Comparative analysis of two-component signal transduction systems of *Bacillus cereus*, *Bacillus thuringiensis* and *Bacillus anthracis*

Mark de Been,1,2,3 Christof Francke,1,2 Roy Moezelaar,1,4 Tjakko Abee1,3 and Roland J. Siezen1,2,5

Correspondence
Mark de Been
mdebeen@cmbi.ru.nl

1Wageningen Centre for Food Sciences (WCFS), Wageningen, The Netherlands
2Centre for Molecular and Biomolecular Informatics (CMBI), Radboud University, PO Box 9101, 6500 HB Nijmegen, The Netherlands
3,4Laboratory of Food Microbiology3 and Food Technology Centre4, Wageningen University and Research Centre, Wageningen, The Netherlands
5NIZO food research BV, Ede, The Netherlands

Members of the *Bacillus cereus* group are ubiquitously present in the environment and can adapt to a wide range of environmental fluctuations. In bacteria, these adaptive responses are generally mediated by two-component signal transduction systems (TCSs), which consist of a histidine kinase (HK) and its cognate response regulator (RR). With the use of *in silico* techniques, a complete set of HKs and RRs was recovered from eight completely sequenced *B. cereus* group genomes. By applying a bidirectional best-hits method combined with gene neighbourhood analysis, a footprint of these proteins was made. Around 40 HK-RR gene pairs were detected in each member of the *B. cereus* group. In addition, each member contained many HK and RR genes not encoded in pairs (‘orphans’). Classification of HKs and RRs based on their enzymic domains together with the analysis of two neighbour-joining trees of these domains revealed putative interaction partners for most of the ‘orphans’. Putative biological functions, including involvement in virulence and host–microbe interactions, were predicted for the *B. cereus* group HKs and RRs by comparing them with those of *B. subtilis* and other micro-organisms. Remarkably, *B. anthracis* appeared to lack specific HKs and RRs and was found to contain many truncated, putatively non-functional, HK and RR genes. It is hypothesized that specialization of *B. anthracis* as a pathogen could have reduced the range of environmental stimuli to which it is exposed. This may have rendered some of its TCSs obsolete, ultimately resulting in the deletion of some HK and RR genes.

INTRODUCTION

The *Bacillus cereus* group consists of Gram-positive, spore-forming bacteria. It includes *B. cereus*, a species often associated with food-borne disease, *Bacillus thuringiensis*, which is used as a biological pesticide worldwide, and *Bacillus anthracis*, a pathogen of warm-blooded animals that can cause the often fatal disease anthrax. Members of the *B. cereus* group form a highly homogeneous subdivision within the genus *Bacillus* and it has been proposed that *B. cereus*, *B. thuringiensis* and *B. anthracis* are in fact varieties of the same species (Daffonchio et al., 2000; Helgason et al., 2000). However, *B. anthracis* and *B. thuringiensis* differ from *B. cereus* by containing plasmid-encoded specific toxins and a capsule (*B. anthracis* only) (Okinaka et al., 1999; Schnepf et al., 1998) and recent studies have shown that *B. anthracis* is rather monomorphic, whereas there is large diversity within *B. cereus* and *B. thuringiensis* (Bavykin et al., 2004; Hill et al., 2004; Priest et al., 2004).

Members of the *B. cereus* group are ubiquitously present in the environment and can adapt to a wide range of environmental conditions (Abee & Wouters, 1999; Jensen et al., 2003; Kotiranta et al., 2000). This raises the question of how these organisms are able to monitor these conditions and respond to them. In bacteria, sensing and adapting to environmental fluctuations is generally mediated by two-component signal transduction systems (TCSs) (Parkinson & Kofoid, 1992; Stock et al., 1989). These systems have been...
shown to monitor a wide variety of conditions, including nutrient deprivation, cold/heat shock, osmotic stress, low pH and many others (Aguilar et al., 2001; Bearer et al., 1998; Jung & Altendorf, 2002; Sun et al., 1996). In addition, TCSs have been shown to initiate important adaptive responses, such as sporulation, biofilm formation and chemotaxis (Jiang et al., 2000; Lyon & Novick, 2004; Szurmunt & Ordal, 2004). TCSs consist of a sensor histidine kinase (HK) and its cognate response regulator (RR), which are often encoded on adjacent genes. A typical HK contains an N-terminal, membrane-associated sensor domain and a C-terminal, cytosolic H-box and HATPase domain. Together, these cytoplasmic domains make up the phosphotransferase domain. A typical RR is a cytosolic protein consisting of an N-terminal receiver domain and a C-terminal DNA-binding domain. Upon sensing specific environmental stimuli the HATPase domain mediates autophosphorylation of the HK at a conserved histidine residue of the H-box. The histidine-bound phosphoryl group is subsequently transferred onto an aspartic acid residue of the RR receiver domain, leading to activation of the RR. The activated RR then binds to specific regions on the DNA, which leads to the activation/repression of genes involved in adaptive responses (Parkinson & Kofoid, 1992; Stock et al., 1989). Besides the prototypical TCSs, in which the phosphoryl group is transferred to the RR in a single step, more complex signal transduction systems also occur in bacteria. In these so-called phosphorelays, activation of the RR by the HK occurs through a multitude of phosphoryl transfer steps (Appleye et al., 1996; Burbulys et al., 1991; Posas et al., 1996).

Although the B. cereus group has received much attention in the past few years and many B. cereus group genomes have recently been sequenced and published (Han et al., 2006; Ivanova et al., 2003; Rasko et al., 2004; Read et al., 2002, 2003), hardly any research has been done on TCSs in this bacterial group. Only recently, a number of HKs has been reported to initiate sporulation in this bacterial group. Only recently, a number of HKs has been shown to initiate sporulation in B. cereus (van Schaik et al., 1996). Besides the prototypical TCSs, in which the phosphoryl group is transferred to the RR in a single step, more complex signal transduction systems also occur in bacteria. In these so-called phosphorelays, activation of the RR by the HK occurs through a multitude of phosphoryl transfer steps (Appleye et al., 1996; Burbulys et al., 1991; Posas et al., 1996).

Although the B. cereus group has received much attention in the past few years and many B. cereus group genomes have recently been sequenced and published (Han et al., 2006; Ivanova et al., 2003; Rasko et al., 2004; Read et al., 2002, 2003), hardly any research has been done on TCSs in this bacterial group. Only recently, a number of HKs has been shown to initiate sporulation in B. cereus (van Schaik et al., 1996). Since so little is known about two-component signal transduction in the B. cereus group, we initiated a computational analysis to predict which kind of TCSs are present in each member of this group and, more importantly, to predict the differences between the members of this group regarding these signal transduction systems.

