

RESEARCH ARTICLE

# Impact of CodY protein on metabolism, sporulation and virulence in *Clostridioides difficile* ribotype 027

Nadine Daou<sup>1#a</sup>, Yuanguo Wang<sup>2#b</sup>, Vladimir M. Levдикov<sup>3</sup>, Madhumitha Nandakumar<sup>4#c</sup>, Jonathan Livny<sup>5</sup>, Laurent Bouillaut<sup>1#d</sup>, Elena Blagova<sup>3</sup>, Keshan Zhang<sup>2#e</sup>, Boris R. Belitsky<sup>1</sup>, Kyu Rhee<sup>4</sup>, Anthony J. Wilkinson<sup>3</sup>, Xingmin Sun<sup>6</sup>, Abraham L. Sonenshein<sup>1\*</sup>

**1** Department of Molecular Biology and Microbiology, Tufts University School of Medicine, Boston, MA, United States of America, **2** Department of Infectious Disease and Global Health, Cummings School of Veterinary Medicine, Tufts University, North Grafton, MA, United States of America, **3** Structural Biology Laboratory, Department of Chemistry, University of York, York, United Kingdom, **4** Department of Medicine, Division of Infectious Diseases, Weill Cornell Medical College, New York, NY, United States of America, **5** Broad Institute of MIT and Harvard, Cambridge, MA, United States of America, **6** Department of Molecular Medicine, Morsani College of Medicine, University of South Florida, Tampa, FL, United States of America

#a Current address: Axcella Health, Inc., Cambridge, MA, United States of America

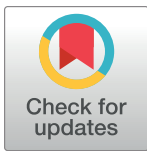
#b Current address: Department of Molecular Medicine, Morsani College of Medicine, University of South Florida, Tampa, FL, United States of America

#c Current address: Seres Therapeutics, Cambridge, MA, United States of America

#d Current address: ConsortiaTX, Inc., Woburn, MA, United States of America

#e Current address: State Key Laboratory of Veterinary Etiological Biology, National Foot and Mouth Disease Reference Laboratory, Lanzhou Veterinary Research Institute, Chinese Academy of Agricultural Sciences, Lanzhou, China

\* [linc.sonenshein@tufts.edu](mailto:linc.sonenshein@tufts.edu)



**OPEN ACCESS**

**Citation:** Daou N, Wang Y, Levдикov VM, Nandakumar M, Livny J, Bouillaut L, et al. (2019) Impact of CodY protein on metabolism, sporulation and virulence in *Clostridioides difficile* ribotype 027. PLoS ONE 14(1): e0206896. <https://doi.org/10.1371/journal.pone.0206896>

**Editor:** Michel R. Popoff, Institut Pasteur, FRANCE

**Received:** November 13, 2017

**Accepted:** September 25, 2018

**Published:** January 30, 2019

**Copyright:** © 2019 Daou et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files. In addition, the raw data for the RNA-seq experiments are available to the public through the NIH Sequence Read Archive (SRA) as BioProject PRJNA438155. The website is [https://www.ncbi.nlm.nih.gov/sra?linkname=bioproject\\_sra\\_all&from\\_uid=438155](https://www.ncbi.nlm.nih.gov/sra?linkname=bioproject_sra_all&from_uid=438155).

**Funding:** All research reported in this manuscript was funded by grants from the US National Institutes of Health. The funder provided support in the form of salaries for most of the authors and the

## Abstract

Toxin synthesis and endospore formation are two of the most critical factors that determine the outcome of infection by *Clostridioides difficile*. The two major toxins, TcdA and TcdB, are the principal factors causing damage to the host. Spores are the infectious form of *C. difficile*, permit survival of the bacterium during antibiotic treatment and are the predominant cell form that leads to recurrent infection. Toxin production and sporulation have their own specific mechanisms of regulation, but they share negative regulation by the global regulatory protein CodY. Determining the extent of such regulation and its detailed mechanism is important for understanding the linkage between two apparently independent biological phenomena and raises the possibility of creating new ways of limiting infection. The work described here shows that a *codY* null mutant of a hypervirulent (ribotype 027) strain is even more virulent than its parent in a mouse model of infection and that the mutant expresses most sporulation genes prematurely during exponential growth phase. Moreover, examining the expression patterns of mutants producing CodY proteins with different levels of residual activity revealed that expression of the toxin genes is dependent on total CodY inactivation, whereas most sporulation genes are turned on when CodY activity is only partially diminished. These results suggest that, in wild-type cells undergoing nutrient limitation, sporulation genes can be turned on before the toxin genes.

purchase of materials and supplies, but did not have any role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The specific roles of all the authors are articulated in the 'author contributions' section. Some of the authors are now employees of commercial organizations, but at the time that they contributed to the research described in this manuscript they were university employees. R01 GM042219, National Institute of General Medical Sciences (ALS, AJW, KR). K01 DK092352, National Institute of Diabetes and Digestive and Kidney Diseases (XS).

**Competing interests:** The authors have declared that no competing interests exist.

## Introduction

*Clostridioides difficile*, a spore-forming bacterial pathogen, is now the primary cause of antibiotic-associated diarrhea, an infection usually acquired during stays in healthcare facilities. At present, >500,000 patients in the USA are newly diagnosed with *C. difficile* infection (CDI) each year and 75,000–175,000 cases of recurrent CDI are seen, leading to an increase in healthcare costs of \$4.2 billion [1–3]. While current treatments cure almost 90% of primary infections, recurrent infection is so high that thousands of patients are on long-term antibiotic treatment and more than 25,000 CDI patients die each year [1, 4]. The ability of the bacterium to be so effective in causing disease depends on two critical aspects of its biology. Pathogenic *C. difficile* produces at least two potent toxins, TcdA and TcdB, that cause major disruptions of the mammalian intestinal tract [5]. In addition, the bacteria form spores in the GI tract, a process that makes a significant fraction of the infecting cells resistant to all antibiotics and provides a way for this oxygen-sensitive anaerobe to survive in the environment. Spore formation appears to be the primary factor responsible for the unusual frequency of recurrent *C. difficile* infection; the spores released from the GI tract are difficult to kill, survive in the environment and are able to cause a new round of infection in patients who have completed their antibiotic treatment. Discovering factors that influence toxin production or spore formation or both is critical to our understanding of infection and has the potential to lead to novel ways of preventing and treating *C. difficile* infection.

The genes that encode the toxins have no known impact on sporulation. Similarly, most sporulation proteins are not known to influence toxin synthesis directly. (There is some evidence that the sporulation regulator Spo0A contributes to toxin gene regulation, but the effect varies from strain to strain [6–8]). The toxin genes are transcribed at very high levels by RNA polymerase containing the toxin gene-specific sigma factor TcdR [9]. Surprisingly, a null mutation in *tcdR* causes a decrease in sporulation efficiency and the heat-resistance of the resulting spores [10], implying that one or more genes normally transcribed by TcdR-containing RNA polymerase encode proteins that affect spore formation. The conditions that lead to TcdR activity are complex. During exponential growth in rich medium, the *tcdR* gene is transcribed at low levels by  $\sigma^A$ -containing RNA polymerase, the primary form of RNA polymerase in growing cells. When cells reach stationary phase, the *tcdR* gene is much more highly expressed, first from the  $\sigma^A$ -dependent promoter and then, as TcdR accumulates, from a TcdR-dependent promoter [11, 12]. A third sigma factor,  $\sigma^D$  (the motility sigma factor), also contributes to *tcdR* gene expression [13]. When TcdR accumulates, it directs RNA polymerase to the promoters of the toxin genes, leading to high-level toxin gene expression [9].

Although they may or may not influence each other directly, both sporulation and toxin synthesis are turned on when cells experience nutrient limitation, e.g., when cells in laboratory culture in a complex medium reach stationary phase. At least two global regulatory proteins are known to contribute to this form of regulation. CcpA, the regulatory protein that is activated when rapidly metabolizable sugars are in excess, is a repressor of the *tcdR* and toxin genes and represses sporulation as well [14]. The global regulator, CodY, is also a strong repressor of the *tcdR* gene [15, 16], as well as an inhibitor of sporulation [17]. Nawrocki *et al.* [17] found that in the ribotype 027 strain UK1, a null mutation in the *codY* gene increases spore formation 1000-fold in cells growing in 70:30 medium. CodY is found in nearly all low G+C Gram-positive bacteria and is activated by binding of one of the branched-chain amino acids (BCAAs; isoleucine, leucine, valine) [18] and, in most species, GTP [19–23]. The responsiveness of CodY to BCAAs implies that the cell has evolved to assess accumulation of these amino acids as a major indicator of nutrient availability; a likely reason is that the BCAAs are not only used for protein synthesis but are also the precursors of branched-chain fatty acids,

the principal membrane fatty acids of most Gram-positive bacteria [24, 25]. The GTP level in the cell influences the ability to synthesize RNA, but is also a general indicator of amino acid availability; when the intracellular concentration of any amino acid drops below the level needed for its incorporation into protein, the GTP pool drops significantly due to the conversion of GTP to pppGpp via the stringent response [26, 27]. Thus, when certain nutrients become limiting, both toxin and sporulation genes can potentially be turned on.

The CodY protein represses transcription of hundreds of genes and activates transcription of dozens of other genes in many Gram-positive species [15, 16, 28–43]. The direct target genes tend to be involved in multiple aspects of nutrient uptake, metabolism and virulence; some target genes encode other regulators that are the direct activators or repressors of the apparent CodY targets. A previous microarray analysis of genes whose expression is altered by a *codY* null mutation in the *C. difficile* ribotype 012 strain JIR8074 (also known as 630E) identified 146 genes that are negatively regulated >4-fold by CodY and 19 genes that are positively regulated [15].

