller, 1978), co-infections appear to be the norm in the field under natural conditions. This makes understanding the competitive process at this level of selection all the more important.

To address the issue of the within-host competitive fitness of a fast-acting virus genotype, we chose to study a baculovirus lacking the *egt* gene. This recombinant virus (in future called vEGTDEL) was derived from the AcMNPV Wt L1 strain by deletion of a part of the ecdysteroid UDP glucosyltransferase ORF (O'Reilly and Miller, 1989, see also Figure 2b). The product of this gene (EGT) inactivates ecdysteroids by conjugation with glucose, which inhibits molting. The virus kills hosts faster than the wild-type virus and shows reduced yield in single genotype infections (Cory *et al.*, 2004). Deletion of the *egt* gene is a commonly used approach for constructing recombinant baculovirus and is often combined with the incorporation of second gene, such as an insect selective toxin, which increases speed of kill further (Inceoglu *et al*., 2005).

In order to evaluate the within-host competitiveness of vEGTDEL, direct competition experiments with its parental wild-type virus were performed (Elena and Lenski, 2003). A calibrated quantitative real-time PCR (qPCR) assay (e.g. Zwart *et al*., 2008 [2]) for the Wt L1 and vEGTDEL genotypes was developed to determine the ratio of the two genotypes in individual larval cadavers. Single passage experiments with a range of mixtures of the Wt L1 and vEGTDEL polyhedra used as inoculum were performed. Serial passage experiments with a 1:1 mixture of polyhedra were also performed. Both sorts of experiments were carried out as: (1) Different baculovirus genotypes can be co-occluded in the same polyhedron if the host is infected with

multiple genotypes (Hamblin *et* al., 1990; Bull *et* al., 2001). Since the initial innoculum used in these experiments was a mixture of single genotype polyhedra, using these two set ups also allows us to explore if co-occlusion – presumably occurring after the first passage in the serial passage experiment - has any effect on the persistence of the genotypes. (2) The serial passage setup is inherently more complex, due to - for example - co-occlusion and differences between larval cohorts. As such, it offers a good test of the results measured in single passage experiments.

2. Methods:

2.1 Insects and viruses

Trichoplusia ni larvae were reared as described for *Spodoptera exigua* (Smits *et al*., 1988), with minor modifications: (1) in the artificial diet used, the cornmeal was replaced by an equal amount of extra wheat germ (2) adult moths were fed a 5% sucrose solution instead of only water. The Wt L1 virus and vEGTDEL were amplified separately by inoculating 30 fourth instar *T. ni* with a high dose (approx. 100 x LD_{99}), and collecting larval cadavers upon death and storing them at –20° C. Polyhedra were purified from cadavers as described by Zwart *et al.* (2008 [2]).

2.2 Electron microscopy

Electron microscopy of thin sections of polyhedra of Wt L1 and vEGTDEL viruses was performed as described by Zwart *et al.* (2008 [2]). The number of nucleocapsids per occlusion derived virus (ODV) was counted for approximately 100 suitably oriented ODV for each genotype. The Kolmogorov-Smirnoff test (KST) was used to test for differences between the two viruses in the number of nucleocapsids per ODV. KST tests for differences in the whole probability density function of nucleocapsids per ODV, not just different means.

More than 100 randomly selected polyhedron cross sections were digitally photographed for each virus. The number of ODV per polyhedron cross section was counted and the surface area was measured using AnalySIS software (Olympus Soft Imaging Solutions). The Mann-Whitney *U*-test was used to test for differences in mean cross-sectional area and mean number of ODV per polyhedron cross-section between the Wt L1 and vEGTDEL viruses.

2.3 qPCR

For specific detection of the Wt L1 virus, the forward primer 5'-GTCGTCGTGGAAGCGTTTGCC-3' and the reverse primer 5'-TCGGCCAAACCGTAGCCAGG-3' were used. For detection of the vEGTDEL, the forward primer 5'-CGTTACGGTCGTCAAGCCCAAACTGTTTG-3' and the reverse primer 5'-TCGAATTACGCGTTCTGTGAATTTGATGGC-3' were used. Note that the vEGTDEL primers can in principle also detect the Wt L1 virus, but will give a much larger product (2 kbp, see Figure 2B). Since vEGTDEL contains no unique DNA sequences, the only way to discriminate between the two viruses with a PCR-based assay was to use primers for the amplicon spanning the junction of the deletion in the *egt* ORF, and to keep elongation periods in the PCR program as short as possible (10 seconds). When PCR was performed on DNA from pure vEGTDEL, only the short PCR product corresponding to vEGTDEL could be observed on an agarose gel. No product was observed when the Wt L1 virus was used (data not shown). During qPCR runs, the "melting curve" peak observed was consistently corresponded with the short PCR product of vEGTDEL, and no peaks of higher dissociation temperatures were ever observed. Thus, the PCR for vEGTDEL appeared to be specific to this genotype.

Wt L1 and vEGTDEL PCR products from a regular PCR reaction were electrophorised in 1% agarose gels, excised from the gel and subsequently purified using the DNA Extraction Kit (Fermentas). The purified DNA was cloned into the pGEM-T easy vector (Promega), and its identity checked by PCR and restriction enzyme analysis. Plasmid DNA concentration was quantified by photospectrometry and gel electrophoresis. A dilution with 10^7 plasmids μl^{-1} containing the respective fragment (Wt L1, vEGTDEL) was used as a calibrator sample in qPCR reactions. Quantification of virus genotype ratios by qPCR was performed as described elsewhere (Zwart *et al*., 2008 [2]), with minor modifications: (1) an annealing temperature of 58° C instead of 60° C was used and (2) an elongation time (72° C) of 10 seconds was used.

Prior to analysis of samples, mixtures of polyhedra were used to calibrate the qPCR assay, as described (Zwart *et al*., 2008 [2]). Briefly, pure Wt L1 and vEGTDEL polyhedra, and a range of mixtures (1:10^4 to 10^4:1) with 10-fold intervals were quantified by qPCR. The polyhedra input ratio was corrected for differences in mean number of ODV per polyhedron (see Results 3.1; ODV per polyhedron Wt L1:vEGTDEL = 1.830). Regression analysis on log transformed genotype ratios was performed to determine the accuracy of the PCR assay. The calibration experiment was performed twice. The regression line obtained from the combined data of the PCR

calibration experiments was used to correct the measured ratios and obtain actual genotype ratios, for all subsequent samples analyzed.

For all experimental samples the genotype ratio (R) was calculated, which could subsequently be used to determine the proportion of Wt L1 virus in the sample (f_{wt}):

(1) $$\log(R) = \log\left(\frac{f_{wt}}{1-f_{wt}}\right) = \frac{\log(c_{wt}/c_{ed}) - b}{m}$$

Here c_{wt} and c_{ed} are the measured copy numbers obtained from the qPCR assay, and m and b are the slope and y-intercept of the qPCR calibration regression line.

2.4 Single passage mixed infection of larvae

Newly molted fifth instar *T. ni* larvae were selected by head capsule slippage within a 12 h time frame. The larvae were then kept on artificial diet at 27° C for 24 h before being challenged with mixtures of polyhedra. Polyhedra of the Wt L1 virus and vEGTDEL were quantified by counting with a haemocytometer (20 counts per virus). The polyhedra were then diluted in water to give a final concentration of 10^6 polyhedra ml^{-1}. Wt L1 and vEGTDEL were mixed in ratios of 1:100, 1:10, 1:1, 10:1 and 100:1. Two µl of uncontaminated water, pure virus suspension or one of the mixtures was then pipetted onto a small plug of artificial diet in a 12-well tissue culture plate and allowed to dry briefly, giving a virus dose of 2000 polyhedra per larva (LD_{80}). Twenty-four larvae were taken for the non-virus control, and 36 larvae for each treatment. The synchronized larvae were then added to wells individually, and kept at 27° C for a further 12 h. Only those larvae which had consumed all the diet were individually transferred to new 12-well plates with fresh, uncontaminated artificial diet. Mortality was recorded every 12 h until all larvae had died or pupated, and larval cadavers collected and individually stored in 1.5 ml Eppendorf tubes at -20° C.

Fourteen larvae from each inoculum mixture, and five larvae from the pure Wt L1 and vEGTDEL treatments, were randomly selected for qPCR analysis. Larval weight was determined by weighing the tubes containing larval cadavers, and subtracting the weight of the tube. Larvae were macerated in 500 µl milliQ water. A 1:100 dilution of the resulting virus suspension was counted twice with a haemocytometer to determine polyhedra yield. The larval remains were filtered through cheesecloth #80 (muslin), and DNA was then extracted as described elsewhere (Zwart *et al.*, 2008 [2]). DNA was diluted 1:200 in milliQ water prior to qPCR analysis.

As a tool for the interpretation of the results obtained here, an estimate of the number of virions invading a host larva was made. The mean number of invaders that best describes the frequency of dual genotype infection observed (equation 7, Chapter 4) was found by comparing predicted and observed values for the frequency of dual genotype infection using the sum of squares. This analysis still assumes there is independent action in the invasion of the host (but not that invasion automatically leads to host death), and that a given number of virions entering are required for pathogenesis.

2.5 Serial passage mixed infection experiments

For passaging from individual larvae, a 1:1 starting mixture of polyhedra was used. Per replicate, 10 larvae were challenged as described for the single passage infection experiment, with the same procedure and virus dose. A randomly selected, individually stored larval cadaver was used for polyhedra purification. Each larva was macerated and the remains filtered through cheesecloth. Polyhedra concentrations were quantified by counting in a haemocytometer (at least 2 counts per replicate). These polyhedra were used to initiate the next round of infection. Fifteen replicates were performed: five replicates in an initial experiment and ten replicates in a subsequent experiment. Five serial passages were performed, and the data from the two experiments were combined for analysis. As controls for the passaging experiment, a single replicate of the Wt L1 virus and vEGTDEL was passaged in individual larvae. DNA extraction and qPCR were performed as described above for each replicate and control at every time point.

2.6 Estimating the within-host fitness of vEGTDEL

In order to analyze the qPCR data from the single passage experiment, a Generalized Linear Model (GLM, SAS) was used with the natural logarithm of the virus inoculum ratio as the independent variable, and the observed proportion of Wt L1 virus - calculated from the qPCR-measured virus ratio (Equation 1) - as the dependent variable. A logit link function and a binomial error structure were used. The number of invading virions estimated from the single passage data, as described above, was used for the number of trials for this binomial distribution. As our estimate of the number of invading virions is only an approximation, a value of 20 was used for the number of trials. The scale parameter was estimated from dividing the square root of the deviance by the degrees of freedom.

For the analysis of the serial passage experiments, the deterministic model for

selection between competing species proposed by de Wit was used (de Wit, 1960; Godfray *et al.*, 1997; Georgievska *et al.*, 2005). Given that the log of the virus genotype ratio was used for analysis, this model can be stated as:

(2) $$\log(R_t) = \log(R_0) + t \cdot \log(w)$$

Here w is the selection rate constant, a constant factor by which the virus genotype ratio is multiplied in each passage. In addition, t is the passage number, R_0 is the genotype ratio at passage 0, and R_t is the log genotype ratio at passage t. Using the observed mean R-values over passages 1 through 5, w was calculated by linear regression.

2.7 Estimating the overall fitness of vEGTDEL with respect to yield

A more elaborate model for competition between species (de Wit, 1960) was used to understand the competitive process better, and to explore the combined effect of differences in fitness at the within and between-host levels (using absolute yield of transmission stages as a surrogate for between-host transmission). For each virus, the yield of only that particular virus generated (y_{wt}, y_{ed}) when there is a given proportion of that virus in the inoculum (f_{wt}, f_{ed}) is then:

(3) $$y_{wt} = m_{wt} \frac{k_{wt} f_{wt}}{k_{wt} f_{wt} + f_{ed}}$$

and:

(4) $$y_{ed} = m_{ed} \frac{k_{ed} f_{ed}}{k_{ed} f_{ed} + f_{wt}}$$

Here m_{wt} and m_{ed} are the respective yields of the virus in single genotype infection, and k_{wt} and k_{ed} are crowding coefficients: a constant that describes how the two viruses will perform in direct competition with each other. For individual larvae, the yield of polyhedra for each genotype was calculated by partitioning the total yield (determined by counting polyhedra in a haemocytometer) using the qPCR-determined genotype

ratio. For single genotype treatments, the mean yield of all larval cadavers collected was determined and assumed to be pure, as no contamination was detected in a subset of these samples. Non-linear regression (SPSS 12.0) was then used to estimate m_{wt} and k_{wt}, and subsequently m_{ed} and k_{ed},

3. Results:

3.1 Polyhedra morphology of the Wt L1 and vEGTDEL viruses

From phase-contrast microscopy, it was suspected that there were differences in polyhedron size between the Wt L1 and vEGTDEL viruses (Figure 1). To define the

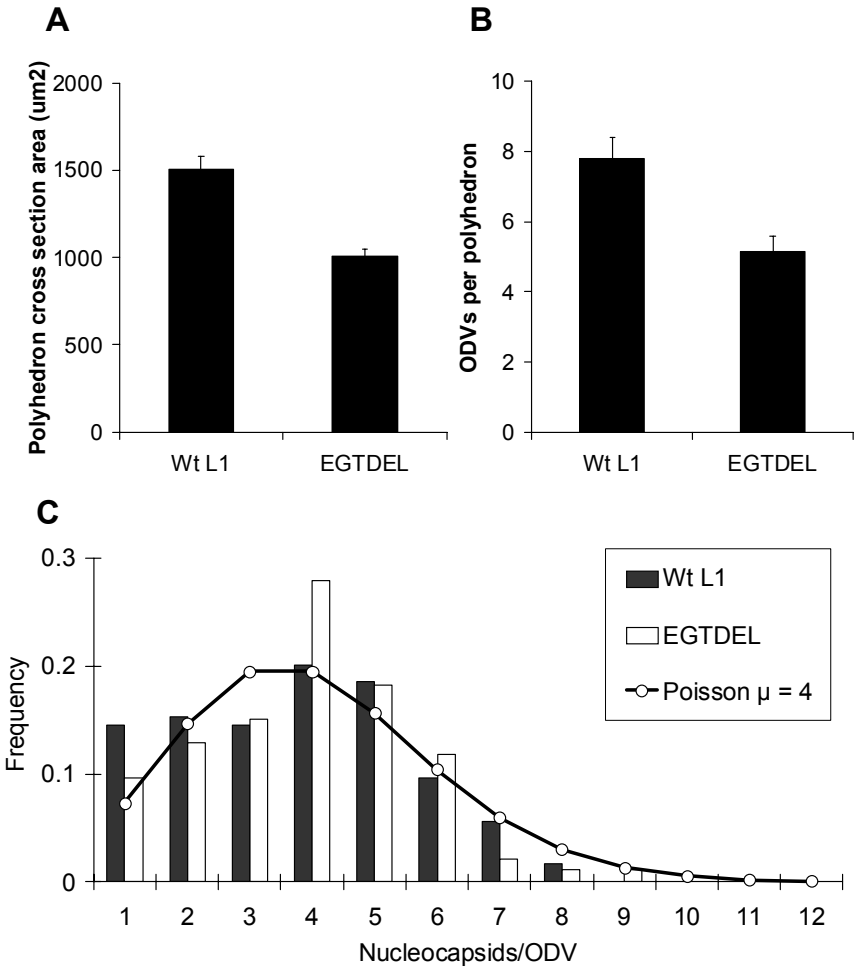


Figure 1: The morphology of polyhedra of the Wt L1 and vEGTDEL viruses. A. The mean polyhedron cross section area in square micrometers, with standard error. B. The mean number of ODV per polyhedron, with standard error. C. The frequency distribution of nucleocapsids per ODV, and a Poisson distribution with a μ (mean) of 4 for comparative purposes.

