Genome-Wide Analysis of Cell Type-Specific Gene Transcription during Spore Formation in Clostridium difficile
Laure Saujet, Fátima C Pereira, Monica Serrano, Olga Soutourina, Marc Monot, Pavel V Shelyakin, Mikhail S Gelfand, Bruno Dupuy, Adriano O Henriques, Isabelle Martin-Verstraete

To cite this version:
Laure Saujet, Fátima C Pereira, Monica Serrano, Olga Soutourina, Marc Monot, et al.. Genome-Wide Analysis of Cell Type-Specific Gene Transcription during Spore Formation in Clostridium difficile. PLoS Genetics, Public Library of Science, 2013, 9 (10), pp.e1003756. 10.1371/journal.pgen.1003756 . pasteur-01370780

HAL Id: pasteur-01370780
https://hal-pasteur.archives-ouvertes.fr/pasteur-01370780
Submitted on 23 Sep 2016
Genome-Wide Analysis of Cell Type-Specific Gene Transcription during Spore Formation in *Clostridium difficile*

Laure Saujet¹,², Fátima C. Pereira³,⁴, Monica Serrano³,⁴, Olga Soutourina¹,², Marc Monot¹, Pavel V. Shelyaakin⁴, Mikhail S. Geldan⁴,⁵, Bruno Dupuy¹, Adriano O. Henriques³, Isabelle Martin-Verstraete¹,²,∗

¹Laboratoire Pathogénèse des Bactéries Anaérobies, Institut Pasteur, Paris, France, ²Université Paris Diderot, Sorbonne Paris Cité, Cellule Pasteur, Paris, France, ³Microbial Development Laboratory, Instituto de Tecnología Química e Biológica, Universidade Nova de Lisboa, Oeiras, Portugal, ⁴Institute for Information Transmission Problems, RAS, Bolshoi Karetny per, 19, Moscow, Russia, ⁵M.V. Lomonosov Moscow State University, Faculty of Biengineering and Bioinformatics, Vorobyevy Gory 1-73, Moscow, Russia

Abstract

*Clostridium difficile*, a Gram positive, anaerobic, spore-forming bacterium is an emergent pathogen and the most common cause of nosocomial diarrhea. Although transmission of *C. difficile* is mediated by contamination of the gut by spores, the regulatory cascade controlling spore formation remains poorly characterized. During *Bacillus subtilis* sporulation, a cascade of four sigma factors, σ³ and σ⁵ in the forespore and σ⁵ and σ⁸ in the mother cell governs compartment-specific gene expression. In this work, we combined genome wide transcriptional analyses and promoter mapping to define the *C. difficile* σ³, σ⁵, σ⁶ and σ⁸ regulons. We identified about 225 genes under the control of these sigma factors: 25 in the σ³ regulon, 97 σ⁵-dependent genes, 50 σ⁵-governed genes and 56 genes under σ⁸ control. A significant fraction of genes in each regulon is of unknown function but new candidates for spore coat proteins could be proposed as being synthesized under σ⁵ or σ⁸ control and detected in a previously published spore proteome. SpoIII D of *C. difficile* also plays a pivotal role in the mother cell line of expression repressing the transcription of many members of the σ⁵ regulon and activating sigK expression. Global analysis of developmental gene expression under the control of these sigma factors revealed deviations from the *B. subtilis* model regarding the communication between mother cell and forespore in *C. difficile*. We showed that the expression of the σ³ regulon in the mother cell was not strictly under the control of σ³ despite the fact that the forespore product SpoIII D was required for the processing of pro-σ⁶. In addition, the σ⁸ regulon was not controlled by σ⁸ in *C. difficile* in agreement with the lack of pro-σ⁸ processing. This work is one key step to obtain new insights about the diversity and evolution of the sporulation process among Firmicutes.

Introduction

*Clostridium difficile*, a Gram positive, anaerobic, spore-forming bacterium is a major cause of nosocomial infections associated with antibiotic therapy and is a major burden to health care services. This enteropathogen can lead to antibiotic-associated diarrhea and pseudo-membranous colitis, a potentially lethal disease. Two large toxins, the enterotoxin TcdA and the cytotoxin TcdB are the main virulence factors required for the development of symptoms of *C. difficile* infection (CDI). *C. difficile* produces highly resistant spores that facilitate the persistence of this bacterium in the environment in particular in aerobic conditions and contaminate hospital environments contributing to the establishment of a reservoir. Transmission of *C. difficile* is further mediated by contamination of the gut by spores as demonstrated recently using a murine model for CDI [1,2]. The disruption of the colonic microflora by antimicrobial therapy precipitates CDI and colonization of the intestinal tract. An early event towards this colonization is the germination process that converts dormant spores into vegetative cells that multiply leading to the production of toxins [3]. Glycine and bile salts like sodium cholate or sodium taurocholate are co-germinants required to induce *C. difficile* spores germination [4]. However, the molecular mechanisms involved in sporulation and germination are still poorly studied in *C. difficile* and our current knowledge on these processes is based mainly on data derived from the studies on Bacillus subtilis.
Author Summary

Clostridium difficile, a major cause of antibiotic-associated diarrhea, produces resistant spores that facilitate the persistence of this bacterium in the environment including hospitals. Its transmission is mediated by contamination of gut by spores. Understanding how this complex developmental process is regulated is fundamental to decipher C. difficile transmission and pathogenesis. The regulatory cascade controlling sporulation that involves four sigma factors, $\sigma^A$, $\sigma^B$, $\sigma^G$, and $\sigma^K$ in the forespore and $\sigma^D$ and $\sigma^E$ in the mother cell remains poorly characterized in C. difficile. By combining transcriptome analysis and promoter mapping, we identified genes expressed under the specific control of each sigma factor. Among sporulation-controlled proteins detected in spore, we can propose candidates for new spore coat proteins important for spore resistance. We also showed differences in the intercompartment communication between forespore and mother cell in C. difficile compared to the Bacillus subtilis model. In C. difficile, we observed that the activation of the $\sigma^D$ regulon was partially independent of $\sigma^D$ and that the $\sigma^K$ regulon was not controlled by $\sigma^K$. Our finding suggests that the C. difficile sporulation process might be more ancestral compared to that of B. subtilis.

At the onset of sporulation in B. subtilis, sporulating cells undergo an asymmetric division which partitions the sporangial cell into a larger mother cell and a smaller forespore (the future spore). The forespore is next wholly engulfed by the mother cell and later the dormant spore is released from the mother cell by lysis [5]. The developmental program of sporulation is mainly governed by the sequential activation of four sigma factors: $\sigma^D$, $\sigma^E$, $\sigma^G$, and $\sigma^K$. Their activity is confined to the forespore for $\sigma^D$ and $\sigma^E$ and to the mother cell for $\sigma^G$ and $\sigma^K$. Compartmentalization of gene expression is coupled to morphogenesis with $\sigma^D$ and $\sigma^K$ becoming active after asymmetric division and $\sigma^E$ and $\sigma^G$ after completion of engulfment of the forespore by the mother cell [5,6,7].

In B. subtilis, coordinated changes in gene expression underlie morphological differentiation in both the predivisional sporangium and later in the two compartments with the existence of communication between the mother cell and the forespore. The response regulator, Spo0A, and a phosphorelay involving five kinases, intermediary phosphorylated proteins and phosphatases control sporulation initiation. The alternative sigma factor, $\sigma^K$, which transcribes spo0A and sigF also controls early sporulation steps. When Spo0A-P level reaches a critical threshold, Spo0A-P activates sporulation genes including spoIE, as well as both the spoIIE-spoIIAB-sigF and the spoIECA-sigE operons encoding $\sigma^D$ and $\sigma^E$, respectively [8]. After its synthesis, $\sigma^D$ is held inactive by the anti-sigma factor SpoIIAB until the phosphatase SpoIIE dephosphorylates the anti-anti sigma factor SpoIIEA leading to the release of an active $\sigma^D$ from SpoIIAB after completion of asymmetric cell division [5]. $\sigma^E$ then transcribes about 50 genes in the forespore [9,10] including spoIIR encoding a secretory protein required for the processing of pro-$\sigma^E$ into active $\sigma^E$ in the mother cell [11]. $\sigma^D$ regulates in turn the expression of mother cell specific genes [9,12,13] and activates sigK expression with the combined activity of the SpoIIDD regulator [14]. In the forespore, $\sigma^E$ also controls sigG transcription. However, $\sigma^D$ becomes active coincidently with the completion of forespore engulfment by the mother cell. Following engulfment completion, the $\sigma^E$-controlled SpoIII A proteins, together with the forespore-specific SpoIQ protein, are required to maintain the potential for macromolecular synthesis in the forespore [6]. The RNA polymerase $\sigma^G$ holoenzyme transcribed 95 genes in the forespore including proteins like SpoIVB or CtpB involved in the processing and activation of $\sigma^K$ the last factor sigma of sporulation [5,6].

The four sporulation specific sigma factors are conserved in Clostridia and are present in C. difficile [15,16,17,18,19,20]. This is also the case for key genes involved in the spore morphogenesis [16,17,20]. In Clostridium acetobutylicum and Clostridium perfringens, the role of sporulation sigma factors in spore morphogenesis has been analyzed [21,22,23,24]. No resistant spores are formed by the sigE and sigG mutants of both Clostridia and by the sigE mutant of C. acetobutylicum whereas the sigE and sigK mutants of C. perfringens are severely defective in their ability to sporulate. Interestingly, sigE and sigF mutants of C. acetobutylicum fail to form the asymmetric division septum [22,24] while a sigK mutant is blocked earlier in sporulation than a sigE mutant in C. perfringens [21]. A detailed study of the morphological changes during the C. difficile spore differentiation process has recently been performed [19]. The C. difficile sigF and sigE mutants are arrested at the asymmetric stage while the sigG mutant is blocked after the completion of engulfment but unlike in B. subtilis, shows deposition of electro-dense coat material around the forespore. These mutants are unable to sporulate [19]. A sigK mutant of C. difficile forms four orders of magnitude fewer heat resistant spores than the isogenic wild-type strain. While showing no signs of coat deposition around the developing spore, this mutant shows accumulation of at least some cortex material unlike the case for B. subtilis [19]. The characterization of the sig sporulation mutants in several Clostridia suggests deviations from the B. subtilis paradigm regarding the function of the sporulation sigma factors [19,20,21,22,24,25].

The regulatory cascade controlling spore formation is also still poorly characterized in Clostridium compared to B. subtilis. The global analysis of the C. acetobutylicum transcriptional program during sporulation has given the dynamic orchestration of expression of the sporulation sigma factors and of key sporulation genes in this Clostridia [26] but the definition of the four-sigma factors regulons remains to be determined. In C. difficile, recent data have been obtained on sporulation initiation. Both Spo0A and $\sigma^D$ are present and required for efficient C. difficile sporulation [27,28]. By contrast, the phosphorelay is absent and the sporulation initiation remains in C. difficile and other Clostridia as a two-component system with Spo0A and associated kinases [18,28,29]. Both $\sigma^K$ and Spo0A directly control the expression of the spoIIEA-spoIIEB-sigF operon while Spo0A binds in vitro to the promoter region of spoIIEA, spoIEE and spoIIEG (in operon with sigE) [27,30]. By contrast, little is known about the molecular mechanisms controlling later steps in the sporulation regulatory cascade in C. difficile. In this work, we combined transcriptome, genome-wide transcriptional start site (TSS) mapping and in silico identification of promoters to define the $\sigma^A$, $\sigma^D$, $\sigma^E$, $\sigma^G$ and $\sigma^K$ regulons. This helped us to propose candidates for new spore coat proteins. Finally, we also identified interesting differences in the regulatory cascade of C. difficile compared to the B. subtilis model. We showed that activation of the $\sigma^D$ regulon was partially independent of $\sigma^D$ and that the $\sigma^K$ regulon was not controlled by $\sigma^K$.

Results and Discussion

I. Analysis of the sporulation pathway regulatory network by global approaches

Overview of the transcriptome data. The C. difficile mutants inactivated for $\sigma^A$, $\sigma^D$, $\sigma^E$ or $\sigma^K$ have been recently
constructed [19]. The sigF and sigE mutants are blocked at stage II of sporulation. The sigG mutant is blocked after the completion of engulfment and the sigK mutant form spores lacking coat material [19]. Thus, the sequential appearance of events seems to be similar to that observed in *B. subtilis*: σ^F^ and σ^K^ control early stages of development and are likely replaced by σ^K^ and σ^K^ later during sporulation. These studies have also shown the confinement of the activities of σ^K^ and σ^K^ to the forespore, and those of σ^K^ and σ^K^ to the mother cell [19]. To investigate more in detail the sporulation regulatory cascade in *C. difficile* and to define the regulons controlled by the four cell type-specific RNA polymerase sigma factors, we compared the expression profiles of the wild-type 630Aerm strain and of the sigF, sigE, sigG or sigK mutant. We performed preliminary tests to determine conditions allowing detection of differential expression between the 630Aerm strain and each mutant inactivated for a sigma factor as described in Materials and Methods. The cells were harvested 14 h after inoculation for the wild-type strain 630Aerm and the sigF or the sigE mutant, after 19 h of growth for strain 630Aerm and the sigG mutant or after 24 h of growth for strain 630Aerm and the sigK mutant. We found 111 genes and 141 genes differentially expressed with a p value <0.05 between the 630Aerm strain and the sigF mutant or the sigE mutant, respectively (Table S1 and S2). While 92 and 19 genes were down and upregulated in the sigF mutant compared to the 630Aerm strain, all the 141 genes were downregulated in the sigE mutant. In addition, 51 and 66 genes were differentially expressed with a p value <0.05 between the 630Aerm strain and the sigG or the sigK mutant in transcriptome, respectively (Table S3 and S4). All of the presumptive sigG-controlled genes were downregulated in the mutant compared to the 630Aerm strain. Finally, 58 and 8 genes were down and upregulated in the sigK mutant relative to the 630Aerm strain. To confirm the results obtained in the microarrays experiments, we performed qRT-PCR on a subset of 10 to 15 genes representative of various cell functions and regulated by each of the cell type-specific sigma factors. The qRT-PCR results confirmed the microarrays data for all of the tested genes (Table S5).

Mapping of transcriptional start sites of genes controlled by sporulation sigma factors. We recently performed a genome wide determination of transcriptional start sites (TSS) for the 630Aerm strain using RNA-seq [31]. We searched in our data for TSS of all the genes found to be controlled by σ^K^, σ^K^ or σ^K^ in the transcriptome analysis. By this global approach, we mapped about 111 TSS upstream of genes positively controlled by at least one of the four sporulation-specific sigma factors (Table S6 and S7). We analyzed these data by an iterative *in silico* strategy as described in Materials and Methods. We further manually analyzed all promoters not found by this *in silico* procedure (<30%). These strategies allowed us to identify promoters for each sigma factor (Table S6, S7 and S8). We identified about 40 promoters likely to be utilized by RNA polymerase associated to one of the forespore sigma factors (σ^K^ or σ^K^) (Table S6). The consensus elements recognized by σ^K^ in *C. difficile* were determined by using 10 mapped promoters of genes specifically controlled by σ^K^ in transcriptome (downregulated in a sigK mutant but not in the sigE, sigG or sigK mutant) (Table S6). The conserved motifs for σ^K^-dependent promoters were defined from a dataset of 30 promoters located upstream of genes specifically controlled by σ^K^ in transcriptome (downregulated in a sigK mutant but not in a sigG mutant) (Table S6). The *in silico* analysis identified 9 of 10 σ^K^-promoters and 23 of 30 σ^K^-promoters. The consensus sequences for *C. difficile* σ^K^ and σ^K^-controlled promoters (Fig. 1A) are very similar to those of *B. subtilis* [10,16] and to each other in their −10 and −35 elements as noted for *B. subtilis* [10]. However, they differ in that a G is conserved on the upstream side of the −10 promoter element in σ^K^-dependent promoters (position 25 in Fig. 1A) whereas for σ^K^-dependent promoters a A is conserved on the downstream side of the −10 element (position 32 in Fig. 1A).

Upstream of genes specifically controlled by σ^K^ (downregulated in a sigE mutant but neither in a sigK mutant nor in a sigG mutant), we identified 47 promoters, 25 of which were in the *in silico* analysis (Table S7 and S8). Moreover, we mapped 24 TSS corresponding to promoters recognized by σ^K^ upstream of σ^K^-controlled genes in transcriptome, 19 of which were found in *silico*. The *C. difficile* consensus elements specific for σ^K^ and σ^K^-controlled promoters are very similar to those determined in *B. subtilis* [12,14,16]. The motif recognized by σ^K^ and σ^K^-shared very similar −10 sequences while the −35 elements differed with an ATA motif for σ^K^ and AC for σ^K^- (Fig. 1B). In *B. subtilis*, the specificity of interaction of these sigma factors with the −35 region sequences is associated with the presence of a glutamine at position 217 of σ^K^ and an arginine in σ^K^-Indeed, the replacement of glutamine 217 in σ^K^-by an arginine allows σ^K^-to recognize σ^K^-controlled genes [32]. Interestingly, σ^K^ and σ^K^-of *C. difficile* contain a glutamine and an arginine at position 218, respectively (Fig. S1). This observation supports our inference that a conserved T or C is an important feature of the −35 element herein identified for *C. difficile* σ^K^-and σ^K^-promoters. To date, less than 10 promoters of sporulation-regulated genes have been characterized in Clostridia, mainly in *C. acetobutylicum* (*spoIIGA, sigG, sigK*) and *C. perfringens* (*spoIIGA, epe, sigK*) [33,34]. This work greatly increased the number of characterized promoters for clostridial sporulation genes. The probability to observe orthologs with high-scoring candidate promoters in many clostridial genomes (Table S8) is highest for genes regulated both in *C. difficile* and *B. subtilis* [9,10,12,14], as these genes likely form the regulon cores.