METHODS

**Sequence information.** Complete genome sequences of the B. cereus group (B. cereus strains ATCC 14579, ATCC 10987 and ZK, B. thuringiensis konkukian and B. anthracis strains Ames, Ames 0581, Sterne and A2012) and B. subtilis 168 were retrieved from the National Center for Biotechnology Information (NCBI) (ftp.ncbi.nlm.nih.gov/genomes/Bacterial/) on 5 October 2004. Sequence information of B. cereus group plasmids was obtained from the NCBI microbial plasmid database (www.ncbi.nlm.nih.gov/genomes/static/ebd_p.html) on 21 July 2005. At this date, one plasmid of B. cereus ATCC 14579 (pBClin15), one of B. cereus ATCC 10987 (pBC10987), five of B. cereus ZK, 12 of B. thuringiensis, six of B. anthracis (3 × pX01, 3 × pX02) and four of B. mycoides were available.

**Sequence analysis.** HMMER 2.3.2 (Durbin et al., 1998) was used for hidden Markov model (HMM) searches against amino acid sequences and a DeCypher hardware-accelerated system (Active Motif) was used to perform HMMER searches against nucleic acid sequences. Protein domain organizations were determined by running HMMER searches against the Pfam database and the SMART (Schultz et al., 1998) HMM databases, using default threshold values, while TMHMM 2.0 (Krogh et al., 2001) was used to detect transmembrane helices. Sequence similarities were detected using the NCBI protein-protein BLAST server (www.ncbi.nlm.nih.gov/blast/bl2seq.cgi) or the NCBI microbial BLAST server (www.ncbi.nlm.nih.gov/sutils/genom_table.cgi). The latter server was used to scan the whole-genome shotgun sequences of B. cereus G9241 and B. anthracis strains A1055, Australia 94, CNEVA-9066, Kruger B, Vollum and Western North America USA6153. Multiple sequence alignments were created with MUSCLE 3.51 (Edgar, 2004) and bootstrapped neighbour-joining (NJ) trees were created with CLUSTAL W 1.83 (Thompson et al., 1994). Trees were visualized with Levels of Orthology through Phylogenetic Trees (LOFT) (R. van der Heijden and others, unpublished results). DNA patterns were detected using PatScan (Douza et al., 1997).

**Identification of HKs and RRs.** The genome and plasmid sequences of the B. cereus group and B. subtilis 168 were searched for genes encoding putative HKs and RRs. To detect these genes, HMMER searches were performed against the protein and nucleic acid datasets of the different genomes and plasmids, using the Pfam HATPase_c (Pfam02518) and Response_reg (Pfam00072) HMMs. The HATPase_c HMM was used to scan for the highly conserved HATPase domain of HKs, while the Response_reg HMM was used to scan for the highly conserved phosphoryl-accepting domain of RRs. Recovered sequences were further scrutinized according to the following criteria: (i) the HATPase domain had to be located in the C-terminus (last 2/3) of the encoded protein and (ii) a putative H-box had to precede the HATPase domain. If no H-box was detected, the H-box was localized by hand. HMMER searches against the B. cereus group nucleic acid datasets were performed to detect HK- and RR-encoding genes for which the ORF prediction was erroneous. In these cases, translation start sites were localized by hand. Frameshifts and/or overlap of more than 75 bp with an existing gene were not allowed.

**Detection of HK-RR gene pairs and ‘orphan’ HK and RR genes.** HKs and their cognate RRs are often encoded on adjacent genes on the DNA. Therefore, all gene clusters containing at least one HK and one RR gene were considered to encode functional HK-RR gene pairs and were thus considered to encode a specific TCS. Single HK and RR genes were categorized as ‘orphans’. The definition of a gene cluster was set as follows: intergenic distances within a cluster had to be less than 300 bp and genes had to lie in the same transcription direction or in a divergent direction on the DNA. A convergent direction was not allowed, since converging genes do not lie in a single operon.

**Detection of orthologous TCSs.** Potential protein orthologues (and in-paralogues) were automatically detected from pairwise species comparisons using INPARANOID 1.35 (Remm et al., 2001). To identify orthologous HK-RR pairs between members of the B. cereus group and B. subtilis, the genomic protein datasets of these species were used as input for INPARANOID. TCSs were regarded as orthologous when both the HKs of TCSs A (species 1) and A’ (species 2) and the RRs of these systems were detected as orthologues (Fig. 1a). When only the HKs of TCSs A and A’ but not the RRs of these systems (or vice versa) were detected as orthologues, the TCSs had to share gene context to be regarded as orthologous systems (Fig. 1b).
The rationale behind this was that gene neighbourhood has been shown to provide strong signals for functional association between gene products within and between species (Dandekar et al., 1998; Overbeek et al., 1999).

**RESULTS AND DISCUSSION**

**Initial identification of HKs and RRs**

The Pfam HMMs HATPase_c and Response_reg were used to recover all TCSs from eight completely sequenced genomes of the *B. cereus* group. The *B. subtilis* genome was scanned in the same way for benchmarking and comparative analysis. As shown in Table 1, 50–58 putative HKs containing a C-terminal HATPase domain preceded by an H-box and 48–52 putative RRs containing a RR receiver domain were detected in the genomes of the *B. cereus* group. In contrast, 35 HKs and 35 RRs were found in *B. subtilis*, which is in agreement with what was found before in this organism (Fabret et al., 1999). Among the total of HK and RR genes detected, 16 had previously been unannotated due to erroneous ORF predictions (gene coordinates are shown in Supplementary Table S1, available with the online version of this paper). For all HKs and RRs detected, the protein domain organization was analysed using TMHMM, Pfam and SMART. The results of these analyses are shown in Supplementary Table S2.

Around 40 HK-RR gene pairs were identified in each genome of the *B. cereus* group, which is about 10 more than the number of pairs found in *B. subtilis*. It is remarkable that, in contrast to *B. subtilis*, the members of the *B. cereus* group contain HK-RR fusion proteins, which have both a HK phosphotransferase domain and a RR phosphoryl-accepting domain. Typically, two fusion proteins were found in each of the three *B. cereus* genomes, whereas only one was found in the *B. thuringiensis* and *B. anthracis* genomes. All HK and RR genes not clustering in HK-RR gene pairs and not encoding fusion proteins were considered ‘orphans’. As many as 10–14 ‘orphan’ HKs and 7–11 ‘orphan’ RRs were found in the members of the *B. cereus* group, compared to six of each in *B. subtilis*. The number of HKs and RRs and their distribution among pairs, fusions and ‘orphans’ was exactly the same for the *B. anthracis* strains, Ames, Ames 0581 and Sterne (Table 1). The numbers shown in Table 1 correspond with those of a recent, more limited, study in which only the genomes of *B. cereus* ATCC 14579, *B. anthracis* A2012 and the draft genome of *B. thuringiensis* Fig. 1.

Detection of orthologous TCSs using INPARANOID. (a) TCS A of species 1 and TCS A' of species 2 are regarded as orthologous because both the HKs of TCSs A and A' and the RRs of these systems are detected as orthologues by INPARANOID. (b) In this situation, only the HKs of TCS A and A' are detected as orthologues. However, systems A and A' are still regarded as orthologous because they share gene context (gene X and gene X' are orthologues).