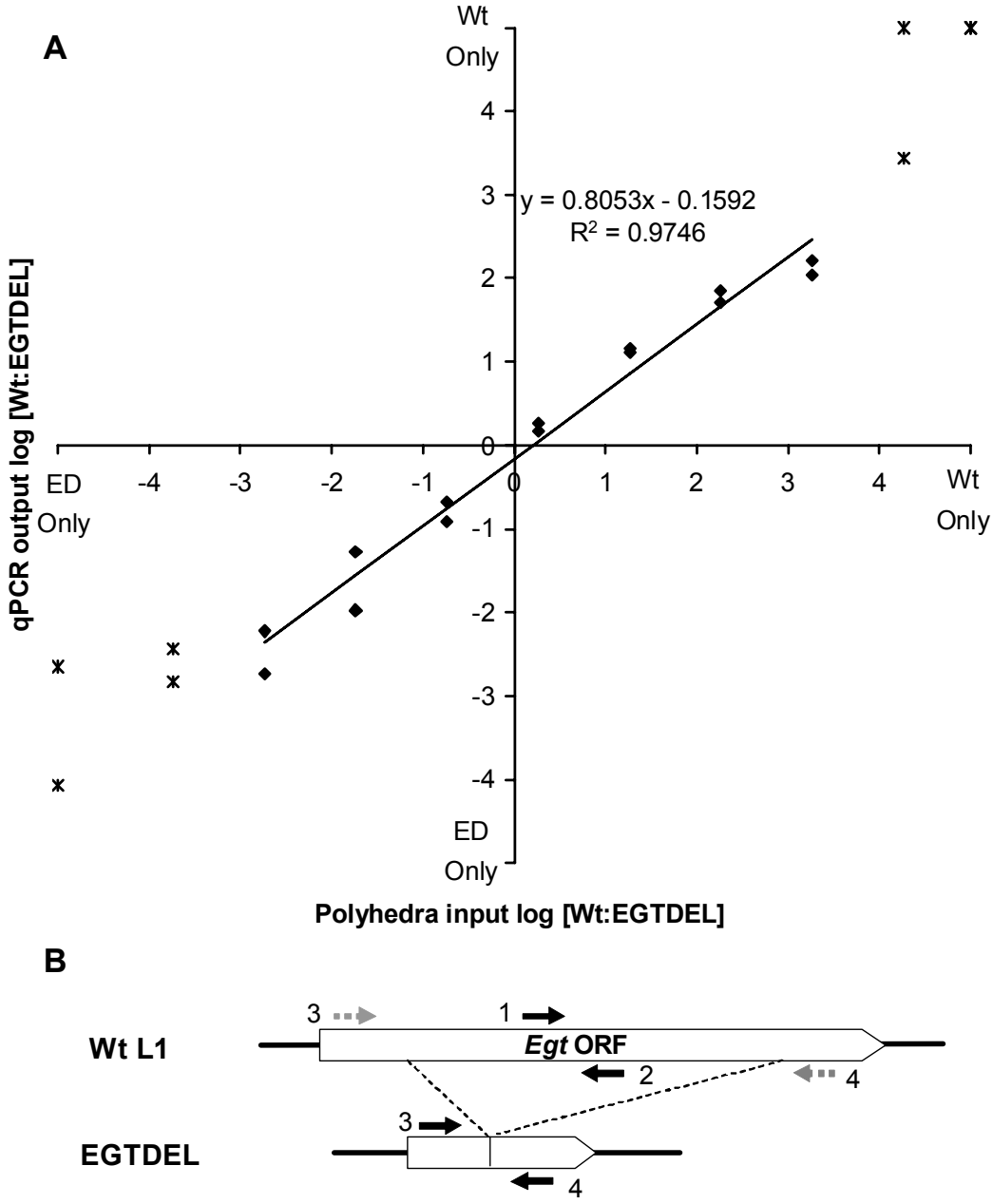


Figure 2: A. The data of the qPCR calibration experiment data. On the x-axis is the log transformed input ratio (Wt L1:vEGTDEL), the polyhedra ratio corrected for the difference in ODV content per polyhedron between the two genotypes. On the y-axis is the log transformed ratio measured by qPCR (Wt L1:vEGTDEL), for two replicate experiments. A regression line is shown, with equation and r^2-value. The black diamonds were included in the regression analysis, whereas the crosses were not because the assay was not satisfactory in this range. B. The location of the primer sets for identification of the Wt L1 and vEGTDEL viruses. PCR primers are denoted by numbered arrows: 1 and 2 are the forward and reverse primers for the Wt virus; 3 and 4 the forward and reverse primers for vEGTDEL detection. Note that although primers 3 and 4 can in principle anneal and amplify a product from the Wt L1 DNA, whether or not this template will be amplified depends on the thermal cycling program used.

starting material for experiments, EM analysis of polyhedra of the two genotypes was performed. The mean polyhedron cross sectional area of the Wt L1 virus was significantly larger than that of the vEGTDEL (Figure 1A; Mann Whitney *U*-test: $Z = -5.747$, N = 208, $P < 0.001$). The difference in cross sectional areas was approx 1.5-fold. The polyhedra of the Wt L1 virus also had significantly more virions per cross section than vEGTDEL polyhedra (Figure 1B, Mann Whitney *U*-test: $Z = -3.400$, N = 208, $P = 0.001$), also approx. 1.5-fold. It was assumed that ODV are isotropically distributed in polyhedra, allowing conversion of the difference in ODV per polyhedron cross-sectional area to the actual difference in ODV per polyhedron. The estimated ODV per polyhedron ratio of Wt L1:vEGTDEL is then the observed ODV per polyhedron cross section ratio (= 1.496), raised to the power 3/2 (= 1.830) to translate to three dimensions.

In previous studies on the original wild type AcMNPV isolate and bacmid-derived viruses there was a zero-truncated Poisson-like distribution of the number of nucleocapsids per virion (Zwart *et al.,* 2008 [2]). A similar result was obtained here for the Wt L1 and vEGTDEL virions (Figure 1C). The distribution was similar for both genotypes (Kolmogorov Smirnoff test: $Z = 0.529$, N = 217, $P = 0.942$).

3.2 qPCR calibration

The qPCR assay performed satisfactorily over a range of six orders of magnitude (1000:1-1:1000; Figure 2). The regression equation allowed us to make a small correction to measured qPCR-values in order to obtain the true genotype ratios. The Wt L1 primers consistently gave a very late fluorescent signal with vEGTDEL alone. Over 10 passages in larvae, the signal remained in the same range (< 1:5000; Wt: vEGTDEL), indicating that it is probably background.

3.3 Estimating the number of infection founders

The ratio of Wt L1:vEGTDEL in individual larval cadavers was determined by means of qPCR (Figure 3). These data suggest that there may be selection for the Wt L1 genotype, as the ratio appears to be shifted towards the Wt L1 virus for many samples. In order to know whether one genotype displaced the other**,** an estimate of the number of infection founders would be valuable as it allows us to predict the role of genetic drift in modulating genotype frequencies during the infection process. The independent action hypothesis (IAH) – which allows for estimation of the number of founders based on the level of host survival – was previously rejected for AcMNPV infection of *T. ni* L5 (Chapter 4). This conclusion could be confirmed with our current

data on a different set of genotypes by using the same test, with the frequency of larvae infected by both genotypes being much higher than predicted by IAH (One-way binomial test (SPSS 12.0); 1:100 (Wt L1:vEGTDEL) to 10:1: $P < 0.001$; 100:1: $P = 0.005$). More virions than predicted by IAH appear to invade host larvae, requiring another method than IAH for estimating the number of founders.

In order to estimate the number of founders, we considered the frequency of dual genotype infection (Chapter 4). However, rather than assume that a single virion can

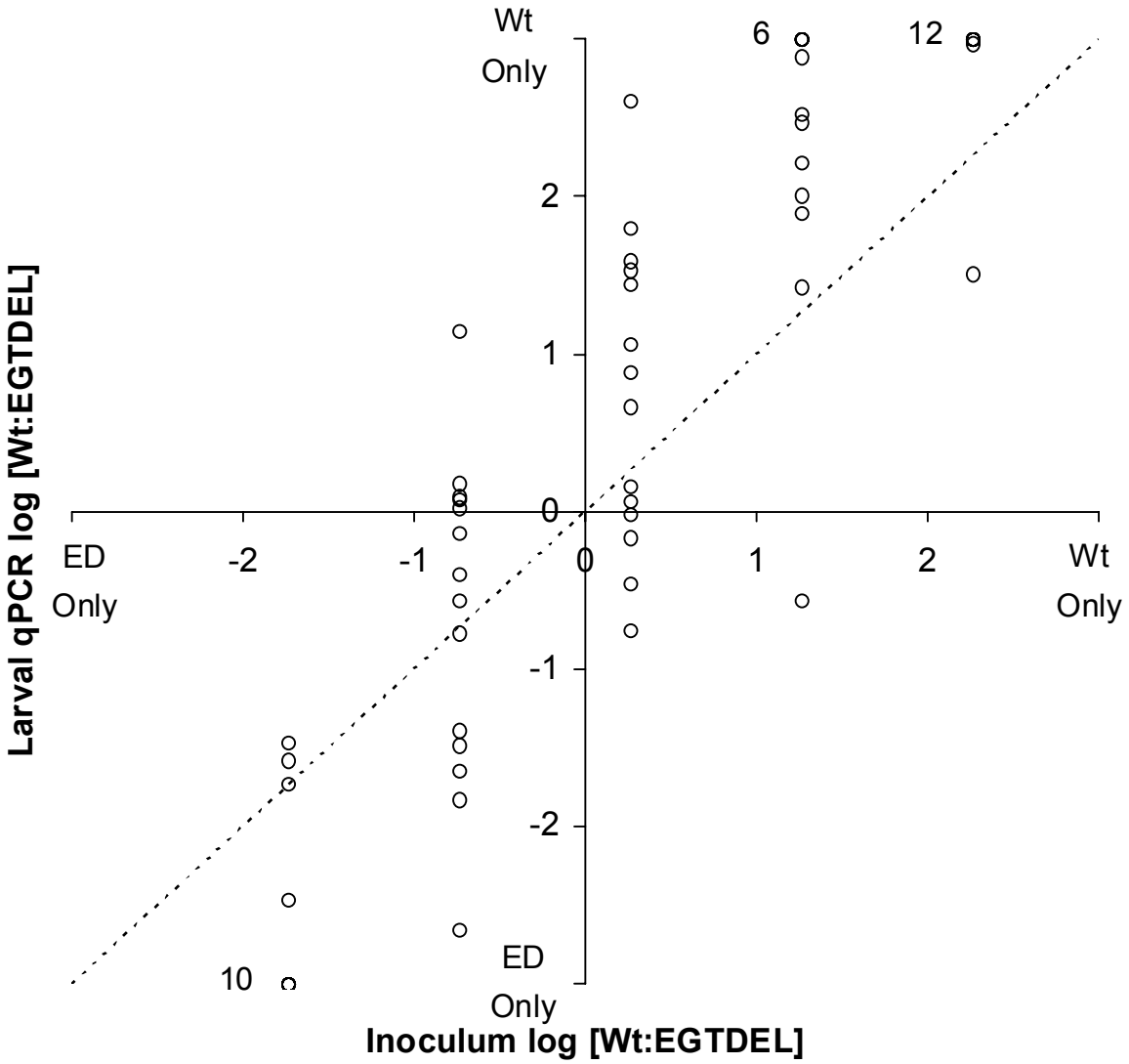


Figure 3: qPCR data from 1:100 through 100:1 polyhedra inoculum mixtures. The log of the ODV ratio in the inoculum (Wt L1: vEGTDEL) on the x-axis. The log ratio Wt L1: vEGTDEL measured in individual larvae by qPCR is on the y-axis, which extends to ±infinity at log ±3. The number of data points at ±infinity are denoted by the numbers to the left of the point. Circles are measurements in individual larvae (N = 14 per treatment), the dotted line denotes a 1:1 relationship between inoculum and qPCR ratio measured.

kill the host, a model that simply tests what number of virion hits best describes the data was used (see Methods 2.5). Using the frequency of dual genotype infections, it was calculated that on average 16.6 virions (i.e. ODV) invaded the host. Given this relatively large number of founders, the chance that at intermediate genotype ratios one of the viruses goes to fixation by genetic drift is small.

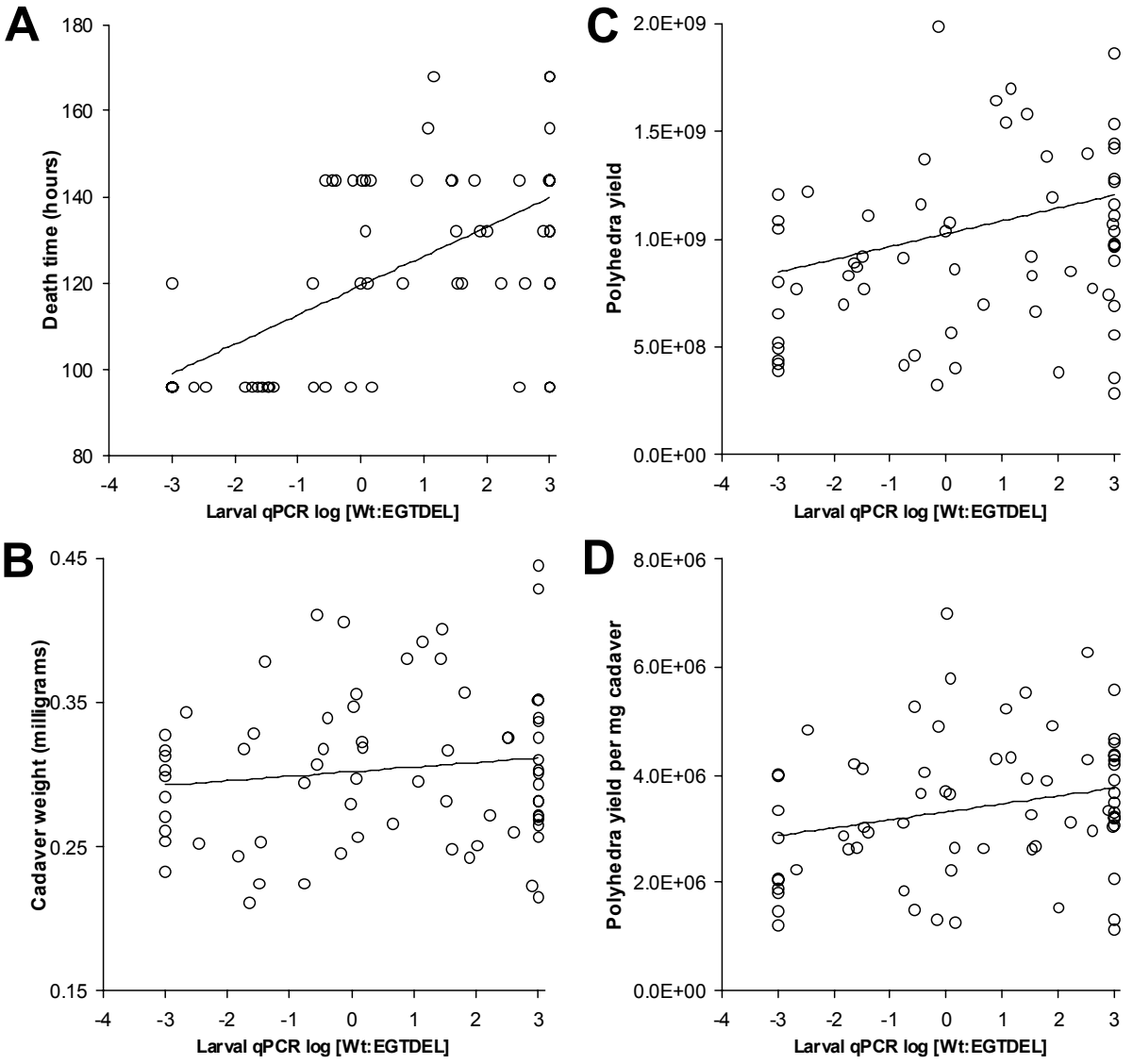


Figure 4: comparison of virus genotype ratio to other infection parameters. On the x-axis is the log virus ratio (Wt L1:EGTDEL), for all panels. On the y-axis is (A) time of death in hours (B) larval cadaver weight in grams (C) polyhedra yield and (D) polyhedra yield per unit cadaver weight. Note that although a linear regression line is plotted for convenience, the data were analyzed with Jonckheere-Terpstra non-parametric regression.

