Stochastic holin expression can account for lysis time variation in the bacteriophage λ

Abhyudai Singh¹ and John J. Dennehy²,3

¹Department of Electrical and Computer Engineering, Biomedical Engineering, Mathematical Sciences, University of Delaware, Newark, DE 19716, USA
²Department of Biology, Queens College, Queens, NY 11367, USA
³The Graduate Center, City University of New York, New York, NY 10016, USA

The inherent stochastic nature of biochemical processes can drive differences in gene expression between otherwise identical cells. While cell-to-cell variability in gene expression has received much attention, randomness in timing of events has been less studied. We investigate event timing at the single-cell level in a simple system, the lytic pathway of the bacterial virus phage λ. In individual cells, lysis occurs on average at 65 min, with an s.d. of 3.5 min. Interestingly, mutations in the lysis protein, holin, alter both the lysis time (LT) mean and variance. In our analysis, LT is formulated as the first-passage time (FPT) for cellular holin levels to cross a critical threshold. Exact analytical formulae for the FPT moments are derived for stochastic gene expression models. These formulae reveal how holin transcription and translation efficiencies independently modulate the LT mean and variation. Analytical expressions for the LT moments are used to evaluate previously published single-cell LT data for λ phages with mutations in the holin sequence or its promoter. Our results show that stochastic holin expression is sufficient to account for the intercellular LT differences in both wild-type phages, and phage variants where holin transcription and the threshold for lysis have been experimentally altered. Finally, our analysis reveals regulatory motifs that enhance the robustness of lysis timing to cellular noise.

1. Introduction

Phenotypic variation is an intrinsic feature of biological systems. Most phenotypic variation stems from genetic or environmental variation, but some phenotypic variation can exist even for identical genotypes in the same environment. This variation often emerges from the stochastic nature of biochemical reactions within a cell. Small fluctuations in transcriptional and translational machinery can lead to large variations in important regulatory molecules, impacting cellular switches, pathways and, ultimately, cell fates [1–5].

The causes and effects of stochastic variation have been best explored in microbial systems such as bacteriophages [6–8], bacteria [9–13], yeast [14–19] and human viruses [20–23]. For example, fluctuations in gene expression make it possible for populations of isogenic λ phages in the same environment to take one of two distinct developmental fates: immediate lysis of the host cell or lysogeny as a prophage in the host genome. This developmental switch is controlled by cytoplasmic levels of bacteriophage λ protein CII [24]. When CII levels are high, the lysogenic pathway is favoured, but when CII levels are low, phages will enter the lytic pathway. While the switch is essentially stochastic, the probabilities of lysis or lysogeny can vary according to environmental conditions and the multiplicity of infection [8,25–27].

Bacteriophage λ also possesses a simpler regulatory network that determines the timing of lysis. Lysis is accomplished by a suite of four genes: S (encodes holin and antiholin), R (encodes endolysin) and Rz/Rz1 (encode spanins). Expression of these genes, and all λ structural and assembly proteins, is under the control of the λ late promoter pR". While expression from promoter pR" is constitutive, transcription terminates upstream of the lysis genes [28]. When the 'decision' is made to pursue the
lytic life cycle, protein Q is produced and acts to prevent premature termination of late mRNA transcription approximately 10–15 min after lysogenisation [29,30]. Following the production of the late mRNA transcript, the lysis proteins are translated by host ribosomes. Holin accumulates evenly in the Escherichia coli plasma membrane until a genetically encoded time when they spontaneously aggregate to form large holes [31]. These holes allow endolysin and the spanins to access and disrupt their targets beyond the periplasmic space [32]. Subsequently, the cell ruptures and phage progeny are released into the surrounding medium.

As hole formation and cell rupture are nearly simultaneous, lysis timing depends on the accumulation of holin in the cell membrane up to a critical threshold [33]. Dennehy and Wang observed the lysis time (LT) of single λ lysogens using a temperature-controlled perfusion chamber mounted on a microscope stage (see electronic supplementary material, figure S1 and movie). Their measurements show that λ lysis occurred on average at 65 min, with an s.d. of 3.5 min (figure 1). Interestingly, mutations in the holin amino acid sequence can significantly change the LT mean and variance [34]. As randomness in the onset of the late promoter expression cannot account for these variations [34,35], stochastic holin expression and accumulation may be key in determining LT differences. Ultimately, the expression and accumulation of holin in the inner membrane depend on several factors: the rate of transcription of late mRNA, the rate of holin protein translation by host ribosomes, the rate of holin insertion into the plasma membrane and the threshold holin concentration required for membrane hole formation.