The diversity and complexity of the CodY regulon raises important issues about how and why the cell uses such a regulator to control multiple, seemingly unrelated processes. A key aspect of our approach was to subject a panel of CodY point mutants with different levels of residual activity to genome-wide transcription analysis by RNA-seq in order to identify genes that are regulated at different levels of CodY activity. Similar approaches have been used to analyze the CodY regulons in *Bacillus subtilis* [30] and *Staphylococcus aureus* [42]. This type of analysis provides a global picture of *C. difficile*'s strategy for altering virulence and sporulation properties in response to various levels of nutrient availability.

## Results

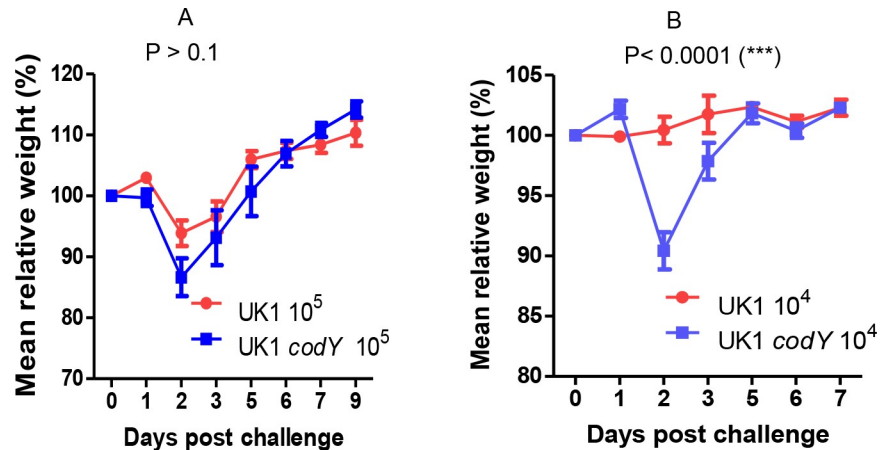
### Virulence of *codY* mutant strains

The documented role of CodY as a repressor of *C. difficile* toxin gene expression in laboratory cultures [15, 16] raised the possibility that a *codY* null mutant strain might overexpress the toxin genes during infection and therefore be hypervirulent. To address this possibility, we tested the virulence of strain LB-CD16, a *codY* null mutant [44] of strain UK1, a ribotype 027 isolate from the Stoke-Mandeville outbreak of 2003 [45]. (The mutation was created by insertion within *codY* of an intron containing an erythromycin-resistance determinant and is noted as *codY::intron::erm*.) In a mouse model of infection in which mice are fed a cocktail of antibiotics before infection [46], strain UK1 and the *codY* null mutant (LB-CD16) induced diarrhea and weight loss in all mice infected at a dose of  $10^5$  spores per mouse (Fig 1A and Table 1); 10% of the mice died after infection by UK1, 20% after infection by the *codY* null mutant. At a dose of  $10^4$  spores per mouse, strain UK1 was greatly reduced in virulence, but the *codY* null mutant caused diarrhea and loss of weight (without death) in nearly all infected mice (Fig 1B and Table 1).

### Toxin titers in parental and mutant strains

To test the relationship between the hypervirulence of strain LB-CD16 (UK1 *codY*) and toxin titer, culture fluids of laboratory cultures that had grown for 24 hrs at 37°C in an anaerobic chamber were tested for toxin titers by ELISA assays. As shown in Fig 2, the culture fluid of the *codY* null mutant contained considerably more TcdA (Toxin A) and TcdB (Toxin B) than did its parent strain. These results were consistent with the much higher level of expression of the toxin genes in laboratory cultures of strain UK1 *codY* (Fig 3).

To verify that toxin overproduction in strain LB-CD16 was a result of the *codY* null mutation, complementation was achieved by integrating a wild-type copy of *codY* contained within



**Fig 1. Infection outcomes induced by strain UK1 and its *codY* mutant derivative.** Groups of mice (n = 10) were infected with spores of strain UK1 and its *codY* null mutant (UK1 *codY*, also known as LB-CD16) and were monitored for weight change at doses of 10<sup>5</sup> spores per mouse (A) or 10<sup>4</sup> spores per mouse (B). Weight changes between the two groups were analyzed by 2-way ANOVA. The p values were calculated based on the differences in weight changes between the mice infected with the wild-type strain and the mice infected with the *codY* null mutant over the entire time course of the experiment (days 0–9 for Fig 1A and days 0–7 for Fig 1B). The impact of these infections on mouse diarrhea and viability is listed in Table 1.

<https://doi.org/10.1371/journal.pone.0206896.g001>

the transposon Tn916 at a non-homologous locus in the *C. difficile* chromosome. As shown in Fig 2, the complemented strain had near-parental levels of toxin proteins. In addition, the vast overexpression of the *tcdA* and *tcdB* genes in the UK1 *codY* mutant strain was greatly reduced in the complemented strain (Fig 3).

### Creation of a family of CodY point mutants with different levels of residual activity

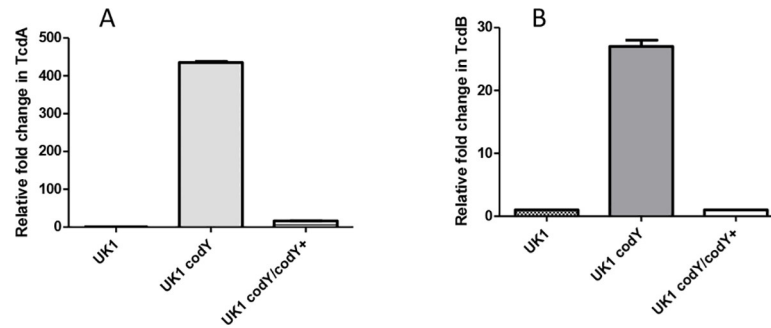
Whereas complete inactivation of CodY in strain UK1 led to about a 1000-fold increase in toxin gene expression at the onset of stationary phase (Fig 3), the level of CodY activity needed to repress the toxin genes has not been determined. To address this question and simultaneously determine the extent to which other CodY target genes are affected by partial inactivation of CodY in strain UK1, we created a family of *codY* point mutants with different levels of residual activity. We then determined the impact of these mutations on gene expression during mid-exponential phase in cells growing in a rich medium. Doing so prevented the potential contribution of changes in nutrient availability that would alter multiple other regulatory mechanisms that are likely to affect many of the CodY-regulated genes.

**Table 1. Diarrhea in mice infected with *codY*<sup>+</sup> and *codY* null mutant strains.**

Strain	Dose of spores per mouse	Mice with diarrhea (%)	Death of infected mice (%)
UK1	10 <sup>5</sup>	100	10
UK1 <i>codY</i>	10 <sup>5</sup>	100	20
UK1	10 <sup>4</sup>	20	0
UK1 <i>codY</i>	10 <sup>4</sup>	90	0

Groups of 10 mice pretreated with antibiotics and infected with the indicated doses of spores of *C. difficile* strains UK1 and UK1 *codY* (LB-CD16) were monitored for survival and the occurrence of diarrheal symptoms over the course of 7 days of infection.

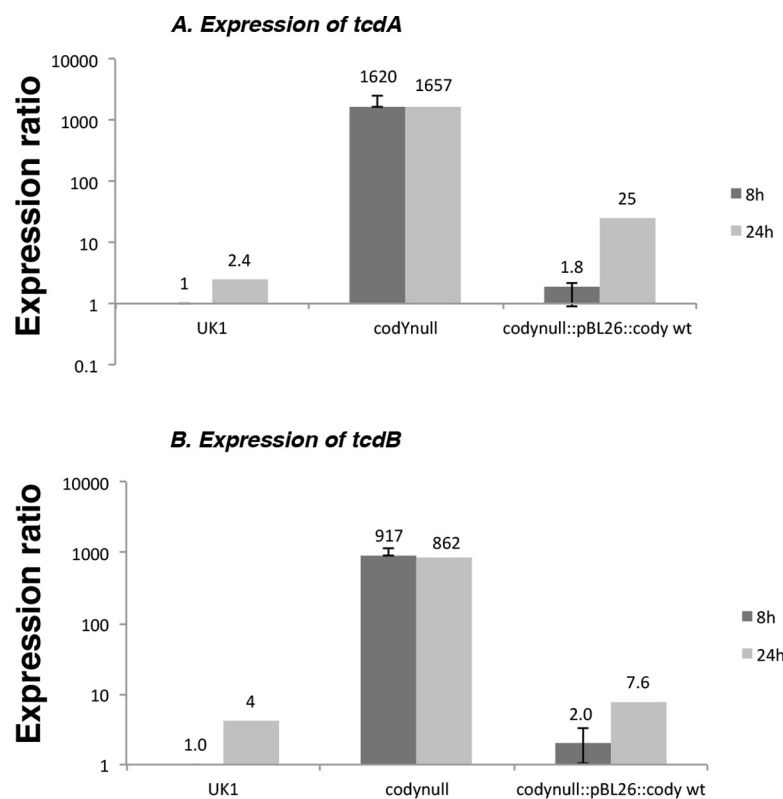
<https://doi.org/10.1371/journal.pone.0206896.t001>



**Fig 2. Toxin titers in cultures of parent and *codY* null mutant strains.** Relative levels of TcdA and TcdB in culture supernatants of UK1, UK1 *codY* (LB-CD16) and UK1 *codY/codY+* (ND-CD10) collected after 24 hrs of bacterial growth were determined by ELISA. Two samples were assayed for each toxin for each strain; toxin levels were averaged and normalized to the values in the wild-type strain (UK1) set at 1.0. Error bars were created for all pairs of samples, but for some pairs the difference was so small that the bars are not visible.