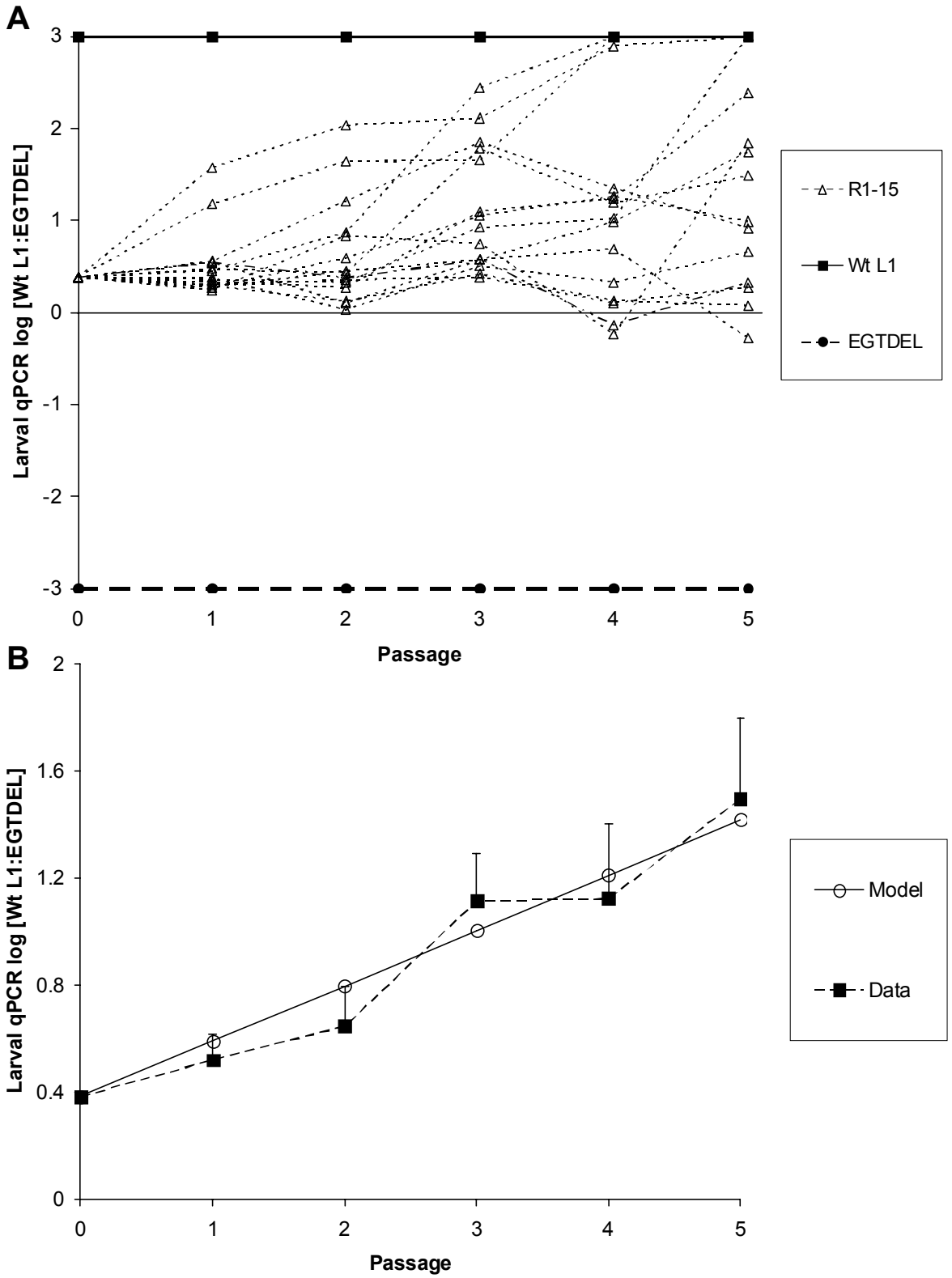


Figure 5: (A) qPCR data of the serial passage experiments in individual larvae. (B) Model fitting of the mean log ratio of individual larvae – with standard error denoted by bars - and the model estimate (see equation 2).

3.4 Estimating the within-host fitness of vEGTDEL from single passage experiments

In order to test if there is selection for the Wt L1 virus in the single passage experiment (Figure 3), a GLM was used (see Methods 2.6). The slope of the regression line obtained was not significantly different from 1 (0.8956 ± 0.1253). The y-intercept was significantly greater than 0 (0.4221 ± 0.2037; χ^2 = 4.29, d.f. = 1, P = 0.0383), which is evidence for selection for the Wt L1 genotype. Transforming the intercept from a log to a natural scale gives a selection rate constant (i.e. w-value, see equation 2) of $10^{0.4221}$ = 1.53.

3.5 Comparison of virus ratio to larval death time, weight and polyhedra yield

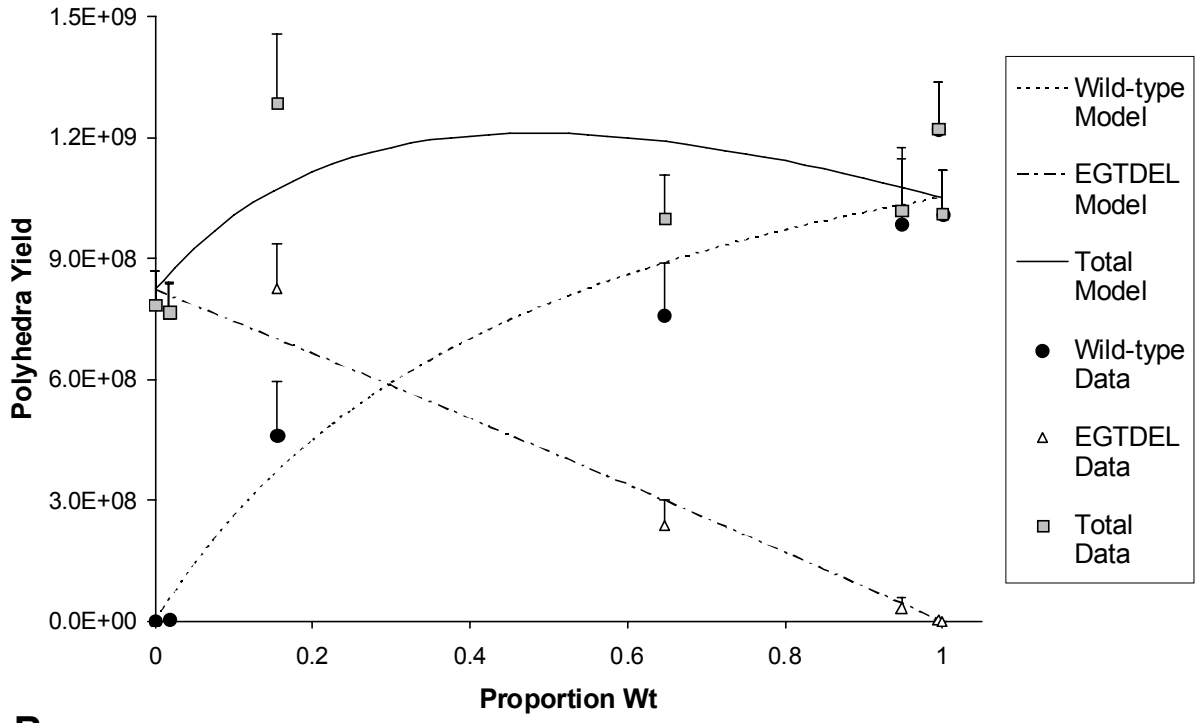
Data on the death time, cadaver weight and polyhedra yield were available for the individual larval cadavers from which DNA was analyzed by qPCR (Figure 4). A Jonckheere-Terpstra non-parametric regression was used to determine if qPCR ratio was related to death time, cadaver weight, yield, and yield per milligram cadaver weight. A significant positive relationship was found for both death time (Stand. JT = 5.243, N = 70, $P < 0.001$), yield (Stand. JT = 1.994, N = 70, P = 0.046) and yield per milligram cadaver weight (Stand. JT = 2.205, N = 70, P = 0.027), but not for cadaver weight (Stand. JT = 0.861, N = 70, P = 0.389).

3.6 Estimating the within-host fitness of vEGTDEL from serial passage experiments

The ratio of virus genotypes (R) was determined for each replicate after every passage, for the serial passaging from individual and pooled larvae (Figure 5A). In 4 out of 15 replicates, the Wt L1 genotype went to fixation, whereas vEGTDEL did not go to fixation in a single replicate. Moreover, the log virus ratio did not drop below ~ -0.3 (approx. 35% Wt virus) in any of the replicates. A simple testing procedure confirmed that there is selection for the Wt L1 genotype over multiple passages (Jonckheere-Terpstra regression; Stand. JT = 3.087, N = 75, P = 0.002.).

A selection rate constant (w) of 1.61 was calculated for the serial passage data using equation 2 (see Methods Section). This estimate of the selection rate constant is remarkably similar to the result of the single passage experiment (w = 1.53). Our data therefore provide very strong evidence for reduced within-host competitive fitness of vEGTDEL, with an approx. 1.6-fold increase of the Wt : vEGTDEL ratio per passage.

A



B

Virus	m_i experimental (x 10^8 polyhedra) mean ± se	m_i model (x 10^8 polyhedra) mean ± se	k_i ± se	r^2
Wt L1	10.52 ± 1.08	10.73 ± 0.69	2.994 ± 1.418	0.939
EGTDEL	7.85 ± 0.95	8.22 ± 0.51	1.052 ± 0.405	0.970

Figure 6: The combined effects of differences in yield and within-host fitness, fitted to the model proposed by de Wit (1960). (A) On the x-axis is the proportion Wt virus in the inoculum; on the y-axis the polyhedra yield. The individual data points are the mean yield for each virus and the total yield, with error bars denoting the standard error of the mean. The dotted lines are the fitted model (see equation 3) for yield of each virus. The solid line is total yield as predicted by the model, which is the sum of the yields for two viruses. (B) Model fitting results: independent experimental data and model predictions for single genotype infection yield (m_i experimental and model), model estimates for the crowding co-efficient k_i, and r^2-values obtained from the non-linear regression.

3.7 Integrating within-host fitness and yield

It has been previously reported that vEGTDEL generates a lower polyhedra yield (O'Reilly and Miller, 1989). In our experiments we found reduced within-host fitness of this virus. To understand the combined effect of these two observations, the more elaborate competition model of de Wit (1960) was used (see Methods 2.7). This model adequately described the data (Figure 6A), and model predictions for the yield in single genotype infection (m_i) were comparable to experimental data (Figure 6B). The *k*-values obtained confirm that the Wt virus is more competitive (k_{wt} = 2.994) than vEGTDEL (k_{egtdel} = 1.052). Since the *k*-values are not reciprocals the interaction is not neutral; some mixtures generate a greater total yield than if each virus fraction were to simply be as productive as in a single genotype infection. If the model is used to predict how the virus genotype ratio would be modulated by a passage in larvae, it becomes apparent that there is frequency-dependent selection (Figure 7). Selection for the wild-type virus is stronger when the frequency of the Wt virus is low (< 0.5). These results confirm and further detail those presented in sections 3.4 and 3.6.

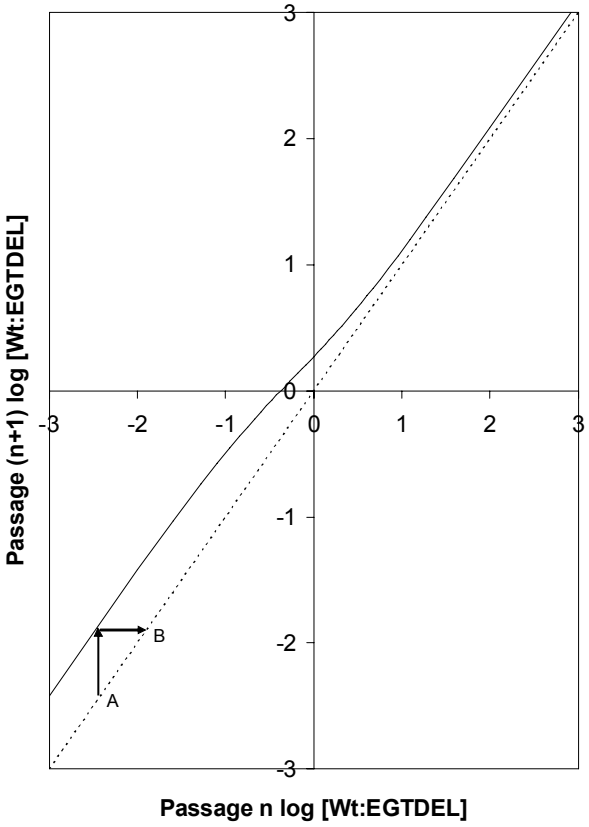


Figure 7: On the x-axis is the log genotype ratio (Wt L1:EGTDEL) at passage n and on the y-axis the log genotype ratio in the virus progeny after one passage of infection (n+1). The dotted line (1:1) indicates no fitness differences (1:1 ratio passage n:[n+1]). The solid black line is the prediction from the 'de Wit' competition model for the change in log genotype ratio over a single passage of infection. Note that this analysis takes into account differences due to within-host competition and total yield generated. As this line lies above the 1:1 line at all ratios, the Wt virus continually has a selective advantage. For example, if the genotype ratio is Point A, within one passage it will be displaced to Point B, as read along the x-axis. Note the frequency dependence: the lower the ratio Wt L1:EGTDEL, the stronger selection for the Wild-type virus will be.

4. Discussion

Although the concept of fitness is of central importance in evolutionary theory, it is generally difficult to define – let alone determine - fitness (e.g. Brommer *et al*., 2002). Direct competition of marked genotypes so as to quantify fitness – often times an evolved versus an ancestral strain – is one of the tools of experimental evolution (Elena and Lenski, 2003; Holland *et al*., 1991). Fitness can then be easily defined and quantitatively determined. Of course, fitness is inextricably linked to the particular set of environmental conditions in which it is measured (Elena and Lenski, 2003). The relevance of direct competition results to assessing fitness in the natural environment must therefore be given due consideration.

We have experimentally quantified the within-host competitive fitness of a fast-acting recombinant baculovirus by means of direct competition in single and serial passage experiments. This setup was tailored for selection based on differences in within-host competitive fitness. For single passage experiments, the frequency of the recombinant virus was determined by qPCR in a randomly selected larval cadaver. For serial passage experiments, a fixed number of polyhedra from a single randomly selected larval cadaver were passaged. This ensured that the probability of passage to the next generation was dependent on the frequency at which a genotype was present in individual larvae, and independent of, for example, the absolute yield and time of kill.

A qPCR-based assay for determining the Wt L1:EGTDEL ratio in larval cadavers was developed and validated (Figure 2). A previously reported qPCR-based assay performed better (Zwart *et al.*, 2008 [2]), but did not have the same constraints for primer design as in the case of Wt L1 and vEGTDEL; in the former case 'tag sequences' were deliberately inserted into the genomes of the viruses used in a well defined locus. Using the qPCR assay developed here, considerable variation between individual larvae in the genotype ratio was observed (Figure 3). This variation probably arises as a consequence of a genetic bottleneck at the level of invasion of the insect host, with a small number of virions achieving invasion of the host. Moreover, given that there is a small number of founders of infection, there will likely also be stochastic effects during virus amplification, as (1) there will be stochastic components to virus spread within the host (e.g. how long will it take before a BV infects a new cell?), and (2) BV production will vary between individual cells and cell types.

The independent action hypothesis (IAH) does not describe the infection process of AcMNPV in late instar larvae (L5), whereas it does in earlier larval instars (L3). In

late instar larvae, there are more dual genotype infections than predicted by IAH (Chapter 4). Our results here confirmed these findings for L5 larvae. As a tool for understanding the implications of our results, we estimated that 16.6 virions were invading host larvae. Accordingly, host larvae challenged with an intermediate ratio of genotypes (1:10 through 10:1) were predominantly co-infected, whereas challenges with more extreme ratios (1:100 and 100) led to mostly single genotype infections (Figure 3). For studying within-host competitiveness, we chose to work with final instar larvae. As a consequence, the number of co-infections at moderate doses was probably maximized, while allowing the use of observed mortality to determine if experiments with different larval cohorts are comparable.

