How to be good at being a virus
Berngruber, Thomas

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2008

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
Chapter 1

Introduction and aim of thesis

‘You can’t always get what you want,
... but if you try sometimes you might find, you get what you need’

(The Rolling Stones)

1.1 Biological organization, biological evolution and biological models

How much is the path of evolution guided by its history, i.e. the evolutionary constraints that result from the established organization of an organism? This question has puzzled evolutionary biologist for centuries and spawned a controversy that is as old as evolutionary biology itself (Gould & Lewontin 1979; Charlesworth et al. 1982; Maynard-Smith et al. 1985; Houston 1997; Wagner 1998; Pigliucci & Kaplan 2000; Pigliucci 2007). Until the present day, this controversy seems to be unresolved - maybe because this fundamental evolutionary question has no general answer. Evolutionary constraints seem to depend on the biological context, i.e. the organism in question, the level of organization considered as well as the evolutionary time frame one is looking at.

Nonetheless, the evolutionary consequences of biological organization cannot be ignored. After all, evolutionary contingency implies that new
organisms are descending from existing organisms. The current organization of a biological organism is therefore the raw material for mutational variation and future evolutionary change. The big question is therefore not whether biological organization matters, but how it can be incorporated into evolutionary models in a meaningful way. In its essence organization is the interaction between the parts of an organism. Yet, what are these parts and how can we define interactions between them?

Representing biological organization is difficult, incorporating it into evolutionary thinking even more so. Organisms consist of numerous parts that show complex interactions. Approaches to reduce the complexity have followed different avenues. Generic models reduce the details of interactions and instead use a statistical description subsuming interactions of a similar kind. Specific models reduce the number of interacting parts by focusing on a specific subset of the organisms, like a specific genetic pathway. Generic and specific models both have their benefits and drawbacks. Generic models are largely independent of a specific biological context and therefore applicable to a wide range of biological systems. On the downside, they can only produce generic conclusions and, therefore, have little predictive power for a specific biological context. In contrast, specific models can predict the path of evolution in quite some detail, but they only apply to a specific group of organisms in a limited evolutionary time frame. In this sense predicting the path of evolution is analogous to forecasting the weather – the longer one projects the current situation into the future, the more uncertainty is attached to the prediction.

1.2 Aim of this thesis

How much generality should we trade in to account for a specific biological context? In my opinion, evolutionary models should incorporate specific real world mechanisms in order to be applicable to real world problems. In viral systems these mechanisms are particularly well known. In this thesis I therefore chose to study the evolutionary
limitations of viruses by mathematical models that are based on known molecular genetic mechanisms. In my view, this integration of molecular genetics and evolutionary theory is the most promising avenue to apply evolutionary models to important problems like the evolution of viral drug resistance, changes in viral virulence or the switch between host organisms.

Molecular genetics has opened the possibility to introduce a meaningful representation of biological organization into evolutionary thinking. The well defined concepts of molecular genetics aid the development of evolutionary models that are stated in real world mechanisms, like protein expression or enzymatic rates, and therefore enable a direct experimental test of theoretical predictions. A specific mechanistic underpinning puts a boundary to the process of evolution, as it allows for an a priori definition of feasible strategies in terms of protein expression, enzymatic activity and binding properties. Given the range of possibilities one can then find the best strategy given the biochemical limitations. Models that are framed in terms of a specific mechanistic organization can therefore arrive at testable predictions on the limitations of viral adaptations or in other words explain why a virus cannot always get what it wants. Knowledge of such limitations is of obvious importance for designing approaches to existing problems of viral evolution. In particular, explicit consideration of mechanistic limitations of viral adaptation might challenge the pessimistic premise that viral adaptation cannot be countered in view of the high rate of evolution due to the enormous size of virus populations.

