Ubiquitous late competence genes in Bacillus species indicate the presence of functional DNA uptake machineries
Kovacs, Akos T.; Smits, Wiep Klaas; Mironczuk, Aleksandra M.; Kuipers, Oscar; Mirończuk, Aleksandra M.

Published in:
Environmental Microbiology

DOI:
10.1111/j.1462-2920.2009.01937.x

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

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

Publication date:
2009

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

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

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

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

Ubiquitous late competence genes in Bacillus species indicate the presence of functional DNA uptake machineries

Ákos T. Kovács,1 Wiep Klaas Smits,1† Aleksandra M. Mirończuk1 and Oscar P. Kuipers1,2*
1Department of Genetics, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, 9751 NN Haren, the Netherlands.
2Kluyver Centre for Genomics of Industrial Fermentation, Delft, the Netherlands.

Summary
Natural competence for genetic transformation, i.e. the ability to take up DNA and stably integrate it in the genome, has so far only been observed in the bacterial kingdom (both in Gram-negative and Gram-positive species) and may contribute to survival under adverse growth conditions. Bacillus subtilis, the model organism for the Bacillus genus, possesses a well-characterized competence machinery. Phylogenetic analysis of several genome sequences of different Bacillus species reveals the presence of many, but not all genes potentially involved in competence and its regulation. The recent demonstration of functional DNA uptake by B. cereus supports the significance of our genome analyses and shows that the ability for functional DNA uptake might be widespread among Bacilli.

Introduction
The Gram-positive bacterium Bacillus subtilis is a soil-dwelling member of the genus Bacillus, which comprises commercially interesting (for instance B. amyloliquefaciens, B. licheniformis and B. stearothermophilus) as well as pathogenic species (such as B. cereus and B. anthracis). It can be isolated from many environments (Vilain et al., 2006) and it is regarded a model organism for Gram-positives, because of its long history of scientific research and good experimental amenability (Sonen shein, 2002). In addition to this, B. subtilis was one of the first bacteria for which cellular differentiation was recognized (spore formation and competence development). As spore formation is accompanied by easily visible morphological changes, much of the early research on cellular differentiation was aimed at the isolation, mapping and classification of sporulation genes (Hoch, 1971; Piggot and Coote, 1976). Groundbreaking work of Anagnostopoulos and Spizizen in the early 1960s established a competence regime for B. subtilis (Spizizen, 1958; Anagnostopoulos and Spizizen, 1961), making it genetically accessible and paving the way for advanced molecular research. As a result, the characterization of competence and competence genes has become one of the longest standing fields of investigation for this bacterium. Since that time, many genes and proteins involved in competence have been characterized (for reviews see Hamoen et al., 2003; Chen and Dubnau, 2004).

DNA uptake apparatus in Bacillus subtilis
The ability to take up naked DNA has been detected in eubacteria, including Gram-positive (e.g. Bacillus, Strep tococcus) and Gram-negative species (e.g. Campylobacter, Haemophilus, Helicobacter, Neisseria, Vibrio) (Lorenz and Wackernagel, 1994). It may help bacteria to survive under adverse conditions such as nutrient limitation (Finkel and Kolter, 2001; Claverys et al., 2006; Palchevskiy and Finkel, 2006). In addition, it has been suggested that the thermophilic traits of the highly transformable bacterium Thermus thermophilus, which allow it to survive under extreme temperatures, were acquired via gene transfer (Averhoff, 2009). The DNA uptake machinery in most competent bacteria is related to Type II secretion systems and Type IV pili (Dubnau, 1999). In the case of Gram-positive organisms, such as B. subtilis, the presence of a thick peptidoglycan layer necessitates the modulation of the cell wall make-up in order for the DNA

Received 3 February 2009; accepted 13 March 2009. *For correspondence. E-mail o.p.kuipers@rug.nl; Tel. (+31) 50 3632093; Fax (+31) 50 3632348. †Present address: Department of Biology, Building 68–540D, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