Through single and serial passage experiments, we show that vEGTDEL has lower within-host fitness relative to the parent Wt L1 genotype. A selection rate constant of 1.53 was estimated based on a single passage experiment with different inoculum mixtures (Figure 3). From a serial passage experiment, a selection rate constant of 1.61 was estimated (Figure 5). It has been previously shown that deletion of the *egt* gene affects other fitness components (O'Reilly *et al*., 1989, Cory *et al*., 2004). These components – virus yield particularly - mainly concern fitness at the between-hosts level. Surprisingly, our data show that vEGTDEL also has reduced fitness at the within-host level. Given that vEGTDEL is capable of autonomous infection of a host insect, we had not anticipated such a marked fitness reduction at the within-host level. The fact that single and serial passage experiments lead to similar estimates of the selection rate constant suggests that co-occlusion of virions with different genotypes in a single polyhedron does not have major effects on competition within the host.

Using the selection rate constants we have estimated and equation 3 it can be estimated that it will take approximately 14.5 (w = 1.61) to 16.2 (w = 1.53) passages for the virus genotype ratio to travel from 1:1 (Wt L1:vEGTDEL) to 1000:1. The 1000:1 ratio is somewhat arbitrary, but given that mixed infections were rarely seen with the 100:1 mixture, at that point (i.e. 1000:1) genetic drift would almost entirely exclude vEGTDEL. Moreover, the probability of fixation becomes higher as a genotype ratio moves further from the 1:1 starting ratio. Since fixation is an irreversible process, genetic drift will decrease the mean time taken until vEGTDEL is purged from the population. A stochastic component is clearly very important in determining the genotype ratio in individual larvae. This means that there is a chance vEGTDEL can be maintained at high frequencies in the population for long periods of time in a small number of instances of 'individual replicates'. Even so, we did not observe fixation of vEGTDEL, or that a replicate dropped below 35% Wt L1 genotype in the serial passage experiment (Figure 5).

In order to put our results in a broader context, we also considered how the genotype ratio in larvae was related to host death time, cadaver weight, total virus yield and virus yield per unit cadaver weight (Figure 4). We observed that death time, polyhedra yield and yield per unit cadaver weight could be predicted from the qPCR ratio. Cory *et al.* (2004) previously reported differences in these parameters in single genotype infections, but how they would be affected by co-infection has not been given consideration. The relationship between genotype ratio and yield per unit cadaver weight is especially intriguing, since it suggests that *egt* primarily modulates how larvae develop and not simply how large they become. A significant relationship between virus genotype ratio and weight was not found, whereas this has been reported for single genotype infections (Cory *et al*., 2004).

Using the genotype ratio and yield data from the single passage experiment, the polyhedra yield per virus could be estimated (Figure 6). The simplest model that can be employed to describe these data is probably the competition model proposed by de Wit (1960), and we found that the model fitted the data well. The crowding co-efficient k_i was greater than 1 for the Wt L1 virus (k_{wt} = 2.994), while it was very close to 1 for vEGTDEL (k_{egtdel} = 1.052). This combination of *k*-values not only confirms within-host selection for the Wt L1 virus (Figures 6), it also provides evidence for frequency-dependent selection: the Wt L1 virus is most strongly selected when its frequency is low (Figure 7). In the serial passage experiments, the proportion of wild-type virus never fell very low (0.35) despite high variation between replicates, as would be expected if this frequency-dependence exists. The model also shows that the proportion wild-type vs. wild-type yield (and therefore also total yield) relationship is not linear. The hump-shaped total yield curve (Figure 6A) probably does not result from intermediate mixtures giving higher yields. A more parsimonious explanation would be that even with low frequencies of Wt L1 virus, systemic EGT levels are high enough to favorably influence host development and increase yield to wild-type virus only levels.

EGT's effect on speed of kill (O'Reilly *et al.,* 1989) makes it easy to propose an explanation for how deletion of this gene leads to reduced fitness with respect to yield: larvae die sooner and have less time to grow, and hence there are simply less host resources available for the virus to convert to transmission stages. Understanding why *egt* also influences fitness at the within host level clearly requires another explanation, as all genotypes present in an individual host will be equally affected by how host development progresses and the time of host death. O'Reilly *et al.* (1995) observed that wild-type and vEGTDEL virus infection proceeded in a markedly different manner when the viruses were engineered to express prothoracicotropic hormone

(PTTH): whereas wild-type infection proceeded normally, the vEGTDEL virus had a 100-fold reduction in pathogenicity. O'Reilly *et al*. (1995) suggested that ecdysteroids can adversely influence viral replication, since PTTH induces ecdysteroid production and vEGTDEL has lost the ability to inactivate ecdysteroids. Similarly, Keeley and Vinson (1975) reported that injection of ecdysone into larvae reduced an NPV's speed of kill and larval mortality. A plethora of mechanisms linking viral replication and ecdysteroids can be proposed, given the importance of the latter in regulating host development. The presence of EGT will in all probability have an effect *in trans* on replication. Consequently, the fitness of the wild-type virus will only be higher than that of vEGTDEL if (1) cells are predominantly infected by a single BV, such as early in the infection process (Zwart *et al*., 2008 [1], see also Bull *et al*., 2001) and (2) the positive effects of the EGT protein on replication are local and not systemic. Of course our results do not provide direct evidence for these assertions, but a coherent and parsimonious explanation of our observations demands them.

Polyhedron morphology appeared to be affected by *egt,* with vEGTDEL making smaller polyhedra with fewer ODV. As we have considered polyhedra purified from larvae, the cause for altered polyhedron morphology could be (1) systemic – since larvae die sooner in the absence of *egt*, in the majority of cells polyhedra will have less time for formation before cellular processes are disturbed by death – and/or (2) cellular – e.g. effects on replication result in altered polyhedron morphology. There are feasible experimental approaches for determining more specifically what the cause of altered morphology could be. For example, if the difference in morphology is determined solely by a cellular process, than vEGTDEL polyhedra generated in insect cell culture should also have altered morphology. The *egt* gene therefore appears to affect not only the timing of host death, but also host development, viral replication and polyhedron morphology. A better understanding of the mechanisms by which these effects of the *egt* gene are achieved may help elucidate how a single gene has evolved to influence these multiple and highly diverse phenotypic traits.

In terms of the consequences of releasing vEGTDEL as an insecticide, our data suggest: (1) when larvae are co-infected by the vEGTDEL and wild type AcMNPV viruses, vEGTDEL will be outcompeted at this level of selection i.e. within the host. This will be a relatively slow process; displacement of vEGTDEL will typically take multiple passages. Importantly, we find no evidence that there is any equilibrium frequency at which vEGTDEL is stably maintained; vEGTDEL is continuously outcompeted until it disappears from the population altogether (e.g. Figure 7). (2) We found evidence for frequency dependent selection for the Wt L1 virus (Figure 7), when including the between-hosts level (the production of transmission stages - absolute

yield). Consequently, when Wt L1 frequency is low, selection for it will be strongest, and hence the fixation of vEGTDEL is made more unlikely. This effect comes about in part as total yield is reduced only when the frequency of the Wt L1 virus is very low, whereas it is restored to wild type levels at low frequencies of the Wt L1 virus (0.1). (3) Given the prevalence of co-infections observed, vEGTDEL will not displace Wt AcMNPV viruses unless the starting frequency of vEGTDEL is very high (e.g. a ratio Wt L1:vEGTDEL < 1:100). (4) The competitive process in L5 will also be less stochastic than in earlier instar larvae at moderate levels of host mortality. Since there is independent action in earlier instar larvae of *T.ni* (Chapter 4), at comparable levels of host mortality (80%) genetic drift and stochastic processes during initial infection will play major roles in determining genotype frequencies in infected hosts. This suggests that at low levels of host mortality, vEGTDEL will more easily go to fixation in early instar larvae – perhaps leading to local 'reservoirs' of the vEGTDEL– while this is less likely to occur in L5 larvae.

Care must be taken in extrapolating conclusions based on our data to other recombinant viruses. It has been reported that the engineered expression of a heterologous marker (β-gal) protein leads to reduced fitness (Huang *et al.*, 1991). Similarly, it has been reported the presence of a toxin (venom of *Euplectrus comstockii*) can adversely effect baculovirus replication *in vivo* (Coudron *et al.*, 1995). On the other hand, the presence of a toxin alone (AaIT) has also been reported not to affect baculovirus competitive fitness *in vivo* (Milks *et al.*, 2001). The different conclusions we reach in this study may be attributed to (1) AaIT not having an effect on viral fitness, as the authors suggest, (2) differences in experimental setup: the competitive process was studied by considering when a passaged replicate went to fixation. Using this methodology alone, our evidence for selection of the wild type virus would have been weak.

What the combined effect of *egt* deletion and toxin expression would be on competitive fitness remains to be seen. Given the likely mechanisms bringing about effects on fitness, interactions which lead to restored fitness do not seem probable. Therefore, those viruses which have both *egt* deleted and expressing a toxin expression are likely to see either (1) a similar reduction in within-host fitness as the vEGTDEL or (2) further reductions in competitive fitness. Thus, from a biological safety perspective, it may be prudent to use recombinants with *egt* deleted – next to further modification enhancing biological control efficacy – simply because of the effects on fitness.

Acknowledgements:

The authors thank Gerrit Gort for GLM analysis of single passage qPCR data, Hanke Bloksma for expert assistance with EM work, Marieke Zwart and Liljana Georgievska for assistance with determining weight and yield, and Marcel Westenberg for helpful advice. JSC would like to thank David O'Reilly for providing the L1 and vEGTDEL viruses and NSERC for funding. MZ was supported by a strategic Ph.D. grant from the C.T. de Wit Graduate School for Production Ecology and Resource Conservation.

Chapter 7

General Discussion

1. Synopsis

A Ph.D. project should be a journey into uncharted territory. It is perhaps ironic that the most tangible output of this inherently myopic peregrination - the actual thesis - is expected to be a harmonious whole with a clear focus. Yet many theses do reach this state – be it through sheer brilliance, discoveries that are fortuitously complementary, or tortuous prose. Given the history of the present project – the motivation for doing experiments and the vision of the whole developed over time – it must be said that contingency has surely played a role, whatever the level of coherence one ascribes to this thesis.

Notwithstanding these considerations, there is a clear theme that binds together the work here described without any convoluted composition: how many viruses initiate a successful infection of a host and what does this mean for competitive processes among viruses within a host? In this thesis, the goal was to address two issues: (1) virus-host and virus-virus interactions that lead to invasion and disease of the host, specifically the number of virus individuals invading a host insect and causing disease (2) the interactions that can occur between virus genotypes during the process of diseasing the host.

The first issue has not been addressed and resolved in any virus – multi-cellular host pathosystem, and here we used baculovirus-insect larvae pathosystems to experimentally investigate it. Specifically, this venture has shown that: (1) A simple, existing model can be used to predict the number of virus individuals initiating an infection in baculovirus-insect host pathosystems characterized by high host susceptibility (Chapters 2 and 4). This model is built on the assumption that virions act independently in invading the host and in inducing pathogenesis and death. (2) However, this model does not apply for late instar and semi-permissive hosts. Therefore, we tested whether extended models satisfactorily describe the infection process in more resistant hosts, and found that a model which assumes differences in host susceptibility is most satisfactory (Chapter 5).

We have also shown that the number of virions getting into the host matters with respect to the interactions occurring between virus genotypes in the host: (1) Virus genotypes that do not support autonomous replication are purged because the small number of virions initiating infection results in low levels of co-infection at the cellular level during initial infection (Chapter 3). (2) A fast acting recombinant baculovirus, near isogenic to a wild type baculovirus, was found to have reduced within-host competitive fitness (Chapter 6). Interpreting the competitive process and

the mechanisms underlying this reduction in fitness both require an understanding of the number of virus individuals initiating the disease process, as investigated in chapters 2 and 4. In summary: one important component of the population genetics of baculoviruses has been quantitatively described and the relevance of this knowledge to understanding two phenomena has been demonstrated.

2. Placing this thesis in a broader context

The baculoviruses and insects constitute a model system highly suited for the work described here: (1) There are numerous potential applications, including biological control of insect pests, making the study of these viruses relevant from a societal perspective. (2) Baculoviruses have been well studied molecularly, and this allows for the use of numerous tools developed and the use of a great wealth of knowledge on the molecular biology and ecology of these viruses. (3) Their biology makes them suited to large scale laboratory and field experiments with minimal resources (e.g. the possibility of quantifying occlusion bodies by light microscopy, the small and easily mass-cultured host species) and regulatory restrictions (safety). Some of the implications of the work described here also extend beyond baculoviruses and their potential applications, as it pertains to our view of virus infections and disease more generally.

Chapter 3 puts forward a case for a general principle: virus populations will be under different selection pressures during the course of infection due to changes in the multiplicity of cellular infection (MOI). One of the ways in which this selection pressure can manifest itself is by affecting the frequencies of genes or genotypes in a population. A virus may have the ability to lose or gain genomic regions through mutation, insertion, deletion and recombination. The phenotypes of these variants can be widely differing. For example, a virus missing large genomic regions may be able to replicate faster than helper (full-length genome) viruses, but only when it is co-infecting cells with a helper virus. Conversely, full-length viruses can replicate and be transmitted independently (i.e. without the need of the gene products of a helper virus). When MOIs are low (e.g. after initial infection of the host but before most viral amplification has occurred), there will be strong selection for the ability to replicate autonomously, and thus for the presence in the viral genome of genes required for replication. This selection comes about as the low MOI results in low levels of cellular co-infection. When MOIs are high (i.e. late in infection), there will be weaker selection for the ability to replicate autonomously, as there is a great deal more co-

infection at the cellular level (see Mukawa and Goto, 2006, for *in vivo* growth curves). Traits linked to transmission (i.e. the production of polyhedra) are only moderately selected for late in infection, as shown by Bull *et al.* (2001, 2003).

This principle - derived from experimental work described in this thesis - may be widely applicable in virology, for two reasons: (1) Deletion mutants unable to replicate autonomously - and even defective interfering genotypes - are found in populations of most viruses (Huang, 1973). (2) Viral disease can be initiated by a small number of individuals (see Chapter 4 for discussion of this point), and the numbers of viruses typically increase as infection progresses. On the other hand, there are at least two limits to its applicability: (1) It will depend on whether there is a dichotomy between genes involved in replication and transmission in a particular viral system. Given the life histories of baculoviruses (e.g. different phenotypes for within host and between hosts spread of the virus) and their large genomes coding numerous genes, this dichotomy could very well be the greatest in baculoviruses. (2) Another limit to the broad applicability of this principle is at which stage of an infection a virus primarily achieves transmission: baculoviruses do this at very end of infection, whilst most other viruses may do this much earlier in infection (influenza, polio, foot-and-mouth disease, etc.). If transmission is achieved early in infection when MOIs are low, there will always be strong selection for the presence of genes involved in replication and transmission. Those viruses that transmit late in the infection process or post mortem (e.g. baculoviruses) can perhaps tolerate defective or parasitic viruses to a greater extent.

In Chapter 4 it was demonstrated for the first time that the independent action hypothesis (IAH; Druett, 1952) derived model of infection describes the baculovirus infection process in susceptible hosts. There are a number of important implications which stem from this work. First, it suggests that conspecific virus individuals can act independently of one another: they neither interfere with each other nor complement each other. One reason this observation is important is that it counterbalances recent work in RNA viruses, which stresses the complimentary nature of different virus genotypes (e.g. Vignuzzi *et al*., 2006). The two concepts are not at all mutually exclusive though: one concerns interactions between individuals of the same genotypes (or, in experimental practice, highly similar genotypes) in the processes of invasion and diseasing of the host, whereas the second concerns itself with the interactions between individuals of different genotypes in bringing about disease. Nonetheless, a balanced perspective for understanding viral fitness and pathogenesis may require incorporation of both concepts. It does remain to be seen whether IAH is broadly applicable to other viral pathosystems.

