Identification of a new regulatory pathway for the K-State in bacillus subtilis

Mathieu Miras

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)

Présentée et soutenue par :
Mathieu MIRAS

Le 28 Avril 2017

Titre :

IDENTIFICATON OF A NEW REGULATORY PATHWAY FOR THE K-STATE IN BACILLUS SUBTILIS

ED BSB : Microbiologie

Unité de recherche :
Laboratoire de Microbiologie et Génétique Moléculaires - UMR 5100

Directeur(s) de Thèse :

Patrice POLARD

Rapporteurs :

Melanie BLOKESCH	Associate Professor EPFL	Rapporteur
Mireille ANSALDI	Directeur de recherche CNRS	Rapporteur

Autre(s) membre(s) du jury :

Patrice POLARD	Directeur de recherche CNRS	Directeur de thèse
Claude GUTIERRez	Professeur UT3 Paul Sabatier	Président du jury
David DUBNAU	Principal Investigator PHRI	Examinateur

TABLE OF CONTENTS

RESUME

Bacillus subtilis, une bactérie Gram-positive présente dans le sol, peut lorsque les nutriments sont en concentrations limitantes, sporuler, former des biofilms ou devenir compétente. La compétence est, chez *B. subtilis*, caractérisée par un arrêt de la division cellulaire, une tolérance aux antibiotiques et l'expression de plus d'une centaine de gènes. L'expression de la compétence, aussi désignée sous le nom de « K-state », est dépendante de la synthèse du facteur de transcription ComK et se fait de façon stochastique résultant en la formation de deux sous-populations bactériennes, non-compétentes et compétentes. L'émergence, à partir de cellules génétiquement identiques, de deux sous-populations distinctes est une stratégie de survie très répandue chez les procaryotes, connue sous le nom de « bet-hedging ».

Bien que les mécanismes de régulation du développement de la compétence ont, chez *B. subtilis*, largement été étudiés au cours des dernières années, la raison pour laquelle les souches non-domestiques sont très peu transformables (1-2% de la population) comparé aux souches domestiques (~15%) reste méconnue. Nous démontrons ici que c'est essentiellement dû à une mutation de transition dans le promoteur du gène *degQ*. Cette mutation diminue la synthèse de DegQ, protéine impliquée dans la régulation de la formation de biofilms, de la synthèse d'exoprotéases et de la transformation génétique. DegQ est une protéine impliquée dans le transfert d'un groupe phosphoryl entre la kinase DegS et son substrat DegU. Une faible quantité de DegQ diminue la concentration en DegU~P ce qui a pour conséquence la désinhibition de l'opéron *srfA* entrainant une accumulation de ComK et l'expression de la compétence. C'est ainsi que, dans les souches domestiques de *B. subtilis*, un plus grand nombre de bactéries atteignent le niveau nécessaire en ComK pour activer une boucle d'auto-activation transcriptionnelle de *comK*. Nous

démontrons aussi que l'activation transcriptionnelle de *srfA* est, dans les souches non-domestiques, transitoire alors que la population bactérienne entre en phase stationnaire de croissance. Ces données indiquent que le développement de la compétence est moins fréquent et plus transitoire dans les souches non-domestiques de *B. subtilis*. De plus, cette limitation du K-state dans les souches non-domestiques est plus importante que précédemment « pensé » probablement dû à la domestication de *B. subtilis* au cours de ces 50 dernières années.

Ce travail reflète non seulement, l'importance de l'utilisation de souches non-domestiques dans la caractérisation des voies de régulation de la compétence chez *B. subtilis*, mais aussi la portée du choix de modèle biologique dans l'étude de phénomènes biologiques complexes.

ABSTRACT

Bacillus subtilis, a Gram-positive soil bacterium, can enter into several developmental pathways such as sporulation, biofilm formation and competence development for DNA transformation when it becomes limited for essential nutrients. During competence, cells do not divide, are tolerant to antibiotics and competent cells express more than a 100 genes. The competent state has been named the K-state after its master regulator ComK.

In *B. subtilis*, the entry into the K-state is stochastically determined by the activation of the transcription factor ComK and occurs, in the domesticated strains of *B. subtilis*, in approximately 15% of the population. The emergence from genetically identical cells of two distinct subpopulations (competent cells and non-competent cells) is known to be a classic survival strategy for bacteria, known as bet-hedging. Regulation of entry into the K-state has been intensively studied and is well understood; however, the reasons why undomesticated isolates of *B. subtilis* are poorly transformable compared to the domesticated strains remained unexplained. We show here that fewer cells enter the K-state, suggesting that some regulatory pathway limiting its expression has been lost in the domesticated backgrounds. We demonstrate that this is largely due to an inactivating point mutation in the *degQ* promoter region resulting in a decrease of the amount of DegQ. DegQ is known to stimulate phosphate transfer from the DegS autokinase to its cognate response regulator DegU. A low level of DegQ thus decreases the concentration of the phosphorylated form of DegU, leading to the de-repression of the *srfA* operon, which increases the amount of ComS leading to the stabilization of ComK. Thus, in domesticated strains of *B. subtilis*, more cells reach the concentration threshold of ComK needed to activate the positive auto-regulatory loop of ComK acting on its own

promoter. We also show that the activation of *srfA* transcription in undomesticated strains is transient, as it is turned off when cells enter the stationary phase.

Taken together, these data indicate that the K-state and transformability are less frequent and more transient in the undomesticated strains of *B. subtilis*. Consideration of the regulatory mechanisms and the fitness advantages and costs of the K-state must from now on take these features into consideration. These results underscore that our understanding of real-life biology requires the use of wild isolates.

ACKNOWLEDGEMENTS

It has been five years since I began working towards my PhD. During these years, I have received support from colleagues, friends and family and I would like to dedicate this work to them.

To my mentor, David Dubnau, you invited me to be a part of your laboratory and I could not imagine doing my thesis work with anyone else. Thank-you for guiding me through the PhD's rough waters, for giving me the freedom and encouragements. Thank-you for your wisdom and your contagious (but precious) optimism: I think I will never forget the sentence "This is gonna work!!". There is no doubt you helped me become a better scientist. Thank-you!

To my thesis committee. Claude Gutierrez, Mireille Ansaldi and Melanie Blokesch, thank-you for being a member of this committee, for coming to Toulouse and for reading this work. Thank-you to Patrice Polard for allowing me to go to Dave's lab by giving me those precious signatures and for guiding me during the last few months of my thesis.

To the "Daffettes". Hedia, Ségolène, Nawel, Fabienne and Annaïk, thank-you for training me as a student and for coming to my defense after all these years. You showed me that good science could be done in a happy and friendly environment and that is something I have been trying to reproduce every day since then. Annaïk, thank-you for recommending me to Dave. I would definitely not be "here" without you!

This work would not have been possible without the excellent undergraduate education I received at the Université Paul Sabatier. I thank the professors and teachers in Departments of Biology for their instruction.

To the Dubnau Lab, past and present: During these five years, you guys have become like family. Thank-you for "enjoying" my eclectic music playlists, for many fun and encouraging conversations during lunch over these years, for enjoying all these sports events (mainly soccer) from the lab and of course, for sharing your stock solutions so kindly #TeamSpirit. Jeanie, thank-you for opening the doors of your house on several occasions, for always sharing your political and scientific opinions with passion… Please, don't give up the fight, I am sure we can "Make this world and our planet great again". Jeanette, thank-you for being my lab "mom", and for making fun of my accent. I really have enjoyed working by your side and have appreciated our conversations and your advices. Val, thank-you for always being helpful, for your insights about this manuscript, and for taking me to my first hockey game. Christine and Micaela, thank-you for bringing new accents in the lab, for these treats brought back from Germany and Italy, for your laughs and your friendship. Miguel and Nico, thank-you for helping me during my first months in America, not only as a PhD student, or for these passionate soccer conversations but also for showing me how to fill a tax return form…

To the PHRI/Rutgers kitchen staff: for making all these plates so I could dedicate my time showing (again and again) Jeanette how to use the Geneious software…

To all, I will miss having you around, and wish you all the best.

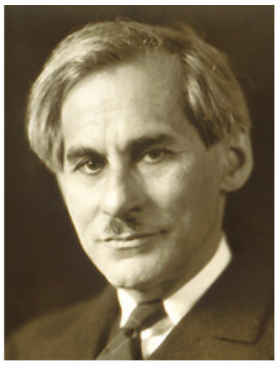
To my friends. Manu, Marina, Beun, Marion, Pierre, JF, Paula, Ced, Theresa, Anne-So, Fabe, Romain, Mathieu, Pablo, thank-you for always being around despite the distance and for making me feel like I never left when I got the chance to see you. To my college mates: Cindy, Saad, Thomas and Thibault. I will not forget all these times we could not stop laughing and of course ALL these "Belote games' we played anytime we got a chance. To friends from the New World: Thomas, JL, JP, Stéphane, Benoit, Iuri, Julien, Romain, Jessie, Charles, Gaétan, Soufiane it's been a great pleasure meeting you, playing soccer and spending time with you. You guys made this American adventure unforgettable.

To my family, and more particularly to my parents. To become a scientist, I think you need to be, over all, curious. And since I am young, you have always tried to open my eyes and mind to people, to new things, to new languages and cultures. I think I would have never become a scientist without all these travels and trivial cards games. Thank-you for always supporting me, for taking time to skype me when it was late at night and for sending me these precious packages filled with surprises, books and treats, especially dark chocolate without which I could have not survived to these cold winters.

INTRODUCTION

A brief history of DNA

The DNA story began in the mid-nineteenth century with the work of a young Swiss physician, Friedrich Miescher. After graduation, he went to Germany to work in one of the first laboratories focusing on the composition of lymphoid cells. As it was difficult to extract material from the lymph glands, Miescher had the idea to accumulate bandages from a nearby clinic to collect the pus. During his experiments with this material, he identified a substance with unexpected properties that did not match those of proteins; this substance was made of oxygen, hydrogen, nitrogen and phosphorus but there was a unique ratio of phosphorus to nitrogen. Since it had been purified from the nucleus of the cells, Miescher named this substance "nuclein," which ended up being partially purified deoxyribonucleic acid (DNA) (Dahm, 2008).



Born in 1869 in Lithuania, Phoebus Levene migrated to the U.S. to practice medicine in the Lower East Side of Manhattan a few years before World War II. As he was interested in the chemical structure of sugars, Levene further characterized Miescher's "nuclein" by showing it was made of a nitrogenous base, a sugar, and a phosphate group and that the different units were connected to each other via a phosphate group (Levene, 1919).

In 1928, the British bacteriologist Frederick Griffith was working on the epidemiology and pathology of *Streptococcus pneumoniae*, the pathogen responsible for pneumonia, and showed in what is now known as "Griffith's Experiment" the first evidence of bacterial transformation. Indeed, he realized he could transfer the virulence trait from a heat-inactivated (dead) virulent strain of *S. pneumoniae* to a live but non-virulent strain of *S. pneumoniae*, simply by mixing both live and dead strains. Unaware of the precise substance involved in the transfer of this new virulence phenotype acquired by the non-pathogenic strain, Griffith named this mysterious agent the "transforming principle" (Griffith, 1928) (Figure 1).

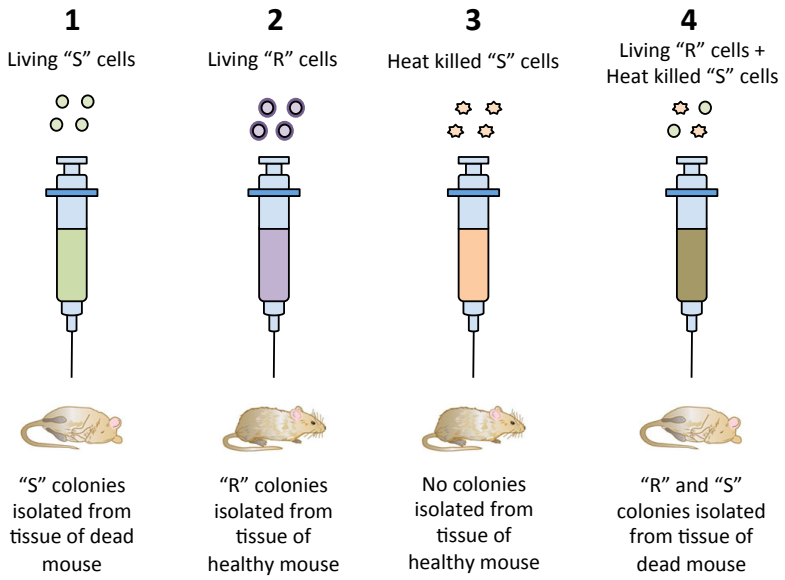
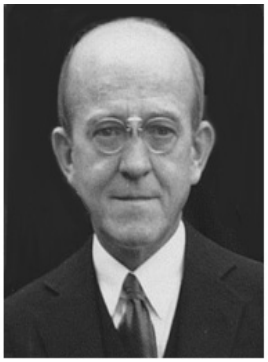


Figure 1: The Griffith Experiment
1- The mice infected by a virulent "S" strain of *S. pneumoniae* are killed. **2 and 3-** The mice infected by either non-virulent "R" strain or heat-killed virulent "S" strain of *S. pneumoniae* survive. **4-** After co-infection by "R" and "S" killed by heat, the mice die. Griffith was then able to isolate both live "R" and live "S" strains of *S. pneumoniae* from the blood of these dead mice. He concluded that the "R" strain had been "transformed" into the lethal "S" strain by a "transforming principle".



In the mid-nineties, the scientific community wondered about a big mystery: in which molecule is the genetic information hidden? Son of a Baptist minister at the Mariners' Temple In New York's Lower East Side, Oswald Theodore Avery, a Canadian doctor, repeated Griffith's original experiment using purified DNA from a virulent strain of *S. pneumoniae* and showed that it was sufficient to transfer the virulence trait to a non-virulent strain (Avery, Macleod, & McCarty, 1944) (Hershey, 1952). Avery was the first to propose that Griffith's transforming principle was actually DNA.

DNA acquisition in prokaryotes

Bacteria can acquire new traits either by mutation or by the acquisition of external genetic material. This external acquisition of DNA is termed horizontal gene transfer (HGT) (de la Cruz & Davies, 2000) (Ochman, Lawrence, & Groisman, 2000), referring to the transfer of genes between organisms in a manner that contrasts with vertical transfer, the inheritance of genes from the parental generation to offspring via sexual or asexual reproduction. HGT has been shown to be an important factor in the evolution of many organisms and the primary reason for the spread of bacterial antibiotic resistance (Koonin, Makarova, & Aravind, 2001). HGT occurs through three mechanisms: conjugation, transduction and transformation (Koonin et al., 2001).

Conjugation

Conjugation is the transfer of DNA from a donor to a recipient by direct physical contact between two cells (Clark & Adelberg, 1962). Many but not all species of bacteria can conjugate and conjugation is possible between cells of the

same species or even between cells of two different species (Trieu-Cuot, Gerbaud, Lambert, & Courvalin, 1985) (Penalva, Moya, Dopazo, & Ramon, 1990) (Kuhsel, Strickland, & Palmer, 1990). A plasmid called the F factor (Fertility factor) is required for conjugation (Lederberg, Cavalli, & Lederberg, 1952). As it carries its own origin of replication and an origin of tranfert, the F plasmid is an episome. It also carries the *tra* locus that incodes for pilin proteins. Only one copy of the F plasmid can be found in a given bacterium. In bacteria there are two "mating types," a donor (or F^+) and a recipient (or F-) and the direction of transfer of genetic material is unidirectional (Lederberg et al., 1952). After the pilus of the donor cell recognizes and binds to specific receptors sites on the cell wall of the recipient cell, a single stranded DNA molecule from the mobile plasmid is transferred from the donor to the recipient and is later converted to double stranded DNA (Wozniak & Waldor, 2010) (Figure 2).

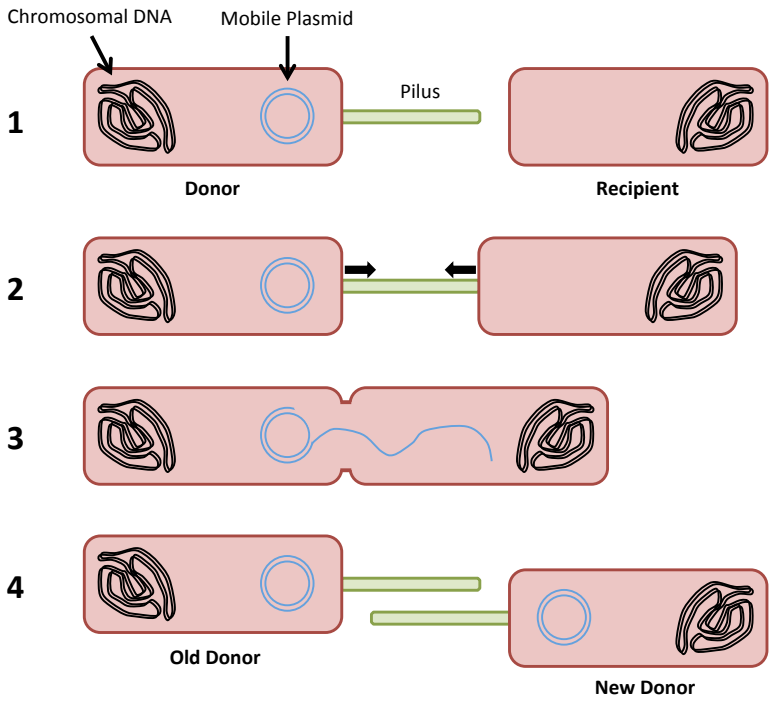


Figure 2: Schematic drawing of bacterial conjugation
1- The donor cell produces a pilus, which is encoded on the mobile plasmid. **2-** The pilus attaches to the recipient cell and brings the two cells in close proximity. **3-** The plasmid is nicked and a single strand of DNA is then transferred to the recipient cell. **4-** After genetic transfer, both cells synthesize the complementary strands to produce a double stranded circular plasmid. Both cells are now donors.

Transduction

Transduction is the transfer of genetic information from a donor to a recipient by means of a bacteriophage and can happen through either the lytic cycle or the lysogenic cycle (Zinder & Lederberg, 1952) (Kresge, Simoni, & Hill, 2011). In the lysogenic cycle, the bacteriophage's genome is integrated in the host genome to form the prophage and will not be expressed. The prophage is then transmitted as well as the rest of the genetic information to the daughter cells (called lysogens) as the bacterium divides. The switch to the lytic cycle can be induced at any time and the prophage DNA is excised from the bacterial chromosome, transcribed, and translated to produce phage elements. The prophage DNA will then be incorporated in the phage particles, which are released by lysis of the host. Because the packaging of prophage DNA in the phage heads is a low fidelity process, it happens that small pieces of the

host chromosomal DNA may be incorporated as well. The phage coat protects its genome in the environment so that transduction, unlike transformation, is not affected by extracellular nucleases. We can assume that, because this mode of gene transfer does not depend on many dedicated bacterial genes, phages played an important role in the ecology and evolution of bacteria. Indeed, by moving pieces of bacterial DNA among themselves, phages probably contributed in making bacteria what bacteria are nowadays (Spizizen, Reilly, & Evans, 1966) (Figure 3).

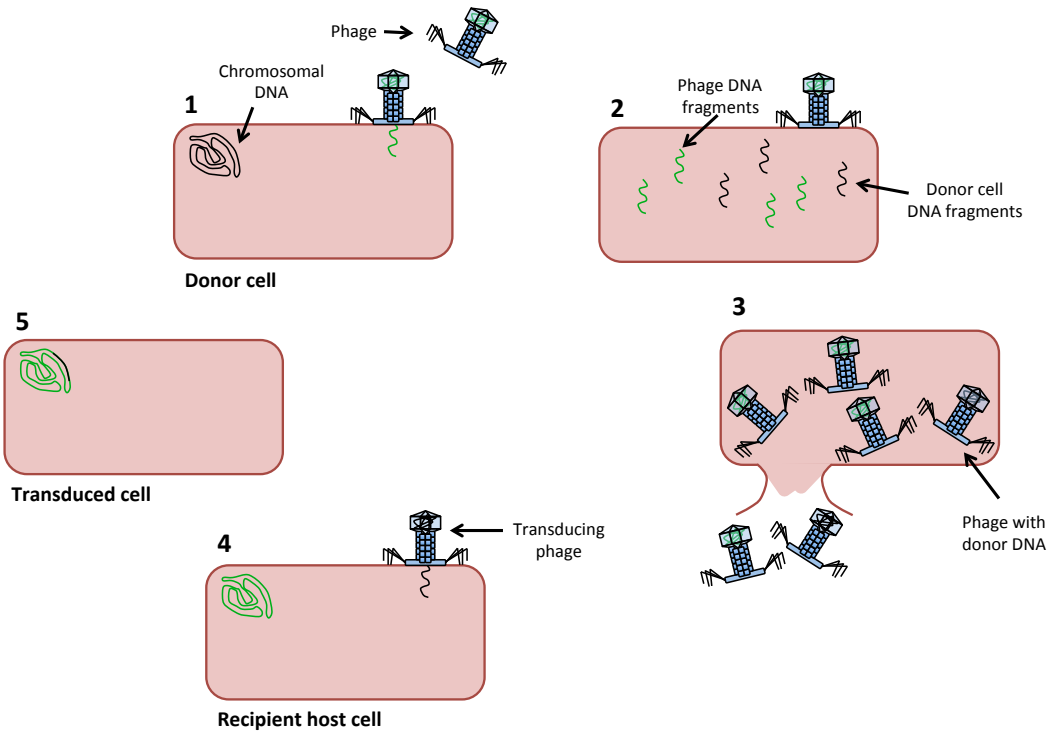


Figure 3: Schematic drawing of bacterial transduction.
1- Phage infects a susceptible bacterial cell. **2-** The host bacterial DNA is hydrolyzed while phage DNA and proteins are produced. **3-** New viral particles are synthesized and will occasionally incorporate phage DNA into the mature virion, instead of phage DNA. The cell lyses and releases the new bacteriophages. **4-** Transducing phages infect new cell (recipient host) and transfer bacterial DNA, but are defective as lytic phages. **5-** New DNA is incorporated into recipient's genome by recombination.

Transformation

Transformation is the gene transfer resulting from the uptake of exogenous (environmental) DNA. Certain bacteria species (e.g. *Bacillus*, *Haemophilus, Neisseria*, *Pneumococcus*) can naturally incorporate DNA into their genomes. Like conjugation systems, transformation systems depend on specialized operons that encode DNA uptake machinery. Proteins encoded by these genes include those necessary for DNA binding, internalization, and possible recombination of DNA with the genome. This process will be discussed in great detail (see "Competence for transformation and the K-state").

Transformation is a conserved phenomenon among bacteria

Cells expressing genes needed for acquiring DNA from the environment are said to be in a state of competence. Since this state is mediated by genes present in the bacterial genome, and is not induced artificially, this form of competence is referred to as genetically-programmed (Erickson, 1970) (Chen, Christie, & Dubnau, 2005). In contrast, other bacteria such as *Escherichia coli* are not able to take up DNA naturally, but the competent state can be induced through the addition of either chemical or physical agents (electroshock), which permit DNA to pass through the cell wall. Most naturally competent bacteria only take up DNA at a specific time during growth. It has been argued that the benefits of genetic competence are that the acquired DNA can be used as a source of nutrition, to repair existing genes, or for the acquisition of new genetic material (Redfield, 1993). Examples of natural bacterial competence are represented in both Gram-negative (i.e. *Neisseria gonorrhoeae*, *Haemophilus influenzae*, *Helicobacter pylori*) and Gram-positive (i.e. *Streptococcus pneumonia*, *Bacillus subtilis*) species (Chen et al., 2005). While in Gram-positive bacteria, DNA must go through the thick peptidoglycan layer of the cell wall and the

cytoplasmic membrane; it must also traverse the highly impermeable outer membrane of Gram-negative bacteria. Therefore, additional steps are involved in the Gram-negative transformation systems, and the initial interaction of DNA with the cell envelope is also different in the two types of bacteria. Despite these differences, the protein machinery required to transport DNA in Gram-positive and Gram-negative bacteria is conserved, and interestingly, is also closely related to protein machines used for molecular secretion (Burton & Dubnau, 2010).

Bacillus subtilis

General introduction

Bacillus subtilis (*B. subtilis*) is a bacterium naturally found in soil and vegetation that belongs to the *Firmicute* phylum of bacteria, which includes Gram-positive bacteria with a low G+C content. Firmicutes have been differentiated from one another based on different characteristics such as the nature of their cell envelope, their aerotolerance (how well they live and grow in oxygen) and their ability to form endospores. Consequently there are seven classes of Firmicutes: the *Erysipelotrichia*, the *Negativicutes*, the *Limnochordia*, the *Tissierellia*, the *Thermolithobacteria*, the *Clostridia* and the *Bacilli*. Because stress and starvation are common in its environment, *B. subtilis* has evolved a set of strategies that allow its survival under these harsh conditions: formation of stress resistant endospores or uptake of external DNA, which allows the bacteria to adapt by recombination. And partly because *B. subtilis* is also a non-pathogenic bacterium for humans it has been used as a model organism to further understand pathogenic microorganisms belonging to the same phylum such as *Streptococcus pneumoniae* (a major cause of pneumonia), *Bacillus anthracis* (agent of anthrax) or *Listeria monocytogenes* (food infection). *B. subtilis* is used in industry to produce antibiotics (subtilin, bacitracin) and secrete several commercial enzymes used in the food industry (amylases for bread production) or the detergent industry (proteases). Also, a strain of *B. subtilis* formerly known as *Bacillus natto* is typically used in Japan to produce a treat called “natto”. This traditional course is made of soybeans fermented with *Bacillus natto* and can be eaten for breakfast.

A short history of *Bacillus subtilis*

B. subtilis was originally discovered more than a hundred and fifty years ago and the domesticated laboratory strains have now been used for more than half a century (Zeigler et al., 2008). *B. subtilis* laboratory strains derive from a single tryptophan-requiring auxotrophic strain, strain 168. In the mid 1900s, two Yale University botanists, Paul Burkholder and Norman Giles, isolated strain 168 after the *B. subtilis* Marburg strain was mutagenized with X-rays (Burkholder & Giles, 1947). Unfortunately the Yale group abandoned its *B. subtilis* experiments to focus on other research interests, and most of its *B. subtilis* collection has been lost. At least five mutants (auxotrophs requiring threonine (strain 23), nicotinic acid (strain 122), or tryptophan (strains 160, 166 and 168)) were preserved from those original experiments, and were transferred to Charles Yanofsky and John Spizizen, who showed that three of them (strains 122, 166 and 168) could be transformed to prototrophy when exposed to DNA from strain 23 (Spizizen, 1958). The highly transformable strain 168 became the subject of many other studies and was therefore disseminated around the world. By the mid 1970s, so many mutants had been developed from strain 168 that a centralized repository, the Bacillus Genetic Stock Center (BGSC), was established in Ohio to maintain them. Interestingly, X-irradiation and domestication brought many changes in *B. subtilis* strain 168 behaviors: a dramatic reduction in its ability to form complex, structured and matrix-adhered colonies known as biofilms, an inability to swarm on solid surfaces and an increased competence for genetic transformation (Branda, Gonzalez-Pastor, Ben-Yehuda, Losick, & Kolter, 2001) compared to its probable ancestor NCIB3610 (hereafter referred to as 3610). One major change was the loss of a plasmid, called pBS32, which encodes a cassette of phage genes (Konkol, Blair, & Kearns, 2013). While

pLS32 is not present in the closest relatives of strain *B. subtilis* 3610, it is found in more distantly related strains (Tanaka & Ogura, 1998).

Cell types in *Bacillus subtilis*

Soil is a variable environment and accordingly *B. subtilis* can adapt by differentiating into distinct subpopulations of specialized cells, which express different traits that may confer a survival advantage in adverse conditions (Lopez, Fischbach, Chu, Losick, & Kolter, 2009) (Lopez, Vlamakis, & Kolter, 2009). Thus, *B. subtilis* is a good model organism for the study of alternative lifestyles. When nutrients are abundant, cells can grow in a free-floating planktonic form. At the onset of stationary phase, or when *B. subtilis* is shifted to poor nutrient conditions, some *B. subtilis* cells adapt by becoming naturally transformable (Chen et al., 2005). Other cells secrete toxins, which kill siblings by inducing cell lysis and the release of cellular components for scavenging (Lopez, Fischbach, et al., 2009). Also, during growth, cells may switch from a sessile to a motile state or vice versa. The sessile lifestyle choice leads to the formation of a multi-cellular, community known as a biofilm. Production of the extracellular matrix, which is essential for biofilm formation, is carried out by a subpopulation of specialized cells in *B.* subtilis (Chai, Chu, Kolter, & Losick, 2008). However, the whole community benefits from the presence of the extracellular matrix, because all of the cells are encased within the matrix in mature biofilms, and thus protected from environmental insults such as antibiotics or phage (Vlamakis, Aguilar, Losick, & Kolter, 2008). As an extreme survival mechanism, the population will produce spores, and the mother cells lyse during the process. These spores are metabolically inactive, resistant to heat, radiation and toxic chemicals, and able to persist over long periods of time without nutrients. If optimal conditions arise and nutrients become available, spores germinate and resume

growth. The development of each of these specialized states is triggered by specific signals. For example, sporulation is proposed to be activated when the intracellular concentrations of essential metabolic molecules, like GTP, decrease (Lopez, Gontang, & Kolter, 2010) but the signal itself remains unknown. Transformation is regulated in part by quorum-sensing, a form of cell-cell communication in which cells secrete signaling molecules that accumulate in the medium as they grow and typically elicit a coordinated response from the population (Hahn & Dubnau, 1991) (Lazazzera, 2000). Overall, although the cells of a *B. subtilis* culture are genetically identical, different triggers activate specific changes in gene expression that often result in distinct subpopulations of cells (Lazazzera, 2000) (Lopez, Vlamakis, & Kolter, 2009).

Competence for transformation and the K-state

Foreword

The timing for competence expression is different among bacteria. *S. pneumoniae*, for example, expresses competence only during early exponential growth (Pakula & Walczak, 1963) while *Neisseria gonorrhoeae* is competent during the entire exponential phase (Sparling, 1966). *B. subtilis* develops competence at the onset of the stationary phase. The percentage of competent cells is also different between bacteria: only 10 to 20% of the cells express competence in domesticated *B. subtilis* while the entire population becomes competent in *S. pneumonia*e. Also, factors triggering the expression of competence are different; as the chemical composition of the media is important for *S. pneumoniae,* whereas nutrient limitation is important for *B. subtilis,* although this has not been studied in any detail (Tomasz, 1966) (Morrison & Baker, 1979).

Competence provides the cell population with an alternative mechanism of survival under environmentally challenging conditions. Indeed, different theories have been proposed on the benefits of genetic transformation, such as the use of DNA to repair damaged genes, the use of DNA as a source of nutrition (carbon, nitrogen and phosphorous) or the use of this exogenous DNA to allow genetic diversity in the population (Finkel & Kolter, 2001). A unique feature of competence in *B. subtilis* is that cells expressing competence do not divide. It has also recently been shown that competent cells are tolerant of antibiotics, a form of persistence (Hahn, Tanner, Carabetta, Cristea, & Dubnau, 2015). Because more than 100 genes are expressed when cells are competent (Hamoen, Smits, de Jong, Holsappel, & Kuipers, 2002) (Berka et al., 2002) (Ogura et al., 2002), and because most of these genes are not needed for transformation, the competent state has been called the K-state, named for

the master regulator ComK (Maamar & Dubnau, 2005). The emergence from genetically identical cells of two distinct subpopulations (competent cells and non-competent cells) is a classical survival strategy for bacteria, known as bet-hedging (Suel, Kulkarni, Dworkin, Garcia-Ojalvo, & Elowitz, 2007) (Veening, Smits, & Kuipers, 2008).

