Noise and Stochasticity in Gene Expression: A Pathogenic Fate Determinant

Mikkel Girke Jørgensen, Renske van Raaphorst, Jan-Willem Veening

Molecular Genetics Group, Groningen Biomolecular Sciences and Biotechnology Institute, Centre for Synthetic Biology, University of Groningen, Groningen, The Netherlands

1Corresponding author. e-mail address: j.w.veening@rug.nl

1 INTRODUCTION

Individual cells in a bacterial population grown in the same environment never exhibit exactly the same phenotype, despite a common genetic identity. This phenomenon is known as phenotypic variation. At the genetic level, this means that individual isogenic cells in a bacterial population show variable gene expression patterns, which translate into changes in protein levels and thus influence the cellular behaviour to a certain extent. Stochastic gene expression, or gene expression ‘noise’, has been proposed as a major source of this variability (reviewed in de Lorenzo & Perez-Martin (1996); Locke & Elowitz (2009); Munsky, Neuert, & van Oudenaarden (2012)). This phenomenon has, until recently, received little scientific attention, in part because classical molecular microbiologists have assumed that isogenic cells respond in a nonfluctuating fashion to a given stimulus and, thus, the traditional techniques used in laboratories to study gene expression rely on pooling of millions of cells and therefore determine the average values for the entire population (Dubnau & Losick, 2006). However, with the emergence of single-cell analytical techniques such as flow cytometry and fluorescence microscopy, it has become clear that phenotypic variation is ubiquitous in nature and occurs in many biological processes.

In general, gene expression is intuitively a ‘noisy’ process with multiple molecular origins. While beneficial for certain traits such as virulence (see the succeeding text), noise in gene expression is altogether an unwanted by-product for the cell and all organisms, both prokaryotes and eukaryotes, tend to reduce and control it in relation to essential processes. An example of such a control mechanism is negative feedback loops (Fraser, Hirsh, Giaever, Kumm, & Eisen, 2004; Dublanche, Michalodimitrakis, Kummerer, Foglierini, & Serrano, 2006; Smits, Kuipers, & Veening, 2006). Strikingly, it was found that autoregulated systems such as self-repression show a decrease in noise levels compared to an unregulated system (Becskei & Serrano, 2000).
2 ORIGINS OF NOISE

To explain the concept of noise, we can start by considering a bacterial cell about to divide. Let us assume that a cell in this state will have all its molecules in Brownian motion. This means that all particles or molecules move randomly around the cell. Therefore, the chance that each new daughter cell will inherit the same number of ribosomes, transcription factors, RNA polymerases and so on is negligible (Munsky et al., 2012). If we scale that up, the chance that two cells out of an entire isogenic population, grown under the same conditions, are identical with respect to the number and composition of their molecules is minute. As a natural consequence of this randomness in the distribution of molecules, there will be cell-to-cell variations in most biological processes in the population. This type of variation is termed extrinsic noise.

The mechanisms causing extrinsic noise in gene expression include, for example, the concentrations of RNA polymerases and ribosomes or regulatory factors that would lead to variations in the level of expression of a given gene between one cell and another but not between two identical genes in the same cell (Swain, Elowitz, & Siggia, 2002; Raser & O’Shea, 2005). Thus, extrinsic noise arises from sources that are global to a single cell but vary from one cell to another. Intrinsic noise on the other hand arises from random fluctuations in the biochemical process of gene expression itself regardless of the presence of extrinsic noise. Let us consider a purely hypothetical population of isogenic bacterial cells, containing exactly the same number and composition of molecules. The amount of protein produced by any given gene would still fluctuate from one cell to the other (Elowitz, Levine, Siggia, & Swain, 2002). This intrinsic noise arises from randomness in the binding of transcription factors to the promoter region, of RNA polymerases to the promoter and of ribosomes to the ribosomal binding site on the messenger RNA. The rate of translation can also contribute to intrinsic noise, since the availability of tRNA for each ribosome and mRNA turnover by ribonucleases are processes that show probabilistic behaviours. The counterpart to protein synthesis, protein degradation, also adds to the pool of intrinsic noise.

Thereby, we can in principle define intrinsic noise as that arising directly from the process of gene expression and extrinsic noise as that arising from changes in the intercellular environment (Elowitz et al., 2002; Swain et al., 2002; Dublanche et al., 2006). ‘Noise’ can therefore be defined as variation in gene expression in a population of isogenic cells. Phenotypic variation and total noise is the sum of all these parts. A typical image of gene expression noise is seen in Figure 6.1A. Here, isogenic cells of Streptococcus pneumoniae are expressing gfp. Even a constitutive promoter generates cell-to-cell variability, which can be observed and quantified by fluorescence microscopy.