© 2009 Society for Applied Microbiology and Blackwell Publishing Ltd
to access a membrane-located receptor, ComE<sub>c</sub>, which contains a DNA binding domain in its C-terminus (Provvedi and Dubnau, 1999). The proteins encoded by the comG operon resemble those of type IV pili (Dubnau, 1997) and are thought to form a structure called the competence pseudopilus (Chen and Dubnau, 2004; Chen et al., 2006). The ComG<sub>a</sub> protein is believed to be a traffic NTPase, providing the energy for the assembly of the competence pseudopilus. Single molecule measurements of DNA translocation into <i>B. subtilis</i> have demonstrated that DNA uptake depends on the transmembrane proton motif force (Maier et al., 2004). The ComG<sub>a</sub> protein localizes to the cell-pole during competence, and hence functional transport is believed to take place at that site (Hahn et al., 2005). ComG<sub>b</sub> is a polytopic membrane protein that is required for pilus assembly. The major pseudopilin is ComG<sub>c</sub>, while ComG<sub>d</sub>, ComG<sub>e</sub> and ComG<sub>f</sub> are minor pseudopilins. All of these proteins are produced as prepilins, and require cleavage of an N-terminal sequence by the prepilin peptidase ComC before assembly into a pseudopilus is possible (Chung and Dubnau, 1995; Chen et al., 2006). Recently it was reported that intramolecular disulfide bonds in the major pseudopilin ComG<sub>c</sub> dependent on the thiol-disulfide oxidoreductase pair BdbCD, are required to stabilize the protein (Meima et al., 2002; Chen et al., 2006). Strikingly, the competence pseudopilus is dispensable for DNA binding in the absence of a cell wall (Provvedi and Dubnau, 1999), supporting the hypothesis that it modulates the cell wall in such a way that the DNA gains access to the ComE<sub>c</sub> receptor protein. The receptor subsequently delivers the DNA to the permease, ComE<sub>c</sub>. Oligomers of the ComE<sub>c</sub> protein, that is essential for DNA uptake (Hahn et al., 1987; Inamine and Dubnau, 1995; Dubnau, 1999), form an aqueous pore through which the DNA is transported into the cell (Draskovic and Dubnau, 2005). Like ComG<sub>c</sub>, the protein contains an intramolecular disulfide bond that is probably introduced by BdbCD (Draskovic and Dubnau, 2005). ComF<sub>a</sub> is a membrane-associated protein (Londono-Vallejo and Dubnau, 1994) that is structurally similar to an ATP-dependent family of helicases and may act as the motor protein for DNA transport (Londono-Vallejo and Dubnau, 1993). Alternatively, its helicase activity may be required for unwinding of incoming double-stranded DNA. However, only single-stranded DNA (ssDNA) was shown to be taken up into the cells (Venema et al., 1965; Dubnau and Cirigliano, 1972). The non-transforming strand is degraded by an unidentified nuclease, which is presumably located on the outside of the membrane (Chen and Dubnau, 2004), and the degradation products are released into the medium (Dubnau and Cirigliano, 1972). The importance of DNA processing for transport is reinforced by the notion that the introduction of double strand breaks (cleavage) is required for efficient uptake, facilitated by the competence-induced nuclease NucA (van Sinderen et al., 1995; Provvedi et al., 2001). In order to reconstitute replicative plasmids or allow recombination of the transforming ssDNA with the host chromosome, the DNA has to be shielded from the action of nucleases by association with a competence-induced ssDNA binding protein (Eisenstadt et al., 1975), which likely corresponds to DprA (Smf) (Mortier Barriere et al., 2007; Tadesse and Graumann, 2007), SsbB (YwpH) (Lindner et al., 2004; Hahn et al., 2005; Morrison et al., 2007) and/or RecA (Lovett et al., 1989; Kidane and Graumann, 2005). Indeed, colocalization of these proteins with ComG<sub>a</sub> and ComF<sub>a</sub> in vivo has been observed at the cell pole using fluorescence resonance energy transfer experiments (Kramer et al., 2007). Plasmids are reconstituted and stably maintained, while ssDNA can be integrated into a homologous double-stranded DNA duplex mediated by the RecA ATPase. Its function depends in part on a helicase/nuclease complex that is formed by the AddAB proteins in <i>B. subtilis</i> (Hajjema et al., 1995; Arnold and Kowalczykowski, 2000; Kidane and Graumann, 2005), although it has been noted that the addAB genes were not identified as ComK-dependent in a DNA array analysis of the ComK regulon (Hamoen et al., 2002). CoiA (YjbF) contributes to the recombination process in a way that is not fully understood (Desai and Morrison, 2006; 2007). Notably, an extensive network of interactions exists between proteins that act after DNA uptake (Kidane and Graumann, 2005; Kramer et al., 2007; Mortier Barriere et al., 2007).