The impact of these factors on LT variation is investigated using a stochastic gene expression model. LT is modelled as the first-passage time (FPT) for cellular holin copy numbers to cross a critical threshold in the stochastic model (figure 2). Exact analytical formulae for the FPT statistical moments predict how LT mean and coefficient of variation (CV) change in response to experimental alterations of the holin transcription efficiency, translational efficiency and the threshold concentration of holin required for lysis. Some of these predictions are confirmed through analyses of LT distribution data obtained by Dennehy & Wang [34] for wild-type phage, and phages carrying mutations in either the holin sequence or the prR' promoter. We discuss the impact of these findings on the evolution of bacteriophage lysis timing as well as the implications for the study of variation in gene expression.

2. Stochastic gene expression model formulation

We denote by \( x(t) \), the number of holin molecules inside the cell at time \( t \), where \( t = 0 \) corresponds to the onset of holin expression. As holin does not degrade over relevant timescales [36], \( x(t) \) increases over time. The goal is to compute the FPT (time taken by a random walker to first reach a defined point) to the threshold \( X \geq 1 \) assuming \( x(0) = 0 \). The CV\(^2 \), defined as

\[
CV^2_{\text{FPT}} := \left( \frac{\langle \text{FPT}^2 \rangle - \langle \text{FPT} \rangle^2}{\langle \text{FPT} \rangle^2} \right)
\]

is used to quantify the randomness in FPT, where the symbol \( \langle \cdot \rangle \) denotes expected value. Let, \( T_i \) be independent and identically distributed random variables denoting the time interval between the \( i - 1 \)th and \( i \)th mRNA transcription event from the promoter.

We consider gene expression in the burst limit: transcribed mRNAs degrade instantaneously after producing a translational burst of protein molecules [37–40]. Consistent with measurements [41], we assume protein bursts are geometrically distributed with a mean translational burst size \( b = k_p/\gamma_m \). Here, \( k_p \) denotes the protein translation rate per mRNA and \( \gamma_m \) is the mRNA degradation rate. Thus, protein count

\[
x(t) = \sum_{i=1}^{n} B_i, \quad b := (B_i), \quad \forall i \geq 1
\]

is an accumulation of independent, identical geometric bursts \( B_i \), \( n \) is the number of transcription events in \([0,t]\) and the inter-burst arrival time distribution is given by \( T_i \).

3. Analytical calculation of first-passage time moments

First, the minimum number of transcriptional events \( N \) required for \( x(t) \) to reach the threshold \( X \) is quantified. Let,

\[
N = \inf \left( n \geq 1 : \sum_{i=1}^{n} B_i \geq X \right), \quad n \in [1, 2, \ldots] \Rightarrow
\]

\[
\text{Probability}(N \leq n) = \text{Probability}\left( \sum_{i=1}^{n} B_i \geq X \right).
\]

Then, the statistical moments for \( N \) are given by (see electronic supplementary material, appendix C)

\[
\langle N \rangle = \frac{X}{b} + 1 \quad \text{and} \quad \langle N^2 \rangle = \frac{X}{b} \left( 1 + \frac{b}{b} \right).
\]

The average number of transcriptional events required for lysis is the lysis threshold \( (X) \) divided by the mean translational burst size \((b)\) plus one. The plus one arises from the fact that, even if the threshold is close to zero, at least one transcription event is required for \( x(t) \geq X \). Given \( N \) (the number of burst events) and \( T_i \) (the time interval between bursts), time taken to reach the threshold is

\[
\text{FPT} = \sum_{i=1}^{N} T_i = \langle \text{FPT} \rangle = \langle N \rangle \langle T_i \rangle
\]

and

\[
\langle (\text{FPT})^2 \rangle = \langle N \rangle \langle (T_i)^2 \rangle + \langle (N^2) \rangle \langle (T_i)^2 \rangle.
\]
Below, distribution for \( T_i \) is determined for two different mRNA production processes: constitutive and bursty transcription.