II. An overview of the four compartment-specific σ^F^, σ^E^, σ^K^-and σ^K^-regulons

We combined transcriptome data and promoter identifications to define the σ^K^-, σ^K^-, σ^K^- and σ^K^-regulons and to determine the pool of genes probably under the direct control of each sigma factor among genes positively controlled by these four sigma factors (Table 1 and 2).

The σ^K^-regulon. Among the 111 genes controlled by σ^K^-in transcriptome, 25 were downregulated in the sigF mutant but not in the sigE, sigG or sigK mutant (Table 1). We identified 10 σ^K^-dependent promoters controlling the expression of 14 genes. The σ^K^-regulon is the smallest of the four-compartment-specific regulons as observed in *B. subtilis* [9,10]. The orthologs of *B. subtilis* genes involved in sporulation regulatory functions, in spore morphogenesis or in germination [5,16,17,20] were downregulated in the sigF mutant compared to strain 630Aerm. σ^K^-positively controlled the expression of *spoIIR*, encoding a signaling protein that in *B. subtilis* triggers the activation of the membrane bound SpoIIGA protease that in turn cleaves the pro-σ^K^-protein and leads to the production of an active σ^K^- factor [5]. The *spoIQ* and *spoIP* genes, the expression of which was also downregulated in the sigF mutant, are involved in the engulfment process. SpoIIP is a component of a molecular machine involved in dissolution of the septal cell wall during engulfment in *B. subtilis* [35,36]. The *B. subtilis* SpoIIP protein interacts with a mother cell-specific protein across the intermediate space, an interaction that helps driving engulfment and that is also central to the formation of a channel linking the mother cell to the forespore [6,37,38]. In addition, σ^K^-was required to transcribe the *gpr* gene encoding a protease that is important for the degradation of small acid-soluble spore proteins.
(SASPs) during germination. The orthologs of these genes are members of the $\sigma^F$ regulon in *B. subtilis* [9,10,16]. On the contrary, *gpr*, *spoIIR* and *spoIIP* are not controlled by $\sigma^F$ in *C. acetobutylicum* [22]. $\sigma^F$ was also required for the expression of genes involved in cell wall metabolism: *slb*, *CD2141* and *CD1229* encoding a spore cortex hydrolising protein, a DAla-DAla carboxypeptidase and a peptidoglycan (PG) glycosyltransferase, respectively. Moreover, $\sigma^F$ positively controlled the expression of 5 genes encoding proteins previously detected in the spore proteome [39]: a glyceraldehyde-3P-dehydrogenase (*CD0580*), three peptidases (*CD2660*, *CD2661* *CD3595*), and an RNA helicase (*CD0761*) (Table 1). Finally, 3 genes encoding proteins of unknown function and 2 genes encoding putative membrane proteins (*CD2686* and *CD2856*) were downregulated in the *sigF* mutant.

**The $\sigma^E$ regulon.** We found 97 genes to be downregulated in a *sigE* mutant but not in a *sigG* mutant or in a *sigK* mutant (Table 2). TSS mapping allowed us to identify 47 $\sigma^E$-dependent promoters controlling the expression of 63 genes (Table S7). $\sigma^F$ controls the largest of the four cell type-specific sporulation regulons as also found for *B. subtilis* [9,12]. Several $\sigma^E$ target genes in *C. difficile* are orthologs of genes known to control engulfment, cortex formation, initiation of spore coat assembly, preparation of the late phase of sporulation or germination in *B. subtilis* (Table 2). This is consistent with the morphological block imposed by a *sigE* insertional mutation, which arrests development at the asymmetric division stage [19]. *spoIID*, which encodes a PG hydrolase required for the engulfment of the forespore by the mother cell was controlled by $\sigma^E$. In *B. subtilis*, SpoIID is associated to SpoIIP and SpoIIM to form the DMP machine necessary for engulfment [35,36]. The *spoIIP* gene was only controlled by $\sigma^F$ in the transcriptome analysis. However, in qRT-PCR experiments, we showed that *spoIIP* expression decreased 10-fold in a *sigE* mutant and 800-fold in a *sigF* mutant. This suggested that this gene was under the control of both sigma factors as observed in *B. subtilis* [10,12]. Also, even though a membrane protein (*CD1221*) sharing weak similarity to the N-terminal part of *B. subtilis* SpoIIM is present in the *C. difficile* genome, its gene was not differentially expressed.

Figure 1. Consensus sequences of $\sigma^F$, $\sigma^E$, $\sigma^G$ or $\sigma^K$-dependent promoters in *C. difficile*. Consensus promoter sequences for $\sigma^E$- and $\sigma^G$-controlled promoters (A) or $\sigma^F$- and $\sigma^K$-controlled promoters (B) in *C. difficile*. The sequence logo was created on the WebLogo website (http://weblogo.berkeley.edu) using promoters mapped in this study and listed in Tables S6 and Table S7: compilation of 10 common motifs upstream of $\sigma^E$-regulated genes, of 30 common elements upstream of $\sigma^F$-regulated genes, of 47 conserved sequences upstream of $\sigma^K$-regulated genes or 24 conserved elements $\sigma^G$-regulated genes. The height of the letters is proportional to their frequency.

doi:10.1371/journal.pgen.1003756.g001
**Table 1.** The forespore line of expression with the identification of the $\sigma^F$ and $\sigma^G$ regulon.

| Gene      | Name  | Function                                      | Expression ratio in transcriptome $^a$ | Promoter $^b$ | Detection in spore $^c$ | operon       |
|-----------|-------|-----------------------------------------------|----------------------------------------|--------------|-------------------------|-------------|
| **Members of the $\sigma^F$ regulon** | sigF/630 erm | | | | | |
| **sporulation** | | | | | | |
| CD0125    | spoIIQ | Stage II sporulation protein Q             | 0.11                                   | $\sigma^F$   | –                       | CD2470-C2468 |
| CD2470    | gpr    | Spore endopeptidase                         | 0.22                                   | $\sigma^F$   | +                       |             |
| CD2469    | spoIIP | Stage II sporulation protein P              | 0.14                                   | –            | –                       |             |
| CD2468    |        | Conserved hypothetical protein              | 0.28                                   | –            | –                       |             |
| CD3564    | spoIIR | Pro-SigE endopeptidase signalling protein   | 0.43                                   | $\sigma^F$   | –                       | CD3564-3563 |
| CD3563    | sleB   | Spore-cortex-lytic protein                  | 0.16                                   | –            | –                       |             |
| CD0783    | spoIVB' | SpoIVB protein, S55 peptidase family        | 0.10                                   | –            | –                       |             |
| **envelopes** | | | | | | |
| CD2141    |        | Serine-type D-Ala-D-Ala carboxypeptidase     | 0.30                                   | –            | –                       |             |
| CD1229    |        | Putative peptidoglycan glycosyltransferase  | 0.42                                   | $\sigma^F$   | –                       |             |
| CD2686    |        | Putative membrane protein                   | 0.33                                   | $\sigma^F$   | –                       |             |
| CD2856    |        | Putative membrane protein                   | 0.46                                   | $\sigma^F$   | –                       |             |
| CD2107    |        | Xanthine/uracil/thiamine permease family    | 0.37                                   | –            | –                       |             |
| CD2102    |        | Putative Na(+)H(+) antiporter               | 0.49                                   | $\sigma^A$   | –                       |             |
| **miscellaneous** | | | | | | |
| CD0580    | gapN   | Glyceraldehyde-3-phosphate dehydrogenase    | 0.36                                   | $\sigma^F$   | +                       |             |
| CD3595    |        | Aminopeptidase                              | 0.55                                   | $\sigma^F$   | +                       |             |
| CD2661    |        | Putative peptidase, M16 family              | 0.47                                   | +            | CD2661-2660             |             |
| CD2660    |        | Putative peptidase, M16 family              | 0.48                                   | +            | CD2661-2660             |             |
| CD0047    | ispD   | 2-C-methyl-D-erythritol 4-P cytidylyltransferase | 0.48 | $\sigma^A$ | –                       | CD0047-CD0048 |
| CD0048    | ispF   | 2-C-methyl-D-erythritol 2,4-cyclophosphate synthase | 0.50 | – | – | |
| CD3220    |        | Putative methyltransferase                  | 0.39                                   | $\sigma^A$   | –                       |             |
| CD0761    |        | Putative ATP-dependent RNA helicase         | 0.46                                   | $\sigma^F$ and $\sigma^A$ | + |             |
| CD1323    | tepA   | Protein export-enhancing factor             | 0.38                                   | –            | +                       |             |
| **proteins of unknown function** | | | | | | |
| CD0347    |        | Conserved hypothetical protein              | 0.35                                   | $\sigma^F$   | –                       | CD0347-0348 |
| CD0348    |        | Fragment of conserved hypothetical protein  | 0.35                                   | –            | –                       |             |
| CD2687    |        | Conserved hypothetical protein              | 0.36                                   | –            | –                       |             |
| **Members of the $\sigma^G$ regulon** | sigG/630 erm | | | | | |
| **sporulation** | | | | | | |
| CD0773    | spoIVC | DPA uptake protein, SpoAC                   | 0.24                                   | $\sigma^G$   | –                       | CD0773-CD775 |

$^a$ Expression ratio in transcriptome

$^b$ Promoter

$^c$ Detection in spore
| Gene | Function | Expression ratio in transcriptome | Detection in spore | Members of the σ^G regulon | Promoter? |
|------|----------|----------------------------------|-------------------|---------------------------|----------|
| snaLR (CD0047) | DPA uptake protein, SpoP | 2.17 | + | + | + |
| spvA (CD1757) | DPA uptake protein, SpoQ | 0.32 | + | + | + |
| spvE (CD1758) | DPA uptake protein, SpoP | 0.02 | + | + | + |
| spoA (CD1290) | Small acid-soluble spore protein alpha | 0.01 | + | + | + |
| spoB (CD1291) | Small acid-soluble spore protein | 0.07 | + | + | + |
| spoVT (CD1294) | Transcriptional regulator, SpoVT | 0.07 | + | + | + |
| ssaA (CD2670) | Small acid-soluble spore protein alpha | 0.00 | + | + | + |
| ssaB (CD2671) | Small acid-soluble spore protein | 0.00 | + | + | + |
| putative stage IVB sporulation protein B, peptidase | 0.60 | + | + | + | + |
| pdaA (CD1430) | Putative h-lactam-biosynthetic deacteylase | 0.06 | + | + | + |
| dacF (CD1431) | D-alanyl-D-alanine carboxypeptidase | 0.11 | + | + | + |
| putative N-acetylmuramoyl-L-alanine amidase | 0.25 | + | + | + | + |
| uppS (CD2110) | Undecaprenyl pyrophosphate synthetase | 0.34 | + | + | + |
| putative membrane protein, DUF81 family | 0.10 | + | + | + | + |
| CD2637-CD2635 | Putative membrane protein | 0.34 | + | + | + |
| CD1799-CD1798 | Putative membrane protein, DUF421 family | 0.10 | + | + | + |
| CD2465 | Putative amino acid/polyamine transporter | 0.43 | + | + | + |
| cotG (CD2466) | Putative manganese catalase | 0.25 | + | + | + |
| rbr (CD2467) | Ruberythrin | 0.05 | + | + | + |
| sodA (CD2468) | Superoxide dismutase (Mn) | 0.36 | + | + | + |
| putative ribosome recycling factor | 0.05 | + | + | + | + |
| ATP-dependent 30S-like NTPase | 0.05 | + | + | + | + |
| putative NADH-azoreductase | 0.05 | + | + | + | + |
| putative C4-dicarboxylate anaerobic carrier | 0.05 | + | + | + | + |
### Table 1. Cont.

| Gene Name | Function | Expression ratio in transcriptome<sup>a</sup> | Promoter<sup>b</sup> | Detection in spore<sup>c</sup> | operon |
|-----------|----------|---------------------------------------------|----------------------|-------------------------------|--------|
| **Members of the σ<sup>G</sup> regulon** |
| CD3489    | Putative oligoendopeptidase F           | sigF<sup>630</sup>Δerm 0.62 | σ<sup>G</sup> | –                             |        |
| CD1676    | *pcp* Pyrrolidone-carboxylate peptidase | sigG<sup>630</sup>Δerm 0.66 | +                   | CD1677-CD1676                  |        |
| CD1677    | Putative membrane protein               | 0.48                     |                     | –                             |        |
| CD2841    | Putative amidohydrolase                 | 0.57                     | σ<sup>G</sup>       | +                             |        |
| **proteins of unknown function** |
| CD0543    | Conserved hypothetical protein          | 0.38                     | σ<sup>G</sup>       | –                             |        |
| CD1297    | Conserved hypothetical protein          | 0.31                     |                     | –                             |        |
| CD1298    | Conserved hypothetical protein          | 0.19                     |                     | –                             |        |
| CD1301.1  | Conserved hypothetical protein          | 0.28                     |                     | –                             |        |
| CD1354    | Conserved hypothetical protein          | 0.22                     | σ<sup>G</sup>       | –                             |        |
| CD1463    | Conserved hypothetical protein          | 0.06                     | σ<sup>G</sup>       | +                             |        |
| CD1880    | Conserved hypothetical protein          | 0.07                     | σ<sup>G</sup>       | +                             |        |
| CD2112    | Conserved hypothetical protein          | 0.03                     | σ<sup>G</sup>       | +                             |        |
| CD2245.1  | Conserved hypothetical protein          | 0.10                     | σ<sup>G</sup>       | +                             |        |
| CD2375    | Conserved hypothetical protein          | 0.14                     | σ<sup>G</sup>       | +                             |        |
| CD2808    | Conserved hypothetical protein          | 0.05                     | σ<sup>G</sup>       | –                             |        |
| CD2809    | Conserved hypothetical protein          | 0.05                     | σ<sup>G</sup>       | +                             |        |
| CD1568    | Conserved hypothetical protein          | 0.60                     | σ<sup>G</sup>       | –                             |        |
| CD3610    | Conserved hypothetical protein          | 0.66                     | σ<sup>G</sup>       | +                             |        |

<sup>a</sup>A gene is considered differentially expressed when the P value is < 0.05 using the statistical analysis described in Materials and Methods.

<sup>b</sup>We search for promoters recognized by the different sigma factors active in forespore upstream of the TSS mapped by RNA-seq. These promoters are listed in Table S6. 3 promoters with the –10 and/or –35 elements less conserved are indicated with a question mark.