**ComK, the master regulator for competence development in <i>Bacillus</i>**

The transcription of the genes that encode the DNA-binding, uptake and -recombination proteins is controlled by the auto-activating competence transcription factor ComK (van Sinderen and Venema, 1994). In <i>B. subtilis</i>, the induction of this protein is strictly regulated at the level of comK transcription as well as post-translationally. Transcription of comK is repressed by binding of AbrB, CodY and Rok to the comK promoter region (Hamoen et al., 2003), whereas the ComK protein is trapped by MecA and targeted for proteolytic degradation by CipCP (Turgay et al., 1998). High cell density is a prerequisite for optimal competence development, as at least two independent quorum sensing pathways induce the production of anti-adaptor protein ComS, which liberates ComK from the proteolytic complex (D’Souza et al., 1994; Hamoen et al., 1995; Turgay et al., 1997; Prepiak and Dubnau, 2007).

Through the use of whole genome DNA microarrays, it was established that ComK is directly or indirectly responsible for the activation of ~100 genes (Berka}
ComK activates transcription by binding through the minor groove to specific sequences, so called K-boxes that are composed of two AT-boxes with the consensus sequence AAAAA-N5-TTTT. The boxes are separated by a spacing of a discrete number of helical turns, which places them on the same side of the DNA-helix (Hamoen et al., 1998).

**Distribution of genes encoding the competence machinery in various Bacilli**

Competence development within the Bacillus genus has so far been described for *B. subtilis*, *B. licheniformis* and *B. amyloliquefaciens* (Spizizen, 1958; Thorne and Stull, 1966; Koumoutsi et al., 2004), but not for *B. cereus*, *B. anthracis* or other Bacillus species. With the availability of the genome sequence of other members of Bacillaceae (Takami et al., 2000; Ivanova et al., 2003; Read et al., 2003; Rasko et al., 2004; Rey et al., 2004; Han et al., 2006), our attention was drawn to the high number of homologues of competence-related genes in other species.

Up to now, the only functional competence machinery that has been characterized in detail is the one of *B. subtilis*. Similarly, the genetic network controlling the differentiation into the competent state has only been thoroughly characterized for this species. To gain insight in the function and evolution of the DNA uptake machinery, we have analysed all the completely sequenced bacterial genomes of Bacillaceae and other closely related bacteria, one Oceanobacillus and two Geobacillus species, as they are deposited in the NCBI database (20 genomes in January 2009). These BLAST analyses revealed the presence of many orthologous genes putatively involved in DNA uptake, and are visualized with Genesis software (Sturn et al., 2002) in Fig. 1.

![Fig. 1. Presence of homologues of competence regulator proteins (A) and competence-related structural proteins (B) in Bacillus and closely related species. Results of BLAST searches were visualized with Genesis 1.6 software: white is absent (with e-value of E-0), dark blue is present (e-value < E-20). BLAST analysis was performed with *B. subtilis* protein sequences against translated protein database of a given genome. Where no hit (or bellow E-05) was found additional TBLASTN was performed. Protein names are indicated on the right. Bsu, *B. subtilis*; Bam, *B. amyloliquefaciens*; Bli, *B. licheniformis*; Bpu, *B. pumilus*; Ban, *B. anthracis*; Bce, *B. cereus*; Bth, *B. thuringiensis*; Bwe, *B. weihenstephanensis*; Oih, *O. iheyensis*; Gka, *G. kaustophilus*; Gth, *G. thermodenitrificans*; Bcl, *B. clausii*; Bha, *B. halodurans*).
Presence of the master regulator, ComK, in members of the Bacillus genus