<https://doi.org/10.1371/journal.pone.0206896.g002>

Our approach was to introduce mutations in the region of CodY that includes the BCAA-binding pocket in order to reduce, but not eliminate, the affinity of CodY for the BCAAs or the extent of CodY conformational change upon BCAA binding. To know which amino acid



**Fig 3. Effect of a *codY* null mutation on *tcdA* and *tcdB* transcription in strain UK1.** Cultures of strains UK1, LB-CD16 (*codY::intron::erm*) and ND-CD10 (*codY::intron::erm codY+*) were grown in CDMM medium and samples were removed at 8 hrs (late exponential growth phase) and 24 hrs (stationary phase). RNA was extracted and assayed for *tcdA* and *tcdB* expression by qRT-PCR (see Materials and Methods). Results for the toxin genes were normalized to those obtained for the *rpoA* gene.

<https://doi.org/10.1371/journal.pone.0206896.g003>

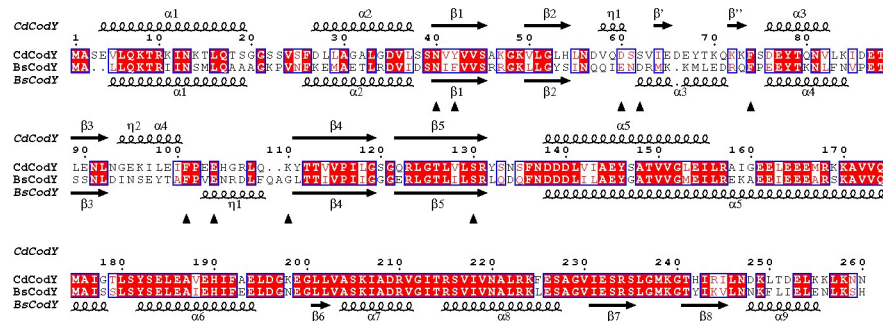
residues of CodY are likely to play a major role in interaction with and response to BCAAs, we determined the crystal structure of the N-terminal half of *C. difficile* CodY (CdCodY).

**a) Structure of the GAF domain of CdCodY.** The cGMP-specific phosphodiesterases-adenylyl cyclases-FhlA (GAF) domain of *B. subtilis* CodY (BsCodY) protein, formed by the N-terminal 155 residues, includes a loop of amino acid residues that interact with BCAAs; such binding induces a change in protein structure leading to increased affinity for CodY-binding sites [47–49]. The GAF domain sequences of *C. difficile* and *B. subtilis* CodY proteins (44% identity) are not nearly as similar as the C-terminal, DNA-binding domains (93% identity) (Fig 4). The crystal structure of the GAF domain of CdCodY bound to isoleucine was solved by molecular replacement and refined against data extending to 1.7 Å resolution, as described in Supplementary Materials and Methods and S1 Table. The CdCodY GAF domain was found to consist of a central, five-stranded, anti-parallel  $\beta$ -pleated sheet ( $\beta$ 3- $\beta$ 4- $\beta$ 5- $\beta$ 1- $\beta$ 2) onto the bottom face of which are packed two  $\alpha$ -helices,  $\alpha$ 2 and  $\alpha$ 4. Extending from the opposite face of the sheet are two extended loops;  $\beta$ 2- $\beta$ 3 meanders across the top of the sheet forming a prominent  $\beta$ -hairpin and a short  $\alpha$ -helix, whereas  $\beta$ 3- $\beta$ 4 does not contain a regular secondary structure. The N-terminal  $\alpha$ 1 helix protrudes strikingly away from the rest of the GAF domain in CdCodY (Fig 5A). Although they differ in sequence, the tertiary structure of the CdCodY GAF domain is similar to that of the GAF domain of BsCodY (Fig 5C); 120 C $\alpha$  atoms can be superposed with a root mean squared positional displacement (rms $\Delta$ ) of 1.5 Å.

In both CdCodY and BsCodY, the effector isoleucine is bound in a pocket on the face of the  $\beta$ -sheet that is distal to the long helices  $\alpha$ 2 and  $\alpha$ 5, but the positions and modes of isoleucine binding are quite different (Fig 5C). The CdCodY binding site is formed by residues from strands  $\beta$ 1 and  $\beta$ 5 from the  $\beta$ -pleated sheet and from the  $\beta$ 2- $\beta$ 3 and  $\beta$ 3- $\beta$ 4 segments of the polypeptide, which embrace the ligand so that it is completely enclosed in the protein interior (Fig 5B). Isoleucine is expected to bind to the protein as a zwitterion. The first of the carboxylate oxygens of isoleucine forms an ion-pairing interaction with the  $\epsilon$ -NH<sub>3</sub><sup>+</sup> of K110 and a charge-dipole interaction with the side chain hydroxyl of S130, with the second carboxylate oxygen forming charge-dipole hydrogen bonds with the phenolic hydroxyl of Y42 and the side chain amide-NH<sub>2</sub> of N40 (Fig 5B). Meanwhile the  $\alpha$ -amino group of the ligand forms an ion-pairing interaction with the side chain carboxylate of E104 and charge-dipole interactions with the hydroxyl of the side chain of S62 and the main chain carbonyl of D60. As shown in Fig 5B, R107 plays an important role in ligand binding since its guanidino moiety forms a two-pronged ion pairing interaction with the side chain of E104. Collectively these interactions fulfill the electrostatic and hydrogen bonding potential of the ligand. The isobutyl side chain, which distinguishes BCAAs from other amino acids, projects into a pocket circumscribed by the aromatic side chains of Y42, F74 and F101, and the aliphatic portions of the side chains of S62, S130 and E104.

In the GAF domain of CdCodY, the isoleucine carboxylate abuts the  $\beta$ -sheet, while in BsCodY it is displaced from the sheet and oriented away from it such that the respective carboxylate carbon atoms are separated by 9 Å in Fig 5C. This structural comparison helps to explain why residues participating in BCAA binding in BsCodY are strongly conserved in the CodY orthologs of some other low G+C Gram-positive bacteria, but not in CodY from *C. difficile* (Fig 4).

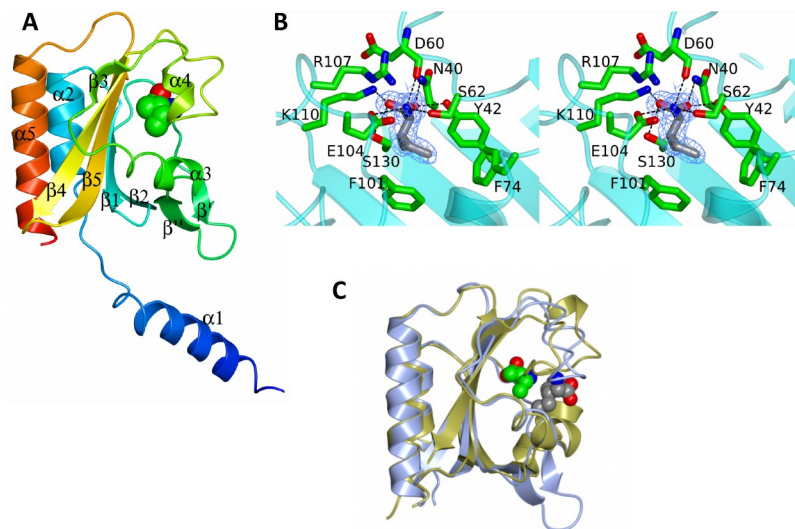
**b) Creation of point mutations in the GAF domain.** To create relatively conservative mutations in the *codY* gene based on the structure determined by crystallography, we modified some amino acid residues, such as F74 and F101, that interact directly with isoleucine, and others, such as E99, P102 and E103, that are nearby, but do not interact directly. Because of the difficulty in creating point mutations in a specific gene in the *C. difficile* UK1 chromosome, we cloned the wild-type or mutant genes within the conjugative transposon Tn916; when



**Fig 4. Alignment of the sequences of CodY from *C. difficile* (Cd) and *B. subtilis* (Bs).** The protein secondary structures for the full-length CdCodY and BsCodY are depicted above and below the sequences, respectively. Sequence identities are indicated by red shading. Residues that form prominent interactions with the effector molecule (isoleucine) in *C. difficile* CodY are labeled with black arrowheads.

<https://doi.org/10.1371/journal.pone.0206896.g004>

introduced by conjugation into *C. difficile* strain LB-CD16 (*codY::intron::erm*), the modified transposon inserted at random sites. In no case was the *codY::intron::erm* mutation altered. Instead the transconjugants have two or more copies of *codY*, one at the normal *codY* locus interrupted by an *erm*-containing intron and at least one additional copy (with or without a GAF domain point mutation) elsewhere on the chromosome. Analysis by qPCR allowed us to choose strains in which the lowest number of copies of Tn916 carrying *codY* had integrated (S2 Fig). The presence of the expected point mutations was verified by amplifying and sequencing the versions of *codY* located within Tn916. The chromosomal locations of the transposons were also determined by sequencing (S2 Table); isolates in which the insertion occurred within a gene were removed from analysis.