### 3.1. Constitutive transcription model

Assuming mRNA production is a Poisson process with rate \( k_m \), \( T_i \), is an exponential random variable with mean \( 1/k_m \), and (3.3) yields

\[
CV_{FPT}^{2} = \frac{b^2 + X + 2bX}{(b + X)^2} \quad \text{and} \quad \langle FPT \rangle = \left(\frac{X}{b} + 1\right) \frac{1}{k_m}, \quad (3.4)
\]

where \( CV_{FPT}^{2} \) is the FPT CV\(^2\). For \( X/b \gg 1 \) (number of transcriptional events required for lysis is large), (3.4) simplifies to

\[
CV_{FPT}^{2} = \frac{1 + 2b}{X} \quad \text{and} \quad \langle FPT \rangle = \frac{X}{bk_m}, \quad b = \frac{k_p}{\gamma_m}. \quad (3.5)
\]

Note that under similar assumptions, we obtain

\[
CV_{FPT}^{2} = \frac{1 + b}{X} \quad \text{and} \quad \langle FPT \rangle = \frac{X}{bk_m}, \quad (3.6)
\]

for deterministic arrival of bursts (i.e. \( \langle T_i^2 \rangle - \langle T_i \rangle^2 = 0 \) in (3.3)). Thus, for \( b \gg 1 \), randomness in the FPT is halved for fixed time intervals between protein bursts compared with exponentially distributed intervals. When \( b \ll 1 \), we obtain the Poisson limit \( CV_{FPT}^{2} = X^{-1} \). The results above show that for a given \( X \), \( CV_{FPT}^{2} \) and \( \langle FPT \rangle \) can be independently controlled via holin translation rate \( (k_p) \) and transcription rate \( (k_m) \). Desired FPT mean and CV can be set by choosing

\[
b = \frac{CV_{FPT}^{2}X - 1}{2} \quad \text{and} \quad k_m = \frac{X}{\langle FPT \rangle}. \quad (3.7)
\]

Equation (3.5) predicts how FPT CV\(^2\) changes with mean as \( k_m \) or \( X \) are experimentally altered (figure 2):

- \( CV_{FPT}^{2} \) is inversely related to \( \langle FPT \rangle \) for perturbations in the lysis threshold. More specifically, increasing \( X \) will increase \( \langle FPT \rangle \) but decrease \( CV_{FPT}^{2} \).
- For alterations in translational efficiency \( k_p \) (and hence, the mean translational burst size \( b \)), increase in \( \langle FPT \rangle \) comes with a decrease in \( CV_{FPT}^{2} \).
- Changes in the transcriptional efficiency \( k_m \) alter \( \langle FPT \rangle \) without affecting \( CV_{FPT}^{2} \).

Next, a more sophisticated mRNA production process is considered, where the promoter switches between different transcriptional states.

### 3.2. Stochastic promoter switching model

Promoter transitions between active and inactive states have been well documented in various genes [42–46]. In the stochastic formulation, the promoter resides in active and inactive states for exponentially distributed time intervals with means \( 1/k_{off} \) and \( 1/k_{on} \), respectively, where \( k_{on} \) and \( k_{off} \) represent transition rates (figure 3, top right). From the active state, mRNA production is a Poisson process with rate \( k_m \). As before, we assume transcribed mRNAs degrade instantaneously after producing a geometrically distributed burst of protein molecules of mean size \( b \). Let, \( T_i \) be independent and identically distributed random variables denoting the time interval between the \( i - 1 \)th and \( i \)th transcriptional event from the active state. Then, the mean and variance for the inter-burst time interval is given by

\[
\langle T_i \rangle = \frac{k_{off} + k_{on}}{k_m k_{on}} \quad \text{and} \quad \langle T_i^2 \rangle - \langle T_i \rangle^2 = \frac{2k_m k_{off} + (k_{on} + k_{off})}{k_m^2 k_{on}}, \quad (3.8)
\]

for all \( i \geq 1 \) (see electronic supplementary material, appendix D). Using (3.3), (3.8) and assuming \( X/b \gg 1 \), we obtain

\[
\langle FPT \rangle = \frac{X}{b} \frac{k_{off} + k_{on}}{k_m k_{on}} \quad (3.9a)
\]

and

\[
\langle FPT^2 \rangle - \langle FPT \rangle^2 = \left(\frac{X}{b}\right) \frac{2k_m k_{off} + (k_{on} + k_{off})^2}{k_m^2 k_{on}^2} + \frac{X + b}{b} \left(\frac{k_{off} + k_{on}}{k_m k_{on}}\right)^2. \quad (3.9b)
\]

As expected, (3.9) reduces to (3.5) for constitutive transcription, i.e. \( k_{on} \gg k_{off} \) and promoter is always in the active state.