In *B. subtilis*, competence development is strictly dependent on the ComK protein. Homologues of the comK gene were present in most species, but could not be detected in *B. clausii* KSMK16 and *B. halodurans* C-125. The presence or absence of comK in these species does not seem to be correlated to the presence of other competence genes. Assuming a similar dependency for competence development on ComK, this suggests that the gene may have been lost recently. In support of this, some other genes are either absent (*nucA*, *nin*, *addB*) or less conserved (*comGD*) compared with species that do harbour a comK gene (Fig. 1). In addition, the comG operon of *B. clausii* and *B. halodurans* contains an inserted tRNA region. Both these species are alkaliphilic and it is possible that diversification into an alkaline environment is not compatible with natural competence. Disruption of competence gene homologues by tRNA genes is reminiscent of the prophage insertion in comK of *Listeria monocytogenes* (Boreeze et al., 2000).

Interestingly, species from the *B. cereus* group of Bacilli (*B. cereus*, *B. anthracis*, *B. thuringiensis* and *B. weihenstephanensis*) contain two genes with homology to ComK (hereafter referred to as ComK1 and ComK2) (Fig. 2B). The putative ComK proteins show 61–62% (ComK1) and 44–48% (ComK2) similarity to *B. subtilis* ComK respectively. The ComK1 proteins, in general, are similar in length to the *B. subtilis* ComK protein, whereas the ComK2 proteins of the *B. cereus* group appear to be C-terminally truncated by 22–32 amino acids (Fig. 3).

Strikingly, the ComK/ComK1 proteins of outliers *G. kaustophilus*, *G. thermodenitrificans* and *O. iheyensis* are similarly 15, 10 and 23 amino acids shorter than the *B. subtilis* ComK protein respectively, raising the possibility that such a specific truncation may be relevant for the function of the protein. *Oceanobacillus iheyensis* ComK2 is 39 amino acids shorter. Previous studies on *B. subtilis* ComK have shown that a 25–35-amino-acid C-terminal truncation is incapable of transcriptional activation of a specific late competence promoter (PcomG) (Susanna et al., 2006), although the protein retained its ability to bind to DNA. Potentially, the multimerization of the ComK protein, presumably required for comG activation in *B. subtilis*, is affected. It will be of great interest to determine the oligomeric state of the various ComK1 and ComK2 proteins. Notably, both ComK1 and ComK2 of *B. cereus* ATCC14579 possess DNA binding activity (A.M. Mirończuk, Á.T. Kovács and O.P. Kuipers, unpublished).

The C-terminal part of the *B. subtilis* ComK protein shows homology to the DNA binding domain to two High...
Fig. 3: Multiple alignment of ComK homologues. The ComK proteins of *B. cereus* ATCC14579 are highlighted with boxes. Black background represents conserved amino acids and grey background represents similar amino acids. Alignment was performed using Clustal W (Thompson et al., 1994), and presented using the Boxshade 3.21 program. The N- and C-terminal deletions analysed by Susanna and colleagues (2006) are marked. Boxed amino acid residues indicate the residues involved in interaction with MecA (Prepiak and Dubnau, 2007). Alpha-helices and beta-sheets of *B. subtilis* ComK protein are indicated with rectangles and arrows under the alignment respectively. For abbreviations of species names see Fig. 1. In the case of *B. anthracis* and *B. licheniformis* strain names are not indicated as ComK protein sequences are identical within the different strains of the same species.
Mobility Group proteins, SRY and TCF-1, that like ComK bind through the minor groove of the DNA and induce bending. However, as both C-terminally truncated *B. subtilis* ComK (Susanna *et al.*, 2006) and ComK2 of *B. cereus* – that lacks this region – have DNA binding activity, it seems unlikely that the region of homology is the sole DNA binding domain of ComK. It is conceivable, however, that this region is required for modulating the DNA topology, a property that may be critical for transcriptional activation by *B. subtilis* ComK (Smits *et al.*, 2007).

A quest for ComK binding sites in the upstream region of putative competence related genes in *B. cereus* ATCC14579 reveals the presence of AT-boxes, although spacing between these sites mostly does not correspond to the spacing in functional K-boxes found in *B. subtilis*. This makes it likely that the *B. cereus* ComK proteins recognize a different sequence or that they bind as a dimer to AT-boxes, rather than as tetramer to K-boxes.