2.1 Transcriptional bursting

Recently, a set of exciting experiments with single-molecule detection has shown that gene expression occurs as transcriptional bursts. In a hallmark paper, Golding and coworkers detected individual mRNA transcripts in individual living cells of
Escherichia coli in real time (see the succeeding text; Golding, Paulsson, Zawilski, & Cox, 2005). Strikingly, the authors found that E. coli cells produced transcripts in short distinct bursts and not at a steady rate as would be expected according to a normal distribution (Golding et al., 2005; Raj & van Oudenaarden, 2008). Based on their data, the authors suggested that transcriptional bursting is a result of an ON/OFF

**FIGURE 6.1**

Image analysis and data output for measuring noise in gene expression. (A) Microscopic image acquisition of an isogenic culture of Streptococcus pneumoniae grown under identical conditions. The cells express GFP from a constitutive promoter; see phase contrast image (left) and GFP signal (right image). Even when grown under identical conditions, the fluorescence (and likely the protein levels) of GFP varies from one cell to another (compare the two cells marked by white arrows). (B) A typical FACS output of the cells in (A). Here, 10,000 S. pneumoniae cells are analysed within a few seconds. From the graphical output, it is clear that the level of GFP can be described in statistical terms such as the mean value $\langle p \rangle$ and a standard deviation $\sigma_p$.

*Escherichia coli* in real time (see the succeeding text; Golding, Paulsson, Zawilski, & Cox, 2005). Strikingly, the authors found that *E. coli* cells produced transcripts in short distinct bursts and not at a steady rate as would be expected according to a normal distribution (Golding et al., 2005; Raj & van Oudenaarden, 2008). Based on their data, the authors suggested that transcriptional bursting is a result of an ON/OFF
model where gene expression randomly switches back and forth between active and inactive transcription. Consequently, in light of these authors’ findings, a simple Poisson stochastic model of gene expression is unlikely to apply. The Poisson expression model, in which transcripts are constitutively expressed at a constant rate and degraded in a first-order reaction, is the simplest model (Li & Xie, 2011; Munsky et al., 2012). Moreover, the transcriptional bursting model observed in *E. coli* reflects findings in eukaryotic cells (Raj & van Oudenaarden, 2008). At the mechanistic level, transcriptional bursting can be explained by promoter kinetic and promoter transition states. Binding of a transcriptional activator to the promoter region, for instance, would likely lead to several RNA polymerases associating with the promoter in successive order, thereby resulting in several rounds of transcription before the activator disassociates from the promoter again. This ON/OFF switching would then result in high and low transcription rates, respectively (Kaern, Elston, Blake, & Collins, 2005). Zenklusen and coworkers found both constitutively expressed and transcriptional bursting genes using *Saccharomyces cerevisiae* as a model organism. These findings reflect a complex nature of gene expression and demonstrate several modes of transcription modulation (Zenklusen, Larson, & Singer, 2008). It should be noted that transcriptional bursting may not be a common source of noise in prokaryotes. In fact, most models concerning prokaryotic gene expression assume that the transition between ON and OFF switching is so fast that the promoters are in steady states (Kaern et al., 2005). In addition, studies in *E. coli* and *Bacillus subtilis* revealed that the stochasticity in protein production was a result of translational bursting and not transcriptional bursting (Ozbudak, Thattai, Kurtser, Grossman, & van Oudenaarden, 2002; Yu, Xiao, Ren, Lao, & Xie, 2006). Transcriptional bursting is thought to be more important in eukaryotic gene expression and the transcriptional bursting observed in bacteria is much weaker and measured only on an inducible gene (Golding et al., 2005; Zenklusen et al., 2008). However, recent single-molecule fluorescence *in situ* hybridization experiments revealed that small RNAs are often bimodally expressed in clonal bacterial populations and, in the ON cells, typically contain between 1 and 10 copy numbers, in line with a bursting model of gene transcription (Shepherd et al., 2013). In eukaryotic systems, chromatin remodelling between open and closed structures correspond to the ON and OFF promoter transition states, and it has become the prominent model for gene expression control (Blake et al., 2006; Chubb, Trcek, Shenoy, & Singer, 2006; Kaufmann & van Oudenaarden, 2007; Zenklusen et al., 2008).