**Table 1. Number of HK-RR pairs, fusions and ‘orphans’ detected in eight *B. cereus* group genomes and *B. subtilis***

<table>
<thead>
<tr>
<th>Species</th>
<th>HKs</th>
<th>RRs</th>
<th>HK-RR gene pairs</th>
<th>HK-RR fusions</th>
<th>‘Orphans’ Hks</th>
<th>‘Orphans’ RRs</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>B. cereus</em> ATCC 14579</td>
<td>55</td>
<td>48</td>
<td>39</td>
<td>2</td>
<td>14</td>
<td>7</td>
</tr>
<tr>
<td><em>B. cereus</em> ATCC 10987</td>
<td>54</td>
<td>49</td>
<td>40</td>
<td>2</td>
<td>12</td>
<td>7</td>
</tr>
<tr>
<td><em>B. cereus</em> ZK</td>
<td>57</td>
<td>52</td>
<td>43</td>
<td>2</td>
<td>12</td>
<td>7</td>
</tr>
<tr>
<td><em>B. thuringiensis</em> konkukian</td>
<td>58</td>
<td>52</td>
<td>44</td>
<td>1</td>
<td>13</td>
<td>7</td>
</tr>
<tr>
<td><em>B. anthracis</em> Ames (0581), Sterne</td>
<td>52</td>
<td>51</td>
<td>41</td>
<td>1</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td><em>B. anthracis</em> A2012</td>
<td>50</td>
<td>50</td>
<td>38</td>
<td>1</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td><em>B. subtilis</em> 168</td>
<td>35</td>
<td>35</td>
<td>29</td>
<td>0</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>
israelensis were scanned for the amount of TCSs (Anderson et al., 2005).

Since main differences between members of the B. cereus group have been attributed to their plasmids (Okinaka et al., 1999; Rasko et al., 2005; Schneff et al., 1998), the DNA of 29 B. cereus group plasmids was also scanned for genes encoding TCSs. Surprisingly, only plasmid pBc10987 of B. cereus ATCC 10987 appeared to encode a TCS, while one ‘orphan’ RR was found on the megaplasmid pE33L466 of B. cereus ZK (results not shown). Apparently, plasmid-encoded features, such as toxin production and host specificity, are not regulated by specific plasmid-encoded TCSs.

**Classification of HKs and RRs**

As a basis for assigning biological functions to the HKs and RRs detected, we made a classification of these proteins. To that end, two bootstrapped NJ trees were constructed, a HK tree and a RR tree. The HK tree was constructed with all B. cereus group and B. subtilis HK phosphotransferase domains, while the RR tree was constructed with all RR receiver domains. In addition to this initial set of sequences, homologous sequences of other bacterial species were included to improve the resolution of both trees (the HK and RR trees are included as Supplementary Figs S1 and S2 with the online version of this paper). Information on the nature of the RR output domains, as identified using Pfam and SMART, was also added. Based on the two trees and the RR output domains detected, we were able to classify the B. cereus group HKs and RRs into the subfamilies described by Grebe & Stock (1999), who discerned HK and RR subfamilies on a similar basis. The results of the classification procedure are shown in Table 2. Analysis of the two trees showed that the receiver domains of all RRs pairing to a HK of a certain class generally clustered together in the same branches of the RR tree. Furthermore, their DNA-binding output domains also roughly fell into distinct groups. For example, all RRs pairing with a class 7 HK contained a NarL-like output domain. These results are in agreement with the findings of Grebe & Stock (1999), who suggested that the HK phosphotransferase domains, the cognate receiver domains and the RR output domains have evolved as integral units.

**Function prediction: a footprint analysis including B. subtilis**

Several classes of TCSs have been shown to function in distinct cellular processes. TCSs consisting of a class 4, 5, 9 and 10 HK are known to be involved in sporulation initiation, C4-dicarboxylate metabolism, chemotaxis and quorum sensing, respectively (Asai et al., 2000; Grebe & Stock, 1999; Jiang et al., 2000; Kaspar & Bott, 2002; Lyon & Novick, 2004; Szurmant & Ordal, 2004; Tanaka et al., 2003; Yamamoto et al., 2000; Zientz et al., 1998). The fact that members of the B. cereus group contain HKs of these classes strongly suggests that some of their TCSs mediate the signals that initiate the processes described above.

To get a more specific functional annotation of the B. cereus group HKs and RRs, they were compared with those of B. subtilis, which is the model Gram-positive organism and for which relatively much is known about the functionality of its TCSs. We used INPARANOID (Remm et al., 2001) to detect protein orthologues. With this program, the protein datasets of the B. cereus group were compared with each other and with the protein dataset of B. subtilis. From the INPARANOID output, we were able to detect HK-RR pairs, fusions and ‘orphans’ shared between the different B. cereus group genomes and between each B. cereus group genome and B. subtilis. The resulting footprint is shown in Table 2. The B. cereus group appeared to share as many as 20 orthologous HK-RR pairs and six ‘orphans’ with B. subtilis. Not all these HKs and RRs were found in every single B. cereus group genome. For example, the well-characterized B. subtilis HK CheA is absent from B. anthracis. In contrast, the well-characterized B. subtilis systems ResED, PhoRP, YycGF, YufLM, LiaSR and components of the B. subtilis sporulation initiation phosphorelay were found in all members of the B. cereus group (see Table 2, column 11 for biological functions). Interestingly, some well-known B. subtilis TCSs appeared to be absent from the B. cereus group. Among these were the systems CssSR, BceSR, DesKR and DegSR.

**Function prediction: TCSs putatively involved in antibiotic resistance/production and virulence**

The B. cereus group HKs and RRs were also compared with those of other bacterial species, using the NCBI BLAST server. Maintaining an E-value cut-off of $1 \times 10^{-15}$, we found a number of B. cereus group TCSs to be similar to systems with a known biological function (Table 2). Among the functionally defined systems, many are known to respond to cell-wall-acting antibiotics or general cell-envelope stresses, such as CesKR, CroSR, VanSR, VanSRb and VraSR (Arthur & Quintiliani, 2001; Comenge et al., 2003; Evers & Courvalin, 1996; Kallipolitis et al., 2003; Kuroda et al., 2000), and many are known to function in lantibiotic production and resistance, such as SpaKR, NisKR, BacSR and SalKR (Engelke et al., 1994; Klein et al., 1993; Neumuller et al., 2001; Upton et al., 2001).