Molecular genetics can benefit from evolutionary thinking, since evolutionary theory necessitates the integration of biochemical aspects into a single overall measure of viral fitness. In other words, evolutionary thinking leads to a consideration of the interaction of biochemical processes throughout the entire life cycle. Tools of evolutionary modeling, in particular life history theory, can therefore provide an integrative approach for novel questions in molecular
genetics, resulting from the study of large scale gene expression profiling. Classical molecular genetics often focuses on cases in which the coupling of a certain biochemical rate has a one-on-one relationship with survival or replication. This understanding can be misleading when one biochemical rate affects multiple aspects of the viral life cycle in opposing directions (Krakauer & Komarova 2003). Even more, the replication rate itself does not need to be an accurate predictor of viral fitness when it has negative side effects on other aspects of the life cycle. For example, fast replicating pathogens that kill their host quickly will not have much opportunity to be transmitted to a new host organism. In such a case, fast replication has a negative side effect on transmission. Selection will therefore not maximize the replication rate. Instead, overall fitness is maximized for a certain intermediate balance between replication and transmission. Evolutionary modeling, and in particular the tools of life history theory are designed to integrate these opposing selective forces throughout the viral life cycle in order to arrive at an overall prediction.

The viral life cycle establishes an important link between viral biochemistry and viral fitness. Gene regulatory interactions of even the simplest organisms are highly complex. Accounting for the full complexity remains a forbidding task for most organisms. Analysis of the viral life cycle breaks down this complexity by focusing on a few components that are crucial for viral fitness. In a first step, a life history model identifies those components of the life cycle that are most relevant for viral fitness. In a second step, these components are then investigated taking account of viral biochemistry. This approach has the advantage that the complexity of the problem is reduced by singling out the biochemical aspects that are most relevant for viral fitness. It is the main aim of this thesis to demonstrate that this modeling approach is capable of identifying and integrating biochemical detail into models of viral evolution in a rather straightforward way. Moreover, these models can arrive at predictions that are stated in biochemical parameters that are accessible to experimental testing.
1.3 The bacteriophage life cycle

Bacteriophages are ideal model systems to investigate the evolutionary limits of viral evolution and their dependence on viral biochemistry. The biochemical key elements shaping the viral life cycle are known in considerable detail. Most globally, the viral life cycle falls into two main steps: replication inside the host cell and transmission between hosts cells. Essentially, viruses are genetic elements that acquired the ability to survive outside of the host cell and to transmit to new host cells by packaging their genome into a protein capsid or lipid vesicle. The ability to survive outside of the host cell is the single property that distinguishes viruses from other selfish genetic elements like transposons or plasmids (Shimotohno et al. 1980; Shimotohno & Temin 1980; Forterre 2006). Phage transmission between host cells depends on diffusion through the medium and adsorption to a new host cell by the aid of surface proteins. In liquid conditions, diffusion is fast and transmission between hosts is proportional to an adsorption rate and the density of host cells and phage particles (Schlesinger 1960). In a spatially structured agar medium viral diffusion can, however, become limiting for transmission to new host cells.

Phage transmission can occur horizontally, i.e. between unrelated hosts, or by vertical transmission from an infected parent cell to daughter cells. Horizontal and vertical transmission each require specific genetic mechanisms for the adaptation of the life cycle. Horizontal transmission requires packaging of the viral genome in a survival stage and mechanisms to leave the host cell. Packaging of the viral genome can either occur in a capsid of viral proteins, or by budding of viral vesicles through the host cell membrane. Assembly of viral particles in the host membrane provides a mechanism for the virus to leave the host cell in a continuous process without destroying it (e.g. Phage M13). In contrast, the release of viral particles which are not assembled in the cell membrane requires the destruction of the host cell through lysis mediated by specific lysis proteins (Young et al. 2000). Viruses that destroy the host cell in order to transmit are called lytic.
The decision to destroy the host cell though the expression of lysis genes is one of the most important aspects of the viral life cycle, since destruction of the host cell also destroys the machinery of phage replication. The time interval between infection of the host cell and its destruction is called lysis time. The number of phage offspring which is produced during that interval is called burst size. The reproductive output of lytic bacteriophages depends on the interplay of lysis time and burst size. During viral replication within the host cell the number of newly produced phage particles increases with time (Doermann 1952; Loeb & Zinder 1961). Prolonged lysis time therefore increases burst size, but at the same time increases generation time and delays viral spread and infection of new host cells. These two consequences of lysis timing have opposing effects on viral fitness. Bacteriophage lysis timing should therefore evolve to an intermediate optimum that balances its cost and benefits.