A second implication of the work presented in this thesis is that infections of a host initiated by the invasion of a single virus can routinely occur, provided hosts are exposed to doses that result in moderate or low mortality. While there have been previous reports that the number of individuals initiating an infection can be small (e.g. Sacristan *et al.*, 2004; Ali *et al.*, 2006), our methodology places this observation in a quantitative framework. Although this is a significant issue, one should be weary of too wide extrapolations. For example, the theory as applied here assumes a highly homogenous host population which is challenged with an identical dose per individual. Often, nothing could be further from the truth in a typical field situation, as there is, for example, heterogeneity of pathogen density in space. An extreme example to clarify this issue: in a field situation, half of the hosts sampled may be dead due to a virus infection. However, half of the hosts were exposed to large numbers of viruses, while the other half was not exposed at all. It can be appreciated that the proposed relationship between host survival and the number of invading virions is absolutely moot in this situation. Thus the limitations of the IAH model must be given due consideration. Nevertheless, the implications of the possibility that single virions can initiate infection are still important. This growing body of work provides a challenge, particularly to the modelling community: given that infections initiated by a single virion can regularly occur, how will this affect virus evolution at the level of the entire virus population?

Chapter 5 puts forward a model specifically for understanding the infection process in more resistant final instar (L5) larvae. The model used in Chapter 4 (IAH) is a special case of the more general model presented in Chapter 5, which allows for differences in host susceptibility. One is inclined to believe that the model for the infection process in L5 may be the more widely applicable than the model presented in Chapter 4. As host defences become more complex and effective, it may be there is a greater propensity for the assumption of a constant invasion probability to fail in describing the infection process. Given the expectation that in many pathosystems IAH is likely to fail, there is a need for exactly these sorts of extended models.

3. Will a better understanding of the population genetics of the baculoviruses make the world a better place?

While the previous section was concerned exclusively with the scientific context of this thesis, the original intent of the project was – in part – to explore the possibilities for exploiting natural baculovirus diversity for the purposes of biological control. The

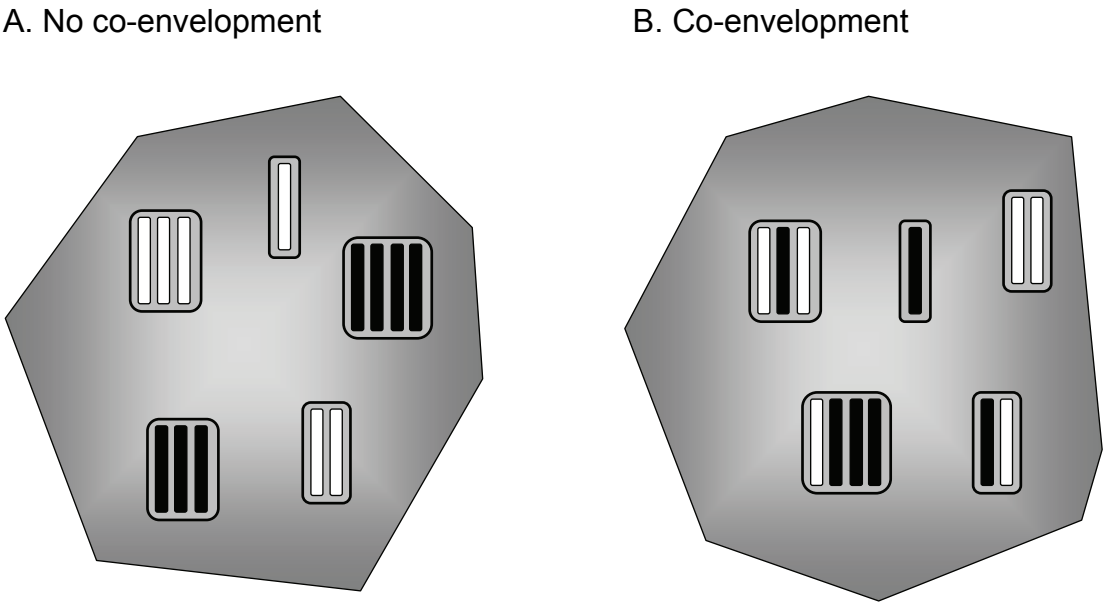


Figure 1: Does co-envelopment occur? It is known that multiple baculovirus genotypes can be contained in a single polyhedron. On the other hand, whether multiple genotypes can be contained in an occlusion derived virus (ODV) is not known. Panel A illustrates a polyhedron in which ODV are not co-enveloped, with black and white nucleocapsids representing different genotypes. Panel B illustrates a polyhedron in which co-envelopment does occur.

question of the immediate societal utility of a thesis is legitimate but not requisite. There are at least four areas in which the work described in this thesis is perhaps of some utility.

One of the questions that originally drove this research was how baculovirus speed of kill could be modulated by experimental evolution to make baculoviruses more competitive with chemical insecticides. One of the key issues unknown at the time - and addressed by our subsequent work - was the number of virions invading the host. Given the knowledge acquired, how would one design a setup to select for a faster or slower killing baculovirus? For a setup to select such viruses from standing genetic variation (i.e. starting with a natural virus isolate in which numerous genotypes – presumably with differing speeds of kill – were present), a low dose would be beneficial (e.g. $LD_{50\text{-}80}$ for *Spodoptera exigua* L3), as the resulting small virus effective population size would generate heterogeneity between infected hosts and allow for selection of the desired genotypes based the course of the infection. Pooling a small number of insect hosts with the desired infection phenotype (e.g. dying

quickly), instead of passaging virus progeny from a single host, could then prevent disrupting the selection process because of variation between hosts (e.g. a host dies fast not because of the genotype that has invaded it, but because it was a highly susceptible individual). This sort of setup was reported to have worked to a limited extent for Cooper *et al.* (2002), where they pooled five larvae infected with very low doses. It must be noted that the authors report that each of their 20 lines (10 fast and 10 slow) had a unique RFLP pattern. This is surprising in light of the above: one would expect the same fast or slow acting genotype to be present in at least a number of replicates of the same selective condition. On the other hand, it is not irreconcilable as the high recombination rate of the virus combined with 10 serial passages in insect hosts may confound the RFLP analysis. This is especially relevant since Cooper *et al.* (2002) were obliged to use larvae obtained from egg masses collected outdoors, and these were probably contaminated to some extent by latent baculovirus infections.

A second contribution of this thesis follows from our understanding that the selection pressure against baculovirus deletion mutants defective in replication is intense (Chapter 3). Within two passages *in vivo* the frequency of the *ie-1* gene - a gene necessary for replication that is lost during serial passage in insect cells - was restored. Even more dramatic is the rescue of polyhedron morphology: within a single passage the polyhedra formed were highly similar to those of a wild type virus. It must be stated that the rapid pace of these changes came as a surprise. What utility is there to be found in these observations? They suggest that if one is working with a virus with a relatively stable genome (e.g. *Autographa californica* MNPV as opposed to *Spodoptera exigua* MNPV; Pijlman *et al.,* 2002), a low number of passages in cultured insect cells followed by a low number of passages *in vivo* will result in a virus population which will not have major imbalances in the frequency at which genes are present. Of course, this process may still affect other genomic characteristics (e.g. Few Polyhedra mutants may be present). Nevertheless, this means that for this system cell culture does not have to be studiously avoided, be it to generate viral populations for experiments or for practical purposes (i.e. biological control). Ironically, one of the technologies we spent a good deal of time developing is the *in vivo* transfection of insect larvae with bacmid DNA (Chapter 2). Looking back, this time-consuming and difficult step was not absolutely necessary. Our data suggest that deletion mutants that would have arisen during transfection of insect cells would have been quickly lost once the virus was introduced to insect larvae. On the other hand, the route we took – transfection of insect larvae - is in principle a more elegant approach.

An important application of this thesis content is in that it may contribute to the assessment of the ecological impact of the release of fast acting recombinant

baculoviruses. Chapter 6 describes a series of experiments which are particularly relevant to this problem, and draws heavily from the concepts derived in Chapter 4. As the discussion of these issues is already rather long, the issue will only be touched upon here. Chapter 6 concludes that a fast acting recombinant baculovirus (vEGTDEL) has reduced within host competitive fitness, but that it will be only slowly purged from the population if only the within host level is considered. The implications of this result must be considered within the context of previous results (Cory *et al.*, 1994; see Chapter 6 for further discussion), which suggest that other fitness component of fast acting baculoviruses are also reduced. One could argue that this conclusion only complicates discussions about whether it should be permitted to use these viruses. The decreased fitness of these fast acting viruses – also observed in our study - could of course be an argument in favour of their use. On the other hand, the intensity of the selection we observed (Chapter 6), combined with the apparent stochasticity in the invasion and competition processes, could be used to argue against the use of recombinant viruses: if they are used, they will be present in populations for long periods of time. This discussion thus brings us to what is ultimately a matter of values: are we satisfied if we have good evidence that our activities will not lead to a major disruption and new steady states? Or must we only tolerate 'small' perturbations of natural systems? (So far as we can speak in terms of 'small' as industrialized societies!) There are no easy answers to these questions, and these will never be free of value judgements. Perhaps it is enough to simply state that science cannot exist in a vacuum and at the same time it should be independent. An accompanying normative plea for or against a particular action on the part of a scientist is acceptable and perhaps even desirable, so long as it is clearly differentiated from the reporting of actual scientific results.

The work described in Chapters 4 and 5 offers a new perspective on (1) the processes of pathogen invasion of a host, (2) the subsequent occurrence of disease, and (3) pathogen population genetics. If one is to take a critical perspective on this work, it could be stated that it is performed in a single model system and nothing is proven as such: we only offer a plausible explanation of how it might all work in a virus-insect pathosystem. On the other hand, if the ideas are 'big' enough, the system used to test them becomes irrelevant, and it is the demonstration of an important principle (i.e. independent action in pathogen invasion and pathogenesis) which is central. Therefore, this work may prompt those working in other fields - particularly the dynamics of human infectious diseases – to at least consider the implications of these ideas, and ultimately formulate, test and perhaps even exploit suitable models of infection.

4. What next?

Given the ample possibilities for questions to be addressed in scientific research, it can be a difficult task to formulate a focussed outlook. This problem is perhaps greatest for research driven by fundamental questions: interesting issues are virtually infinite! To limit the scope of this outlook, this final section will focus on what needs to be done in order to have a basic framework for understanding the role that genetic drift plays during the baculovirus infection process *in vivo*. First, the model for infection of L5 larvae (Chapter 5) must be tested. Although the model leads to an improved description of the infection process in L5, both in terms of dose-response and dual genotype invasions, it is still merely hypothetical. Some proposals for this are suggested in the discussion section in Chapter 5.

Second, MNPVs have ODV (occlusion derived virus) with multiple nucleocapsids. It is not known whether or not a single ODV can contain nucleocapsids of multiple genotypes (which will be referred to as 'co-envelopment'), as discussed in the introduction (See Figure 1). Conversely, it has been shown that co-occlusion occurs: ODV of different genotypes can be present in the same polyhedron (Hamblin *et al.*, 1990; Bull *et al.*, 2001). Although some have made the assertion that co-envelopment occurs (e.g. Smith and Crook, 1998; Simon *et al.*, 2006), we are not aware of any empirical evidence that supports this claim. In fact, rows of nucleocapsids are formed in the viral stroma; if some adjacent nucleocapsids are cleaved at random and then enveloped, this could also explain the observed distribution of nucleocapsids per ODV (Chapter 2). Note that for the work described in this thesis we could avoid this issue by using polyhedra derived from hosts infected with a single genotype; all virions can then only contain nucleocapsids of this genotype. However, there are experimental strategies for determining whether co-envelopment occurs, for example: (1) IAH adequately describes the infection process in *S. exigua* L3 larvae. Consequently, at very low doses (e.g. LD_{20}) infection will typically be initiated by a single virion. If larvae are then challenged with a low dose of co-occluded polyhedra containing two genotypes (as opposed to a mixture of single genotype polyhedra as in Chapter 4), there are two possibilities: (i) if ODV contain only one genotype (i.e. no co-envelopment occurs), the frequency of dual genotype infection is similar to that predicted by IAH for a mixture of single genotype polyhedra (i.e. most larvae have been invaded by only one genotype), or (ii) if most ODV contain nucleocapsids of both genotypes (i.e. co-envelopment occurs, and nucleocapsids of the two genotypes are randomly distributed in ODV), the frequency of dual genotype infection is much

higher than predicted by IAH for a mixture of single genotype polyhedra. The expected frequency for this scenario can be estimated – under the assumptions of IAH – since the distribution of nucleocapsids in ODV is known (see Chapter 2). (2) Baculoviruses with large genomic deletions generate shorter nucleocapsids (Kool *et al.*, 1991). Consequently, if co-occluded polyhedra containing a full-length virus and a deletion mutant are generated, then electron microscopy based tomography may reveal whether long (full-length) and short (deletion mutant) nucleocapsids are present in the same ODV. Conceptually this approach is simple and gives a definitive answer; in practice it will be technically demanding but may very well be feasible.

Third, during viral amplification there will be stochastic processes which affect genotype frequencies. In other words, if one virion of genotype A and three virions of genotype B invade a host, this ratio will not be exactly conserved during the amplification due to chance processes. This is illustrated by, for example, the analysis of single passage mixed infection experiments in Chapter 6: there was considerably more variation in genotype ratios than would be predicted if a binomial process at the level of invasion was the only source of variation. These changes in genotype frequency will be most significant when the number of invading virions is small. Since IAH describers the infection process in susceptible larvae, there can be a very small number of infection founders. Hence, these effects will probably have an important role in determining the genotype frequencies coming in a single larva. However, how much variation is generated in this process is unknown. This issue could be avoided in the work described here by considering the frequency of dual genotype infection (Chapters 4, 5 and 6), and not attempting to predict the distribution of genotype frequencies in dual genotype infected larvae. This is an interesting issue, however, and it is required for a full description of baculovirus population genetics. As there are many possible strategies to investigate this process (both theoretically – seen Renshaw *et al.*, 1991 - and empirically), only a simple suggestion will be made: BV of a mixture of two genotypes could be injected into larvae, and the genotype ratio in polyhedra obtained from dead larva quantified. Based on the frequency of dual genotype infection, an estimate of the number of infectious units actually received by the host can be made (this will probably be different from PFU in cell culture, for example). Subsequently, one can estimate the variation of the genotype ratio in dual genotype infected hosts.

On a philosophical note: I will personally not be surprised if the above questions remain unanswered for the time being: they are not extremely pressing from an applied point of view, and it is to be expected that others in the field of virus evolution will pursue the testing of their own hypotheses. Solely testing hypotheses is a rather dull endeavour with a somewhat negative inclination: 'simply' checking whether anything is wrong. It is a testament to the human spirit and ingenuity that this is often not the way we direct our undertakings – in science and in life.

Acknowledgements:

I would like to thank Just and Ellen Vlak, Wopke van der Werf, and Arjan de Visser for their comments.

References

Ali, A., Li, H. Y., Schneider, W. L., Sherman, D. J., Gray, S., Smith, D. & Roossinck, M. J. 2006 Analysis of genetic bottlenecks during horizontal transmission of Cucumber mosaic virus. *Journal of Virology* **80**, 8345-8350.

Arends, H. M., Winstanley, D. & Jehle, J. A. 2005 Virulence and competitiveness of *Cydia pomonella* granulovirus mutants: parameters that do not match. *Journal of General Virology* **86**, 2731-2738.

Ben-Ami, F., Regoes, R. R. & Ebert, D. 2008 A quantitative test of the relationship between parasite dose and infection probability across different host–parasite combinations. *Proc R Soc Lond B Biol Sci* **275**, 853-859.

Bianchi, F., Snoeijing, I., van der Werf, W., Mans, R. M. W., Smits, P. H. & Vlak, J. M. 2000 Biological activity of SeMNPV, AcMNPV, and three AcMNPV deletion mutants against *Spodoptera exigua* larvae (Lepidoptera : Noctuidae). *Journal of Invertebrate Pathology* **75**, 28-35.

Bianchi, F., Vlak, J. M., Rabbinge, R. & van der Werf, W. 2002 Biological control of beet armyworm, Spodoptera exigua, with baculoviruses in greenhouses: Development of a comprehensive process-based model. *Biological Control* **23**, 35-46.

Bonsall, M. B., O'Reilly, D. R., Cory, J. S. & Hails, R. S. 2005 Persistence and coexistence of engineered baculoviruses. *Theoretical Population Biology* **67**, 217-230.