**Fig 5. Structure of the GAF domain of *C. difficile* CodY.** A. Ribbon tracing of chain A color ramped from the N-terminus (blue) to the C-terminus (red). The secondary structure elements are labeled. The isoleucine cofactor is shown as spheres and colored by atom: carbon, green; nitrogen, blue; oxygen, red. B. Stereo view of the ligand binding site with the isoleucine effector shown in cylinder format and with carbon atoms colored grey.  $2F_o - F_c$  electron density associated with the effector in the refined structure is contoured at  $1\sigma$  and shown in light blue. Surrounding protein residues are shown in cylinder format with their carbon atoms colored in green. C. Comparison of the GAF domains of CodY from *C. difficile* and *B. subtilis* shown as blue and gold ribbons respectively with the isoleucine ligands shown as spheres with grey and green carbon atoms respectively. The different positions and orientations of the effector molecules are apparent.

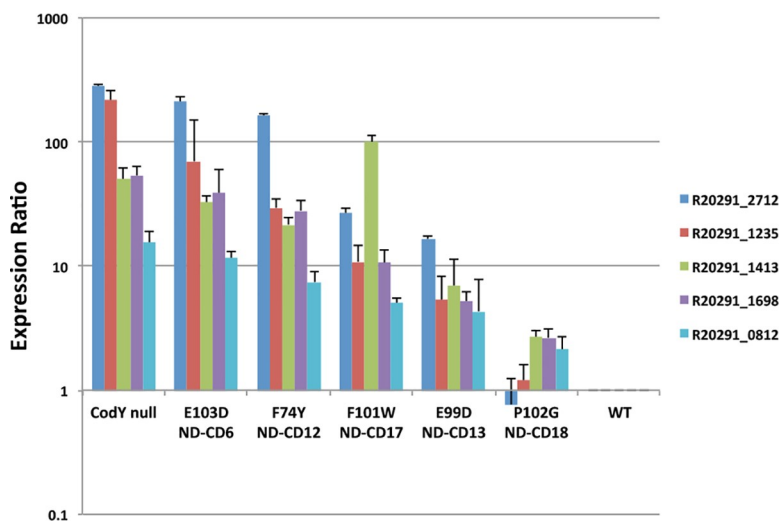
<https://doi.org/10.1371/journal.pone.0206896.g005>

### Altered gene expression in the panel of *codY* mutant strains

To assess the effects of the various point mutations on CodY activity, RNA was extracted from cells grown in tryptose-yeast extract (TY) medium to mid-exponential phase ( $A_{600} = 0.4-0.6$ ). After removal of DNA, the RNA was assayed for five specific transcripts by qRT-PCR. The genes were chosen based on previous microarray analysis of CodY-regulated genes in strain JIR8094 [15]. The results, shown in Fig 6, indicate that the five genes, all of which are repressed by CodY, were derepressed to different extents in the panel of mutants, suggesting that the mutants had different levels of residual CodY activity.

As shown in S3 Fig, all strains used for the experiments below (except for the original null mutant) produced significant levels of CodY protein. Unlike the case in *B. subtilis*, however, the *codY* gene of *C. difficile* is repressed by CodY. As a result, the concentration of CodY protein in most or all point mutants was higher than in wild-type cells (S3 Fig). Thus, the extent of gene regulation mediated by various point mutants is a complex result of both the mutant protein's intrinsic activity and the extent to which it is overexpressed.

Based on their apparent intermediate levels of residual CodY activity, RNA from the wild-type strain, the *codY::intron::erm* null mutant and three of the point mutant strains (ND-CD12 *codY::intron::erm codY* [F74Y]; ND-CD13 *codY::intron::erm codY* [E99D] and ND-CD17 *codY::intron::erm codY* [F101W]) was then subjected to RNA-seq analysis. Of the 3505 protein-encoding genes annotated for this strain, DESeq analysis [50] identified 522 and 79 genes with more than a 3-fold higher or lower transcript abundance, respectively, and with an adjusted P value <0.05 in the null mutant compared to the parent strain. Transcript levels in all strains were then converted to RPKMO values (reads per kb of gene length per million reads aligning to all annotated ORFs in sample) [51] (see S4 Table for all values obtained). To further reduce the likelihood of erroneous identification of CodY-regulated genes, we excluded genes whose highest average normalized transcript level was less than 2 RPKMO in either the wild-type strain (for positively regulated genes) or the *codY* null mutant strain (for negatively



**Fig 6. Effects of *codY* point mutations on expression of five *C. difficile* genes.** Three-to-four individual cultures of wild-type strain UK1, its *codY* null mutant and five different *codY* point mutant strains were grown in TY medium to an  $OD_{600} = 0.4-0.6$ . RNA was extracted and assayed by qRT-PCR for the potential target genes R20291\_2712 (encoding a peptidase), R20291\_1235 (encoding chloromuconate cyclisomerase), R20291\_1413 (*ilvC*), R20291\_1698 (encoding a cell surface protein) and R20291\_0812 (*glgC*). Counts were normalized to those for the *rpoA* gene and presented relative to the counts obtained with the wild-type strain.

<https://doi.org/10.1371/journal.pone.0206896.g006>



regulated genes) and genes whose RPKMO value for one of the null mutant samples was within the range of wild-type samples. This filtering reduced the numbers of genes over- and underexpressed in the *codY* null mutant to 495 and 57, respectively (S3 Table). These genes may not represent the entire CodY regulon, since the exclusion of genes whose transcript level was affected less than 3-fold by a *codY* null mutation is arbitrary and is likely to exclude some genes whose expression is managed by other proteins in addition to CodY (see below). The impact of the three point mutations on gene expression as measured by RNA-seq was similar to that seen by qRT-PCR (Fig 6) for the five genes tested using both analyses (see S3–S7 Tables for RNA-seq data for the genes in Fig 6).

The derepressed genes (S4A Table) included more than 200 nutrient transport and metabolism genes (including amino acid biosynthesis genes), 14 peptidase and protease genes, 12 regulatory proteins, 47 known sporulation genes, the genes of the major toxin locus (*tcdR tcdBEA*) and many genes of unknown function. The most highly underexpressed genes in the *codY* null mutant were the glucitol-sorbitol metabolism operon, pyruvate/formate metabolism genes and cysteine metabolism genes (S4B Table).

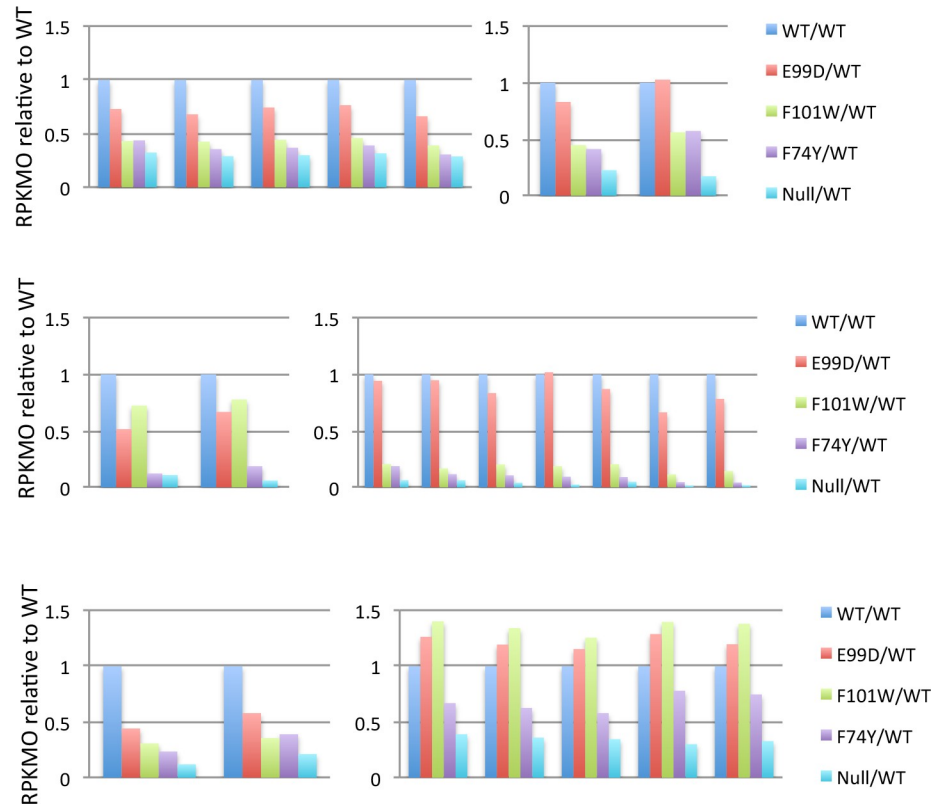
When the transcript levels of both negatively and positively regulated CodY target genes were compared in the panel of point mutants, the overall pattern was consistent with the assumption that the strain expressing the E99D mutant form of CodY is closest to the wild-type for both negative and positive regulation, that the strain expressing the F74Y mutant form is closest to the null mutant (despite its own high level of expression, S3 Fig), and that the strain expressing the F101W mutant form has intermediate residual activity. This pattern can be seen in Fig 7 for examples of the 57 positively regulated genes and in Table 2 and Figs 8 and 9 for the negatively regulated genes. The full sets of positively and negatively regulated genes for the panel of *codY* mutants can be found in S4–S7 Tables.

The transcript levels of some genes positively regulated by CodY were significantly reduced in all of the point mutants, whereas the transcript levels of other genes were reduced in only one or two of the point mutants (Fig 7), suggesting that for some genes the mutant forms of CodY are still able to serve as activators of transcription.