### 3 Measuring Noise

In recent years, it has become possible to detect and analyse gene expression at the single-cell and single-molecule level. These advances in the experimental protocols have been pivotal in detecting and quantifying the variability of gene expression. In particular, the development of green fluorescent protein (GFP) as a reporter to count the number of molecules—either mRNA or protein—has been essential. Today, the
use of GFP (or its variants), coupled to single-cell imaging, is the method of choice for studying noise in gene expression (Figure 6.1A).

Most techniques used in molecular biology require pooling of a vast number of cells, resulting in an averaging that is not representative for individual single cells (for instance, samples taken from bacterial cultures for downstream analyses such as microarray and proteomics). By studying population heterogeneity, it has become clear that each cell in a population, even when grown at identical conditions, is unique. Depicting gene expression as a single average value based on a population sample is therefore somewhat misleading and must be interpreted with care. Rather, gene expression in a population is more accurately represented by a distribution with associated statistical properties such as standard deviation and variance (Nevozhay, Adams, Van Itallie, Bennett, & Balazsi, 2012; Figure 6.1B). Fluorescence-based techniques are the most direct tools for determining gene expression noise, either by measuring protein or mRNA levels in single cells (Figure 6.1A; Larson, Singer, & Zenklusen, 2009).

Today, several types of fluorophores are available to visualize cell population heterogeneity. The principal methods for the detection of fluorescent signals are fluorescence-activated cell sorting (FACS) and microscopy-based cell imaging. Both techniques measure the fluorescence emitted at the single-cell level. FACS has the advantage of analysing several thousands of cells at once (Figure 6.1B). However, cell sorting is less sensitive and does not facilitate the detection of low-abundance signals (Larson et al., 2009). Microscopy-based cell imaging, on the other hand, is more sensitive with a high dynamic range, and single cells can be tracked over time and generations (de Jong, Beilharz, Kuipers, & Veening, 2011). Data acquisition, conversely, requires several time series of pictures and fewer cells are analysed compared to FACS (Larson et al., 2009).

Measuring noise for a given gene usually involves promoter fusion to a fluorescent protein like GFP and then determining the total level of fluorescence signal per cell either by flow cytometry or microscopy. FACS analysis immediately generates a fluorescent value for each cell including statistical properties such as mean signal and standard deviation. Microscopy images on the other hand require additional analysis. Several software tools exist to quantify the level of fluorescent signal emitted by each cell from microscopic images, for example, ImageJ (Schneider, Rasband, & Eliceiri, 2012), MicrobeTracker (Sliusarenko, Heinritz, Emonet, & Jacobs-Wagner, 2011), and Schnitzcells (Young et al., 2012). Although the use of flow cytometry and microscopy in combination with a fluorescent reporter protein facilitates the monitoring of gene expression at the single-cell level, it should be borne in mind that the correct folding and maturation of these proteins themselves can also contribute to the total noise.

The detection of single proteins is a challenging task as it requires the recording of each single protein in a cell. This is made difficult as proteins diffuse within the cytoplasm making the signal spread and image acquisition problematical. In a milestone paper, Yu and coworkers overcame this difficulty by making a fusion protein of a yellow (YFP) variant of GFP to the membrane-bound Tsr protein as a reporter to monitor the lac promoter activity (Yu et al., 2006). The membrane
anchoring slowed the diffusion rate, thereby making it possible to study single proteins in live cells by fluorescence microscopy. The use of photobleaching allowed the authors to monitor single-protein production, and they observed that proteins were produced as bursts from a single mRNA molecule (Yu et al., 2006; Larson et al., 2009).

Another way to detect noise in gene expression is by two-photon fluorescence fluctuation microscopy. In brief, this technique measures the intensity fluctuations of signals from a fluorescent protein inside a cell at each pixel in a set of fast scanned images, which are then deconvolved, allowing for counting of the molecular brightness and determination of the absolute number of fluorescent proteins diffusing inside cells. Using this novel approach, Ferguson and coworkers recently showed that a B. subtilis glycolytic promoter driving GFP showed strong transcriptional bursting and, surprisingly, that for highly ‘bursty’ promoters, negative feedback does not suppress the noise (Ferguson et al., 2012).