We could also identify TCSs putatively involved in virulence and host–microbe interactions. Among these were TCSs 24, 25 and 26, which are similar to LisKR of *Listeria monocytogenes*, ArlSR of *Staphylococcus aureus* and CiaHR of streptococci (Table 2). LisKR plays an important role in cellular responses of *L. monocytogenes* to ethanol, pH, hydrogen peroxide and antimicrobials, but also contributes to the virulence potential of this organism ( Cotter et al., 1999, 2002). ArlSR mediates the expression of many genes involved in autolysis, cell division and virulence (Liang et al., 2005) and CiaHR has been suggested to regulate maintenance of the cell envelope (e.g. modifications of
Table 2. Classification, footprint analysis and function prediction of the *B. cereus* group HKs and RR

Column 1 contains the codes referring to the *B. cereus* group HKs and RRs. A translation of these codes to NCBI codes can be found in Supplementary Table S3. Columns 2 and 3 show the classification of HKs and RRs, respectively. The classification into HK and RR subfamilies was based on the classification described by Grebe & Stock (1999). Null, RR does not contain an output domain. Columns 4–9 show the HKs and RR detected in each genome of the *B. cereus* group. *Bce*; *B. cereus*; *Bth*; *B. thuringiensis*; *Ban*; *B. anthracis*; strain Ames (0581) and Sterne; §strain A2012; □ HK-RR pair; □ 'orphan' HK; □ 'orphan' RR; □ HK-RR fusion protein; □ tyrosine kinase; □ N-terminally truncated HK; □ RR with truncated output domain. Orthologous HK-RR pairs and 'orphans' detected in *B. subtilis* are in bold in column 10. All other, homologous, HK-RR pairs and 'orphans' are in normal type. %, amino acid identity; *, conserved gene neighbourhood with the corresponding *B. cereus* group HK and RR gene(s). Species name abbreviations: *Bli*, *Bacillus licheniformis*; *Bsu*, *Bacillus subtilis*; *Cpe*, *Clostridium perfringens*; *Efa*, *Enterococcus faecalis*; *Efc*, *Enterococcus faecium*; *Eco*, *Escherichia coli*; *Lpl*, *Lactobacillus plantarum*; *Lmo*, *Listeria monocytogenes*; *Lla*, *Lactococcus lactis*; *Sau*, *Staphylococcus aureus*; *Smu*, *Streptococcus mutans*; *Spn*, *Streptococcus pneumoniae*. Column 11 shows the biological functions predicted for the *B. cereus* group HKs and RRs. References: CesKR, Kallipolitis et al. (2003); CroSR, Comenge et al. (2003); VanSR, Arthur & Quintiliani (2001); VanSRa, Evers & Courvalin (1996); YvrGH, Serizawa et al. (2005); SpaKR, Klein et al. (1993); NisKR, Engelke et al. (1994); BacSR, Neumuller et al. (2001); LysKR, Cotter et al. (1999, 2002); ArlSR, Liang et al. (2005); CiaHR, Guenzi et al. (1994), Mascher et al. (2003b), Throup et al. (2000); LcoSR, Liu et al. (2002); ResED, Nakano et al. (1996); SrrBA, Yarwood et al. (2001); PhoRP, Sun et al. (1996); YycGF, Fabret & Hoch (1998); VicKR, Dubrac & Msadek (2004), Martin et al. (1999), Mohedano et al. (2005); GlnKL, Satomura et al. (2005); YycQP, YsdKJ, Lmo; CroSR, 34/41, Efa; VanSR, 31/41, Efc Cell wall stress response, antibiotic resistance03 1a OmpR VanSR, 35/44, Lmo; CroSR, 35/48, Efa; VanSR, 36/47, Efc Cell wall stress response, antibiotic resistance07 1a OmpR VanSR, 35/44, Lmo; CroSR, 35/48, Efa; VanSR, 36/47, Efc Cell wall stress response, antibiotic resistance08 1a OmpR VanSR, 35/44, Lmo; CroSR, 35/48, Efa; VanSR, 36/47, Efc Cell wall stress response, antibiotic resistance09 1a OmpR Unknown10 1a – Unknown11 1a OmpR Unknown12 1a OmpR Unknown13 1a OmpR Unknown14 1a OmpR Unknown15 1a OmpR Unknown16 1a OmpR Unknown17 1a OmpR Unknown18 1a OmpR Unknown19 1a OmpR Unknown20 1a Null Unknown21 1a OmpR Unknown22 1a OmpR Unknown

