Published in final edited form as:

Res Microbiol. 2010 December ; 161(10): 830–837. doi:10.1016/j.resmic.2010.08.002.

Guidelines for nomenclature assignment of Ger receptors

Christian A Ross and Ernesto Abel-Santos

Christian A Ross: rossc6@unlv.nevada.edu; Ernesto Abel-Santos: ernesto.abelsantos@unlv.edu

^aDepartment of Chemistry, University of Nevada, Las Vegas, Las Vegas, NV, 89154, USA

^bSchool of Life Science, University of Nevada, Las Vegas, Las Vegas, NV, 89154, USA

Abstract

Some members of the orders *Bacillales* and *Clostridiales* form dormant spores when subjected to environmental stress. These resistant spores return to normal vegetative growth upon encountering nutrients. A family of membrane-bound proteins called germination (Ger) receptors is tasked with detecting metabolites that serve as germination signals. During the characterization of tri-partite Ger receptor proteins from the genera *Bacillus* and *Clostridium*, we found numerous nomenclature inconsistencies. In this work, we issued guidelines to remediate this problem. We generated a sample of sequenced *Bacillus*, *Clostridium* and related endospore-forming genera. Ger receptor proteins were recovered from these genomes by PSI-BLAST. The resulting Ger receptor protein sequences were then clustered by neighbor-joining based on sequence alignment. Inconsistencies in Ger receptor labeling were noted and new names proposed. Finally, a systematic approach for naming of new Ger receptors was designed.

Keywords

Ger receptors; *Bacillus*; *Clostridium*; Spore germination

1. Introduction

In the Gram-positive spore-formers of the orders *Bacillales* and *Clostridiales*, endospores are formed when survival conditions are suboptimal for vegetative growth. These environmentally hardened spores remain viable for long periods until conditions improve. It is the function of germination (Ger) receptors to detect small molecules signals that trigger the emergence of spores from dormancy. Each Ger receptor is generally expressed from a tricistronic *ger* receptor operon. The tripartite receptors form a protein family, some members of which have established roles in recognizing specific germinants (Atluri et al., 2006; Fisher Hanna, 2005; Hornstra et al., 2006). The Ger receptor complex is believed to be composed of two putative transmembrane proteins termed subunit A and subunit B (e.g., GerAA and GerAB) and a third, potentially membrane anchored protein termed subunit C (e.g., GerAC) (Ross Abel-Santos, 2010).

© 2010 Elsevier Masson SAS. All rights reserved.

Correspondence to: Christian A Ross, rossc6@unlv.nevada.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Ger receptors were first named in *Bacillus subtilis* after a concerted effort by geneticists to map and phenotype mutants in the spore germination pathway (Moir et al., 1979). The first studied Ger receptor responded to a single germinant, L-alanine, and was designated GerA (Yasuda Tochikubo, 1984). Two other Ger receptors were shown to act cooperatively to trigger germination with either L-asparagine supplemented with glucose, fructose, and potassium ions (GFK) or L-alanine supplemented with GFK (Irie et al., 1996). These two receptors were designated GerB and GerK, respectively.

The first designations for genes encoding Ger receptors in *B. subtilis*, *gerA* and *gerB*, were regarded as temporary (Moir et al., 1979), but have remained in use without regards for phylogenetic relationships or functional identity. Furthermore, *B. subtilis* also possesses two hypothetical *ger* receptor operons, *yfkQRT* and *yndDEF* (Paidhungat Setlow, 2000) that encode proteins with no assigned function.

Among the *Bacillus cereus* species group, six *ger* receptor operons (*gerA*, *gerH*, *gerK*, *gerL*, *gerS*, and *gerY*) are found in the sequenced *B. anthracis* Ames genome (Fisher Hanna, 2005; Read et al., 2003) and a seventh (*gerX*) is located on the virulence plasmid pX01. The closely related *B. cereus* strain E33L has seven chromosomally located *ger* receptor operons. Because of inconsistent nomenclature, seven labels (*gerA*, *gerB*, *gerH*, *gerI*, *gerK*, *gerL* and *gerQ*) have been entered into GenBank (Han et al., 2006). An additional *ger* receptor partial operon located on plasmid pE33L466 has also been labeled *gerK*. Another strain of the same species, *Bacillus cereus* ATCC 14579, was found to have seven *ger* receptor operons (Ivanova et al., 2003) and they were labeled *gerG*, *gerI*, *gerK*, *gerL*, *gerQ*, *gerR* and *gerS* (Hornstra et al., 2006). The difference in nomenclature in evolutionarily related *ger* receptor operons creates confusion in the identification of individual genes.