Similarly, the transcript levels of some genes negatively regulated by CodY were markedly increased in one or two or three point mutants whereas the transcript levels of others were not significantly elevated in any of the point mutants (Fig 9). The similarity in the patterns of transcript abundance seen for all the genes of a given operon argues that these patterns are a true indication of the specific impact of the point mutants on transcript levels of genes and operons.

The CodY-regulated genes that are overexpressed in the null mutant but not at all in any of the point mutants include those that encode aconitase (*acn*) and isocitrate dehydrogenase (*icd*) (Fig 9K). The simplest interpretation of this result is that such genes are so strongly repressed by CodY that they are only expressed when CodY is almost totally inactivated. This would occur when the concentrations of the BCAAs and/or GTP in the cell become too low to maintain the minimal concentration of ligand-bound CodY needed to bind to the relevant CodY targets. An extension of this interpretation is that, in cells that are consuming the nutrients that supply ILV and other amino acids, the Krebs Cycle genes would be among the last of the CodY regulon components to be turned on.

This hypothesis is complicated, however. As the intracellular levels of ILV and/or GTP drop, repression by CodY of the ILV biosynthetic pathway and ILV transporters is relieved, resulting in at least partial restoration of the intracellular ILV concentrations. This assumption was tested by analysis of the intracellular concentrations of valine, leucine and intermediates in their biosynthesis. (Intermediates in isoleucine biosynthesis were not analyzed.) As shown in Fig 10, the accumulation of valine, leucine and most intermediates in their biosynthesis in



**Fig 7. RNA-seq analysis of genes and operons positively regulated by CodY.** Transcript levels (RPKMO values relative to wild-type, which was set at 1.0) in the *codY* null mutant and each of three point mutants are shown for examples of the 57 genes whose transcript levels were >3-fold lower in the *codY* null mutant than in the wild-type strain. The different colored patterns for each gene indicate the relative transcript levels in each of the mutant strains. Sections A-F present the results for all the genes of six different clusters. In clusters A and B, the gene names are their numbers in the R20291 genome.

<https://doi.org/10.1371/journal.pone.0206896.g007>

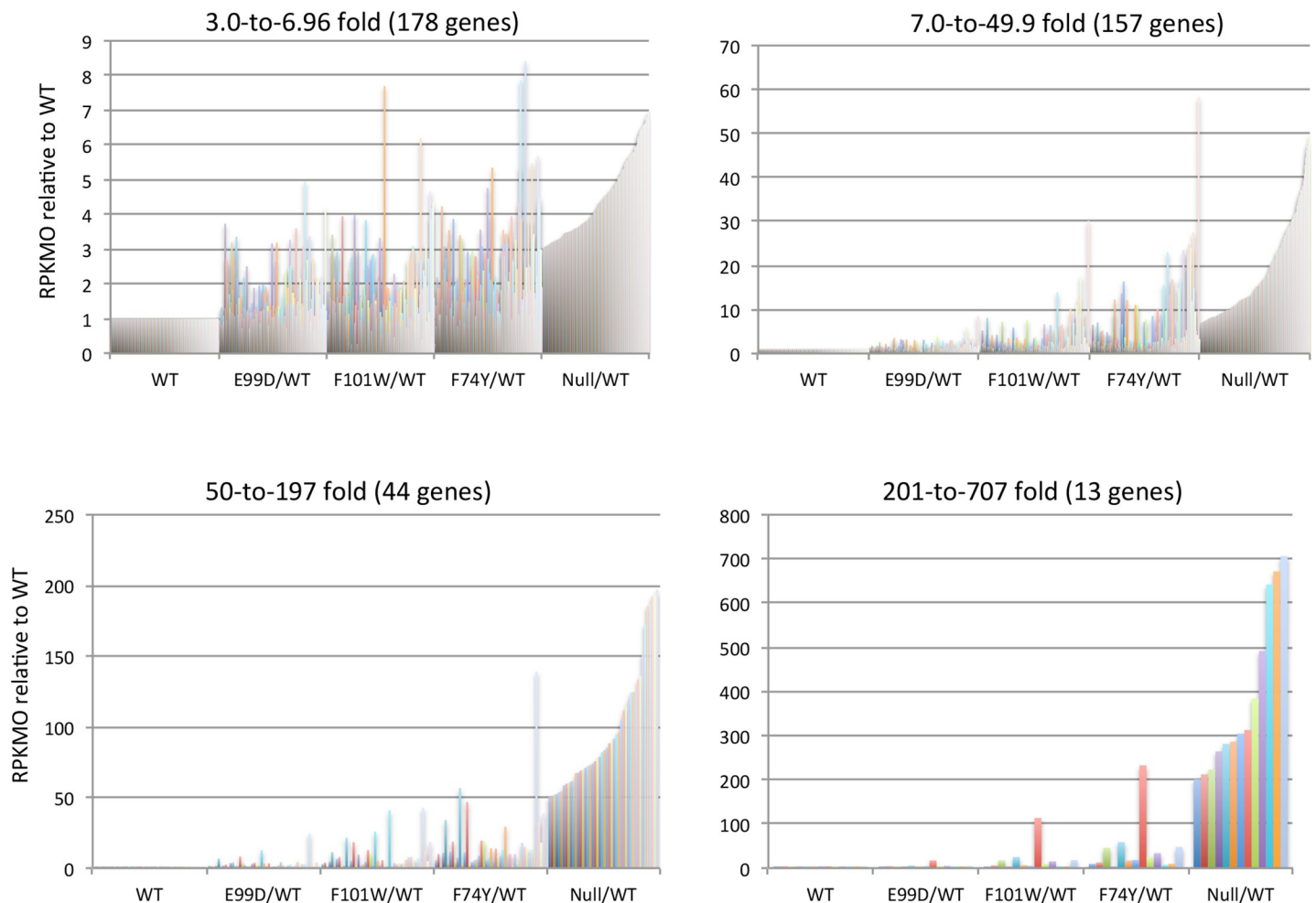
the panel of *codY* mutant cells is consistent with the increased expression of the biosynthesis genes as CodY activity decreases. Note that the measurement of isopropylmalate does not distinguish between 2-isopropyl- and 3-isopropyl- and that the leucine value includes both

**Table 2. Overexpressed genes in *codY* mutant strains.**

Overexpression compared to WT	<i>codY</i> mutation			
	Null	F74Y	F101W	E99D
>3-fold	495	257	173	98
>5-fold	333	149	105	39
>10-fold	210	91	57	26
>20-fold	135	43	40	20
>50-fold	71	17	28	15
>100-fold	40	7	23	14

As indicated in the first line of data, 495 genes were overexpressed >3-fold in the *codY* null mutant compared to the WT. Of those 495 genes, 333 were overexpressed >5-fold, 210 were overexpressed >10-fold, etc. Members of the group of genes overexpressed in the null mutant were overexpressed to various extents in the individual point mutants. The ethanolamine gene cluster (CDR20291\_1828–1846) has such a low level of expression in the *codY* null mutant that 13 of the 19 genes were excluded from the listing. However, this gene cluster was very highly overexpressed in the point mutants. The pattern of expression is presented in detail below.

<https://doi.org/10.1371/journal.pone.0206896.t002>



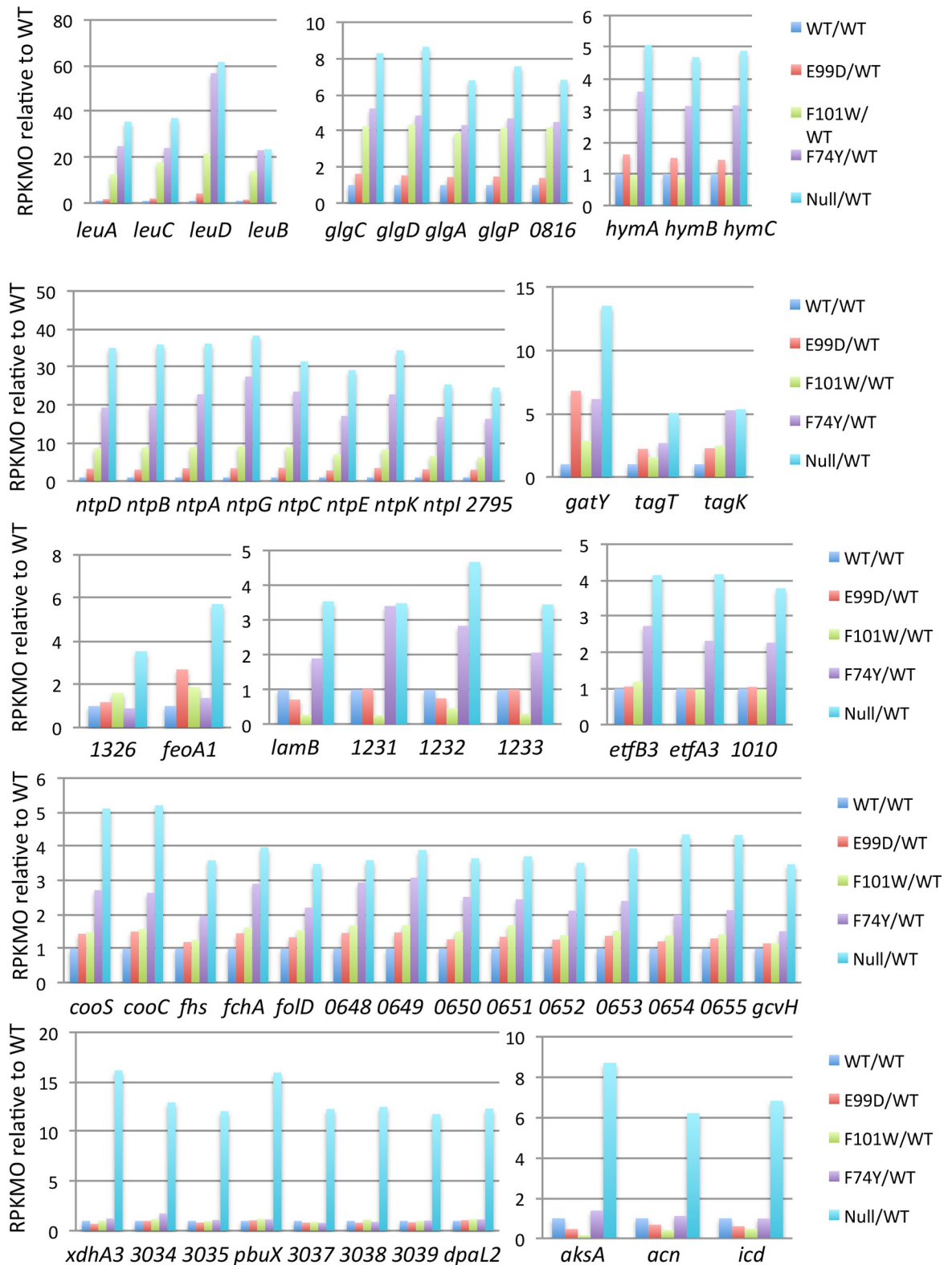
**Fig 8. RNA-seq analysis of genes negatively regulated by CodY.** Of the 495 genes whose transcript levels (measured as RPKMO) were >3-fold higher in the *codY* null mutant than in the wild-type strain, the relative levels for 392 genes are shown for the null mutant and three point mutants. The genes were divided into subgroups according to the level of overexpression in the null mutant.