In the absence of promoter switching, decreasing transcriptional efficiency is predicted to increase \( \langle FPT \rangle \) without affecting \( CV_{FPT}^{2} \) (figure 2). However, inclusion of stochastic promoter

---

**Figure 2.** Predicted \( CV^2 \) versus mean trends for the FPT (arbitrary units) in the constitutive transcription model. LT is captured by the FPT for protein copy numbers to cross a critical threshold (a). Stochastic expression ensures that the threshold is reached at different times in individual cells, causing cell-to-cell variations in lysis timing. Decreasing holin transcriptional efficiency increases the mean FPT without altering the CV (b). By contrast, decreasing the holin translational efficiency or increasing the lysis threshold increases the mean FPT but decreases the CV. (Online version in colour.)
switching complicates the relationship between \( k_{\text{FPT}} \) and \( CV_2^{\text{FPT}} \). Qualitative \( CV_2^{\text{FPT}} \) versus \( k_{\text{FPT}} \) trends for perturbations in the transcriptional efficiency (measured by either \( k_{\text{on}} \), \( k_{\text{off}} \) or \( k_{\text{m}} \)) are shown in figure 3. Interestingly, modulating the transcriptional efficiency can have different effects on \( CV_2^{\text{FPT}} \) depending on the underlying parameter that is being altered. For example, decreasing \( k_{\text{on}} \) or \( k_{\text{m}} \) will both increase \( k_{\text{FPT}} \), but the former is predicted to increase \( CV_2^{\text{FPT}} \), while the latter decreases \( CV_2^{\text{FPT}} \) (figure 3). As in the previous section, \( FPT^{CV_2} \) changes inversely with the mean for alterations in \( X \) or \( b \).

4. Measured lysis time variation consistent with model predictions

If LT variation is indeed a result of noisy holin expression, then randomness in event timing should vary with mean as predicted by (3.5) or (3.9), i.e. \( CV_2^{\text{FPT}} \) should decrease with increasing \( FPT^{\text{CV_2}} \) as the lysis threshold is perturbed. Before we test this prediction, quantification of \( CV_2^{\text{FPT}} \) and \( FPT^{\text{CV_2}} \) from single-cell LT measurements is discussed.

4.1. Connecting lysis time to first-passage time

LT can be related to the FPT as \( LT = \text{FPT} + T_{pR'} \), where \( T_{pR'} \) is the time for the onset of transcription from the late promoter. As \( T_{pR'} \) is shown to be uncorrelated with FPT [35],

\[
(FPT) = (LT) - (T_{pR'}) \quad \text{and} \quad CV_2^{\text{FPT}} = \frac{\sigma_{LT}^2 - \sigma_{T_{pR'}}^2}{((LT) - (T_{pR'}))^2},
\]

where \( \sigma_I \) denotes the s.d. in \( I \in \{LT, T_{pR'}\} \), and \( (T_{pR'}) \) is taken to be 15 min [29,30]. As \( T_{pR'} \) variation only accounts for a small fraction of LT variability [34,35], we assume \( \sigma_{T_{pR'}} = 0 \). Non-zero values of \( \sigma_{T_{pR'}} \) do not change the \( CV_2^{\text{FPT}} \) versus \( FPT^{\text{CV_2}} \) trend described below. For wild-type \( \lambda \) (see electronic supplementary material, appendix B),

\[
(LT) \approx 65 \text{ min} \quad \text{and} \quad \sigma_{LT} \approx 3.5 \text{ min},
\]

which implies from (4.1),

\[
(FPT) \approx 50 \text{ min} \quad \text{and} \quad CV_2^{\text{FPT}} \approx 0.005.
\]