**Ref. HK** | **RR** | **Bce** | **Bce** | **Bce** | **Bth** | **Ban** | **Ban** | **Orthologous (B. subtilis) and homologous systems,** | **Predicted function**
---|---|---|---|---|---|---|---|---|---
01 | la | OmpR | 14579 | 10987 | ZK | konk. | | | CPE0235-6*, 29/49, Cpe | Virulence, carbohydrate uptake/metabolism
02 | la | OmpR | | | | | | CesKR, 33/45, Lmo; CroSR, 34/41, Efa; VanSR, 31/41, Efc Cell wall stress response, antibiotic resistance
03 | la | OmpR | 37/47 | 36/45 | | | | CesKR, 30/42, CroSR, 39/50, Efa; VanSR, 36/46, Efc Cell wall stress response, antibiotic resistance
04 | la | OmpR | 35/47 | 36/46 | | | | CesKR, 35/47, Lmo; CroSR, 39/57, Efa; VanSR, 36/46, Efc Cell wall stress response, antibiotic resistance
05 | la | OmpR | 37/46 | 36/47 | | | | CesKR, 37/51, Lmo; CroSR, 39/47, Efa; VanSR, 36/47, Efc Cell wall stress response, antibiotic resistance
06 | la | OmpR | 35/44, Lmo; CroSR, 35/44, Efa; VanSR, 36/47, Efc Cell wall stress response, antibiotic resistance
07 | la | OmpR | 30/38, Lmo; CroSR, 30/38, Efa | | | | | | VanSRa, 25/42, Efa | Cell wall stress response, antibiotic resistance
08 | la | OmpR | | | | | | SpaKR*, 24/38, Bsu | Lantibiotic production/resistance
09 | la | OmpR | | | | | | SpaKR*, 31/52, Bsu; NisKR*, 25/40, Lla | Lantibiotic production/resistance
10 | la | OmpR | | | | | | YvrGH, 38/57, Bsu | Cell envelope maintenance
11 | la | OmpR | | | | | | BacSR*, 54/72, Bli | Unknown
12 | la | OmpR | | | | | | YcbML*, 58/65, Bsu; BacSR*, 28/45, Bli | Unknown
13 | la | OmpR | | | | | | BacSR*, 54/72, Bli | Unknown
14 | la | OmpR | | | | | | Lmo; CroSR, 34/41, Efa; VanSR, 31/41, Efc Cell wall stress response, antibiotic resistance
15 | la | OmpR | | | | | | Lmo; CroSR, 34/41, Efa; VanSR, 31/41, Efc Cell wall stress response, antibiotic resistance
16 | la | OmpR | | | | | | Lmo; CroSR, 34/41, Efa; VanSR, 31/41, Efc Cell wall stress response, antibiotic resistance
17 | la | OmpR | | | | | | Lmo; CroSR, 34/41, Efa; VanSR, 31/41, Efc Cell wall stress response, antibiotic resistance
18 | la | OmpR | | | | | | Lmo; CroSR, 34/41, Efa; VanSR, 31/41, Efc Cell wall stress response, antibiotic resistance
19 | la | OmpR | | | | | | Lmo; CroSR, 34/41, Efa; VanSR, 31/41, Efc Cell wall stress response, antibiotic resistance
20 | la | Null | | | | | | Unknown | Unknown
21 | la | OmpR | | | | | | Unknown | Unknown
22 | la | OmpR | | | | | | Unknown | Unknown
<table>
<thead>
<tr>
<th>Ref.</th>
<th>HK class</th>
<th>RR output</th>
<th>Bce</th>
<th>Bce</th>
<th>Bce</th>
<th>Bth</th>
<th>Ban†</th>
<th>Ban‡</th>
<th>Orthologous (B. subtilis) and homologous systems, id % HK/id % RR</th>
<th>Predicted function</th>
</tr>
</thead>
<tbody>
<tr>
<td>23</td>
<td>la OmpR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>YkoHG, 30/46, Bsu; LisKR, 26/41, Lmo; ArlSR, 27/43, Sau</td>
<td>Cell envelope maintenance, virulence</td>
</tr>
<tr>
<td>24</td>
<td>la OmpR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>YkoHG, 28/48, Bsu; LisKR, 34/56, Lmo; ArlSR, 31/48, Sau</td>
<td>Cell envelope maintenance, virulence</td>
</tr>
<tr>
<td>25</td>
<td>la OmpR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CiaHR, 34/44, Spn; CiaHR, 28/45, Smu; YkoHG, 26/40, Bsu</td>
<td>Cell envelope maintenance, virulence</td>
</tr>
<tr>
<td>26</td>
<td>la OmpR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LcoSR, 35/53, Lla</td>
<td>Metal resistance/uptake/metabolism</td>
</tr>
<tr>
<td>27</td>
<td>la OmpR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ResED*, 47/70, Bsu; SrrBA*, 30/65, Sau</td>
<td>Aerobic/anaerobic respiration, virulence</td>
</tr>
<tr>
<td>28</td>
<td>la OmpR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>PhoRP*, 48/73, Bsu</td>
<td>Phosphate uptake/metabolism</td>
</tr>
<tr>
<td>29</td>
<td>la OmpR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>GlnKL*, 54/73, Bsu; YvcQP*, 46/69, Sau</td>
<td>Amino acid (Gln) uptake/metabolism</td>
</tr>
<tr>
<td>30</td>
<td>la OmpR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>YbdKJ*, 50/67, Bsu</td>
<td>Unknown</td>
</tr>
<tr>
<td>31</td>
<td>la OmpR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>YvcQP*, 34/60, Bsu</td>
<td>Cell wall stress response, antimicrobial resistance</td>
</tr>
<tr>
<td>32</td>
<td>la OmpR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>YxdKJ*, 38/55, Bsu</td>
<td>Cell wall stress response, antimicrobial resistance</td>
</tr>
<tr>
<td>33</td>
<td>la OmpR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>YxdKJ, 40/52, Bsu; YvcQP, 32/48, Bsu</td>
<td>Cell wall stress response, antimicrobial resistance</td>
</tr>
<tr>
<td>34</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>KinA, 33, Bsu</td>
<td>Sporulation initiation</td>
</tr>
<tr>
<td>35</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>KinC, 36, Bsu</td>
<td>Sporulation initiation</td>
</tr>
<tr>
<td>36</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>KinE, 33, Bsu</td>
<td>Sporulation initiation</td>
</tr>
<tr>
<td>37</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>KinE, 34, Bsu</td>
<td>Sporulation initiation</td>
</tr>
<tr>
<td>38</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>KinE, 35, Bsu</td>
<td>Sporulation initiation</td>
</tr>
<tr>
<td>39</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>KinC, 27, Bsu; KinE, 34, Bsu</td>
<td>Sporulation initiation</td>
</tr>
<tr>
<td>40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>KinB, 36, Bsu</td>
<td>Sporulation initiation</td>
</tr>
<tr>
<td>41</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>KinD*, 38, Bsu</td>
<td>Sporulation initiation</td>
</tr>
<tr>
<td>42</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>KinB, 40, Bsu</td>
<td>Sporulation initiation</td>
</tr>
<tr>
<td>43</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>KinE, 37, Bsu</td>
<td>Sporulation initiation</td>
</tr>
<tr>
<td>44</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>KinB, 32, Bsu</td>
<td>Sporulation initiation</td>
</tr>
<tr>
<td>45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>KinB, 26, Bsu</td>
<td>Sporulation initiation</td>
</tr>
<tr>
<td>46</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>YufLM*, 43/52, Bsu; DcuSR, 38/44, Eco</td>
<td>C4-dicarboxylate (citrate) uptake/metabolism</td>
</tr>
<tr>
<td>47</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>YvfTU*, 43/62, Bsu; DesKR, 37/62, Bsu</td>
<td>C4-dicarboxylate (citrate) uptake/metabolism</td>
</tr>
<tr>
<td>48</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>YfJK*, 21/38, Bsu</td>
<td>C4-dicarboxylate (malate) uptake/metabolism</td>
</tr>
<tr>
<td>49</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>YdfHI, 36/60, Bsu</td>
<td>Membrane fatty acid saturation/desaturation</td>
</tr>
<tr>
<td>50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LiaSR*, 42/64, Bsu</td>
<td>Unknown</td>
</tr>
<tr>
<td>51</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ComPA, 26/38, Bsu</td>
<td>Natural competence</td>
</tr>
<tr>
<td>52</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LiaSR, 35/60, Sau</td>
<td>Unknown</td>
</tr>
<tr>
<td>53</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DecSR, 31/29, Bsu; CitAB, 32/36, Eco</td>
<td>C4-dicarboxylate (citrate) uptake/metabolism</td>
</tr>
<tr>
<td>54</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CitST*, 44/48, Bsu</td>
<td>C4-dicarboxylate (citrate) uptake/metabolism</td>
</tr>
<tr>
<td>55</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>YvfTU*, 43/62, Bsu; DesKR, 37/62, Bsu</td>
<td>C4-dicarboxylate (malate) uptake/metabolism</td>
</tr>
<tr>
<td>56</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Membrane fatty acid saturation/desaturation</td>
<td>Unknown</td>
</tr>
<tr>
<td>57</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>YdfHI, 36/60, Bsu</td>
<td>Unknown</td>
</tr>
<tr>
<td>58</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LiaSR*, 42/64, Bsu</td>
<td>Cell wall stress response, antimicrobial resistance</td>
</tr>
<tr>
<td>59</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ComPA, 26/38, Bsu</td>
<td>Cell wall stress response, antimicrobial resistance</td>
</tr>
<tr>
<td>60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LiaSR, 35/60, Sau</td>
<td>Cell wall stress response, antimicrobial resistance</td>
</tr>
<tr>
<td>61</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>YdfHI, 36/60, Bsu</td>
<td>Cell wall stress response, antimicrobial resistance</td>
</tr>
</tbody>
</table>

**Table 2. cont.**
peptidoglycan), virulence and repression of competence
(Guenzi et al., 1994; Mascher et al., 2003b; Throup et al., 2000). Because of their similarity with LisKR, ArlSR and CiaHR, it is conceivable that TCSs 24, 25 and 26 of the B. cereus group also play a role in virulence. However, the TCSs described above influence many different processes, indicating that their primary function is, for example, to maintain the cell envelope, which has great influence on the virulence potential of an organism.