Overview of the transformation process in *B. subtilis*

The proteins that are essential for DNA-uptake are encoded by three different operons, *comE* (Albano, Hahn, & Dubnau, 1987) (Hahn, Inamine, Kozlov, & Dubnau, 1993) (Inamine & Dubnau, 1995), *comF* (Londono-Vallejo & Dubnau, 1993) (Londono-Vallejo & Dubnau, 1994) and *comG* (Albano, Breitling, & Dubnau, 1989) (Albano & Dubnau, 1989) (Briley et al., 2011) as well as the gene product of *comC* (Mohan, Aghion, Guillen, & Dubnau, 1989) (Chung & Dubnau, 1995). Other genes are needed for the processing and integration of DNA following uptake. The process of transformation can be subdivided in four distinct steps. First, double-stranded DNA (dsDNA) from the environment is bound to competent cells. Previous studies showed that there is no sequence preference and that there are approximately 50 binding sites per cell (Dubnau, 1991). It was proposed that a "competence pseudopilus," made up of pilin subunits encoded by the *comG* operon which are processed by the protease ComC, traverses the cell wall and makes contact with the exogenous dsDNA. However, while the competence pseudopilus is required for DNA binding, none of its components have been shown to have DNA binding properties (Chen & Dubnau, 2004), suggesting that some unknown DNA binding protein exists. Following the association of DNA with the cell, the ds-DNA is non-specifically cleaved by the membrane-bound endonuclease NucA, where the average size of the fragments is 11-18kb (Provvedi, Chen, & Dubnau, 2001) (Figure 4). Next, the dsDNA is transported

through the plasma membrane via the aqueous ComEC channel and one strand is degraded by an unknown nuclease. ComFA was proposed to be a DNA translocase and/or helicase, which may provide energy for the movement of the DNA through the ComEC channel (Londono-Vallejo & Dubnau, 1993) (Londono-Vallejo & Dubnau, 1994) (Dubnau & Cirigliano, 1972) (Lacks, Greenberg, & Neuberger, 1975). In *B. subtilis* there is no evidence that the complementary strand is degraded while the other strand is transported through the channel, as has been shown in *S. pneumoniae* (Mejean & Claverys, 1993). Also, an ortholog of the endonuclease that degrades the non-transforming strand in *S. pneumoniae* is lacking in *B. subtilis.* Finally, once in the cytoplasm, the single stranded DNA (ssDNA) is integrated into the genome by homologous recombination (Fernandez, Ayora, & Alonso, 2000). This final step requires the involvement of competence-induced proteins DprA, RecA and SSB. Indeed, DprA (DNA processing protein A) presumably binds the ssDNA to protect it from degradation by nucleases, and also recruits the recombinase RecA (Yadav et al., 2013) (Yadav, Carrasco, Serrano, & Alonso, 2014) (Lenhart, Schroeder, Walsh, & Simmons, 2012). RecA polymerizes on ssDNA and promotes a homology search along chromosomal DNA (Yadav et al., 2013) (Yadav et al., 2014) to form a recombination heteroduplex intermediate with the host chromosome (Figure 4).

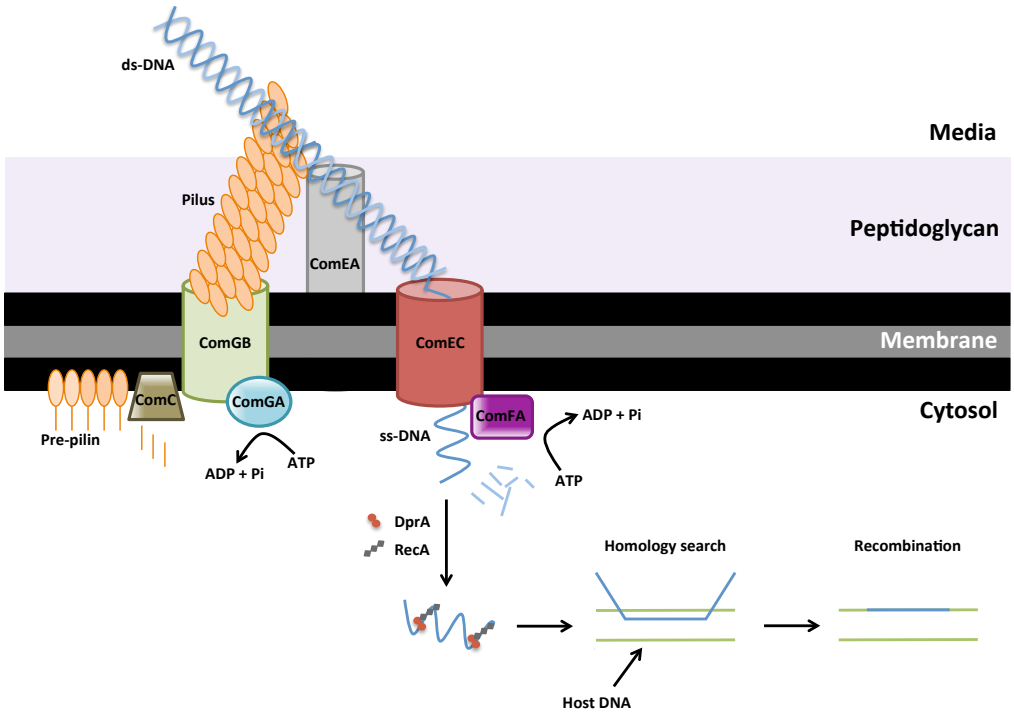


Figure 4: DNA uptake machinery in *B. subtilis*
ComG pre-pilin subunits are first translated as integral membrane proteins with a cytoplasmic leader peptide. They are then processed by the pre-pilin peptidase ComC before being translocated out of the membrane where they oligomerize into a pilus. The energy necessary for the movement of the pilus through the conserved membrane protein ComGB is provided by the associated ATPase ComGA. When the double stranded DNA (dsDNA) makes contact with the pilus, it is believed that ComEA delivers the DNA to the membrane channel ComEC where one strand of DNA is likely degraded by an unknown nuclease. This single-stranded DNA (ssDNA) internalization is driven by the ATP-dependent translocase ComFA. Cytoplasmic ssDNA is bound by DprA, which recruits RecA to allow homologous recombination with the host DNA.

K-state regulation in *B. subtilis*

Competence development in *B. subtilis* has become an important model for bistable gene expression where the expression of a key regulatory protein (ComK) is tightly regulated to prevent the entire population from becoming competent. In *B. subtilis*, competence development is the result of a dramatic increase in the cellular concentration of the transcriptional regulator, ComK (Hamoen, Van Werkhoven, Bijlsma, Dubnau, & Venema, 1998; van Sinderen et al., 1995; van Sinderen & Venema, 1994). ComK is referred to as the "master regulator of competence," because it is both necessary and sufficient for the transcriptional activation of all of

the K-state genes. ComK is a 22kDa transcription factor that binds as a tetramer or dimer of dimers to A/T rich regions within its targets promoter genes that contain "ComK boxes" (van Sinderen et al., 1995) (Hamoen et al., 1998). But, in addition to competence regulation, ComK also activates the transcription of genes involved in many other different processes such as cell shape determination, cell division, transcriptional regulation, transport, protein synthesis and stress responses (Ogura et al., 2002) (Berka et al., 2002). As noted already, ComK regulates more than genes required for transformation, and the "K-state," which can be broadly described as a global, persistence-like, bet-hedging, expression state (Berka et al., 2002) (Hahn et al., 2015). Thus, the K-state enhances fitness by a means distinct from transformability (Maamar & Dubnau, 2005).

Regulation of *comK*

As mentioned previously, cells expressing ComK are fully dedicated to enter the K-state and stop DNA replication and cellular division. This represents a risk for the cell population. To avoid such a situation, the expression of ComK is tightly regulated during growth. During exponential growth, the transcription of *comK* is inhibited via the direct binding of the transcriptional repressors AbrB (Hamoen et al., 2003), CodY (Serror & Sonenshein, 1996) and Rok (Hoa, Tortosa, Albano, & Dubnau, 2002) (Figure 4). In addition to this transcriptional repression, ComK stability is also regulated. Any ComK that is made will be degraded by a second regulation system, ClpC/ClpP mediated proteolysis (Turgay, Hahn, Burghoorn, & Dubnau, 1998). The adaptor protein MecA directly interacts with ComK and delivers it to the ATP-dependent chaperone ClpC, where it is unfolded and then completely degraded by ClpP (Figure 5) (Turgay, Hamoen, Venema, & Dubnau, 1997) (Turgay

et al., 1998). These negatively acting mechanisms are relieved as the cells approach stationary phase, which is when the small peptide ComS is produced. ComS binds directly to the ComK-binding site of MecA releasing ComK from degradation (Turgay et al., 1997) (Prepiak & Dubnau, 2007). The expression of *comS* is a consequence of quorum-sensing mechanisms taking place in late exponential phase. Once released from degradation, the basal level of ComK exceeds a threshold in some cells and *comK* expression will then be amplified by a positive feedback loop in which ComK acts on its own promoter (P*comK*). The resulting burst of ComK synthesis triggers the expression of the transformation apparatus genes and the other K-state genes. (Figure 5).

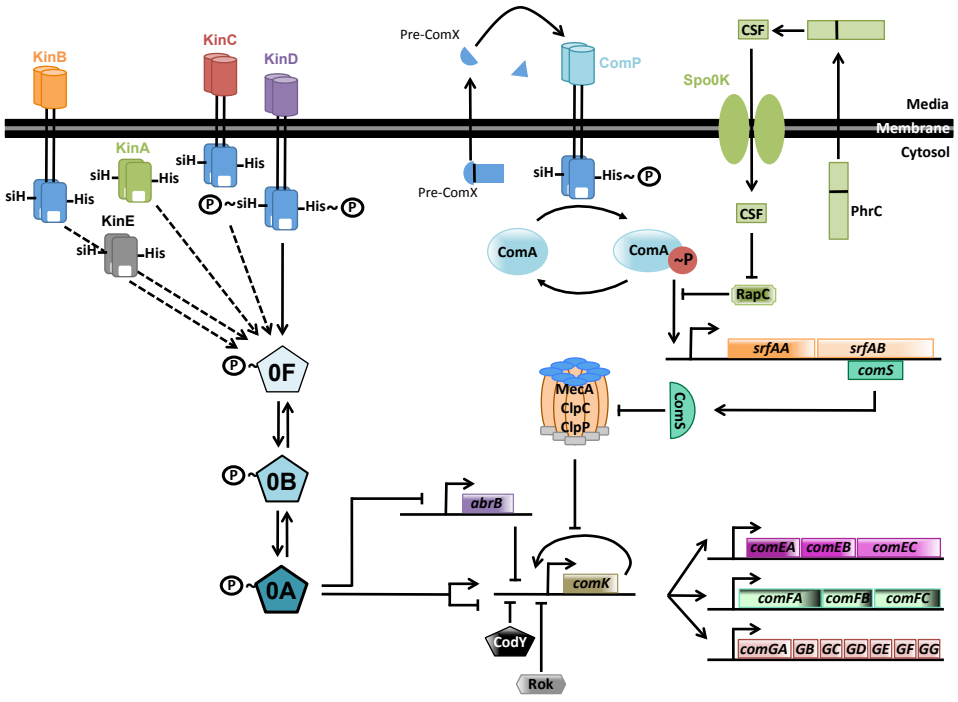


Figure 5: Overview of the competence regulation in *B. subtilis*
A represention of the relevant effectors involved in the expression of competence for DNA uptake. On the left is the phosphorelay composed of five histidine kinases KinA-E, Spo0F (0F), Spo0B (0B) and Spo0A (0A) ultimately leading to the phosphorylation of the master regulator Spo0A. Depending on its concentration in the cells, Spo0A-P will activate and repress directly or indirectly the *comK* promoter. To the right, the synthesis and processing of the two qurorum sensing pheromones ComX and CSF are shown, which leads to the stabilization of ComK through the activation of the transcription of the *srfA* operon and the synthesis of ComS peptide. ComS, in turn, blocks ComK degradation by the MecA-ClpC-ClpP complex. Once ComK is stabilized, it autoactivates its own transcription, and triggers the expression of the *com* operons, which encode the proteins to assemble the DNA uptake apparatus.

Regulation of *comK* expression by quorum sensing

In strain 168 derivatives the maximum of K-state expression is reached two hours after entering in the stationary phase because of the accumulation of quorum sensing pheromones. Quorum-sensing is a cell-cell communication mechanism that monitors cell density. Secreted molecules accumulate in the medium and activate a cognate receptor at high concentration, triggering an intracellular response (Lazazzera, 2000). In *B. subtilis*, two quorum-sensing peptides are known to regulate the expression of *comS*: ComX (Magnuson, Solomon et al. 1994) and competence stimulating factor, CSF (Solomon, Magnuson, Srivastava, & Grossman, 1995).

ComX is a signaling peptide that regulates competence through the ComP-ComA two-component regulatory system (Hahn & Dubnau, 1991) (Solomon et al., 1995). Synthesized as a 55 amino acid precursor, pre-ComX is cleaved and modified by ComQ, to generate its mature form consisting in a 10 amino acid peptide with a hydrophobic isoprenoid modification on a tryptophan and a cyclization event (Weinrauch, Msadek, Kunst, & Dubnau, 1991) (Ansaldi, Marolt, Stebe, Mandic-Mulec, & Dubnau, 2002). The mature ComX is secreted and when present in high enough concentrations, binds to ComP, a membrane histidine kinase. Binding activates the autophosphorylation of a conserved histidine residue in the cytoplasmic domain of ComP (Piazza, Tortosa, & Dubnau, 1999). The phosphoryl group is transferred to the N-terminal regulatory domain of its cognate response regulator ComA (Weinrauch et al., 1991) allowing its binding to DNA and the activation of the *srfA* operon that encodes ComS (Figure 5) (Nakano, Xia, & Zuber, 1991) (Nakano & Zuber, 1991) (Roggiani & Dubnau, 1993).

CSF (Competence and Sporulation Factor) is the second signaling peptide that regulates competence through quorum-sensing. CSF comes from the last five codons of the gene product of *phrC* (Solomon, Lazazzera, & Grossman, 1996) and is exported via an unknown mechanism after cleavage to produce a pentapeptide (ERGMT) without any post-translational modifications. Secreted CSF is internalized back into the cell through the oligopeptide permease Spo0K (Rudner, LeDeaux, Ireton, & Grossman, 1991) (Solomon et al., 1995) (Lazazzera, Solomon, & Grossman, 1997). This peptide then prevents the dephosphorylation of ComA-P by inhibiting the phosphatase RapC (Figure 5) (Solomon et al., 1995).

Regulation of *comK* basal expression: "the uptick"

Another cause of temporal regulation comes from the tight regulation of *comK* basal expression. The basal expression of *comK*, measured in the absence of ComK autoregulation, increases gradually with growth, reaching a maximum as cells enter the stationary phase, and then declines (Leisner, Stingl, Radler, & Maier, 2007) (Mirouze, Desai, Raj, & Dubnau, 2012). This uptick in basal expression is the result of an increasing concentration of the phosphorylated form of the master regulator Spo0A. Spo0A-P levels are controlled by a multi-component phosphorelay (further described below). Spo0A-P affects the *comK* promoter both directly and indirectly. Spo0A-P represses the expression of *abrB* thus relieving the repression of AbrB on the *comK* promoter (Figure 6A) (Strauch, Webb, Spiegelman, & Hoch, 1990). Secondly, early on when Spo0A-P levels are low, it binds to three high affinity sites within the *comK* promoter and activates *comK* transcription by antagonizing the repressive effect of Rok (Mirouze et al., 2012). As Spo0A-P levels accumulate in the cells, it binds to two lower affinity sites within the *comK* promoter, shutting off the

transcription of *comK*. Thus, Spo0A-P establishes a temporal gate referred as to a "window of opportunity" for cells to enter the K-state, by mediating a transient uptick in *comK* expression (Figure 6A) (Mirouze et al., 2012).

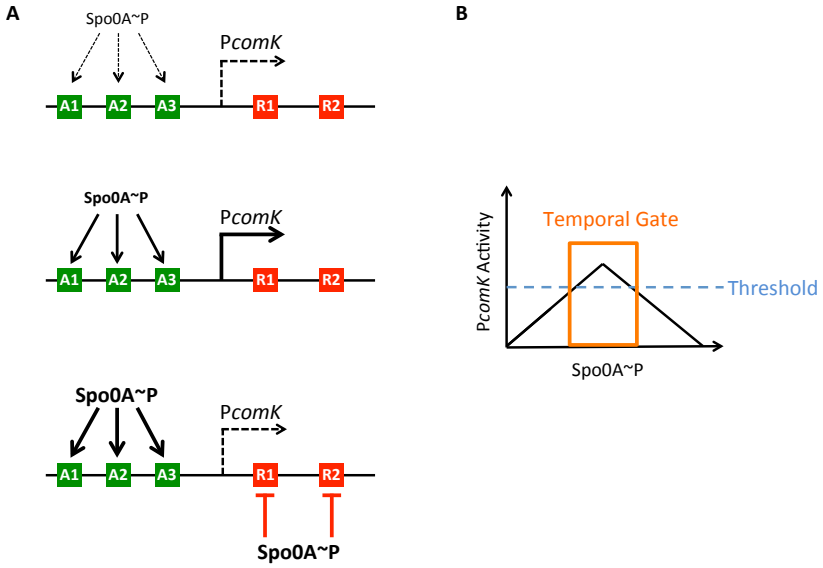


Figure 6: *comK* uptick regulation in *B. subtilis*
A- Low levels of Spo0A-P bind to three low affinity sites (A1-A3) and activate basal *comK* transcription by antagonizing the repressor Rok. At higher concentrations, Spo0A-P binds to repressor sites (R1-2) that are downstream of the transcription start site, and block gene expression. **B-** Schematic representation of the window of opportunity for competence development in *B. subtilis* defined by the levels of Spo0A-P and their effects on the *comK* promoter.

In a small percentage of cells the basal level of ComK is high enough to allow expression of the DNA uptake apparatus and other K-state genes, because the expression is above a threshold and ComK has been further stabilized through the quorum-sensing produced ComS. ComS is made in all cells in a population, whereas noise in the basal expression allows only some cells to exceed the threshold for positive autoregulation. Thus the temporal gate provided by the uptick results in a "window of opportunity" for K-state expression explaining both the timing of expression and its bistable nature (Mirouze et al., 2012) (Smits et al., 2005). The

accumulation of ComX contributes to the timing by ensuring that the K-state can only occur when the cell density is high (Figure 6B).

The escape from the K-state

When the cell population is facing favorable environmental conditions the non-competent cells resume growth, whereas competent cells are delayed at least 90 minutes before they start growing again. This is because growth is inhibited by ComGA, Maf and MreB, all of which are expressed in the K-state under ComK control. The late competence protein ComGA is known to inhibit cell elongation by preventing the formation of FtsZ rings but this regulation mechanism remains unknown. ComGA prevents the degradation of ppGpp by binding to RelA (Hahn et al., 2015) and the accumulation of this small molecule inhibits cell elongation. It is possible that this is the mechanism of inhibition of FtsZ rings because these structures do not form if cell-size remains small. The highly conserved protein Maf is also involved in the regulation of cell division in competent cells where Maf acts downstream of ComGA (Briley et al., 2011). *comGA* null mutant cells remain blocked in division in a later stage (Haijema, Hahn, Haynes, & Dubnau, 2001). When the cells become competent ComGA accumulates at the cell poles where it co-localizes with other proteins to form the DNA uptake machinery (Hahn, Maier, Haijema, Sheetz, & Dubnau, 2005). The actin-like protein MreB also plays a role in the delay in growth observed for competent cells (Mirouze, Ferret, Yao, Chastanet, & Carballido-Lopez, 2015). During vegetative growth, MreB localizes along the sidewalls and promotes cell elongation whereas, in competent cells, MreB relocalizes at the cell poles with ComGA. After 120 minutes, MreB and ComGA co-localization is lost, allowing MreB to relocalize along the sidewalls to reinitiate elongation suggesting that

ComGA sequesters MreB in competent cells to prevent cell elongation and therefore escape from competence. Thus ComGA may work in two ways to delay the resumption of growth. This delay in growth for the competent cells may be a "check-point" allowing the cells to repair the chromosome after recombination and before resuming growth. Also, it is likely that the delay has been selected because it confers antibiotic tolerance (persistence) (Nester & Stocker, 1963) (Haijema et al., 2001; Johnsen, Dubnau, & Levin, 2009) (Briley et al., 2011; Hahn et al., 2015) (Yuksel, Power, Ribbe, Volkmann, & Maier, 2016).

Two component systems in bacteria

General introduction

Free-living organisms modulate gene expression in response to environmental changes. To do so, they require sensors to detect physical and/or chemical signals as well as regulators to change the levels of gene products. Many of these signaling systems have been described as two-component systems, since they depend upon the interaction of two regulatory proteins: a sensor kinase and a response regulator (Henner, Ferrari, Perego, & Hoch, 1988) (Giraldo, Andreu, & Diaz-Orejas, 1998; Jiang, Shao, Perego, & Hoch, 2000).

Sensor kinases typically have two functional domains: an N-terminal stimulus detection domain (or input domain, which is often transmembrane) and a C-terminal autokinase domain. The autokinase domain contains a phosphotransferase subdomain (with a histidine that becomes phosphorylated) as well as an ATP-binding subdomain (Figure 7A). Because of the variety of signals, input domains are very heterogeneous in amino acid sequence and size compared to autokinase domains, which are of

similar length and show many conserved amino acids suggesting a common evolutionary origin (Mitrophanov & Groisman, 2008). When the input domain of the kinase is activated, it performs an autophosphorylation reaction, transferring a phosphoryl group from ATP to a specific histidine residue within the phosphotransferase sub-domain. The phosphoryl group is then transferred from the kinase to an aspartate residue on the response regulator's receiver domain (Figure 7A). This typically triggers a conformational change that activates the response regulator's effector domain, which in turn produces the cellular response to the signal, usually by activating (or repressing) expression of target genes (Mitrophanov & Groisman, 2008) (Capra & Laub, 2012). While the majority of effector domains have DNA-binding activity to regulate the transcription of specific genes, some have an enzymatic domain or no C-terminal domain at all, i.e. in the chemotaxis system and in the Spo0A phosphorelay (Stock, Robinson, & Goudreau, 2000).

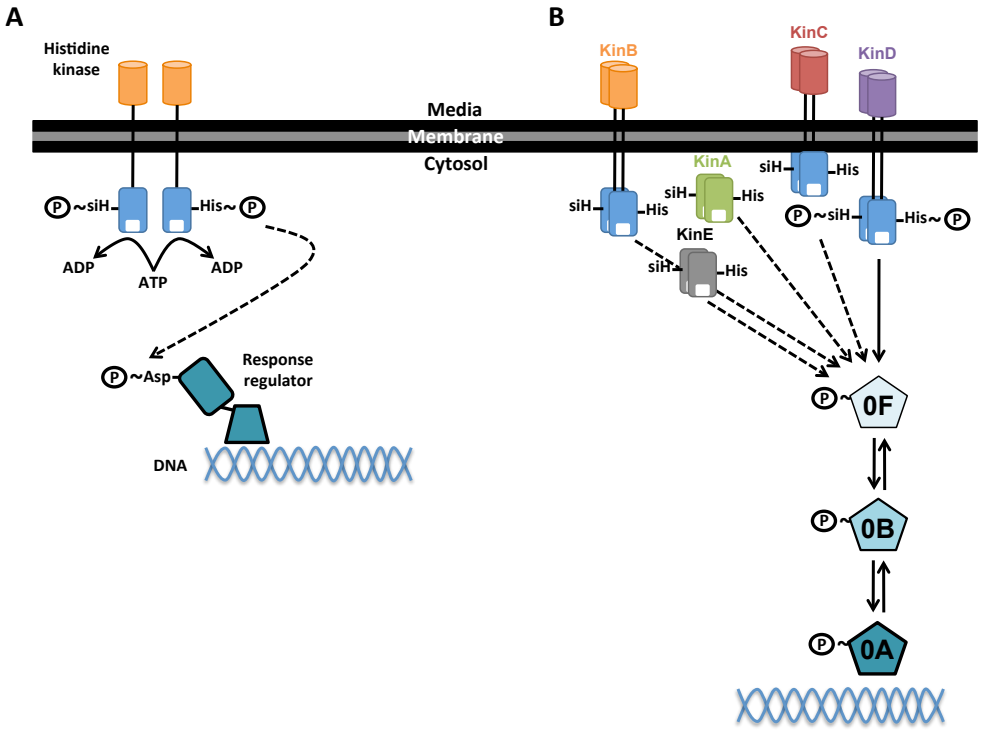


Figure 7: Two component systems in bacteria

A- Two component system arecomposed of a histidine kinase and its cognate response regulator. After activation by signal binding to the input domain of the kinase, a histidine residue is phosphorylated within the autokinase domain. The phosphoryl group is then transferred from the histidine to an aspartate residue on the response regulator's receiver domain, which triggers a conformational change that activates the response regulator's effector domain. **B-** Phosphorelays are evolved two component systems. Shown here is the phosphorelay in *B. subtilis* leading to the phosphorylation of Spo0A (here 0A). Input signals can be sensed by a pool of five kinases (KinA, B, C, D and E) leading to the phosphorylation cascade where Spo0F and Spo0B (0F and 0B respectively) are used as intermediates between these kinases and their cognate response regulator Spo0A.

The number of two-component systems present in a bacterial genome is highly correlated with genome size as well as ecological niche; bacteria that occupy niches with frequent environmental fluctuations indeed have more histidine kinases and response regulators (Capra & Laub, 2012). Furthermore it has been shown that bacteria can acquire new two-component systems through gene duplication or by horizontal gene transfer (Alm, Huang, & Arkin, 2006). In most cases, response regulator genes are located in the same operon as their cognate histidine kinase.

The phosphorelay

To respond to multiple signal inputs, bacteria may use more complex types of two-component-based systems called phosphorelays. In these systems, the response regulator is not directly phosphorylated by the sensor kinase but instead additional regulatory and phosphotransferase domains may be involved. For example, the sensor kinase may transfer the phosphoryl group to a single domain response regulator that will subsequently transfer it to a second phosphotransferase domain, which then serves as the primary phosphoryl donor to the response regulator. In *B. subtilis*,the "sporulation phosphorelay" is one such example, where the sensor kinases KinA-E transfer a phosphoryl group to the transcription factor Spo0A through Spo0F and Spo0B (Burbulys, Trach, & Hoch, 1991) (Figure 7B). In this case, the different domains are on different proteins, but for other phosphorelays, the sensor kinase and the additional domains can form a multi-domain protein, as in the BvgS sensor kinase of *Bordetella pertussis* (Uhl & Miller, 1996).

The DegS-DegU Two Component System in *B. subtilis*

The DegS-DegU two-component system in B. subtilis is involved in the control of many cellular processes, including exoprotease production and the K-state (Dahl, Msadek, Kunst, & Rapoport, 1992) (Msadek, Kunst, Klier, & Rapoport, 1991). DegS is a 44 kDa protein with a histidine kinase domain, but without a transmembrane segment, suggesting that DegS is either cytosolic or associated to the membrane through another protein. The signal that activates DegS has not yet been identified, but recently (Hsueh et al., 2011) (Cairns, Marlow, Bissett, Ostrowski, & Stanley-Wall, 2013) it has been proposed that flagellar rotation or basal body assembly may serve as signals that inhibit DegS phosphorylation. The *degU* and *degS*

genes constitute an operon that contains three promoters, which will be discussed in more detail later. (Figure 8A).

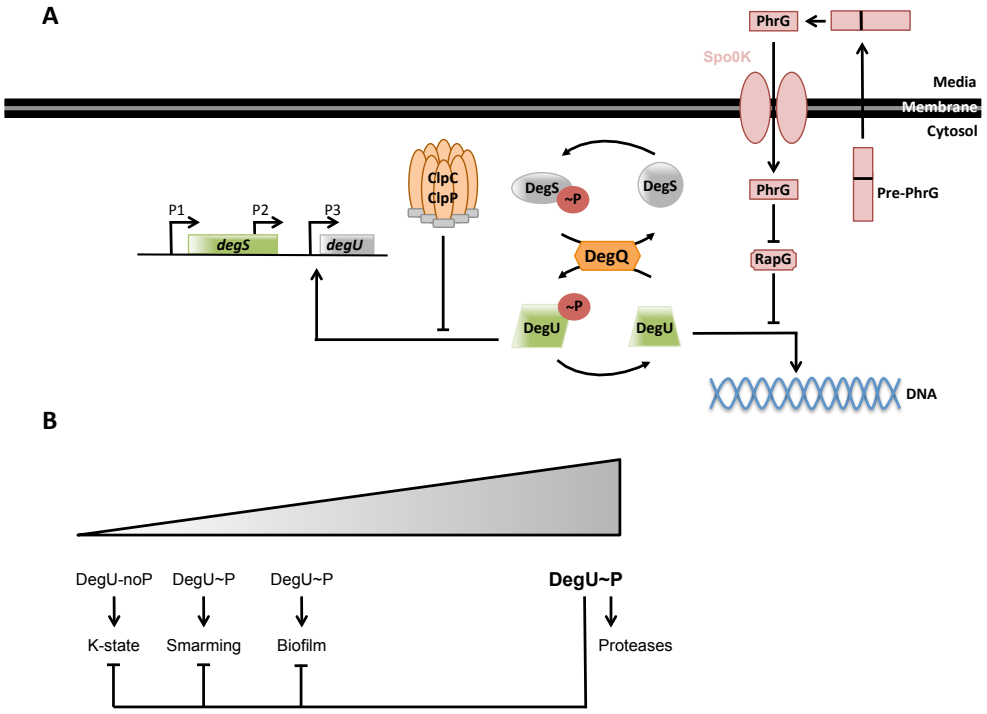


Figure 8: The DegS-DegU two component system in *B.subtilis.*
A- Schematic representation of the regulation of the response regulator DegU by its cognate kinase DegS. DegQ facilitates the phosphotransfer from DegS to DegU by an unknown mechanism. The cell regulates the level of DegU-P by interrupting the positive feedback loop of DegU-P on the P3 promoter of *degU*. DegU DNA binding is also regulated by the phosphatase RapG which is inhibited when PhrG accumulates in the cell. **B-** Representation of the different effects of DegU-P depending on its concentration in the cell. Unphosphorylated DegU activates competence, while a low and intermediate level of DegU-P triggers swarming and biofilm formation. Instead, a high concentration of DegU-P in the cell activates the synthesis of exoproteases and inhibits competence, swarming motility and biofilm formation.

