Gene position within a long transcript as a determinant for stochastic switching in bacteria

Jan-Willem Veening* and Oscar P. Kuipers**
Molecular Genetics Group, Groningen Biomolecular Sciences and Biotechnology Institute, Kluiver Centre for Genomics of Industrial Fermentation, University of Groningen, Haren, the Netherlands.

Summary

How cultures of genetically identical cells bifurcate into distinct phenotypic subpopulations under uniform growth conditions is an important question in developmental biology of relevance even to relatively simple developmental systems, such as spore formation in bacteria. A growing Bacillus subtilis culture consists of either cells that are motile and can swim or cells that are non-motile and are chained together. In this issue of Molecular Microbiology, Cozy and Kearns show that the probability of a cell to become motile depends on the position of the sigD gene within the long (27 kb) motility operon. sigD encodes the alternative sigma factor σ^D that, together with RNA polymerase, drives expression of genes required for cell separation and the assembly of flagella. sigD is the penultimate gene of the B. subtilis motility operon and, in the control strain approximately, 70% of the cells are motile. When sigD was moved upstream within the operon, a larger fraction of cells became motile (up to 100%). This study highlights that the position of a gene within an operon can have a large impact on the control of gene expression. Furthermore, it suggests that RNA polymerase processivity or mRNA turnover can play important roles as sources of noise in bacterial development, and that gene position might be an unrecognized and possibly widespread mechanism to regulate phenotypic variation.

With the use of single cell techniques, more and more examples of non-genetic phenotypic variability within clonal bacterial populations are being unveiled (Veening et al., 2008). While most of this cell-to-cell variability was previously masked by bulk assays, it is now recognized that many bacteria employ epigenetic mechanisms to bifurcate into phenotypically different subpopulations. A good model to study heterogeneous development is the Gram-positive Bacillus subtilis. This soil bacterium can display a number of phenotypes including competence for DNA uptake, spore development, protease or antimicrobial compound production and swimming motility (Veening et al., 2008). Interestingly, within genetically identical B. subtilis populations, grown under identical conditions, only part of the population exhibits one of these phenotypes. For instance, only part of an isogenic B. subtilis culture forms highly resistant endospores (Errington, 1993; Chung et al., 1994).

One of the best-understood heterogeneous pathways at the molecular level is competence development in B. subtilis. The entry into the competent state, or K-state, depends on a noise-driven switch involving the master regulator ComK (Suel et al., 2006; 2007; Maamar et al., 2007). Once a threshold level of ComK is reached inside a single cell, a positive feedback loop kicks in and more ComK is produced (Maamar and Dubnau, 2005; Smits et al., 2005). This subsequently leads to activation of genes whose products are required for the uptake and integration of exogenous DNA into the chromosome. After a transient period, cells exit from the K-state and resume normal vegetative growth or go on to engage in another developmental pathway such as sporulation (Leisner et al., 2007; Cagatay et al., 2009). Unlike sigD, comK is not present within an operon structure but is present as a mono-cistron. However, comK is regulated by at least five transcriptional regulators. Furthermore, ComK turnover is highly controlled by a quaternary complex regulating ComK proteolysis via the Clp protease system (Leisner et al., 2008). Importantly, it was shown that stochastic fluctuations in the levels of comK mRNA set the probability of becoming competent (Maamar et al., 2007). Thus, a number of
non-exclusive mechanisms can generate phenotypic heterogeneity.

In the current issue of *Molecular Microbiology*, Cozy and Kearns show that the probability for *B. subtilis* cells to produce flagella and swim in liquid cultures is largely determined by the position of the *sigD* gene within the long (27 kb) *fla/che* motility operon (Cozy and Kearns, 2010). The operons under control of *σ^D*-containing RNA polymerase include genes encoding cell-wall remodeling enzymes (so-called autolysins) required for cell separation after cell division and include genes required for the production of functional flagella, such as the gene encoding the flagellin filament protein (Hag). Thus, cells that are ON for *σ^D* separate, produce flagella and become motile whereas cells OFF for *σ^D* remain attached to their siblings and form chains, causing them to be non-motile (Fig. 1A). It was previously established that cells that are motile contain active *σ^D* protein, as motile cells produce Hag, as determined using GFP fusions (Kearns and Losick, 2005). In the current study, Cozy and Kearns physically separated chained cells from motile cells by centrifugation and showed that the chained cells indeed do not express *sigD* and lack *σ^D* protein. Thus, the bifurcation into swimming and chained cells seems to depend largely on the levels of *σ^D* protein in individual cells. Whether other factors besides *σ^D* protein levels are also important in the motility decision still remains unknown.

**Fig. 1.** Heterogeneity in *Bacillus subtilis* motility.

A. Cells with high levels of *σ^D*-protein produce autolysins and flagella and swim freely in liquid medium (blue cells). Cells with low levels of *σ^D* protein remain chained and do not swim (white cells).