As we ignore \( T_{pR'} \) variation in these calculations, the reported values for \( CV_2^{\text{FPT}} \) are an upper bound. Assuming constitutive transcription from the wild-type \( pR' \) and a lysis threshold \( X \) of 1500 holin molecules [36], we obtain the following estimates for the translational burst size and the transcription rate from (3.5) and (4.3)

\[
b \approx 3 \quad \text{and} \quad k_{m} \approx 10 \text{ min}^{-1}.
\]

Noisy holin expression with these values is sufficient to generate intercellular LT differences similar to data (figure 4). These parameter values are consistent with previous estimates of holin transcription/translation efficiencies measured using bulk assays [36].

4.2. Mutations in the holin sequence

Mutations in the holin amino acid sequence can both increase and decrease LT mean and s.d. compared with wild-type [34]. These mutations may alter holin self-affinity or membrane insertion rate, thus affect the total cellular holin threshold.

Figure 3. Predicted \( CV_2 \) versus mean trends for the FPT (arbitrary units) in the stochastic promoter switching model. Schematic of the stochastic gene expression model where promoter randomly toggles between active and inactive states with rates \( k_{\text{on}} \) and \( k_{\text{off}} \) (top right). Transcription only occurs from the active state at a rate \( k_{\text{m}} \). Qualitative trends for the FPT \( CV_2 \) as a function of the mean, for alterations in \( k_{\text{on}} \) (top left), \( k_{\text{off}} \) (bottom right) and \( k_{\text{m}} \) (bottom left). Trends obtained from analysis of (3.9). (Online version in colour.)
required for lysis. For example, mutant holin with a stronger membrane binding affinity than the wild-type will have a larger fraction of the expressed holin present in the plasma membrane. Consequently, the critical holin concentration required for hole formation will be reached with less total holin expression. In our model, this corresponds to having a lower threshold $X$, and hence, a shorter LT compared with the wild-type.

Single-cell LT measurements for six holin amino acid sequence mutations (see electronic supplementary material, appendix B) reveal that shorter LTs are associated with higher CV to (figure 5). Thus, as predicted by a stochastic model of holin expression, results show an inverse relationship between CV and LT (figure 5). The inverse correlation between CV and LT is consistent with a stochastic model where LT differences arise due to cell-to-cell variations in holin expression rates between cells.

4.3. Mutations in the late promoter

Previous work measured LT variation for four lambda variants carrying mutations in the late promoter pR\' TATA box region [34]. Only two of these mutants showed a significant decrease in promoter strength, which increased both the LT mean and CV (see electronic supplementary material, appendix B). For example,

$$(\text{FPT}) \approx 95 \text{ min and } CV^{2}_{\text{FPT}} \approx 0.03,$$  

(4.6)

for promoter mutant SYP208 [34]. Note a twofold increase in (FPT), and a sixfold increase in $CV^{2}_{\text{FPT}}$, compared with their respective values in wild-type lambda (see (4.3)). The observed increase in $CV^{2}_{\text{FPT}}$ cannot be explained by a reduced transcription rate $k_{\text{on}}$ in the constitutive model, as in this case CV is invariant of $k_{\text{on}}$, (figure 2). However, as discussed below, these results can be explained by a model that includes random promoter transitions between active and inactive states (figure 3).

Results in §3.2 show that in the stochastic promoter switching model (figure 3, top right), reducing transcription from the active state (decreasing $k_{\text{on}}$) increases (FPT) but decreases $CV^{2}_{\text{FPT}}$. By contrast, enhancing the stability of the inactive state (decreasing $k_{\text{off}}$) increases both (FPT) and $CV^{2}_{\text{FPT}}$. Finally, for decreasing active-state stability (increasing $k_{\text{off}}$), $CV^{2}_{\text{FPT}}$ exhibits an inverted U-shaped profile with increasing (FPT). Thus, an increase in both (FPT) and $CV^{2}_{\text{FPT}}$ in response to a reduced holin transcription is consistent with promoter mutations decreasing $k_{\text{on}}$ or increasing $k_{\text{off}}$ in the promoter switching model (figure 3). A recent study has shown that promoter mutations in the TATA box region primarily modulate the transcription burst frequency $k_{\text{on}}$ [48]. Assuming constitutive transcription from the wild-type pR\' ($k_{\text{on}} \gg k_{\text{off}}$), mutations that significantly decrease $k_{\text{on}}$ will induce additional promoter switching noise in holin expression.
expression. Consequently, weaker non-constitutive late promoter expression makes cell lysis longer, and more unpredictable compared with wild-type \( \lambda \).