DegU, a 25 kDa protein that belongs to the LuxR-FixJ family, has a helix-turn-helix DNA domain at its C-terminus that recognizes AT-rich octamers. DegU is a response regulator activated by phosphorylation of a conserved aspartate residue within N-terminal domain by its cognate kinase DegS (Ogura, Shimane, Asai, Ogasawara, & Tanaka, 2003). Importantly, the small protein DegQ stimulates

phosphate transfer from phosphorylated *DegS* (DegS~P) to DegU (Figure 8A) (Kobayashi, 2007b).

DegU is also regulated direct interaction by RapG, which inhibits DegU's DNA binding. This inhibition is relieved by the accumulation of the small peptide PhrG, that prevents RapG binding (Ogura et al., 2003) (Figure 8A). When unphosphorylated, DegU is known to regulate the *comK* promoter by stimulating ComK binding to its own promoter when concentrations are low (Hamoen, Van Werkhoven, Venema, & Dubnau, 2000). No other targets of unphosphorylated DegU are known, except possibly the *fla-che* promoter (Tsukahara & Ogura, 2008) (Mordini et al., 2013), whereas phosphorylated DegU is known to trigger the expression of many genes, including *aprE*, *nprE* (degradation of proteins), *sacB*, *sacX* (sucrose metabolism) and *degQ*, and repress that of *wapA* (WapA carries a C-terminal toxin domain which is deployed to inhibit the growth of neighboring cells) (Msadek et al., 1991) (Crutz & Steinmetz, 1992) (Dartois, Debarbouille, Kunst, & Rapoport, 1998). For example, when levels of DegU-P are low, it activates the *fla-che* operon, which is critical for motility (Hamoen et al., 2000) (Tsukahara & Ogura, 2008). Low levels of DegU~P also promotes complex colony architecture during biofilm formation, but higher DegU~P concentrations inhibits it (Verhamme, Kiley, & Stanley-Wall, 2007). When levels of DegU-P are high, *degU* transcription is activated by an autoregulatory loop (Kobayashi, 2007b) (Ogura & Tsukahara, 2010), but the AAA+ (**A**TPase **a**ssociated with diverse cellular **a**ctivities) protease ClpCP may specifically degrade the DegU-P, leading to modulation of DegU autoactivation (Ogura & Tsukahara, 2010). The hyperphosphorylation of DegU results in overproduction of degradative enzymes and prevents K-state development (Hahn, Luttinger, & Dubnau, 1996). The genetic context of the DegS-DegU two-component system and its role in the

regulation of competence development in *B. subtilis* will be discussed later in this manuscript (Figure 8B).

Biofilms

Introduction to biofilms

Biofilm formation is a universal trait among bacteria, and biofilms can be found on many diverse natural or artificial surfaces (Hall-Stoodley, Costerton, & Stoodley, 2004) (Stewart & Franklin, 2008). Since biofilms are problematic in many man-made settings, they have been studied intensively during the past decade. Biofilms confer resistance to antimicrobial and antibacterial agents (Mah & O'Toole, 2001) (Stewart, 2002) (Davies, 2003). There is also an industrial interest in characterizing biofilms to possibly exploit them for bioremediation, which is a waste-management technique that involves the use of organisms to remove or neutralize pollutants from a contaminated site, or as a potential source of energy in the form of microbial fuel cells (Singh, Paul, & Jain, 2006) (Logan, 2009). Natural biofilms likely contain mixtures of different microbial species, but much work has focused on single species biofilms.

Description of *B. subtilis* biofilms

Biofilms are communities of surface-associated microorganisms encased in a self-produced extracellular matrix. The *B. subtilis* matrix is primarily composed of exopolysaccharide (EPS) and proteins. The major EPS component of all *B. subtilis* biofilms is synthesized by the products of the *epsABCDEFGHIJKLMNO* operon (Branda, Vik, Friedman, & Kolter, 2005) (Kearns, Chu, Branda, Kolter, & Losick, 2005). Another extracellular polymer, γ-poly-DL-glutamic acid (PGA), is produced in large amounts by some strains and can enhance formation of submerged biofilms, but

PGA is not required for wrinkled-colony morphology or for pellicle formation (Branda, Chu, Kearns, Losick, & Kolter, 2006). In addition to EPS, many other proteins are found in biofilms, and two important structural proteins have been so far described: the translocation-dependant antimicrobial spore component (TasA) and the biofilm surface layer protein BslA (Kobayashi & Iwano, 2012). So far, TasA has been described to form long fibers attached to the cell through TapA (TasA anchoring and assembly protein) but this theory has recently been challenged (not yet published). These two proteins are encoded by the *tapA-sipW-tasA*, *sipW* encoding for the type I signal peptidase W which processes both TasA and TapA (Stover & Driks, 1999).

B. subtilis biofilms have been studied as colonies at an air-agar interface, floating biofilms that form at the air-liquid interface (also termed pellicles) and, in the case of certain domesticated strains, submerged, surface-adhered biofilms that form at the liquid-solid interface. The capacity to form robust biofilms has largely been lost in the descendants of the laboratory strain 168, and is best studied on natural isolates, such as 3610, the probable parent of 168 (McLoon, Guttenplan, Kearns, Kolter, & Losick, 2011). In the laboratory, matrix production and biofilm formation are promoted when cells are grown in a defined media (MsGG).

Biofilm life cycle

It has been shown that biofilms contain different cell types, in that cells producing matrix, expressing motility or sporulating are all found in a same biofilm in *B. subtilis* (Vlamakis et al., 2008). Using time-lapse microscopy, it was showed that at the early stage of biofilm formation on a solid surface, most cells produce flagella and were motile (Vlamakis et al., 2008). Later, the number of motility-expressing cells decreases, and the few remaining motile cells are located at the edge and the base of

the biofilm once it is formed. This is in part because EpsE, not only participates in matrix formation but also acts as a clutch to repress flagellar rotation by interacting with the flagella motor switch protein, FliG (Blair, Turner, Winkelman, Berg, & Kearns, 2008) (Guttenplan, Blair, & Kearns, 2010). Matrix production is then activated and requires Spo0A-P. While the number of matrix-producing cells declines after 24 hours, sporulating cells have been shown to arise from the matrix producers and to locate preferentially in aerial projections from the biofilm surface, called "fruiting bodies" (Figure 9) (Branda et al., 2001).

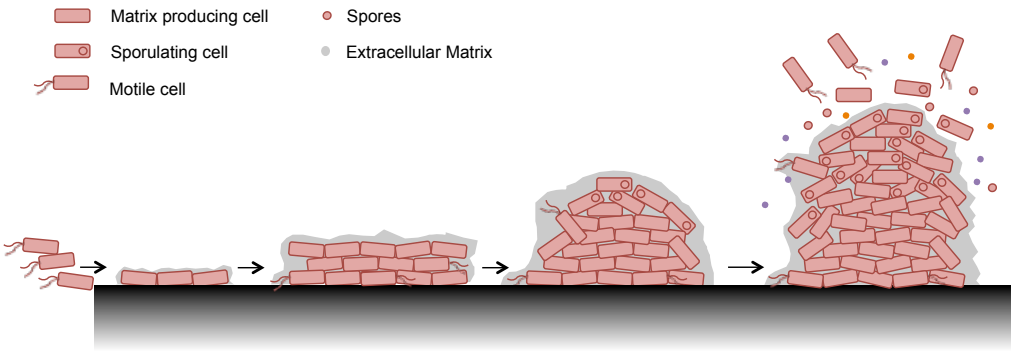


Figure 9: The life cycle of a *B. subtilis* biofilm. The formation of a biofilm occurs in several stages, comprising the development, maturation and disassembly of the bacterial community. At the initiation of biofilm formation, motile cells with flagella differentiate into non-motile, matrix-producing cells that form long chains which are encased in extracellular matrix. In mature biofilms, matrix-producing cells may sporulate. As time progresses, the extracellular matrix will break down, which allows the cells to disperse in the environment. The exact signals and mechanism for biofilm dispersal remain unknown. It is important to note that although functionally distinct cell types exist within the biofilm, these cells are genetically identical, and differentiation into a specific cell type is not terminal and can be altered when environmental conditions change.

It is important to note that although functionally distinct cell types exist within the biofilm, these cells are genetically identical, and differentiation into a specific cell type is not terminal and can be altered when environmental conditions change. Given all of the components that are necessary to assemble the matrix, *B. subtilis* has evolved a complex regulatory network to coordinate expression of matrix genes in

response to the shifting environmental conditions that it encounters in its natural environment.

The Spo0A pathway

As its name suggests, Spo0A was first discovered as a gene required for the sporulation pathway. Further studies revealed that Spo0A is an essential transcriptional factor for all the adaptation pathways in *B. subtilis* by controlling the expression of more than a hundred promoters, which differ in their Spo0A-P binding affinities. Spo0A-P induces biofilm formation when present at intermediate levels, but as the biofilm matures and the concentration of Spo0A-P increases in a subpopulation of cells, sporulation will be triggered (Fujita, Gonzalez-Pastor, & Losick, 2005). As previously described, Spo0A-P also regulates the K-state.

Spo0A-P promotes biofilm formation by inhibiting the action of two major repressors of the *epsA-O* and the *tapA* operons, SinR and AbrB (Figure 10). The derepression of SinR via Spo0A-P is indirect,. Spo0A-P directly increases the transcription of *sinI*. SinI is an antagonist of the matrix gene repressor SinR, via protein-protein interactions that inhibit SinR DNA binding. SinI is a SinR paralog, that lacks the N-terminal DNA binding domain but contains a C-terminal oligomerization domain similar to that of SinR. Furthermore, SinR activity is regulated by SlrR. SlrR is another SinR paralog, but contains both domains found in SinR. Its transcription is repressed by SinR (Chu et al., 2008). When *slrR* expression is derepressed by SinR inactivation through SinI, induced SlrR binds to SinR and reprograms SinR to repress expression of motility-promoting genes (see below) (Vlamakis et al., 2008) (Chai et al., 2008). The second repressor of matrix production AbrB is directly repressed by very low concentrations of Spo0A-P. In addition to its

inhibitory effect on the *epsA-O* and the *tapA* operons, AbrB has been shown to repress the expression of the regulatory protein SlrR and the matrix protein BslA (Chu et al., 2008) (Chai, Kolter, & Losick, 2009). The presence of the two matrix production repressors SinR and AbrB with overlapping targets indicates that the regulation of biofilm formation is tightly regulated in order to coordinate expression of the matrix genes.

The SinR-SlrR regulation switch

The SinR-SlrR complex serves as a switch between biofilm formation and motility, which is why, as previously mentioned, both SinR and AbrB control *slrR* gene expression. At low concentrations of SlrR, *lytABC* and *lytF* (genes encoding for autolysins, which are proteins involved in the separation cell chains), and *hag* (the gene encoding flagellin) expression are not repressed (Figure 10). But when the level of SlrR is high enough, the activity of SinR decreases in the cell, allowing the expression of matrix coding genes and the development of biofilms. Additionally, the SinR-SlrR complex represses the expression of *hag*, *lytABC* and *lytF,* as cell chaining is essential for the onset of biofilm formation (Chai, Kolter, & Losick, 2010) (Chai, Norman, Kolter, & Losick, 2010). The switch between the low level to the high level of SlrR is dependent on SinI-mediated inhibition of SinR, which is directly controlled by Spo0A-P (Figure 10). The SinR-SlrR switch remarkably leads to epigenetically heritable changes (Vlamakis, Chai, Beauregard, Losick, & Kolter, 2013).

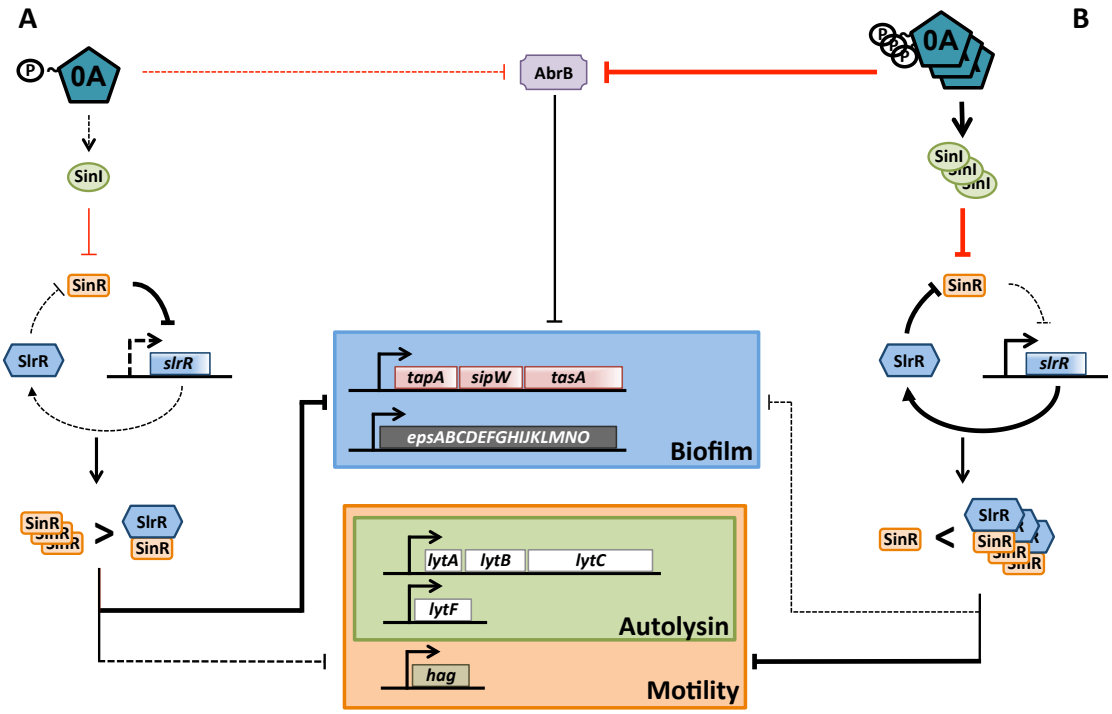


Figure 10: The SinI-SinR-SlrR switch between motile and sessile life
A- When the concentration of Spo0A-P is low, motility genes and autolysin are expressed while biofilm formation is inhibited as the *tasA-sipW-tapA* and *epsA-O* operons are repressed. Without SinI, the repression of SinR low, and therefore SlrR is not present in highe enough concentrations to inhibit SinR. So, SinR represses biofilm formation, and cells are motile **B-** The concentration of Spo0A-P increases to some threshold level, and the SinR/SinR-SlrR ratio is reversed, by lowering the amount of free SinR in the cells. The repression of matrix production is relieved, while the synthesis of autolysin and motility is inhibited allowing the development of biofilm formation.

Biofilms and the DegS-DegU pathway

As previously described, DegU is a major regulator that is involved in the regulation of many cellular adaptations such as the K-state, motility and secretion of degradative enzymes. Furthermore, a *degU* mutant cannot form normal biofilms because of the loss of the surface hydrophobicity protein BslA (Kobayashi & Iwano, 2012) (Kobayashi, 2007a) (Verhamme, Murray, & Stanley-Wall, 2009). BslA is a small extracellular protein that, with TasA and exopolysaccharide, facilitates the assembly of the matrix within a biofilm (Ostrowski, Mehert, Prescott, Kiley, & Stanley-Wall, 2011). BslA forms a hydrophobic layer on the surface of the *B. subtilis*

biofilm (Kobayashi & Iwano, 2012) (Hobley et al., 2013). Transcription of *bslA* has been shown to be indirectly activated by the transcription factor Rok (Kovacs & Kuipers, 2011), is directly inhibited by AbrB (Verhamme et al., 2009) and by an intermediate level of DegU-P during biofilm formation. Additionaly, a *degU* null mutant cannot synthesize the PGA polymer encoded by the *pgs* operon (Stanley & Lazazzera, 2005), which in some bacilli is a prominent feature of biofilms. DegU-P, in addition to its activator function at an intermediate level, can also inhibit biofilm formation when its concentration is high (above some threshold) (Verhamme et al., 2007) (Figure 8B).

Quorum Sensing in biofilms

One of the first molecules identified as an inducer of matrix gene expression was surfactin. In addition to its surfactant and anti-microbial activities, surfactin triggers the phosphorylation of Spo0A through the membrane-localized kinase KinC. Here, instead of responding to direct binding of surfactin, KinC is activated by the function of the molecule. Surfactin is a lipopeptide that inserts into the membrane and results in potassium leakage, which activates KinC by an unknown mechanism and leads to the expression of matrix genes (Lopez et al., 2010). Surprisingly, surfactin is only produced by a sub-population of cells, and the cells that produce surfactin do not respond to the molecule. Since the surfactin producers do not respond to the signal they make and the signal is unidirectional, the surfactin production in *B. subtilis* can be described as a paracrine signaling system (Lopez, Vlamakis, Losick, & Kolter, 2009). It contrasts with previously described quorum-sensing systems where every cell in a population is thought to produce and respond to the signaling molecule.

Making a biofilm, a hallmark of undomesticated *B. subtilis* strains

As many laboratories work with the domesticated strain 168 or its derivatives, the ability of *B. subtilis* to form robust biofilms has only been studied and characterized at the beginning of this century (Branda et al., 2006). This work has been carried out in the "wild" strain 3610, a strain believed to be the closest relative of the laboratory strain 168 as these two strains share a great similarity in genome sequence. A few mutations that are responsible for the phenotypic differences between these two strains have been identified. Interestingly, single point mutations in four genes (*sfp*, *epsC*, *swrA* and *degQ*) and the lack of the *rapP* gene are responsible for the defect in matrix production in the domesticated strain. The *sfp* gene encodes a broad-substrate-specificity phosphopantetheinyl transferase, which is involved in the production of surfactin and a point mutation in strain 168 impairs its function. The second point mutation is in *epsC*, which results in a decrease of exopolysaccharide synthesis. The third mutation contributing to the impairment of biofilm formation is located within *swrA*, which is known to be essential for swarming motility and poly-γ-polyglutamic acid synthesis (Stanley & Lazazzera, 2005). The last point mutation is in the promoter sequence of *degQ*, a protein involved in the regulation of the DegS-DegU pathway that is important for the regulation of biofilm formation (Verhamme et al., 2009) (Marlow et al., 2014). As noted above, DegQ facilitates the transfer of a phosphoryl group from DegS-P to DegU. The regulatory gene *rapP*, encoded on the plasmid pLS32 is no longer present in strain 168. RapP has also recently been shown to be involved in the regulation of biofilm formation (Omer Bendori, Pollak, Hizi, & Eldar, 2015) (Parashar, Konkol, Kearns, & Neiditch, 2013) is unclear.

THESIS PURPOSE

It is well known that all the developmental adaptations (competence, sporulation, biofilm formation) in *B. subtilis* share regulatory proteins (Spo0A, DegS/DegU, SinI/SinR...). These adaptations have so far been studied in optimal media with selected *B. subtilis* strains such as the laboratory strain 168 for the study of competence, and the undomesticated 3610 strain for the study of biofilm formation. In a given environment, while all the cells have the same genome, some cells will induce one adaptation while others will develop another. Because biofilms contain sporulating and non-sporulating cells, and presumably competent cells as well, they appear to be the perfect context in which to study the relationships between all the environment adaptations in *B. subtilis*. It is now clear that strain 168 has been selected over decades for its high expression of competence for transformation, loosing at the same time its ability to form biofilms. To study competence development in a biofilm context it is therefore important to first understand competence regulation in undomesticated isolates of *B. subtilis*.

In what follows I will refer to the state of competence for transformation as the K-state because it is induced by the transcription factor ComK. The use of this term serves as a reminder that this state of gene expression involves more than just transformation and is a state of dormancy distinct from spores that confers tolerance to antibiotics.

I began this scientific exploration by observing that most of the undomesticated, biofilm forming strains of *Bacillus subtilis* are poorly transformable: (i) the number of cells entering into the K-state is lower in undomesticated strains compared to the laboratory strain 168; (ii) the ensemble rate of transcription of ComK, the master regulator of competence, is significantly lower in wild isolates (PS216 and 3610) than in the reference strain 168 while the uptick (explained below)

is the same in these three backgrounds. This difference suggests that an important aspect of regulation has been lost by domestication and that our understanding of K-state control is therefore incomplete. This work is devoted to correcting this deficiency.

The main results of this study are presented here in three different chapters. The first chapter introduces the publication of the work we have done on the characterization of a genomic sequence of a new strain of *B. subtilis*: the PS216 strain is indeed another undomesticated strain of *B. subtilis* that has been isolated in Slovenian soil that form robust and structured biofilms.

The second chapter shows the work done on the characterization of the basic network for competence regulation in the domesticated strain 168 and two undomesticated strains of *B. subtilis*: the ancestral strain 3610 and the neo-characterized PS216 strain. By comparing the genome sequences of these three different strains and studying the effect of competence gene knock outs on the expression of transformability and the K-state, we concluded that the core circuitry for competence development is identical in domesticated and undomesticated strains. We also confirmed that, in comparison to PS216 and 3610, the laboratory strain 168 carries in its genome a previously identified point mutation in the promoter of *degQ* that will be of great interest in this study.

In the third chapter we show that this point mutation is involved in the increased level of ComK in the undomesticated strain 168 and that its effect is mediated by regulation of the phosphorylation state of DegU-P, a response regulator known to be involved in the control of motility and biofilm formation. We discovered that because of this point mutation in the promoter region of *degQ*, ComK is

stabilized in the laboratory strain compared to the undomesticated strain 3610. We finally showed that, after screening other candidates, *rapP*, a gene encoded by a plasmid in strain 3610 is also involved in the regulation of competence development in this isolate, as it turns down the expression of the *srfA* operon and therefore limits the percentage of cells entering into the K-state.

CHAPTER I

Genome sequence of the *Bacillus subtilis* biofilm-forming transformable strain PS216

Genome sequence of the *Bacillus subtilis* biofilm-forming transformable strain PS216

Russell Durrett[1], Mathieu Miras[2,3], Nicolas Mirouze[2,4], Apurva Narechania[5], Ines Mandic-Mulec[6] and David Dubnau*[2]

[1] Institute for Computational Biomedicine, Weill Medical College of Cornell University, New York, NY 10021, USA

[2]Public Health Research Institute, New Jersey Medical School, Newark, NJ 07103, USA

[3]Université de Toulouse (Toulouse III), 118 route de Narbonne, 31062 Toulouse, France

[4] UMR1319 Micalis, Bat. Biotechnologie (440), I.N.R.A., Domaine de Vilvert, 78352 Jouy-en-Josas Cedex, France

[5]Sackler Institute for Comparative Genomics, American Museum of Natural History, New York, New York, USA 10024

[6]Biotechnical Faculty, University of Ljubljana, 1000 Ljubljana, Slovenia

*To whom correspondence should be addressed,

ABSTRACT

PS216, a strain of *Bacillus subtilis* isolated in Slovenia has been sequenced. PS216 is transformable and forms robust biofilms, making it useful for the study of competence regulation in an undomesticated bacterium.

RESULTS

Bacillus subtilis is the most studied Gram-positive model organism (Sonenshein, 2002 #4258). It has become apparent that the standard reference strain 168 has been modified by decades of inadvertent selection in the laboratory, thereby acquiring a high frequency of transformability and losing the ability to form biofilms (McLoon, 2011 #4080). Recently, "undomesticated" strains, notably NCIB3610, have been investigated intensively because of their ability to form robust biofilms (McLoon, #4080). However, NCIB3610 is poorly transformable, limiting its usefulness for the study of genetic competence and compromising its ability to be manipulated genetically. *B. subtilis* PS216, which was isolated in Slovenia from sandy soil, forms robust biofilms and is more transformable than NCIB3610. Based on phylogenetic analysis of three concatenated protein coding genes (*dnaJ*, *gyrA* and *rpoB*), PS216 is most closely related to *B. subtilis* ssp. *subtilis* and belongs to a clade demarcated as the putative ecotype 10 (Stefanic, #4727). Strain PS216 resides in the same quorum-sensing pherotype group as 168 (Tortosa, 2001 #1706;Stefanic, 2009 #4201).

MATERIALS AND METHODS

The genome sequence of *B. subtilis* PS216 was generated as described in (Koren, #4728). Briefly, 274 Mb of PacBio long-read data was error-corrected with 150 bp MiSeq data using the pacBioToCA pipeline, resulting in approximately 71 Mb of corrected long reads that were then assembled by the Celera assembler. This assembly contained 146 contigs, 90% of the assembly in 26 contigs larger than 42Kb.

The initial assembly yielded a total of 112 single nucleotide changes compared to the reference strain 168 (NC_000964) and 140 SNPs compared to strain NCIB3610 (NZ_CM000488). SNPs were identified in SAMtools (Li, 2009 #4726) using short read alignments generated by BWA (Li, 2009 #4725). Sequencing of PCR products confirmed four of the nucleotide changes in genes of interest (*oppD*, *comP*, *degQ*, *sigH*). Of the 112 nucleotides that differed between strains 168 and PS216, 27 were identical in sequence between PS216 and NCIB3610. These include the confirmed nucleotide changes in *degQ*, *oppD* and *sigH*.

Notably, no large plasmids were detected in PS216 such as the one present in NCIB3610 (McLoon, 2011 #4080). We used NUCmer (Kurtz, 2004 #4721) and ABACAS (Assefa, 2009 #4723) to order and orient the contigs with respect to the reference, an analysis which revealed that both the 20,521 bp ICEBs1 element (Lee, #4701) and the 134,385 bp SPß temperate bacteriophage present in 168 (Lazarevic, 1999 #4719) were missing from PS216. The absence of the latter two elements was verified by sequencing a PCR product that crossed the two insertion sites.

We anticipate that this sequence information for PS216 will facilitate comparative studies of development and physiology in *Bacillus*.

NUCLEOTIDE SEQUENCE ACQUISITION NUMBER.

This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession AQGR00000000. The version described in this paper is the first version, AQGR01000000.

ACKNOWLEDGEMENTS

We appreciate discussions with Barry Kreiswirth. This work was supported by NIH grant 5R01GM057720-43 awarded to DD and by Slovenian ARRS grants J4-3631 and JP4-116, awarded to IMM.

REFERENCES

1. **Sonenshein AL, Hoch JA, Losick R.** 2002. *Bacillus subtilis* and its closest relatives. From genes to cells. ASM Press, Washington, DC.

2. **McLoon AL, Guttenplan SB, Kearns DB, Kolter R, Losick R.** 2011. Tracing the domestication of a biofilm-forming bacterium. J. Bacteriol. **193:**2027–2034.

3. **Stefanic P, Decorosi F, Viti C, Petito J, Cohan FM, Mandic-Mulec I.** 2012. The quorum sensing diversity within and between ecotypes of *Bacillus subtilis*. Environ. Microbiol. **14:**1378–1389.

4. **Tortosa P, Logsdon L, Kraigher B, Itoh Y, Mandic-Mulec I, Dubnau D.** 2001. Specificity and genetic polymorphism of the *Bacillus* competence quorum-sensing system. J. Bacteriol. **183:**451–460.

5. **Stefanic P, Mandic-Mulec I.** 2009. Social interactions and distribution of *Bacillus subtilis* pherotypes at microscale. J. Bacteriol. **191:**1756–1764.

6. **Koren S, Schatz MC, Walenz BP, Martin J, Howard JT, Ganapathy G, Wang Z, Rasko DA, McCombie WR, Jarvis ED, Phillippy AM.** 2012. Hybrid error correction and *de novo* assembly of single-molecule sequencing reads. Nat. Biotechnol. **30:**693–700.

7. **Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup.** 2009. The sequence alignment/map format and SAMtools. Bioinformatics **25:**2078–2079.

8. **Li H, Durbin R.** 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics **25:**1754–1760.

9. **Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, Salzberg SL.** 2004. Versatile and open software for comparing large genomes. Genome Biol. **5:**R12.

10. **Assefa S, Keane TM, Otto TD, Newbold C, Berriman M.** 2009. ABACAS: algorithm-based automatic contiguation of assembled sequences. Bioinformatics **25:**1968–1969.

11. **Lee CA, Thomas J, Grossman AD.** 2012. The *Bacillus subtilis* conjugative transposon ICEBs1 mobilizes plasmids lacking dedicated mobilization functions. J. Bacteriol. **194:**3165–3172.

12. **Lazarevic V, DüsterhoÅNft A, Soldo B, Hilbert H, Mauël C, Karamata D.** 1999. Nucleotide sequence of the *Bacillus subtilis* temperate bacteriophage SPbetac2. Microbiology **145**(Pt 5):1055–1067.

CHAPTER II

Domesticated and undomesticated strains of *Bacillus subtilis* share the same basic network for competence development

The firefly luciferase: a powerful tool for gene expression studies

To further investigate if the core circuitry for competence regulation was the same in the undomesticated backgrounds compared to the reference strain 168, we looked at the effects on *comK* or *comGA* expression of knock outs (KO) of the main regulators of competence determined in the reference strain. To follow the transcriptional activity of these two genes in *B. subtilis*, we chose the luciferase gene as a reporter. In the luciferase reaction, light is emitted when the luciferase acts on the luciferin substrate. Photon emission can be detected by a properly equipped light sensitive apparatus such as a plate reader. Possibly because the luciferase half life is low, this is a powerful tool to study transcription rates *in vivo* compared to reporters like ß-galactosidase that measure he accumulation of gene product (Figure 1A) (Gould & Subramani, 1988).

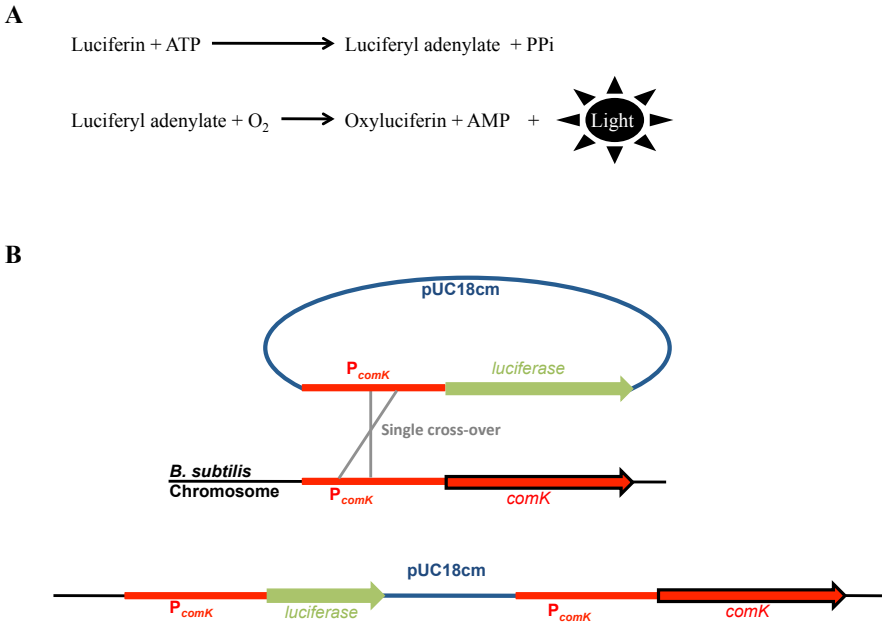


Figure 1: The luciferase as a reporter gene to measure genes rate of transcription
(A) Enzymatic reaction.
(B) Genetic organization of the gene reporter.

To study the transcriptional activity of *comK* (and *comGA*), the firefly luciferase coding sequence has been cloned downstream of the *comK* and *comGA* promoters and inserted in the chromosome by homologous recombination (Figure 1B). This was done in two natural isolates, PS216 and NCIB3610.