<sup>c</sup>Proteins that are detected associated with the spore by a proteomic approach [39].

doi:10.1371/journal.pgen.1003756.t001
Table 2. The mother cell line expression with the identification of the $\sigma^E$ and $\sigma^K$ regulons and the regulation by SpoIIID of $\sigma^E$-controlled genes.

| Gene        | Function                              | Expression ratio in transcriptome | Promoter | Detected in spore | operon     |
|-------------|---------------------------------------|-----------------------------------|-----------|-------------------|------------|
|             |                                       | $\text{sigE/630} \cdot \text{erm}$ | $\text{spoIIID/630} \cdot \text{erm}$ |          |                   |            |
| **sporulation** |                                      |                                   |           |                   |            |
| CD0124      | spoIIID peptidoglycane hydrolase SpoIID | 0.24                              | $\sigma^E$ | –                 |            |
| CD1192      | spoIIIAA ATP-binding stage III sporulation protein | 0.05                              | 3.85$^d$ | $\sigma^E$ | – $\text{spoIIIAA-spoIIAH}$ |
| CD1193      | spoIIIBA Stage III sporulation protein AB | 0.05                              | 3.8       | –                 |            |
| CD1194      | spoIIIBAC Stage III sporulation protein AC | 0.03                              | 4.9       | –                 |            |
| CD1195      | spoIIIBAD Stage III sporulation protein AD | 0.07                              | 3.4       | –                 |            |
| CD1196      | spoIIIBAE Stage III sporulation protein AE | 0.09                              | 2.4       | –                 |            |
| CD1197      | spoIIIBAF Stage III sporulation protein AF | 0.08                              | 3.6       | –                 |            |
| CD1198      | spoIIIBAG Stage III sporulation protein AG | 0.02                              | 7.75$^d$ | $\sigma^E$ | –            |
| CD1199      | spoIIIBAH Stage III sporulation ratchet engulfment protein | 0.05                              | 3.3       | –                 |            |
| CD0126      | spoIIID Transcriptional regulator SpoIIID | 0.04                              | $\sigma^E$ | –                 |            |
| CD2629      | spoIIB Assembly of cortex and coat layers | 0.02                              | 3.5       | $\sigma^E$ | +           |
| CD3567      | sipL Functional homolog to spoIVD       | 0.05                              | 3         | $\sigma^E$ | +           |
| CD2443      | Conserved hypothetical protein YqfC-like | 0.05                              | $\sigma^E$ | +                 | CD2443-CD2442 |
| CD2442      | spoIV Stage IV sporulation protein YqfD-like | 0.14                              | –         |                   |            |
| CD1230      | sigK Fragment of sigma-K factor         | 0.13                              | $\sigma^E$ and $\sigma^K$ | –             |            |
| CD1234      | putative phage protein, skin element  | 0.22                              | $\sigma^E$ | –                 |            |
| CD1016      | cwID N-acetyl-L-alanine amidase         | 0.14                              | $\sigma^E$ | –                 |            |
| CD2246      | cspBA serine germination protease      | 0.11                              | $\sigma^E$ | +                 | cspAB-cspC |
| CD2246      | cspC serine germination protease      | 0.16                              | +         |                   |            |
| CD3542      | spmA Spore maturation protein A        | 0.16                              | $\sigma^E$ | –                 | spmA-spmB  |
| CD3541      | spmB Spore maturation protein B        | 0.11                              | 2         | –                 |            |
| CD1511      | cotB Spore outer coat layer protein CotB | 0.05                              | 3.5       | $\sigma^E$ | +           |
| CD0213      | Putative spore coat protein            | 0.42                              | –         |                   | CD0214-CD0213 |
| CD0214      | Conserved hypothetical protein         | 0.23                              | +         |                   |            |
| CD1320      | Putative peptidase, M16 family        | 0.25                              | –         |                   | CD1320-CD1322 |
| CD1321      | Putative sporulation protein           | 0.23                              | –         |                   |            |
| CD1322      | daspG Aspartokinase I                  | 0.32                              | –         |                   |            |
| CD2641      | Putative sporulation protein           | 0.13                              | $\sigma^E$ | –                 | CD2641-CD2643 |
| CD2640      | nrdR Transcriptional regulator, NrdR family | 0.23                              | –         |                   |            |
| CD2639      | Putative cytotoxic factor              | 0.21                              | –         |                   |            |
| CD0782      | Putative sporulation protein YunB      | 0.32                              | –         |                   |            |
| CD1168      | YlbJ-like protein, spore cortex formation | 0.18                              | –         |                   |            |
| CD3455      | C-terminal protease, homolog of CtpB   | 0.25                              | $\sigma^E$ | –                 |            |
| Members of the \( \sigma^E \) regulon | Function | Expression ratio in transcriptome\(^a\) | Promoter\(^b\) | Detected in spore\(^c\) | operon |
|--------------------------------------|----------|---------------------------------|----------------|----------------|-------|
| CD3494 | Putative spore protein | \( \text{sigE} \times \text{erm} \) | \( \sigma^E \) | – | CD3494-CD3493 |
| CD3493 | Putative membrane protein | \( \text{sigE} \times \text{erm} \) | \( \sigma^E \) | – | – |
| \text{envelopes} | | | | | |
| CD3464 | Conserved hypothetical protein | \( \text{sigE} \times \text{erm} \) | \( \sigma^E \) | – | CD3464-CD3463 |
| CD3463 | \( \text{dr}2 \) | Alanine racemase 2 | 0.05 | + | – |
| CD2761 | N-acetylmuramoyl-L-alanine amidase | \( \text{sigE} \times \text{erm} \) | 0.40 | – | – |
| CD2833 | Putative calcium-transporting ATPase | \( \text{sigE} \times \text{erm} \) | 0.22 | \( \sigma^E \) | – |
| CD0760 | Putative Ca\(2^+\)/Na\(+\) antiporter | \( \text{sigE} \times \text{erm} \) | 0.20 | \( \sigma^E \) | – |
| CD3483 | Putative zinc/iron permease | \( \text{sigE} \times \text{erm} \) | 0.34 | \( \sigma^E \) | – |
| CD2445 | Transmembrane signaling protein, TspO | \( \text{sigE} \times \text{erm} \) | 0.07 | \( \sigma^E \) | – |
| CD0131 | Putative membrane protein | \( \text{sigE} \times \text{erm} \) | 0.18 | – | – |
| CD0314 | Putative membrane protein | \( \text{sigE} \times \text{erm} \) | 0.15 | \( \sigma^E \) | – |
| CD1301 | Putative membrane protein | \( \text{sigE} \times \text{erm} \) | 0.33 | – | – |
| CD1416 | Putative membrane protein | \( \text{sigE} \times \text{erm} \) | 0.20 | \( \sigma^E \) | – |
| CD1928 | Putative membrane protein | \( \text{sigE} \times \text{erm} \) | 0.14 | \( \sigma^E \) | – |
| CD1929 | Putative membrane protein | \( \text{sigE} \times \text{erm} \) | 0.27 | – | CD1928-CD1929 |
| CD1940 | Putative membrane protein | \( \text{sigE} \times \text{erm} \) | 0.08 | \( \sigma^E \) | – |
| CD2800 | Putative membrane protein | \( \text{sigE} \times \text{erm} \) | 0.11 | \( \sigma^E \) | – |
| CD3636 | Putative membrane protein | \( \text{sigE} \times \text{erm} \) | 0.42 | – | CD3636-CD3635 |
| CD3635 | Conserved hypothetical protein | \( \text{sigE} \times \text{erm} \) | 0.29 | – | – |
| \text{metabolism} | | | | | |
| CD3638 | Conserved hypothetical protein | \( \text{sigE} \times \text{erm} \) | 0.21 | \( \sigma^E \) | – |
| CD3637 | NADPH-dependent FMN reductase | \( \text{sigE} \times \text{erm} \) | 0.27 | – | CD3638-CD3637 |
| CD2429.1 | 4Fe-4S ferredoxin | \( \text{sigE} \times \text{erm} \) | 0.27 | \( \sigma^E \) | – |
| CD2429 | Flavodoxin/ferredoxin oxidoreductase, alpha | \( \text{sigE} \times \text{erm} \) | 0.31 | – | CD2429.1-CD2428 |
| CD2428 | Flavodoxin/ferredoxin oxidoreductase, beta | \( \text{sigE} \times \text{erm} \) | 0.38 | – | – |
| CD3251 | Putative dehydrogenase | \( \text{sigE} \times \text{erm} \) | 0.25 | \( \sigma^E \) | – |
| CD3258 | Iron hydrogenase | \( \text{sigE} \times \text{erm} \) | 0.08 | \( \sigma^E \) | – |
| CD2000 | \( \text{isp} \) | Intracellular serine protease | \( \text{sigE} \times \text{erm} \) | 0.18 | \( \sigma^E \) | – |
| CD3652 | Putative peptidase, M1 family | \( \text{sigE} \times \text{erm} \) | 0.12 | + | – |
| CD3521 | Putative peptidase T, M20B family | \( \text{sigE} \times \text{erm} \) | 0.33 | \( \sigma^E \) | + |
| CD1085 | Putative membrane protein | \( \text{sigE} \times \text{erm} \) | 0.26 | \( \sigma^E \) | – |
| CD1086 | Putative peptidase, M20D family | \( \text{sigE} \times \text{erm} \) | 0.26 | – | – |
| Gene              | Function                                      | Expression ratio in transcriptome | Promoter | Detected in spore | operon       |
|-------------------|-----------------------------------------------|-----------------------------------|----------|-------------------|--------------|
|                   |                                               | sigE30 \erm                       |          |                   |              |
|                   |                                               | spoIIDI30 \erm                    |          |                   |              |
| **Members of the σ^E regulon** |                                               |                                   |          |                   |              |
| CD1555            | Putative amino acid permease                   | 0.22                              |          |                   |              |
| CD1746            | gIC                                           | 0.20                              | σ^E      |                   |              |
| CD1259            | bniQ-I                                        | 0.39                              |          |                   |              |
| CD2439            | Diacylglycerol kinase/undecaprenol kinase      | 0.45                              |          |                   |              |
| CD1068            | Polysaccharide biosynthesis protein            | 0.23                              |          |                   |              |
| CD1257            | Polysaccharide deacetylase                    | 0.11                              | 2.1      |                   |              |
| CD1248            | Polysaccharide deacetylase                    | 0.09                              | 2.3^d    | σ^E               |              |
| CD1319            | Polysaccharide deacetylase                    | 0.06                              | 3.2^d    | σ^E               | +            |
| CD10982           | ubIA                                          | 0.39                              |          |                   |              |
| **miscellaneous** |                                               |                                   |          |                   |              |
| CD1235            | ssb                                           | 0.03                              | 3.6^d    | σ^E               |              |
| CD1167            | recV                                          | 0.26                              | σ^E      |                   |              |
| CD2864            | Putative hydrolase                            | 0.08                              | 4        | σ^E               |              |
| CD1298            | Putative ATP/GTP-binding protein              | 0.19                              | σ^E      |                   |              |
| CD3462            | mazE                                          | 0.05                              |          |                   | CD3462-CD3461|
| CD2865            | Putative antitoxin EndoAI                     | 0.11                              |          |                   |              |
| CD0757            | mazF                                          | 0.25                              | 2.2^d    | σ^E               | +            |
| CD1611            | mazF                                          | 0.45                              |          |                   |              |
| CD2637            | mazF                                          | 0.37                              |          |                   |              |
| **proteins of unknown function** |                                               |                                   |          |                   |              |
| CD1019            | Conserved hypothetical protein, DUF1256       | 0.15                              | σ^E      |                   |              |
| CD2121            | Conserved hypothetical protein                 | 0.23                              |          |                   |              |
| CD2374            | Conserved hypothetical protein                 | 0.23                              |          |                   |              |
| CD1884            | Conserved hypothetical protein                 | 0.11                              | σ^E      |                   |              |
| CD1726            | Conserved hypothetical protein                 | 0.05                              | 3.9^d    | σ^E               |              |
| CD3457            | Conserved hypothetical protein                 | 0.05                              | 2.6      | σ^E               |              |
| CD3465            | Conserved hypothetical protein                 | 0.40                              |          |                   | CD3466-CD3465|
| CD1066            | Conserved hypothetical protein                 | 0.16                              |          |                   |              |
| CD3522            | Conserved hypothetical protein                 | 0.09                              | σ^E      |                   | +            |
| CD2816            | Conserved hypothetical protein                 | 0.11                              | σ^E      |                   |              |
### Table 2. Cont.

| Gene | Function | Expression ratio in transcriptome<sup>a</sup> | Promoter<sup>b</sup> | Detected in spore<sup>c</sup> | operon |
|------|----------|-----------------------------------------------|----------------------|-------------------------------|--------|
|      |          | sigE630Δerm | spoIID630Δerm |                  |        |
| **Members of the σ<sup>e</sup> regulon** | | | | | |
| CD3234 | Conserved hypothetical protein | 0.09 | | | |
| CD1930 | Conserved hypothetical protein | 0.27 | | + | |
| CD1063 | Conserved hypothetical protein | 0.11 | 3 | σ<sup>e</sup> | − |
| **Members of the σ<sup>e</sup> regulon** | | | | | |
|      | | sigE630Δerm | sigK630Δerm |                  |        |
| **sporulation** | | | | | |
| CD1613 | cotA | Spore outer coat layer protein CotA | 0.10 | 0.05 | σ<sup>e</sup> | + |
| CD0596 | Conserved hypothetical protein | 0.02 | | σ<sup>e</sup> ?<sup>a</sup> | + |
| CD0597 | cotF | Spore coat peptide assembly protein | | | | |
| CD0598 | cotCB | Spore-coat protein CotCB manganese catalase | 0.18 | 0.01 | + | |
| CD2399 | Conserved hypothetical protein | 0.14 | 0.01 | σ<sup>e</sup> ?<sup>a</sup> | + | CD2399-CD2401 |
| CD2400 | cotR2 | Spore peptide assembly protein CotR2 | 0.10 | 0.01 | + | |
| CD2401 | cotD | Spore coat protein CotD manganese catalase | 0.11 | 0.01 | + | |
| CD1433 | cotE | Spore coat protein CotE peroxiredoxin/chitinase | 0.14 | 0.03 | σ<sup>e</sup> | + | |
| CD0551 | sleC | Spore cortex-lytic enzyme pre-pro-form | 0.15 | 0.14 | σ<sup>e</sup> | + | |
| CD0332 | bclA1 | Putative exosporium glycoprotein | 0.21 | 0.08 | σ<sup>e</sup> ?<sup>a</sup> | + | |
| CD3230 | bclA2 | Putative exosporium glycoprotein | 0.12 | 0.04 | σ<sup>e</sup> ?<sup>a</sup> | − | |
| CD3349 | bclA3 | Exosporium glycoprotein BclA3 | 0.01 | 0.02 | σ<sup>e</sup> ?<sup>a</sup> | − | |
| CD2968 | dpaA | Dipicolinate synthase subunit A | 0.30 | 0.08 | σ<sup>e</sup> | − | CD2968-CD2967 |
| CD2967 | dpaB | Dipicolinate synthase subunit B | 0.29 | 0.09 | − | |
| CD3569 | Sporulation-specific protease, YabG-like | 0.22 | 0.08 | σ<sup>e</sup> | − | |
| **envelopes** | | | | | |
| CD0119 | glmM | Phosphoglucomamine mutase | 0.49 | | − | |
| CD0120 | glmS | Glucosamine-fructose-6-P aminotransferase | 0.52 | | σ<sup>e</sup>/σ<sup>e</sup> | − | |
| CD2664 | mreE | UDP-N-acetylMuramyl-tripeptide synthetase | 0.45 | 0.42 | σ<sup>e</sup>/σ<sup>e</sup> | − | |
| CD1307 | Putative LD-transpeptidases | 0.37 | 0.48 | + | |
| CD1884 | Putative N-acetylMuramoyl-L-alanine amidase | 0.21 | 0.13 | σ<sup>e</sup> | − | |
| CD1898 | Putative N-acetylMuramoyl-L-alanine amidase | 0.12 | 0.09 | σ<sup>e</sup> | + | CD1898-CD1897 |
| CD1897 | Conserved hypothetical protein | 0.13 | | | |
| CD0514 | cwpV | Cell surface protein | 0.63 | 0.20 | + | |
| CD1518 | feoA | Ferrous iron transport protein | 0.05 | | σ<sup>e</sup> | − | CD1518-CD1517 |
| CD1517 | feoB | Ferrous iron transport protein B | 0.05 | | − | |
| CD1904 | ABC-type transport system, permease | 0.26 | 0.13 | − | |
| CD1927 | yfC | ABC transporter component, ATP-binding | 0.3 | 0.08 | σ<sup>e</sup> | − |
### Table 2. Cont.

| Members of the σK regulon | sigE/630 Δerm/s/g0/630 Δerm |
|---------------------------|------------------------------|
| CD1891 Fragment of ABC-type transport system | 0.26 0.08 σK − |
| CD2720 Putative transporter | 0.16 |
| CD902 Putative cation efflux protein | 0.29 0.17 σK − |
| CD2144 Putative membrane protein | 0.46 0.12 σK − |
| CD2346 Putative membrane protein | 0.24 |
| CD2459 Putative glucokinase, ROK family | 0.41 σK − CD2459-CD2458 |
| CD2458 Transporter, Major Facilitator Superfamily | 0.39 |
| **miscellaneous** | |
| CD3032 Putative PLP-dependent transferase | 0.32 0.11 + |
| CD939 Putative hydrolase, HAD superfamily | 0.38 |
| CD3350 Putative glycosyl transferase, family 2 | 0.09 |
| CD1951 Putative Acyl-CoA N-acyltransferase | 0.19 |
| CD2749 Putative DNA helicase, UvrD/REP type | 0.32 0.05 |
| CD2537 Membrane-associated 5'-nucleotidase/phosphoesterase | 0.15 0.22 |
| CD1845 Membrane protein Tn1549 | 0.14 0.17 − CD1845-CD1846 |
| CD1846 Conjugative transposon protein Tn1549 | 0.15 0.20 − |
| CD564 Putative ATP-dependent protease | 0.39 σK − |
| **proteins of unknown function** | |
| CD9196 Fragment of conserved hypothetical protein | 0.18 0.18 − |
| CD9896 Conserved hypothetical protein | 0.05 0.02 σK − |
| CD1063.1 Conserved hypothetical protein | 0.17 0.21 σK + |
| CD1063.2 Conserved hypothetical protein | 0.01 0.04 σK − CD1063.2-CD1063.3 |
| CD1063.3 Conserved hypothetical protein | 0.01 0.04 − |
| CD1065 Conserved hypothetical protein | 0.06 0.30 σK ?* |
| CD1067 Conserved hypothetical protein | 0.00 0.03 σK + |
| CD1133 Conserved hypothetical protein | 0.12 0.16 σK + |
| CD1581 Conserved hypothetical protein | 0.01 0.01 + |
| CD2055 Conserved hypothetical protein | 0.33 |
| CD2409 Conserved hypothetical protein | 0.11 σK − |
| CD3380 Conserved hypothetical protein | 0.12 0.12 σK + |
| CD3613 Conserved hypothetical protein | 0.04 σK + |
| CD3620 Conserved hypothetical protein, similar to YmaF | 0.34 0.04 + |

*A gene is considered differentially expressed when the P value is <0.05 using the statistical analysis described in Materials and Methods.