Fluorescence in situ hybridization (FISH) is a cytogenetic technique used to detect and quantify single mRNA (or DNA) species within a cell. FISH uses DNA probes complementary to the specific target mRNA. As the probes are conjugated to a fluorescent dye, fluorescent microscopic imaging allows the quantification and localization of the mRNA. Usually, several probes complementary to the same mRNA are used to enhance the signal-to-noise ratio (Femino, Fay, Fogarty, & Singer, 1998; Larson et al., 2009; Trcek et al., 2012). However, the cells have to be fixed, limiting the method to visualizing only the mRNA at a snapshot in time. Conversely, the MS2 system based on the RNA-binding phage protein MS2 fused to a fluorescent protein like GFP allows the trafficking of mRNA molecules in live cells. Using a reporter mRNA with multiple stem–loop structures recognized by the MS2 fusion protein provides a molecular beacon, which can be visualized by fluorescent microscopy. The method benefits from using living cells as the mRNA reporter molecule is tracked and quantified (Golding et al., 2005; Querido & Chartrand, 2008; Trcek et al., 2012).

Recently, a set of RNA aptamers have been designed that bind fluorophores (Paige, Wu, & Jaffrey, 2011). These RNA–fluorophore complexes can be used to visualize and localize mRNAs in live cells. One such RNA–fluorophore complex, called Spinach, closely resembles the fluorescence properties of GFP (Paige et al., 2011) and such RNA-adapter approaches are rapidly becoming a popular combination to study RNA dynamics in live cells (Armitage, 2011) and we foresee that they will be used to study noise in gene expression.

4 ENGINEERING NOISE

Noise in gene expression has long been predicted and several studies have quantified, measured and modelled noise in many genetic systems. The field of research has expanded rapidly and attracted scientists ranging from geneticists and biophysicists to theoretical mathematicians. From a system biologist’s point of view, measuring
and quantifying noise is not so much the goal as to seek to model and engineer the level of noise. Constructing robust gene circuits is a challenging task and optimizing a gene network usually requires minimizing heterogeneity (Kaern, Blake, & Collins, 2003). Also, for biotechnological applications, it might be desirable to reduce heterogeneity in production of a commercially valuable molecule, for instance.

Several parameters are useful when discussing phenotypic variability and gene expression noise. The distribution of gene expression of a single gene can be described by a mean value of expression denoted \( \langle p \rangle \) with a standard deviation \( \sigma_p \); see Figure 6.1B. The relative standard deviation \( \sigma_p/\langle p \rangle \) is sometimes used as a measure of noise. However, the Fano factor \( (\sigma_p^2/\langle p \rangle) \), or phenotypic noise strength, is a more commonly used measurement of noise. This is because the relative standard deviation changes as the mean value changes, whereas the phenotypic noise strength is less sensitive to changes in the mean value. The Fano factor is thus a noise measurement that directly correlates with the width of the population distribution (Thattai & van Oudenaarden, 2001; Ozbudak et al., 2002; Kaern et al., 2003). Another important measure is the coefficient of variance, \( \sigma_p^2/\langle p \rangle^2 \), and this is important in relation to engineering of noise since it gives a measure of the signal-to-noise ratio (Kaern et al., 2003).

In a landmark paper, Ozbudak and colleagues investigated the biochemical contribution to stochastic gene expression using *Bacillus subtilis* as a model organism (Ozbudak et al., 2002). The authors fused the *gfp* gene in front of an IPTG-inducible promoter and quantified the noise level by FACS analysis. By adding various amounts of IPTG and making targeted mutagenesis in either the promoter region or the ribosomal binding site, the authors were able to determine the source of the noise. In one set of experiments, the transcriptional rate was changed by varying the IPTG concentration and the authors found that the transcriptional efficiency did not significantly affect noise strength. In contrast, by making point mutations in the ribosomal binding site, the authors found a strong positive correlation with translational efficiency (Ozbudak et al., 2002). This study provides the engineer with tools to modify the level of noise for a single gene. Increasing the rate of translation, for instance, increases the noise strength and, conversely, decreasing the rate of translation leads to a reduction in noise strength. If, for example, the desired outcome in protein production is to stay constant, changing the rate of translation can be counterbalanced by changing the rate of transcription. In the model organism *Escherichia coli*, it has been observed that key regulatory proteins display reduced translational rates, thereby minimizing noise at the protein level (the ultimate measure of gene expression) (Ozbudak et al., 2002; Raser & O’Shea, 2005).

In a recent study, Mutalik, Guimaraes, Cambray, Lam, et al. (2013) showed that the identity of variation in genetic elements is more complex. By constructing a full combination library of different promoters and 5′ untranslated regions (UTR) fused to either *gfp* or *rfp*, the researchers monitored the amounts of mRNA and protein for all combination. This systematic approach revealed that the 5′ UTR containing the Shine–Dalgarno element is a key contributing factor to variance (Mutalik, Guimaraes, Cambray, Mai, et al., 2013). Control of the way genetic elements are
structured allow a reliable, establishment of basic principles for genetic designs (Mutalik, Guimaraes, Cambray, Lam, et al., 2013).