Brommer, J. E., Merila, J. & Kokko, H. 2002 Reproductive timing and individual fitness. *Ecology Letters* **5**, 802-810.

Bull, J. C., Godfray, H. C. J. & O'Reilly, D. R. 2001 Persistence of an occlusion-negative recombinant nucleopolyhedrovirus in *Trichoplusia ni* indicates high multiplicity of cellular infection. *Applied and Environmental Microbiology* **67**, 5204-5209.

Bull, J. C., Godfray, H. C. J. & O'Reilly, D. R. 2003 A few-polyhedra mutant and wild-type nucleopolyhedrovirus remain as a stable polymorphism during serial coinfection in *Trichoplusia ni*. *Applied and Environmental Microbiology* **69**, 2052-2057.

Burden, J. P., Griffiths, C. M., Cory, J. S., Smith, P. & Sait, S. M. 2002 Vertical transmission of sublethal granulovirus infection in the Indian meal moth, *Plodia interpunctella. Molecular Ecology* **11**, 547-555.

Burden, J. P., Hails, R. S., Windass, J. D., Suner, M. M. & Cory, J. S. 2000 Infectivity, speed of kill, and productivity of a baculovirus expressing the itch mite toxin txp-1 in second and fourth instar larvae of Trichoplusia ni. *Journal of Invertebrate Pathology* **75**, 226-236.

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

Carrasco, P., Daros, J. A., Agudelo-Romero, P. & Elena, S. F. 2007 A real-time RT-PCR assay for quantifying the fitness of tobacco etch virus in competition experiments. *Journal of Virological Methods* **139**, 181-188.

Chao, L. 1990 Fitness of RNA virus decreased by Muller's Ratchet. *Nature* **348**, 454-455.

Christou, P., Capell, T., Kohli, A., Gatehouse, J. A. & Gatehouse, A. M. R. 2006 Recent developments and future prospects in insect pest control in transgenic crops. *Trends in Plant Science* **11**, 302-308.

Coleman, T.F. & Li Y. 1996 An interior trust region approach for nonlinear minimization subject to bounds. *SIAM Journal on Optimization* **6**, 418-445.

Coleman, T.F. & Li Y. 1994 On the convergence of reflective Newton methods for large-scale nonlinear minimization subject to bounds. *Mathematical Programming* **67**, 189-224.

Cooper, D., Cory, J. S., Theilmann, D. A. & 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.

Cooper, V. S., Reiskind, M. H., Miller, J. A., Shelton, K. A., Walther, B. A., Elkinton, J. S. & Ewald, P. W. 2002 Timing of transmission and the evolution of virulence of an insect virus. *Proc. R. Soc. B* **269**, 1161-1165.

Cory, J. S., Clarke, E. E., Brown, M. L., Hails, R. S. & O'Reilly, D. R. 2004 Microparasite manipulation of an insect: the influence of the egt gene on the interaction between a baculovirus and its lepidopteran host. *Functional Ecology* **18**, 443-450.

Cory, J. S., Green, B. M., Paul, R. K. & Hunter-Fujita, F. 2005 Genotypic and phenotypic diversity of a baculovirus population within an individual insect host. *Journal of Invertebrate Pathology* **89**, 101-111.

Cory, J. S., Hirst, M. L., Williams, T., Hails, R. S., Goulson, D., Green, B. M., Carty, T. M., Possee, R. D., Cayley, P. J. & Bishop, D. H. L. 1994 Field trial of a genetically improved baculovirus insecticide. *Nature* **370**, 138-140.

Cory J. S. & J. H. Myers 2003. The ecology and evolution of insect baculoviruses. *Annual Review of Ecology, Evolution and Systematics*, **34**, 239-272.

Coudron, T. A., W. C. Rice, M. R. Ellersieck, & R. E. Pinnell. 1995. Mediated pathogenecity of the baculovirus HzSNPV by the venom from *Euplectrus comstockii* (Hymenoptera, Eulophidae). *Biological Control* **5**, 92-98.

Cuevas, J. M., Moya, A. & Sanjuan, R. 2005 Following the very initial growth of biological RNA viral clones. *Journal of General Virology* **86**, 435-443.

Dai, X. J., Hajos, J. P., Joosten, N. N., van Oers, M. M., Ijkel, W. F. J., Zuidema, D., Pang, Y. & Vlak, J. M. 2000 Isolation of a *Spodoptera exigua* baculovirus recombinant with a 10 center dot 6 kbp genome deletion that retains biological activity. *Journal of General Virology* **81**, 2545-2554.

de Gooijer, C. D., Koken, R. H. M., van Lier, F. L. J., Kool, M., Vlak, J. M. & Tramper, J. 1992 A Structured Dynamic-Model for the Baculovirus Infection Process in Insect-Cell Reactor Configurations. *Biotechnology and Bioengineering* **40**, 537-548.

de la Iglesia, F. & Elena, S. F. 2007 Fitness declines in Tobacco etch virus upon serial bottleneck transfers. *Journal of Virology* **81**, 4941-4947.

De Roode, J. C., M. E. H. Helsinki, M. A. Anwar and A. F. Read. 2005. Dynamics of multiple infection and within-host competition in genetically diverse malaria infections. *Am. Nat.* **166**, 531-542.

De Roode, J. C., R. Culleton, A. S. Bell and A. F. Read. 2004. Competitive release of drug resistance following drug treatment of genetically diverse *Plasmodium chabaudi* infections. *Malaria Journal* **3**, 33.

de Wit, C.T. 1960. On Competition. *Versl. Landbouwk. Onderzoek*. **66**.8.

de Gooijer, C. D., Vanlier, F. L. J., Vandenend, E. J., Vlak, J. M. & Tramper, J. 1989 A Model for Baculovirus Production with Continuous Insect Cell-Cultures. *Applied Microbiology and Biotechnology* **30**, 497-501.

Domingo, E., Martinezsalas, E., Sobrino, F., Delatorre, J. C., Portela, A., Ortin, J., Lopezgalindez, C., Perezbrena, P., Villanueva, N., Najera, R., Vandepol, S., Steinhauer, D., Depolo, N. & Holland, J. 1985 The quasispecies (extremely heterogeneous) nature of viral RNA genome populations – biological relevance - a review. *Gene* **40**, 1-8.

Druett, H. A. 1952 Bacterial invasion. *Nature* **170**, 288-288.

Dushoff, J.& G. Dwyer. 2001. Evaluating the risks of engineered viruses: modelling pathogen competition. *Ecological Applications* **11**, 1602-1609.

Ebert, D. & Weisser, W. W. 1997 Optimal killing for obligate killers: the evolution of life histories and virulence of semelparous parasites. *Proc R Soc B* **264**, 985-91.

Elena, S. F. & Lenski, R. E. 2003 Evolution experiments with microorganisms: The dynamics and genetic bases of adaptation. *Nature Reviews Genetics* **4**, 457-469.

Engelhard, E. K., Kammorgan, L. N. W., Washburn, J. O. & Volkman, L. E. 1994 The insect tracheal system - a conduit for the systemic spread of *Autographa californica* M nuclear polyhedrosis virus. *PNAS* **91**, 3224-3227.

Engelhard, E. K. & Volkman, L. E. 1995 Developmental resistance in 4th-instar *Trichoplusia ni* orally inoculated with *Autographa californica* M Nuclear Polyhedrosis-Virus. *Virology* **209**, 384-389.

England, L. S., Holmes, S. B. & Trevors, J. T. 1998 Persistence of viruses and DNA in soil. *World Journal of Microbiology & Biotechnology* **14**, 163-169.

Erlandson, M., Newhouse, S., Moore, K., Janmaat, A., Myers, J. & Theilmann, D. 2007 Characterization of baculovirus isolates from *Trichoplusia ni* populations from vegetable greenhouses. *Biological Control* **41**, 256-263.

Federici, B. A. 1997 *Baculovirus Pathogenesis*. The Baculoviruses (L.K. Miller, Ed.). New York: Plenum Press.

Flipsen, J. T. M., Martens, J. W. M., Vanoers, M. M., Vlak, J. M. & Vanlent, J. W. M. 1995 Passage of *Autographa californica* nuclear polyhedrosis virus through the midgut epithelium of *Spodoptera exigua* larvae. *Virology* **208**, 328-335.

Funk C.J., Braunagel S. C., and Rohrman, G.F. 1997 Baculovirus Structure. *The Baculoviruses* (Ed. Miller. L.K.) New York: Plenum Press.

Gander, W. & Gautschi W. 2000. Adaptive quadrature – revisited. *BIT* **40**, 84-101.

Gandon, S., Jansen, V. A. & van Baalen, M. 2001 Host life history and the evolution of parasite virulence. *Evolution Int J Org Evolution* **55**, 1056-62.

Georgievska, L., Velders, R., Dai, X., Bianchi, F.J.J.A., Werf, W. van der, & Vlak, J.M. 2005. Competition in insect larvae between wild-type baculovirus (*Spodoptera exigua* nucleopolyhedrovirus) and a marked recombinant with enhanced speed of action. *IOBC/WPRS Bulletin* **28**, 141-145.

Giannini, C., Giannelli, F., Monti, M., Careccia, G., Marrocchi, M. E., Laffi, G., Gentilini, P. & Zignego, A. L. 1999 Prevalence of mixed infection by different hepatitis C virus genotypes in patients with hepatitis C virus-related chronic liver disease. *Journal of Laboratory and Clinical Medicine* **134**, 68-73.

Gibbs, M. J., Armstrong, J. S. & Gibbs, A. J. 2001 Recombination in the hemagglutinin gene of the 1918 "Spanish flu". *Science* **293**, 1842-1845.

Godfray, H. C. J., Oreilly, D. R. & Briggs, C. J. 1997 A model of nucleopolyhedrovirus (NPV) population genetics applied to co-occlusion and the spread of the few polyhedra (FP) phenotype. *Proc R Soc B* **264**, 315-322.

Goulson, D. & Cory, J. S. 1995 Responses of Mamestra-Brassicae (Lepidoptera, Noctuidae) to Crowding - Interactions with Disease Resistance, Color Phase and Growth. *Oecologia* **104**, 416-423.

Goulson, D. & Gory, J. S. 1995 Sublethal effects of baculovirus in the cabbage moth, *Mamestra brassicae*. *Biological Control* **5**, 361-367.

Grove, M. J. & Hoover, K. 2007 Intrastadial developmental resistance of third instar gypsy moths (*Lymantria dispar* L.) to L-dispar nucleopolyhedrovirus. *Biological Control* **40**, 355-361.

Hails R. S., P. Hernández-Crespo, S. M. Sait, C. A. Donnelly, B. M. Green and J. S. Cory. 2002. Transmission patterns of natural and recombinant baculoviruses. *Ecology* **83**, 906-916.

Hajos, J. P., Zuidema, D., Kulcsar, P., Heldens, J. G. M., Zavodszky, P. & Vlak, J. M. 1998 Recombination of baculovirus DNA following lipofection of insect larvae. *Archives of Virology* **143**, 2045-2050.

Hamblin, M., Vanbeek, N. A. M., Hughes, P. R. & Wood, H. A. 1990 Coocclusion and persistence of a baculovirus mutant lacking the polyhedrin gene. *Applied and Environmental Microbiology* **56**, 3057-3062.

Hawtin, R. E., Zarkowska, T., Arnold, K., Thomas, C. J., Gooday, G. W., King, L. A., Kuzio, J. A. & Possee, R. D. 1997 Liquefaction of Autographa californica nucleopolyhedrovirus-infected insects is dependent on the integrity of virus-encoded chitinase and cathepsin genes. *Virology* **238**, 243-253.

Heldens, J. G. M., van Strien, E. A., Feldmann, A. M., Kulcsar, P., Munoz, D., Leisy, D. J., Zuidema, D., Goldbach, R. W. & Vlak, J. M. 1996 *Spodoptera exigua* multicapsid nucleopolyhedrovirus deletion mutants generated in cell culture lack virulence in vivo. *Journal of General Virology* **77**, 3127-3134.

Hernandez-Crespo, P., Hails, R. S., Sait, S. M., Green, B. M., Carty, T. M. & Cory, J. S. 1999 Response of hosts of varying susceptibility to a recombinant baculovirus insecticide in the field. *Biological Control* **16**, 119-127.

Herniou, E. A., Luque, T., Chen, X., Vlak, J. M., Winstanley, D., Cory, J. S. & O'Reilly, D. 2001 Use of whole genome sequence data to infer baculovirus phylogeny. *Journal of Virology* **75**, 8117-8126.

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

Herniou, E. A., Olszewski, J. A., O'Reilly, D. R. & Cory, J. S. 2004 Ancient coevolution of baculoviruses and their insect hosts. *Journal of Virology* **78**, 3244-3251.

Hitchman, R. B., Siaterli, E. A., Nixon, C. P. & 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. & Strauss, M. 1995 Efficient gene transfer into human hepatocytes by baculovirus vectors. *PNAS* **92**, 10099-10103.

Holland, J. J., Delatorre, J. C., Clarke, D. K. & Duarte, E. 1991 Quantitation of relative fitness and great adaptability of clonal populations of RNA viruses. *Journal of Virology* **65**, 2960-2967.

Hoover, K., Grove, M. J. & Su, S. Z. 2002 Systemic component to intrastadial developmental resistance in Lymantria dispar to its baculovirus. *Biological Control* **25**, 92-98.

Hoover, K., Washburn, J. O. & Volkman, L. E. 2000 Midgut-based resistance of Heliothis virescens to baculovirus infection mediated by phytochemicals in cotton. *Journal of Insect Physiology* **46**, 999-1007.

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

Huang, A.S., 1973. Defective interfering viruses. *Annual Review of Microbiology* **27**, 101-117.

Huang, Y. S., Bobseine, K. L., Setzer, R. W. & Kawanishi, C. Y. 1991 Selection kinetics during serial cell culture passage of mixtures of wild type *Autographa californica* nuclear polyhedrosis virus and its recombinant AC360-Beta-gal. *Journal of General Virology* **72**, 2653-2660.

Hughes, D. S., Possee, R. D. & 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, P. R. & Wood, H. A. 1981 A synchronous peroral technique for the bioassay of insect viruses. *Journal of Invertebrate Pathology* **37**, 154-159.

Il'inykh, A. V. & Ul'yanova, E. G. 2005 Latency of baculoviruses. *Biology Bulletin* **32**, 496-502.

Inceoglu, A. B., Kamita, S. G. & Hammock, B. D. 2006 Genetically modified baculoviruses: A historical overview and future outlook. *Advances in Virus Research* **68**, 323-360.

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

Jurkovicova, M. 1979 Activation of latent virus infections in larvae of *Adoxophyes orana* (Lepidoptera, Tortricidae) and *Barathra brassicae* (Lepidoptera, Noctuidae) by foreign polyhedra. *Journal of Invertebrate Pathology* **34**, 213-223.

Keeley, L. L., and S. B. Vinson. 1975. Beta-ecdysone effects on development of nucleopolyhedrosis in *Heliothis* spp. *Journal of Invertebrate Pathology* **26**, 121-123.

King, L. A. & Possee, R. D. 1992 *The Baculovirus Expression System*. Cambridge, U.K.: University Press.

Knell, J. D. & Summers, M. D. 1981 Investigation of genetic heterogeneity in wild isolates of *Spodoptera frugiperda* nuclear polyhedrosis virus by restriction endonuclease analysis of plaque-purified variants. *Virology* **112**, 190-197.

Kool, M., Voncken, J. W., Vanlier, F. L. J., Tramper, J. & Vlak, J. M. 1991 Detection and analysis of *Autographa californica* nuclear polyhedrosis-virus mutants with defective interfering properties. *Virology* **183**, 739-746.

Kovacs, G. R., Choi, J., Guarino, L. A. & Summers, M. D. 1992 Functional dissection of the *Autographa californica* nuclear polyhedrosis virus *immediate early-1* transcriptional regulatory protein. *Journal of Virology* **66**, 7429-7437.

Lee, H. H. & Miller, L. K. 1978 . Isolation of genotypic variants of *Autographa californica* nuclear polyhedrosis virus. *Journal of Virology* **27**, 754-767.

Lenhoff, R. J., Luscombe, C. A. & Summers, J. 1998 Competition in vivo between a cytopathic variant and a wild-type duck hepatitis B virus. *Virology* **251**, 85-95.