With the expansion of the number of Ger receptors characterized and the availability of a substantial number of completely sequenced bacterial genomes, the limitations of the existing designations have become apparent. In many cases, inconsistent or confusing notations have been applied. For example, directly orthologous Ger receptors in closely related *Bacillus* species have been given different names (e.g. the proteins encoded by the *gerH* operon in *B. anthracis* and the *gerI* operon in *B. cereus* are almost identical). Furthermore, designations have become largely independent of their evolutionary and functional roles. For example, *gerK* has been used to designate *Clostridium* and *Bacillus* *ger* receptor operons that have little sequence homology.

In the present study, we introduce a systematic approach to clarify the nomenclature of Ger receptors. Our goal is to make a minimum number of nomenclature changes, to avoid changing protein names that already reflect functional roles. Furthermore, we attempt to reflect consistent evolutionary relationships as best as they can be inferred at present. In order to systematize the Ger receptor nomenclature, we selected a representative sample of sequenced *Bacillus*, *Clostridium* and related endospore formers. We then extracted proteins encoded by Ger receptor operons through PSI-BLAST searches, and clustered them via protein sequence alignment and neighbor-joining. After reviewing the Ger receptor names included in the associated Genbank records, we proposed a series of nomenclature changes, provided a list of Entrez accession numbers for these receptors and their orthologues, and suggested guidelines for future Ger receptor annotations.

2. Materials and methods

2.1. Construction of germination protein database

Protein PSI-BLAST was performed on October 28, 2008, using GerAA (NP_391185), GerAB (NP_391186), GerAC (NP_391187) as queries. The results were used as seeds to

develop a position-specific weight matrix (PSWM) model to search the Genbank sequenced genome database. Data from PSWM were used to develop a protein family (Altschul et al., 1997). The result of each PSI-BLAST iteration was filtered against the list of organisms in Table S1. PSI-BLAST results with E-values of less than $1e-15$ were discarded. PSI-BLAST was continued until convergence was achieved for all three Ger receptor subunits (no new sequences discovered with subsequent searches).

2.2. Protein alignment

Each family of related proteins from the PSI-BLAST search was imported into MEGA 4.0 software (Tamura et al., 2007). Proteins were aligned with ClustalW using the Gonnet matrix. The aligned proteins were then trimmed to remove leading and trailing short peptides not common to all members of the family. Sequences that were incomplete or aligned poorly were not included in phylogenetic reconstruction. The resultant trimmed alignments were then used to construct neighbor-joining trees in MEGA software using homogenous Poisson correction for estimating distances (Nei Kumar, 2000). Complete deletion was used for missing sites. Bootstrap values were calculated from 1,000 replicate trees.

2.3. Procedure for nomenclature assignment in *B. cereus* ATCC14579 and *B. megaterium* QM B1551

Identification of *ger* receptor operons encoding the GerG, GerR and GerL receptors was accomplished by nucleotide BLAST against the NCBI genome sequence of *B. cereus* ATCC 14579 (NC_004722) using the specificity regions of primers provided (Hornstra et al., 2006) as in silico probes. The remaining Ger receptors were retrieved from *B. cereus* ATCC 14579 by BLAST using an amino acid sequence of *B. subtilis* GerAA. Identification of the *B. megaterium* QM B1551 GerU receptor, GerVB, and GerWB proteins was accomplished by BLAST using the amino acid sequence of *B. subtilis* GerAA.

Presumptive identification of Ger receptor protein orthologues from the list of organisms in Table S1 was accomplished by protein BLAST. To confirm membership in each orthologous grouping, protein sequences were imported into the existing alignment, realigned using ClustalW, and a new neighbor-joining tree constructed as described above.

3. Results and discussion

3.1. Number of receptors

BLAST searches conducted with GerAA from *B. subtilis* and limited to one sample strain from each sporulating bacteria species with a finished sequenced genome (listed in Supplementary Table S1) resulted in 139 hits in the NCBI RefSeq database, 20 of which were identified as SpoVAF-like proteins. Similarly, 125 and 112 sequences were discovered using GerAB, and GerAC as seeds. Use of alternate Ger receptor proteins (GerXA, GerXB, GerXC from pXO1) as seed sequences did not affect the results. In all, the number of encoded Ger receptors varies between species from one to seven distinct operons. For example, the *B. anthracis* 'Ames strain' appears to contain four intact chromosomally encoded *ger* operons, with one additional operon located on virulence plasmid pXO1 (in strain 'Ames ancestor'). Additionally, the *B. anthracis* 'Ames' genome contains two non-functional *ger* receptor operons, *gerA* and *gerY*. These cryptic operons contain frame shift mutations in either the B or C subunits.