Another virulence-associated system of the B. cereus group might be the class 10 TCS 68. In Gram-positive bacteria, class 10 TCSs are known as quorum-sensing systems. They function as intercellular communication modules that use small peptides as signalling molecules. After processing, the peptides are exported and sensed by other cells via the sensory domains of the HK. In this way, distinct cellular processes are generated in a cell-density-dependent manner (Lyon & Novick, 2004). A well-known example of a quorum-sensing system is AgrACDB of S. aureus. The propeptide AgrD is processed and secreted by AgrB and is then sensed by the HK AgrC. The phosphoryl group is transferred to the RR AgrA, which mediates transcription of 

\[ \text{agrACDB} \]

and RNAIII from two promoters. RNAIII is an intracellular effector that targets the production of virulence factors (Tegmark et al., 1998). Other known agr-like systems are ComDE of streptococci and FsrCA of Enterococcus faecalis, both involved in virulence (Lyon & Novick, 2004), and LamCA of Lactobacillus plantarum, which mediates the production of surface proteins and cell adherence (Sturme et al., 2005). Like many agr-like modules, TCS 68 of the B. cereus group might function as a quorum-sensing system, regulating the production of virulence factors and mediating host–microbe interactions. Analysis of the B. cereus genomes did not reveal a putative signalling-peptide-encoding gene nor an agrB-like gene in the near vicinity of the TCS genes, but it has to be pointed out that, across species, the signalling peptides and the AgrB-like processing enzymes often share low sequence similarity, making it difficult to detect novel ones with in silico techniques (Lyon & Novick, 2004).

Other important virulence regulators of the B. cereus group may be ResED, PhoRP and YycGF (TCSs 29, 31 and 32, respectively). They form a group of TCSs that is highly conserved in the low-G+C Gram-positives. In S. aureus, ResED (SrrBA) represses the production of staphylococcal exotoxin and surface-associated virulence factors under low-oxygen conditions (Yarwood et al., 2001), while YycGF has been shown to maintain the cell envelope (Fabret & Hoch, 1998; Martin et al., 1999; Throup et al., 2000). ResED and YycGF have also been implicated in the regulation of respiration, phosphate uptake and maintenance of the cell envelope (Mohedano et al., 2005; Nakano et al., 1996; Sun et al., 1996). Moreover, YycGF (VicKR) has been shown to be essential in a number of organisms (Yarwood et al., 2001). ResED and YycGF have also been implicated in the regulation of virulence factors in several pathogens. In S. aureus, the resED locus is involved in the production of extracellular virulence factors and maintenance of the cell envelope (Mohedano et al., 2005; Nakano et al., 1996; Sun et al., 1996). Moreover, YycGF (VicKR) has been shown to be essential in a number of organisms (Yarwood et al., 2001). ResED and YycGF have also been implicated in the regulation of virulence factors in several pathogens. In S. aureus, the resED locus is involved in the production of extracellular virulence factors and maintenance of the cell envelope (Mohedano et al., 2005; Nakano et al., 1996; Sun et al., 1996). Moreover, YycGF (VicKR) has been shown to be essential in a number of organisms (Yarwood et al., 2001). ResED and YycGF have also been implicated in the regulation of virulence factors in several pathogens. In S. aureus, the resED locus is involved in the production of extracellular virulence factors and maintenance of the cell envelope (Mohedano et al., 2005; Nakano et al., 1996; Sun et al., 1996). Moreover, YycGF (VicKR) has been shown to be essential in a number of organisms (Yarwood et al., 2001). ResED and YycGF have also been implicated in the regulation of virulence factors in several pathogens. In S. aureus, the resED locus is involved in the production of extracellular virulence factors and maintenance of the cell envelope (Mohedano et al., 2005; Nakano et al., 1996; Sun et al., 1996). Moreover, YycGF (VicKR) has been shown to be essential in a number of organisms (Yarwood et al., 2001).
antigens (Dubrac & Msadek, 2004). Because of the implicated role of the above-mentioned systems in virulence and because of the apparent conservation of their RR binding sites across species (Dubrac & Msadek, 2004), we scanned the B. cereus group genomes with the B. subtilis binding sites for ResD [5’-TA(T/A)T(C/T)TG(T/G)(A/C)-3’], PhoP [5’-T(T/A/C)ACA-N3 to N7-T(T/A/C)ACA-3’] and YycF [5’-TG(T/A)T(A/T)/C)N3-TGT(A/T)A(T/C)-3’] (Howell et al., 2003; Makita et al., 2004). Just as in B. subtilis, the ResD, PhoP and YycF binding sites were detected upstream of genes involved in respiration (e.g. resB), phosphate transport (e.g. pstA, pstC) and cell division (e.g. ftsE, ftsX), respectively (results not shown). Interestingly, we detected putative ResD binding sites 54 bp upstream of the haemolysin II-encoding gene of B. cereus ATCC 14579, 85 bp upstream of the haemolysin A-encoding gene of all B. cereus group genomes and 55 bp upstream of the capsule-encoding gene (capA) of plasmid pXO2. We did not find any putative PhoP or YycF binding sites upstream of genes clearly involved in virulence. The results suggest that ResED might regulate the virulence-associated genes described above. We are currently working on a more extended promoter analysis, which may shed light on the complicated transcriptional network of these RRs.

Another system that is possibly involved in virulence is TCS 01. This TCS is similar to a system of unknown function (CPE0235/CPE0236) of Clostridium perfringens 13 (Table 2). In both the HK and the RR tree, the phosphoryl-transferring domains of these systems clustered closely together in distinct branches, indicating that the TCSs are highly related. Furthermore, the genes encoding the TCSs appeared to share strong gene neighbourhood conservation. Given these data, we conclude that these TCSs are specific for the B. cereus group and C. perfringens and that the shared genes lie in one operon with the TCS genes. Based on the neighbouring genes, which encode putative (carbohydrate) transport systems, and the fact that C. perfringens is a notorious pathogen of humans and animals, these TCSs might be virulence-associated, functioning in the breakdown of host tissues and the subsequent import of nutrients.

The predicted functions of the B. cereus group TCSs, as revealed by the comparative analyses, are shown in column 11 of Table 2. Column 10 and the table legend give information on detected gene context conservation.