5. Discussion

The timing of cellular events, such as mitosis, apoptosis, developmental differentiation and reversion from latency, frequently depends on the concentrations of essential regulator molecules. Stochastic gene expression causes variation in regulator molecule concentrations even for identical cells induced at the same time. Consequently, molecular threshold concentrations will be reached at different times in different cells, giving rise to significant variation in the timing of critical cellular events across a homogeneous cell population. The holin lysis system of bacteriophage \( \lambda \) provides a simple system to study how event timing is regulated at the single-cell level. The main contributions of this work are as follows:

(i) LT was determined by the FPT for cellular holin levels to reach a critical threshold. Exact analytical formulae for the FPT mean and CV\(^2\) were derived for different stochastic gene expression models. These formulae are given by (3.4) and (3.9) for the constitutive transcription model and the stochastic promoter switching model, respectively.

(ii) FPT derivations reveal that for a fixed lysis threshold, LT CV\(^2\) is minimized by reducing the translation burst size \( b \). As discussed later, mechanisms for reducing \( b \) are embedded in \( \lambda \)'s lytic pathway in order to ensure precision in lysis timing.

(iii) Using (3.4) and (3.9), predictions on how biologically meaningful parameters (transcription rate, translation rate, lysis threshold, etc.) impact LT statistics were generated. These predictions are summarized in figure 2 (for the constitutive transcription model) and figure 3 (for the stochastic promoter switching model).

(iv) Analysis of data from [34] confirms model predictions. In particular, when the threshold for lysis is altered through holin sequence mutations, increases in mean LT are associated with decreases in LT CV\(^2\). By contrast, when holin lysis is alternated through holin promoter mutations, increases in mean LT are associated with increases in CV\(^2\).

Below, we discuss the implications of some of these findings in further detail.

5.1. Independent control of lysis time mean and variation

Formulae for the FPT moments reveal how transcriptional/translation efficiencies independently modulate the mean and variation in lysis timing. For constitutive transcription, FPT CV\(^2\) increases linearly with the translational burst size \( b \), but is independent of the transcription efficiency (figure 2). Precision in the FPT can be achieved by setting \( b \) sufficiently small. When \( b \ll 1 \), the limit of noise suppression for a given lysis threshold \( X \) is attained: CV\(^2\)\(_{FPT} \approx X^{-1} \). This limit corresponds to a Poisson arrival process for holin production. Once \( b \) is chosen for a desired CV\(^2\)\(_{FPT} \), holin transcription efficiency can be adjusted to have any mean FPT.

Note that \( b = k_p / \gamma_m \) is proportional to the translation efficiency \( k_p \), where \( \gamma_m \) is the mRNA degradation rate. In \( \lambda \), all lysis and structural proteins are translated from the same poly-cistronic mRNA transcript. Hence, reducing \( b \) by making \( k_p \) small is a more viable option than increasing \( \gamma_m \) as the latter would affect the translational burst sizes of structural proteins. S gene, which encodes holin, is reported to have a low translational efficiency [36]. As discussed later, holin translation bursts are further reduced through production of a second protein, antiholin, from the S mRNA. Given the direct relationship between CV\(^2\)\(_{FPT} \) and \( b \), mechanisms in place for minimizing \( b \) are essential for buffering randomness in lysis timing.

Our results on the FPT are complementary to previous work quantifying protein copy number variations in stochastic gene expression models. These studies have also shown that a reduced translation efficiency is critical for minimizing stochastic fluctuations in protein levels [13,38]. However, there are qualitative differences in the formulae for the FPT and protein copy number statistical moments. For example, the FPT CV\(^2\) is proportional to the translation efficiency \( k_p \). By contrast, for \( b \gg 1 \), the steady-state protein copy number CV\(^2\) is independent of \( k_p \) [13,38,49].