3.2. Phylogeny of the germination receptor families

Neighbor-joining trees showing an estimated phylogeny of the GerAA, GerAB, and GerAC protein families are shown in Figs. S1–S3, respectively (supplementary material). The three

families are very similar in organization, suggesting a high degree of coevolution, as has been observed previously (Igarashi Setlow, 2005; Moir, 2006; Paidhungat Setlow, 1999). In the case of the GerAA-like family among the *Bacillus* genus, we found that these proteins fell into four distinct clades: the GerA/L (Fig. 1), the GerH(I)/Y (Fig. 2), the GerS(K) (Fig. 2) and the non S-like GerK proteins (Fig. 3). The GerX receptor does not consistently cluster within any other *Bacillus* groupings and thus form a unique clade. The proteins within each cluster, except in rare cases, are generally more closely related to orthologous proteins from other *Bacillus* species than they are to proteins within the same organism from other clusters. As an example, *B. anthracis* GerHA and GerYA form a distinct cluster from GerSA and GerKA, as do GerLA and GerAA. These groupings also hold true for the B- and C-subunits of Ger receptors. For example, GerHB and GerYB cluster (Fig. S2), as do GerHC with GerYC (Fig. S3).

From these results, we have defined the following archetypical Ger receptor clade names: GerA, GerI, GerK, GerY, GerS, and GerX. These denominations will serve to define individual clusters and will serve as the basis for naming related Ger receptors.

3.3. Receptors related to GerA from *B. subtilis*

The *gerA*, *gerB* and hypothetical *yndDEF* operons from *B. subtilis* encode proteins that are all closely related and thus cluster together (Fig. 1 and Fig. S1–S3). Functionally, the GerA receptor in *B. subtilis* is associated with germination in response to L-alanine, whereas the GerB receptor, together with the GerK receptor, responds to L-asparagine and sugars (Atluri et al., 2006; Cabrera-Martinez et al., 2003). Interestingly, single amino acid changes in the GerB receptor allow it to respond to D-alanine (Paidhungat Setlow, 1999). Thus, the GerA and GerB receptors are related both phylogenetically and functionally. Although it is evident that the GerB receptor is in the same clade as the GerA receptor, we decided to keep the designation GerB to minimize nomenclature changes, since this receptor has been characterized extensively under that name.

The nearest GerA paralogues within the *B. cereus* species group are encoded by an uncharacterized operon also called *gerA*. In the case of the *B. anthracis* 'Ames' strain, the *gerA* operon carries a frameshift likely rendering it non-functional. We find that the truncated receptor encoded by the *gerA* operon from members of the *B. cereus* species group are more closely related to the Ger receptor hypothetically encoded by *B. subtilis yndDEF* than to those encoded by either the *B. subtilis gerA* or *gerB* operons. Furthermore, orthologous proteins to those encoded by the *gerA* and *yndDEF* operons are found in *B. licheniformis*, while *gerB* and *gerA*-encoded orthologues are found in *B. amyloliquifaciens*. This suggests that the *gerA* and *gerB* loci are the result of duplication events in the *B. subtilis* species group after divergence from the *B. cereus* species group.

Since a single species can contain more than one Ger receptor from a grouping, we suggest using numerical designation to distinguish between closely related receptors. For example, since *B. subtilis* possess three different GerA-type receptors, we suggest retaining the labels GerA and GerB to minimize nomenclature changes, but to relabel the hypothetical YndDEF as GerA2. The GerA receptor in the *B. cereus* species group would also become GerA2 to emphasize its relationship to YndDEF. Under this regime, *B. cereus* species will generally lack GerA and GerB receptors, as there are no clear orthologues for these receptors in this species group. In cases where there are more than one Ger receptor from the same clade (Christie et al., 2010), the different proteins should be differentiated by using subscript numerical labels.

The GerL receptor from *B. cereus*, *B. anthracis* and *B. thuringiensis* also clusters with the GerA group, albeit distantly. Interestingly, GerL receptors are also associated with

germination in the presence of L-alanine (Barlass et al., 2002; Fisher Hanna, 2005). Since the GerL receptor is paralogous, yet more distant to the GerA, GerB, and hypothetical YndDEF receptor cluster, the designation GerL was retained for this group to reduce nomenclature changes of previously characterized receptors.