As an alternative to horizontal transmission, lysogenic bacteriophage can transmit from parent to offspring through vertical transmission. Parent-to-offspring transmission requires the survival and growth of host cells. In order to ‘keep its host alive’ lysogenic phages therefore actively repress the viral genes for replication and host lysis, keeping the host in a state called lysogenic infection. Additionally, lysogenic phages ensure their vertical transmission through integration of their genome into the host genome by specialized enzymes called integrases (Groth & Calos 2004). A lysogenic phage that is integrated in the host genome is referred to as pro-phage. Lysogenic phage can switch to a lytic cycle upon external triggers that indicate poor survival conditions for the host cell like UV damage, nutrient starvation or general stress response. This switch from the lysogenic to lytic conditions under poor host survival can be compared to a strategy of ‘leaving a sinking ship’.

Viral replication within the host cell determines the rate at which the resources of the host cell are exploited. Molecular mechanisms that regulate this stage of the life cycle have important implications for viral
virulence, e.g. for degree of damage a virus causes to host fecundity and longevity. Benign phages like phage M13 and phage λ, for example, which do not lyse the host cell, evolved a suite of mechanisms to control their own rate of replication within the host cell, in order to keep the host alive. Phage M13, for example limits its own replication by the production of large amounts of protein P5 that covers the single stranded form of M13 and prevents a conversion to the double stranded DNA replication stage (Baas 1985). Likewise, bacteriophage λ actively represses its replication genes by the virulence repressor cI in order to ensure host survival during vertical transmission (Oppenheim et al. 2005). These, self-repression mechanisms that regulate of viral genome replication demonstrate that viruses are able to control their own level of virulence and, hence, adapt it to the environmental conditions.

Replication within the host cell falls into two distinct stages that have specific requirements on the timing and balance of viral gene expression. During early infection primarily the viral replication proteins are expressed. Accordingly, there is no production of new viral particles. This period is called eclipse time. During late infection viral packaging and lysis proteins are produced, viral particles are assembled and the number of phage particles inside the host cell increases. This phase is called maturation time. The timing and balance of viral protein expression over the early and late infection cycle is essential for efficient production of viruses. The expression of a single gene can therefore have conflicting consequences for early and late viral infection. This can lead to opposing selection forces on the expression of a particular gene during early and late infection and limit the evolution of viral gene regulation.
1.4 The environment determines selection on the viral life cycle

Viruses are obligatorily dependent on their host. The evolutionary success of viruses is therefore closely linked to the ability to transmit between host organisms. The conditions for viral transmission, in turn, strongly depend on the environment, e.g. the availability of hosts, the structure of the host population and competition for these hosts with other viral strains. The viral environment determines the relative importance of the transmission stage in comparison to other viral life stages. Accordingly, a change in the environment alters this balance and creates selection on viral life-history characteristics like virulence and transmissibility.

An important example is the adaptation of bacteriophage lysis timing to the host density in the environment (Wang et al. 1996). Theory predicts that a change of host density changes the relative contribution of replication and transmission to overall viral fitness over the entire life cycle and should therefore select for altered lysis timing. This adaptation of lysis timing to the host density has indeed been demonstrated in experimental evolution with phage RB69 (Abedon et al. 2003).