Effect of different competence regulators KO on K-state development

Because of the essential role of Spo0A in the regulation of competence development via the regulation of the uptick (Mirouze et al., 2012), we first examined the effect of a *spo0A* KO on the expression of *comK* and *comGA*. Figures 2A and 2B show that the rate of transcription of both *comK* and *comGA* is greatly reduced in the *spo0A* null mutants compared to the wild type, which is consistent with the known role of Spo0A in the domesticated strain.

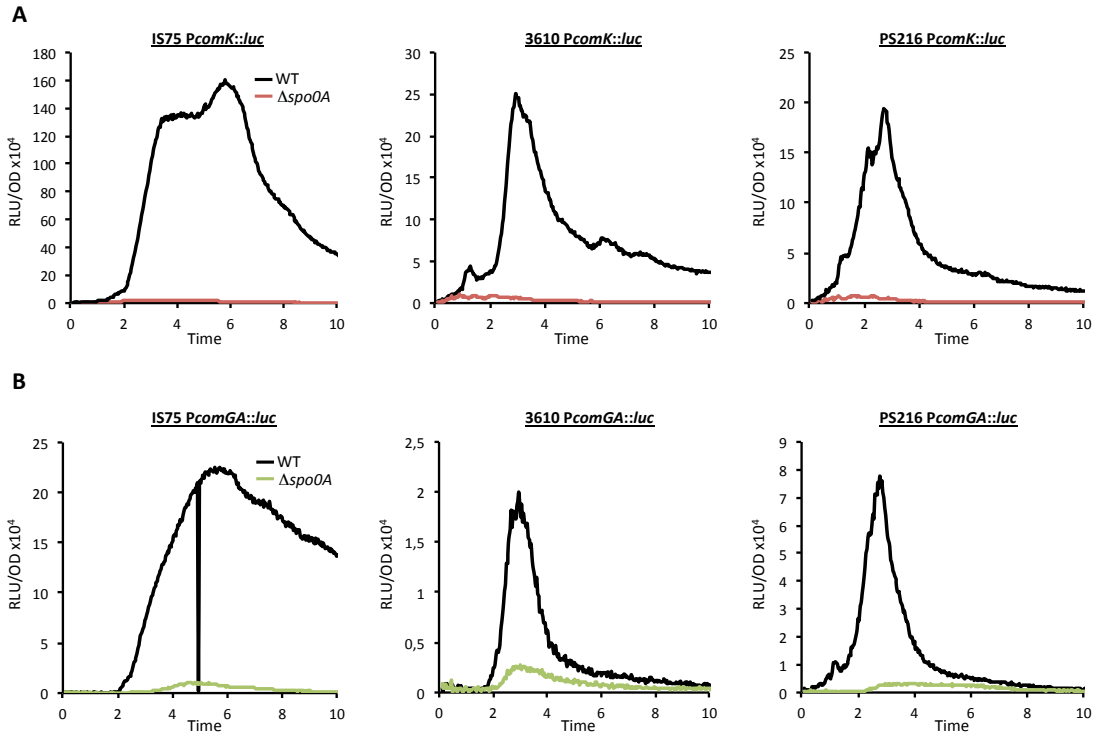


Figure 2: Effect of Δ*spo0A* on expression of *comK* (A) and *comGA* (B) in IS75, 3610 and PS216.

On the other hand, in all three backgrounds, a *rok* KO dramatically increases the transcriptions of *comK* and *comGA*, as expected for this repressor of ComK (Figures 3A and 3B).

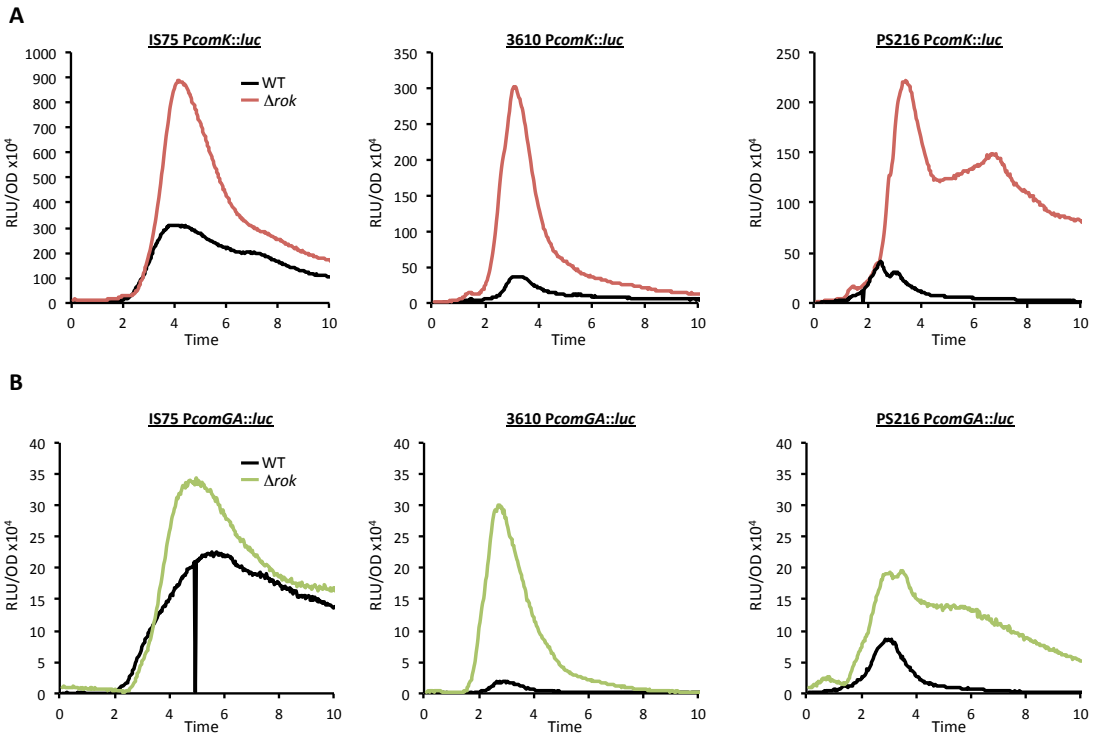


Figure 3: Effect of Δ*rok* on expression of *comK* (A) and *comGA* (B) in IS75, 3610 and PS216.

We also investigated the effect of early competence gene KOs on the expression of *comK* and *comGA*. For example a *comS* KO has a severe effect on the expression of *comGA* in the three backgrounds; in the *comS* null mutants, ComS cannot prevent ComK degradation via the MecA/ClpC/ClpP system (Figure 4B). We observed the same phenomenon when looking at the expression of *comK*: the rate of transcription of *comK* is lowered in the *comS* mutant compared to the wild type in the three strains because ComK is not available to activate its own promoter (Figure 4A). However we can distinguish some residual expression that coincides with the uptick: the *comS* KO disrupts the autoregulatory loop of ComK on P*comK* as ComK is degraded but doesn't have any effect on the uptick regulation of P*comK*.

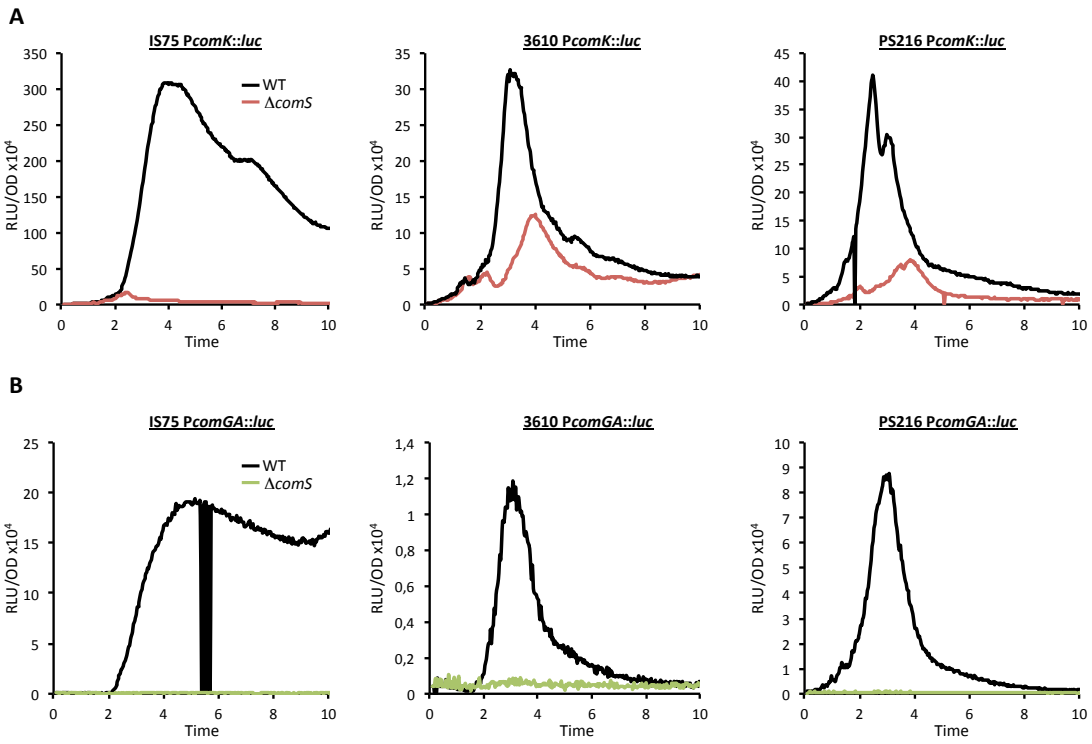


Figure 4: Effect of Δ*comS* on expression of *comK* (A) and *comGA* (B) in IS75, 3610 and PS216.

Contrary to the undomesticated strains, the laboratory strain 168 does not produce surfactin because of a mutation in *sfp* (a gene involved in surfactin production). To determine if the production of surfactin explains the lower transformability of the wild isolates compared to strain 168 we monitored the expression of *comK* and *comGA* in a surfactin null mutant. Eliminating the expression of surfactin has no effect on the expression of either *comK* or *comGA* in the undomesticated strains: the rate of transcription of our two reporter genes is comparable in the mutants and the wild types in 3610 and PS216 (Figure 5A and 5B).

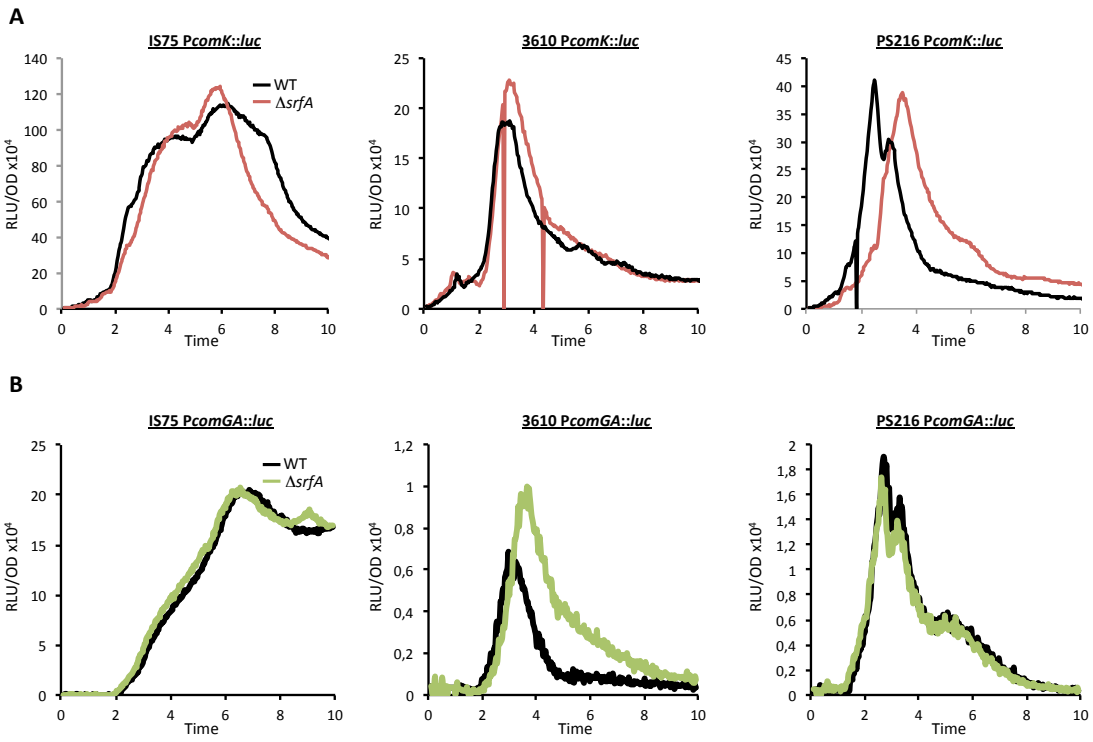


Figure 4: Effect of Δ*srfA* on expression of *comK* (A) and *comGA* (B) in IS75, 3610 and PS216.

All together, these results suggest that the basic circuitry for competence development is the same in the laboratory strain 168 and the two undomesticated strains tested, 3610 and PS216.

Strains Table

Strain Number	Background[a]	Description[b]
IS75	168 derivative	*his leu met*
BD6655	*B. subtilis* (IS75)	P*comK*::*luc* (Cm), Δ*spo0A* (Kan)
BD6656	*B. subtilis* (3610)	P*comK*::*luc* (Cm), Δ*spo0A* (Kan)
BD6657	*B. subtilis* (PS216)	P*comK*::*luc* (Cm), Δ*spo0A* (Kan)
BD6667	*B. subtilis* (IS75)	P*comGA*::*luc* (Cm), Δ*spo0A* (Kan)
BD6668	*B. subtilis* (3610)	P*comGA*::*luc* (Cm), Δ*spo0A* (Kan)
BD6669	*B. subtilis* (PS216)	P*comGA*::*luc* (Cm), Δ*spo0A* (Kan)
BD6460	*B. subtilis* (IS75)	P*comK*::*luc* (Cm), Δ*rok* (Kan)
BD6464	*B. subtilis* (3610)	P*comK*::*luc* (Cm), Δ*rok* (Kan)
BD6462	*B. subtilis* (PS216)	P*comK*::*luc* (Cm), Δ*rok* (Kan)
BD6658	*B. subtilis* (IS75)	P*comGA*::*luc* (Cm), Δ*rok* (Kan)
BD6659	*B. subtilis* (3610)	P*comGA*::*luc* (Cm), Δ*rok* (Kan)
BD6660	*B. subtilis* (PS216)	P*comGA*::*luc* (Cm), Δ*rok* (Kan)
BD6604	*B. subtilis* (IS75)	P*comK*::*luc* (Cm), Δ*comS* (Tet)
BD6605	*B. subtilis* (3610)	P*comK*::*luc* (Cm), Δ*comS* (Tet)
BD6606	*B. subtilis* (PS216)	P*comK*::*luc* (Cm), Δ*comS* (Tet)
BD6607	*B. subtilis* (IS75)	P*comGA*::*luc* (Cm), Δ*comS* (Tet)
BD6608	*B. subtilis* (3610)	P*comGA*::*luc* (Cm), Δ*comS* (Tet)
BD6609	*B. subtilis* (PS216)	P*comGA*::*luc* (Cm), Δ*comS* (Tet)
BD6652	*B. subtilis* (IS75)	P*comK*::*luc* (Cm), Δ*srfA* (Kan)
BD6653	*B. subtilis* (3610)	P*comK*::*luc* (Cm), Δ*srfA* (Kan)
BD6654	*B. subtilis* (PS216)	P*comK*::*luc* (Cm), Δ*srfA* (Kan)
BD6664	*B. subtilis* (IS75)	P*comGA*::*luc* (Cm), Δ*srfA* (Kan)
BD6665	*B. subtilis* (3610)	P*comGA*::*luc* (Cm), Δ*srfA* (Kan)
BD6666	*B. subtilis* (PS216)	P*comGA*::*luc* (Cm), Δ*srfA* (Kan)

[a]The IS75 derivatives are all *his leu met* auxotrophs.
[b]All of the fusion constructs are inserted by single crossover at the native loci.

CHAPTER III

A DegU-P and DegQ-dependent regulatory pathway for the K-state in *Bacillus subtilis*

A DegU-P and DegQ-Dependent Regulatory Pathway for the K-state in *Bacillus subtilis*

Mathieu Miras[1,2] and David Dubnau[1]*

[1]Public Health Research Institute Center, New Jersey Medical School, Rutgers University, Newark, NJ, USA,

[2]Laboratoire de Microbiologie et Génétique Moléculaires, Université de Toulouse, Toulouse, France

ABSTRACT

The K-state in the model bacterium *Bacillus subtilis* is associated with transformability (competence) as well as with growth arrest and tolerance for antibiotics. Entry into the K-state is determined by the stochastic activation of the transcription factor ComK and occurs in about ~15% of the population in domesticated strains. Although the upstream mechanisms that regulate the K-state have been intensively studied and are well understood, it has remained unexplained why undomesticated isolates of B. subtilis are poorly transformable compared to their domesticated counterparts. We show here that this is because fewer cells enter the K-state, suggesting that a regulatory pathway limiting entry to the K-state is missing in domesticated strains. We find that loss of this limitation is largely due to an inactivating point mutation in the promoter of degQ. The resulting low level of DegQ decreases the concentration of phosphorylated DegU, which leads to the de-repression

of the *srfA* operon and ultimately to the stabilization of ComK. As a result, more cells reach the threshold concentration of ComK needed to activate the auto-regulatory loop at the comK promoter. In addition, we demonstrate that the activation of *srfA* transcription in undomesticated strains is transient, turning off abruptly as cells enter the stationary phase. Thus, the K-state and transformability are more transient and less frequently expressed in the undomesticated strains. This limitation is more extreme than appreciated from studies of domesticated strains. Selection has apparently limited both the frequency and the duration of the bistably expressed K-state in wild strains, likely because of the high cost of growth arrest associated with the K-state. Future modeling of K-state regulation and of the fitness advantages and costs of the K-state must take these features into account.

INTRODUCTION

The transcription factor ComK (van Sinderen et al., 1995) directly activates more than 100 genes (Berka et al., 2002; Hamoen et al., 2002; Ogura et al., 2002). While about 20 of these mediate the uptake, processing and integration of exogenous DNA resulting in transformation (Burton and Dubnau, 2010), the roles of the remaining ~80 genes are poorly understood. Because these genes are not needed for transformation (J. Hahn, unpublished) the expression of ComK must result in phenotypes beyond competence that presumably enhance fitness. In fact, ComK also induces a period of growth arrest during which the cells that express ComK exhibit antibiotic tolerance (Nester and Stocker, 1963; Haijema et al., 2001; Johnsen et al., 2009; Briley et al., 2011; Hahn et al., 2015; Yuksel et al., 2016). This persistent state has been called the K-state, to emphasize that ComK regulates more than competence for transformation (Berka et al., 2002).

A long recognized and remarkable feature of the K-state is that it is expressed in a minor fraction of a clonal population, in a more or less all or nothing fashion (Nester and Stocker, 1963; Singh and Pitale, 1967; Hadden and Nester, 1968; Haseltine- Cahn and Fox, 1968; Maamar and Dubnau, 2005; Smits et al., 2005). Entry into the K-state is determined stochastically and studies in domesticated strains derived from the indole-requiring auxotrophic strain 168 (Spizizen, 1958) have ascribed this stochastic determination to noise in the basal expression of the *comK* promoter (Maamar et al., 2007; Süel et al., 2007). When the noisy expression of *comK* causes a subpopulation of cells to exceed a threshold level of ComK, two dimers of this protein bind cooperatively to the *comK* promoter (van Sinderen and Venema, 1994; Hamoen et al., 1998), activating a positive feedback loop and the rapid transition of these cells into the K-state, where ComK activates downstream genes.

The frequency of these activation events is extremely low during exponential growth and then rises as a culture approaches the stationary phase of growth. This temporal control has two principal causes. First, growing cultures secrete ComX, a quorum sensing pheromone, which accumulates and ultimately causes the phosphorylation of the response-regulator protein ComA (Magnuson et al., 1994). ComA-P then binds to and activates the promoter of the *srfA* operon, which encodes the small protein ComS (Nakano and Zuber, 1991; Nakano et al., 1991; Roggiani and Dubnau, 1993; Hamoen et al., 1995). ComS in turn competes with ComK for binding to the MecA-ClpC-ClpP protease, which rapidly degrades ComK during growth (Turgay et al., 1997, 1998; Prepiak and Dubnau, 2007). Stabilization occurs toward the end of exponential growth when high cell densities have produced sufficient ComX levels. A second cause of temporal regulation derives from the exquisitely controlled kinetics of *comK* basal expression (Leisner et al., 2007; Mirouze et al.,

2012). Ensemble measurements show that the average basal rate of *comK* transcription increases gradually during growth, reaches a maximum as cells depart from exponential growth (T0), and then declines. This uptick in basal expression is due to a gradual increase in the phosphorylated form of the master regulator Spo0A as cells approach stationary phase (Mirouze et al., 2012). Spo0A is phosphorylated as a consequence of a phosphorylation cascade in which several kinases transfer phosphoryl groups to Spo0F (Burbulys et al., 1991). These groups are then passed to the phosphotransfer protein Spo0B and finally to Spo0A. Low levels of Spo0A-P directly activate the basal expression of *comK*, while higher levels bind to repressive operator sites so that the rate of *comK* expression decreases (Mirouze et al., 2012). Thus, Spo0A- P opens and then closes a temporal gate for transitions to the K-state.

Although K-state regulation has been well characterized, there was reason to believe that our understanding was lacking. It has been observed that the transformability of undomesticated isolates of *Bacillus subtilis* and its close relatives is much lower than that of the domesticated derivatives of 168 (Cohan et al., 1991). In fact the model undomesticated isolate NCIB3610 (hereafter 3610), is poorly transformable, although it is very closely related to the wild parent of 168. The poor transformability is due in part to comI, a gene that is absent in 168-derivatives. Interestingly, the ComI protein appears to decrease the uptake of DNA, without affecting K-state expression (Konkol et al., 2013). Together, these observations suggest that some regulatory feature has been lost in the domesticated strains and that our appreciation of K-state regulation is consequently incomplete.

In the present study, we have shown that a known promoter mutation in *degQ* in domesticated strains (Yang et al., 1986; McLoon et al., 2011) is primarily responsible for this difference in transformability. It is of interest that this mutation

also contributes to the failure of strain 168 derivatives to form robust biofilms (McLoon et al., 2011). A consequence of this mutation is that the response regulator DegU is poorly phosphorylated in domesticated strains (Stanley and Lazazzera, 2005; Kobayashi, 2007). This deficit in DegU-P derepresses P*srfA*. This causes ComK to be stabilized, allowing more cells to pass the threshold for *comK* auto-activation, thereby increasing the fraction of K-state cells. It is known that unphosphorylated DegU is required for K-state expression because it helps ComK bind to its own promoter, thus acting as a priming protein when the ComK concentration is low (Hamoen et al., 2000). Thus, our present results show that the regulation of the K-state in undomesticated strains requires the proper ratio of phosphorylated and unphosphorylated DegU, in accordance with the view that this protein acts as a rheostat for development (Verhamme et al., 2007).

MATERIALS AND METHODS

Microbiological Methods

Bacterial strains are listed in Supplementary Table S1. The backgrounds used for all experiments were either IS75, a derivative of strain 168, PS216 (an undomesticated strain of B. subtilis isolated in Slovenia and kindly provided by Inés Mandic-Mulec) (Durrett et al., 2013), 3610Δ*com*I or 3610 *com*IQ12L (both gifts from Dan Kearns). The *com*IQ12L mutation abolishes ComI activity (Konkol et al., 2013), removing a block in DNA uptake. Constructs were introduced into IS75 by transformation (Albano et al., 1987) and into 3610Δ*com*I and PS216 by transduction using bacteriophage SPP1 (Cozy and Kearns, 2010). An exception was for the swapping of the *degQ* alleles, which was carried out by transformation, as described below. Bacterial growth was at 37◦C in competence medium (Albano et al., 1987)

unless otherwise specified. Antibiotic selections were carried out on Lysogeny Broth (LB) agar plates (Cozy and Kearns, 2010) containing ampicillin (100 μg ml−1), spectinomycin (100 μg ml−1), erythromycin (5 μg ml−1), kanamycin (Kan) (5 μg ml−1) or chloramphenicol (5 μg ml−1). In some cases selection was for erythromycin (Ery) (1 μg ml−1) plus lincomycin (20 μg ml−1). Solid media were solidified by the addition of 1.5% agar. Transformation frequencies were determined using genomic DNA isolated from a leucine prototroph.

Luciferase Assays

Light output from luciferase reporter constructs was measured as previously described (Mirouze et al., 2011). Briefly, strains were first grown in LB medium for 2 h. Cells were then centrifuged and resuspended in fresh competence medium (Albano et al., 1987), adjusting all the cultures to an OD600 of 2. These pre-cultures were then diluted 20-fold in fresh competence medium and 200 μl was distributed in duplicate in the wells of a 96-well black plate (Corning Incorporated Costar). Ten microliters of luciferin was added to each well to reach a final concentration of 1.5 mg/ml (4.7 mM). The cultures were incubated at 37◦C with agitation in a PerkinElmer Envision 2104 Multilabel Reader equipped with an enhanced sensitivity photomultiplier for luminometry. The temperature of the clear plastic lid was maintained at 38◦C to avoid condensation. Relative Luminescence Units (RLU) and OD600 were measured at 1 min intervals after two 30-s shaking steps. The data were processed using a script written in MATLAB, exported to Excel and plotted as RLU/OD versus time from the beginning of growth.

SDS-PAGE and Immunoblotting

Cell pellets were resuspended in STM buffer (50 mM Tris pH 8.0, 25% sucrose, 50 mM NaCl, 5 mM MgCl2) containing 300 μg ml−1 lysozyme and incubated at 37◦C for 5 min. The volume of STM was normalized to the turbidity measurement of the culture (determined in a Klett colorimeter) when the sample was collected. Sample buffer (final concentration of 20 mM Tris HCl pH 6.8, 10% glycerol, 1% SDS, 0.01% bromophenol blue, 2% 2-mercaptoethanol) was added to the samples, which were then incubated at 100°C for 5 min. Samples were separated by electrophoresis in 12% Tris-tricine SDS polyacrylamide gels (Schägger and von Jagow, 1987). The proteins were transferred to nitrocellulose membranes (Millipore) using a Trans Blot Turbo semidry transfer apparatus (Bio-Rad). Primary antiserum raised in rabbits against ComK was used at a dilution of 1:5000. Signal was detected using a secondary dilution of 1:10,000 goat anti- rabbit antiserum conjugated to horseradish-peroxidase (HRP) followed by visualization using Enhanced Chemiluminescence (ECL) Prime Western Blot Detection Reagent (GE Healthcare) according to the manufacturers' instructions. Images were recorded by using a Thermo Scientific MyECL Imager and band intensities were measured using ImageJ software (Schneider et al., 2012).

degQ Allele Swapping with pMiniMAD2

To swap the *degQ* promoter between IS75 and 3610 at the native loci, we used the pMiniMAD2 cloning strategy as described (Cozy and Kearns, 2010; Mukherjee et al., 2013). A 2 kb fragment containing *degQ* from strain 168 or 3610 was amplified using primers (HinDIII-degQ9: 5′GCAGCAAAGCTTCTGCGATTTCCGGATAAAAGAACATAATAATCCCAG-3′ and BamHI-degQ10: 5′-GCAGCAGGATC

CGCAGCCTGCTTCTTATATGCTGATCG-3′). The amplicons, which carried the *degQ* wild and mutant alleles near their centers, were cloned into the HinDIII and BamHI sites on pMiniMAD2, to produce the plasmids pED1932 and pED1931, which carried the wild-type and mutant *degQ* alleles, respectively. These plasmids were used to transform IS75 (with pED1932) and 3610 (with pED1931) where they integrated by single crossover events. Plasmid-free strains carrying the swapped *degQ* promoters were isolated (Cozy and Kearns, 2010; Mukherjee et al., 2013) to create IS75 *degQ*3610 (BD7454) and 3610 *degQ*IS75 (BD7445). The presence of the swapped promoters was confirmed by sequencing of PCR products amplified from the chromosome, carried out by Eton Biosciences (Union, NJ, USA).

The Phyper-spank-*degQ* Construction

The primers degQ15 (5′-TTAGTCGACAGCTAGCCACCATAC ACAATTCATTGATCTTTCA-3′) and degQ16 (5′-CTTGCAT GCGGCTAGCTACTCGTTAATCCTACTGTATACAAGGA-3′) were used to amplify a 676 bp sequence containing the *degQ* gene without its promoter. The amplicon was inserted into the Phyper-spank vector pED1870 that had been cut with NheI, using the In-Fusion HD cloning kit (Clontech, Inc.), as per the manufacturer's instructions. The resulting plasmid was integrated into the desired host strains by transformation with single crossover events, by selection for Kan resistance. pED1870 carries the Phyper-spank, lacI and lacO sequences. This and all other constructs were verified by sequencing. The resulting plasmid creates a strain in which *degQ* is under Phyper-spank at its native locus.

Microscopy

Cells were harvested, diluted into PBS (81 mM Na_2HPO_4 + 24.6 mM Na_2HPO_4 + 100 mM NaCl) and 1 μl of each culture was placed on a pad of 1% agarose made up in 0.5X TAE buffer. Images were collected using a Nikon Eclipse Ti inverted microscope equipped with an Orca Flash 4.0 digital camera (Hamamatsu), with a Nikon TIRF 1.45 NA Plan Neoflur 100 oil immersion objective. NIS-Elements AR (v 4.40, Nikon) software was used to collect and analyze images, which were then imported into Photoshop to configure the images for publication. Fluorescence intensities were determined using the automated General Analysis tool of NIS-Elements.

RESULTS

Undomesticated *Bacillus subtilis* Strains Express the K-state in Few Cells

It has been observed that the transformability of undomesticated B. subtilis strains is generally lower than that of the domesticated strain 168 and its derivatives (Cohan et al., 1991). Because the K-state is expressed bistably, this difference may reflect differing efficiencies of transformation per cell or differing frequencies of K-state cells in the population. We therefore determined the percentage of competence-expressing cells during growth in five randomly chosen natural isolates: the B. subtilis subsp. subtilis RO-OO-2 and RO-FF-1 strains, the B. subtilis subsp. spizizenii RO-E-2 [all three isolated in the Mojave desert (Cohan et al., 1991)], the commonly used model strain 3610 (Branda et al., 2006; McLoon et al., 2011) and B. subtilis PS216, isolated in Slovenia (Durrett et al., 2013). Throughout this study we have used comI mutants of 3610. ComI does not affect the expression of *comK* but does inhibit the uptake of transforming DNA (Konkol et al., 2013). In each case, after verifying that the regulatory sequences upstream of *comK* were the same in all the strains (not

shown), *comK* promoter (P*comK*) fusions to the genes encoding Cyan Fluorescent Protein (CFP) or Green Fluorescent Protein (GFP) were integrated by single reciprocal recombination and the percentages of *comK*-expressing cells were enumerated microscopically and compared to that of the 168- derivative IS75, which is the reference domesticated strain used throughout this study. We determined that the time of maximal transformability for PS216 and 3610 was attained after 240 min of growth and after 315 min for IS75 (Supplementary Figure S1). The results in Figure 1 were therefore obtained with samples taken at those times, which correspond to about T0 and T2 for the undomesticated and domesticated strains, respectively. The growth curves for all the strains under these conditions were similar (not shown). The results summarized in Figure 1 show that the samples from the undomesticated isolates contain significantly fewer competence-expressing cells (0.2–5%) than the domesticated strain sample (15.4%).