5.2. An incoherent feed-forward circuit minimizes lysis time stochasticity

The \( \Lambda \) S gene encodes two proteins with opposing function: holin and antiholin [50–52]. These proteins are produced in a 2:1 ratio, i.e. for every two holins, there is one antiholin molecule. These two proteins differ only in that antiholin has an extra two residues, methionine and lysine, at the N-terminus. The positively charged lysine prevents antiholin from accessing holin’s characteristic three transmembrane domain configuration in the E. coli inner membrane [53]. Experimental evidence suggests that antiholin binds holin and sequesters it in the cytoplasm, preventing it from participating in hole formation [54,55]. This creates an incoherent feed-forward circuit in \( \Lambda \)'s lytic pathway [56].

Preliminary analysis provides some clues on the functioning of this feed-forward circuit. Let, holin and antiholin translation from the mRNA be Poisson processes with rates \( k_h \) and \( k_{ah} \), respectively. Then, the probability of producing \( i \) holin and \( j \) antiholin molecules from the mRNA in time interval \([0, t]\) will be

\[
\frac{(k_h t)^i (k_{ah} t)^j}{i! j!} \exp \left( - (k_h + k_{ah}) t \right), \quad i, j \in \{0, 1, 2, \ldots \}. \quad (5.1)
\]

We denote by \( x \) and \( y \), the total number of holin and antiholin molecules synthesized by an individual mRNA before it decays. Random variables \( x \) and \( y \) would be correlated, because an mRNA that lives longer than average by random chance would create more of both holin and antiholin. Assuming each mRNA lives for an exponentially distributed time interval with mean \( 1 / \gamma_m \), the joint probability distribution of \( x \) and \( y \) is

\[
\text{Probability}(x = i, y = j) = \gamma_m \int_0^\infty \frac{(k_h t)^i (k_{ah} t)^j}{i! j!} \exp \left( - (k_h + k_a + \gamma_m) t \right) dt
\]

\[
= \frac{k_h^i k_{ah}^j \gamma_m (i + j)!}{i! j! (k_h + k_{ah} + \gamma_m)^{i+j+1}}, \quad i, j \in \{0, 1, 2, \ldots \}. \quad (5.2)
\]
Consider holin–antiholin interaction in a limiting case: antiholin rapidly binds to holin making holin inactive in an irreversible reaction. Then the active holin translational burst size (number of active holin made per mRNA) would be \(x - y\) (if \(x > y\)) or 0 (if \(x \leq y\)). Equation (5.2) yields the following mean translational burst size \(b\) for active holin

\[
b = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} (i - j) \frac{k_m^j y_m (i + j)!}{n! (k_a + k_{ah} + y_m)^{i+j+k}}
\]

\[\text{(5.3a)}\]

and

\[
\approx \frac{k_a - k_{ah}}{y_m} \quad \text{if} \quad k_a \gg k_{ah}, \quad \text{(5.3b)}
\]

which decreases with increasing antiholin production rate \(k_{ah}\). Recall from (3.5), that for a fixed \(X\) and \(\text{FPT}\), \(\text{CV}^2_{\text{FPT}}\) is minimized by reducing \(b\) (see equation (3.5)). Thus, antiholin may enhance precision in lysis timing by lowering \(b\), and making the active holin production process less bursty. Removing antiholin \((k_{ah} = 0)\) would increase \(b\), which is predicted to decrease \(\text{FPT}\) but increase \(\text{CV}^2_{\text{FPT}}\) (equation (3.5)). Indeed, \(\lambda\) variant IN62 (no antiholin expression) lysed 10 min faster but with a twofold higher \(\text{CV}^2_{\text{FPT}}\) compared with wild-type [34]. Our results show that as in other systems [57–59], antiholin-mediated incoherent feed-forward loop is a key regulatory motif in \(\lambda\)’s lytic pathway for buffering LT from stochastic gene expression.

5.3. Evolution of lysis timing

In life-history theory, the timing of lysis is viewed as a trade-off between present and future reproduction. Assuming a linear rate of progeny assembly [60], postponing lysis increases progeny number, but delays infection of new hosts. In other words, increasing fecundity also increases generation time. Several studies have theoretically and experimentally examined the optimal timing of lysis [60–66]. These studies reach similar conclusions: when host density and quality are high, early lysing genotypes will be selectively favoured as they will have more grandchildren than competing genotypes. Conversely, poor host quality and low host densities should favour longer LTs.