Not only the density, but also the structure of the host population markedly influences the evolution of viral life history. In a spatially structured host population transmission to uninfected hosts is limited when infected individuals form spatial clusters (Boots et al. 2004). Furthermore virulent viruses can drive the host locally to extinction and therefore cut of their route for future transmission. A spatially structured host population therefore generally favor the evolution towards lowered pathogen virulence (van Baalen & Sabelis 1995; van Baalen & Rand 1998; Johnson & Boerijst 2002), but also viral infectivity (Boots 2007).
Competition between viruses is another important factor for the evolution of viral life history. Generally competition of parasites during co-infection of the same host is predicted to select for an increased replicating rate and an associated increase in virulence (van Baalen & Sabelis 1995). However, competition between strains can be affected by mechanisms other than resource competition, like for example direct interference between viral strains. If viruses interfere with each other’s replication, the production of viable offspring and the level of virulence may be reduced (Turner & Chao 1999; Chao et al. 2000; Brown et al. 2002).

1.5 Constraints imposed by molecular mechanisms

Viral adaptation to the prevailing environment through changes in the viral life history requires biochemical changes that alter the timing and balancing of the production of viral proteins and their enzymatic activity. Biochemical constraints that limit changes in the timing of the life-cycle can hamper optimal adaptation to a given environment. Constraints occur when the change in a biochemical aspect that is beneficial to one part of the life-cycle has negative side effects on other parts.

Overlapping genes are obvious examples for a mechanistic link between different aspects of the viral life cycle. For instance, gene A of bacteriophage ΦX174 can either be translated as a complete protein, or as a C-terminal fragment of gene A (called gene A*). Even though both proteins are encoded for by a single stretch of the viral genome they fulfil different tasks. Whereas gene A is essential for viral DNA replication, gene A* is affecting lysis and causes the shutdown of host DNA replication (Funk & Snover 1976; Colasanti & Denhardt 1985; Baas 1985). A mutation in gene A can therefore change several parameters of the life cycle simultaneously. In bacteriophage MS2 the lysis gene has an overlapping region with the coat and replication genes. Accordingly, changes in the RNA sequence of the lysis gene can
have consequences for the viral replication and packaging process, simultaneously.

Besides overlapping genes, the nucleotide genome of single-stranded viruses has a secondary structure that affects multiple processes like ribosome binding, gene regulation and the initiation of viral packaging. Mutations that affect viral secondary structure can therefore alter multiple aspects of viral reproduction simultaneously. A particularly illustrative example is the central operator loop of the RNA phage MS2 which is (1) part of the over-lapping lysis and replicase gene, forms (2) the translation termination hairpin for the coat gene which controls the expression of lysis proteins. It furthermore contains (3) the Shine-Dalgarno sequence and start-codon of the replicase gene, is (4) engaged in the control of replicase expression through the Min-Jou RNA-RNA interaction and forms (5) the nucleation point of the viral packaging process (Peabody 1997; Licis et al. 2000). The multifunctionality of the central MS2 operator loop leads to strong selection for the maintenance of the secondary RNA structure. Due to the importance of the secondary structure the realm of viable mutations in sequence space is confined to sequences that maintain this secondary RNA structure (Olsthoorn et al. 1994; Licis et al. 1998). Such feasible mutational paths have been studied extensively in computational models of RNA folding (Fontana & Schuster 1998; Schuster & Fontana 1999). These models predict the occurrence of extensive ‘neutral networks’, e.g. adjacent positions in genotype space that represent a particular RNA fold, and eventual nucleotide changes that radically change the RNA fold (Van Nimwegen & Crutchfield 2000).