**HK-RR fusion proteins**

Although many HKs and RRs could be assigned putative biological functions, the function of a large number is still completely unknown. For instance, it is unclear what role the two HK-RR fusion proteins fulfil and whether they interact with other HKs and/or RRs. In general, HK-RR fusion proteins are involved in more complex phosphorelays (Appleby et al., 1996). Fusion protein 20, found in all members of the B. cereus group investigated, might function in a phosphorelay similar to the Sln1-Ypd1-Ssk1 phosphorelay of Saccharomyces cerevisiae (Posas et al., 1996). Activation of the protein probably results in phosphoryl transfer from its HK phosphotransferase domain to its own RR receiver domain. Subsequent steps may include phosphoryl transfer to the H-box of a second protein and, finally, to the RR receiver domain of a third protein that carries a RR output domain. Fusion protein 58, which was only found in B. cereus, is probably not involved in such a phosphorelay. The fact that it contains a DNA-binding domain suggests that it functions as a single unit. However, TMHMM predicted the protein to be membrane-bound (Supplementary Table S2), which seems to conflict with its putative role as a transcriptional regulator. Typically, fusion protein 58 does not share any sequence similarity with other HK-RR fusion proteins, indicating that it is unique for B. cereus. To shed light on the biological role of the two HK-RR fusion proteins, we are currently investigating these B. cereus proteins in our laboratory.

**Matching of ‘orphans’**

*In silico* detection of HKs and RRs in members of the B. cereus group revealed a relatively large number of ‘orphans’. To uncover the signal transduction routes in which these ‘orphans’ are involved, we compared the NJ trees (the HK and RR trees described above) of the interacting domains and coupled ‘orphans’ on the basis of cognate clustering within these trees. This method was successfully employed by us before (C. Francke and others, unpublished results) and it has been shown that HKs and RRs that are known to interact fall into corresponding phylogenetic subfamilies (Grebe & Stock, 1999, Korete et al., 2000).

For most ‘orphans’, a putative partner HK or RR could be predicted. For example, the distribution of the ‘orphan’ class 1a HK 10 in the HK tree was identical to that of the ‘orphan’ RR 69 in the RR tree, suggesting that HK 10 and RR 69 act together in a TCS (Fig. 2a). The fact that RR 69 contains an OmpR output domain strengthens this assignment, as class 1a HKs generally act with RRs containing these DNA-binding domains.

The largest group of ‘orphans’ that were matched to partner proteins was the group of class 4 HKs (Fig. 2b). In B. subtilis, these HKs have been shown to act in the sporulation initiation phosphorelay, transferring a phosphoryl group to the ‘orphan’ RR Spo0A via the ‘orphan’ single-domain RR Spo0F and the phosphotransferase Spo0B. The multi-component structure of this transduction route provides for many levels of regulation, including the input of several environmental signals by the different HKs (Burbulys et al., 1991; Jiang et al., 2000). Orthologues of Spo0F, Spo0B and Spo0A were found in all members of the B. cereus group, indicating that these species use a similar phosphorelay. While B. subtilis contains five class 4 HKs (KinA, B, C, D and E), members of the B. cereus group contain a larger number of these HKs, suggesting that they contain an even more extended system with more signal inputs. In B. anthracis, nine class 4 HKs were detected, while as many as 14 were...
detected in *B. cereus* ATCC 14579. The HK tree shows that all these HKs clustered within or close to branches containing one of the *B. subtilis* sporulation HKs. In fact, they only clustered in branches containing HKs of species known to form endospores. Class 4 HK 39 clustered closest to HKs of non-spore-forming bacteria, such as AtoS of *Escherichia coli*. However, overexpression of a HK 39 orthologue in *B. thuringiensis* EG1351 has been shown to bypass sporulation defects and a spo0F mutation in different *B. thuringiensis* strains (Malvar et al., 1994). In addition, it has recently been shown that HKs 39, 40, 48 (KinD orthologue), 49 (KinB orthologue) and 50 are capable of inducing sporulation in *B. anthracis* (Brunsing et al., 2005).

In addition to predicting putative partners for the ‘orphans’ described above, putative partners were found for ‘orphan’ RR 73 (LytT homologue) and 74 (CheV orthologue). However, these RRs were not matched to ‘orphan’ HKs, but to HKs already found in HK-RR pairs (Fig. 2c, d). In the RR tree, RR 73 clustered close to a branch containing RR 65 (LytT orthologue). Since RR 73 also contains a LytTR output domain, we hypothesize that the class 8 HK 65 (LytS orthologue) is not only capable of phosphorylating its cognate RR 65, but can also transfer a phosphoryl group to RR 73. The fact that the RR 73-encoding gene shares gene context with LytST orthologues of other species (e.g. TTE0871/TTE0870 of *Thermoanaerobacter tengcongensis* MB4) and the fact that it clusters with genes putatively involved in cell envelope maintenance, the confirmed function of LytST (Bruskill & Bayles, 1996), further strengthens this prediction. Similarly, the ‘orphan’ RR 74 (CheV orthologue) was matched to the class 9 HK 67 (CheA orthologue). In the RR tree, RR 74 clustered with CheV of *B. subtilis*, which is known to accept a phosphoryl group from the chemotactic signal modulator CheA (Szurmant & Ordal, 2004). Since HK 67 clustered together with CheA in the HK tree, it is likely that phosphoryl transfer from CheA to CheV occurs in *B. cereus* and *B. thuringiensis*. In *B. anthracis*, a frameshift mutation has probably rendered cheA non-functional (Fig. 3c), leaving CheV and CheY (the RR that pairs with CheA) as ‘orphans’. In addition to cheA, the cheV gene of *B. anthracis* also carries a frameshift mutation, encoding a putative CheV protein without a CheW domain. This suggests that the complete chemotaxis system of *B. anthracis* is non-functional. The fact that *B. anthracis* carries truncations in other genes of the flagellar gene cluster (Read et al., 2003), and the fact that most

![Fig. 2. Matching of ‘orphans’.](http://mic.sgmjournals.org)}
B. anthracis strains are non-motile (Turnbull, 1999), strengthens this hypothesis.

Besides CheY and CheV in B. anthracis, other ‘orphans’ could not be matched to putative partners. For example, using the methods described above, we could not find a putative partner for the ‘orphan’ RR RsbY (RR 75), which is responsible for activating the alternative sigma factor $\sigma^B$ in B. cereus (van Schaik et al., 2005).

**Differences in TCSs within the B. cereus group**

As already mentioned, differences were found within the B. cereus group regarding the number of HK-RR fusion proteins, the number of sporulation HKs and the chemotaxis machinery. In addition, other remarkable differences were found. Strikingly, a number of TCSs appeared to be truncated in all four B. anthracis strains (for examples, see Fig. 3). Besides the truncation in CheV, truncations were found in the B. anthracis TCSs 02, 34, 38, 43, 53 and 63. These systems were regarded as truncated since their HK sensory or their RR output domains are reduced by at least 50 amino acids as compared to their orthologues in the other B. cereus group genomes. Two other systems (TCSs 09 and 36) were not regarded as truncated in B. anthracis, but they differ by having a slightly shorter RR output (TCS 09) or HK sensory domain (TCS 36). Closer analysis of the B. anthracis genome sequences showed that the truncations were not caused by such trivialities as gene annotation errors. Moreover, the fact that the truncations were found in all four B. anthracis strains reduces the chance of sequencing errors as the cause for finding these truncations.