B. Heterogeneity within *B. subtilis* biofilms. Part of the population contributes to the production of resources required for biofilm formation (green and yellow cells) and part of the population has preassembled flagella and might be ON for *σ^D* (green and blue cells). Whether the green cells are really ON for *σ^D* and exist within dense biofilms, or whether they represent an intermediate state between swarming and biofilm-forming cells with pre-assembled flagella requires experimental verification. Only the blue cells that do not express the biofilm protein EpsE (the molecular clutch for flagella) but are ON for *σ^D* can swim freely.
Interestingly, the authors found that the probability of switching to one state or the other is influenced by the genetic location of the $\text{sigD}$ gene within the long motility operon. They observed a decrease in transcript abundance when the gene location was closer to the end of the operon. This means that genes at the 3' end of the 27 kb long operon, such as $\text{sigD}$, are expressed to a lesser extent than those located at the 5' end. This was verified by both in vivo and in vitro methods. Fluorescence microscopy demonstrated that the 5' end of the motility operon was uniformly expressed in all cells (both swimming and chained cells), while the 3' end was only expressed in motile single cells. This was verified by again separating chained cells from free swimming cells using centrifugation and analysing the abundance of a set of genes within the motility operon using quantitative reverse transcriptase PCR. The molecular origin(s) for reduced abundance of the 3' end of the transcript of the large motility operon still remains unknown, but possible sources include reduced RNA polymerase processivity, weak terminator structures within the operon or specific 3' mRNA degradation.

While RNA processing is a common process in eukaryotes, the (specific) degradation of RNA in prokaryotes is only recently receiving more attention and it is now established that RNA degradation can play an important role in the control of gene expression in bacteria (condon, 2007). Future research should identify whether the motility operon transcript is specifically degraded at its 3' end or whether other processes are at play. In any case, gene location within a (long) operon might be another source of stochastic gene expression and should be taken into account when studying the origins of phenotypic variation. For instance, other long or complex operons (e.g. the $\text{eps}$, $\text{sigB}$, $\text{srfA}$ and $\text{resD}$ operons of $\text{B. subtilis}$), containing regulator genes and showing heterogeneous expression would be worth inspecting in the same way to determine whether the location effect is more general.

Another interesting observation made by Cozy and Kearns is that the physiological state of the cell (swim or chained) seems to be epigenetically inherited. Thus, if a cell switched to become a swimmer, the chance is high that its progeny will also be swimmers. This was visualized by using a strain that lacks the major autolysins and is predominantly present in the chained form. In this mutant background, sets of $\sigma^D$-ON cells could be found next to $\sigma^D$-OFF cells. This indicates that there must be some sort of positive feedback mechanism within the system to ‘memorize’ the physiological state. In fact, the authors did identify a new $\sigma^D$-dependent promoter within the long motility operon that might be involved in this mechanism. Thus, once a cell reaches a certain threshold level of $\sigma^D$-protein, the positive feedback loop kicks in and even more $\sigma^D$-protein is made, locking the cell in its ON state.

What could be the physiological relevance for $\text{B. subtilis}$ to diversify into chained and swimming cells? To become motile, more than 30 different proteins need to be properly produced and assembled to produce a functional flagellum (Chevance and Hughes, 2008), which is a time- and energy-consuming process. Thus, upon the stimulus to become motile, it would take a cell at least an hour before it can respond. Heterogeneity in the population ensures that part of the population is already equipped with pre-assembled flagella so that at least part of the clonal population can rapidly respond to take advantage of cell motility.

In nature, most bacteria, including $\text{B. subtilis}$, can be found in dense multilayered biofilms. What could be the role of $\sigma^D$ heterogeneity within biofilms? Biofilms add another level of complexity, as only part of the $\text{B. subtilis}$ population contributes to the formation of the biofilm (Chai et al., 2007; Vlamakis et al., 2008). For instance, Vlamakis et al. showed that motile cells (Hag ON) can switch to biofilm-producing cells but do not express the yqxM biofilm gene simultaneously with hag (Vlamakis et al., 2008). Thus, only part of the population is ON for $\sigma^D$ and only part of the population contributes to biofilm formation (Fig. 1B). Interestingly, the Kearns lab also recently found that one of the genes that is activated in biofilm formers encodes EpsE, a protein that acts as a molecular clutch to inhibit rotation of flagella (Blair et al., 2008). However, this was shown in swarming cultures of $\text{B. subtilis}$ cells and not within dense biofilms. Thus, cells that are presumably ON for $\sigma^D$ are not able to swim away from the microbial community when they are also contributing to forming the biofilm (Fig. 1B). Whether these types of cells ($\sigma^D$ ON, biofilm ON) actually exist within dense biofilms awaits further experimentation. EpsE might be particularly important to inhibit motility of pre-assembled flagella in cells that have been long turned OFF for $\sigma^D$ but still contain functional (durable) flagella. In any case, this type of regulation ensures that the structural integrity of the biofilm remains intact and helps to stabilize the biofilm. When cells have finished their role in the biofilm and already have fully formed flagella, they can rapidly respond and swim away. Alternatively, when cells have finished swimming and swimming and found a decent new niche, they can rapidly switch to the biofilm-forming mode without disrupting attachment to the surface by motion caused by rotating flagella. The fact that more than two different phenotypes can coexist creates an enormous potential for phenotypic variation in an isogenic population (Veening et al., 2008), enabling the population to fine tune responses under highly variable conditions in a competitive environment.
Acknowledgements

We thank Ákos Kovács for useful suggestions on the manuscript. J. W. V. was supported by startup funds from the University of Groningen and by a Marie-Curie Reintegration grant. O. P. K. is supported by the research programme of the Kuyver Centre for Genomics of Industrial Fermentation, which is part of the Netherlands Genomics Initiative/Netherlands Organization for Scientific Research.

References