Resources of the host cell are finite. Production of viral components is therefore limited by host-cell resources. Over-allocation into the replication of viral genomes, for example can hamper the production of viral proteins and therefore drastically reduce the production of viable viral offspring (Katanaev et al. 1996).
Additional constraints for the viral intra-cellular development result from the competition between the processes of genome replication, genome translation and genome packaging. The processes of genome replication and translation are mutually exclusive since replication moves along the viral genome in the 3’ to 5’ direction, whereas translation occurs in the 5’ to 3’ direction. Therefore the queuing of ribosomes on the viral genome interferes with genome replication (Eigen et al. 1991; van Duin & Tsareva 2004; Regoes et al. 2005). Furthermore, the process of genome packaging is competing with genome replication and translation. A rate of packaging that above an optimal level can therefore hamper viral genome replication and reduce viral fitness (Krakauer & Komarova 2003). Due to this interdependence of genome replication, translation and packaging evolutionary change in on aspect of the reproductive cycle will affect other aspects simultaneously. A prediction of optimal viral development therefore relies on the integration of these processes.

Biochemical constraints can also occur when a single protein affects multiple aspects of the life cycle. One example is the function of the virulence repressor protein cI in the life-cycle of phage λ. The cI protein represses the replication and host lysis of the lysogenic bacteriophage λ, after it has integrated into the host genome, thereby enabling vertical transmission (Oppenheim et al. 2005). At the same time the cI helps to defend the host cell against the take-over of a competing λ phage, by repressing the replication of this competitor inside the cell. Such a defence is called superinfection inhibition. The dual role of the repressor protein cI in the control of virulence and superinfection inhibition leads to a mechanistic coupling between these two traits.

Another example in which one protein effects multiple parts of the life cycle is the receptor-destroying enzyme of influenza virus. This receptor-destroying enzyme has the seemingly paradoxical task to destroy the target receptor on the host cell surface that is used for viral
attachment. The activity of receptor destroying enzymes is, however, required to release of viral particle from cell mucus and from the host cell after the process of viral budding. Since the detrimental and beneficial effects are both caused by one and the same protein, the costs and benefits of receptor destruction are unavoidably linked.

1.6 Overview of the thesis: How do genetic constraints limit the adaptation of the viral life-cycle?

In order to illustrate how biochemical constraints affect the path of viral evolution I studied four different classes of viruses and their particular biochemical limitations in the evolution of the viral life cycle.

In chapter 2 I disturbed viral gene expression balance and investigated how the virus can adapt its life cycle to this disturbance. For this purpose I cloned and overexpressed each of the four genes of MS2 from a plasmid leading to an excess supply of these genes during viral infection. Even though the virus receives one of its proteins ‘for free’, this disturbance of the viral gene balance is disastrous for the production of viral offspring. However, through adaptation to the conditions of gene over-expression the virus is able to restore some of its reproduction, supposedly through adaptation of its own gene regulation. The outcome of adaptation demonstrates some peculiar consequences of the evolution of viral protein balance. Adaptation to an excess of the coat gene led to an offspring production which exceeds the offspring produced by a wildtype infection in undisturbed condition whereas the adaptation to an excess of the replication protein is particularly difficult for the virus. Adaptation to the over-abundance of the lysis gene enabled the virus to reach a wildtype burst size at a much shorter lysis time and therefore greatly improved viral fitness. This indicates that lysis time of phage MS2 cannot evolve freely but is constrained by the expression and efficiency of the lysis protein.
Whereas chapter 2 studies the balance of viral replication within the host cell, chapter 3 is focused on the optimal strategy for transmission between host cells in a spatially structured host population. In bacteriophage the strategy of transmission is determined by the rate of attachment to the host cell. In a liquid medium it always pays to attach to a host when it is encountered. Therefore in liquid medium the rate of attachment should increase to its maximum. In a spatially structured host population, however, maximal attachment rate is not necessarily beneficial. When a phage grows on a spatial lawn of host bacteria it kills all bacteria in a focal area that is termed plaque. In the center of the plaque the absence of bacteria prohibits any further replication. Therefore it can pay for a phage to decide ‘not to infect’ and instead diffuse further out to the border of the plaque. By this mechanism the short-term benefit of immediate replication can be outweighed by the long-term benefit of future plaque growth. For this reason the phage should evolve lowered rates of host attachment in a spatial medium. I investigated this effect in an individual-based computer simulation and in experimental evolution with bacteriophage ΦX174. For this purpose I followed the evolution of attachment rates and plaque size of a phage strain that was previously adapted to a spatially structured host population into the conditions of a liquid environment. As expected adaptation to the liquid environment selected for an increased rate of attachment and at the same time decreased the ability to form plaques in a spatially structured host population. These experimental results were in line with my theoretical predictions. Yet, when I studied the growth curve of the evolved lines I discovered that the lines evolved in the liquid environment had a lower fitness than the ancestor in the liquid environment. An unknown factor must have led to a decrease in fitness during the course of adaptation.