3.4 The GerI(H), GerY and GerQ receptors in *B. cereus* and related spp

The GerI, GerY, and GerQ receptors are all closely related among members of the *B. cereus* species group included in this study. The GerI and GerH receptor components are homologous and the names GerI and GerH have even been used interchangeably within the same receptor. For example, GerHA (YP_038778) is encoded in an operon with GerIB, (YP_038779) in *B. thuringiensis*. A similar situation is found for GerHA (YP_086059) and GerIB (YP_086060) in *B. cereus*. The GerH receptor from *B. anthracis* is associated with germination in the presence of aromatic amino acids and inosine as co-germinants (Weiner et al., 2003). Similarly, the GerI receptor from *B. cereus* is necessary for inosine-mediated germination (Clements Moir, 1998). Thus, this Ger receptor family seems to be involved in the recognition of aromatic compounds.

The nomenclature ambiguity is further complicated by the application of the GerIB label to a *B. thuringiensis* B-subunit protein (YP_037655) that is encoded by an operon that also encodes an A-subunit protein labeled GerKA (YP_037654) and a C-subunit protein labeled GerQC (YP_037656). Moreover, all three subunits of this Ger receptor appears to be most closely related to GerS from *B. anthracis*. We suggest that those proteins that are more closely related to the GerI receptor in *B. cereus* ATCC 10876 (AF067645) be labeled GerI. These would include many Ger receptors currently labeled GerH. Those proteins previously labeled GerI which cluster with the GerS receptor from *B. anthracis* 'Ames ancestor' (AAP27387–AAP27389), should be labeled GerS1.

GerQ receptors have also been implicated in inosine recognition. GerIA in *B. thuringiensis* serovar *konkukian* is nearly identical to a GerQA protein found in *B. cereus* ATCC 14579 (not shown), suggesting this comprises a separate grouping distinct from other GerI receptors. Similarly, GerQ proteins from *B. thuringiensis* and *G. thermodenitrificans* do not clearly fit into either GerY or GerI clusters, although they are clearly members of the larger GerY/I clade. Relabeling these GerI, to which they are most closely aligned, may create confusion with other GerH and GerI proteins from these organisms, suggesting that the GerQ designation be maintained. In the case of the *B. thuringiensis* GerI, it should be considered to be a member of the GerQ grouping to avoid confusion with GerH (now GerI).

The GerY receptor has no identified function, but clusters closely with those in the GerI and GerQ receptor groups. The proteins of the GerA receptor from *B. thuringiensis* (YP_035020-YP_035022), along with those of the GerA receptor of *B. cereus* E33L, cluster closer to the GerY receptor from *B. anthracis* than with any *B. subtilis* GerA homologue. This suggests that these proteins are mislabeled and are really members of the GerY group.

3.5. Receptors labeled GerK and GerS

The clustering of GerK receptors presents a serious difficulty as the use of the GerK label has been applied to a large number of germination receptor proteins without clear phylogenetic rationale. These genes form deeply branching groups and their phylogeny is unclear, with low bootstrap values between similarly labeled GerK proteins (Supplementary Fig. S1–S3).

Thus, there are a number of GerK-like proteins in public databases with differing levels of similarity to canonical *B. subtilis* GerK receptor subunits. For consistency, the receptor designation GerK would be reserved for those proteins relatively close in homology to GerK

in *B. subtilis* (Fig. 3). It should be mentioned that hypothetical A- and C-subunit proteins encoded by the *yfkQ* and *yfkR* loci in *B. subtilis* weakly cluster to the GerK group, but the corresponding B-subunit encoded by *yfkT* does not. As the functionality of the hypothetical YfkQTR receptor has not been established, and its similarity to other proteins is weak, its designation is left untouched. Other Ger receptors currently labeled GerK in the *Bacillus* and closely related genera should be renamed according to the archetypal group where they most closely cluster.

The GerS receptor has been shown to respond to aromatic ring structures in *B. anthracis* (Ireland Hanna, 2002). Even though aromatic compounds activate both GerS and GerI receptors, these operons are phylogenetically distinct and hence are designated by independent nomenclatures. Furthermore, GerS receptors cluster into two distinct subgroups that we designated GerS1 and GerS2. In *B. cereus* E33L, two different Ger receptors have been labeled GerK. The subunit proteins from one receptor (YP_245633, YP_245636-YP_245637) are more closely related to the *B. subtilis* GerK receptor, and should retain the label GerK. The subunit proteins from the second Ger receptor (YP_082150-YP_082153) are most similar to GerS receptors, and thus should be relabeled GerS1. Related GerK-labeled receptors in other *B. cereus* group members would be better labeled as GerS2 or similar (Fig. 2).