<https://doi.org/10.1371/journal.pone.0206896.g008>

leucine and isoleucine. The accumulation of amino acids is presumably a result of derepression of both the biosynthesis operons and transporters.

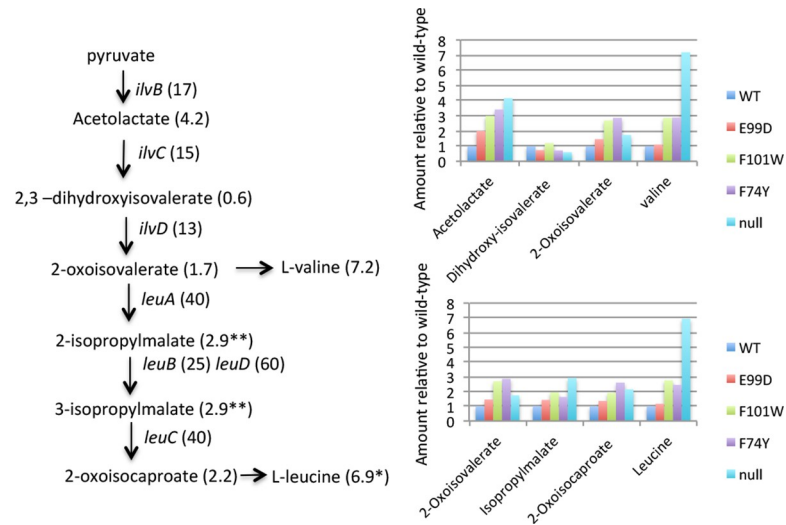
### CodY-mediated regulation of virulence genes

The vast hyperexpression of the major toxin genes (*tcdA*, *tcdB*) in exponential phase cells of the *codY* null mutant of strain UK1 (Fig 11) is consistent with its increased virulence in mice (Fig 1), although other factors must be involved as well. Many other genes are known to contribute to virulence [52]. For instance, the genes encoding the ADP-ribosylating toxin CdtAB were overexpressed 2.5–3.0 fold in the *codY* null mutant, even though the expression of CdtR, the positive regulator of *cdtAB* [53, 54], was not affected by *codY* mutations. Interestingly, the *tcdA* and *tcdB* genes are positively regulated by CdtR in two ribotype 027 strains [55]. In contrast, transcription of other virulence factor genes, such as the *dlt* operon [56], *fbpA* [57], *groEL* [58], *fliC* and *fliD* [59], the *pdaV-sigV* gene cluster [60], zinc- and collagen-binding protein genes [61–63], and a lipoprotein gene [64] was not affected substantially by a *codY* null mutation. In a cluster of 39 surface protein-related genes that includes *uppS*, *tuaG*, *manC*,



**Fig 9. RNA-seq analysis of individual metabolic operons negatively regulated by CodY.** Examples of the effects of the *codY* null mutation and three point mutations on individual negatively regulated operons are shown. For each gene, the average RPKMO value for wild-type cells was assigned a value of 1.0 and average transcript levels in each mutant strain were calculated relative to that in the wild-type. The different colored patterns for each gene indicate the relative transcript levels in each of the mutant strains. Each of the A-K parts of the figure present separate gene clusters.

<https://doi.org/10.1371/journal.pone.0206896.g009>

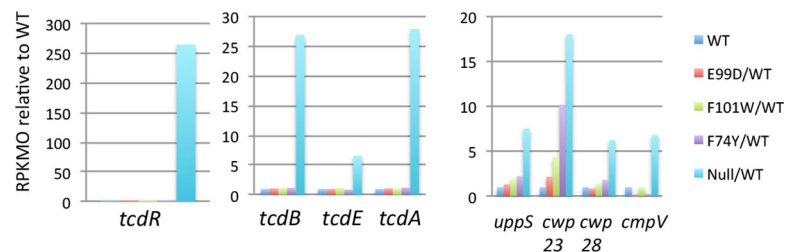


**Fig 10. Impact of *codY* mutations on expression of the genes and biochemical intermediates of the BCAA biosynthesis pathway.** On the left side of the figure, the pathway from pyruvate to valine and leucine is shown. The genes that code for the various enzymes are in italics. The numbers in parentheses indicate the concentrations of the various metabolites in the *codY* null mutant relative to the wild-type. The right side of the figure shows the results of metabolomics analysis for a panel of *codY* mutants relative to wild-type (the latter set at 1.0 for all metabolites). The pathway to valine is shown above the pathway to leucine. The asterisk associated with the concentration of leucine indicates that the concentration is actually the sum of leucine and isoleucine, which were not separated in this analysis. The double asterisk for 2- and 3-isopropylmalate indicates that 2.9 is the sum of the two inseparable isomers.

<https://doi.org/10.1371/journal.pone.0206896.g010>

*pgm2*, *mviN*, *cwp84*, *cwp66*, *secA2* and *slpA* [65–67], only the *uppS* gene was significantly derepressed (7.5-fold) in a *codY* null mutant (Fig 11). Of 22 other surface protein-encoding genes (Cwp8–Cwp29) [65], the genes encoding CwpV (R20291\_0440), Cwp23 (R20291\_1698) and Cwp28 (R20291\_1911) were derepressed 6.8-, 18.0- and 6.2-fold, respectively, in the *codY* null mutant (Fig 11).

The intracellular concentration of c-di-GMP also has an impact on *C. difficile* virulence. Overexpression of the di-GMP cyclase gene CD630\_14200 causes a decrease in motility [68] and in expression of the *tcdA* and *tcdB* genes [69]. This effect appears to be due to inactivation by c-di-GMP of SigD, the RNA polymerase sigma factor that activates flagellar genes and the *tcdR* gene [13, 69]. Of the 31 genes encoding di-GMP cyclases or c-di-GMP phosphodiesterases in strain R20291 [70], only one, R20291\_0884 (corresponding to CD630\_10280),



**Fig 11. RNA-seq analysis of virulence genes.** The relative transcript levels (RPKMO) of eight virulence-associated genes in the *codY* null mutant and three *codY* point mutants is shown. All genes displayed here were overexpressed >3-fold in the *codY* null mutant. The different colored patterns for each gene indicate the relative transcript levels in each of the mutant strains. In part A, the genes of the major toxin locus are shown in their genetic order. In part B, other virulence genes are grouped, but are not genetically linked to each other.

<https://doi.org/10.1371/journal.pone.0206896.g011>

encoding a di-GMP cyclase, was overexpressed (15.9-fold) in the *codY* null mutant of strain UK1 (S4 Table). None of these genes was underexpressed in the *codY* null mutant. Since overexpression of a di-GMP cyclase would be expected to limit virulence, the effect of CodY on R20291\_0884 is unlikely to contribute to the *codY* mutant strain's hypervirulence. Interestingly, a di-c-GMP hydrolase gene, CD630-15150, is repressed by CodY in strain 630 [15, 71], but not in UK1 (R20291\_1364; S3 Table).