The truncations in the putative genes encoding HKs 02, 34, 38 and 63 and RR 53 presumably render their corresponding proteins non-functional, since no sensory domains are left in the HKs and no output domain is left in the RR. Interestingly, many of the truncated TCSs are similar to systems known to respond to cell-wall-acting antibiotics or cell-envelope stresses in general (TCSs 02, 36, 38 and 63). Since a distinguishing feature of B. anthracis is its susceptibility to penicillin (Turnbull, 1999), it is possible that one (or more) of these TCSs contributes to penicillin resistance in B. cereus and B. thuringiensis and that it is indeed non-functional in B. anthracis. Recent work has shown that penicillin-susceptible B. anthracis strains contain silent $\beta$-lactamase genes, while these genes are active in penicillin-resistant members of the B. cereus group (Chen et al., 2003, 2004). Given these data, it is plausible that one or more of the non-truncated TCSs in B. cereus and B. thuringiensis provide a route for activation of the $\beta$-lactamase genes, while their truncated orthologues in B. anthracis are unable to activate these genes.

Although the truncations in the B. anthracis TCSs may indicate the inactivity of these systems, it has to be mentioned that next to the genes encoding the truncated HKs, putative genes encoding the ‘missing’ sensory domains were found. For example, we found that the putative gene upstream of the truncated HK 63 gene actually encodes the two ‘missing’ GAF domains (Fig. 3b). The presence of putative genes encoding the ‘missing’ sensory domains leaves open the possibility that the truncated HKs are part of functional systems. It is conceivable that these HKs can somehow interact with the proteins containing their ‘missing’ sensory domains, thereby forming three-component systems. An example of such a system might be YycHGF of B. subtilis. YycH, which is located external to the cell membrane, has been proposed to function as an extracellular sensor that confers its activity to the HK YycG (Szurmant et al., 2005). Another possibility is that the truncated HKs are relieved from sensory constraints and

---

**Fig. 3.** Examples of truncated and degraded HKs in B. anthracis. Upper genes are of B. cereus ATCC 14579. Lower genes are corresponding orthologues in B. anthracis Sterne. (a) The gene encoding the putative sporulation HK 43 is truncated in B. anthracis. However, the gene is probably still functional, since the part encoding the two PAS domains and the enzymic HK domains is still intact. (b) The truncation in the gene encoding HK 63 of B. anthracis has probably rendered this gene non-functional, since the translated HK would have no sensory domains left. (c) A frameshift between the H-box and the HATPase-encoding parts of cheA (HK 67 in B. cereus) has probably rendered this gene non-functional in B. anthracis.
are therefore more active than their non-truncated orthologues.

In addition to the truncated TCSs, some systems appeared completely absent from all four B. anthracis strains. Among these were, as already mentioned, fusion protein 58 (also absent from B. thuringiensis) and CheA, but also ComPA (TCS 59), the system that regulates natural competence in B. subtilis (Lyon & Novick, 2004), and the two putative sporulation HKs 46 and 47. Fragments of some of the corresponding genes were still found in the B. anthracis genomes. cheA, for example, is disrupted by a frameshift, separating the H-box- from the HATPase-encoding sequence (Fig. 3c).

To examine the nature of the TCSs described above in other B. cereus group genomes, their HK and RR protein sequences were compared to the whole-genome shotgun sequences of B. cereus G9241 and B. anthracis strains A1055, Australia 94, CNEVA-9066, Kruger B, Vollum and Western North America USA6153. As shown in Supplementary Table S4, not all the TCS truncations/deletions detected in B. anthracis strains Ames, Ames 0581, Sterne and A2012 were found in the six additional B. anthracis strains. Perhaps most remarkable was the detection of a complete cheV gene in all the newly sequenced B. anthracis genomes. However, all the new strains (except A1055) do contain a disrupted cheA gene, indicating that their chemotaxis machinery is non-functional. Except for the cheV gene, most of the TCS truncations/deletions were found in the new strains, indicating that some TCSs are generally degraded in or completely absent from B. anthracis.

Concluding remarks

In this paper we describe the results of an in silico comparative analysis of the TCSs of the B. cereus group. With the use of Pfam HMMs, 50–58 HKs and 48–52 RRs were detected in each member of the B. cereus group. A footprint analysis of these HKs and RRs, including those of B. subtilis, revealed which of these proteins are shared between the different members of the B. cereus group, which ones are specific for certain members and which ones are shared between the B. cereus group and B. subtilis. In addition, we were able to assign putative interaction partners for most of the ‘orphan’ HKs and RRs detected by using a congruence-of-trees analysis.

The combination of these in silico techniques revealed interesting differences within the B. cereus group. For example, the fact that B. anthracis contains fewer class 4 ‘orphan’ HKs than B. cereus and B. thuringiensis indicates that its sporulation initiation machinery is somewhat less fine-tuned than this mechanism is in its closest relatives. Besides the reduced number of sporulation HKs, other TCS genes appeared to be absent from or truncated in B. anthracis. If the truncated genes are indeed non-functional, the effective number of B. anthracis TCSs would be drastically reduced compared to the number in B. cereus and B. thuringiensis. This would suggest that B. anthracis is less capable of processing extracellular signals than its close relatives, which may proliferate in more fluctuating environments. It has been proposed that B. anthracis evolved as a pathogen of warm-blooded animals early in the evolution of the B. cereus group, while the other members of this group kept exploiting more fluctuating environments (e.g. invertebrate guts, plant rhizospheres and supplemented soils) (Jensen et al., 2003; Turnbull, 1999). B. anthracis might have a more specialized pathogenic lifecycle than the other members of the B. cereus group. It probably survives in the environment mainly in the form of dormant endospores. Upon ingestion by herbivores, spores germinate to form toxin-producing vegetative cells that kill the host. Death of the host results in the release of large numbers of B. anthracis cells into the environment. These cells probably sporulate immediately upon contact with air, completing the B. anthracis life cycle (Jensen et al., 2003; Rasko et al., 2005). Specialization of B. anthracis as a pathogen could have reduced the range of environmental stimuli to which it is exposed. This might have rendered some TCSs obsolete, ultimately resulting in the inactivation of HK and RR genes. This hypothesis is in agreement with earlier results, which showed that bacteria that inhabit relatively stable host environments generally encode fewer signalling systems than environmental bacteria with the same genome size (Galperin, 2005).

With this work, we provide the first in-depth analysis of the complete TCS arsenal of the B. cereus group. By scanning different B. cereus group genomes for HK- and RR-encoding genes, we have gained insights into the capacity of these organisms to adapt to changes in their environment. The results presented here provide a basis for future research on signal transduction mechanisms in the B. cereus group.

ACKNOWLEDGEMENTS

The authors would like to thank Richard Notebaart and Maarten Mols for useful discussions and Bernadet Renckens for gene context conservation analyses.

REFERENCES


Okinaka, R. T., Cloud, K., Hampton, O. & 12 other authors (1999). Sequence and organization of pXO1, the large Bacillus anthracis plasmid harboring the anthrax toxin genes. J Bacteriol 181, 6509–6515.