An important factor contributing to noise is the so-called finite number effect. Basically, the hypothesis predicts that stochastic effects and noise are more prominent when only a few molecules of a process are present. Thus, increasing the number of molecules for a chemical reaction is an effective way of reducing noise (Smits et al., 2006).

Transcriptional bursting, discussed earlier, may not necessarily result in an overall heterogeneous protein pool within a population, since its effects can be buffered by reducing the rate of protein degradation (Raj, Peskin, Tranchina, Vargas, & Tyagi, 2006). Therefore, when constructing a gene regulatory circuit, system noise can be reduced at the protein level by slow decay rates.

Recently, it was suggested that one source of noise that is important for phenotypic variation is transcription fidelity (Gordon et al., 2009). It was found that E. coli cells that supposedly display reduced fidelity of RNA polymerase (RNAP) activity, thereby increasing the mistakes in transcribing DNA into RNA, had a perturbed switching frequency of an artificial bistable switch (Gordon et al., 2009). This implies that bistable switches can be used to identify factors that alter processivity and/or fidelity of transcription in vivo. For instance, if the transcription of a gene encoding a nonabundant transcriptional regulator is more frequently paused, resulting in the production of less regulator protein, this would have a significant impact on the fraction of cells displaying the phenotype controlled by this regulator. Alternatively, if RNAP pauses less frequently and is more processive, more regulator protein will be produced.

A powerful way to control the level of noise in gene expression is through feedback mechanisms. For instance, it has been proposed and shown experimentally that negative feedback loops (autoregulation) provide robustness and stability in gene networks, thereby reducing noise (Becskei & Serrano, 2000). Using E. coli as a model organism, by measuring the coefficient of variance, these authors showed that the degree of variability for an autoregulated system is lower compared to an unregulated system (Becskei & Serrano, 2000). Conversely, positive feedback is known to increase noise (Kaern et al., 2003) and positive feedback (or double-negative feedback) is a main driving force for creating bistable switches, that is, the existence of two stable expression states (Ferrell, 2002; Smits et al., 2006; Ghosh, Banerjee, & Bose, 2012).

5 NOISE AND HETEROGENEITY IN GENE EXPRESSION

At first sight, it might appear counterintuitive that the regulatory systems of important bacterial phenotypes, such as sporulation, biofilm formation and even virulence, rely on stochasticity (Veening, Smits, & Kuijpers, 2008; Eldar & Elowitz, 2010). However, for microbes to respond to changes in their microenvironment, they need features like noise in gene expression and as such noisy gene expression driving certain traits might have been selected for during evolution as part of bet-hedging or division of labour.
strategies. To demonstrate the power of noise-regulated gene circuits, we take a more detailed look at the heterogeneous expression of certain virulence factors.

In order to invade a host, a pathogen needs to overcome many hurdles, like travelling from one part of the body to another through varyingly hostile environments. Resilience to these changing environments requires rapid adaptation. Pathogens use numerous strategies to keep one step ahead of their host, such as the exchange of DNA or a high mutation rate, enabling the pathogen to acquire advantageous traits. However, when a pathogen needs to adapt rapidly, mutation is not efficient enough and the pathogen needs a more sophisticated means of adaptation of the population, for example in the form of division of labour. Pathogens appear to be able to accomplish this by using noise-regulated or noisy gene circuits.

Because of the complex nature of a pathogen’s natural environment, from the many interspecies interactions to the number of changing environments a pathogen sometimes must go through, this is a relatively underresearched field. However, for an increasing number of pathogens, ranging from well-known bacteria like *E. coli* to the malaria parasite *P. falciparum*, the last few years have seen an increase in identification of noise-regulated gene circuits that play a role in virulence and pathogenesis (Table 6.1; Butala et al., 2012; Rovira-Graells et al., 2012).

In the following section, we discuss how two pathogenic bacteria employ heterogeneous gene expression when invading a host: *Salmonella enterica* serovar Typhimurium, (short: *Salmonella* Typhimurium) and *Streptococcus pneumoniae*.

### 6 BISTABLE EXPRESSION OF PNEUMOCOCCAL PILI

It is not well understood why, in most instances, *S. pneumoniae* lives as a harmless commensal organism, but that can suddenly turn into a dangerous pathogen. In fact *S. pneumoniae* is one of the most important human pathogens, responsible for the deaths of nearly 1 million children each year (O’Brien et al., 2009). The consensus is that invasion of a host by *S. pneumoniae* starts with nasopharynx colonization. Disease occurs when *S. pneumoniae* is able to travel to otherwise sterile parts of the body, such as the bloodstream, the inner ear or the lungs. Important virulence factors and pathogenicity islands are necessary for colonization of the host (Weiser, 2010).