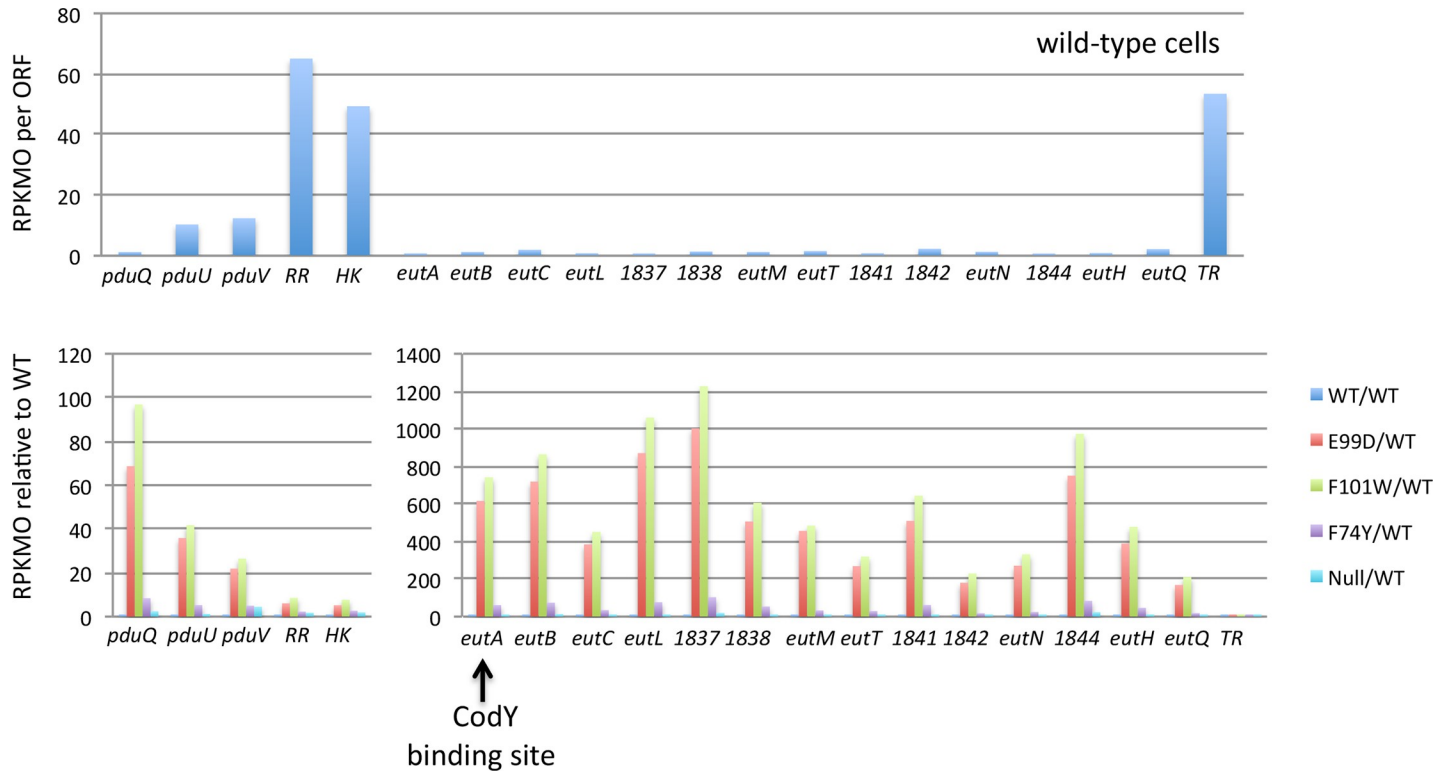
### Complex regulation of the ethanolamine utilization and other pathways

Although the *codY* point mutant strains had intermediate residual CodY activity with respect to the parent strain and the *codY* null mutant for more than 390 genes (Fig 8), some genes were most highly expressed in the point mutants. A major example is the large gene cluster that encodes the proteins of ethanolamine metabolism (CDR20291\_1828–1846) (Fig 12 and S8 Table). In *C. difficile* strain 630 $\Delta$ erm, all of these genes are expressed at a high level at 2 hrs after the end of exponential growth [72]. Most of the genes in the *eut* cluster were expressed at so low a level in the *codY* null mutant that they were not included among the 495 genes whose transcripts are increased >3-fold by the null mutation (S3 Table). All the genes of this cluster, however, proved to be very highly expressed in all three *codY* point mutants (S8 Table). In addition, 7 genes, including the *ilvCB* operon (CDR20291\_1412–1414) and a cellobiose transport operon (CDR20291\_2927–2928), were much more highly expressed in the F101W mutant than in the *codY* null mutant.

The mechanism responsible for the unusual effect of the *codY* point mutations is not known, but the general pattern is similar to that seen for the *B. subtilis braB* gene [73]. CodY is a direct repressor of *braB*, but is also a repressor of *scoC*, which encodes a second repressor of *braB*. Thus, as CodY activity diminishes, direct repression of *braB* is initially relieved, but a further decrease in CodY activity leads to ScoC accumulation and restoration of repression. The identities of the putative second regulators of *eut*, *ilvCB* and cellobiose genes are unknown. In the case of *eut*, two genes within the cluster (R20291\_1831 and 1832) appear to encode a response regulator and histidine kinase, but they are likely to be positive regulators of *eut* gene expression. An apparent transcription factor encoded immediately downstream of the *eut* gene cluster (R20291\_1847) is also unlikely to be the predicted *eut* regulator, since its own expression is only slightly affected by *codY* mutations.

### Regulation by CodY of sporulation gene expression

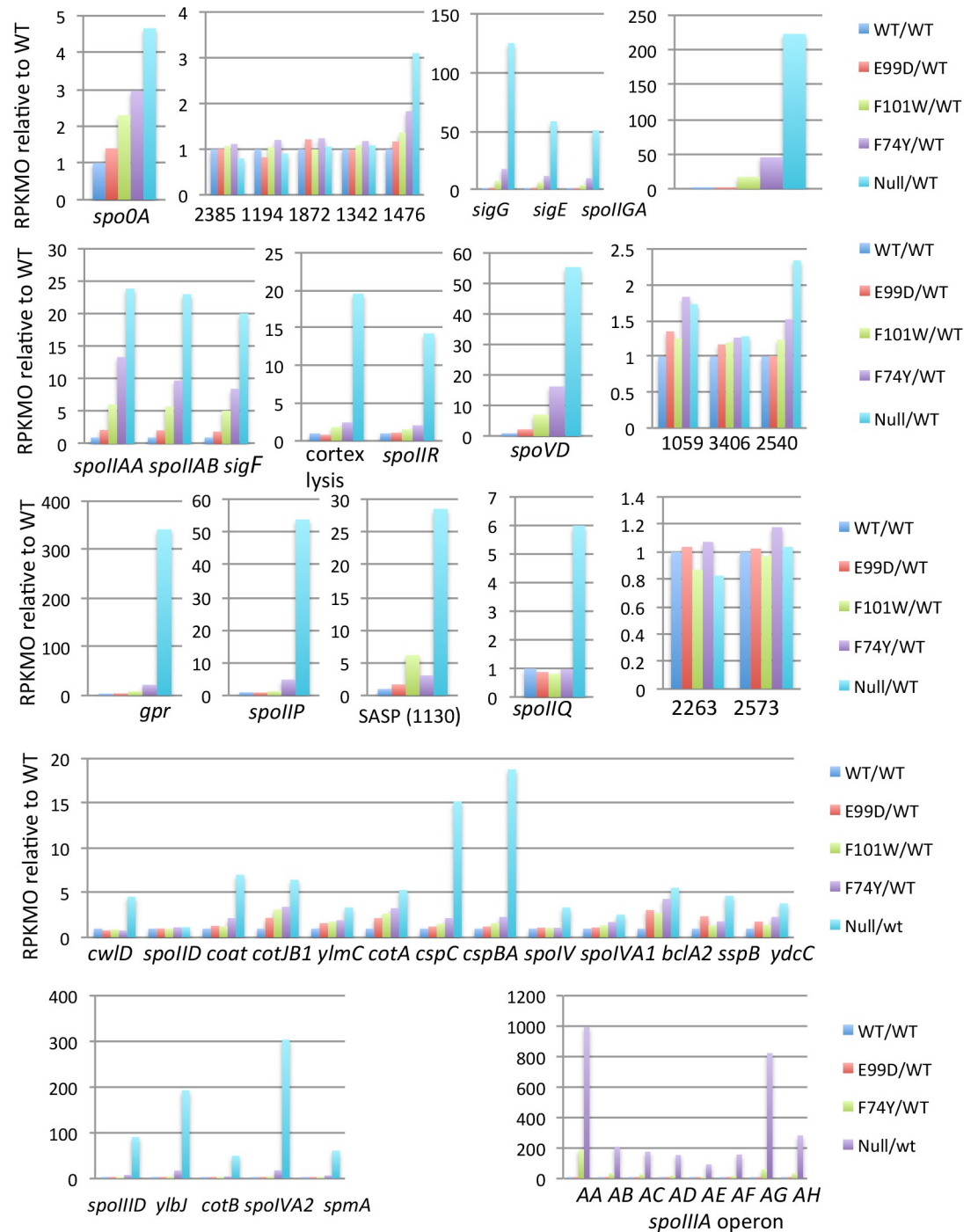
Endospore formation is a type of bacterial differentiation that is restricted to a handful of bacterial genera; it has been studied principally in the genera *Bacillus* and *Clostridium* and a few of their close relatives. When faced with certain kinds of nutrient limitation, cells of spore-forming species stop dividing and create two compartments within the same cytoplasmic membrane. One compartment, the forespore, is the precursor of the mature spore and is surrounded by the mother cell. The forespore and mother cell express a large number of compartment-specific genes. When the developing spore has reached maturity, the mother cell lyses, releasing the spore into the environment. Both the initiation and the various stages of spore formation are dependent on specific transcriptional regulators. Five different RNA polymerase sigma factors,  $\sigma^H$ ,  $\sigma^F$ ,  $\sigma^E$ ,  $\sigma^G$ , and  $\sigma^K$ , are needed for transcription of key genes, some of which are transcribed before internal cell division and others in the separate compartments [74, 75]. The Spo0A protein is a positive regulator of a large number of genes at the onset of sporulation; its own synthesis depends on  $\sigma^H$  and it activates transcription of the genes that encode  $\sigma^F$  and  $\sigma^E$  [76–79].