The solitary GerQ designation in *B. cereus* E33L is used in an operon encoding proteins labeled GerKA, GerIB and GerQC (YP_084868-YP_084870). Each of these subunits clusters within the GerS2 receptor group, and should so be named (Fig. 2, Supplementary Fig. S1–S3).

3.6. The GerX receptor

The GerX receptor is encoded by the virulence plasmid pXO1 in *B. anthracis* and was named for germination, pXO1 related (Guidi-Rontani et al., 1999). The GerX receptor is known to mediate germination within murine macrophages (Hu et al., 2007). The GerX receptor is unusual in that it is more closely related to Ger receptors occurring outside the *Bacillus* genus than it is to any other Ger receptor found within this genus. GerXA (YP_016488.2) appears more deeply branching than other *Bacillus* Ger receptor A-subunits, suggesting either loss of more closely related proteins within *Bacillus* species or a possible horizontal transfer event. The nearest BLAST hits for GerXA outside of *B. anthracis* or *B. cereus* are in the recently sequenced thermophile *Anoxybacillus flavithermus* WK1, *Brevibacillus brevis*, and *Geobacillus* spp. These observations are similar for the other GerX receptor subunits, GerXB and GerXC.

3.7. Clostridial receptors

A cursory glance at the clustering of Ger receptors in the *Clostridium* genus reveals that there are nearly as many different receptor clades as have been seen in the better studied *Bacillus* genus; however, the existing annotation is sparser.

There appear to be at least four well defined Ger receptor clusters in the *Clostridium* genus, currently defined by GerB1 in *C. tetani*, GerKA2 in *C. kluyveri*, GerKA3 in *C. kluyveri* and GerKA in *C. tetani* E88 (most similar to SpoVAF).

Unfortunately, the germination responses in the *Clostridium* genus have been studied in less detail than in the *Bacillus* genus. As a result, there is a lack of functional information for *Clostridium* Ger receptors. This makes it difficult to assign appropriate nomenclature to the grouping of *Clostridium* Ger receptor proteins. However, it is clear that the use of GerK and/or GerB labels for such a diverse group with little similarity to the canonical GerK and GerB receptors is inappropriate. We suggest using the temporary label GerC to denote Ger

receptors in *Clostridium* species. Appropriate numerical designations can then be used to denote to which group a new Ger receptor clusters. Thus, proteins clustering with GerB1 in *C. tetani* will be designated GerC1, homologues to GerKA2 in *C. kluyveri* will be designated GerC2, homologues of GerKA3 in *C. kluyveri* will be designated GerC3 and homologues of GerKA in *C. tetani* E88 will be designated GerC4.

3.8. Assignment of nomenclature for Ger operons of *B. cereus* ATCC 14579 and *B. megaterium* QM B1551

B. cereus ATCC 14579 encodes seven Ger receptors (GerG, GerI, GerK, GerL, GerQ, GerR and GerS) (Hornstra et al., 2006). As discussed above, these designations were applied arbitrarily. To assess the robustness of the nomenclature assignment rules proposed, we tested each Ger receptor sequence from *B. cereus* ATCC 14579 to determine their individual phylogenetic relationship with the Ger receptor clades determined above. Based on these relationships, we changed GerG to GerA2, GerK to GerS2, GerR to GerY, and GerS to GerS1 to better reflect individual Ger receptor evolutionary kinship. GerI, GerL and GerQ correctly clustered with their group, and their designations were not modified (See Table 1). These assignments are consistent with previously published reports (Hornstra et al., 2006).

B. megaterium QM B1551 encodes a plasmid-based Ger receptor (GerU) and two monocistronic B-subunit proteins (GerVB and GerWB) (Christie et al., 2010; Christie Lowe, 2007). Similar to *B. subtilis* GerK protein (Atluri et al., 2006), all three *B. megaterium* B-subunits have been shown to be involved in glucose recognition (Christie et al., 2010). As expected due to functional similarities, GerUABC, GerVB, and GerWB all cluster with the GerK archetypical receptors. Thus, we suggest renaming GerU as GerK₁, GerVB as GerKB₂, and GerWB as GerKB₃. The subscript numeral will then refer to multiple proteins from the same clade encoded in the same genome.