Li, W. H. 1997 *Molecular Evolution*. Sunderland, Massachusetts: Sinnauer Associates.

Lo, H. R. & Chao, Y. C. 2004 Rapid titer determination of baculovirus by quantitative real-time polymerase chain reaction. *Biotechnology Progress* **20**, 354-360.

Lopez-Ferber, M., Simon, O., Williams, T. & Caballero, P. 2003 Defective or effective? Mutualistic interactions between virus genotypes. *Proc R Soc B* **270**, 2249-2255.

Luckow, V. A., Lee, S. C., Barry, G. F. & 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.

Mackay, I. M., Arden, K. E. & Nitsche, A. 2002 Real-time PCR in virology. *Nucleic Acids Research* **30**, 1292-1305.

Maeda, H., Fujimoto, C., Haruki, Y., Maeda, T., Kokeguchi, S., Petelin, M., Arai, H., Tanimoto, I., Nishimura, F. & Takashiba, S. 2003 Quantitative real-time PCR using TaqMan and SYBR Green for *Actinobacillus actinomycetemcomitans, Porphyromonas gingivalis, Prevotella intermedia*, tetQ gene and total bacteria. *Fems Immunology and Medical Microbiology* **39**, 81-86.

Marks, H., van Duijse, J. J. A., Zuidema, D., van Hulten, M. C. W. & Vlak, J. M. 2005 Fitness and virulence of an ancestral White Spot Syndrome Virus isolate from shrimp. *Virus Research* **110**, 9-20.

Meynell, G. G. & Stocker, B. A. D. 1957 Some hypotheses on the aetiology of fatal infections in partially resistant hosts and their application to mice challenged with *Salmonella paratyphi* B or *Salmonella typhimurium* by intraperitoneal injection. *Journal of General Microbiology* **16**, 38-58.

Milks, M. L., Leptich, M. K. & Theilmann, D. A. 2001 Recombinant and wild-type nucleopolyhedroviruses are equally fit in mixed infections. *Environmental Entomology* **30**, 972-981.

Miller, L. K. 1997 Baculovirus Pathogenesis. *The Baculoviruses* (L.K. Miller, Ed.) New York: Plenum Press.

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

Moya, A., Elena, S. F., Bracho, A., Miralles, R. & Barrio, E. 2000 The evolution of RNA viruses: A population genetics view. *PNAS* **97**, 6967-6973.

Moxon, E. R. & Murphy, P. A. 1978 *Hemophilus influenzae* bacteremia and meningitis resulting from survival of a single organism. *PNAS* **75**, 1534-1536.

Mukawa, S. & Goto, C. 2006 In vivo characterization of a group II nucleopolyhedrovirus isolated from Mamestra brassicae (Lepidoptera : Noctuidae) in Japan. *Journal of General Virology* **87**, 1491-1500.

Munoz, D., Castillejo, J. I. & Caballero, P. 1998 Naturally occurring deletion mutants are parasitic genotypes in a wild-type nucleopolyhedrovirus population of *Spodoptera exigua*. *Applied and Environmental Microbiology* **64**, 4372-4377.

Myers, J. H., Malakar, R. & Cory, J. S. 2000 Sublethal nucleopolyhedrovirus infection effects on female pupal weight, egg mass size, and vertical transmission in gypsy moth (Lepidoptera : Lymantriidae). *Environmental Entomology* **29**, 1268-1272.

Olkin, I., Gleser, L.J. and Derman, C. 1994 *Probability Models and Applications*. New York: Macmillan.

O'Reilly, D. R., Kelly, T. J., Masler, E. P., Thyagaraja, B. S., Robson, R. M., Shaw, T. C. & Miller, L. K. 1995 Overexpression of *Bombyx mori* prothoracicotropic hormone using baculovirus vectors. *Insect Biochemistry and Molecular Biology* **25**, 475-485.

O'Reilly, D. R. & Miller, L. K. 1989 A baculovirus blocks insect molting by producing ecdysteroid UDP-glucosyl transferase. *Science* **245**, 1110-1112.

O'Reilly, D. R. & Miller, L. K. 1991 Improvement of a baculovirus pesticide by deletion of the *EGT* gene. *Bio-Technology* **9**, 1086-1089.

Pas, S. D., Noppornpanth, S., van der Eijk, A. A., de Man, R. A. & Niesters, H. G. M. 2005 Quantification of the newly detected lamivudine resistant YSDD variants of Hepatitis B virus using molecular beacons. *Journal of Clinical Virology* **32**, 166-172.

Payne, C. C. 1986 Insect pathogenic viruses as pest control agents. *Fortschr. Zool.* **32**, 183-200.

Pijlman, G. P., Dortmans, J., Vermeesch, A. M. G., Yang, K., Martens, D. E., Goldbach, R. W. & Vlak, J. M. 2002 Pivotal role of the non-hr origin of DNA replication in the genesis of defective interfering baculoviruses. *Journal of Virology* **76**, 5605-5611.

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

Pijlman, G. P., van Schijndel, J. E. & Vlak, J. M. 2003 Spontaneous excision of BAC vector sequences from bacmid-derived baculovirus expression vectors upon passage in insect cells. *Journal of General Virology* **84**, 2669-2678.

Pineda, F. J. R., F. Moscardi, T. Luque, J. A. Olszewski and B. M. Ribiero. 2003. Inactivation of the *ecdysteroid UDP-transferase* (*egt*) gene of *Anticarsia gemmatalis* nucleopolyhedrovirus (AgMNPV) improves its virulence towards its insect host. *Biological Control* **27**, 336-344.

Plymale, R. C., Felton, G. W. & Hoover, K. 2007 Induction of systemic acquired resistance in cotton foliage does not adversely affect the performance of an entomopathogen. *Journal of Chemical Ecology* **33**, 1570-1581.

Ponchel, F., Toomes, C., Bransfield, K., Leong, F. T., Douglas, S. H., Field, S. L., Bell, S. M., Combaret, V., Puisieux, A., Mighell, A. J., Robinson, P. A., Inglehearn, C. F., Isaacs, J. D. & Markham, A. F. 2003 Real-time PCR based on SYBR-Green I fluorescence: An alternative to the TaqMan assay for a relative quantification of gene rearrangements, gene amplifications and micro gene deletions. *BMC Biotechnology* **3**.

Possee, R. D., Hirst, M., Jones, L. D., Bishop, D. H. L. & Cayley, P. J. 1993 *Field Tests of Genetically Engineered Baculoviruses*. Opportunities for Molecular Biology in Crop Protection. Farnham: The British Crop Protection Council.

Ravva, S. V. & Stanker, L. H. 2005 Real-time quantitative PCR detection of *Mycobacterium avium* subsp paratuberculosis and differentiation from other mycobacteria using SYBR Green and TaqMan assays. *Journal of Microbiological Methods* **63**, 305-317.

Reeson, A. F., Wilson, K., Gunn, A., Hails, R. S. & Goulson, D. 1998 Baculovirus resistance in the noctuid Spodoptera exempta is phenotypically plastic and responds to population density. *Proc. R. Soc. B* **265**, 1787-1791.

Renshaw, E. 1991 *Modelling Biological Populations in Space and Time.* Cambridge: Cambridge University Press.

Ridout, M. S., Fenlon, J. S. & Hughes, P. R. 1993 A Generalized One-Hit Model for Bioassays of Insect Viruses. *Biometrics* **49**, 1136-1141.

Sacristan, S., Malpica, J. M., Fraile, A. & Garcia-Arenal, F. 2003 Estimation of population bottlenecks during systemic movement of Tobacco mosaic virus in tobacco plants. *Journal of Virology* **77**, 9906-9911.

Sait, S. M., Begon, M. & Thompson, D. J. 1994 The effects of a sublethal baculovirus infection in the Indian meal moth, *Plodia interpunctella*. *Journal of Animal Ecology* **63**, 541-550.

Schmittgen, T. D., Zakrajsek, B. A., Mills, A. G., Gorn, V., Singer, M. J. & Reed, M. W. 2000 Quantitative reverse transcription-polymerase chain reaction to study mRNA decay: Comparison of endpoint and real-time methods. *Analytical Biochemistry* **285**, 194-204.

Scholte, E. J., K. Ng'habi, J. Kihonda, W. Takken, K. Paaijmans, S. Abdulla, G. F. Killeen et al. 2005. An entomopathogenic fungus for control of adult African malaria mosquitoes. *Science* **308**, 1641-1642.

Simon, O., Williams, T., Caballero, P. & Lopez-Ferber, M. 2006 Dynamics of deletion genotypes in an experimental insect virus population. *Proceedings of the Royal Society B-Biological Sciences* **273**, 783-790.

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

Smith, D. R., Adams, A. P., Kenney, J. L., Wang, E. & Weaver, S. C. 2008 Venezuelan equine encephalitis virus in the mosquito vector *Aedes taeniorhynchus*: infection initiated by a small number of susceptible epithelial cells and a population bottleneck. *Virology* **372**, 176-186.

Smith, G. E., Vlak, J. M. & Summers, M. D. 1983 Physical analysis of *Autographa californica* nuclear polyhedrosis virus transcripts for polyhedrin and 10,000-molecular-weight protein. *Journal of Virology* **45**, 215-225.

Smith, I. R. L. & Crook, N. E. 1988 In vivo isolation of baculovirus genotypes. *Virology* **166**, 240-244.

Smits, P. H., Vandevrie, M. & Vlak, J. M. 1986 . Oviposition of beet armyworm (Lepidoptera, Noctuidae) on greenhouse crops. *Environmental Entomology* **15**, 1189-1191.

Spieker, J. O., Yuill, T. M. & Burgess, E. C. 1996 Virulence of six strains of duck plague virus in eight waterfowl species. *Journal of Wildlife Diseases* **32**, 453-460.

Sun, X. L., Sun, X. C., Bai, B. K., van der Werf, W., Vlak, J. M. & Hu, Z. H. 2005 Production of polyhedral inclusion bodies from *Helicoverpa armigera* larvae infected with wild-type and recombinant HaSNPV. *Biocontrol Science and Technology* **15**, 353-366.

Sun, X. L., Sun, X. C., van der Werf, W., Vlak, J. M. & Hu, Z. H. 2004 Field inactivation of wild-type and genetically modified *Helicoverpa armigera* single nucleocapsid nucleopolyhedrovirus in cotton. *Biocontrol Science and Technology* **14**, 185-192.

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

Theilmann, D. A. B., G.W.; Bonning, B.; Jehle, J.; O'Reilly, D.R.; Rohrmann, G.F.; Thiem, S. & Vlak, J.M. 2005 Baculovirus Taxonomy in *Virus Taxonomy, Classification and Nomenclature of Viruses.* San Diego, U.S.: Elsevier Academic Press.

Thomas, E. D., Reichelderfer, C. F. & Heimpel, A. M. 1972 Accumulation and persistence of a nuclear polyhedrosis virus of the cabbage looper in the field. *Journal of Invertebrate Pathology* **20**, 157-164.

Turner, P. E. & Chao, L. 1999 Prisoner's dilemma in an RNA virus. *Nature* **398**, 441-443.

Vail, P. V., Sutter, G., Jay, D. L. & Gough, D. 1971 Reciprocal Infectivity of Nuclear Polyhedrosis Viruses of Cabbage Looper and Alfalfa Looper. *Journal of Invertebrate Pathology* **17**, 383-8.

van Lent, J., Wellink, J. & Goldbach, R. 1990 Evidence for the involvement of the 58k and 48k proteins in the intercellular movement of cowpea mosaic virus. *Journal of General Virology* **71**, 219-223.

van Lier, F. L. J., van den Hombergh, J. P. T. W., de Gooijer, C. D., den Boer, M. M., Vlak, J. M. & Tramper, J. 1996 Long-term semi-continuous production of recombinant baculovirus protein in a repeated (fed-)batch two-stage reactor system. *Enzyme and Microbial Technology* **18**, 460-466.

van Lier, F. L. J., van der Meijs, W. C. J., Grobben, N. G., Olie, R. A., Vlak, J. M. & Tramper, J. 1992 Continuous beta-galactosidase production with a recombinant baculovirus insect cell system in bioreactors. *Journal of Biotechnology* **22**, 291-298.

van Oers, M. M. 2006 Vaccines for viral and parasitic diseases produced with baculovirus vectors. In *Insect Viruses: Biotechnological Applications*, vol. 68, pp. 193.

Vaughn, J. L., Goodwin, R. H., Tompkins, G. J. & McCawley, P. 1977 Establishment of 2 cell lines from insect *Spodoptera frugiperda* (Lepidoptera-Noctuidae). *In Vitro-Journal of the Tissue Culture Association* **13**, 213-217.

Vignuzzi, M., Stone, J. K., Arnold, J. J., Cameron, C. E. & Andino, R. 2006 Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. *Nature* **439**, 344-348.

Von Magnus, P. 1954 Incomplete forms of influenza virus. *Advances in Virus Research* **2**, 59-79.

Washburn, J. O., Kirkpatrick, B. A., Haas-Stapleton, E. & Volkman, L. E. 1998 Evidence that the stilbene-derived optical brightener M2R enhances Autographa californica M nucleopolyhedrovirus infection of Trichoplusia ni and Heliothis virescens by preventing sloughing of infected midgut epithelial cells. *Biological Control* **11**, 58-69.

Washburn, J. O., Lyons, E. H., Haas-Stapleton, E. J. & Volkman, L. E. 1999 Multiple nucleocapsid packaging of *Autographa californica* nucleopolyhedrovirus accelerates the onset of systemic infection in *Trichoplusia ni. Journal of Virology* **73**, 411-416.

Washburn, J. O., Trudeau, D., Wong, J. F. & Volkman, L. E. 2003 . Early pathogenesis of *Autographa californica* multiple nucleopolyhedrovirus and *Helicoverpa zea* single nucleopolyhedro virus in *Heliothis virescens*: a comparison of the 'M' and 'S' strategies for establishing fatal infection. *Journal of General Virology* **84**, 343-351.

Wickham, T. J., Davis, T., Granados, R. R., Hammer, D. A., Shuler, M. L. & Wood, H. A. 1991 Baculovirus defective interfering particles are responsible for variations in recombinant protein-production as a function of multiplicity of infection. *Biotechnology Letters* **13**, 483-488.

Yin, J. L., Shackel, N. A., Zekry, A., McGuinness, P. H., Richards, C., Van der Putten, K., McCaughan, G. W., Eris, J. M. & Bishop, G. A. 2001 Real-time reverse transcriptase-polymerase chain reaction (RT-PCR) for measurement of cytokine and growth factor mRNA expression with fluorogenic probes or SYBR Green I. *Immunology and Cell Biology* **79**, 213-221.

Zhou, M. Z., X. L. Sun, X. C. Sun, J. M. Vlak, Z. H. Hu, & W. van der Werf. 2005. Horizontal and vertical transmission of wild-type and *recombinant Helicoverpa armigera* single-nucleocapsid nucleopolyhedrovirus. *Journal of Invertebrate Pathology* **89**, 165-175.

Zwart, M. P., Erro, E., van Oers, M. M., de Visser, J. A. G. M. & Vlak, J. M. 2008 Low multiplicity of infection in vivo results in purifying selection against baculovirus deletion mutants. *Journal of General Virology* **89**, 1220-1224.

Zwart, M. P., van Oers, M. M.., Cory, J. S., van Lent, J. W. M., van der Werf, W. & Vlak, J. M. 2008 Development of a quantitative real-time PCR to determine genotype frequencies for studies in baculovirus population biology. *Journal of Virological Methods* **148**, 146-154.