*We search for promoters recognized by the different sigma factors active in mother cell upstream of the TSS mapped by RNA-seq. These promoters are listed in Table S7.

*Proteins that are detected associated with the spore by a proteomic approach [39].

*Using the consensus sequence recognized by Spo0IId in B. subtilis [14], we identified a putative Spo0IId binding motif upstream of these genes.

*σK? indicated the presence of sequences similar to σK consensus elements upstream of these genes but the TSS was not mapped (see Table S7).

DOI:10.1371/journal.pgen.1003756.t002
between the 630Δerm strain and the sigE or sigF mutant in our conditions. In addition, $\sigma^E$ activated the expression of the spoIVA octacistronic operon encoding an ATPase and membrane proteins that localize to the outer forespore membrane in B. subtilis, and integrate with SpoI IQ a novel type of secretion system [6,37,30].

In B. subtilis, the synthesis of the protective envelopes that encase the spore, the coat and the cortex PG, is initiated under $\sigma^E$ control [12]. In C. difficile, the expression of spoIVA and sipL encoding an ortholog of the SpoIVA morphogenetic ATPase and a functional homolog to B. subtilis SpoVID was downregulated in the sigE mutant. These proteins are required for proper spore coat localization around the forespore in C. difficile [40]. The expression of colB encoding a recently identified coat protein [41] was also specifically regulated by $\sigma^E$. Accordingly, a $\sigma^E$-dependent promoter was mapped upstream of spoIVA, sipL, and cotB (Table S7). Moreover, an ortholog of the colD gene, encoding a N-acetylmuramoyl-L-alanine amidase required for the synthesis of muramic $\varepsilon$-lactam, a specific cortex PG component [42], was controlled by $\sigma^E$. On the contrary, the second enzyme of this pathway, the PdaA deacetylase, was produced in the forespore (see below). The synthesis of other polysaccharide deacetylases (CD3257, CD3248, CD1319), of a polysaccharide biosynthetic enzyme (CD1608), of an alanine racemase (Ahr2) and of a N-acetylmuramoyl-L-alanine amidase (CD2761) was also under $\sigma^E$ control. CD1168 encoding an YblJ-like protein, and the CD2443-CD2442 operon encoding YgcC and YgdD-like proteins were controlled by $\sigma^E$. The inactivation of ygcC, ygdD and yblJ in B. subtilis reduces heat resistance of the spore and prevents the development of refractility, both phenotypes attributed to defects in cortex formation [12,13]. In addition, the expression of the spoK-spmB operon involved in spore core dehydratation and heat resistance in B. subtilis and C. perfringens [43] was reduced in the sigE mutant. We observed a positive control by $\sigma^G$ of the synthesis of CD2963, an homolog of a bacterioferritin involved in iron storage and oxidative stress protection in the anaerobe Desulfovibrio vulgaris [44].

Several genes under $\sigma^E$ control in B. subtilis are required for spore germination because they are involved in modifications of the cortex or in coat assembly. In C. difficile, $\sigma^E$ was required for the transcription of the csp operon encoding serine proteases. In C. perfringens and C. difficile, CapB is necessary for pro-SleC maturation to form the spore cortex lytic enzyme SleC during germination [45,46,47] while CapG is a germinant receptor for taurocholate [48]. The promoter of the C. difficile csp operon matched the consensus for $\sigma^E$-dependent promoters while sleC was transcribed from a $\sigma^K$-dependent promoter (see below). We also observed a positive control by $\sigma^E$ of three other proteases (CD2000, CD3455, CD1320). CD3455 shares 38% identity with CspB, a protease required for $\sigma^E$ activation in B. subtilis [6] and the Isp ortholog of CD2000, is involved in sporulation in Bacillus thuringiensis [49]. Presumably, these proteases also participate in the spore formation process in C. difficile.

Another major function under $\sigma^E$ control in B. subtilis is the maintenance of a sufficient level of metabolic activity to enable the completion of sporulation [9,12]. $\sigma^E$ also controlled the expression of several metabolic genes in C. difficile (Table 2). Among them are genes encoding proteins likely to be required for oxidoreductation pathways and energy metabolism: an oxido-reductase (CD2429-1-CD2429-CD2426), a FMN reductase (CD3637), a dehydrogenase (CD3251) and an iron hydrogenase (CD3258). In addition, $\sigma^E$ activated the expression of genes encoding peptides (CD3652, CD1080, CD3521) or amino acids perreas (CD1746, CD1555 and CD1259). Moreover, the CD2833 and CD0760 genes encoding a Ca$^{2+}$-transporting ATPase and a Ca$^{2+}$/Na$^+$ antipporter, respectively, were downregulated in the sigE mutant. These proteins are probably important for the import or export of Ca$^{2+}$, a key element of the sporulation and germination processes in endosporing Firmicutes [47].

Other C. difficile $\sigma^E$ controlled genes do not belong to any of the categories above. Of particular interest, $\sigma^E$ controlled the expression of the mazF-maeE operon, encoding the unique toxin-antitoxin (TA) system of C. difficile [50]. In B. subtilis, a different TA system, SpoIISA-SpoISB synthesized in the mother cell, is involved in sporulation [51]. MazE cleaves mRNA while the SpoIISA toxin is targeted to the cell membrane. The role of MazEF in the sporulation process remains to be characterized. Finally, 18 genes encoding proteins of unknown function were downregulated in the sigE mutant compared to the 630Δerm strain. Only 15 members of the $\sigma^E$ regulon are detected in the spore proteome [39]. This included the SpoIVA and SipL proteins, the CotB coat protein, the Csp-like proteases, the YgcL-like protein and three proteins of unknown function (CD2024, CD3522 and CD1930).

The $\sigma^Ge$ regulon. From the transcriptome experiments, we identified 50 genes controlled by $\sigma^G$ but not by $\sigma^E$ (Table 1). The TSS mapping performed by RNA-seq allowed us to identify 30 $\sigma^G$-dependent promoters controlling the expression of 33 genes. Several $\sigma^G$-controlled genes encode proteins sharing similarities with proteins involved in spore resistance in B. subtilis [5,16,52]. Dipicolinic acid (DPA) is a major spore component important for spore resistance. The expression of the spoIVA operon whose products are required for the import of DPA into the forespore from the mother cell [53], decreased in a sigG mutant (Table 1). A C. perfringens spoIVA mutation prevents the accumulation of DPA and reduces spore viability [47]. As observed in B. subtilis, we also showed the requirement for $\sigma^E$ for the expression of all the genes encoding SASPs; these proteins bind to the forespore chromosome protecting the DNA from damage [52]. Indeed, the spoA (CD2688) and spoB (CD3249) genes encoding alpha/beta-type SASP and two other genes annotated as SASPs (CD1290 and CD3220) were expressed under the direct $\sigma^E$ control (Table 1). Moreover, CD0684 encoding an ATP-dependent protease sharing similarity with FisI, a protein involved in protein quality control and stress resistance in eubacteria, was downregulated in a sigG mutant. In addition, genes encoding proteins presumably involved in mitigating oxidative stress were also controlled by $\sigma^G$. CD1567 (cotG), CD1631 (sodA) and CD2045 encode a catalase, a superoxide dismutase and a rubrythrin, respectively. Interestingly, SodA and CotG has recently been detected at the C. difficile spore surface [54]. These results suggested that proteins synthesized under the control of $\sigma^G$ are located to the spore surface in C. difficile.

Expression of several genes required for cell wall synthesis and cortex formation was controlled by $\sigma^E$. Thus, expression of CD1470 (pdaA) encoding a N-acetylmuramic acid deacetylase involved in the formation of muramic $\varepsilon$-lactam in the spore cortex PG and indirectly required for efficient spore germination [42] was downregulated in a sigG mutant. In addition, dacF, encoding a D-alanyl-D-alanine carboxypeptidase regulating the degree of cross-linking of the spore PG as well as sipB and uppS coding for an amidase and an undecaprenyl pyrophosphate synthetase, respectively were also positively controlled by $\sigma^G$.

Finally, the transcriptome analysis revealed 16 $\sigma^E$-controlled genes coding for proteins of unknown function. One of the most important aspects of spore biology is the mechanism by which dormant spores sense a suitable environment for germination and trigger the process through the specific recognition of germinants by germination receptors. These receptors are located in the spore inner membrane and are synthesized in the forespore under $\sigma^G$. 

PLOS Genetics | www.plosgenetics.org

October 2013 | Volume 9 | Issue 10 | e1003756
control in *B. subtilis* [9,10]. No known homologs of the *B. subtilis* *gerA*, *gerB* and *gerK* operons as well as of the *gerA*, *gerKB*, *gerK1* and *gerK2* genes of *C. perfringens* are found in the *C. difficile* genome [35,36]. However, a bile acid germinant receptor CsgG has been recently identified [40]. Nine genes encoding probable membrane proteins (CD0792, CD0793, CD1677, CD1789, CD2051, CD2645, CD2655, CD2636, CD3551.1) were downregulated in the *sigG* mutant and these proteins may be involved in germination. 55% of our σ^K^-controlled genes encode proteins associated to the spore in proteome [39]. This is in agreement with the crucial role of σ^K^ in the assembly of the spore surface layers, spore maturation and mother cell lysis in *C. difficile* as well as in other endospore forming Firmicutes [19,57].

### III. The forespore line of gene expression

Our global approaches allowed to obtain new insights into the forespore line of gene expression, which is governed by the RNA polymerase sigma factors, σ^F^ and σ^G^, and the DNA-binding transcriptional regulator, SpoVT.

**Control of sigF and sigG expression.** The spoIIGA-spoIAB-sigF operon, which is responsible for the synthesis and the activation/inactivation of σ^F^, is transcribed by the RNA polymerase σ^A^ holoenzyme and is positively controlled by Spo60 [27]. Like in *B. subtilis* and other spore formers, the *spoIIGA*, *sigE* and *sigG* genes are clustered in *C. difficile* [16,20]. The *B. subtilis* *sigG* gene is transcribed from two promoters, one located upstream of *spoIIGA* (σ^F^) and the second located upstream of *sigG* transcribed first by σ^F^ and later by σ^G^ after engulfment [5]. In both *B. subtilis* and *C. acetobutylicum*, three transcripts are detected for the *spoIIGA-sigE-sigG* locus, corresponding to a *spoIIGA-sigE-sigG*, a *spoIIGA-sigE* and a *sigG* transcript [5,33,62]. In *C. difficile*, a σ^G^ dependent promoter (GTGACA and TATAAT boxes) was mapped by RNA-seq upstream of *spoIIGA* (TSS at position 3052270 in the genome of strain 630). A promoter was also mapped upstream of *sigG* both by RNA-seq (Table S6) and by 5′-RACE [27]. The associated consensus motifs clearly correspond to those found for the forespore sigma factors (Fig. 1A) but appear closer to that found for σ^P^ promoters as the −10 element contains a G at position 23 and lacks the A conserved at position 32 for σ^G^ dependent promoters (Fig. 1A). However, sigG expression was not downregulated in the sigF mutant in the transcriptome analysis as well as by qRT-PCR, under our conditions. In *B. subtilis*, sigG expression also appears not to be regulated in transcriptome analyses comparing sigF^F^ and sigF^F^ strains [9,10]. By contrast, sigG expression is downregulated in a sigF mutant in *C. acetobutylicum* [22] suggesting some differences in the control of sigG expression. In *C. difficile*, the absence of control of sigG expression by σ^F^ strongly suggests that sigG is transcribed from at least two promoters, one in front of *spoIIGA* recognized by σ^F^, and one just upstream of *sigG* recognized by σ^P^ and/or by σ^A^ studied with a PsigG-SNAP transcriptional fusion indicate that this second promoter responsible for the forespore-specific transcription of sigG is dependent on σ^A^ [19]. A more complete analysis of the fine tuning of *spoIIGA-sigE-sigG* locus expression will require further investigations.

It is worth noting that the expression of many sigG-controlled genes (82% of σ^G^ regulon members) including those for key spore proteases like SpoVA, SpSA, SpSb and PdaA was strongly reduced in a sigF mutant in transcriptome (Table 1). The absence of detection of the complete σ^G^ regulon in the sigF mutant might be due to the timing of sampling (14 h) that is probably not optimal for sigG targets. Moreover, inactivation of sigF also eliminated fluorescence from a PsigF-SNAP fusion [19]. Thus, σ^A^ directly or indirectly controls expression of the σ^G^ regulon. One possibility is that the σ^F^ protein is abolished in a sigF mutant as observed in *C. acetobutylicum* and *C. perfringens* [22,23]. Alternatively, the positive regulation of σ^G^ target genes by σ^A^ could be mediated through modulation of σ^A^ activity. In *B.
subtilis, for instance, two anti-sigma factors, SpoIAB and CsfB (Gin) negatively regulate σ^d activity. CsfB, in particular, plays an important role in σ^g regulation in the forespore, and is produced under σ^g control [63,64]. No Gin ortholog is present in the C. difficile genome and the phenotype expected if σ^g would be required for the synthesis of an anti-σ^g factor would be the opposite. However, we cannot exclude that proteins synthesized under σ^g control might be necessary for σ^g activity. Yet another possibility is that both σ^f and σ^g were involved in the transcription of most σ^g-target genes. In B. subtilis, σ^f and σ^g have overlapping promoter specificities and some genes are under the dual control of these sigma factors [9,10]. This could be also the case in C. difficile. Finally, σ^f might control σ^g targets through an as yet unknown regulatory factor.

The role of SpoVT in sporulation. In B. subtilis, two transcriptional regulators participate in the forespore regulatory network, RsαA and SpoVT. SpoVT controls the synthesis of about half of the members of the σ^g regulon [10]. RsαA is absent from the genome of C. difficile and several Clostridia while an ortholog of SpoVT (55% identity with SpoVT of B. subtilis) is present.

Interestingly, we found that spoVT expression was controlled by both σ^f and σ^g in transcriptome (Table 1). We confirmed by qRT-PCR a 280- and a 14-fold decrease of spoVT transcription in a sigF or in a sigG mutant compared to strain 630Aerm, respectively (Table 3). We also showed that the expression of spoVT was restored to the wild-type level in the sigG mutant complemented by sigG while its expression was 19-fold higher in the sigF mutant complemented by sigF compared to the wild-type strain (Table S9). Upstream of the spoVT TSS, we identified a promoter resembling both the σ^f and σ^g consensus elements with the presence of a G at position 23 and an A at position 32 (Table S6). So, spoVT might be a direct target of both σ^f and σ^g because contrary to other σ^g-controlled genes, σ^g inactivation caused a much more important decrease of spoVT expression than σ^f inactivation.

To test the possible role of SpoVT in sporulation in C. difficile, we constructed a spoVT mutant (Fig. S2). A complemented strain, CDIP263 carrying a multi-copy allele of spoVT under the control of its native promoter was also obtained. To determine the impact of SpoVT inactivation on spore morphogenesis, samples of the strain 630Aerm, of the spoVT mutant and of the complemented strain CDIP263 (spoVT::erm, pMTL84121-spoVT) were collected and labeled with the DNA stain DAPI and the lipophilic membrane dye FM4-64. Phase contrast and fluorescence microscopy experiments showed that the spoVT mutant was able to complete the engulfment process (Fig. 2). However, this mutant formed phase dark immature spores, which were not released from the sporangial cells, clearly suggesting that the cortex is absent or highly reduced. The wild-type phenotype was restored in the complemented strain CDIP263. We further tested the ability of this mutant to sporulate. After 72 h of growth in SM medium, the 630Aerm strain or the spoVT mutant containing the plasmid pMTL84121-spoVT produced 4×10^6 and 10^7 heat-resistant spores/ml, respectively. In contrast, no clones were obtained after an heat treatment for the spoVT mutant. This was probably due to the lack of production of heat-resistant spores as suggested by the phase dark spores phenotype observed for this mutant (Fig. 2). In conclusion, SpoVT is required for mature spore formation in C. difficile but the phenotype of the spoVT mutant of C. difficile differs from that of the same mutant of B. subtilis. Indeed, the spoVT mutant of C. difficile formed phase dark spores instead of phase bright spores [65] and the spoVT mutant of B. subtilis has a reduced ability to sporulate [65] while we were unable to detect heat resistant spores for the mutant of C. difficile.