3.9. Conclusions

In this study, we created systematic rules for the assignment of nomenclature to Ger receptors. A summary of suggested nomenclature changes is shown in Fig. 1 and Table 1. As expected, the suggested changes allow consistency in labeling both in phylogenetic and functional analysis. This removes ambiguity and avoids confusion caused by inconsistent labeling. Furthermore, newly identified receptors may be assigned a putative function by clustering with characterized Ger receptors.

Although Ger receptors in the *Clostridium* genus are as phylogenetically varied as those in the *Bacillus* genus, there is not enough genomic and functional information to clearly define their groupings. As more *Clostridium* genomes are sequenced and more germination receptors are characterized, we expect to further define the Ger receptor clades in *Clostridium* species.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors want to thank Prof. Eduardo Robledo for helpful discussions. E Abel-Santos was supported by USDA grant number 2010-65119-20603 and by the National Science Foundation under CHE-0957400. C Ross was supported by NIH grant number P20 RR-016464 from the INBRE Program of the National Center for Research Resources.

References

- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl Acids Res* 1997;25:3389–3402. [PubMed: 9254694]
- Atluri S, Ragkousi K, Cortezzo DE, Setlow P. Cooperativity between different nutrient receptors in germination of spores of *Bacillus subtilis* and reduction of this cooperativity by alterations in the GerB receptor. *J Bacteriol* 2006;188:28–36. [PubMed: 16352818]
- Barlass PJ, Houston CW, Clements MO, Moir A. Germination of *Bacillus cereus* spores in response to L-alanine and to inosine: the roles of gerL and gerQ operons. *Microbiol* 2002;148:2089–2095.
- Cabrera-Martinez R-M, Tovar-Rojo F, Vepachedu VR, Setlow P. Effects of overexpression of nutrient receptors on germination of spores of *Bacillus subtilis*. *J Bacteriol* 2003;185:2457–2464. [PubMed: 12670969]
- Christie G, Gotzke H, Lowe CR. Identification of a receptor subunit and putative ligand-binding residues involved in the *Bacillus megaterium* QM B1551 spore germination response to glucose. *J Bacteriol*, JB 2010:4317–4326.
- Christie G, Lowe CR. Role of chromosomal and plasmid-borne receptor homologues in the response of *Bacillus megaterium* QM B1551 spores to germinants. *Journal of Bacteriology* 2007;189:4375–4383. [PubMed: 17434971]
- Clements MO, Moir A. Role of the gerI operon of *Bacillus cereus* 569 in the response of spores to germinants. *J Bacteriol* 1998;180:6729–6735. [PubMed: 9852021]
- Fisher N, Hanna P. Characterization of *Bacillus anthracis* germinant receptors *in vitro*. *J Bacteriol* 2005;187:8055–8062. [PubMed: 16291679]
- Guidi-Rontani C, Pereira Y, Ruffie S, Sirard J-C, Weber-Levy M, Mock M. Identification and characterization of a germination operon on the virulence plasmid pXOI of *Bacillus anthracis*. *Mol. Microbiol* 1999;33:407–414. [PubMed: 10411756]
- Han CS, Xie G, Challacombe JF, Altherr MR, Bhotika SS, Bruce D, Campbell CS, Campbell, et al. Pathogenomic sequence analysis of *Bacillus cereus* and *Bacillus thuringiensis* isolates closely related to *Bacillus anthracis*. *J Bacteriol* 2006;188:3382–3390. [PubMed: 16621833]
- Hornstra LM, de Vries YP, Wells-Bennik MH, de Vos WM, Abee T. Characterization of germination receptors of *Bacillus cereus* ATCC 14579. *Appl Environ Microbiol* 2006;72:44–53. [PubMed: 16391023]
- Hu H, Emerson J, Aronson AI. Factors involved in the germination and inactivation of *Bacillus anthracis* spores in murine primary macrophages. *FEMS Microbiol Lett* 2007;272:245–250. [PubMed: 17521404]
- Igarashi T, Setlow P. Interaction between individual protein components of the GerA and GerB nutrient receptors that trigger germination of *Bacillus subtilis* spores. *J Bacteriol* 2005;187:2513–2518. [PubMed: 15774895]
- Ireland JA, Hanna PC. Amino acid- and purine ribonucleoside-induced germination of *Bacillus anthracis* Δ Sterne endospores: gerS mediates responses to aromatic ring structures. *J Bacteriol* 2002;184:1296–1303. [PubMed: 11844758]
- Irie R, Fujita Y, Kobayashi M. Nucleotide sequence and gene organization of the *gerK* spore germination locus of *Bacillus subtilis* 168. *J Gen Appl Microbiol* 1996;42:141–153.
- Ivanova N, Sorokin A, Anderson I, Galleron N, Candelon B, Kapatral V, Bhattacharyya A, Reznik G, et al. Genome sequence of *Bacillus cereus* and comparative analysis with *Bacillus anthracis*. *Nature* 2003;423:87–91. [PubMed: 12721630]
- Moir A. How do spores germinate? *J. Appl. Microbiol* 2006;101:526–530. [PubMed: 16907803]
- Moir A, Lafferty E, Smith DA. Genetics analysis of spore germination mutants of *Bacillus subtilis* 168: the correlation of phenotype with map location. *J Gen Microbiol* 1979;111:165–180. [PubMed: 110906]
- Nei, M.; Kumar, S. *Molecular Evolution and Phylogenetics*. 1 edn. Oxford University Press; 2000.
- Paidhungat M, Setlow P. Isolation and characterization of mutations in *Bacillus subtilis* that allow spore germination in the novel germinant D-alanine. *J Bacteriol* 1999;181:3341–3350. [PubMed: 10348844]