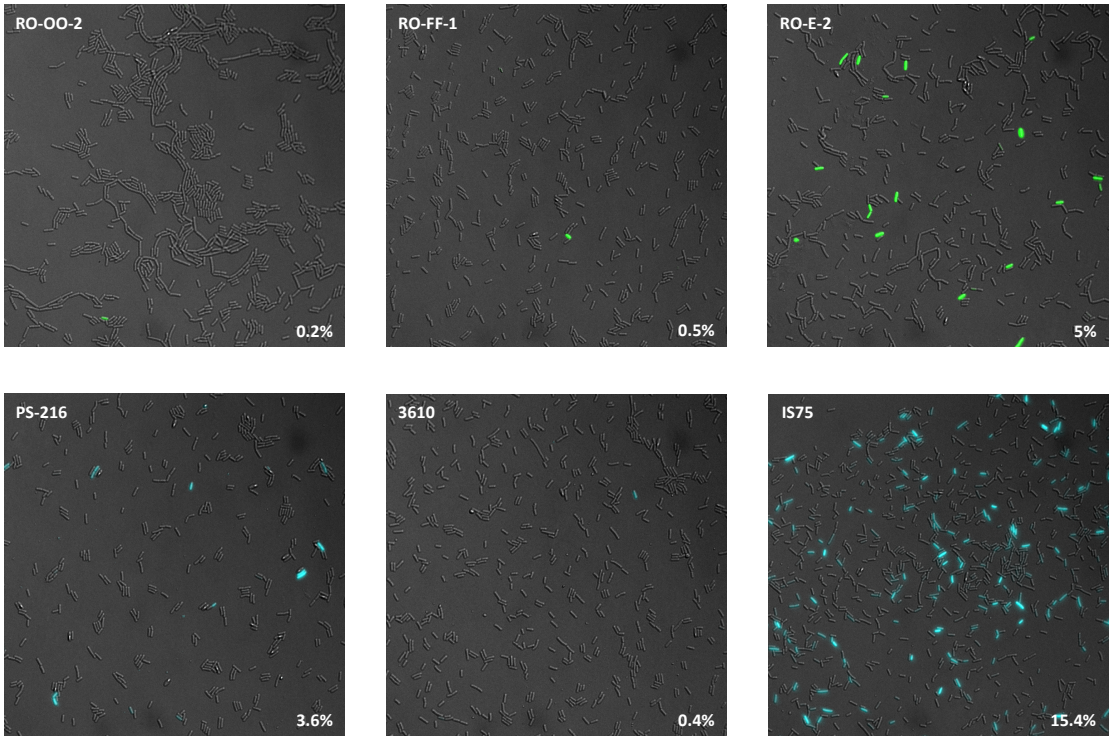


Figure 1: Undomesticated strains express the K-state with low frequency. Shown are representative images of the indicated strains expressing P*comK* fusions to the genes encoding GFP (top row) or CFP (bottom row). The fields were selected to show at least one expressing cell. The measured frequencies of K-state cells are indicated in the lower right of each panel. As explained in the text, the domesticated strain (IS75) was imaged at T2 and the undomesticated strains at T0. Strain numbers are presented in Supplementary Table S1.

To compare the transcription rates from P*comK* in strains 3610 and PS216 to that of the domesticated strain, we utilized promoter fusions to firefly luciferase, which reports transcription rates rather than the accumulation of a gene product (Mirouze et al., 2011). The fusions were integrated by single-reciprocal recombination placing the reporter gene under control of the normal *comK* regulatory sequences. The peak rates of transcription from P*comK* in the domesticated strain and in both 3610 and PS216 approximately reflect the frequencies of K-state cells in each population (Figure 2). The peak transcription rates in the undomesticated strains are about 9.6- and 5.7-fold lower than in the IS75 background, while the fractions of competence- expressing cells differ by 38- and 4.2-fold, respectively. We do not expect an exact correspondence between the peak rates of transcription and transformation frequencies; light output from luciferase does not inform us about the amount of ComK synthesized, the activation of the downstream genes needed for transformation and the assembly of the transformation machinery. In the domesticated strain, the transcription rate remains elevated even as the cultures enter stationary phase (Figure 2A). In 3610 and PS216, the rates decline rapidly from a maximum reached at about T0 (Figures 2B,C). These data demonstrate that unknown mechanisms limit the probability of transitions to the K-state and cause the expression of *comK* to cease abruptly. Interestingly, these mechanisms have been lost in the domesticated strains.

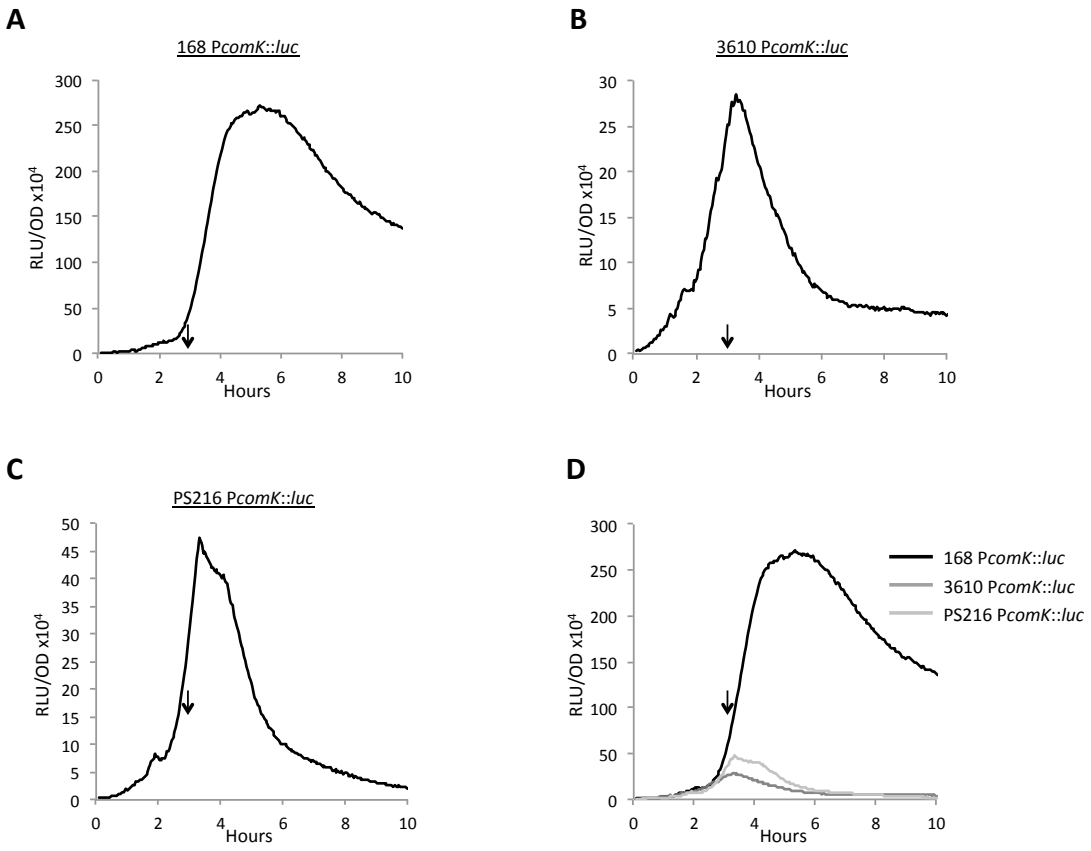


Figure 2: Expression from P*comK* is higher in IS75 (A) than in the undomesticated strains 3610 **(B)** and PS216 **(C)**. **(D)** The three curves plotted on the same scale. The vertical arrows on panels **(A–C)** point to T0. Strain numbers are presented in Supplementary Table S1.

A *degQ* Mutation in the Domesticated Strains Causes Increased *comK* Expression

To identify gene(s) that limit transitions to the K-state in undomesticated strains, we transformed PS216 with DNA from a domesticated strain that carries a fusion of the *comG* promoter to lacZ, linked to a kanamycin (Kan) resistance determinant. It was reasoned that selection for this marker on plates containing Kan and 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) would select for transformable cells and might transfer an unlinked mutation from the domesticated strain, capable of conferring increased expression from P*comG*. Indeed, dark blue colonies were observed, representing about 1% of the total number of KanR colonies, a frequency consistent with "congression," the simultaneous transformation of B.

subtilis by unlinked markers. Although PS216 colonies are normally mucoid, all of these blue colonies were not. Mucoidy in B. subtilis indicates the production of poly-γ-glutamic acid, due to the expression of the pgs operon and is dependent on the presence of phosphorylated DegU and the small protein, DegQ (Stanley and Lazazzera, 2005). DegQ increases the net transfer of a phosphoryl group from the histidine kinase DegS to its cognate response regulator protein DegU (Kobayashi, 2007). Because derivatives of strain 168 carry a promoter mutation that markedly decreases the expression of *degQ* (Yang et al., 1986; McLoon et al., 2011), we reasoned that the blue colonies might have inherited this mutation. The *degQ* promoters of several blue transformants were sequenced and indeed all had the T→C promoter mutation at position −10 that exists in 168 derivatives.

To determine whether this mutation was responsible for the high *comK* expression of domesticated strains, we swapped the wild-type and mutant *degQ* alleles between the domesticated and undomesticated strains, using pMiniMAD2. For this, and for all subsequent experiments reported here, we have used 3610, which has been widely adopted as a model undomesticated B. subtilis, rather than the less intensively studied PS216. Figure 3A shows the expression profiles of P*srfA*, P*comK*, and P*comG* luciferase fusions in the IS75 background, for strains carrying the indigenous mutant ($degQ^{IS75}$) and wild-type ($degQ^{3610}$) alleles.

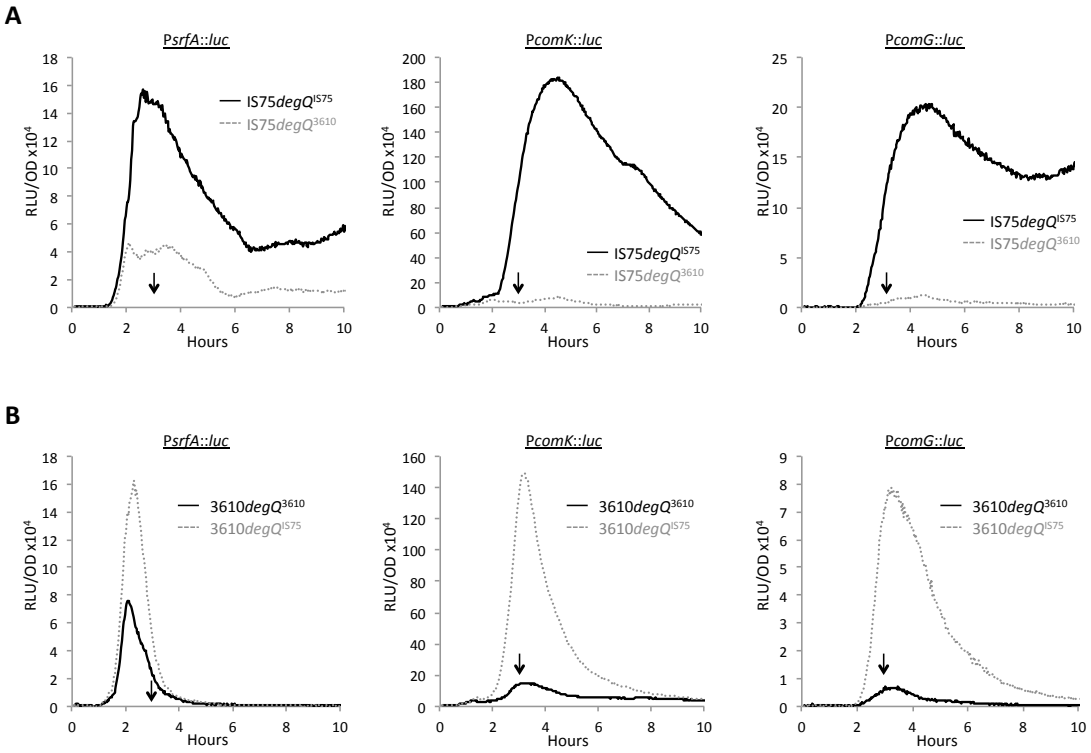


Figure 3: Effect of the mutant (*degQ*^{IS75}) and wild-type (*degQ*^{3610}) *degQ* alleles on expression of *srfA*, *comK*, and *comG* in IS75 (A) and 3610 (B). In each panel, the data from strains with indigenous and swapped alleles are indicated by solid and dashed lines, respectively. The vertical arrows in each panel point to T0. Strain numbers are presented in Supplementary Table S1.

Clearly the introduction of the wild-type *degQ* allele resulted in a marked diminution of the transcription rates of each reporter so that the peak levels were similar to those of 3610 (compare to Figure 3B). The converse was also true (Figure 3B); introduction of the *degQ*IS75 allele into the 3610 background increased the maximum transcription rates of all three promoters to approximately the levels measured in IS75 (Figure 3A). A *degQ* knockout was also tested (Supplementary Figure S2) using a P*comG*-luc reporter in the 3610 background. As expected, a large increase in *comG* transcription was noted. The effects of *degQ* allele swapping shown in Figure 3 can thus be explained most economically by a repressing effect of DegU-P on P*srfA*, which would decrease the amount of ComS, destabilizing ComK. In fact it has been shown that mutations resulting in an elevated level of DegU-P do indeed

depress the transcription of *srfA* in the domesticated background (Hahn and Dubnau, 1991).

The Sharp Decrease in Transcription in 3610 Is Not Due to DegQ

Somewhat surprisingly, the characteristic sharp decreases in the transcription rates of *srfA* and of *comK* in the 3610 background are not affected by the introduction of the mutant *degQ* promoter (Figure 3). Nor are the sustained high transcription rates from these promoters in the domesticated background obliterated when the wild-type *degQ* allele is introduced. This is seen clearly when the peak values of the swapped and un-swapped strains are normalized (Supplementary Figure S3). The comK transcription patterns can be explained by differences in *srfA* transcription if we make the simplifying assumption that ComS is unstable. Thus, when the rate of comS transcription decreases sharply, ComK would no longer be protected from degradation, and the transcription of *comK* and of *comG* would decrease. In accord with this reasoning, the transcription rates of *srfA*, *comK* and *comG* begin to drop after 2.3, 3.1, and 3.1 h of growth, respectively (Figure 3). The converse would apply in the domesticated strains; sustained transcription from the *srfA* promoter would provide a steady supply of ComS, stabilizing ComK. To test this idea, we measured the transcription rates of *srfA* in three different backgrounds. As noted above, PS216 is intermediate between 3610 and IS75 in transformability, in comK expression and in the percentage of cells that enter the K-state. Supplementary Figure S4 shows that while transcription from P*srfA* drops to zero in 3610, it decreases to an intermediate level in PS216. This can be seen clearly when the three curves are normalized to the same peak value (Supplementary Figure S4B). These comparisons are in agreement

with the hypothesis that the kinetics of *srfA* transcription underlie the strain differences in transformability.

The DegQ Effect on K-state Expression Is Mediated by Increased DegU-P

If the presence of the mutant *degQ* allele affects competence expression because of its depressing effect on DegU phosphorylation, it should be mimicked, at least in 3610, by inactivation of *degS*, which encodes the DegU-cognate histidine kinase. The results obtained in the 3610 background support a role for DegU-P as a regulator of *srfA* and hence of *comK* and *comG* expression (Figure 4). The *degS* knockout strain reaches a higher rate of *srfA* transcription than the wild-type and presumably produces more ComS (Figure 4B). Consequently ComK is stabilized, more *comK* transcription takes place and *comG* transcription is actually increased (Figure 4B). In contrast with this straightforward result, Figure 4A reveals that the inactivation of *degS* has no effect on *srfA* transcription in the IS75 background. This is not unexpected, because in this background, which is naturally lacking in DegQ, the amount of DegU-P is low and does not limit *srfA* transcription. However, the expression from comK is affected, decreasing in peak value about twofold. This relatively minor effect of *degS* inactivation can be explained as follows. Unphosphorylated DegU is known to increase the affinity of ComK for its own promoter (Hamoen et al., 2000). Because DegU-P activates one of the promoters that drive degU expression (Ogura and Tsukahara, 2010), it is likely that the inactivation of *degS* decreases the amount of DegU and thus compromises ComK binding to P*comK*. The expression from P*comG* is also affected (Figure 4A) but less than that from PcomK, suggesting that ComK is normally produced in excess.

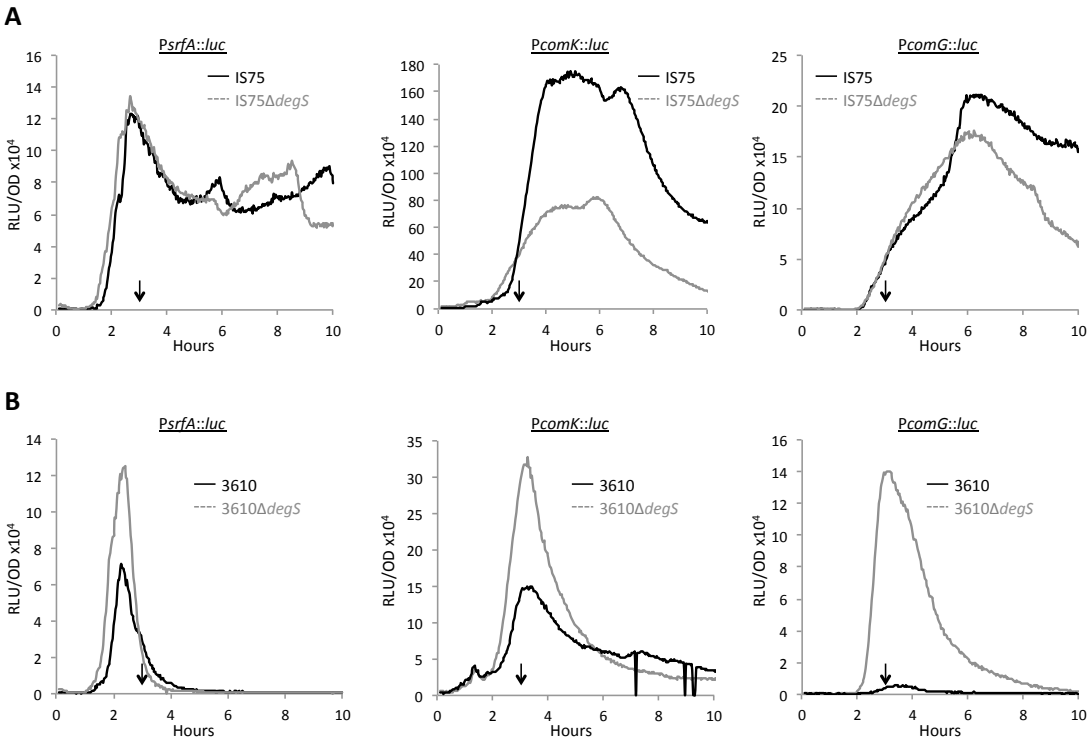


Figure 4: Effect of *degS* on expression of *srfA*, *comK*, and *comG* in IS75 (A) and 3610 (B). The vertical arrows in each panel point to T0. Strain numbers are presented in Supplementary Table S1.

Further evidence that the depressing effect on *srfA* expression in 3610 is due to phosphorylated DegU was obtained using a *degUD56N* allele. The *DegUD56N* protein cannot be phosphorylated. Supplementary Figure S5 shows that the *DegUD56N* mutant strain is de-repressed for *srfA* transcription similarly to the *degS* inactivated strain (compare to Figure 4B). As observed when the *degQ* alleles were swapped (Figure 3; Supplementary Figure S3), no effect of either the D56N or *degS* knockout mutations was observed on the sharp downturn in *srfA* transcription in the 3610 background.

DegQ Regulates the Stability of ComK

If the effects of DegU-P on P*srfA*, and hence on comS transcription underlie the differences in K-state expression, we would expect to see this reflected in measurements of ComK stability. Because the amount of ComK in the presence of

elevated DegQ is quite low, we resorted to the following strategy to investigate its stability. A strain was constructed that carries a copy of *comK* under control of the xylose-inducible Pxyl promoter inserted in the ectopic amyE locus of IS75. The native *comK* gene was left intact so that the effects of induction due to the addition of xylose would be amplified by the auto- activation of *comK* transcription. Also present was a copy of *degQ* under control of the isopropyl-β-D-thiogalactoside (IPTG)-inducible Phyper-spank promoter, located in the chromosome at the *degQ* locus. This strain was grown in the continuous presence of xylose, with and without IPTG. When late log phase was reached, rifampicin and puromycin were added to terminate mRNA and protein synthesis, samples were collected thereafter at regular intervals, and Western blotting was used to monitor the decay of ComK (Figure 5).

As predicted, the stability of ComK was markedly reduced when *degQ* was induced. In the uninduced culture the half-life of ComK is well in excess of 20 min; extrapolation would suggest it to be at least 60 min. The half-life of ComK in the presence of induced DegQ is 7– 8 min. As expected, the initial amount of ComK was also reduced in the induced culture. These data confirm that the enhanced transformability of the domesticated strain can be ascribed to the increased stability of ComK, caused by the de-repression of *srfA* (*comS*).

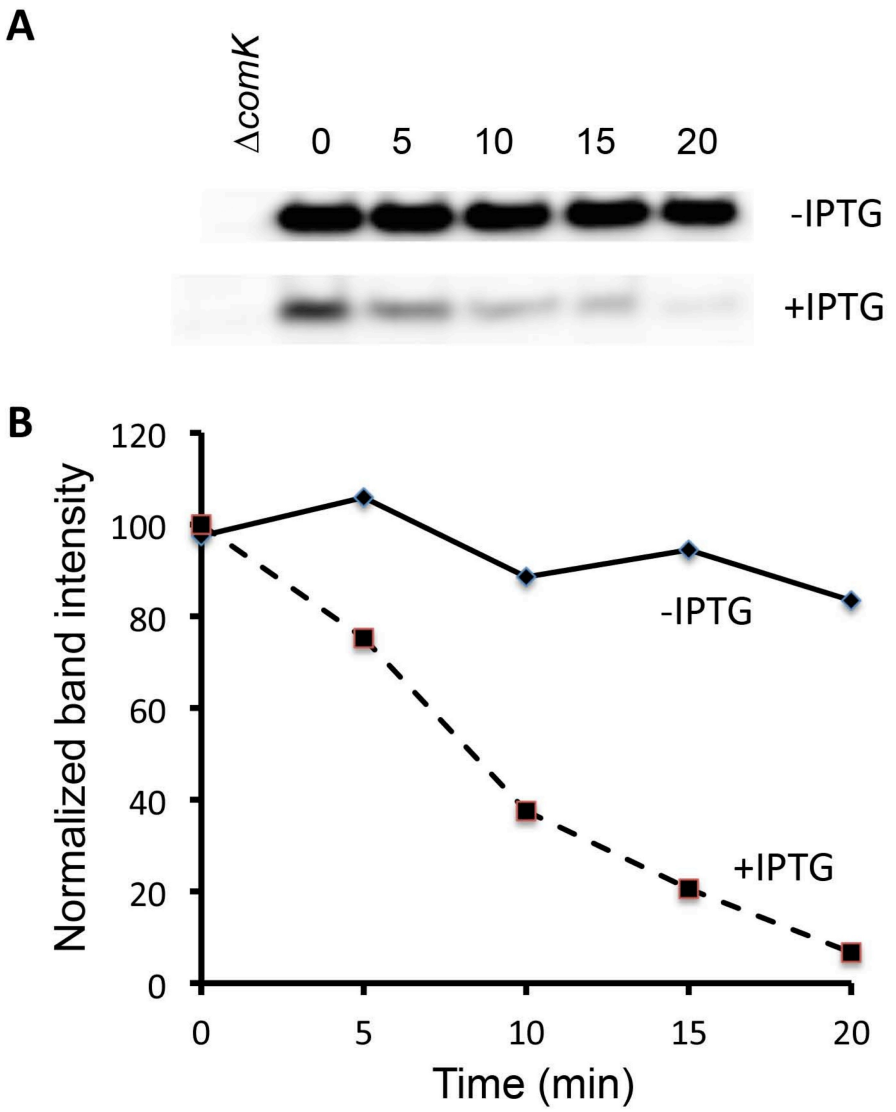


Figure 5: DegQ lowers ComK stability. A strain carrying Pxyl-*comK* and Phyper-spank-*degQ* (BD8288) was grown in the presence of xylose and with and without IPTG. At T−1, puromycin and rifampicin, each at 200 μg/ml were added. At the indicated times, samples were withdrawn for Western blotting with anti-ComK antiserum. **(A)** The -IPTG gel was exposed for a shorter time to compensate for its higher initial intensity of the ComK band. Extracts from a *comK* strain were included to assist in the identification of the correct band. The ±IPTG images were from the same gel and placed one above the other **(B)** Bands intensities were quantified from digitized images using ImageJ software and plotted against time after the addition of the inhibitors.

Single Cell Expression in *degQ*-Swapped Strains

Ensemble measurements of *srfA* and *comG* transcription show that their expression is correlated, and influenced by the levels of DegU-P. Figure 6 shows this

behavior on a single cell level. For these experiments, cells expressing promoter fusions of genes encoding mCherry and GFP to P*srfA* and P*comG*, respectively, were examined microscopically. Unlike firefly luciferase these reporters reflect the accumulation of gene products. These images were collected for the wild-type IS75 and 3610 strains as well as for strains in which the *degQ* alleles had been swapped. It is obvious that the GFP expressed from P*comG* accumulated in more cells in the IS75 background than in 3610, as expected. Also as expected, the intensity of the mCherry fluorescence expressed from P*srfA* was much higher in IS75 than in the undomesticated strain. When the $degQ^{IS75}$ allele was introduced into 3610, the expression from P*srfA* increased and the frequency of P*comG*-expressing cells also increased. Nevertheless, the mCherry fluorescence in IS75 is still more intense than that of the 3610 $degQ^{IS75}$ strain, despite the fact that the maximum rates of *srfA* expression are the same in both strains (Figure 3). This probably reflects the sustained *srfA* transcription rate in IS75 and the sharp decrease in the 3610 background. As expected, when $degQ^{3610}$ was swapped into IS75 the mCherry fluorescence decreased to a level lower than that of wild-type IS75 and even lower than that of 3610 $degQ^{IS75}$, consistent with the fourfold higher amplitude of the P*srfA* transcription rate curve in 3610 $degQ^{IS75}$ than in IS75 $degQ^{3610}$ (Figure 3). These data demonstrate that the results presented above for promoter transcription rates were reflected on the single-cell level.

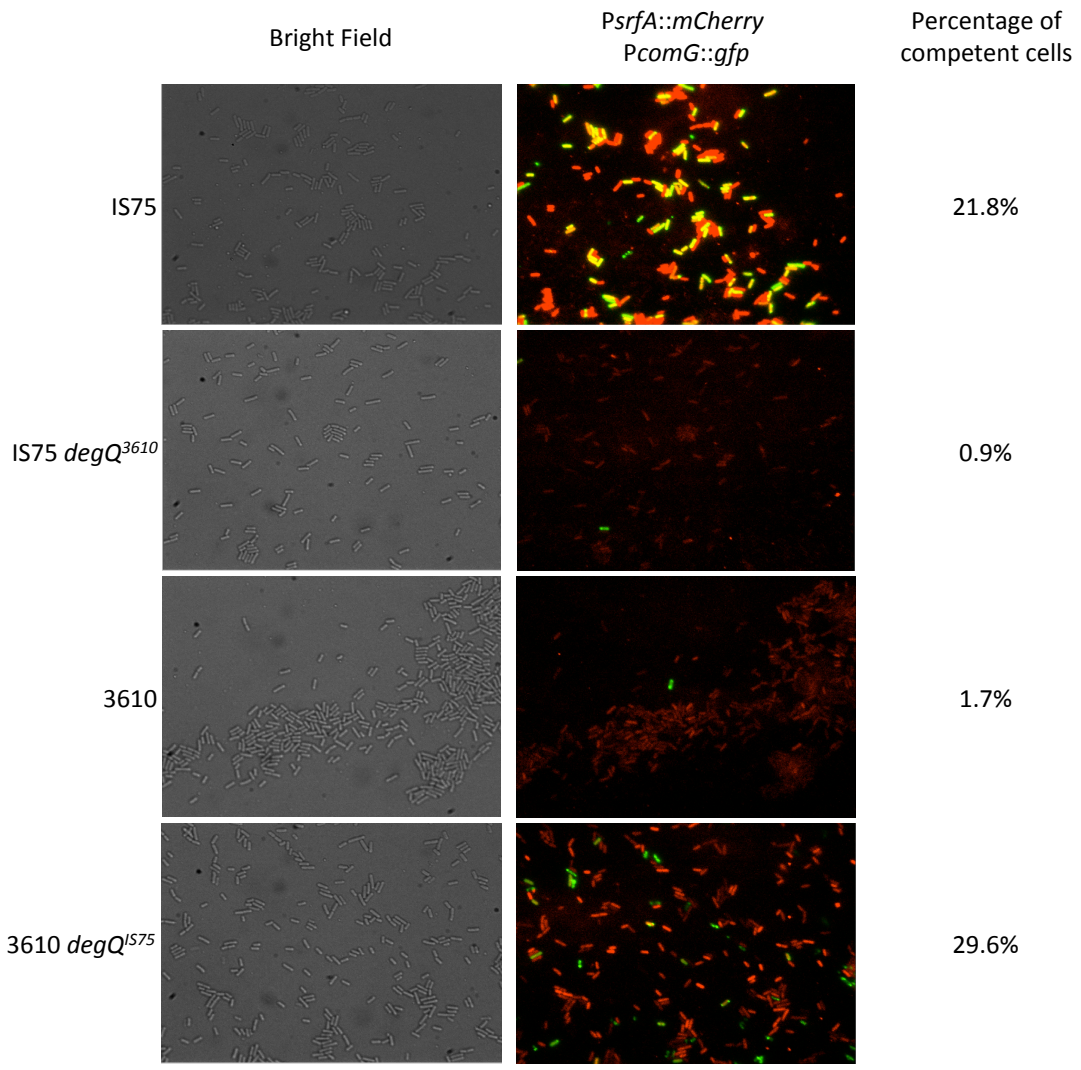


Figure 6: Single-cell expression of P*srfA*-mCherry and P*comG*-gfp. The indicated strains, all of which carried these two fusion constructs, were grown to the time of maximum K-state expression and samples were taken for microscopy. Representative images are shown. The 3610 and IS75 degQ3610 images were selected to include at least one K-state cell each. On the right are the percentages of K-state cells determined by examining at least 1200 cells for each strain. Strain numbers are presented in Supplementary Table S1.