Then, we used qRT-PCR to measure the impact of SpoVT inactivation on the expression of selected σ^f targets after 20 h of growth. We showed that the expression of sigA and sigB decreased 70- and 12-fold, respectively in the spoVT mutant compared to strain 630Aerm (Table S11). A positive effect of SpoVT on sigA and sigB expression is also observed in B. subtilis [63,64]. The decreased expression of the sigA and sigB genes is probably not sufficient to explain the phenotypes observed for the spoVT mutant but we did not see a reproducible differential expression for other tested genes in our conditions. However, the expression of spoHR and gpr, two members of the σ^g regulon, was upregulated 4- and 9-fold in a spoVT mutant compared to strain 630Aerm earlier during sporulation (15 b) (Table S11). This negative effect might be due either to a direct control by SpoVT playing the role of RsαA in the negative control of σ^f targets or to an indirect effect associated with the blocked sporulation in the spoVT mutant.

No other genes encoding transcriptional regulators or proteins with DNA binding motifs were regulated by σ^f or σ^g in our transcriptome study (Table 1). Even though we cannot exclude that other regulatory proteins are at play, SpoVT is probably the key, if not the only, ancillary regulator of forespore gene expression. Further work will be required to analyse more precisely the morphology of the spoVT mutant and to understand at the molecular level the complex role of SpoVT in the regulatory network controlling sporulation.

IV. The mother cell line of gene expression

As discussed above, sigE transcription is probably initiated at the σ^f-dependent promoter located upstream of spoIIG4, which forms an operon with sigE and maybe also with sigG. In B. subtilis, σ^f positively controls the expression of spoIIG and gerR encoding regulatory proteins. GerR acts as a repressor of certain early σ^g-controlled genes and as an activator of some σ^f-dependent genes [67]. SpoIIG is an ambivalent regulator, acting as a repressor as well as an activator of a late class of σ^g-controlled genes [14]. SpoIIG and σ^f jointly activate the expression of sigK, and σ^k in turn drives production of another ancillary regulator, GerE [14]. In C. difficile, the SpoIIG regulator is present while GerR and GerE are absent. The SpoIIG protein of C. difficile shares 64% identity with SpoIIG from B. subtilis with conservation of the HTH motif and of the basic DNA binding motif located near the C-terminus of the protein (Fig. S3) [68].

Characterization of the SpoIIID regulon. The expression of spoIIID was strongly curtailed in a sigE mutant (25- and 800-fold in transcriptome and qRT-PCR) whereas the spoIIID expression was restored when a plasmid pMTL84121 containing the sigE gene with its promoter was introduced into the sigE mutant (Table S9). A TSS corresponding to a σ^g-controlled promoter (ATA-NCATATA) was mapped upstream of spoIIID as observed in B. subtilis. It is interesting to note that in C. perfringens the expression of spoIIID is not controlled by σ^f [21].

In B. subtilis, SpoIIID positively or negatively regulates almost half of the σ^f target genes [14]. To test the possible role of SpoIIID in the mother cell regulatory network in C. difficile, we constructed a spoIIID mutant (Fig. S2) and compared the expression profiles of the 630Aerm strain and of the spoIIID mutant after 15 h of growth in SM medium. We found 96 genes differentially expressed with a p value <0.05 between the spoIIID mutant and the wild-type strain (Table S10). 12 and 84 genes were down and upregulated in the spoIIID mutant, respectively. We then performed qRT-PCR on a subset of genes regulated by SpoIIID in our transcriptome. The qRT-PCR results confirmed the microarrays data for the tested genes (Table S5).

First, 47 genes that are not under the control of sporulation sigma
Factors in our transcriptome analyses were regulated by SpoIIID either positively (10 genes) or negatively (37 genes) (Table S10). Most of these genes encode proteins involved in metabolism. Among the genes under the negative control of spoIIID, 28 were bona fide members of the $\sigma^E$-regulon (Table 2). Therefore, a principal function of SpoIIID is to inhibit the transcription of about 30% of the genes transcribed by $\sigma^E$ as observed in *B. subtilis* [14]. SpoIIID repressed the expression of the spoIIIA operon and of the spoIIB, spoIVA, and spoIVA genes. In *B. subtilis*, spoIIIA and spoIVA are direct SpoIIID targets [14]. By contrast, the spoIID gene, which is a direct SpoIIID target in *B. subtilis*, was not regulated by SpoIIID in transcriptome and qRT-PCR experiments in our conditions. In addition, the expression of genes encoding a transmembrane signaling protein (CD2445), an iron hydrogenase (CD3258), a bacterioferritin (CD2865), three polysaccharide deacetylases (CD3257, CD3248, and CD1319) and of several genes of unknown function was positively controlled by $\sigma^E$ and negatively regulated by SpoIIID (Table 2). By using the consensus sequence recognized by SpoIIID in *B. subtilis* [14], we identified a putative SpoIIID binding motif upstream of 11 of these genes (Table 2) suggesting that they might be direct SpoIIID targets in *C. difficile*. However, the characterization of direct and indirect SpoIIID targets and the experimental identification of the DNA binding motif of SpoIIID will deserve further investigations.

### Table 3. Control of $\sigma^G$ or $\sigma^K$ target genes by both $\sigma^E$ and $\sigma^F$.

| Gene | Function | Expression ratio |
|------|----------|------------------|
| CD1486 | Putative ribosome recycling factor | 0.10 0.26 0.17 |
| CD1631 | Superoxide dismutase (Mn) | 0.06 0.21 0.32 |
| CD1880 | Conserved hypothetical protein | 0.07 0.21 0.25 |
| CD2349 | Small, acid-soluble spore protein beta | 0.01 0.04 0.07 |
| CD3511 | Putative membrane protein | 0.11 0.31 0.18 |
| CD0773 | DPA uptake protein, SpoVAC | <0.01 0.06 0.02 |
| CD0792 | Putative membrane protein, DUF81 family | <0.01 0.02 0.01 |
| CD2123 | Stage IVB sporulation protein B, peptidase | 0.04 0.08 0.12 |
| CD1290 | Putative small acid-soluble spore protein | <0.01 0.05 0.01 |
| CD1430 | Putative d-lactam-biosynthetic deacteylase | <0.01 0.14 0.03 |
| CD2375 | Conserved hypothetical protein | 0.03 0.24 0.05 |
| CD2465 | Putative amino acid/polyamine transporter | 0.18 0.15 0.19 |
| CD2536 | Putative membrane protein | <0.01 0.13 0.03 |
| CD2588 | Small, acid-soluble spore protein alpha | <0.01 0.05 0.02 |
| CD3499 | Transcriptional regulator, SpoVT | <0.01 0.06 0.07 |
| CD0332 | Putative exosporium glycoprotein | 0.43 0.21 0.08 |
| CD0598 | Spore-coat CotCB manganese catalase | 0.18 0.18 0.01 |
| CD0896 | Conserved hypothetical protein | 0.21 0.05 0.02 |
| CD1067 | Conserved hypothetical protein | 0.07 0.00 0.03 |
| CD1581 | Conserved hypothetical protein | 0.03 0.01 0.01 |
| CD2664 | Conserved hypothetical protein | 0.13 0.10 0.05 |
| CD3230 | Putative exosporium glycoprotein | 0.16 0.12 0.04 |
| CD3349 | Exosporium glycoprotein BclA3 | 0.12 0.01 0.02 |
| CD3620 | Conserved hypothetical protein | 0.35 0.34 0.04 |
| CD551 | Spore cortex-lyic enzyme pre-pro-form | 0.34 <0.01 0.01 |
| CD1133 | Conserved hypothetical protein | 0.40 0.02 0.10 |
| CD1333 | Spore coat protein CotE peroxiredoxin/chitinase | 0.32 0.01 0.01 |
| CD2401 | Spore coat CotD manganese catalase | 0.24 0.01 <0.01 |

A gene is considered differentially expressed when the P value is <0.05 using the statistical analysis described in Materials and Methods. qRT-PCR were performed as described in Materials and Methods. doi:10.1371/journal.pgen.1003756.t003

PLOS Genetics | www.plosgenetics.org 16 October 2013 | Volume 9 | Issue 10 | e1003756
transcriptome experiment (Table S10). To determine if sigK itself and other members of the \( \sigma^K \) regulon were under SpoIIID control, we performed qRT-PCR experiments (Table S11) using RNA extracted after 15 h of growth or after 24 h, a time where the \( \sigma^K \) targets are more highly expressed. We found that sigK expression was 25-fold and 250-fold downregulated in the spoIIID mutant after 15 h and 24 h of growth, respectively. This clearly indicated that the expression of sigK was positively controlled by SpoIIID in C. difficile as observed in B. subtilis [14]. In B. subtilis, SpoIIID is also required for the expression of spoIVCA encoding the site-specific recombinase involved in the excision of the \( \sigma^K \) element that creates a functional sigK gene [69]. In contrast, we did not observe a control of CD1231 encoding the specific recombinase of the \( \sigma^K \) element by SpoIIID but also by SigE both in transcriptome and qRT-PCR experiments. This result strongly suggests a different mechanism of control for the \( \sigma^K \) excision in C. difficile compared to B. subtilis. We also showed that eleven bona fide \( \sigma^K \) targets such as cotA, cotCB, cotD, cotE, sleC, bclA1, bclA2, bclA3, CD3580, CD1067 and CD1133 were downregulated 30- to 1000-fold in the spoIIID mutant as compared to the parental strain after 24 h of growth (Table S11). The positive regulation of members of the \( \sigma^K \) regulon by SpoIIID might be due either to a direct binding of this regulator to their promoter regions or to an indirect effect mediated through the control of sigK transcription by SpoIIID.

To investigate the role of SpoIIID in sporulation, we examined the morphology of the spoIIID mutant. Samples of the 630Δerm strain, of the spoIIID mutant and of the complemented strain CDIP262 carrying a multicopy allele of spoIIID under the control of its native promoter (spoIIID::erm, pMTL84121-spoIIID) were collected and labeled with the DNA stain DAPI and the lipophilic membrane dye FM4-64. Phase contrast and fluorescence microscopy showed that the spoIIID mutant completed engulfment of the forespore by the mother cell and formed partially refractile spores with irregular shapes and positioning (Fig. 2). The wild-type phenotype with the production of free spores was restored in the strain CDIP262 containing a copy of spoIIID on a plasmid (Fig. 2). We also tested the ability of the spoIIID mutant to sporulate. After 72 h of growth in SM medium, the 630Δerm strain produced \( 9 \times 10^5 \) heat-resistant spores/ml while a titer of \( 3 \times 10^2 \) heat-resistant spores/ml was obtained for the spoIIID mutant. When we complemented the spoIIID mutant with a pMTL84121-spoIIID plasmid, the titer of heat resistant spores increased to \( 9 \times 10^5 \) heat-resistant spores/ml.
spores increased to $5 \times 10^7$/ml. Interestingly, both the morphology and the sporulation efficiency of the spoIID mutant is reminiscent of the oligosporogenous phenotype obtained for the sigK mutant of C. acetobutylicum [19] in agreement with the dependency on SpoIIID of the transcription of sigK. The morphology and the sporulation efficiency of the spoIID mutant of C. difficile are similar to those of the B. subtilis mutant [70,71]. As demonstrated in B. subtilis [14], SpoIIID in C. difficile plays a pivotal role in the mother cell line of gene expression switching off the transcription of many members of the $\sigma^k$ regulon and switching on the expression of $\sigma^K$ and of members of the $\sigma^F$ regulon.

Control of sigK transcription. In addition to the positive control of sigK expression by SpoIID, its expression also decreased 8- and 600-fold in a sigE mutant, in transcriptome and qRT-PCR experiments, respectively. The introduction of a plasmid pMTLB4121 containing the sigE gene with its promoter into the sigE mutant restored sigK expression (Table S9). Interestingly, we mapped two TSS, 67 nt (P1) and 26 nt (P2) upstream of the translational start site of sigK (Table S7). A canonical −10 box (CATATTAT) for mother cell sigma factors is located upstream of the TSS corresponding to P1 while either a TTA sequence 15 bp upstream from the −10 box or a TTT motif with a more classical 16–18 bp spacing between the −10 and −35 elements could be proposed for a −35 motif. So, this promoter resembles a $\sigma^K$-dependent promoter (Fig. 1B and Table S7). Upstream of the second TSS (P2), we found a consensus for $\sigma^K$-dependent promoters with a −10 box (CATATAT) and a −35 box (AC) (Table S7). We note that in B. subtilis, following the initial transcription of sigK under the command of $\sigma^k$ and SpoIID, an autoregulatory loop is established, which is responsible for about 60% of sigK transcription [72]. In C. difficile, sigK is likely first transcribed by $\sigma^k$-associated RNA polymerase at P1 and then by $\sigma^K$-associated RNA polymerase at P2 probably later during sporulation. In any event, lending support to the idea that $\sigma^k$ has a crucial role in sigK transcription, 65% of the $\sigma^k$-controlled genes were positively controlled by $\sigma^k$ including genes encoding key components of the spore surface layers (catGB, catD, catE, colD, sbcC, bclA1, bclA2, bclA3). Most of these genes were much more downregulated in the sigK mutant than in the sigE mutant (Table 2) suggesting that the timing of sampling of the sigE mutant (14 h) is probably not optimal for $\sigma^k$ targets. This might explain the absence of detection of the complete $\sigma^k$ regulon in the sigE mutant. Importantly, with the exception of ndkK, we did not identify other genes for putative transcriptional regulators that were downregulated in sigKE or sigK mutant (Table 2). Overall, our data indicates that in C. difficile, the mother cell line of gene expression is deployed according to a hierarchical regulatory cascade simpler than in B. subtilis [14].

The regulation of mother cell sigma factors synthesis in C. difficile also differs from other Clostridia. In C. perfringens, a biphasic pattern of sigK expression (early and late in sporulation) is observed and $\sigma^k$ and $\sigma^F$ coregulate the expression of each other [21]. In C. botulinum, sigK is expressed at the onset of stationary phase and $\sigma^k$ positively controls the expression of sigE and spo0A [73]. In C. difficile, we did not observe a control of spo0A, sigK or bona fide members of the $\sigma^k$ regulon by $\sigma^K$ (Table 2). In both C. perfringens and C. botulinum, sporulation is arrested at an early stage in a sigK mutant and not at a late stage as observed in B. subtilis and C. difficile [5,19]. It is also worth noting that the sigK gene is disrupted by a skin element in B. subtilis and C. difficile but not in C. acetobutylicum, C. perfringens and C. botulinum, suggesting that excision of this element may be an important factor in controlling the timing of $\sigma^k$ activity in C. difficile.

V. Communication between the forespore and the mother cell

A hallmark of sporulation in B. subtilis is the existence of cell-cell signaling pathways that link the forespore and mother cell-specific lines of gene expression. Because these pathways operate at critical morphological stages of sporulation, the result is the coordinated deployment of the two lines of gene expression, in close register with the course of morphogenesis [5,6]. Indeed, $\sigma^F$ is required for the activation of pro-$\sigma^k$ into $\sigma^k$ in the mother cell and $\sigma^k$ in turn is necessary to activate $\sigma^F$ in the forespore. Finally, $\sigma^k$ is required for the activation of pro-$\sigma^F$ into $\sigma^F$ in the mother cell [3,6]. We obtained new information on the intercompartment signaling pathway in C. difficile.

Absence of a strict control of the $\sigma^k$ regulon by $\sigma^F$. In B. subtilis, the $\sigma^k$ regulon is indirectly controlled by $\sigma^F$, which is required for the proteolytical activation of pro-$\sigma^k$. Surprisingly, we did not find any global control of the $\sigma^k$ regulon by $\sigma^F$ in C. difficile (Table S1 and Table S2). We confirmed by qRT-PCR that the expression of bona fide $\sigma^k$ targets such as spoIIA, spoIVA, spoIID or CD2064 did not significantly decrease in a sigF mutant (Fig. S4).