We expect phage fitness will be maximized if the LT CV is minimized around the optimum LT for any particular environment. Indeed, we find that wild-type \(\lambda\) has a lower LT CV and a higher fitness than all laboratory-generated mutant strains [34,60]. Furthermore, \(\lambda\)’s lytic pathway exhibits the ingredients required for precision in timing: low \(S\) gene translational efficiency [36] and antiholin synthesis to further reduce the active holin translational burst size. At first, production of antiholin seems energy inefficient as it inactivates holin, translated from the same mRNA. Our analysis and data suggest that antiholin may reduce LT variations and is reminiscent of the classical trade-off between energy efficiency and noise buffering in cellular systems [67,68].

Our results predict how the mean LT should evolve under selection, while minimizing LT CV. In principle, mean LT can be modulated through mutations in the following regions: the \(pR\) promoter sequence, the holin leader sequence and the holin coding sequence. Because of the polycistronic nature of the \(pR\) promoter sequence, we expect it can be ruled out. Presumably, mutations in \(pR\) will have undesirable pleotropic effects on the expression of \(\lambda\)’s structural and regulatory proteins. We predict that selection for shorter LTs (i.e. when host quality and density is high) will favour mutations that increase translational efficiency, and not those that lower the threshold concentration of holin required for lysis (\(X\)). From (3.5), we see that, when translational burst size \((b)\) is small, \(\text{CV}^2 \approx 1/X\), and more sensitive to changes in \(X\) rather than \(b\). So decreasing mean LT by lowering \(X\) will increase \(\text{CV}^2\). By contrast, selection for longer LTs will favour mutations that increase \(X\) because this will decrease LT \(\text{CV}^2\), while mutations decreasing translational efficiency will not substantially change \(\text{CV}^2\). These predictions can be tested by genetic sequencing of \(\lambda\) phage following experimental evolution under high or low host densities.

5.4. Burst size implications of lysis timing variation

Given that progeny are assembled at a linear rate, we expect that differences in LT will be reflected by differences in progeny burst size (total progeny produced in a single infection). Consequently, we should observe a direct relationship between LT stochasticity and burst size stochasticity. Seminal work by Delbrück showed that phage burst size can vary significantly from cell-to-cell, even after controlling for cell size [69]. Interestingly, measured viral burst size CV is an order of magnitude larger than the LT CV (see electronic supplementary material, appendix E). Recent single-cell measurements also show a broad viral burst size distribution for vesicular stomatitis virus (VSV), an animal virus that can occasionally infect humans. Data show a 300-fold cell-to-cell difference in the number of VSV progeny produced per cell, and these differences cannot be attributed to viral genetic variation [70,71]. Phage \(\lambda\) provides an ideal system to study how viral burst size variations arise from LT variations and noisy expression of structural genes. Understanding how non-genetic factors affect viral fitness at the single-cell level has important implications for health and disease.

6. Summary

In summary, cell-to-cell LT differences arise due to the stochastic transcription and translation of the lysis timekeeper, holin. We derive analytical formulae for the mean and \(\text{CV}^2\) of the time it takes for holin to reach a critical concentration in the inner membrane. These formulae are used to generate predictions regarding the impact of biologically relevant parameters, such as transcription and translation rates, on LT mean and \(\text{CV}^2\). We evaluate these predictions using previously published data on single-cell variation in LT [34]. Our analysis confirms that stochastic holin expression can account for variation in LT. Furthermore, we find that the holin lysis system contains several features designed to minimize noise.

Acknowledgements. We thank Andrew Rutenberg, Gillian Ryan, Ing-Nang Wang, Pak-Wing Fok, Pavol Bokes, Ryan Zurakowski, Khem Ghusinga and Ryland Young for discussion on some of the ideas contained herein.

Funding statement. J.J.D. was supported by grants from the Professional Staff Congress of the City University of New York and the National Science Foundation (Division of Environmental Biology Awards 0804039 and 1148879, and Division of Molecular and Cellular Biosciences Award 0918199). A.S. was supported by the National Science Foundation grant no. DMS-1312926, University of Delaware Research Foundation (UDRF) and Oak Ridge Associated Universities (ORAU).