- Paidhungat M, Setlow P. Role of Ger proteins in nutrient and nonnutrient triggering of spore germination in *Bacillus subtilis*. *J Bacteriol* 2000;182:2513–2519. [PubMed: 10762253]
- Read TD, Peterson SN, Tourasse N, Baillie LW, Paulsen IT, Nelson KE, Tettelin H, Fouts DE, et al. The genome sequence of *Bacillus anthracis* Ames and comparison to closely related bacteria. *Nature* 2003;423:81–86. [PubMed: 12721629]
- Ross C, Abel-Santos E. The Ger receptor family in sporulating bacteria. *Cur. Issues Molec. Biol* 2010;12:147–158.
- Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4.0. *Mol Biol Evol* 2007;24:1596–1599. [PubMed: 17488738]
- Weiner MA, Read TD, Hanna PC. Identification and characterization of the *gerH* operon of *Bacillus anthracis* endospores: a differential role for purine nucleosides in germination. *J Bacteriol* 2003;185:1462–1464. [PubMed: 12562819]
- Yasuda Y, Tochikubo K. Relation between D-glucose and L- and D-alanine in the initiation of germination of *Bacillus subtilis* spores. *Microbiol Immunol* 1984;28:197–207. [PubMed: 6427562]

Fig. 1.

The GerA/L clade. *B. subtilis* GerAA and proteins related by amino acid homology; a section of the complete neighbor-joining tree shown in Fig. S1. The percentage of replicate trees showing this clustering under bootstrap testing (1000 replicates) is shown next to the branches. Distance is given in number of amino acid substitutions per site, with 341 positions used to calculate distances.

Fig. 2.

The GerH(I)/Y and GerS clades. *B. cereus* GerHA, GerIA, and other proteins related by amino acid homology; a section of the complete neighbor-joining tree shown in Fig S2. The percentage of replicate trees showing this clustering under bootstrap testing (1,000 replicates) is shown next to the branches. Distance is given in number of amino acid substitutions per site, with 341 positions used to calculate distances.

Figure 3.

The GerK clade. *B. subtilis* GerKA and other proteins related by amino acid homology; a section of the complete neighbor-joining tree shown in Fig. S3. The percentage of replicate trees showing this clustering under bootstrap testing (1000 replicates) is shown next to the branches. Distance is given in number of amino acid substitutions per site, with 341 positions used to calculate distances.

Table 1

List of proposed gene cluster assignments

Organism	Genbank accessions	Current designation			Cluster
		A subunit	B subunit	C subunit	
<i>B. subtilis</i> 168	NP_391185.1				
	NP_391186.1	GerAA	GerAB	GerAC	A
	NP_391187.1				
	NP_391461.1				
	NP_391462.1	GerBA	GerBB	GerBC	B
	NP_391463.1				
	NP_388252.1				
	NP_388254.1	GerKA	GerKB	GerKC	K
	NP_388253.1				
	NP_388660.1				
	NP_388657.1	YfkQ	YfkT	YfkR	N/C
	NP_388659.1				
	NP_389658.1				
	NP_389659.1	YndD	YndE	YndF	A2
	NP_389660.1				
<i>B. anthracis</i> Ames	NP_843166.1				
	NP_843165.1	GerKA	GerKB	GerKC	S2
	NP_843164.1				
	NP_843235.1				
	NP_843236.1	GerLA	GerLB	GerLC	L
	NP_843237.1				
	NP_843286.1	GerYA	N/E	GerYC	Y
	NP_843285.1				
	NP_845469.1	GerAA	N/E	N/E	A2
	NP_845901.1				
	NP_845902.1	GerSA	GerSB	GerSC	S
	NP_845903.1				