This is a major difference relative to the B. subtilis sporulation regulatory network [74] and indicates that $\sigma^k$ is not strictly required for $\sigma^F$ functionality in C. difficile. One possibility is that the SpoIIGA mediated proteolytical activation of pro-$\sigma^k$ is not fully dependent on $\sigma^F$. To test this hypothesis, we detected in several strains (630Δerm, spo0A, sigK and sigE strains) the $\sigma^F$, $\sigma^k$ and pro-$\sigma^k$ polypeptides by Western-blotting using antibodies raised against either $\sigma^F$ or $\sigma^k$ (Fig. 3). In strain 630Δerm, we observed $\sigma^k$ production while two forms corresponding to $\sigma^F$ and pro-$\sigma^k$ were detected with an anti-$\sigma^k$ antibody. In a spo0A mutant, neither $\sigma^F$, pro-$\sigma^k$ nor $\sigma^k$ were detected (Fig. 3). This result is in agreement with the strong decrease of expression of the spoIIA operon in a spo0A mutant [27]. As shown by qRT-PCR, spo0A inactivation also led to a 240-fold decrease of sigF expression indicating that this regulator controls pro-$\sigma^k$ synthesis. It has also been shown that Spo0A directly binds with low affinity to the spoIGA and spoIIA promoter regions [30]. In a sigE mutant, we did not detect $\sigma^F$ or pro-$\sigma^k$ with an anti-$\sigma^k$ antibody while we specifically detected $\sigma^F$ with an anti-$\sigma^k$ antibody. Interestingly, in a sigF mutant in which $\sigma^k$ was absent, we detected pro-$\sigma^k$ but also the processed $\sigma^k$ form in a reduced quantity compared to the situation in a wild-type strain (Fig. 3). The expression of the $\sigma^k$ regulon in the sigF mutant shows that even reduced the level of active $\sigma^F$ in this mutant is sufficient to allow the transcription of $\sigma^k$-controlled genes. This clearly demonstrates that an active $\sigma^k$ protein can be produced in the absence of $\sigma^F$ in C. difficile contrary to B. subtilis [75]. In C. perfringens and C. acetobutylicum, neither pro-$\sigma^k$ nor $\sigma^k$ protein are produced in a sigF mutant [22,23] suggesting a diversity in the mode of forespore control of $\sigma^k$ activity among spore forming Firmicutes.

The role of SpoIIR in the regulatory cascade. In B. subtilis, $\sigma^F$ drives production of the signaling protein SpoIIR, which is secreted across the forespore inner membrane into the intermembrane space, where it stimulates the SpoIIGA-dependent pro-$\sigma^k$ processing in the mother cell [6,11,76]. The results exposed above raised the possibility that SpoIIR is dispensable for pro-$\sigma^k$ processing in C. difficile, or that expression of the spoIIR gene is partially independent on $\sigma^F$. To determine the role of SpoIIR in C. difficile, we inactivated spoIIR. A complemented strain, CDIP246, carrying a multicopy allele of spoIIR under the control of its native promoter was also constructed. We first examined the
Support the idea that expression of the pMTL84121-spoIIR mutant compared to strain 630

It follows that the expression of spoIIR mutant has to occur in part independently of \( \sigma^F \). We first showed that the expression of \( \sigma^F \) increased 2.5- and 30-fold in a sigF mutant in transcriptome and by qRT-PCR (Table S1 and Table S3) while the introduction of pMTL84121-sigF into the sigF mutant increased spoIIR expression 9-fold compared to strain 630Aerm (Table S9). So, the spoIIR transcription is under \( \sigma^F \) control. We also constructed a transcriptional fusion between the spoIIR promoter and the SNAP tag [19]. This P::spoIIR-SNAP\(^{E_{Cd}}\) fusion was introduced by conjugation into strain 630Aerm and a spo0A, sigF or sigE mutant. Labeling with the fluorescent substrate TMR allowed localization of P::spoIIR-SNAP\(^{E_{Cd}}\) expression. When examined by fluorescence microscopy, most of the cells of strain 630Aerm showed fluorescence in the forespore as expected for a gene under \( \sigma^F \) control (Fig. 5). spoIIR expression was not controlled by \( \sigma^E \) since a fluorescence signal was detected in the small compartments of disporic cells in a sigE mutant. Interestingly, a fluorescence of the spoIIR-SNAP\(^{E_{Cd}}\) fusion was still observed in the sigF mutant suggesting that some expression of spoIIR occurs in the absence of \( \sigma^E \). In a spo0A mutant, the fluorescence completely disappeared and the spoIIR expression strongly decreased by qRT-PCR. So, the residual expression observed in the sigF mutant is likely Spo0A-dependent. Presumably, the expression of spoIIR in the sigF mutant allows the production of sufficient SpoIIR to trigger pro-\( \sigma^E \) processing. It is worth noting that spoIIR is not downregulated in a sigF mutant in C. acetobutylicum and only partially under \( \sigma^A \) control in C. difficile (Fig. 5). Also interestingly, in B. subtilis, the forced expression of spoIIR in pre-divisional cells, still allows pro-\( \sigma^E \) processing to occur, even in the absence of \( \sigma^A \) [11,78]. It is possible that the activation of \( \sigma^A \) in ancestral spore formers occurred independently of the forespore. The expression of spoIIR under the exclusive control of \( \sigma^F \) may have appeared later and...
would allow a better coordination between the forespore and mother cell lines of gene expression.

The regulation of $\sigma^G$-dependent genes by $\sigma^K$. In B. subtilis, most of the $\sigma^G$ activity occurs after engulfment completion. In addition, the expression of $\sigma^G$ target genes in the engulfed forespore depends upon $\sigma^K$ activation in the mother cell at least in part through synthesis of the SpoIIIA proteins [5,6,37,38]. In the transcriptome analysis, we observed that the expression of 6 genes belonging to the $\sigma^G$ regulon (spoB, CD1486, CD1631, CD1880, CD2465 and CD3551.1) was downregulated in the sigE mutant (Table 3). We also examined by qRT-PCR if a sigE mutation affected the expression of several other $\sigma^{G'}$ targets. The expression of 9 additional $\sigma^{G'}$-controlled genes decreased in a sigE mutant (Table 3). Interestingly, SpoIID also repressed the expression of 12 members of the $\sigma^G$ regulon (spoA, spoB, spoVT, CD1631, CD1463, CD2112, CD2245.1, CD2375, CD2380, CD2809, CD3489, CD3610) in transcriptome (Table S10). This clearly indicated that both $\sigma^K$ and SpoIID participated in the control of $\sigma^{G'}$-target genes in C. difficile. In B. subtilis, $\sigma^{G'}$ activity is dependent on the SpoIIA-SpoIIE channel [6,37,38]. The SpoIIIQ proteins are encoded by the octacistronic spoIIIAG-AH operon, which is expressed in the mother cell under the direct control of $\sigma^K$ and SpoIID [9,12,14]. As noted above, the spoIII4 operon is a member of the $\sigma^K$ and spoIII3 regulons also in C. difficile (Table 2 and Table S10) and is transcribed from two promoters recognized by $\sigma^K$ located upstream of the spoIII4 and spoIII3G cistrons (Table S7), as observed in B. subtilis [79]. It thus seems possible that control of forespore-specific gene expression by $\sigma^K$ and SpoIID involves similar mechanisms in C. difficile. However, we note that fluorescence from a PspoA-SNAP<sup>Cd</sup> fusion is detected in the forespore compartment of sigE mutant cells [19]. This indicates that a strict requirement for $\sigma^K$ is not observed for the expression of C. difficile $\sigma^{G'}$-targets contrary to what is shown in B. subtilis suggesting a less tight control of $\sigma^K$ on gene expression in the forespore.

The absence of control of the $\sigma^K$ regulon by $\sigma^{G'}$. In B. subtilis, $\sigma^G$ regulates the expression of the $\sigma^K$ regulon in the mother cell through the control of pro-$\sigma^K$ processing [5,6]. A sigG mutant is blocked just after engulfment completion, and does not show any signs of assembly of the surface layers around the forespore. Surprisingly, no $\sigma^K$ target genes were downregulated in a sigG mutant in the transcriptome analysis and this lack of effect of the sigG mutation was confirmed by qRT-PCR for eight $\sigma^K$-controlled genes (cotA, cotCB, cotD, bclA1, bclA2, bclA3, CD1067, CD3620). Thus, synthesis of the spore coat proteins belonging to the $\sigma^K$ regulon (CotA, CotCB, CotD, CotE, CotF) is $\sigma^{G'}$-independent. This result is in agreement with the phenotype of the C. difficile sigG mutant that shows deposition of some coat material around the engulfed forespore [19]. Interestingly, mutations that bypass the need for $\sigma^G$ for pro-$\sigma^K$ processing, or a pro-less allele of the $\sigma^G$ gene in B. subtilis allow expression of $\sigma^K$ targets [80]. In C. difficile, the pro-sequence of $\sigma^K$ is absent and no homologs of SpoIVFB, the membrane-embedded protease that processes pro-$\sigma^K$ in B. subtilis and of two other membrane proteins (SpoIVFA and BoiA) that form a complex with and control the localization and activity of SpoIVFB, can be found [15,17,25,81]. Thus, $\sigma^K$ activity is not controlled through processing of a pre-protein in C. difficile and this may be related to the absence of $\sigma^{G'}$ control.
However, both the synthesis and the processing of pro-σK are also σF
dependent in C. perfringens [23]. Therefore, the absence of control of σK by σF is not restricted to Clostridia lacking a pro-σK protein.

Intriguingly, C. difficile codes for two homologs of the B. subtilis signaling protein SpoIVB. CD1215 and CD0783 share 37% and 36% identity with SpoIVB, respectively. CD0783 and CD1215 belong to the σF and σG regulons, respectively (Table 1). In B. subtilis, SpoIVB is a protease synthesized under the control of σF and σG. Similarly, spoIP is transcribed by σG and σF in B. subtilis, whereas two spoIP genes are present in C. perfringens and transcribed by σG and σK, respectively [82]. SpoIVB is secreted to the intermembrane space where it cleaves SpoVFA releasing SpoVF and allowing pro-σK processing [5,6]. However, B. subtilis SpoIVB has two distinct developmental functions: pro-σK processing control and another role essential for spore formation probably corresponding to the cleavage of SpoIIG [83,84]. It is possible that the C. difficile SpoIVB-like proteins retained the second role, which might correspond to an ancestral function.

Unexpectedly, 14 σF-target genes were downregulated in a sigF mutant compared to strain 630Aerm in transcriptome (cotA, cotCB, belA1, belA2, belA3, cagV, CD0896 CD1067, CD1581, CD1891, CD2184, CD2537, CD2664, CD5620). We further showed by qRT-PCR that 4 additional σF-target genes (sleC, cotD, cotE and CD1133) were positively controlled by σF (Table 3) suggesting a more global regulation of σK targets by σF. This effect might be mediated through an indirect control by σF of the synthesis or the activity of σK itself or of another regulator of members of the σK regulon. Another hypothesis might be that the σK activity requires engulfment completion a process blocked in a sigF mutant but not in a sigG mutant. Interestingly, the synthesis of pro-σK and σG in C. perfringens is also abolished in a sigF mutant but not in a sigG mutant [23] suggesting a switch in the forespore control of σK-activity from σG to σF. However, a sigK mutant of C. perfringens is blocked early in sporulation, at the asymmetric division stage [21]. The molecular mechanisms involved in the σF-dependent control of late mother cell-specific gene expression remain to be determined.

Conclusion. Global approaches combining transcriptome and TSS mapping allow us to have a view of the expression pattern during the early and late stages of sporulation in both forespore and mother cell and to define the four compartment-specific sigma regulons in C. difficile. We identified 25, 97, 50 and 56 members for the σF, σG, σK and σM regulons, respectively and in each regulon, we found key representatives of the homologous regulons of B. subtilis as proposed previously [16]. However, while larger than the evolutionary conserved core machinery for endosporulation, of about 145 genes [15,17], this set of C. difficile sporulation genes (around 225 genes) corresponds to about half the number of genes under the control of cell type-specific sigma factors of B. subtilis [9,10,12]. A more dynamic study with a detailed kinetic analysis would provide additional members for each regulon but it is also probable that the presumably more ancestral type of sporulation process proposed for Clostridia [15,17,18,20] involves a smaller collection of proteins. This has already been observed for the initiation of sporulation where the complex signaling transduction pathway involving in B. subtilis a phosphorelay that modulates Spo0A activity is replaced by a simple two-component system [18,20,29]. In general, the most significant variations between the B. subtilis and C. difficile sporulation process are observed at the interface with their environment: the signal transduction pathway triggering sporulation initiation, composition of the coat shell and the germination-activating pathways [16,18,20,47]. Indeed, the germination signals and germination receptors differ among spore forming Firmicutes and only a few spore coat layer proteins are conserved in Bacilli and Clostridia [47,57]. In each regulon, a significant fraction of genes encodes proteins with unknown function but our work offer new insights about the role of some of them. Some σF-controlled membrane proteins might be involved in germination, which remain poorly characterized in C. difficile [40,55]. Three proteins of unknown function belonging to the σK regulon (CD0214, CD1930, CD3522) and several proteins synthesized under σG control (CD1063, CD1067, CD1133, CD3580 and CD3613) that are detected in the spore proteome [39] are probably new spore coat components. Finally, we found several oxygen detoxification proteins that might help in the long-term survival of clostridial spores as recently proposed by Galperin et al [17]. This probably favors the dissemination of spores of strictly anaerobic Clostridia in aerobic environment, a crucial step for persistence and transmission of pathogenic Clostridia [1,2].

In this work, we also exposed important deviations from the B. subtilis paradigm in C. difficile. Both the global analysis of the program of gene expression under the control of the four cell type-specific sigma factors and the morphological characterization of the corresponding mutants (this work, [19]) indicate that coupling between gene expression and morphogenesis is less tight in C. difficile than in B. subtilis (Fig. 6). First, the σK regulon in the C. difficile mother cell is not strictly under σG control despite the fact that the forespore product SpoIR is required for pro-σG processing. The residual spoIR expression in the sigF mutant might be responsible for this less strict connection. Second, the tight coordination between σF and σG activities observed in B. subtilis, is absent in C. difficile since σG activity does not depend on σK as clearly shown by the morphology of a sigK mutant [19]. In the absence of a σK precursor, the rearrangement of the sigK gene associated to excision of the skin element may be the only way to control the timing of σK activity in C. difficile [16,20,25] and this event does not appear to be under σG control. However, a control of the forespore on σF target genes seems to be maintained through a σG-dependent regulation.

Recent data discussed above also suggest differences in the regulatory network controlling sporulation among Clostridia. This is especially the case for the impact of each sigma factor inactivation on the synthesis of the others and for the role of σK [21,23,73,85]. The timing of sigK expression, the phenotype of sigK mutants and/or some σG targets differ. In addition, with the exception of C. difficile, σG activity is controlled through processing in Clostridia while the insertion of a skin element in the sigK gene is found only in C. difficile strains [23]. The rather low probability to observe orthologs in clostridial genomes for many C. difficile regulon members also suggests a moderate conservation of the sporulation sigma-factor regulons among Clostridia. This work gives new insights about the diversity and evolution of the sporulation process. The sporulation in C. difficile might reflect a more ancestral way of sporulation while a more sophisticated system of developmental control would have been gradually introduced during evolution.

Materials and Methods

Bacterial strains, growth conditions and sporulation assay
C. difficile strains and plasmids used in this study are presented in Table 4. C. difficile strains were grown anaerobically (10% H2, 10% CO2, and 80% N2) in TY or in Brain Heart Infusion (BHI, Difco), which was used for conjugation. Sporulation medium (SM) [86] was used for sporulation assays. SM medium (pH at 7.4) contained per liter: 90 g Bacto tryptone, 5 g Bacto peptone, 1 g (NH4)2SO4,
1.5 g Tris Base. When necessary, cefoxitin (Cfx; 25 mg/ml), thiamphenicol (Tm; 15 mg/ml) and erythromycin (Erm; 2.5 mg/ml) were added to C. difficile cultures. E. coli strains were grown in Luria-Bertani (LB) broth. When indicated, ampicillin (100 mg/ml) or chloramphenicol (15 mg/ml) was added to the culture medium.

Sporulation assay were performed as follows. After 72 h of growth in SM medium, 1 ml of culture was divided into two samples. To determine the total number of cells, the first sample was serially diluted and plated on BHI with 0.1% taurocholate (Sigma-Aldrich). Taurocholate is required for the germination of C. difficile spores [4]. To determine the number of spores, the vegetative bacteria of the second sample were heat killed by incubation for 30 min at 65°C prior to plating on BHI with 0.1% taurocholate. The percentage of sporulation was determined as the ratio of the number of spores/ml and the total number of bacteria/ml (×100).

Construction of C. difficile strains

The ClosTron gene knockout system [87] was used to inactivate the spoIID (CD0126), spoVT (CD3499) and spoIR (CD3564) genes giving strains CDIP224 (CD0126::erm), CDIP227 (CD3499::erm) and CDIP238 (CD3564::erm). Primers to target the group II intron of pMTL007 to these genes (Table S12) were designed by the Targetron design software (http://www.sigmaaldrich.com). The PCR primer sets were used with the EBS universal primer and intron template DNA to generate by overlap extension PCR, a 533-bp product that would facilitate intron retargeting to CD0126, CD3499 or CD3564. The PCR products were cloned between the HindIII and BsrGI sites of pMTL007 to yield pDIA6123 (pMTL007::Cdi-CD0126-39s), pDIA6120 (pMTL007::Cdi-CD3499-157a) and pMS459 (pMTL007::Cdi-CD3564-38a). The presence and orientation of the insert in pMS459 was verified by DNA sequencing using the pMTL007-specific primers pMTL007-F and pMTL007-R. pDIA6120, pDIA6123 and pMS459 were introduced into E. coli HB101 (RP4) and the resulting strains subsequently mated with C. difficile 630Derm [88]. C. difficile transconjugants were selected by sub-culturing on BHI agar containing Tm and Cfx and then plated on BHI agar containing Erm. Chromosomal DNA of transconjugants was isolated as previously described. PCR using the ErmRAM primers (ErmF and ErmR) confirmed that the Erm resistant phenotype was due to the splicing of the group I intron following integration. In order to verify the integration of the Ll.LtrB intron into the right gene targets, we performed PCR with two primers flanking the insertion site in CD0126 (LS184-LS185), CD3499 (LS186-LS187) or CD3564 (IMV649-LS113) and in one hand with the intron primer EBSu and in other hand with a primer in CD0126 (LS184), CD3499 (LS187) or CD3564 (LS113).