We next sought to determine whether the K-state was expressed preferentially in cells that had accumulated more mCherry, expressed from P*srfA*. For this, we measured the average pixel intensities of mCherry in K-state cells, identified by their GFP signals, compared to the intensities of non-K- state cells. Equal numbers of K-state and non-K-state cells were selected from each microscope field to minimize inter-field bias. Table 1 shows that at the time of these measurements, K-state cells in

both the 3610 and IS75 backgrounds do not appear to exhibit a noticeably different mCherry signal than the non-K- state cells. This would suggest that variation in transcription from P*srfA* may not be an important determinant of K-state transitions. However, this conclusion must be tempered by two considerations. First, the decision to enter the K-state was made, on the average, before these measurements were made. Second, the mCherry signal does not necessarily reflect the concentration of ComS.

RapP Influences the Sharp Downturn in *srfA* Expression in 3610

As noted above (Figure 3), the sharp downturn in *srfA*, and hence in *comK* and *comG* transcription, is clearly not dependent on DegQ and remains unexplained. We have considered the possibility that the differing kinetics of *srfA* transcription in 3610 and IS75 are influenced by the absence of a functional copy of *swrA* in 168- derivatives, because SwrA has been reported to modulate the binding of DegU-P to some promoters (Ogura and Tsukahara, 2012). To test this we used pMiniMAD2 to swap the functional and mutant *swrA* alleles between the two strains. No effect on *srfA* transcription was observed (not shown). Similarly, we wondered whether surfactin, the product of the *srfA* operon, would exert an effect on the kinetics of *srfA* transcription. Accordingly, we inactivated the *srfA* operon in the 3610 strain and no effect on P*srfA* transcription was observed (not shown). Another strain difference is that 3610 carries a large plasmid, which encodes RapP, a phosphatase that acts on Spo0F-P (Parashar et al., 2013). It has been reported that the inactivation of *rapP* mitigates the sharp downturn in *srfA* transcription (Parashar et al., 2013; Omer Bendori et al., 2015). This result has been verified in our hands; we have consistently observed a shoulder in the *srfA* transcription rate curves in the *rapP* knockout strain,

and we have found that the transcription rate does not drop to zero as it does in the *rapP*+ parent (Supplementary Figure S6A). Although the inactivation of *rapP* mitigates the downturn in *srfA* transcription in 3610, it does not phenocopy the sustained transcription observed in IS75 (compare Supplementary Figure S6A with Figure 4A). Clearly other genes must be involved that differ between the domesticated and undomesticated strains.

Given the more sustained transcription of *srfA* in the 3610 *rapP* strain, we would expect the transcription of *comK* to increase due to stabilization of ComK. Instead, the transcription of *comK* was reduced (Supplementary Figure S6B). This unexpected effect was caused by a depressing effect of *rapP* inactivation on the basal level of *comK* transcription. This is shown by the expression of P*comK* when auto-regulation is abrogated due to the absence of a functional copy of *comK* (Supplementary Figure S6C). In the *comK rapP* background, the basal expression of *comK* first increases as in the *rapP*+ strain but then abruptly declines.

These data have an interesting implication; although PS216, like IS75, lacks a *rapP* ortholog, the basal level expression of *comK* is similar in amplitude and overall kinetics in IS75, 3610, and PS216 (Supplementary Figure S7). It appears that selection has used both RapP-dependent and independent mechanisms to maintain the balanced increase and decrease in basal *comK* expression.

DISCUSSION

The first important conclusion of this study is that the frequency of K-state cells in the population is controlled in natural isolates by a pathway that regulates the amount of DegU-P, providing one more illustration of the importance of using

undomesticated strains as a way to approximate real-life biology (McLoon et al., 2011). This previously unrecognized pathway for K-state regulation, summarized in Figure 7, acts by controlling the transcription of the *srfA* operon, which in turn affects the stability of ComK. Although noise in the basal expression of *comK* selects cells for competence, it appears that the instability of ComK in undomesticated strains effectively decreases the fraction of cells that exceed the threshold level of ComK needed to activate the auto-induction of *comK* transcription. This pathway is lost in 168-derivatives due to a promoter mutation that reduces the transcription of *degQ*. Since this mutation is present in all of the sequenced 168 derivatives, it must have been present in 168 itself, and is responsible for the choice of the indole- requiring 168 strain as a highly transformable subject for further investigation (Spizizen, 1958). We do not know how the elevated level of DegU-P that is present in 3610 acts to decrease the transcription from P*srfA*, nor do we know whether DegU-P acts directly on this promoter. Although *degS* and *degU* mutations that cause the accumulation of very high levels of DegU-P were known to inhibit the K-state (Msadek et al., 1990; Hahn and Dubnau, 1991), we show here for the first time that this mechanism is biologically relevant; DegU plays both positive and negative roles in regulation of the K-state in undomesticated strains. It is conceivable that the mucoidy, which can accompany increased DegU-P may itself compromise the binding of DNA to K-state cells, thereby contributing to the low transformability of undomesticated strains apart from the effect of DegU-P on P*srfA*.

Interestingly, P*degQ* is activated by ComA-P (Msadek et al., 1991), suggesting the following attractive mechanism. As the pheromone ComX accumulates, both comS and *degQ* would be transcribed. ComS would then stabilize ComK, helping to activate the auto-regulation of *comK* transcription, while DegQ

would accumulate, increasing the concentration of DegU-P and eventually shutting down *srfA* transcription. However, as shown above (Figure 3B), the sharp down turn in *srfA* transcription that occurs in 3610 is not dependent on DegQ. Instead it is partly dependent on RapP, a phosphatase that acts on Spo0F- P, presumably restricting the rate of Spo0A-P accumulation. We conclude that this pleasing model does not seem to be true, at least under laboratory conditions.

Unphosphorylated DegU binds to sequences upstream from *comK*, helping ComK bind to its own promoter (Hamoen et al., 2000). We have shown here that DegU-P plays an important role in 3610, PS216 and presumably in other natural isolates, restricting the expression of *srfA* and hence of *comK*. Because DegU and DegU-P have respective positive and negative effects on K-state expression, the ratio of their concentrations must be critically controlled. Because *srfA* is essential for biofilm formation (Lopez et al., 2009), swarming motility (Kearns et al., 2004) and surface spreading (Kinsinger et al., 2003), the balance of phosphorylated and unphosphorylated DegU is also important for these forms of development, beyond the role of DegU-P in activating the expression of the hydrophobin BslA for biofilms (Hobley et al., 2013). In particular, swarming requires DegU-P (Verhamme et al., 2007). Biofilm formation is likewise inhibited when the amount of DegU-P is too high (Kobayashi, 2007; Verhamme et al., 2007). The mechanisms and upstream signals that control the level of DegU phosphorylation are not clear, but are certainly complex (Jers et al., 2011; Cairns et al., 2015) and important to elucidate.

As described above, Spo0A phosphorylation sets a temporal gate, opening and closing a window of opportunity for transitions to the K-state. Although phosphorylated and unphosphorylated DegU also play both positive and negative roles in the K-state, they may not control a temporal gate, because neither the

inactivation of *degS* (Figure 4B) nor introduction of the non- phosphorylatable mutant form of DegU (D56N) (Supplementary Figure S5) affect the sharp downturn in *srfA* transcription in 3610. Perhaps instead, the ratio of DegU and DegU-P is simply adjusted in response to upstream signals to help set the probability of K-state development.

The inactivation of the RapP phosphatase in 3610 markedly decreases the basal expression from P*comK*, thus reducing transitions to the K-state (Supplementary Figure S6). It is possible that RapP has an activity aside from its ability to dephosphorylate Spo0F-P and that an unknown gene in this strain down-regulates the basal transcription of *comK* in the absence of RapP. However, because Spo0A-P is known to repress the basal expression of *comK*, it seems more likely that in the absence of RapP, excess phosphorelay activity increases the production of Spo0A-P and represses P*comK*. This is consistent with the kinetics of the basal expression shown in Supplementary Figure S7, which displays an initial rise identical to that in the rap+ strain, followed by sharp repression. However, regardless of the presence or absence of *rapP*, the basal expression of *comK* in IS75, PS216, and 3610 are similar (Supplementary Figure S7). Apparently, additional factors in IS75 and PS216 must serve to control the phosphorelay in the absence of RapP. Selection, whether in the laboratory or in nature, precisely modulates the formation of Spo0A-P, probably because of the multiple and dramatic consequences imposed by the presence of too much or too little of this key transcription factor. We have shown elsewhere that 168-derivatives carry a mutation in *sigH*, which decreases the activity of the phosphorelay compared to that of 3610 (Dubnau et al., 2016). Using pMiniMAD2, we have swapped *sigH*168 into a 3610 derivative lacking *rapP* and found that the basal level of *comK* transcription is still reduced (not shown). We suggest that some other

difference between IS75 and 3610 is responsible for retarding Spo0A-P accumulation under K-state conditions.

Bacillus subtilis K-state cells are growth-arrested (Nester and Stocker, 1963; Haijema et al., 2001). The K-state appears to be a persistent state in which a trade off between the costs of growth-arrest and the expression of the competence machinery are balanced by fitness benefits due to tolerance in the face of environmental insults, e.g., exposure to antibiotics (Nester and Stocker, 1963; Johnsen et al., 2009; Hahn et al., 2015; Yuksel et al., 2016), as well as the ability to acquire useful genetic information. In the face of this trade-off, bet-hedging due to the bistable expression of the K-state, presumably helps to maximize fitness (Dubnau and Losick, 2006; Veening et al., 2008). It seems reasonable to assume that the frequency of K-state expression must be adjusted to obtain the biggest fitness advantage. Selection appears to have set the basal expression of *comK* and the kinetics of *srfA* expression to keep the frequency and timing of transitions to the K-state near an optimum. As noted above it has done this by at least two distinct mechanisms; RapP-dependent and RapP-independent. What appear to be important for fitness are the final kinetic behaviors of *srfA* and of *comK* basal expression. Although the transition frequency varies among undomesticated strains, it appears to be lower than in 168-derivatives, suggesting that the optimal bet-hedging frequency lies in a range below ~15% (Figure 1).

Not only do the expressions of *srfA*, *comK*, and *comG* decline abruptly in PS216 and 3610, but so does transformability (Supplementary Figure S1). In other words competence, as traditionally defined, is restricted in these two strains to the time of entry to stationary phase, reaching a maximum at about T0 and decreasing thereafter. Perhaps this rapid decrease is due to the turnover of competence proteins,

which might provide resources for the eventual growth of cells that are hunkered down in a persistent state. ComK expression is not terminated abruptly in IS75, so this turnover might have been overlooked. This would imply the existence of an additional K-state-related mechanism that has been lost during domestication, masking the natural transience of transformability.

AUTHOR CONTRIBUTIONS

Both MM and DD contributed to the planning and interpretation of experiments and to the writing of this manuscript. MM carried out the experimental work.

FUNDING

This work was supported by NIH grant GM057720.

ACKNOWLEDGMENTS

We thank all the members of our laboratory for constant discussions, suggestions and advice. We thank Dan Kearns and Inés Mandic-Mulec for their kind gifts of strains.

REFERENCES

Albano, M., Hahn, J., and Dubnau, D. (1987). Expression of competence genes in *Bacillus subtilis*. J. Bacteriol. 169, 3110–3117. doi: 10.1128/jb.169.7.3110-3117. 1987

Berka, R. M., Hahn, J., Albano, M., Draskovic, I., Persuh, M., Cui, X., et al. (2002). Microarray analysis of the *Bacillus subtilis* K-state: genome-wide expression changes dependent on ComK. Mol. Microbiol. 43, 1331–1345. doi: 10.1046/j. 1365- 2958.2002.02833.x

Branda, S. S., Chu, F., Kearns, D. B., Losick, R., and Kolter, R. (2006). A major protein component of the *Bacillus subtilis* biofilm matrix. Mol. Microbiol. 59, 1229–1238. doi: 10.1111/j.1365-2958.2005.05020.x

Briley, K. Jr., Prepiak, P., Dias, M. J., Hahn, J., and Dubnau, D. (2011). Maf acts downstream of ComGA to arrest cell division in competent cells of B. subtilis. Mol. Microbiol. 81, 23–39. doi: 10.1111/j.1365-2958.2011. 07695.x

Burbulys, D., Trach, K. A., and Hoch, J. A. (1991). Initiation of sporulation in B. subtilis is controlled by a multicomponent phosphorelay. Cell 64, 545–552. doi: 10.1016/0092-8674(91)90238-T

Burton, B., and Dubnau, D. (2010). Membrane-associated DNA transport machines. Cold Spring Harb. Perspect. Biol. 2:a000406. doi: 10.1101/cshperspect. a000406

Cairns, L. S., Martyn, J. E., Bromley, K., and Stanley-Wall, N. R. (2015). An alternate route to phosphorylating DegU of *Bacillus subtilis* using acetyl phosphate. BMC Microbiol. 15:78. doi: 10.1186/s12866-015-0410-z

Cohan, F. M., Roberts, M. S., and King, E. C. (1991). The potential for genetic exchange by transformation within a natural population of *Bacillus subtilis*. Evolution 45, 1383–1421. doi: 10.2307/2409888

Cozy, L. M., and Kearns, D. B. (2010). Gene position in a long operon governs motility development in *Bacillus subtilis*. Mol. Microbiol. 76, 273–285. doi: 10.1111/j.1365- 2958.2010.07112.x

Dubnau, D., and Losick, R. (2006). Bistability in bacteria. Mol. Microbiol. 61, 564–572. doi: 10.1111/j.1365-2958.2006.05249.x

Dubnau, E. J., Carabetta, V. J., Tanner, A. W., Miras, M., Diethmaier, C., and Dubnau, D. (2016). A protein complex supports the production of Spo0A-P and plays additional roles for biofilms and the K-state in *Bacillus subtilis*. Mol. Microbiol. 101, 606–624. doi: 10.1111/mmi.13411

Durrett, R., Miras, M., Mirouze, N., Narechania, A., Mandic-Mulec, I., and Dubnau, D. (2013). Genome Sequence of the *Bacillus subtilis* biofilm- forming transformable strain PS216. Genome Announc. 1:e288-13. doi: 10.1128/ genomeA.00288-13

Hadden, C., and Nester, E. W. (1968). Purification of competent cells in the *Bacillus subtilis* transformation system. J. Bacteriol. 95, 876–885.

Hahn, J., and Dubnau, D. (1991). Growth stage signal transduction and the requirements for *srfA* induction in development of competence. J. Bacteriol. 173, 7275–7282. doi: 10.1128/jb.173.22.7275-7282.1991

Hahn, J., Tanner, A. W., Carabetta, V. J. I., Cristea, M., and Dubnau, D. (2015). ComGA-RelA interaction and persistence in the *Bacillus subtilis* K-state. Mol. Microbiol. 97, 454–471. doi: 10.1111/mmi.13040

Haijema, B. J., Hahn, J., Haynes, J., and Dubnau, D. (2001). A ComGA-dependent checkpoint limits growth during the escape from competence. Mol. Microbiol. 40, 52–64. doi: 10.1046/j.1365-2958.2001.02363.x

Hamoen, L. W., Eshuis, H., Jongbloed, J., Venema, G., and van Sinderen, D. (1995). A small gene, designated comS, located within the coding region of the fourth amino acid-activation domain of *srfA*, is required for competence development in *Bacillus subtilis*. Mol. Microbiol. 15, 55–63. doi: 10.1111/j.1365-2958.1995. tb02220.x

Hamoen, L. W., Smits, W. K., de Jong, A., Holsappel, S., and Kuipers, O. P. (2002). Improving the predictive value of the competence transcription factor (ComK) binding site in *Bacillus subtilis* using a genomic approach. Nucleic Acids Res. 30, 5517–5528. doi: 10.1093/nar/gkf698

Hamoen, L. W., Van Werkhoven, A. F., Bijlsma, J. J., Dubnau, D., and Venema, G. (1998). The competence transcription factor of *Bacillus subtilis* recognizes short A/T-rich sequences arranged in a unique, flexible pattern along the DNA helix. Genes Dev. 12, 1539–1550. doi: 10.1101/gad.12.10.1539

Hamoen, L. W., Van Werkhoven, A. F., Venema, G., and Dubnau, D. (2000). The pleiotropic response regulator DegU functions as a priming protein in competence development in *Bacillus subtilis*. Proc. Natl. Acad. Sci. U.S.A. 97, 9246–9251. doi: 10.1073/pnas.160010597

Haseltine-Cahn, F., and Fox, M. S. (1968). Fractionation of transformable bacteria from competent cultures of *Bacillus subtilis* on renografin gradients. J. Bacteriol. 95, 867–875.

Hobley, L., Ostrowski, A., Rao, F. V., Bromley, K. M., Porter, M., Prescott, A. R., et al. (2013). BslA is a self-assembling bacterial hydrophobin that coats the

Bacillus subtilis biofilm. Proc. Natl. Acad. Sci. U.S.A. 110, 13600–13605. doi:10.1073/pnas.1306390110

Jers, C., Kobir, A., Sondergaard, E. O., Jensen, P. R., and Mijakovic, I. (2011). *Bacillus subtilis* two-component system sensory kinase DegS is regulated by serine phosphorylation in its input domain. PLoS ONE 6:e14653. doi: 10.1371/ journal.pone.0014653

Johnsen, P. J., Dubnau, D., and Levin, B. R. (2009). Episodic selection and the maintenance of competence and natural transformation in *Bacillus subtilis*. Genetics 181, 1521–1533. doi: 10.1534/genetics.108.09 9523

Kearns, D. B., Chu, F., Rudner, R., and Losick, R. (2004). Genes governing swarming in *Bacillus subtilis* and evidence for a phase variation mechanism controlling surface motility. Mol. Microbiol. 52, 357–369. doi: 10.1111/j.1365-2958.2004.03996.x

Kinsinger, R. F., Shirk, M. C., and Fall, R. (2003). Rapid surface motility in *Bacillus subtilis* is dependent on extracellular surfactin and potassium ion. J. Bacteriol. 185, 5627–5631. doi: 10.1128/JB.185.18.5627-5631. 2003

Kobayashi, K. (2007). Gradual activation of the response regulator DegU controls serial expression of genes for flagellum formation and biofilm formation in *Bacillus subtilis*. Mol. Microbiol. 66, 395–409. doi: 10.1111/j.1365-2958.2007. 05923.x

Konkol, M. A., Blair, K. M., and Kearns, D. B. (2013). Plasmid-encoded ComI Inhibits competence in the ancestral 3610 strain of Bacillus subtilis. J. Bacteriol. 195, 4085–4093. doi: 10.1128/JB.00696-13

Leisner, M., Stingl, K., Radler, J. O., and Maier, B. (2007). Basal expression rate of *comK* sets a 'switching-window' into the K-state of *Bacillus subtilis*. Mol. Microbiol. 63, 1806–1816. doi: 10.1111/j.1365-2958.2007. 05628.x

Lopez, D., Vlamakis, H., Losick, R., and Kolter, R. (2009). Paracrine signaling in a bacterium. Genes Dev. 23, 1631–1638. doi: 10.1101/gad.181 3709

Maamar, H., and Dubnau, D. (2005). Bistability in the *Bacillus subtilis* K-state (competence) system requires a positive feedback loop. Mol. Microbiol. 56, 615–624. doi: 10.1111/j.1365-2958.2005.04592.x

Maamar, H., Raj, A., and Dubnau, D. (2007). Noise in gene expression determines cell fate in *Bacillus subtilis*. Science 317, 526–529. doi: 10.1126/science.114 0818

Magnuson, R., Solomon, J., and Grossman, A. D. (1994). Biochemical and genetic characterization of a competence pheromone from *B. subtilis*. Cell 77, 207–216. doi: 10.1016/0092-8674(94)90313-1

McLoon, A. L., Guttenplan, S. B., Kearns, D. B., Kolter, R., and Losick, R. (2011). Tracing the domestication of a biofilm-forming bacterium. J. Bacteriol. 193, 2027–2034. doi: 10.1128/JB.01542-10

Mirouze, N., Desai, Y., Raj, A., and Dubnau, D. (2012). Spo0A~P imposes a temporal gate for the bimodal expression of competence in *Bacillus subtilis*. PLoS Genet. 8:e1002586. doi: 10.1371/journal.pgen.100 2586

Mirouze, N., Prepiak, P., and Dubnau, D. (2011). Fluctuations in spo0A transcription control rare developmental transitions in *Bacillus subtilis*. PLoS Genet. 7:e1002048. doi: 10.1371/journal.pgen.1002048

Msadek, T., Kunst, F., Henner, D., Klier, A., Rapoport, G., and Dedonder, R. (1990). Signal transduction pathway controlling the synthesis of a class of degradative

enzymes in *Bacillus subtilis*: expression of the regulatory genes and analysis of mutations in *degS* and *degU*. J. Bacteriol. 172, 824–834.

Msadek, T., Kunst, F., Klier, A., and Rapoport, G. (1991). DegS-DegU and ComP-ComA modulator-effector pairs control expression of the *Bacillus subtilis* pleiotropic regulatory gene *degQ*. J. Bacteriol. 173, 2366–2377. doi: 10.1128/jb. 173.7.2366- 2377.1991

Mukherjee, S., Babitzke, P., and Kearns, D. B. (2013). FliW and FliS function independently to control cytoplasmic flagellin levels in *Bacillus subtilis*. J. Bacteriol. 195, 297–306. doi: 10.1128/JB.01654-12

Nakano, M. M., Magnuson, R., Myers, A., Curry, J., Grossman, A. D., and Zuber, P. (1991). *srfA* is an operon required for surfactin production, competence development, and efficient sporulation in *Bacillus subtilis*. J. Bacteriol. 173, 1770–1778. doi: 10.1128/jb.173.5.1770-1778.1991

Nakano, M. M., and Zuber, P. (1991). The primary role of ComA in establishment of the competent state in *Bacillus subtilis* is to activate the expression of *srfA*. J. Bacteriol. 173, 7269–7274. doi: 10.1128/jb.173.22.7269-7274.1991

Nester, E. W., and Stocker, B. A. D. (1963). Biosynthetic latency in early stages of deoxyribonucleic acid transformation in *Bacillus subtilis*. J. Bacteriol. 86,785–796.

Ogura, M., and Tsukahara, K. (2010). Autoregulation of the *Bacillus subtilis* response regulator gene *degU* is coupled with the proteolysis of DegU- P by ClpCP. Mol. Microbiol. 75, 1244–1259. doi: 10.1111/j.1365-2958.2010. 07047.x

Ogura, M., and Tsukahara, K. (2012). SwrA regulates assembly of *Bacillus subtilis* DegU via its interaction with N-terminal domain of DegU. J. Biochem. 151, 643–655. doi: 10.1093/jb/mvs036

Ogura, M., Yamaguchi, H., Kobayashi, K., Ogasawara, N., Fujita, Y., and Tanaka, T. (2002). Whole-genome analysis of genes regulated by the *Bacillus subtilis* competence transcription factor ComK. J. Bacteriol. 184, 2344–2351. doi: 10. 1128/JB.184.9.2344- 2351.2002

Omer Bendori, S., Pollak, S., Hizi, D., and Eldar, A. (2015). The RapP- PhrP quorum-sensing system of *Bacillus subtilis* strain NCIB3610 affects biofilm formation through multiple targets, due to an atypical signal- insensitive allele of RapP. J. Bacteriol. 197, 592–602. doi: 10.1128/JB. 02382-14

Parashar, V., Konkol, M. A., Kearns, D. B., and Neiditch, M. B. (2013). A plasmid-encoded phosphatase regulates *Bacillus subtilis* biofilm architecture, sporulation, and genetic competence. J. Bacteriol. 195, 2437–2448. doi: 10.1128/ JB.02030- 12

Prepiak, P., and Dubnau, D. (2007). A peptide signal for adapter protein-mediated degradation by the AAA+ protease ClpCP. Mol. Cell 26, 639–647. doi: 10.1016/ j.molcel.2007.05.011

Roggiani, M., and Dubnau, D. (1993). ComA, a phosphorylated response regulator protein of *Bacillus subtilis*, binds to the promoter region of *srfA*. J. Bacteriol. 175, 3182–3187. doi: 10.1128/jb.175.10.3182-3187. 1993

Schägger, H., and von Jagow, G. (1987). Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. Anal. Biochem. 166, 368–379. doi: 10.1016/0003-2697(87)90587- 2

Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. Nat. Methods 9, 671–675. doi: 10.1038/nmeth. 2089

Singh, R. N., and Pitale, M. P. (1967). Enrichment of *Bacillus subtilis* transformants by zonal centrifugation. Nature 213, 1262–1263. doi: 10.1038/21312 62a0

Smits, W. K., Eschevins, C. C., Susanna, K. A., Bron, S., Kuipers, O. P., and Hamoen, L. W. (2005). Stripping *Bacillus*: ComK auto- stimulation is responsible for the bistable response in competence development. Mol. Microbiol. 56, 604–614. doi: 10.1111/j.1365-2958.2005. 04488.x

Spizizen, J. (1958). Transformation of biochemically deficient strains of *Bacillus subtilis* by deoxyribonucleate. Proc. Natl. Acad. Sci. U.S.A. 44, 1072–1078. doi: 10.1073/pnas.44.10.1072

Stanley, N. R., and Lazazzera, B. A. (2005). Defining the genetic differences between wild and domestic strains of *Bacillus subtilis* that affect poly-gamma-dl-glutamic acid production and biofilm formation. Mol. Microbiol. 57, 1143–1158. doi: 10.1111/j.1365-2958.2005.04746.x

Süel, G. M., Kulkarni, R. P., Dworkin, J., Garcia-Ojalvo, J., and Elowitz, M. B. (2007). Tunability and noise dependence in differentiation dynamics. Science 315, 1716–1719. doi: 10.1126/science.1137455

Turgay, K., Hahn, J., Burghoorn, J., and Dubnau, D. (1998). Competence in *Bacillus subtilis* is controlled by regulated proteolysis of a transcription factor. EMBO J. 17, 6730–6738. doi: 10.1093/emboj/17.22.6730

Turgay, K., Hamoen, L. W., Venema, G., and Dubnau, D. (1997). Biochemical characterization of a molecular switch involving the heat shock protein ClpC, which controls the activity of ComK, the competence transcription factor of *Bacillus subtilis*. Genes Dev. 11, 119–128. doi: 10.1101/gad.11. 1.119

van Sinderen, D., Luttinger, A., Kong, L., Dubnau, D., Venema, G., and Hamoen, L. (1995). *comK* encodes the competence transcription factor, the key regulatory

protein for competence development in *Bacillus subtilis*. Mol. Microbiol. 15, 455–462. doi: 10.1111/j.1365-2958.1995.tb02259.x

van Sinderen, D., and Venema, G. (1994). *comK* acts as an autoregulatory control switch in the signal transduction route to competence in *Bacillus subtilis*. J. Bacteriol. 176, 5762–5770. doi: 10.1128/jb.176.18.5762-5770. 1994

Veening, J. W., Smits, W. K., and Kuipers, O. P. (2008). Bistability, epigenetics, and bet-hedging in bacteria. Annu. Rev. Microbiol. 62, 193–210. doi: 10.1146/ annurev. micro.62.081307.163002

Verhamme, D. T., Kiley, T. B., and Stanley-Wall, N. R. (2007). DegU co-ordinates multicellular behaviour exhibited by *Bacillus subtilis*. Mol. Microbiol. 65, 554–568. doi: 10.1111/j.1365-2958.2007.05810.x

Yang, M., Ferrari, E., Chen, E., and Henner, D. J. (1986). Identification of the pleiotropic sacQ gene of *Bacillus subtilis*. J. Bacteriol. 166, 113–119. doi: 10. 1128/jb.166.1.113- 119.1986

Yuksel, M., Power, J. J., Ribbe, J., Volkmann, T., and Maier, B. (2016). Fitness trade-offs in competence differentiation of *Bacillus subtilis*. Front. Microbiol. 7:888. doi: 10.3389/fmicb.2016.00888

CONFLICT OF INTEREST STATEMENT

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

SUPPLEMENTAL DATA

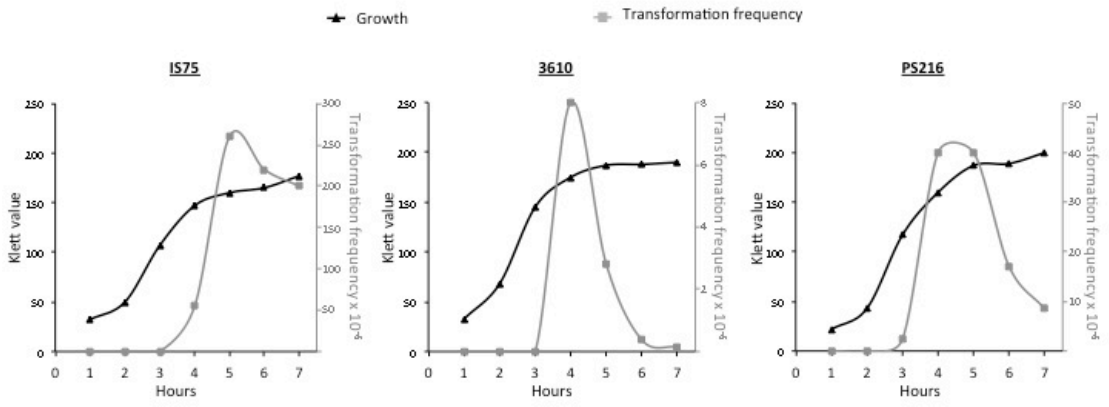


Figure S1: The timing and duration of transformation is different in the domesticated strain IS75 and the two undomesticated strains 3610 and PS216. Transformation efficiency (grey) and turbidity (black) were determined every hour for 7 hours. Note that while the turbidity scale, measured in a Klett colorimeter, is the same for the three panels (left axes), the transformation efficiencies are plotted on different scales.

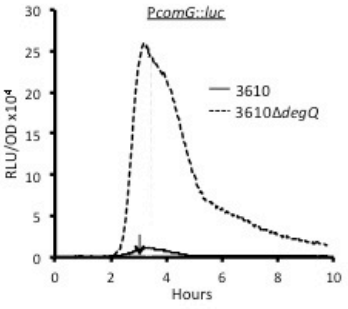


Figure S2: Effect of *ΔdegQ* on expression from P*comG* in the 3610 background. The vertical arrow points to T_0.