Organism	Genbank accessions	Current designation			Cluster
		A subunit	B subunit	C subunit	
	NP_847178.1				
	NP_847179.1	GerHA	GerHB	GerHC	I
	NP_847180.1				
<i>B. anthracis</i> A2084 'Ames ancestor' plasmid pXO-1	YP_016488.2				
	YP_016487.2	GerXA	GerXB	GerXC	X
	YP_016489.2				
	YP_034893.1	GerKA	N/E	GerKC	S2
	YP_034891.1				
	YP_034965.1				
	YP_034966.1	GerLA	GerLB	GerLC	L
	YP_034967.1				
	YP_035022.1	GerAA	N/D	GerAC	Y
	YP_035020.1				
	YP_035021.1				
	YP_037212.1				
	YP_037211.1	GerIA	GerIB	GerIC	Q
	YP_037210.1				
	YP_037228.1				
<i>B. thuringiensis</i> serovar konkukian str. 97-27	YP_037227.1	GerAA	GerAB	GerAC	A2
	YP_037226.1				
	YP_037654.1				
	YP_037655.1	GerKA	GerIB	GerQC	S
	YP_037656.1				
	YP_038778.1				
	YP_038779.1	GerHA	GerHB	GerHC	I
	YP_038780.1				
<i>B. amyloliquefaciens</i> FZB42	YP_001420017.1				
	YP_001420019.1	GerKA	GerKB	GerKC	K

Organism	Genbank accessions	Current designation			Cluster
		A subunit	B subunit	C subunit	
	YP_001420018.1				
	YP_001422580.1				
	YP_001422581.1	GerAA	GerAB	GerAC	A
	YP_001422582.1				
	YP_001422853.1				
	YP_001422854.1	GerBA	GerBB	GerBC	B
	YP_001422855.1				
	YP_001422866.1				
	YP_001422867.1	GerKA1	GerKB1	GerKC1	S2
	YP_001422868.1				
<i>B. cereus</i> E33L	YP_082152.1				
	YP_082151.1	GerKA	GerKB	GerKC	S
	YP_082150.1				
	YP_082225.1				
	YP_082226.1	GerLA	GerLB	GerLC	L
	YP_082227.1				
	YP_082270.1				
	YP_082268.1	GerA	GerB	GerAC	Y
	YP_082269.1				
	YP_084440.1				
	YP_084439.1	GerAA	GerBB	GerBC	A2
	YP_084438.1				
	YP_086059.1				
	YP_086060.1	GerHA	GerHB	GerHC	I
	YP_086061				
	YP_084868.1				
	YP_084869.1	GerKA	GerIB	GerQC	S2
	YP_084870.1				
<i>B. cereus</i> plasmid pE33L466	YP_245637.1	GerKA	GerKB	GerKC	N/C

Organism	Genbank accessions	Current designation			Cluster
		A subunit	B subunit	C subunit	
<i>B. cereus</i> ATCC14579	YP_245633.1				
	YP_245636.1				
	YP_245634.1	N/E	N/D	N/E	N/C
	NP_830451.1				
	NP_830450.1	GerKA	GerKB	GerKC	S2
	NP_830449.1				
	NP_834435.1				
	NP_834436.1	GerIA	GerIB	GerIC	GerIC
	NP_834437.1				
	NP_83306.1				
	NP_83307.1	GerSA	GerSB	GerSC	S1
	NP_83308.1				
	NP_830574.1				
	NP_830573.1	GerRA	GerRB	GerRC	Y
	NP_830572.1				
	NP_832842.1				
	NP_832841.1	GerQA	GerQB	GerQC	Q
	NP_832840.1				
	NP_832854.1				
	NP_832853.1	GerGA	GerGB	GerGC	A2
	NP_830517.1				
	NP_830518.1	GerLA	GerLB	GerLC	L
<i>B. megaterium</i> QM B1551	YP_003565875.1				
	YP_003565877.1	GerUA	GerUB	GerUC	K ₁
	YP_003565876.1				
	YP_003565878.1	N/E	GerVB	N/E	K ₂
	YP_003562283.1	N/E	GerWB	N/E	K ₃

N/E: Protein not encoded. N/D: Gene is present, but the protein has not been designated in the database. N/C: Protein sequence does not cluster with any of the designated clades. Current designation derived from the /gene or /product field of the Genbank record with the exception of GerG, GerR and GerL from *B. cereus* ATCC14579 (see methods)