List of Abbreviations

AcMNPV	*Autographa californica* MNPV: the Alphabaculovirus type species, and the virus species used for all experiments described in this thesis.
Bacmid	A full-length, biologically active viral genome, genetically modified to contain an insertion with: (1) a bacterial replicon and selection marker, allowing for propagation as a plasmid in *Escherichia coli* and (2) transposition sites for insertion of an expression cassette (Chapter 1).
bp	Base pair: a complimentary pair of nucleotides
BV	Budded virus: a single, enveloped nucleocapsid that spreads the infection within the host.
EGTDEL	Recombinant AcMNPV missing the *egt* gene, derived from the Wt L1 genotype.
EM	Electron microscopy
IAH	Independent action hypothesis: formulated by Druett (Nature 170, 288-288) to understand dose-response relationships. The idea that (1) each pathogen individual has a probability of invading the host and causing disease and (2) pathogen individuals act independently of one another. See pages 43-44 for details.
h	hours
kbp	kilo bp
MOI	Multiplicity of infection: the mean number of virions (or infectious units) present per cell.
MNPV	Multiple nucleocapsid (multicapsid) NPV: Alphabaculoviruses with multiple nucleocapsids per ODV.
NPV	Nucleopolyhedrovirus: old genus name for Alphabaculoviruses
NTC	Non template control: a control condition for PCR without any

	template DNA added.
OB	The horizontal transmission stage of a baculovirus: a proteinacious body containing one (Betabaculovirus) or multiple (Alphabaculovirus) ODV.
ODV	Occlusion derived virus: the virions embedded in the occlusion body.
ORF	Open reading frame: a genomic region potentially coding for a protein.
PCR	Polymerase chain reaction: the *in vitro* amplification of a DNA template.
PDF	Probability density function
Polyhedra	The occlusion body of the Alphabaculoviruses. Singular form is polyhedron.
qPCR	Quantitative real-time PCR: a PCR reaction in which the initial number of template copies can be determined quantitatively by following the reaction dynamics in real-time.
SNPV	Single nucleocapsid NPV: Alphabaculoviruses with single nucleocapsids per ODV.
SSQ	Sum of squares
vPolh	A bacmid constructed here, with repaired expression of *polyhedrin* (Chapter 2).
vPolhA/B	vPolh bacmid, also carrying a type A or B 100 bp PCR recognition sequence.

Summary

This thesis explores the population genetics of the baculovirus infection process and the consequences for virus evolution. Using *Autographa californica* multicapsid nucleopolyhedrovirus (AcMNPV) and lepidopteran insect larvae as a model system, we attempt to characterize (1) elemental virus-host and virus-virus interactions pertinent to virus invasion and disease of the host, focussing largely on the number of virus individuals invading a host insect and including the first strong experimental evidence that a single virion can cause disease in a host animal, and (2) the interactions that can occur between virus genotypes during the process of diseasing the host, specifically competition.

The development of two technologies was necessary to address these questions. First, a method for generating clonal baculovirus populations containing molecular tag sequences was devised: insect larvae were transfected with full-length baculovirus genomic DNA modified to allow replication in *Escherichia coli* (bacmids). Second, a quantitative real-time PCR (qPCR) based assay was developed to measure the frequency of a baculovirus genotype in a mixed population.

In order to estimate the number of pathogen individuals invading the host and causing death, a basic probabilistic model that links this number to host survival was developed and experimentally tested. This model is based on the independent action hypothesis (IAH): the notion that (1) every pathogen individual has a fixed probability of invading the host, (2) if a pathogen individual invades the host this will irrevocably lead to host death, and (3) that pathogen individuals act independently. Here the model was specifically developed to predict how often hosts are invaded by two pathogen genotypes, when being challenged by a pathogen population consisting of two genotypes. Model predictions were tested using two near-identical bacmid-derived virus genotypes, larvae of three host species, and the qPCR assay. IAH model predictions were confirmed in early instar (L3) larvae of permissive host species (*Spodoptera exigua* and *Trichoplusia ni*). This strongly suggests that in these instances it is the action of one invading virion which leads to host death. Model predictions did not hold for late instar larvae (L5) of permissive host species and both early and late larvae of a semi-permissive species (*Mamestra brassicae*), which show greater resistance to AcMNPV. In these instances there was a too high frequency of invasions with both genotypes, suggesting that more virions had invaded these larvae to cause disease than predicted by IAH.

In one such instance where the IAH based model predictions were incorrect (*S. exigua* L5), the infection process was characterised in detail. Four probabilistic models for the infection process were postulated. The model that best fit the data - predicting both the shallow dose-response relationship and the frequency of dual genotype invasions – was a model which assumed variation in susceptibility between hosts. A surprising conclusion is that IAH can not be rejected in these instances; variation in host susceptibility cannot only lead to smoother dose-response relationship, but also a higher frequency of dual genotype invaded hosts.

The remaining work described in this thesis concerns competition between virus genotypes in host insect larvae. Serial passaging of a baculovirus in insect cells often leads to the rapid generation of viral variants with large genomic deletions. The fate of these deletion mutants when a virus population is reintroduced into insect larvae was investigated: those variants missing a gene necessary for independent replication were quickly purged from the viral population. To explain this observation, it was suggested that the number of virions initially invading the host was so small that co-infection of both complete and defective viruses at the cellular level rarely occurred. Hence viruses incapable of autonomous replication are lost from the progeny virus population.

Finally, competition in the host insect between a fast acting recombinant baculovirus - missing the *egt* gene (vEGTDEL) - and a wild type virus was studied. Single and serial passage experiments were performed and the frequency of the wild type and vEGTDEL determined by qPCR. It had been previously reported that fast acting viruses had reduced fitness with respect to (1) the viral progeny produced (i.e. the yield of occlusion bodies) and (2) transmission of the virus from insect to insect. Here it was found that the fitness of vEGTDEL virus was also reduced at the within host level. Moreover, selection against vEGTDEL appears to be frequency dependent: the higher the frequency of vEGTDEL, the stronger selection against it appears to be. The mechanism resulting in this selection against vEGTDEL was not elucidated.

The findings in this thesis show that quantitative models for ecological phenomena such as competition, facilitation, drift and founder effects are powerful concepts for explaining the interaction between a host insect and viruses. The resulting understanding and quantitative models are indispensible tools when predicting and studying the population genetics of viruses in man-made or natural ecosystems.

Samenvatting

In dit proefschrift wordt verslag gedaan van het onderzoek van de populatiegenetica van DNA virussen in relatie tot het infectieproces en de consequenties hiervan voor de evolutie van virussen in het algemeen en baculovirussen in het bijzonder. Als modelsysteem is gebruikgemaakt van *Autographa californica* multicapsid nucleopolyhedrovirus (AcMNPV), een insectenvirus, en van larven van Lepidoptera (vlinders en motten). In dit onderzoek is getracht de wisselwerkingen tussen virusdeeltjes onderling en tussen virusdeeltjes en de gastheer, in het tot stand brengen van ziekte te begrijpen. Hierbij is met name gekeken naar het aantal virusdeeltjes dat daadwerkelijk ziekte veroorzaakt en is voor het eerst in dierlijke systemen aangetoond dat onder bepaalde omstandigheden een enkel virusdeeltje ziekte kan veroorzaken. Daarnaast zijn ook de interacties tussen virusgenotypen in de gastheer, vooral competitie tussen virusgenotypen, bestudeerd.

De ontwikkeling van twee nieuwe technieken was nodig voor dit onderzoek. Allereerst is er een manier gevonden om genetisch identieke viruspopulaties met moleculaire merkers te creëren. Hiervoor zijn virusinfecties in insectenlarven op gang gebracht door het inspuiten van specifiek gemerkt, recombinant baculovirus DNA, gemaakt in bacteriën. Ten tweede is met behulp van kwantitatieve 'real-time' PCR (qPCR) een methode ontwikkeld voor het bepalen van het relatief voorkomen van virusgenotypen in gemengde virale populaties.

Om een schatting te maken van het aantal virusdeeltjes dat ziekte veroorzaakt, is een kansmodel opgesteld. Dit kansmodel volgt uit de 'independent action hypothesis' (IAH, ofwel onafhankelijke werkingshypothese). Deze hypothese stelt (1) dat ieder virusdeeltje een vaste kans heeft om de gastheer te doden - in onze benadering gaat het dan om het virion, (2) dat, wanneer een virusdeeltje in de gastheer is binnengedrongen, dit onherroepelijk leidt tot de dood van de gastheer, en (3) dat alle virusdeeltjes een onafhankelijke werkingswijze hebben. Het kansmodel voorspelt, aan de hand van het optreden van ziekte, hoe vaak een gastheer zal worden binnengedrongen door twee virusgenotypen, gegeven het feit dat de gastheer is blootgesteld aan een viruspopulatie dat twee genotypen bevat. De modelvoorspellingen zijn getoetst door middel van het gebruik van twee vrijwel identieke virusgenotypen, insectenlarven van drie soorten en de qPCR methode. Bij proeven met jonge (derde) stadia (L3) van vatbare gastheersoorten (*Spodoptera exigua, Trichoplusia ni*) kwamen de experimentele data overeen met de modelvoorspellingen. Dit wekt sterk de indruk dat de werking van één

virusdeeltje ziekte in de gastheer kan veroorzaken. De modelvoorspellingen klopten echter niet in oude (vijfde) stadia (L5) van vatbare gastheersoorten en in vroege en late stadia van een minder vatbare gastheersoort (*Mamestra brassicae*). In deze gevallen waren beide virusgenotypen in hogere mate aanwezig dan voorspeld door het model en dit duidt erop dat er dan meer virusdeeltjes de gastheer moeten binnendringen om ziekte te veroorzaken.

Voor een geval waarin het IAH-model geen goede voorspelling gaf (*S. exigua* L5), is het infectieproces nader gekarakteriseerd door middel van vervolgproeven. Hierbij is gekeken naar de dosisafhankelijkheid en naar de mate waarin beide virusgenotypen de gastheer waren binnengedrongen. Vervolgens zijn vier kansmodellen opgesteld die het proces mogelijk konden beschrijven. Het model, dat de experimentele data het beste omschreef, vooronderstelde alleen dat er variatie in vatbaarheid was in de gastheerpopulatie. Een verrassende uitkomst is dat de IAH-hypothese niet kan worden verworpen, maar dat variatie in vatbaarheid leidt tot een verhoging van de mate waarin beide virusgenotypen binnendringen.

Het volgende onderdeel in dit proefschrift betreft competitie tussen virusgenotypen in de gastheer. Seriële passage van een baculovirus in insectencellen leidt vaak tot het snelle ontstaan van virusvarianten met grote genomische deleties (m.a.w. het verliezen van erfelijk materiaal). Wat er gebeurt met deze virusvarianten als een viruspopulatie wordt geherintroduceerd in insectenlarven is onderzocht. Varianten die een gen misten dat essentieel is voor virale replicatie verdwenen binnen enkele passages uit de populatie. Deze waarneming wekt de indruk dat het aantal binnendringende virusdeeltjes zeer klein is. Met andere woorden, het lijkt erop dat vroeg in de infectie deze defecte virusvarianten vrijwel geen enkele cel weten te infecteren waar ook een compleet virusdeeltje aanwezig is. Omdat zij niet in staat zijn zelf te repliceren, verdwijnen deze varianten onder deze omstandigheden in een insect snel uit viruspopulaties.

Als laatste experimenteel onderdeel van dit proefschrift is gekeken naar de competitie in gastheren tussen een sneldodend genetisch gemodificeerd baculovirus (vEGTDEL) en een wild-type virus. Proeven met het (serieel) passeren van viruspopulaties zijn uitgevoerd en het voorkomen van de beide virussen is vastgesteld met behulp van de qPCR methode. Al eerder was door anderen vastgesteld dat deze gemodificeerde virussen een verminderde fitness hebben met betrekking tot (1) de hoeveelheid virus die vrijkomt bij infectie en (2) de overdracht van het virus van insect naar insect. Hierbij kwam naar voren dat bij gemengde infecties vEGTDEL door het competitief proces in de gastheer wordt verdrongen door het wild-type virus. De selectie tegen vEGTDEL in het voordeel van het wild-type lijkt ook frequentieafhankelijk te zijn:

hoe meer vEGTDEL aanwezig is, hoe sterker de selectie in het nadeel van het wild-type virus is. Het mechanisme van deze selectie is niet opgehelderd.

De resultaten, beschreven in dit proefschrift, laten zien dat kwantitatieve modellen van ecologische fenomenen, zoals competitie en genetische ‘drift’, krachtige concepten zijn voor het omschrijven van de wisselwerkingen tussen virusdeeltjes onderling en tussen virusdeeltjes en de insectgastheren. Het verbeterde begrip van deze processen en de kwantitatieve modellen die hiervoor zijn opgesteld om deze processen te beschrijven zijn uiterst bruikbaar voor het voorspellen en bestuderen van de populatiegenetica van virussen in kunstmatige of natuurlijke ecosystemen.

PE&RC PhD Education Certificate

With the educational activities listed below the PhD candidate has complied with the educational requirements set by the C.T. de Wit Graduate School for Production Ecology and Resource Conservation (PE&RC) which comprises of a minimum total of 32 ECTS (= 22 weeks of activities)

Review of Literature (5.6 ECTS)

- Using experimental evolution to explore baculovirus diversity: theoretical considerations and applications (2004/2005)

Laboratory Training and Working Visits (2.9 ECTS)

- Baculovirus/host population dynamics and field work; University of British Columbia (2006)
- Population genetics of baculoviruses; Oxford University (2007)

Post-Graduate Courses (2.8 ECTS)

- Course in virology; Mol. Med. Post Grad. School, Erasmus University, Rotterdam (2004)
- Scientific writing; Wageningen Graduate Schools (2007)

Deficiency, Refresh, Brush-up Courses (2.8 ECTS)

- Population ecology; Crop and Weed Ecology Group, Wageningen University (2006)

Competence Strengthening / Skills Courses (2.6 ECTS)

- Time planning and project management; Wageningen Graduate Schools (2005)
- PhD Competence assessment; Wageningen Graduate Schools (2006)
- Laboratory use of isotopes; AMD, WUR

Discussion Groups / Local Seminars and Other Scientific Meetings (6.8 ECTS)

- Tracks of Evolution / Experimental Evolution Discussion Group Meetings & Symposia (2004-2007)
- Dutch Annual Virology Symposium (2004-2007)
- Biorad Real-time PCR workshop Wageningen (2005)

PE&RC Annual Meetings, Seminars and the PE&RC Weekend (2.0 ECTS)

- Introduction Weekend (2004)
- PE&RC Days (2004-2007)

International Symposia, Workshops and Conferences (7.1 ECTS)

- Current Themes in Ecology (2004-2007)
- European Society for Evolution Biology (2005)
- Society for Invertebrate Pathology (2007)

Curriculum Vitae

Mark Zwart was born on the 14th of October, 1979, in Monteria, Colombia. He lived there for five years before moving to Ecuador. Here he attended the Nate Saint Memorial School in Shell-Mera, and the Alliance Academy in Quito. At the age of 14, Mark moved to the Netherlands with his family. Here he completed his secondary education (IGCSE and International Baccalaureate) at the Alberdingk Thijm College in Hilversum.

Mark continued his education at the University College Utrecht, where he completed a Bachelor of Science, focusing on the life sciences. His Bachelor thesis was in cognitive neuroscience and focused on categorical and coordinate processing of visual spatial information. Subsequently, he went on to due a pilot program for the new Life Sciences Master at Utrecht University. His first research internship was at the Hubrecht Institute, researching the role of a murine gene in gastrulation. He subsequently did an internship at the Dutch Tropical Institute, developing a sero-epidemiological test for malaria. For this work he spent some time in Chiang Mai, Thailand.

In January 2004, Mark began his Ph.D. thesis work at the Laboratories of Virology and Genetics of Wageningen University, under the supervision of Prof. Just Vlak. Although he set out to use experimental evolution to modulate the virulence of baculoviruses, the project soon evolved to address more fundamental questions concerning virus population genetics. Specifically, he worked on understanding the number of viruses initiating disease, and the ramifications thereof for virus evolution. As of June 2008, Mark has been at the Quantitative Veterinary Epidemiology Group of Wageningen University, working on adaptive pathogen dynamics under Prof. Mart de Jong.

Mark met Marieke Verboom when studying in Utrecht, and they were married in 2003. They have two daughters: Jade (2 years old) and Joy (1 month old). They lived in Utrecht until recently, and currently reside in Ederveen. Mark enjoys reading, travelling, and music.

The study presented in this thesis was carried out at the Laboratory of Virology, the Crop and Weed Ecology Group and the Laboratory of Genetics of Wageningen University, The Netherlands, and was financially supported by a strategic grant from Wageningen University Graduate School “Production Ecology and Resource Conservation”.