### The regulation of ComK levels

Recently, it was found that the C-terminus of ComK contains a region that is required and sufficient for the interaction with MecA (Prepiak and Dubnau, 2007). This FMLYPK motif can be found in the more closely related *B. subtilis*, *B. amyloliquefaciens*, *B. licheniformis* and *B. pumilus*, but not in any other ComK homologues (Fig. 3 and Prepiak and Dubnau, 2007). This indicates either (i) that a different interaction site is responsible for controlling ComK levels in these species or (ii) that it does not involve a MecA homologue. In either case, the regulatory events upstream of *comK* expression may differ significantly from that of *B. subtilis*. One possibility is that the regulation does not involve quorum sensing-dependent production of an anti-adaptor protein, such as ComS. In support of this, *comS* could not be identified in the *srfA* genes of the *B. cereus* group. Moreover, genes from the quorum sensing pathway could not unambiguously be identified in most species (Fig. S1). It has to be noted that the prediction of small open reading frames (ORFs) is not trivial and the *comS* gene may be located in other regions of the chromosome.

Not surprisingly considering their function in other pathways (Hamoen *et al.*, 2003), the pleiotropic regulators that directly or indirectly control *comK* transcription (e.g. DegU, CodY, AbrB and Spo0A) were almost universally identified in all species (see Fig. S1). As noted before, Rok may have been acquired recently, as it is only present in the *B. subtilis/amiloliquefaciens/pumilus/licheniformis* group (Albano *et al.*, 2005). Whether these regulators have a role similar to the one in *B. subtilis* competence development remains to be established.

### Structural genes for DNA uptake machinery in various *Bacillus* species

For all species analysed, homologues for most of the structural genes encoding DNA uptake machinery could be identified. Two striking exceptions are the *comFB* and *comG**EFG** genes.

The *comFB* gene was present only in *B. subtilis* and its close relatives *B. amyloliquefaciens*, *B. licheniformis* and *B. pumilus*, whereas the downstream *comFC* gene was present in all species analysed. The first gene of the *comF*-operon encodes a helicase-like protein essential for transformation. The function of ComFb is unknown, although transformation frequencies are slightly reduced in a *comFb* mutant (Londono-Vallejo and Dubnau, 1993). Possibly, ComFb is an auxiliary protein, acquired after the branching of the *B. subtilis/amiloliquefaciens* group (Fig. 2a). It is interesting that a *comFB* homologue is also missing in the more distantly related *Streptococcus pneumoniae* *comF* operon (Berge *et al.*, 2002).

The putative ComG-operons demonstrate some interesting features. Whereas ComG**AB** homologues (the ATPase and the polytopic membrane protein respectively) are present in all species analysed and ComG**CD** (similar to pilins) could also be confidently identified in at least the *B. cereus* lineage (on the basis of genome location, homology and size of the encoded protein), no homologues of the ComG**EFG** proteins (minor pilins) were detected (e-value > E-1) in any other species than *B. subtilis*, *B. amyloliquefaciens*, *B. licheniformis* and *B. pumilus* (Fig. 1). In *B. subtilis*, all seven ORFs of the *comG* operon are required for efficient transformation (Chung and Dubnau, 1998), and it is conceivable that the absence of these genes explains the fact that no functional DNA uptake has been reported for these organisms. However, the genes downstream of *comG**A*** are conserved within the *B. cereus* group and it is plausible that they encode functional homologues of the *B. subtilis* ComG**EFG** proteins (Fig. 4). In spite of the lower homology of these proteins to *B. subtilis* ComG**EFG** at the level of whole protein, the N-terminal regions locally show a higher level of conservation (Fig. S2). The N-terminal region of minor pilins contains the conserved prepilin-like domain. Alternatively, ComG**ABCD** might be sufficient in these organisms. Notably, the *comG**A*** gene is truncated in *B. cereus* Cyto toxis strain and in contrast to the conserved DNA sequence within the *comG* operon no clear *orf* for *comG**D*** can be assigned, while the allotted *orf* for *comG**D*** is elongated (Fig. 4).