The search for this factor is the topic of chapter 4. Ultimately, we had to conclude that the loss in burst size of the evolved ΦX174 strains was caused by a peculiarity in the bacterial host strain that we used for phage evolution. Our search revealed that this strain, E.coli C122 mutT,
harbors a hitherto unknown lysogenic bacteriophage that apparently interfered with the reproduction of phage ΦX174. Although unintended, the presence of this phage illustrates that viral competition can have important implications for the outcome of viral evolution. Viral strategies to deal with such competition are the main theme of chapter 4.

In chapter 5 I investigate the evolutionary role of molecular mechanisms that defend the host cells of a non-virulent, vertically transmitting virus against the take-over of a virulent, horizontally transmitting mutant. These mechanisms are called superinfection inhibition. In phage λ, superinfection inhibition is closely linked to the control of viral virulence, since virulence of the resident virus and replication of a superinfecting competitor are both controlled by the same protein (the cI protein). Due to this mechanistic link, increased virulence is associated with the cost of increased susceptibility to superinfection. This detrimental side effect of virulence can enable the persistence of vertically transmitting viruses or lead to the evolutionary co-existence of a vertically transmitting defense specialist and a horizontally transmitting attack specialist. Strikingly, the mechanisms that link virulence to superinfection inhibition also occur in other benign viruses of completely different origin (Retroviruses). The link between virulence and superinfection might therefore play a crucial role in the maintenance of the benign viral state in general.

In chapter 6 we study the constraints in the evolution of influenza virus that result from the opposing effects of the viral receptor-destroying enzyme on the viral life cycle. Influenza virus requires the receptor-destroying activity to prevent the agglomeration of viral particles on the host cell surface and in the cell mucus. However, receptor-destroying activity has an important side effect as it destroys the target receptor of viral binding to the host cell and therefore hampers the infection process. Due to these side effects the activity of receptor destroying enzymes should not evolve to a maximum, but
rather to an intermediate optimum. This optimum depends on the abundance and biochemical properties of the receptor and can lead to conditions that require viral specialization to one tissue or host at the expense of replication in another tissue or host. Since tissue specificity is an important factor for viral virulence and the route of viral transmission, the mechanisms of exclusive tissue specificity can offer a plausible mechanism for the dichotomy between virulent but poorly transmittable strains and non-virulent by highly transmittable strains of influenza virus.

These case studies lead to a row of observations that could lead to applications for to specific problems of viral evolution. Chapter 2, for example, shows that adaptation to the overexpression of the viral replication genes is particularly difficult for the virus. Furthermore, the expression of the lysis gene seems to be an important limitation for the viral life cycle. Knowledge of these bottlenecks can be invaluable in the design of anti-viral drugs. Chapter 3 shows that spatial confinement of epidemics might provide means to select for reduced levels of viral infectivity. Chapter 4 points out that self-imposed limitation of viral genome replication is crucial in the evolution of viral virulence. Chapter 5 describes the limitations of viral adaptation to anti-viral drugs that inhibit the activity of the viral neuraminidase protein. In all four case studies I demonstrate that modeling of the viral life cycle can aid to identify specific molecular targets that are the bottleneck of viral reproduction. Potentially, these targets might become the basis of future anti-viral strategies which are resilient to viral adaptation and drug resistance.

1.7 References