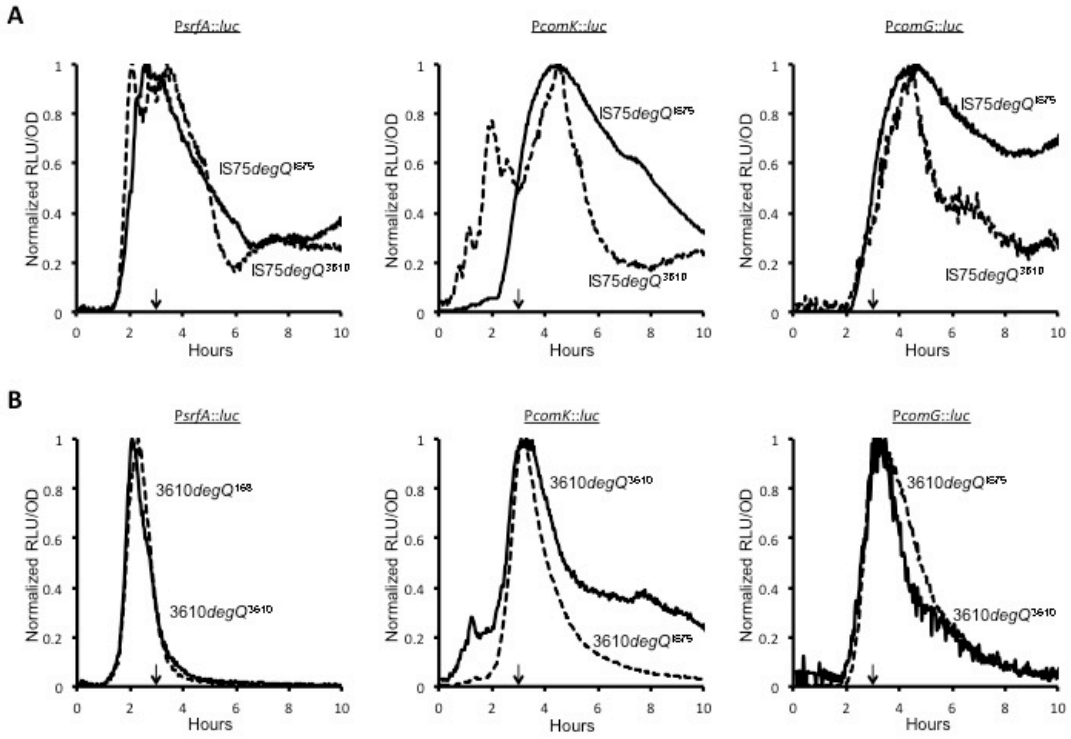


Figure S3: Normalized curves for *srfA, comK* and *comG* transcription rates in the *degQ* swapped strains in the IS75 (A) and 3610 (B) backgrounds. To more graphically compare the shapes of the rate curves, the data from Fig. 2 has been normalized by the maximum value for each curve. The vertical arrows in each panel point to T_0.

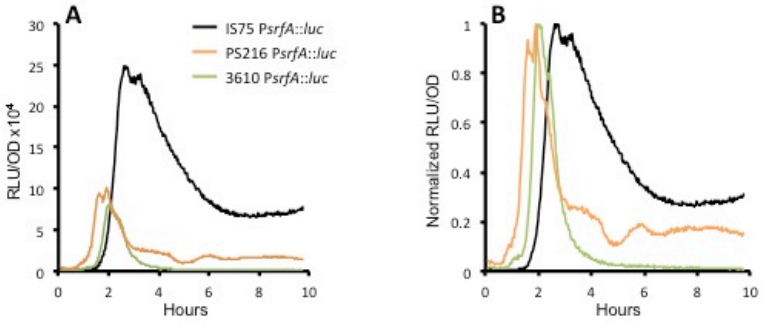


Figure S4:Transcription rate data for P*srfA* in the IS75, PS216 and 3610 backgrounds, plotted together (A). Panel B shows the data from panel A, normalized using the maximum of each curve. In this experiment, PS216 reached T_0 earlier than the other two strains.

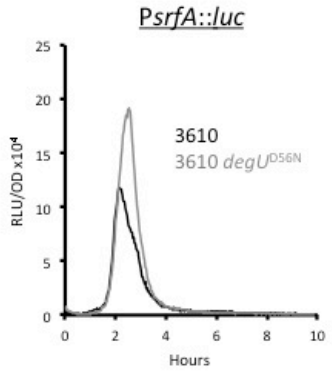


Figure S5: Transcription rates of *srfA-luc* in the 3610 wild-type and *degU*^{D56N} strains

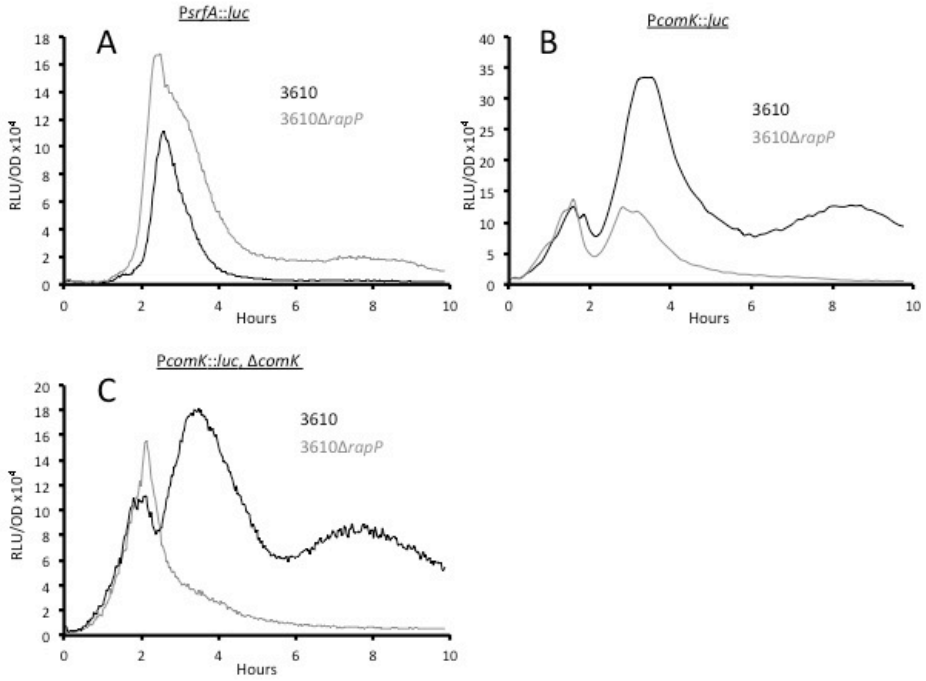


Figure S6 | Effect of the Δ*rapP* mutation on transcription rates from P*srfA*, P*comK* and P*comK* in the Δ*comK* backgrounds.

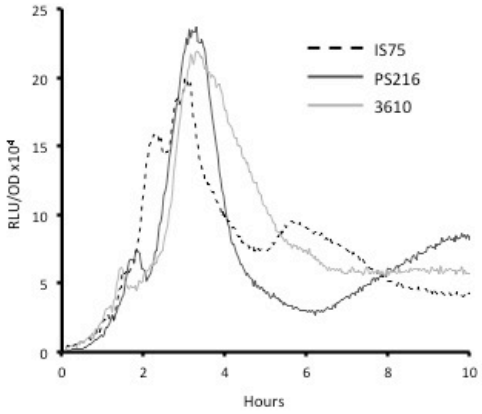


Figure S7: The basal expression from P*comK* is comparable in IS75, PS216 and 3610. All three strains carry P*comK*-*luc* and Δ*comK*.

Table S1 Strains		
Strain Number	**Background**[a]	**Description**[b]
IS75	168 derivative	*his leu met*
BD3338	*B. subtilis* RO-OO-2	P*comK*::*gfp* (Cm)
BD3339	*B. subtilis* RO-FF-1	P*comK*::*gfp* (Cm)
BD3341	*B. subtilis* RO-E-2	P*comK*::*gfp* (Cm)
BD4374	*B. subtilis* (IS75)	P*comK*::*cfp* (Kan)
BD4773	*B. subtilis* (IS75)	P*comK*::*luc* (Cm)
BD4893	*B. subtilis* (IS75)	Δ*comK* (Spc), P*comK*::*luc* (Cm)
BD6432	*B. subtilis* (PS216)	P*comK*::*luc* (Cm)
BD6434	*B. subtilis* (PS216)	P*comK*::*cfp* (Kan)
BD6437	*B. subtilis* (PS216)	Δ*comK* (Spc), P*comK*::*luc* (Cm)
BD6438	*B. subtilis* (3610)	Δ*comI*, Δ*comK* (Spc), P*comK*::*luc* (Cm)
BD6439	*B. subtilis* (3610)	Δ*comI*, P*comK*::*luc* (Cm)
BD6441	*B. subtilis* (3610)	Δ*comI*, P*comK*::*cfp* (Kan)
BD7125	*B. subtilis* (3610)	Δ*comI*, P*comG*::*luc* (Cm), Δ*degQ* (Tet)
BD7447	*B. subtilis* (3610)	Δ*comI*, PcomK::*luc* (Cm), $degQ^{IS75}$
BD7448	*B. subtilis* (3610)	Δ*comI*, P*comG*::*luc* (Cm), $degQ^{IS75}$
BD7456	*B. subtilis* (IS75)	P*comK*::*luc* (Cm), $degQ^{3610}$
BD7457	*B. subtilis* (IS75)	P*comG*::*luc* (Cm), $degQ^{3610}$
BD8276	*B. subtilis* (IS75)	P*srfA*::*luc* (Cm), $degQ^{3610}$
BD8277	*B. subtilis* (3610)	Δ*comI*, PsrfA::*luc* (Cm), $degQ^{IS75}$
BD8278	*B. subtilis* (IS75)	P*srfA*::*luc* (Cm), Δ*degS* (Kan)
BD8279	*B. subtilis* (IS75)	P*comK*::*luc* (Cm), Δ*degS* (Kan)
BD8280	*B. subtilis* (168)	P*comG::luc* (Cm), Δ*degS* (Kan)
BD8281	*B. subtilis* (3610)	Δ*comI*, P*srfA*::*luc* (Cm), Δ*degS* (Kan)
BD8282	*B. subtilis* (3610)	Δ*comI*, P*comK*::*luc* (Cm), Δ*degS* (Kan)
BD8283	*B. subtilis* (3610)	Δ*comI*, P*comG*::luc (Cm), Δ*degS* (Kan)
BD8284	*B. subtilis* (IS75)	P*srfA*::*mCherry* (Kan), P*comG*::*gfp* (Cm)
BD8285	*B. subtilis* (IS75)	P*srfA*::*mCherry* (Kan), P*comG*::*gfp* (Cm), $degQ^{3610}$
BD8286	*B. subtilis* (3610)	Δ*comI*, P*srfA*::*mCherry* (Kan), P*comG*::*gfp* (Cm)
BD8287	*B. subtilis* (3610)	*comI*, P*srfA*::*mCherry* (Kan), P*comG*::*gfp* (Cm), $degQ^{IS75}$
BD8288	*B. subtilis* (IS75)	P*xyl-comK* (Ery), P*hyper-spank-degQ* (Kan)
BD8289	*B. subtilis* (3610)	Δ*rapP*, P*srfA*::*luc* (Cm)
BD8290	*B. subtilis* (3610)	Δ*rapP*, P*comK*::*luc* (Cm)
BD8291	*B. subtilis* (3610)	Δ*rapP*, Δ*comK* (Spc), P*comK*::*luc* (Cm)
BD8292	*B. subtilis* (3610)	Δ*comI*, P*srfA*::*luc* (Cm), *degUD56N* (Kan)

[a]The IS75 derivatives are all *his leu met* auxotrophs. [b]All of the fusion constructs and the P*hyper-spank-degQ* construct are inserted by single crossover at the native loci. The P*xyl-comK* construct is at the *amyE* locus.

THESIS SUMMARY

CONCLUSION

We have presented data supporting the idea that domestication led to the loss of a regulation pathway for the K-state in *B. subtilis*. The standard laboratory strains of *B. subtilis*, derived from strain 168, have been selected during more than fifty years of use for ease of manipulation as useful genetic models. This selection is responsible for (i) the high frequency of transformability and (ii) the loss of ability to form biofilms (McLoon et al., 2011). The inability to form biofilms is attributed to five genomic mutations/deletions. *sfp*, *epsC*, *swrA* and *degQ* are mutated while the extrachromosomally encoded *rapP* gene is missing due to loss of the pLS32 plasmid (McLoon et al., 2011).

Over the past decade, the undomesticated strain 3610, considered as the parent of strain 168, has been intensively investigated because of its ability to form biofilms and is widely viewed as an "undomesticated model" of *B. subtilis*. Although 3610 is genetically really closely related to the domesticated 168 derivative, IS75, it is poorly transformable. This phenotype for DNA transformation is partially due to the presence of *comI*, a gene encoded by pLS32, a plasmid absent from the laboratory strain. (Konkol et al., 2013) showed that ComI decreases the uptake of DNA without interfering with the expression of the K-state. Because 3610 expresses the K-state poorly, these observations suggest that the domesticated strain IS75 has lost some regulatory feature of competence and that, therefore, our understanding of the K-state regulation in *B. subtilis* is incomplete.

The undomesticated B. subtilis PS216 was isolated from a Slovenian soil sample and has been shown to form structured and robust biofilms. Also, it exhibits a transformation frequency and K-state expression intermediate between that of IS75 and 3610. All three strains, IS75, 3610 and PS216 belong to the same quorum-sensing phenotype group and therefore respond to the same ComX quorum sensing

pheromone (Tortosa et al., 2001) (Stefanic & Mandic-Mulec, 2009). We have characterized the genomic sequence of PS216, and have attempted to understand the differences in frequency of K-state expression in the three strains in order to uncover the regulatory features that are missing in IS75.

Our initial analysis confirmed, unsurprisingly, that the core circuitry for competence regulation is the same in both domesticated and undomesticated strains of *B. subtilis* as inactivation of key competence regulators, e.g. *comK*, *mecA*, *comA* and *comS*, has the same effect on K-state expression in the three backgrounds.

We then designed a genetic approach to identify genes reducing K-state development in undomesticated strains of *B. subtilis*. The undomesticated *B. subtilis* PS216 strain was transformed with chromosomal DNA from strain IS75 carrying a *comG* promoter *lacZ* fusion linked to a kanamycin-resistance marker. PS216 was used because it is more transformable than 3610. We selected for kanamycin resistance and screened for blue colonies as an indication that an unlinked allele from IS75 that increases expression from the *comG* promoter had been transferred by "congression" (simultaneous transformation of *B. subtilis* by unlinked markers). 1% of the kanamycin resistant colonies developed a blue color, a frequency consistent with congression in *B. subtilis*. Although PS216 colonies are normally mucoid, all the blue colonies were non-mucoid. This difference in mucoidy expression, pointed to a difference in the *B. subtilis* "Deg" regulatory system. Indeed, the mucoidy phenotype in *B. subtilis* indicates the production of poly-γ-glutamic acid through the expression of the *pgs* operon, that is dependent on DegU~P and DegQ (Stanley & Lazazzera, 2005). As previously reported, one of the genetic differences between domesticated 168 strains (e.g. IS75) and the undomesticated strains 3610 and PS216 is a mutation in the promoter region of *degQ*, which negatively affects the transcription of *degQ*

(Yang, Ferrari, Chen, & Henner, 1986) (McLoon et al., 2011). All the PS216 blue colonies were then shown to have the *degQ* promoter mutation from strain IS75 suggesting that the "Deg" pathway down regulates K-state regulation in undomesticated strains of *B. subtilis*.

To confirm the role of the *degQ* promoter mutation in K-state regulation we swapped the *degQ* allele between strains IS75 and 3610. The rates of transcription of *srfA*, *comK* and *comG* were all increased dramatically in strain 3610 with the mutant *degQ* allele and decreased in strain IS75 with the wild-type *degQ* allele. We also showed that the *degQ* swap has an effect on the percentage of competent cells as the number of cells entering the K-state is increased or decreased in the swapped strains compared to the wild types in the 3610 and IS75 backgrounds, respectively.

degQ is described in the literature as a modulator of the phosphorylation state of DegU, which increases the transfer of a phosphoryl group from the histidine kinase DegS to its cognate response regulator DegU (Kobayashi, 2007b). To determine if the differences in the K-state previously observed is a consequence of a different level of DegU~P in the cells, we looked at the effect of a *degS* KO on the expression of competence genes (*srfA*, *comK* and *comGA*). These experiments confirmed that the *degQ* effect on the K-state expression is mediated by the expected changes in DegU~P.

We then tried to identify how the level of DegU~P can modulate the expression of the K-state in *B. subtilis*. Our lab has shown that a stable form of DegU~P, which is dephosphorylated at a slow rate, decreases the expression of the early *srfA* operon, suggesting that a high level of DegU~P might inhibit *srfA* expression. Because the *comS* open reading frame is embedded in the *srfA* operon, and because ComS protects ComK from degradation by the ATP-dependent

MecA/ClpC/ClpP protease, we hypothesized that an increased level of DegU~P would decrease ComK stability by repressing the expression of *srfA* (and therefore *comS*). Data presented in this work confirm that the enhanced transformability of the domesticated strain IS75 can be explained by the increased stability of ComK, caused by the derepression of *srfA* (*comS*).

We have shown here that DegU~P plays an important role in the undomesticated strains of *B. subtilis* by inhibiting the expression of *srfA* and therefore the K-state. It was shown previously that unphosphorylated DegU binds to sequences upstream from ComK helping ComK to bind to its own promoter at the onset of competence development. Because DegU and DegU~P have both positive and negative effects on the expression of the K-state it appears that the ratio of their concentrations must be tightly regulated. The different mechanisms and upstream signals controlling the phosphorylation of DegU are not yet fully understood. Our results revealed that a relatively high level of DegU~P inhibits the K-state by repressing the expression of *srfA*. The molecular mechanism of this repression has not been characterized. Indeed, while we know DegU is able to directly bind DNA, we don't know if DegU~P is able to bind directly to the *srfA* promoter region to inhibit its expression.

Competence for DNA transformation in *B. subtilis* is a consequence of noise in the basal expression of ComK. Competent cells are produced when the threshold level of ComK needed to activate the auto induction of *comK* transcription is exceeded by basal (ComK-independent) *comK* transcription. This basal level of *comK* transcription is set by changing amounts of Spo0A~P. Here we have shown that the probability that a cell will transition to the K-state is also determined by the ratio of DegU to DegU-P, which determine the binding affinity of ComK to its own promoter

and the stability of ComK, respectively. This work highlights a previously unknown pathway for K-state regulation: the control of the transcription of the *srfA* operon through DegU~P affects the stability of ComK which in turn affects the fraction of cells entering the K-state in undomesticated strains of *B. subtilis*. The working model of this pathway is presented in the figure X.

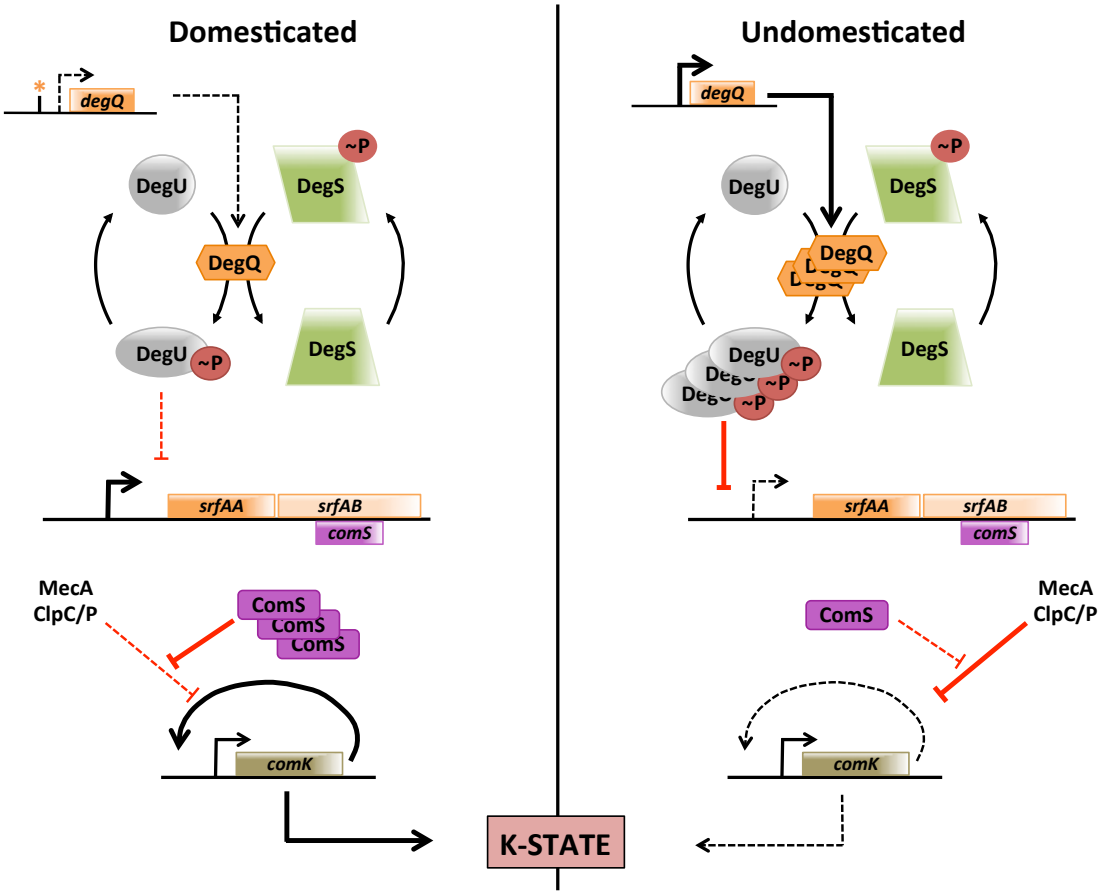


Figure X: K-state regulation in domesticated and undomesticated strains of *B. subtilis*.

In domesticated strains of *B. subtilis*, a mutation in the promoter region of *degQ* decreases the rate of transcription of *degQ* (Yang et al., 1986). This low level of DegQ decreases the phosphorylation of DegU through its cognate kinase DegS which inhibits the repression effect of DegU~P at the *srfA* promoter level. As a result, ComS is synthetized and inhibits the degradation of ComK through the proteolytic complex

MecA/ClpC/ClpP, allowing ComK to bind to its own promoter and develop competence. On the other hand, the undomesticated strains of *B. subtilis* do not carry the mutation in the *degQ* promoter region resulting in a higher synthesis of DegQ resulting in the accumulation of DegU~P. As a result, *srfA* transcription is repressed and so is competence expression.

This work provides one more illustration of the importance of using wild isolates for the study of real-life biology.

REFERENCES

Albano, M., Breitling, R., & Dubnau, D. A. (1989). Nucleotide sequence and genetic organization of the *Bacillus subtilis* comG operon. *J Bacteriol, 171*(10), 5386-5404.

Albano, M., & Dubnau, D. A. (1989). Cloning and characterization of a cluster of linked *Bacillus subtilis* late competence mutations. *J Bacteriol, 171*(10), 5376-5385.

Albano, M., Hahn, J., & Dubnau, D. (1987). Expression of competence genes in Bacillus subtilis. *J Bacteriol, 169*(7), 3110-3117.

Alm, E., Huang, K., & Arkin, A. (2006). The evolution of two-component systems in bacteria reveals different strategies for niche adaptation. *PLoS Comput Biol, 2*(11), e143. doi:10.1371/journal.pcbi.0020143

Ansaldi, M., Marolt, D., Stebe, T., Mandic-Mulec, I., & Dubnau, D. (2002). Specific activation of the *Bacillus* quorum-sensing systems by isoprenylated pheromone variants. *Mol Microbiol, 44*(6), 1561-1573.

Avery, O. T., Macleod, C. M., & McCarty, M. (1944). Studies on the Chemical Nature of the Substance Inducing Transformation of *Pneumococcal* Types : Induction of Transformation by a Desoxyribonucleic Acid Fraction Isolated from *Pneumococcus* Type Iii. *J Exp Med, 79*(2), 137-158.

Berka, R. M., Hahn, J., Albano, M., Draskovic, I., Persuh, M., Cui, X., . . . Dubnau, D. (2002). Microarray analysis of the *Bacillus subtilis* K-state: genome-wide expression changes dependent on ComK. *Mol Microbiol, 43*(5), 1331-1345.

Blair, K. M., Turner, L., Winkelman, J. T., Berg, H. C., & Kearns, D. B. (2008). A molecular clutch disables flagella in the *Bacillus subtilis* biofilm. *Science, 320*(5883), 1636-1638. doi:10.1126/science.1157877

Branda, S. S., Chu, F., Kearns, D. B., Losick, R., & Kolter, R. (2006). A major protein component of the *Bacillus subtilis* biofilm matrix. *Mol Microbiol, 59*(4), 1229-1238. doi:10.1111/j.1365-2958.2005.05020.x

Branda, S. S., Gonzalez-Pastor, J. E., Ben-Yehuda, S., Losick, R., & Kolter, R. (2001). Fruiting body formation by *Bacillus subtilis*. *Proc Natl Acad Sci U S A, 98*(20), 11621-11626. doi:10.1073/pnas.191384198

Branda, S. S., Vik, S., Friedman, L., & Kolter, R. (2005). Biofilms: the matrix revisited. *Trends Microbiol, 13*(1), 20-26. doi:10.1016/j.tim.2004.11.006

Briley, K., Jr., Dorsey-Oresto, A., Prepiak, P., Dias, M. J., Mann, J. M., & Dubnau, D. (2011). The secretion ATPase ComGA is required for the binding and transport of transforming DNA. *Mol Microbiol, 81*(3), 818-830. doi:10.1111/j.1365-2958.2011.07730.x

Burbulys, D., Trach, K. A., & Hoch, J. A. (1991). Initiation of sporulation in B. subtilis is controlled by a multicomponent phosphorelay. *Cell, 64*(3), 545-552.

Burkholder, P. R., & Giles, N. H., Jr. (1947). Induced biochemical mutations in *Bacillus subtilis*. *Am J Bot, 34*(6), 345-348.

Burton, B., & Dubnau, D. (2010). Membrane-associated DNA transport machines. *Cold Spring Harb Perspect Biol, 2*(7), a000406. doi:10.1101/cshperspect.a000406

Cairns, L. S., Marlow, V. L., Bissett, E., Ostrowski, A., & Stanley-Wall, N. R. (2013). A mechanical signal transmitted by the flagellum controls signalling in *Bacillus subtilis*. *Mol Microbiol, 90*(1), 6-21. doi:10.1111/mmi.12342

Capra, E. J., & Laub, M. T. (2012). Evolution of two-component signal transduction systems. *Annu Rev Microbiol, 66*, 325-347. doi:10.1146/annurev-micro-092611-150039

Chai, Y., Chu, F., Kolter, R., & Losick, R. (2008). Bistability and biofilm formation in *Bacillus subtilis*. *Mol Microbiol, 67*(2), 254-263. doi:10.1111/j.1365-2958.2007.06040.x

Chai, Y., Kolter, R., & Losick, R. (2009). Paralogous antirepressors acting on the master regulator for biofilm formation in *Bacillus subtilis*. *Mol Microbiol, 74*(4), 876-887. doi:10.1111/j.1365-2958.2009.06900.x

Chai, Y., Kolter, R., & Losick, R. (2010). Reversal of an epigenetic switch governing cell chaining in *Bacillus subtilis* by protein instability. *Mol Microbiol, 78*(1), 218-229. doi:10.1111/j.1365-2958.2010.07335.x

Chai, Y., Norman, T., Kolter, R., & Losick, R. (2010). An epigenetic switch governing daughter cell separation in *Bacillus subtilis*. *Genes Dev, 24*(8), 754-765. doi:10.1101/gad.1915010

Chen, I., Christie, P. J., & Dubnau, D. (2005). The ins and outs of DNA transfer in bacteria. *Science, 310*(5753), 1456-1460. doi:10.1126/science.1114021

Chen, I., & Dubnau, D. (2004). DNA uptake during bacterial transformation. *Nat Rev Microbiol, 2*(3), 241-249. doi:10.1038/nrmicro844

Chu, F., Kearns, D. B., McLoon, A., Chai, Y., Kolter, R., & Losick, R. (2008). A novel regulatory protein governing biofilm formation in *Bacillus subtilis*. *Mol Microbiol, 68*(5), 1117-1127. doi:10.1111/j.1365-2958.2008.06201.x

Chung, Y. S., & Dubnau, D. (1995). ComC is required for the processing and translocation of *comGC*, a pilin-like competence protein of *Bacillus subtilis*. *Mol Microbiol, 15*(3), 543-551.

Clark, A. J., & Adelberg, E. A. (1962). Bacterial conjugation. *Annu Rev Microbiol, 16*, 289-319. doi:10.1146/annurev.mi.16.100162.001445

Crutz, A. M., & Steinmetz, M. (1992). Transcription of the *Bacillus subtilis* sacX and sacY genes, encoding regulators of sucrose metabolism, is both inducible by sucrose and controlled by the DegS-DegU signalling system. *J Bacteriol, 174*(19), 6087-6095.

Dahl, M. K., Msadek, T., Kunst, F., & Rapoport, G. (1992). The phosphorylation state of the DegU response regulator acts as a molecular switch allowing either degradative enzyme synthesis or expression of genetic competence in *Bacillus subtilis*. *J Biol Chem, 267*(20), 14509-14514.

Dahm, R. (2008). Discovering DNA: Friedrich Miescher and the early years of nucleic acid research. *Hum Genet, 122*(6), 565-581. doi:10.1007/s00439-007-0433-0

Dartois, V., Debarbouille, M., Kunst, F., & Rapoport, G. (1998). Characterization of a novel member of the DegS-DegU regulon affected by salt stress in *Bacillus subtilis*. *J Bacteriol, 180*(7), 1855-1861.

Davies, D. (2003). Understanding biofilm resistance to antibacterial agents. *Nat Rev Drug Discov, 2*(2), 114-122. doi:10.1038/nrd1008

de la Cruz, F., & Davies, J. (2000). Horizontal gene transfer and the origin of species: lessons from bacteria. *Trends Microbiol, 8*(3), 128-133.

Dubnau, D. (1991). Genetic competence in *Bacillus subtilis*. *Microbiol Rev, 55*(3), 395-424.

Dubnau, D., & Cirigliano, C. (1972). Fate of transforming DNA following uptake by competent *Bacillus subtilis*. Formation and properties of products isolated from transformed cells which are derived entirely from donor DNA. *J Mol Biol, 64*(1), 9-29.

Erickson, R. J. (1970). New ideas and data on competence and DNA entry in transformation of *bacillus subtilis*. *Curr Top Microbiol Immunol, 53*, 149-199.

Fernandez, S., Ayora, S., & Alonso, J. C. (2000). *Bacillus subtilis* homologous recombination: genes and products. *Res Microbiol, 151*(6), 481-486.

Finkel, S. E., & Kolter, R. (2001). DNA as a nutrient: novel role for bacterial competence gene homologs. *J Bacteriol, 183*(21), 6288-6293. doi:10.1128/JB.183.21.6288-6293.2001

Fujita, M., Gonzalez-Pastor, J. E., & Losick, R. (2005). High- and low-threshold genes in the Spo0A regulon of *Bacillus subtilis*. *J Bacteriol, 187*(4), 1357-1368. doi:10.1128/JB.187.4.1357-1368.2005

Giraldo, R., Andreu, J. M., & Diaz-Orejas, R. (1998). Protein domains and conformational changes in the activation of RepA, a DNA replication initiator. *EMBO J, 17*(15), 4511-4526. doi:10.1093/emboj/17.15.4511

Gould, S. J., & Subramani, S. (1988). Firefly luciferase as a tool in molecular and cell biology. *Anal Biochem, 175*(1), 5-13.

Griffith, F. (1928). The Significance of *Pneumococcal* Types. *J Hyg (Lond), 27*(2), 113-159.

Guttenplan, S. B., Blair, K. M., & Kearns, D. B. (2010). The EpsE flagellar clutch is bifunctional and synergizes with EPS biosynthesis to promote *Bacillus subtilis* biofilm formation. *PLoS Genet, 6*(12), e1001243. doi:10.1371/journal.pgen.1001243

Hahn, J., & Dubnau, D. (1991). Growth stage signal transduction and the requirements for *srfA* induction in development of competence. *J Bacteriol, 173*(22), 7275-7282.