One important virulence factor is the type 1 pilus, which *S. pneumoniae* can use for adherence to the surface of the upper respiratory tract during colonization. The reports about the role of pili in virulence are somewhat ambiguous. In murine models, pili have been reported to be important for virulence, while in humans, they do not appear to be associated with increased virulence (Basset et al., 2007). However, after the introduction in 2000 of a vaccine targeting the pilus, there was a rapid decrease of strains with pilus genes isolated in hospitals. At the present time, the percentage of strains able to form pilus is back to its old value, around 25%, suggesting that pili can be advantageous for colonization in humans (Regev-Yochay et al., 2010).
As an additional layer of complication, it has recently been shown that the genes encoding pili formation are heterogeneously expressed (Basset et al., 2011; De Angelis et al., 2011). Pili formation is encoded by the rlrA pathogenicity island, containing the regulator RlrA, three cell-anchored surface proteins and three sortases. The surface proteins, RrgA, B and C, contain C-terminal sorting terminals, suggesting that the sortases on the pathogenicity island could be used for this purpose. Not much is known about the regulation of the genes on this island, but it has been shown that the bistability of pilus expression can be altered by changing the expression of RlrA (Basset et al., 2012). RlrA activates expression of the whole pathogenicity island, including itself, but is repressed by one of the surface proteins.

### Table 6.1 Noise-Regulated/Heterogeneously Expressed (Black) and Possible Noise-Regulated (Grey) Virulence Factors and Traits Contributing to Pathogenicity in Bacteria and Other Pathogens

<table>
<thead>
<tr>
<th>Organism</th>
<th>System</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Escherichia coli</td>
<td>Production of bacteriotoxin colisin</td>
<td>Butala et al. (2012)</td>
</tr>
<tr>
<td>Mycobacterium tuberculosis</td>
<td>Persistence and antibiotic resistance</td>
<td>Wakamoto et al. (2013)</td>
</tr>
<tr>
<td>Pseudomonas aeruginosa</td>
<td>Responsiveness to quorum-sensing signals</td>
<td>Kohler, Buckling, and van Delden (2009)</td>
</tr>
<tr>
<td></td>
<td>Metabolic state and antibiotic resistance in biofilms</td>
<td>Williamson et al. (2012)</td>
</tr>
<tr>
<td>Salmonella enterica serovar Typhimurium</td>
<td>Type 3 secretion system</td>
<td>Ackermann et al. (2008) and Diard et al. (2013)</td>
</tr>
<tr>
<td></td>
<td>Flagellar expression</td>
<td>Cummings, Barrett, Wilkerson, Fellnerova, and Cookson (2005) and Stewart and Cookson (2012)</td>
</tr>
<tr>
<td>Streptococcus pneumoniae</td>
<td>Type 1 pilus</td>
<td>Basset et al. (2012) and De Angelis et al. (2011)</td>
</tr>
<tr>
<td></td>
<td>Production of pneumolysin</td>
<td>Ogguniyi, Grabowicz, Briles, Cook, and Paton (2007)</td>
</tr>
<tr>
<td></td>
<td>Expression of capsule</td>
<td>Lysenko, Lijek, Brown, and Weiser (2010)</td>
</tr>
<tr>
<td>Vibrio cholerae</td>
<td>Expression of the toxin-coregulated pilus (TCP)</td>
<td>Nielsen et al. (2010)</td>
</tr>
<tr>
<td>Candida glabrata (fungi)</td>
<td>Production of adhesin Epa1</td>
<td>Halliwell, Smith, Muston, Holland, and Avery (2012)</td>
</tr>
<tr>
<td>Candida albicans (fungi)</td>
<td>Expression of transcription regulator Efg1</td>
<td>Pierce and Kumamoto (2012)</td>
</tr>
<tr>
<td>Plasmodium falciparum (protozoa)</td>
<td>Expression of genes involved in host–parasite interactions</td>
<td>Rovira-Graells et al. (2012)</td>
</tr>
</tbody>
</table>
**RrgB** (Figure 6.2A). Only one repressor outside the island has been identified up until now: MgrA (Hemsley, Joyce, Hava, Kawale, & Camilli, 2003). How bistability in pilus formation is established at the molecular level is currently not known, but it is tempting to speculate that noise in gene expression activates the RlrA autostimulatory loop in a stochastic manner and that the combined action of noise and positive feedback is essential in setting up the observed phenotypic variation. Pili on individual cells might be formed by pulses of gene expression in which the RlrA autostimulatory loop is responsible for rapid synthesis of the pilus and the MgrA repressor acts to dampen and switch off gene expression similar to competence development in *B. subtilis*, which is governed by such a noisy excitable bistable switch (Suel, Garcia-Ojalvo, Liberman, & Elowitz, 2006).