To complement the spoIR mutant, the spoIR gene with its promoter (~169 to +791 from the translational start site), was

Figure 6. Model of the regulatory network controlling sporulation in C. difficile. The upper part of the figure represents the cell before asymmetric division. The parallel vertical lines represent the two membranes separating the forespore (right) and the mother cell (left). The four sigma factors of sporulation are encircled (by oval boxes). Precursor protein of σE is indicated as pro-σE. Square boxes correspond to transcriptional regulators. Proteins associated with the membrane or located into the intermembrane space are illustrated as embedded in the parallel vertical lines. Black solid arrows and white arrow indicate activation at the transcriptional level or at the level of protein activity, respectively. Broken arrows with a question mark represent mechanisms that are not yet fully understood. Dotted arrows indicate a possible direct transcriptional activation or repression. The σE target genes under the negative control of SpoIID are inside a pink box. doi:10.1371/journal.pgen.1003756.g006
amplified by PCR using oligonucleotides IMV642 and IMV641 (Table S12). To complement the spoIHD and the spoVT mutants, the spoIHD gene with its promoter (−117 to +520 from the translational start site) and the spoVT gene with its promoter (−165 to +731 from the translational start site), were amplified by PCR using oligonucleotides LS283 and IMV641 to produce plasmids pDIA6132 and pDIA6133 (Table S12). The PCR fragments were cloned into the pDIA6133, which was transferred by conjugation into the C. difficile spoIHD mutant giving strain CDIP246 (Table 4). Similarly the plasmids pDIA6132 and pDIA6133 were transferred by conjugation into the spoVT or spoIHD mutant giving strains CDIP263 and CDIP262.

To construct a PspoIR-SNAP fusion reporter strain, a 470 bp fragment encompassing the region upstream of the spoIIR gene was amplified by PCR using primers PspoIR-SNAP EcoRI F and PspoIR-SNAP XhoI Rev (Table S12). The PCR product was inserted between the EcoRI and XhoI sites of pFT47 [19] to produce plasmids pDIA6132 (spoVT), pDIA6133 (spoIHD) and pDIA6135 (spoIIR). Using the E. coli HB101 (RP4) strain containing pDIA6133 as donor, this plasmid was transferred by conjugation into the C. difficile 630Δerm, spoIHD::erm, sigF::erm or

| Strains | Genotype | Origin |
|---------|----------|--------|
| E. coli | F− mcrA D(mrr-hsdRMS-mcrBC) f80 lacZΔM15 ΔlacY74 deoR, recA1 araD139 D(ara-leu)7697 gaiK ρpsI(StrR) endA1 nupG | Invitrogen |
| BLD1 (DE3) | F− ampR gal dcm kan hsdS(B− m−) λDE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5] | New England Biolabs |
| HB101 (RP4) | supF4 αα14 galK2 lacY1 Δ (gpt-proA 62 rpsL20 StrR) xyl-5 mtl-1 recA13 Δ (mcrC-mrr) hsdS(B− m−) RP4 | laboratory stock |

| Strains | Genotype | Origin |
|---------|----------|--------|
| C. difficile | 630Δerm | [88] |
| CDIP3 | 630Δerm spo0A::erm | [27] |
| CDIP224 | 630Δerm spo0I::erm | This work |
| CDIP227 | 630Δerm spo0I::erm | This work |
| CDIP238 | 630Δerm spo0I::erm | This work |
| CDIP246 | 630Δerm spo0I::erm pMTL84121-spoIIR | This work |
| CDIP262 | 630Δerm spo0I::erm pMTL84121-spoIIR | This work |
| CDIP263 | 630Δerm spo0I::erm pMTL84121-spoIIR | This work |
| AHD533 | 630Δerm sigI::erm | [19] |
| AHD534 | 630Δerm sigI::erm | [19] |
| AHD535 | 630Δerm sigI::erm | [19] |
| AHD548 | 630Δerm sigI::erm pMTL84121-sigI | [19] |
| AHD549 | 630Δerm sigI::erm pMTL84121-sigI | [19] |
| AHD550 | 630Δerm sigI::erm pMTL84121-sigI | [19] |
| AHD551 | 630Δerm sigI::erm pMTL84121-sigI | [19] |
| AHD606 | 630Δerm PapoIR-SNAP<sup>CD</sup> | This work |
| AHD622 | 630Δerm spo0I::erm PapoIR-SNAP<sup>CD</sup> | This work |
| AHD623 | 630Δerm spo0I::erm PapoIR-SNAP<sup>CD</sup> | This work |
| AHD624 | 630Δerm spo0I::erm PapoIR-SNAP<sup>CD</sup> | This work |

| Plasmids | pET28a-sigF-his<sub>6</sub> | This work |
|----------|---------------------------|--------|
| pET35 | pET28a-sigE-his<sub>6</sub> | This work |
| pFT47 | pMTL84121-SNAP<sup>CD</sup> | [19] |
| pMTL84121-SNAP<sup>CD</sup> | pFT47-PapoIR-SNAP<sup>CD</sup> | This work |
| pMS459 | pMTL8007:Cdi-DD3564-38a | This work |
| pMS462 | pMTL8007:Cdi-DD3564-38a | This work |
| pMS462 | pMTL8007:Cdi-DD3564-38a | This work |
| pDIA6123 | pMTL8007:Cdi-DD3564-38a | This work |
| pDIA6124 | pMTL8007:Cdi-DD3564-38a | This work |
| pDIA6132 | pMTL8007:Cdi-DD3564-38a | This work |
| pDIA6133 | pMTL8007:Cdi-DD3564-38a | This work |
| pDIA6135 | pMTL8007:Cdi-DD3564-38a | This work |

Table 4. Strains and plasmids used in this study.
**sigE::erm** strains by conjugation. Transconjugants were selected on BHI agar plates containing Tm and Cfx followed by plating on BHI agar containing Tn.

RNA extraction, quantitative real-time PCR and transcriptional start site mapping

For RNA-seq experiment allowing TSS mapping, total RNA was isolated from *C. difficile* 630 strain grown in TY medium after 4 h and 10 h of growth or under starvation conditions that correspond to a 1 h resuspension of exponentially growing cells (6 h of growth) into phosphate-buffered saline (PBS: 137 mM NaCl, 10 mM Phosphate, 2.7 mM KCl, pH 7.4) [31]. The Tobacco Acid Pyrophosphatase (TAP)+/− library construction and high-throughput sequencing was realized on a mixed sample combining RNAs extracted from these three different growth conditions as previously described [31]. After Illumina sequencing, the reads were mapped to the *C. difficile* genome using Bowtie [90] then converted into BAM files with the Samtools [91]. The data were visualized at a strand-specific manner using COV2HTML (http://mmnonot.eu/COV2HTML/). All TSS detected by RNA-seq were inspected manually. Deep sequencing data are available at https://mmnonot.eu/COV2HTML/visualisation.php?str_id=-14.

To perform transcriptional analysis for each sigma factor, we first tested the impact of their inactivation on the expression of two target genes at different times to optimize conditions. For this purpose, we used bona fide targets of these sigma factors in *B. subtilis* [16]: *gpr* and *spsIR* for σ^F_+, *spolIIA* and *spolIVA* for σ^F_- and *spolB* for σ^N_. For σ^E_, we used cotCD and cok encoding recently described *C. difficile* spore coat proteins [41]. This preliminary test allowed us to define a time for maximal differential expression for these targets between wild-type and mutant strains for each sigma factor. To study σ^F_+ or σ^F_- dependent control, we harvested cells of strain 630Δerm, the *sigE* and the *sigF* mutants after 14 h of growth in SM medium. The strain 630Δerm and the *sigG* or the *sigF* mutant were harvested at 19 h (630Δerm, *sigG* mutant) and 24 h (630Δerm, *sigF* mutant) of growth in SM medium. For the *spsIR* and *spsIIA* mutants, we harvested cells after 14 h and 15 h of growth in SM medium, respectively. The culture pellets were resuspended in RNApro solution (MP Biomedicals) and RNA extracted using the FastRNA Pro Blue Kit, according to the manufacturer’s instructions. The RNA quality was determined using RNA 6000 Nano Reagenis (Agilent). For the *spsI* mutant and the 630Δerm strain, RNA previously obtained were used to test the impact of SpooA inactivation on gene expression [27].

Quantitative real-time PCR (qRT-PCR) analysis was performed as previously described [27]. The primers used for each marker are listed in Table S12. In each sample, the quantity of cDNAs of a gene was normalized to the quantity of cDNAs of the DNApolIII gene. The relative change in gene expression was recorded as the ratio of normalized target concentrations (ΔΔCt) [92].

**Microarray design, DNA-array hybridization and transcriptome analysis**

The microarray of *C. difficile* 630 genome (GEO database accession number GPL10556) was designed as previously described [27]. Transcriptome was performed using for each condition four (630Δerm compared to the mutant inactivated for each sigma factor) or two different RNA preparations (630Δerm compared to *spolIIA*). Labeled DNA hybridization to microarrays and array scanning were done as previously described [27]. The slides were analyzed using R and limma software (Linear Model for Microarray Data) from Bioconductor (www.bioconductor.org).

We corrected background with the ‘normexp’ method, resulting in strictly positive values and reducing variability in the log ratios for genes with low levels of hybridization signal. Then, we normalized each slide with the ‘loess’ method [93]. To test for differential expression, we used the bayesian adjusted t-statistics and we performed a multiple testing correction of Benjamini and Hochberg based on the false discovery rate (FDR) [94]. A gene was considered as differentially expressed when the p-value is <0.05. The complete data set was deposited in the GEO database with a series record accession number GSE43202.

**In silico analysis of promoters**

We analyzed the TSS data by the following in silico iterative strategy. For σ^F_+ and σ^N_+ we first worked with a reduced training set corresponding mainly to promoters located upstream of genes having an orthologous gene controlled by the same sigma factor in *B. subtilis* [16], or genes encoding experimentally characterized spore coat proteins for σ^E_+ [41] or genes encoding proteins associated with the spore (Table S8). Half-sites (boxes) were then manually re-aligned to maximize similarity and to construct recognition profiles for each sigma factor. The score of a candidate promoter was defined as the sum of positional nucleotide weights W(b,j), with an additional term V(k) to account for preferred interbox spacer length: S(b_1…b_mX_j…b_n-bar+b_TOTAL) = σ_1+m_m+n*n_m+n T b_{i=1}^n V(k), where b_1…b_m b_{n-bar} and n are nucleotides in the promoter boxes, n is the count of nucleotide, m and n are the distal and proximal box lengths, respectively. The positional nucleotide weights were defined as W(b,j)=log (N(b,j)+0.5)−0.25σ_{A=G=C=T} log (N(b,j)+0.5), where N(b,j) is the count of nucleotide b at alignment position j, and box length weight was defined as V(k)=[log (M(k)+0.5)−1/(k_{max}−k_{min}+1)]σ_{k}=k_{min}−k_{max} log (M(j)+0.5), where k_{max} and k_{min} are the maximal and minimal observed spacer length, respectively, and M(k) is the count of spacer length k. The necessary procedures were implemented in SignalX [95]. These profiles determined with the training set were used to score all identified TSS. Candidate promoters scoring higher than the lowest-scoring promoter in the training set were retained for further analysis (Table S8). A powerful approach, as for transcription factors, is to apply comparative techniques based on the assumptions that functional regulatory sites should be conserved in related species. Genome sequences from *C. difficile* six closest relatives from the Clostridium genus (*C. saccharolyticum, C. thermocellum, C. tetani, C. cellulolyticum, C. ljungdahlii, C. klyveri*) and *B. subtilis* 168 were downloaded from GenBank. We searched the intergenic regions of the six Clostridia using the profiles defined for *C. difficile* (Table S8). As the exact TSS positions in these species are unknown, we analyzed regions (−100 +10) relative to annotated start codons. The data on *B. subtilis* promoters were obtained from DBTBS [96]. Finally, for each candidate promoter, we calculated the number of genomes also having a candidate promoter upstream of an orthologous gene. Orthologs were defined using the bidirectional best hit criterion implemented in GenomeExplorer [95].

**Purification of σ^F-His_6 and σ^E-His_6 for antibody production**

The coding sequences of *sigF* and *sigE* were amplified by PCR using primer pairs CDSigF-pET28aFw/CDSigF-pET28aRev and CDSigE-pET28aFw/CDSigE-pET28aRe, respectively. The resulting fragments were introduced between the Ncol and Xhol sites of pET28a to produce pFT33 (σ^F-His_6) and pFT34 (σ^E-His_6). BL21(DE3) derivatives carrying pFT34 or pFT35 were grown in
LB to an OD$_{600}$ of 0.5 and induced with 1 mM IPTG for 4 h. The cells were then harvested by centrifugation at 4000 g, resuspended in 20 mM phosphate, 1 mM phenylmethyl-sulfonyl fluoride (PMSF), 10 mM imidazole, and lysed using a French pressure cell (18000 lb/in$^2$). After centrifugation, the supernatant (or, for the case of $\sigma^B$, the sediment after solubilization with 8M Urea for 30 min) was loaded onto a 1 ml Histrap column (Amersham Pharmacia Biotech). The bound proteins were eluted with a discontinuous imidazole gradient. The purified proteins were used for the production of rabbit polyclonal antibodies (Davids Biotechnologie GmbH).

**Whole cell lysates and immunoblot analysis**

To prepare *C. difficile* whole cell extracts, 10 ml samples were withdrawn from SM cultures at the desired times following inoculation and the cells collected by centrifugation (4000 g, 10 min, at 4°C). Cells were lysed in 1 ml buffer (10 mM Tris pH 8.0, 10 mM MgCl$_2$, 0.5 mM EDTA, 0.2 mM NaCl, 10% glycerol, 1 mM PMSF) using a French pressure cell. Samples (15 μg of protein) were resolved by 12% SDS-PAGE, transferred to a nitrocellulose membrane (BioRad), and subjected to immunoblot analysis as described previously [97]. Antibodies against $\sigma^E$ and $\sigma^B$ were used at a 1:2000 dilution. A rabbit secondary antibody conjugated to horseradish peroxidase (Sigma) was used at a 1:5000 dilution. The immunoblots were developed with enhanced chemiluminescence reagents (Amersham Pharmacia Biotech).

**Microscopy and image analysis**

1 ml samples were withdrawn from SM cultures at the desired times following inoculation, and the cells collected by centrifugation (4000 g, 10 min, at 4°C). The cells were washed with 1 ml of PBS and resuspended in 0.1 ml of PBS supplemented with the membrane dye FM1-43 (10 μg·ml$^{-1}$) and the DNA stain DAPI (4′,6-diamidino-2-phenylindole; 50 μg·ml$^{-1}$) (Molecular Probes, Invitrogen). For SNAP staining, 1 ml samples were stained for 30 min with 50 nM SNAP-Cell TMR-Star (New England Biolabs) as described [19]. Cells were washed four times by centrifugation (4000 g, 5 min) and resuspended in 1 ml of PBS. Following washing, the cells were resuspended in 1 ml of PBS supplemented with the membrane dye Mitotracker Green (0.5 μg·ml$^{-1}$) (Molecular Probes, Invitrogen). Cells were mounted on 1.7% agarose coated glass slides and imaged in a Leica DM6000B microscope as previously described [64]. Fluorescent signals were visualized with a phase contrast Uplan F1 100× objective and captured with a CCD Andor Ixon EM$^{-}$ camera (Andor Technologies). Images were acquired and analyzed using the Metamorph software suite version 5.8 (Universal Imaging).

**Supporting Information**

**Figure S1** Alignment of the $\sigma^B$ and $\sigma^E$ proteins of *B. subtilis* and *C. difficile*. The amino acid sequences from $\sigma^B$ and $\sigma^E$ of *B. subtilis* and *C. difficile* are aligned. The amino acids conserved in these four sequences are indicated by a star. The region 4.2, which may interact with the $\sim$35 regions of their cognate promoters is underlined. In *B. subtilis*, the specificity of interaction of these sigma factors with the $\sim$35 region sequences (a T for $\sigma^E$ and a C for $\sigma^B$) is associated with the presence of a glutamine at position 217 of $\sigma^B$ and of an arginine in $\sigma^E$ [32]. These amino acids are indicated in red and blue, respectively.