The *comG* operon of the alkaliophilic *Bacillus* strains is disrupted by a tRNA island after *comG**A***. However, ComG**D*** proteins are encoded by genes downstream of the island, and at least in the case of *B. clausii* the genes downstream of *comG**D*** show strong homology to the *B. cereus* group of...
downstream genes (BC4236-BC4235 genes in *B. cereus* ATCC14579 respectively) (Fig. 4). In fact, the genomic arrangement of genes homologue to the *comGABCDEFG* is conserved even in *Streptococcus* species (*cglABCDEFG* in *S. pneumoniae*) (Peterson et al., 2004). 

**Induction of competence for DNA uptake in Gram-negative and Gram-positive bacteria**

The conditions that trigger natural competence can vary greatly, and may include medium/nutrient limitation, growth phase, cell density and other stresses (Claverys et al., 2006). Additionally, it may depend on a given substrate in the environment, as was recently described for *Vibrio cholerae* (chitin) (Meibom et al., 2005) or occur only during a short period [e.g. *Streptococci* (Claverys et al., 2006) or *Acinetobacter baylyi* (Porstendorfer et al., 2000; Friedrich et al., 2001)]. As a result, establishing a competence regime can be a daunting task and this may in part explain the lack of successful transformation of species harbouring homologues of competence genes. 

The mechanism of regulation of competence development generally diverges more than the DNA uptake machinery (Claverys and Martin, 2003), consistent with our findings in this study. Hence, several laboratories have focused on artificially inducing competence through the controlled expression of the putative key regulator. The ability to induce the expression of late competence genes has been shown for *Streptococcus pyogenes* (Woodbury et al., 2006), *Lactococcus lactis* (Wydau et al., 2006) and *Streptococcus thermophilus* (Blomqvist et al., 2006). However, only in the latter it was accompanied by...
In general, because of its dramatic impact on the physiology of the cells (Hajjema et al., 2001), competence for genetic transformation commonly only develops under certain conditions – the so-called competence regime. In fact, Neisseria gonorrhoeae is the only organism reported to demonstrate constitutive competence (Sparling, 1966). Within the genus Bacillus, a competence regime has been established for B. subtilis, B. licheniformis and B. amyloliquefaciens (Spizizen, 1958; Thorne and Stull, 1966; Koumoutsi et al., 2004). There is, however, a great interest in the other members, from both a medical and biotechnological point of view. Bacillus cereus and B. anthracis, for instance, are causative agents of common food poisoning and anthrax respectively. Bacillus halodurans and clausii have both been isolated on the basis of their biotechnological potential. Artificially induced genetic competence, depending on the expression of a functional DNA uptake apparatus, can facilitate molecular studies on these bacteria.

Bacilli grow under diverse environmental conditions and group members are commonly observed in soil (Vilain et al., 2006), in the rhizosphere of a variety of plants (Fall et al., 2004), in food samples (B. cereus (Schoeni and Wong, 2005)), in insect guts (B. thuringiensis (Jensen et al., 2003)) or in an aquatic milieu (de Barros Soares et al., 2003). These natural habitats provide complex conditions that are not easily reproduced in the laboratory. It is therefore likely that the number of Bacillus species identified as competent is an underrepresentation of the actual occurrence of natural competence in this genus.

Recently, it has been demonstrated that B. cereus has the ability to take up DNA (Mironczuk et al., 2008). Upon the overproduction of B. subtilis ComK protein consistently low levels of transformation with both chromosomal or plasmid DNA were obtained. These results are significant for two reasons. First, it shows that the gene reservoir of B. cereus ATCC14579 is sufficient for the uptake and integration of DNA from the environment. This indicates either that the lack of homologues of the B. subtilis comGEFG operon is functionally complemented in B. cereus or that these are proteins not needed for functional DNA uptake in this organism. Second, as the observed levels of competence are much lower than obtained by the equivalent overexpression of ComK in B. subtilis, it is expected that other factors (regulatory and/or structural) will be able to augment efficient DNA uptake.