Recently, the pilus protein RgrB has been tested as a vaccine target in murine models, showing that this vaccine does protect against *S. pneumoniae* heterogeneously expressing pili. However, not much is known about the function of heterogeneous expression *in vivo*, nor is it known if there is any difference in expression patterns during the different stages of infection. One possible explanation for the effectiveness of the vaccine is that the pili are always expressed at a certain ratio and that this ratio is maintained actively, which, in immunized mice, eventually leads to the eradication of the pathogen. However, Moschioni and colleagues noted that it could also be the case that pilus expression is more highly expressed at early stages of infection (see Figure 6.2B) (Moschioni et al., 2012). This supports the hypothesis that pilus-expressing cells could initiate colonization by adhesion to the epithelium. When the host is primed against pilus-expressing cells, the colonization would be inhibited, even though the pilus-expressing phenotype would only be important at this very early stage (Figure 6.2C).

**FIGURE 6.2**

(A) Pathogenicity island for pili formation is mostly self-regulated. RlrA regulates itself and all six other genes (blue). RrgB inhibits RlrA at the protein level. MgrA, a regulator outside the pathogenicity island, is known to either directly or indirectly repress RlrA activity. (B) Hypothetical role for pili expression in early colonization. Pili-expressing cells (yellow, 1) can attach to the nasopharynx epithelium, stimulating biofilm formation (2). When there is a chance, nonexpressing cells (red) can invade other parts of the body. (C) When the host is primed with RrgB, macrophages recognize the pili-expressing cells, and *S. pneumoniae* will not be able to colonize the nasopharynx.
**7 COOPERATIVE VIRULENCE IN *SALMONELLA ENTERICA* S. TYPHIMURIUM**

*Salmonella* Typhimurium is the number one cause of food poisoning in Western countries, causing around one million cases of illnesses in the United States every year. The pathogen is shown to be remarkably adaptive, being able to invade a large range of host organisms, and, within the host, has to go through numerous different environments.

*Salmonella* Typhimurium invades the host through the Peyer’s patches, aggregations of lymphoid tissue in the lowest part of the small intestine. During invasion, *Salmonella* secretes flagellin through the type 3 secretion system (T3SS), which helps to outcompete the natural commensals living both in the Peyer’s patches and in the small intestine and evoke an inflammatory response. When the pathogen is taken up by phagocytes, it remains viable and is transported to systemic tissue. Interestingly, in murine models, *Salmonella* Typhimurium isolates show heterogeneous expression of the genes involved in flagella formation. The distribution of this heterogeneity is strikingly different depending on the location of the pathogen. In mouth infections, 100% of the *Salmonella* Typhimurium population is flagellated, but in the Peyer’s patches, only part of the population is (Stewart, Cummings, Johnson, Berezow, & Cookson, 2011). Finally, in systemic tissue, flagella formation is repressed.

Heterogeneity of flagellar expression is known to be regulated by the interplay between the flagella master regulator complex FlhD4C2, its antagonist YdiV and the regulator FliZ (Saini et al., 2010; Stewart et al., 2011; Wada, Tanabe, & Kutsukake, 2011; Moest & Meresse, 2013). A simplified representation of the regulatory circuit is shown in Figure 6.3A (green). FlhD4C2 activates the expression of the fliAZ operon encoding FliZ and sigma factor 28 (σ28 or FliA), which in turn regulates the downstream expression of flagellar genes. FliZ stimulates its own (and thereby σ28) expression and simultaneously represses expression of ydiV, the product of which blocks the function of FlhD4C2. Together, this results in a combination of a double-negative feedback and self-stimulation, which meets the characteristics of a noise-sensitive gene circuit (Smits et al., 2006).