**Figure S2** Inactivation of the *spoIIIR*, *spoIID* and *spoVT* genes in *C. difficile* using the Clostron system. A: Schematic representation of gene inactivation by a type II Intrion with an associated RAM. The group II intron (brace), originally in pMT1007 (top), carries a RAM element (yellow) interrupting an ermB determinant (blue). The intron was targeted to the sig gene of interest (black; middle) by altering the IBS, EBS1 and EBS2 sequences (grey and white stripes; top) by overlapping PCR. Splicing out of the td group I intron from the ermB gene in the RAM restores a functional marker allowing positive selection of mutants following intron integration. Primers used to confirm the integration and orientation of the type II intron are also indicated (bottom). Genetic organisation of the *C. difficile* chromosome in the vicinity of *spoIR*, *spoIID* and *spoVT*. The red arrow indicates the point of insertion of the re-targeted type II introns used for gene disruption. The extent of the DNA fragment present in the indicated replicative plasmids used for in trans complementation of the insertional mutations is shown below each of the genetic maps, except for the *sigK* gene (see also Fig. 6). B: Chromosomal DNA of *E. coli* *C. difficile* conjugants and of strain 630Δerm were screened by PCR using primer pairs RAM-F/R to confirm splicing out of the group I intron in the mutant (lane 1 and 2). To verify the integration of the LLRIR intron into the right gene targets, we further performed PCR using chromosomal DNA of strain 630Δerm and of each mutant (lane 3 and 4) with two primers flanking the insertion site in *CD0126-spoIID* (LS184-LS185), *CD3499-spoVT* (LS186-LS187) or *CD3564-spoIR* (MV649-LS113). Finally, we also performed PCR using chromosomal DNA of strain 630Δerm and of each mutant (lane 5 and 6) with the intron primer EBSu in one hand and with a primer in *CD0126-spoIID* (LS184) in *CD3499-spoVT* (LS187) or *CD3564-spoIR* (LS113) in other hand. Chromosomal DNA from the 630Δerm strain corresponded to lane 1, 3 and 5 while chromosomal DNA of each mutant corresponded to lane 2, 4 and 6. The smart ladder (Eurogentec) was used as a molecular weight marker.

C: Southern blot analysis of genomic DNA from *C. difficile* 630Δerm, *spoIR*, *spoIID*, and *spoVT* mutant strains with an intron probe. Chromosomal DNA (6 μg in each reaction) was digested with *HindIII*. Southern blot analyses were performed using Amersham ECL Direct Nucleic Acid labelling and detection reagents, in accordance with the manufacturer’s guidelines and visualised using Super Signal West Femto Maximum Sensitivity Substrate (Thermo Scientific). The probe was produced by PCR using OBD522 and OBD523 primers (Table S12), designed within the group II intron sequence.

**Figure S3** Alignment of the SpoIIID regulator of several Bacilli and Clostridia. The amino acid sequences from SpoIIID of *B. subtilis* (bs), *B. cereus* (bc), *B. anthracis* (ba), *C. acetobutylicum* (ca), *C. perfringens* (cp) and *C. difficile* (cd) are aligned. The amino acids conserved in these four sequences are indicated by a star. The two regions essential for DNA binding are indicated: an helix-turn-helix motif (HTH) and a basic region near the C-terminus part of the protein [68].

**Figure S4** Control of $\sigma^E$ targets by $\sigma^B$ and $\sigma^E$. Total RNAs were extracted from *C. difficile* 630Δerm strain, the *sigF* mutant and the *sigE* mutant grown in SM medium for 14 h. After reverse transcription, specific cDNAs were quantified by qRT-PCR using DNA PolIII gene for normalization (See Materials and Methods). The expression ratio of strain 630Δerm/*sigF* and 630Δerm/*sigE* were indicated in white and black, respectively. Error bars correspond to standard deviation from at least two biological replicates.
Table S1  List of genes controlled by $\sigma^F$ in transcriptome. A gene is considered differentially expressed between the 630Δerm strain and the sigF mutant when the P value is $<$0.05 using the statistical analysis described in Materials and Methods. We did not include genes with a fold-change $<$2-fold. However, some genes with a fold-change less than 2-fold were included when they appeared to be in the same transcription unit with regulated genes for which the fold-change was $\geq$2 or when they were regulated by other sigma factors of sporulation.

Table S2  List of genes controlled by $\sigma^E$ in transcriptome. A gene is considered differentially expressed between the 630Δerm strain and the sigE mutant when the P value is $<$0.05 using the statistical analysis described in Materials and Methods. We did not include genes with a fold-change $<$2-fold. However, some genes with a fold-change less than 2-fold were included when they appeared to be in the same transcription unit with regulated genes for which the fold-change was $\geq$2 or when they were regulated by other sigma factors of sporulation.

Table S3  List of genes controlled by $\sigma^G$ in transcriptome. A gene is considered differentially expressed between the 630Δerm strain and the sigG mutant when the P value is $<$0.05 using the statistical analysis described in Materials and Methods.

Table S4  List of genes controlled by $\sigma^K$ in transcriptome. A gene is considered differentially expressed between the 630Δerm strain and the sigK mutant when the P value is $<$0.05 using the statistical analysis described in Materials and Methods. We did not insert in this list genes with a fold-change $<$2-fold. However, some genes had a fold-change less than 2-fold but were included because they appeared to be in the same transcription unit with regulated genes for which the fold-change was $\geq$2 or because they were regulated by other sigma factors of sporulation.

Table S5  Validation of microarrays data by qRT-PCR on selected genes. qRT-PCR experiments were performed on two different RNA preparations for each mutant. The results presented corresponded to the mean of at least two independent experiments.

Table S6  Identification of promoters recognized by $\sigma^F$ and $\sigma^G$. The transcriptional start sites and the $-10$ and $-35$ boxes are indicated in red and blue, respectively. The position of $5'$ start was identified by RNA-seq analysis with indicated score corresponding to the read length (51 bases) coverage ratio for TAP+ and TAP- samples. For $\sigma^F$ and $\sigma^G$, the genes underlined are those validated in silico analysis (see Materials and Methods) and listed in Table S8. TSS can be visualized at https://mmnonot.eu/COV2HTML/visualisation.php?str_id=-14.

Table S7  Identification of promoters recognized by $\sigma^E$ and $\sigma^K$. The transcriptional start sites and the $-10$ and $-35$ boxes are indicated in red and blue, respectively. The position of $5'$ start was identified by RNA-seq analysis with indicated score corresponding to the read length (51 bases) coverage ratio for TAP+ and TAP- sample. For $\sigma^E$ and $\sigma^K$, the genes underlined are those validated in silico analysis (see Materials and Methods) and listed in Table S8. TSS can be visualized at https://mmnonot.eu/COV2HTML/visualisation.php?str_id=-14.

Table S8  In silico validation of promoters located upstream of genes identified in transcriptome as regulated by $\sigma^F$, $\sigma^E$, $\sigma^G$ or $\sigma^K$. We analyzed the promoters identified by TSS mapping by an iterative in silico strategy (see Materials and Methods). The training sets of genes used to first construct the recognition profiles for $\sigma^F$, $\sigma^E$, $\sigma^G$ and $\sigma^K$ were highlighted in yellow. For each promoter the score was obtained as defined in materials and methods. Using the same profiles, we searched the intergenic regions (positions $-$100 to +10 relative to start codons) of six closely related Clostridium species, C. saccharolyticum, C. thermoaceticum, C. tetani, C. cellulolyticum, C. ljungdahlii, C. kluyveri. We also analyzed the B. subtilis promoters using the DBTBS database [96]. The absence of an orthologous gene is indicated by 0. When an orthologous gene is present the score and the locus-tag are indicated.

Table S9  Complementation of the mutants inactivated for the sigma factors of sporulation. qRT-PCR experiments were performed on two different RNA preparations for each mutant and each complemented strain. Cells were harvested after 14 h of growth for the strain 630Δerm, the sigE and sigF mutants and the sigK mutant containing pMTL84121-sigF and the sigE mutant containing pMTL84121-sigE, after 20 h of growth for the strain 630Δerm, the sigG mutant and the sigG mutant containing pMTL84121-sigG and after 24 h of growth for the strain 630Δerm, the sigK mutant and the sigK mutant containing pMTL84121-sigKmin. The results presented correspond to the mean of at least two independent experiments.

Table S10  List of genes controlled by SpoIIID in transcriptome. RNA was extracted from strain 630Δerm strain and the spoIID mutant after 15 h of growth in SM medium. A gene is considered differentially expressed between the 630Δerm strain and the spoIID mutant when the P value is $<$0.05 using the statistical analysis described in Materials and Methods.

Table S11  Control of expression of sporulation genes by SpoIIID, SpoVT or SpoIIR. Cells of the 630Δerm strain and of the spoIID mutant were harvested after 15 h or 24 h of growth in SM medium. Cells of the 630Δerm strain and of the spoVT mutant were harvested after 15 h or 20 h of growth in SM medium. Cells of the 630Δerm strain and of the spoIIR mutant were harvested after 14 h of growth in SM medium. qRT-PCR experiments were performed on two different RNA preparations. The results presented corresponded to the mean of at least two independent experiments. NR = not regulated.

Table S12  List of oligonucleotides.

Acknowledgments

We thank Rita Tomé for help with the SNAP labeling experiments. We thank Emilie Camiade for helpful discussions and Isabelle Poquet for her comments on the manuscript. We are grateful to Patrick Stragier for a critical reading of the manuscript and helpful advice.

Author Contributions

Conceived and designed the experiments: IMV AOH BD MSG OS. Performed the experiments: LS MS FCP OS. Analyzed the data: LS FCP AOH BD OS MM PVS MSG IMV. Contributed reagents/materials/analysis tools: LS FCP MS OS MM PVS MSG BD AOH IMV. Wrote the paper: LS MSG BD AOH IMV.
59. Nugroho FA, Yamamoto H, Kobayashi Y, Sekiguchi J (1999) Characterization of a new sigma-K-dependent peptidoglycan hydrolase gene that plays a role in Bacillus subtilis mother cell lysis. J Bacteriol 181: 6230–6237.

60. Burns DA, Heap JT, Minton NP (2010) ScE is essential for germination of Clostridium difficile spores in nutrient-rich medium supplemented with the bile salt taurocholate. J Bacteriol 192: 657–664.

61. Gerdes K, Christensen SK, Lohner-Olesen A (2005) Prokaryotic toxin-antitoxin stress response loci. Nat Rev Microbiol 3: 371–382.

62. Harris LM, Welker NE, Papoutsakis ET (2002) Northern, morphological, and fermentation analysis of spo04 inactivation and overexpression in Clostridium acetobutylicum ATCC 824. J Bacteriol 184: 3586–3597.

63. Rhayat L, Duperrier S, Carballido-Lopez R, Pellegrini O, Stragier P (2009) Direct and indirect control of late sporulation genes by GerR of Bacillus subtilis. J Bacteriol 191: 3406–3413.

64. Serrano M, Hobot J, Cutting S (1996) A compartmentalized regulator of developmental gene expression in Bacillus subtilis. J Bacteriol 178: 4500–4507.

65. Ramirez-Peralta SJ, Stewart KA, Thomas SK, Setlow B, Chen Z, et al. (2012) Effects of the SpoVTR regulatory protein on the germination and germination protein levels of spores of Bacillus subtilis. J Bacteriol 194: 3417–3425.

66. Cangiano G, Mazzone A, Bacciagali L, Istrati R, Eichenberger P, et al. (2010) Cloning and characterization of a Bacillus subtilis transcription factor SpolIID allows a monomer to bind DNA. J Bacteriol 192: 1596–1606.

67. Kunkel B, Losick R, Stragier P (1990) The Bacillus subtilis gene for the development transcription factor sigma K is generated by excision of a dispensable DNA element containing a sporulation recombinase gene. Genes Dev 4: 525–535.

68. Errington J (1995) Bacillus subtilis sporulation: regulation of gene expression and control of morphogenesis. Microbiol Rev 57: 1–33.

69. Yoshii H, Ibara K, Nishimoto T, Sakai H, Komano T (1995) Cloning and characterization of a Bacillus thuringiensis homolog of the spoIID gene from Bacillus subtilis. Gene 15: 23–29.

70. Oke V, Losick R (1995) Multilevel regulation of the sporulation transcription factor sigma K in Bacillus subtilis. J Bacteriol 175: 7341–7347.

71. Kirk DG, Dahlsten E, Zhang Z, Korkeyla H, Lindstrom M (2012) Involvement of Clostridium beijerinckii ATCC 3502 sigma factor K in early-stage sporulation. Appl Environ Microbiol 78: 4590–4596.

72. Favaret P, Eichenberger P, Losick R, Youngman P (2000) The transcriptional profile of early to middle sporulation in Bacillus subtilis. Proc Natl Acad Sci U S A 97: 8063–8068.

73. Jonas RM, Haldenwang WG (1989) Influence of spo mutations on sigma E synthesis in Bacillus subtilis. J Bacteriol 171: 5209–5220.

74. Karow ML, Glaser P, Piggot PJ (1995) Identification of a gene, spoID, that links the activation of sigma E to the transcriptional activity of sigma F during sporulation in Bacillus subtilis. Proc Natl Acad Sci U S A 92: 2012–2016.

75. Eilar A, Chazy VK, Xenopoulos P, Foutes ME, Loson OC, et al. (2009) Partial penetration facilitates developmental evolution in bacteria. Nature 460: 510–514.

76. Zhang L, Higgins ML, Piggot PJ, Karow ML (1996) Analysis of the role of prespore gene expression in the compartmentalization of mother cell-specific gene expression during sporulation of Bacillus subtilis. J Bacteriol 178: 2813–2817.

77. Guillot C, Moran CP, Jr. (2007) Essential internal promoter in the spoFH locus of Bacillus subtilis. J Bacteriol 189: 7101–7109.

78. Cutting S, Oke V, Driks A, Losick R, Lu S, et al. (1990) A forespore checkpoint for mother cell gene expression during development in B. subtilis. Cell 62: 239–250.

79. Traag B, Puglise A, Eisen J, Losick R (2013) Gene Conservation among Endospore-Forming Bacteria Reveals Additional Sporulation Genes in Bacillus subtilis. J Bacteriol 195: 253–260.

80. Dworkin J, Losick R (2005) Developmental commitment in a bacterium. Cell 121: 401–409.

81. Jiang X, Rubio A, Chiba S, Podgano K (2005) Enflagellated-regulated proteolysis of SpoIQ: evidence that dual checkpoints control sigma activity. Mol Microbiol 58: 102–115.

82. Oke V, Schepetov M, Cutting S (1997) SpoIVB has two distinct functions during spore formation in Bacillus subtilis. Mol Microbiol 23: 223–230.

83. Wang Y, Li X, Mao Y, Blaschek HP (2012) Genome-wide dynamic transcriptional profiling in Clostridium beijerinckii NCIMB 8052 using single-nucleotide resolution RNA-Seq. BMC Genomics 13: 102.

84. Wilson KH, Kennedy MJ, Fekety FR (1982) Use of sodium taurocholate to enhance spore recovery on a medium selective for Clostridium difficile. J Clin Microbiol 15: 443–446.

85. Heap JT, Pennington OJ, Cartman ST, Carter GP, Minton NP (2007) The Clostrion: a universal gene knock-out system for the genus Clostridium. J Microbiol Methods 70: 452–464.

86. Hussain HA, Roberts AP, Mullany P (2005) Generation of an erythromycin-sensitive derivative of Clostridium difficile strain 630 (630Dehalo) and demonstration that the conjugal transposon Tn1661DelE enters the genome of this strain at multiple sites. J Med Microbiol 54: 137–141.

87. Wang Y, Li X, Mao Y, Blaschek HP (2012) Genome-wide dynamic transcriptional profiling in Clostridium beijerinckii ATCC 8052 using single-nucleotide resolution RNA-Seq. BMC Genomics 13: 102.

88. Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10: R25.

89. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, et al. (2009) The Sequence Alignment/Map format and SAMtools. Bioinformatics 25: 2078–2079.

90. Navak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2ΔΔC(T) Method. Methods 25: 402–408.

91. Smyth GK, Speed T (2005) Normalization of cDNA microarray data. Methods 31: 263–273.

92. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J Roy Statist Soc Ser: 289–300.

93. Mironov A, Vinokurova N, Gelfand M (2000) GenomeExplorer: software for transcriptional regulation in Bacillus subtilis. J Bacteriol 182: 1483–1493.

94. Wang Y, Li X, Mao Y, Blaschek HP (2012) Genome-wide dynamic transcriptional profiling in Clostridium beijerinckii NCIMB 8052 using single-nucleotide resolution RNA-Seq. BMC Genomics 13: 102.

95. Heap JT, Pennington OJ, Cartman ST, Carter GP, Minton NP (2007) The Clostrion: a universal gene knock-out system for the genus Clostridium. J Microbiol Methods 70: 452–464.

96. Hussain HA, Roberts AP, Mullany P (2005) Generation of an erythromycin-