Hahn, J., Inamine, G., Kozlov, Y., & Dubnau, D. (1993). Characterization of comE, a late competence operon of *Bacillus subtilis* required for the binding and uptake of transforming DNA. *Mol Microbiol, 10*(1), 99-111.

Hahn, J., Luttinger, A., & Dubnau, D. (1996). Regulatory inputs for the synthesis of ComK, the competence transcription factor of *Bacillus subtilis*. *Mol Microbiol, 21*(4), 763-775.

Hahn, J., Maier, B., Haijema, B. J., Sheetz, M., & Dubnau, D. (2005). Transformation proteins and DNA uptake localize to the cell poles in *Bacillus subtilis*. *Cell, 122*(1), 59-71. doi:10.1016/j.cell.2005.04.035

Hahn, J., Tanner, A. W., Carabetta, V. J., Cristea, I. M., & Dubnau, D. (2015). ComGA-RelA interaction and persistence in the *Bacillus subtilis* K-state. *Mol Microbiol, 97*(3), 454-471. doi:10.1111/mmi.13040

Haijema, B. J., Hahn, J., Haynes, J., & Dubnau, D. (2001). A ComGA-dependent checkpoint limits growth during the escape from competence. *Mol Microbiol, 40*(1), 52-64.

Hall-Stoodley, L., Costerton, J. W., & Stoodley, P. (2004). Bacterial biofilms: from the natural environment to infectious diseases. *Nat Rev Microbiol, 2*(2), 95-108. doi:10.1038/nrmicro821

Hamoen, L. W., Kausche, D., Marahiel, M. A., van Sinderen, D., Venema, G., & Serror, P. (2003). The *Bacillus subtilis* transition state regulator AbrB binds to the -35 promoter region of *comK*. *FEMS Microbiol Lett, 218*(2), 299-304.

Hamoen, L. W., Smits, W. K., de Jong, A., Holsappel, S., & Kuipers, O. P. (2002). Improving the predictive value of the competence transcription factor (ComK) binding site in *Bacillus subtilis* using a genomic approach. *Nucleic Acids Res, 30*(24), 5517-5528.

Hamoen, L. W., Van Werkhoven, A. F., Bijlsma, J. J., Dubnau, D., & Venema, G. (1998). The competence transcription factor of *Bacillus subtilis* recognizes short A/T-rich sequences arranged in a unique, flexible pattern along the DNA helix. *Genes Dev, 12*(10), 1539-1550.

Hamoen, L. W., Van Werkhoven, A. F., Venema, G., & Dubnau, D. (2000). The pleiotropic response regulator DegU functions as a priming protein in competence development in *Bacillus subtilis*. *Proc Natl Acad Sci U S A, 97*(16), 9246-9251. doi:10.1073/pnas.160010597

Henner, D. J., Ferrari, E., Perego, M., & Hoch, J. A. (1988). Location of the targets of the hpr-97, sacU32(Hy), and sacQ36(Hy) mutations in upstream regions of the subtilisin promoter. *J Bacteriol, 170*(1), 296-300.

Hershey, A. D. (1952). Inheritance in bacteriophage. *Ann N Y Acad Sci, 54*(6), 960-962.

Hoa, T. T., Tortosa, P., Albano, M., & Dubnau, D. (2002). Rok (YkuW) regulates genetic competence in *Bacillus subtilis* by directly repressing comK. *Mol Microbiol, 43*(1), 15-26.

Hobley, L., Ostrowski, A., Rao, F. V., Bromley, K. M., Porter, M., Prescott, A. R., . . . Stanley-Wall, N. R. (2013). BslA is a self-assembling bacterial hydrophobin that coats the *Bacillus subtilis* biofilm. *Proc Natl Acad Sci U S A, 110*(33), 13600-13605. doi:10.1073/pnas.1306390110

Hsueh, Y. H., Cozy, L. M., Sham, L. T., Calvo, R. A., Gutu, A. D., Winkler, M. E., & Kearns, D. B. (2011). DegU-phosphate activates expression of the anti-sigma factor FlgM in *Bacillus subtilis*. *Mol Microbiol, 81*(4), 1092-1108. doi:10.1111/j.1365-2958.2011.07755.x

Inamine, G. S., & Dubnau, D. (1995). ComEA, a *Bacillus subtilis* integral membrane protein required for genetic transformation, is needed for both DNA binding and transport. *J Bacteriol, 177*(11), 3045-3051.

Jiang, M., Shao, W., Perego, M., & Hoch, J. A. (2000). Multiple histidine kinases regulate entry into stationary phase and sporulation in *Bacillus subtilis*. *Mol Microbiol, 38*(3), 535-542.

Johnsen, P. J., Dubnau, D., & Levin, B. R. (2009). Episodic selection and the maintenance of competence and natural transformation in *Bacillus subtilis*. *Genetics, 181*(4), 1521-1533. doi:10.1534/genetics.108.099523

Kearns, D. B., Chu, F., Branda, S. S., Kolter, R., & Losick, R. (2005). A master regulator for biofilm formation by *Bacillus subtilis*. *Mol Microbiol, 55*(3), 739-749. doi:10.1111/j.1365-2958.2004.04440.x

Kobayashi, K. (2007a). Bacillus subtilis pellicle formation proceeds through genetically defined morphological changes. *J Bacteriol, 189*(13), 4920-4931. doi:10.1128/JB.00157-07

Kobayashi, K. (2007b). Gradual activation of the response regulator DegU controls serial expression of genes for flagellum formation and biofilm formation in *Bacillus subtilis*. *Mol Microbiol, 66*(2), 395-409. doi:10.1111/j.1365-2958.2007.05923.x

Kobayashi, K., & Iwano, M. (2012). BslA(YuaB) forms a hydrophobic layer on the surface of *Bacillus subtilis* biofilms. *Mol Microbiol, 85*(1), 51-66. doi:10.1111/j.1365-2958.2012.08094.x

Konkol, M. A., Blair, K. M., & Kearns, D. B. (2013). Plasmid-encoded ComI inhibits competence in the ancestral 3610 strain of *Bacillus subtilis*. *J Bacteriol, 195*(18), 4085-4093. doi:10.1128/JB.00696-13

Koonin, E. V., Makarova, K. S., & Aravind, L. (2001). Horizontal gene transfer in prokaryotes: quantification and classification. *Annu Rev Microbiol, 55*, 709-742. doi:10.1146/annurev.micro.55.1.709

Kovacs, A. T., & Kuipers, O. P. (2011). Rok regulates yuaB expression during architecturally complex colony development of *Bacillus subtilis* 168. *J Bacteriol, 193*(4), 998-1002. doi:10.1128/JB.01170-10

Kresge, N., Simoni, R. D., & Hill, R. L. (2011). The molecular genetics of bacteriophage: the work of Norton Zinder. *J Biol Chem, 286*(25), e4-5.

Kuhsel, M. G., Strickland, R., & Palmer, J. D. (1990). An ancient group I intron shared by eubacteria and chloroplasts. *Science, 250*(4987), 1570-1573.

Lacks, S., Greenberg, B., & Neuberger, M. (1975). Identification of a deoxyribonuclease implicated in genetic transformation of *Diplococcus pneumoniae*. *J Bacteriol, 123*(1), 222-232.

Lazazzera, B. A. (2000). Quorum sensing and starvation: signals for entry into stationary phase. *Curr Opin Microbiol, 3*(2), 177-182.

Lazazzera, B. A., Solomon, J. M., & Grossman, A. D. (1997). An exported peptide functions intracellularly to contribute to cell density signaling in *B. subtilis*. *Cell, 89*(6), 917-925.

Lederberg, J., Cavalli, L. L., & Lederberg, E. M. (1952). Sex Compatibility in Escherichia Coli. *Genetics, 37*(6), 720-730.

Leisner, M., Stingl, K., Radler, J. O., & Maier, B. (2007). Basal expression rate of *comK* sets a 'switching-window' into the K-state of *Bacillus subtilis*. *Mol Microbiol, 63*(6), 1806-1816. doi:10.1111/j.1365-2958.2007.05628.x

Lenhart, J. S., Schroeder, J. W., Walsh, B. W., & Simmons, L. A. (2012). DNA repair and genome maintenance in *Bacillus subtilis*. *Microbiol Mol Biol Rev, 76*(3), 530-564. doi:10.1128/MMBR.05020-11

Levene, P.A. (1919) The structure of yeast nucleic acid: IV. Ammonia hydrolysis. *J Biol Chem* 40:415-424

Logan, B. E. (2009). Exoelectrogenic bacteria that power microbial fuel cells. *Nat Rev Microbiol, 7*(5), 375-381. doi:10.1038/nrmicro2113

Londono-Vallejo, J. A., & Dubnau, D. (1993). comF, a *Bacillus subtilis* late competence locus, encodes a protein similar to ATP-dependent RNA/DNA helicases. *Mol Microbiol, 9*(1), 119-131.

Londono-Vallejo, J. A., & Dubnau, D. (1994). Mutation of the putative nucleotide binding site of the *Bacillus subtilis* membrane protein ComFA abolishes the uptake of DNA during transformation. *J Bacteriol, 176*(15), 4642-4645.

Lopez, D., Fischbach, M. A., Chu, F., Losick, R., & Kolter, R. (2009). Structurally diverse natural products that cause potassium leakage trigger multicellularity in *Bacillus subtilis*. *Proc Natl Acad Sci U S A, 106*(1), 280-285. doi:10.1073/pnas.0810940106

Lopez, D., Gontang, E. A., & Kolter, R. (2010). Potassium sensing histidine kinase in *Bacillus subtilis*. *Methods Enzymol, 471*, 229-251. doi:10.1016/S0076-6879(10)71013-2

Lopez, D., Vlamakis, H., & Kolter, R. (2009). Generation of multiple cell types in *Bacillus subtilis*. *FEMS Microbiol Rev, 33*(1), 152-163. doi:10.1111/j.1574-6976.2008.00148.x

Lopez, D., Vlamakis, H., Losick, R., & Kolter, R. (2009). Paracrine signaling in a bacterium. *Genes Dev, 23*(14), 1631-1638. doi:10.1101/gad.1813709

Maamar, H., & Dubnau, D. (2005). Bistability in the *Bacillus subtilis* K-state (competence) system requires a positive feedback loop. *Mol Microbiol, 56*(3), 615-624. doi:10.1111/j.1365-2958.2005.04592.x

Mah, T. F., & O'Toole, G. A. (2001). Mechanisms of biofilm resistance to antimicrobial agents. *Trends Microbiol, 9*(1), 34-39.

Marlow, V. L., Porter, M., Hobley, L., Kiley, T. B., Swedlow, J. R., Davidson, F. A., & Stanley-Wall, N. R. (2014). Phosphorylated DegU manipulates cell fate differentiation in the *Bacillus subtilis* biofilm. *J Bacteriol, 196*(1), 16-27. doi:10.1128/JB.00930-13

McLoon, A. L., Guttenplan, S. B., Kearns, D. B., Kolter, R., & Losick, R. (2011). Tracing the domestication of a biofilm-forming bacterium. *J Bacteriol, 193*(8), 2027-2034. doi:10.1128/JB.01542-10

Mejean, V., & Claverys, J. P. (1993). DNA processing during entry in transformation of *Streptococcus pneumoniae*. *J Biol Chem, 268*(8), 5594-5599.

Mirouze, N., Desai, Y., Raj, A., & Dubnau, D. (2012). Spo0A~P imposes a temporal gate for the bimodal expression of competence in *Bacillus subtilis*. *PLoS Genet, 8*(3), e1002586. doi:10.1371/journal.pgen.1002586

Mirouze, N., Ferret, C., Yao, Z., Chastanet, A., & Carballido-Lopez, R. (2015). MreB-Dependent Inhibition of Cell Elongation during the Escape from Competence in *Bacillus subtilis*. *PLoS Genet, 11*(6), e1005299. doi:10.1371/journal.pgen.1005299

Mitrophanov, A. Y., & Groisman, E. A. (2008). Signal integration in bacterial two-component regulatory systems. *Genes Dev, 22*(19), 2601-2611. doi:10.1101/gad.1700308

Mohan, S., Aghion, J., Guillen, N., & Dubnau, D. (1989). Molecular cloning and characterization of comC, a late competence gene of *Bacillus subtilis*. *J Bacteriol, 171*(11), 6043-6051.

Mordini, S., Osera, C., Marini, S., Scavone, F., Bellazzi, R., Galizzi, A., & Calvio, C. (2013). The role of SwrA, DegU and P(D3) in fla/che expression in *B. subtilis*. *PLoS One, 8*(12), e85065. doi:10.1371/journal.pone.0085065

Morrison, D. A., & Baker, M. F. (1979). Competence for genetic transformation in *pneumococcus* depends on synthesis of a small set of proteins. *Nature, 282*(5735), 215-217.

Msadek, T., Kunst, F., Klier, A., & Rapoport, G. (1991). DegS-DegU and ComP-ComA modulator-effector pairs control expression of the *Bacillus subtilis* pleiotropic regulatory gene *degQ*. *J Bacteriol, 173*(7), 2366-2377.

Nakano, M. M., Xia, L. A., & Zuber, P. (1991). Transcription initiation region of the *srfA* operon, which is controlled by the comP-comA signal transduction system in *Bacillus subtilis*. *J Bacteriol, 173*(17), 5487-5493.

Nakano, M. M., & Zuber, P. (1991). The primary role of comA in establishment of the competent state in *Bacillus subtilis* is to activate expression of *srfA*. *J Bacteriol, 173*(22), 7269-7274.

Nester, E. W., & Stocker, B. A. (1963). Biosynthetic Latency in Early Stages of Deoxyribonucleic Acidtransformation in *Bacillus Subtilis*. *J Bacteriol, 86*, 785-796.

Ochman, H., Lawrence, J. G., & Groisman, E. A. (2000). Lateral gene transfer and the nature of bacterial innovation. *Nature, 405*(6784), 299-304. doi:10.1038/35012500

Ogura, M., Shimane, K., Asai, K., Ogasawara, N., & Tanaka, T. (2003). Binding of response regulator DegU to the aprE promoter is inhibited by RapG, which is counteracted by extracellular PhrG in *Bacillus subtilis*. *Mol Microbiol, 49*(6), 1685-1697.

Ogura, M., & Tsukahara, K. (2010). Autoregulation of the *Bacillus subtilis* response regulator gene degU is coupled with the proteolysis of DegU-P by ClpCP. *Mol Microbiol, 75*(5), 1244-1259. doi:10.1111/j.1365-2958.2010.07047.x

Ogura, M., Yamaguchi, H., Kobayashi, K., Ogasawara, N., Fujita, Y., & Tanaka, T. (2002). Whole-genome analysis of genes regulated by the *Bacillus subtilis* competence transcription factor ComK. *J Bacteriol, 184*(9), 2344-2351.

Omer Bendori, S., Pollak, S., Hizi, D., & Eldar, A. (2015). The RapP-PhrP quorum-sensing system of *Bacillus subtilis* strain NCIB3610 affects biofilm formation through multiple targets, due to an atypical signal-insensitive allele of RapP. *J Bacteriol, 197*(3), 592-602. doi:10.1128/JB.02382-14

Ostrowski, A., Mehert, A., Prescott, A., Kiley, T. B., & Stanley-Wall, N. R. (2011). YuaB functions synergistically with the exopolysaccharide and TasA amyloid fibers to allow biofilm formation by *Bacillus subtilis*. *J Bacteriol, 193*(18), 4821-4831. doi:10.1128/JB.00223-11

Pakula, R., & Walczak, W. (1963). On the nature of competence of transformable *streptococci*. *J Gen Microbiol, 31*, 125-133. doi:10.1099/00221287-31-1-125

Parashar, V., Konkol, M. A., Kearns, D. B., & Neiditch, M. B. (2013). A plasmid-encoded phosphatase regulates *Bacillus subtilis* biofilm architecture, sporulation, and genetic competence. *J Bacteriol, 195*(10), 2437-2448. doi:10.1128/JB.02030-12

Penalva, M. A., Moya, A., Dopazo, J., & Ramon, D. (1990). Sequences of isopenicillin N synthetase genes suggest horizontal gene transfer from prokaryotes to eukaryotes. *Proc Biol Sci, 241*(1302), 164-169. doi:10.1098/rspb.1990.0081

Piazza, F., Tortosa, P., & Dubnau, D. (1999). Mutational analysis and membrane topology of ComP, a quorum-sensing histidine kinase of *Bacillus subtilis* controlling competence development. *J Bacteriol, 181*(15), 4540-4548.

Prepiak, P., & Dubnau, D. (2007). A peptide signal for adapter protein-mediated degradation by the AAA+ protease ClpCP. *Mol Cell, 26*(5), 639-647. doi:10.1016/j.molcel.2007.05.011

Provvedi, R., Chen, I., & Dubnau, D. (2001). NucA is required for DNA cleavage during transformation of *Bacillus subtilis*. *Mol Microbiol, 40*(3), 634-644.

Redfield, R. J. (1993). Genes for breakfast: the have-your-cake-and-eat-it-too of bacterial transformation. *J Hered, 84*(5), 400-404.

Roggiani, M., & Dubnau, D. (1993). ComA, a phosphorylated response regulator protein of *Bacillus subtilis*, binds to the promoter region of srfA. *J Bacteriol, 175*(10), 3182-3187.

Rudner, D. Z., LeDeaux, J. R., Ireton, K., & Grossman, A. D. (1991). The spo0K locus of *Bacillus subtilis* is homologous to the oligopeptide permease locus and is required for sporulation and competence. *J Bacteriol, 173*(4), 1388-1398.

Serror, P., & Sonenshein, A. L. (1996). CodY is required for nutritional repression of *Bacillus subtilis* genetic competence. *J Bacteriol, 178*(20), 5910-5915.

Singh, R., Paul, D., & Jain, R. K. (2006). Biofilms: implications in bioremediation. *Trends Microbiol, 14*(9), 389-397. doi:10.1016/j.tim.2006.07.001

Smits, W. K., Eschevins, C. C., Susanna, K. A., Bron, S., Kuipers, O. P., & Hamoen, L. W. (2005). Stripping *Bacillus*: ComK auto-stimulation is responsible for the bistable response in competence development. *Mol Microbiol, 56*(3), 604-614. doi:10.1111/j.1365-2958.2005.04488.x

Solomon, J. M., Lazazzera, B. A., & Grossman, A. D. (1996). Purification and characterization of an extracellular peptide factor that affects two different developmental pathways in *Bacillus subtilis*. *Genes Dev, 10*(16), 2014-2024.

Solomon, J. M., Magnuson, R., Srivastava, A., & Grossman, A. D. (1995). Convergent sensing pathways mediate response to two extracellular competence factors in *Bacillus subtilis*. *Genes Dev, 9*(5), 547-558.

Sparling, P. F. (1966). Genetic transformation of *Neisseria gonorrhoeae* to streptomycin resistance. *J Bacteriol, 92*(5), 1364-1371.

Spizizen, J. (1958). Transformation of Biochemically Deficient Strains of *Bacillus Subtilis* by Deoxyribonucleate. *Proc Natl Acad Sci U S A, 44*(10), 1072-1078.

Spizizen, J., Reilly, B. E., & Evans, A. H. (1966). Microbial transformation and transfection. *Annu Rev Microbiol, 20*, 371-400. doi:10.1146/annurev.mi.20.100166.002103

Stanley, N. R., & Lazazzera, B. A. (2005). Defining the genetic differences between wild and domestic strains of *Bacillus subtilis* that affect poly-gamma-dl-glutamic acid production and biofilm formation. *Mol Microbiol, 57*(4), 1143-1158. doi:10.1111/j.1365-2958.2005.04746.x

Stefanic, P., & Mandic-Mulec, I. (2009). Social interactions and distribution of *Bacillus subtilis* pherotypes at microscale. *J Bacteriol, 191*(6), 1756-1764. doi:10.1128/JB.01290-08

Stewart, P. S. (2002). Mechanisms of antibiotic resistance in bacterial biofilms. *Int J Med Microbiol, 292*(2), 107-113. doi:10.1078/1438-4221-00196

Stewart, P. S., & Franklin, M. J. (2008). Physiological heterogeneity in biofilms. *Nat Rev Microbiol, 6*(3), 199-210. doi:10.1038/nrmicro1838

Stock, A. M., Robinson, V. L., & Goudreau, P. N. (2000). Two-component signal transduction. *Annu Rev Biochem, 69*, 183-215. doi:10.1146/annurev.biochem.69.1.183

Stover, A. G., & Driks, A. (1999). Regulation of synthesis of the *Bacillus subtilis* transition-phase, spore-associated antibacterial protein TasA. *J Bacteriol, 181*(17), 5476-5481.

Strauch, M., Webb, V., Spiegelman, G., & Hoch, J. A. (1990). The SpoOA protein of *Bacillus subtilis* is a repressor of the abrB gene. *Proc Natl Acad Sci U S A, 87*(5), 1801-1805.

Suel, G. M., Kulkarni, R. P., Dworkin, J., Garcia-Ojalvo, J., & Elowitz, M. B. (2007). Tunability and noise dependence in differentiation dynamics. *Science, 315*(5819), 1716-1719. doi:10.1126/science.1137455

Tanaka, T., & Ogura, M. (1998). A novel *Bacillus natto* plasmid pLS32 capable of replication in *Bacillus subtilis*. *FEBS Lett, 422*(2), 243-246.

Tomasz, A. (1966). Model for the mechanism controlling the expression of competent state in *Pneumococcus* cultures. *J Bacteriol, 91*(3), 1050-1061.

Tortosa, P., Logsdon, L., Kraigher, B., Itoh, Y., Mandic-Mulec, I., & Dubnau, D. (2001). Specificity and genetic polymorphism of the *Bacillus* competence quorum-sensing system. *J Bacteriol, 183*(2), 451-460. doi:10.1128/JB.183.2.451-460.2001

Trieu-Cuot, P., Gerbaud, G., Lambert, T., & Courvalin, P. (1985). In vivo transfer of genetic information between gram-positive and gram-negative bacteria. *EMBO J, 4*(13A), 3583-3587.

Tsukahara, K., & Ogura, M. (2008). Promoter selectivity of the *Bacillus subtilis* response regulator DegU, a positive regulator of the fla/che operon and sacB. *BMC Microbiol, 8*, 8. doi:10.1186/1471-2180-8-8

Turgay, K., Hahn, J., Burghoorn, J., & Dubnau, D. (1998). Competence in *Bacillus subtilis* is controlled by regulated proteolysis of a transcription factor. *EMBO J, 17*(22), 6730-6738. doi:10.1093/emboj/17.22.6730

Turgay, K., Hamoen, L. W., Venema, G., & Dubnau, D. (1997). Biochemical characterization of a molecular switch involving the heat shock protein ClpC, which controls the activity of ComK, the competence transcription factor of *Bacillus subtilis*. *Genes Dev, 11*(1), 119-128.

Uhl, M. A., & Miller, J. F. (1996). Central role of the BvgS receiver as a phosphorylated intermediate in a complex two-component phosphorelay. *J Biol Chem, 271*(52), 33176-33180.

van Sinderen, D., Luttinger, A., Kong, L., Dubnau, D., Venema, G., & Hamoen, L. (1995). *comK* encodes the competence transcription factor, the key regulatory protein for competence development in *Bacillus subtilis*. *Mol Microbiol, 15*(3), 455-462.

van Sinderen, D., & Venema, G. (1994). *comK* acts as an autoregulatory control switch in the signal transduction route to competence in *Bacillus subtilis*. *J Bacteriol, 176*(18), 5762-5770.

Veening, J. W., Smits, W. K., & Kuipers, O. P. (2008). Bistability, epigenetics, and bet-hedging in bacteria. *Annu Rev Microbiol, 62*, 193-210. doi:10.1146/annurev.micro.62.081307.163002

Verhamme, D. T., Kiley, T. B., & Stanley-Wall, N. R. (2007). DegU co-ordinates multicellular behaviour exhibited by *Bacillus subtilis*. *Mol Microbiol, 65*(2), 554-568. doi:10.1111/j.1365-2958.2007.05810.x

Verhamme, D. T., Murray, E. J., & Stanley-Wall, N. R. (2009). DegU and Spo0A jointly control transcription of two loci required for complex colony development by *Bacillus subtilis*. *J Bacteriol, 191*(1), 100-108. doi:10.1128/JB.01236-08

Vlamakis, H., Aguilar, C., Losick, R., & Kolter, R. (2008). Control of cell fate by the formation of an architecturally complex bacterial community. *Genes Dev, 22*(7), 945-953. doi:10.1101/gad.1645008

Vlamakis, H., Chai, Y., Beauregard, P., Losick, R., & Kolter, R. (2013). Sticking together: building a biofilm the *Bacillus subtilis* way. *Nat Rev Microbiol, 11*(3), 157-168. doi:10.1038/nrmicro2960

Weinrauch, Y., Msadek, T., Kunst, F., & Dubnau, D. (1991). Sequence and properties of comQ, a new competence regulatory gene of *Bacillus subtilis*. *J Bacteriol, 173*(18), 5685-5693.

Wozniak, R. A., & Waldor, M. K. (2010). Integrative and conjugative elements: mosaic mobile genetic elements enabling dynamic lateral gene flow. *Nat Rev Microbiol, 8*(8), 552-563. doi:10.1038/nrmicro2382

Yadav, T., Carrasco, B., Hejna, J., Suzuki, Y., Takeyasu, K., & Alonso, J. C. (2013). *Bacillus subtilis* DprA recruits RecA onto single-stranded DNA and mediates

annealing of complementary strands coated by SsbB and SsbA. *J Biol Chem, 288*(31), 22437-22450. doi:10.1074/jbc.M113.478347

Yadav, T., Carrasco, B., Serrano, E., & Alonso, J. C. (2014). Roles of *Bacillus subtilis* DprA and SsbA in RecA-mediated genetic recombination. *J Biol Chem, 289*(40), 27640-27652. doi:10.1074/jbc.M114.577924

Yang, M., Ferrari, E., Chen, E., & Henner, D. J. (1986). Identification of the pleiotropic sacQ gene of *Bacillus subtilis*. *J Bacteriol, 166*(1), 113-119.

Yuksel, M., Power, J. J., Ribbe, J., Volkmann, T., & Maier, B. (2016). Fitness Trade-Offs in Competence Differentiation of *Bacillus subtilis*. *Front Microbiol, 7*, 888. doi:10.3389/fmicb.2016.00888

Zeigler, D. R., Pragai, Z., Rodriguez, S., Chevreux, B., Muffler, A., Albert, T., . . . Perkins, J. B. (2008). The origins of 168, W23, and other *Bacillus subtilis* legacy strains. *J Bacteriol, 190*(21), 6983-6995. doi:10.1128/JB.00722-08

Zinder, N. D., & Lederberg, J. (1952). Genetic exchange in Salmonella. *J Bacteriol, 64*(5), 679-699.

AUTHOR: Mathieu MIRAS

TITLE: Identification of a new regulatory pathway for the K-state in *Bacillus subtilis*

THESIS SUPERVISOR: Patrice POLARD

PLACE AND DATE OF PhD DEFENSE:
Université Toulouse III, Paul Sabatier, Toulouse
IBCG's conference room, Friday April 28th 2017, 2pm

ABSTRACT:
Bacteria can acquire new traits either by mutation or by the acquisition of external genetic material. This external acquisition of DNA is termed horizontal gene transfer and can occur through three mechanisms: conjugation, transduction and transformation. Transformation is the gene transfer resulting from the uptake of exogenous (environmental) DNA. The Gram-positive model organism *Bacillus subtilis* develops competence for DNA transformation if, and only if, the master regulator ComK is produced. In domesticated strains (strains used and selected after years of genetic manipulation in laboratory) occurs in approximately 15% of the population. Although competence regulation mechanisms are well understood, the reason why wild isolates (ancestors of the domesticated strains) of *B. subtilis* are poorly transformable (~1-2%) remains unclear. We demonstrate in this work that it is largely due to an inactivating point mutation in the promoter region of *degQ*. DegQ is known to be involved in the regulation of several different developmental pathways such as biofilm formation, exoprotease synthesis and competence development. This discovery introduces a new regulatory pathway in the development of genetic competence that was missing from our understanding. Indeed, all domesticated strains of *B. subtilis* share the *degQ* mutation that down-regulates the synthesis of DegQ and thereafter stabilizes ComK allowing a higher percentage of the population to become competent compared to the wild isolates. It appears to us that domestication and intensive genetic manipulations over the past fifty years are responsible of the loss of the "*degQ* regulation pathway" of competence in *B. subtilis* illustrating the significance of using wild isolates for the study of real-life biology.

KEY WORDS: *Bacillus subtilis*, DNA transformation, regulation, competence, DegU, ComK

AREA OF EXPERTISE: Microbiology

HOST LABORATORY
Laboratoire de Microbiologie et Génétique Moléculaires (LMGM)
Université Toulouse III - Paul Sabatier
118 Route de Narbonne 31062 TOULOUSE Cedex 9 FRANCE

AUTEUR: Mathieu MIRAS

TITRE: Identification of a new regulatory pathway for the K-state in *Bacillus subtilis*

DIRECTEUR DE THESE: Patrice POLARD

LIEU ET DATE DE SOUTENANCE:
Université Toulouse III, Paul Sabatier, Toulouse
Salle de conférence de l'IBCG, le vendredi 28 Avril 2017 à 14h00

RESUME:
Les bactéries peuvent développer de nouvelles caractéristiques par mutation génétique ou par l'acquisition d'un matériel génétique exogène. Ce dernier, aussi appelé transfert de gène horizontal, peut se produire selon trois mécanismes différents : la conjugaison, la transduction et la transformation.
L'organisme modèle, *Bacillus subtilis*, est naturellement transformable (ou compétent) si, et seulement si, le facteur de transcription ComK est produit. Au sein des souches domestiques, environ 15% de la population bactérienne devient compétente. Bien que les mécanismes de régulation de la compétence soient bien caractérisés, la raison pour laquelle les souches non domestiques de *B. subtilis* sont peu transformables (1~2%) reste inconnue. Nous démontrons ici que c'est essentiellement dû à une mutation de transition dans le promoteur du gène *degQ*. La protéine DegQ est impliquée dans la régulation de la formation de biofilms, de la synthèse d'exoprotéases et de la transformation génétique.
Ce travail permet de mettre en évidence une nouvelle voie de régulation dans le développement de la compétence génétique qui était jusque-là méconnue. En effet, toutes les souches domestiques de *B. subtilis* partagent la mutation ponctuelle dans la région du promoteur de *degQ* qui a pour conséquence de diminuer la synthèse protéique de DegQ et de favoriser, à plus long terme, la stabilité de ComK. Ainsi, dans le contexte des souches domestiques, ComK va permettre à un plus grand nombre de bactéries de devenir compétentes.

MOTS-CLES: *Bacillus subtilis*, DNA transformation, regulation, competence, DegU, ComK

DISCIPLINE: Microbiologie

INTITULE ET ADRESSE DE L'UFR OU DU LABORATOIRE :
Laboratoire de Microbiologie et Génétique Moléculaires (LMGM)
Université Toulouse III - Paul Sabatier
118 Route de Narbonne 31062 TOULOUSE Cedex 9 FRANCE