While the importance of YdiV and FliZ for heterogeneity is extensively researched *in vitro*, Stewart and colleagues showed that a ydiV knockout resulted in a fully flagellated population *in vitro* and in systemic sites in mice (Stewart et al., 2011). In the spleen of mice models, however, ydiV mutants and wild-type cells are both unflagellated. Moreover, it has been shown that FlhD4C2 activity is repressed by degradation by the ClpXP protease, influencing heterogeneity as well (Cummings, Wilkerson, Bergsbaken, & Cookson, 2006; Kage, Takaya, Ohya, & Yamamoto, 2008). ClpXP-knockout mutants are hyperflagellated and, interestingly, their virulence is attenuated. Infection with ClpXP-knockouts protects mice from infection with fully virulent Salmonella Typhimurium (Cummings et al., 2006).

These findings, together with the identification of the unique expression pattern of the flagella in the spleen, lead to the hypothesis that flagella are important for the initiation of invasion but that an unflagellated population is needed for perseverance.
FIGURE 6.3
Role of heterogeneous gene expression during *Salmonella* Typhimurium invasion.
(A) Simplified representation of regulation of heterogeneous gene expression of the flagellar proteins (green) and type 3 secretion system (red). Coloured arrows represent gene regulation, black arrows protein–protein interaction. Dashed arrow: not defined whether direct or indirect interaction. The regulatory complex FlhD4C2 activates operon FliAZ containing sigma factor 28 and regulatory protein FliZ. FliZ determines heterogeneous expression of sigma 28 and thereby the flagella formation together with YdiV, which can bind to FldDC. FliZ is also one of the important activators of *hilD* expression. HilD activates *hilA* expression together with RtsA and HilC, which all activate each other and themselves (not depicted for clarity), but HilD is the only essential regulator and thought to influence heterogeneity. (B) Green panel: at early invasion, a flagellated subpopulation reaches the epithelium cells early and triggers the host’s immune response. The nonflagellated population can, in contrast to the flagellated population, survive in macrophages and are thereby transported to systemic tissue. (C) Red panel: the majority of T3SS-expressing cells (orange) invade the host epithelium. When a T3SS mutation (black) occurs, the faster-growing nonvirulent subpopulation and the mutants both benefit from the invasion, but a small fraction of the virulent subpopulation remains. However, when a mutation occurs in a 100% virulent population (D), the slow-growing cells will be outcompeted by the mutants who are unable to withstand the commensal population (blue).
Flagellated organisms have an advantage over unflagellated ones to reach the Peyer’s patches through the GI tract. Only a few *Salmonella Typhimurium* cells are sufficient for invading the Peyer’s patches, where they reproduce and form a heterogeneous population. Once there, the flagellated population triggers the immune response. The unflagellated population is taken up by macrophages just as well as the flagellated population, but it will not be recognized and can thereby travel to systemic sites.

This cooperation model of the role of flagella during a *Salmonella Typhimurium* invasion cannot be reviewed without looking at the formation of the T3SS, the genes for which are also expressed heterogeneously. Even though the regulation of both systems is connected by two important regulators (FliZ and HilD; see Figure 6.3A), heterogeneity of T3SS expression can be regulated separately from flagellar expression. This provokes the thought that several phenotypes of *Salmonella Typhimurium* might exist during invasion.

The molecular mechanism of T3SS regulation and formation is not fully understood, but expression is controlled by the regulators HilD, RtsA and HilC, which activate their own expression and that of each other, and, importantly that of HilA, the key regulator of the pathogenicity island 1 on which the T3SS system is located. However, HilD is the only essential regulator and is important for heterogeneity in T3SS expression. Its gene and activity is regulated by many proteins, of which FliZ is one of the most important by promoting HilD activity at the protein level (Chubiz, Golubeva, Lin, Miller, & Slauch, 2010).

How the phenotypes resulting from these different gene circuits interact is not known. Stewart and Cookson proposed a model where three distinct populations could be found in a host organism (Stewart & Cookson, 2012). Where the flagella and T3SS would mostly be expressed together, they see a distinct role for flagella-only expression in early invasion, when motility is an important trait for colonization but expression of T3SS evokes inflammatory responses.

**CONCLUDING REMARKS**

The use of noise-regulated gene circuits and bistable gene expression is widespread among bacteria and appears to be important for pathogenesis: a growing number of reports show a range of pathogenic organisms utilize noisy gene circuits and/or show bistable gene expression (Table 6.1). Importantly, noise-driven phenotypic variation can set up a small subpopulation of cells, which are already prepared for a specific change in the environment in the future such as the presence of a new host. Most of the time, these noise-driven phenotypes can also be activated in nonactivated cells by the ‘standard’ way, for instance, via sensing mechanisms. These strategies are mostly related to fast adaptation to the different environments and threats inside the host. While mutation, as a means of adaptation, is rigorous and unregulated, noise-mediated gene expression is both fast and reversible.
References