The regulation of transcriptional activation of late competence genes can be conserved between closely related species (Martin et al., 2006) and the artificial induction of the key regulator of competence in S. thermophilus resulted in DNA uptake (Blomqvist et al., 2006). We anticipate therefore that a strategy similar to Mironczuk et al. may induce functional DNA uptake in at least the close relatives of B. cereus, such as B. anthracis and B. thuringiensis. Moreover, the strategy may be applicable to induce competence in natural Bacillus isolates, some of which demonstrate traits such as the ability to form architecturally complex communities of cells (biofilm) that have been lost in laboratory strains (Branda et al., 2001).

Concluding remarks

By screening the genomes of fully sequenced Bacillus species, we have identified genes encoding homologues of the proteins involved in competence in B. subtilis. Our findings suggest that species for which no competence regime has been established so far have the potential to develop natural competence or acquire DNA. Indeed, B. cereus has recently been shown to harbour a functional DNA uptake machinery. The induction of DNA uptake, either naturally or induced, will facilitate molecular genetic studies with these organisms. Moreover, the insights gained from these comparative studies extend our understanding of natural competence in general and competence in Bacillus species specifically and gives directions for further research on the factors required for competence in the Bacillus genus.

Natural competence of Bacilli in the environment like in the rhizosphere of plants might contribute to the genome plasticity observed in Bacilli, similar to conjugation that was demonstrated to occur in the rhizosphere between B. anthracis species (Saile and Koehler, 2006).

Future studies on competence should be aimed at answering whether a single factor is capable of inducing late competence (analogous to ComK in B. subtilis) and what the genetic or environmental factors affecting the activity of these regulators are (e.g. by fusing the...
promoter regions of ComK and/or late competence genes to reporter genes). Moreover, complementation studies will reveal whether putative (functional) homologues such as ComG between in the B. cereus group can support DNA uptake and may lead to the definition of a minimal competence machinery.

Acknowledgements

The authors wish to thank R.W.W. Brouwer for expert assistance with BLAST analysis and R. van Kraanenburg for helpful discussions. A.T.K. was financially supported by PURAC and by Grant 818.02.004 from ALW-NWO Open programma, W.K.S. by Grant 811.35.002 and a Rubicon fellowship of the Netherlands Organization for Scientific Research. This project was carried out within the research programme of the Kluiver Centre for Genomics of Industrial Fermentation which is part of the Netherlands Genomics Initiative/Netherlands Organization for Scientific Research.

References


© 2009 Society for Applied Microbiology and Blackwell Publishing Ltd, Environmental Microbiology, 11, 1911–1922


Supporting information

Additional Supporting Information may be found in the online version of this article:

Fig. S1. Presence of homologues of proteins involved in the regulation of competence in Bacillus and closely related species. Results of BLAST queries were visualized with Genesis 1.6 software: white is absent (with e-value of E-0), dark blue is present (e-value < E-20). BLAST analysis was performed with B. subtilis protein sequences against the translated protein database of a given genome. Question
marks denote small ORFs where identification is uncertain using the available bioinformatic tools that can miss homologues. Where the search yielded no hit (or with an e-value below E-05) an additional TBLASTN was performed. Protein names are indicated on the right. Bsu, *B. subtilis*; Bam, *B. amyloliquefaciens*; Bli, *B. licheniformis*; Bpu, *B. pumilus*; Ban, *B. anthracis*; Bce, *B. cereus*; Bth, *B. thuringiensis*; Bwe, *B. weihenstephanensis*; Oih, *O. iheyensis*; Gka, *G. kaustophilus*; Gth, *G. thermodenitrificans*; Bcl, *B. clausii*; Bha, *B. halodurans*.

**Fig. S2.** Multiple alignments of ComGA homologues. Black background represents conserved amino acids and grey background represents similar amino acids. Alignment was performed using Clustal W ([Thompson et al.](#), 1994), and presented using the Boxshade 3.21 program. For abbreviations of species names see Fig. S1. Conserved domains of *B. subtilis* ComG proteins are indicated with arrows above the alignment. AAA+, ATPase domain (Smart accession number: SM00382); GSPII-F, general secretion pathway domain (PFAM accession number: PF00482); TMS, transmembrane segment; N-met, N-methyl domain often found at N-terminus of pilins and proteins involved in secretion (PFAM accession number: PF07963); SSSeq, signal peptide sequence.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.