**Fig 12. RNA-seq analysis of ethanolamine metabolism gene cluster.** Some CodY-repressed genes had unusual patterns of expression in the panel of *codY* point mutants. Shown here are the genes of the ethanolamine metabolism gene cluster, R20291\_1828 to 1846. In the top panel, the average RPKMO values for each gene in the wild-type strain are shown. In the bottom panel, the transcript levels in the mutants relative to the wild-type (set at 1.0) are presented. The different colored patterns for each gene indicate the relative transcript levels in each of the mutant strains. The top panel indicates that several genes, including those encoding a putative histidine kinase (HK) and response regulator (RR) are expressed at a relatively high level in the parent strain. The lower panel shows that all genes were overexpressed to a greater extent in two point mutants, E99D and F101W, than in the null mutant.

<https://doi.org/10.1371/journal.pone.0206896.g012>

Previous work by Nawrocki *et al.* [17] showed that spore formation and the expression of *spo0A* and the genes that encode sporulation sigma factors  $\sigma^F$ ,  $\sigma^E$  and  $\sigma^G$  are significantly enhanced in the *codY* null mutant (LB-CD16) of strain UK1 compared to wild-type. The effect of the *codY* mutation was overcome by complementation with the wild-type gene [17]. Consistent with this finding, the *spo0A* gene, which encodes the primary factor needed for initiation of sporulation gene expression, was derepressed during exponential growth phase in the *codY* null mutant (Fig 13A), as was a gene encoding a histidine kinase known to activate Spo0A by phosphorylation (R20291\_1476, equivalent to CD630\_15790) [80] (Fig 13B). Four other genes encoding potential Spo0A kinases were not affected by *codY* mutations (Fig 13B). The genes and operons encoding  $\sigma^F$ ,  $\sigma^E$ ,  $\sigma^G$  were also overexpressed >3-fold in the *codY* null mutant (Fig 13C–13E), as were more than 40 other genes that are required for *C. difficile* sporulation or are homologous to known sporulation genes in other bacteria (Table 3 and Fig 13D–13H). The distribution of genes under the control of the various sporulation sigma factors is shown in Fig 13I–13M for  $\sigma^F$ , Fig 13N–13P for  $\sigma^E$ , Fig 14A–14C for  $\sigma^G$  and Fig 14D and 14E for  $\sigma^K$ . The gene encoding  $\sigma^H$  was only overexpressed 1.8-fold (adjusted P value = 0.054) in the *codY* null mutant and the gene encoding  $\sigma^K$  was only overexpressed 2.9-fold; neither gene was included in the list of CodY-regulated *spo* genes (Table 3). Moreover, genes encoding quorum-sensing factors that potentially activate Spo0A phosphorylation (R20291\_2639–2640 and 3187–3189) [81] were neither overexpressed nor underexpressed more than 3-fold in the *codY* null mutant. Fimlaid *et al.* [82] reported that in strain JIR8094 more than 200 stationary phase



**Fig 13. RNA-seq analysis of Spo0A-,  $\sigma^F$ - or  $\sigma^E$ -dependent genes.** (A-H) Expression patterns of Spo0A-dependent (A-H),  $\sigma^F$ -dependent (I-M) and  $\sigma^E$ -dependent (N-P) genes and operons in a panel of *codY* mutants. Transcript levels (RPKMO values) are presented relative to the levels in the wild-type strain (set at 1.0). The different colored patterns within each part of the figure reflect the behavior of individual genes. In parts C, E, F, and P, the genes are from individual clusters. In parts B, H, M, N and O, the genes presented together are not genetically linked. In part P, the expression of the *spoIIIA* operon in strain ND-CD17 was not included, because the transposon carrying the mutant *codY* gene inserted upstream of *spoIIIAA* and reduced expression of the entire operon.

<https://doi.org/10.1371/journal.pone.0206896.g013>



**Table 3. Hyperexpression of *spo* genes in a *codY* null mutant.**

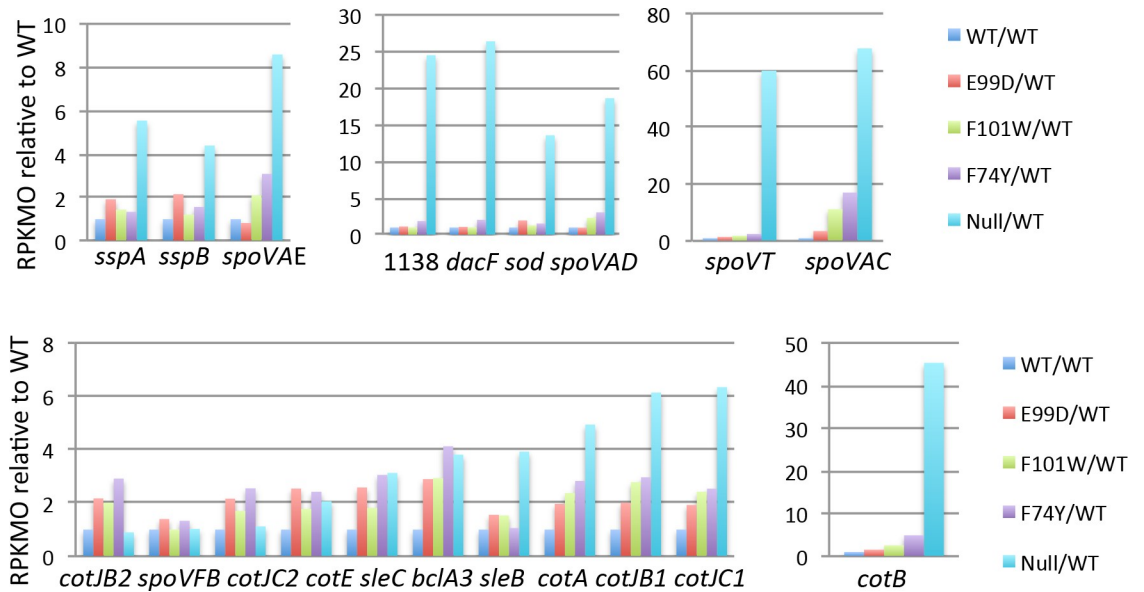
Gene	Expression ratio ( <i>codY</i> null/WT)	Gene	Expression ratio ( <i>codY</i> null/WT)
<i>spo0A</i>	4.7	<i>spoIVA</i>	304
R20291_1476 (Spo0A kinase)	3.5	R20291_0714 (stage IV sporulation protein)	5.1
<i>spoIIA</i> operon: <i>spoIIAA</i>	24	R20291_3400 (spore cortex-lytic enzyme)	21
<i>spoIIAB</i>	23	<i>sspA</i>	5.7
<i>sigF</i>	20	<i>sspB</i>	4.7
<i>spoIIIE</i>	223	<i>cspBA</i>	19
<i>spoIIG</i> operon: <i>spoIIGA</i>	51	<i>cspC</i>	15
<i>sigE</i>	59	<i>gerG</i>	50
<i>spoIIP</i>	61	<i>spoVAC</i>	76
<i>spoIIR</i>	16	<i>spoVAD</i>	21
<i>sigG</i>	125	<i>spoVAE</i>	9.4
<i>spoIIIA</i> operon: <i>spoIIIAA</i>	995	<i>spoVD</i>	60
<i>spoIIIAB</i>	212	<i>spoVS</i>	3.2
<i>spoIIIAC</i>	177	<i>spoVT</i>	670
<i>spoIIIDAD</i>	157	spore coat assembly	50
<i>spoIIIAE</i>	95	R20291_0212 (spore coat protein)	7
<i>spoIIIAF</i>	160	<i>cotJBI</i>	6.5
<i>spoIIIAG</i>	824	<i>cotJCI</i>	6.6
<i>spoIIIAH</i>	284	<i>spmA</i>	61
<i>spoIIIID</i>	91	<i>spmB</i>	14
<i>sleB</i>	4.6	<i>cwlD</i>	4.6
<i>sleC</i>	3.3	<i>bclA2</i>	5.6
<i>gpr</i>	384	<i>bclA3</i>	4.1
		<i>gerS</i>	3.8

The indicated genes have been shown to be required for spore formation in *C. difficile* or are homologs of *B. subtilis* sporulation genes. The expression ratios are based on RPKMO levels as determined by RNA-Seq.

<https://doi.org/10.1371/journal.pone.0206896.t003>

genes depend on Spo0A for their expression. The number of Spo0A-dependent genes in strain JIR8094 was later increased to 434 (A. Shen, personal communication) based on RNA-seq results of Pishdadian et al. [83]. Taking advantage of a table provided by L. Barquist and R. Fagan that matches R20291 genes with those of strain 630, we found that 159 of the 434 Spo0A-dependent genes were overexpressed  $\geq 3$ -fold during exponential growth in the *codY* null mutant (S7 and S8 Tables). Dembek et al. [84] used transposon mutagenesis to identify R20291 genes essential for efficient sporulation. Of the 798 genes in which transposon insertions limited spore formation 4-fold or more, 138 were overexpressed during exponential phase growth in the *codY* null mutant and 14 were underexpressed.

The expression patterns of sporulation operons in the point mutant strains suggested that the level of CodY activity needed to maintain repression of sporulation genes is variable (Figs 13 and 14). The *spo0A* gene was partially activated in all of the point mutants, whereas other sporulation genes required greater inactivation of CodY in order to be expressed, implying either that the level of *spo0A* derepression determines to different extents the level of expression of these genes or that some or all of these genes are also direct targets of CodY.



**Fig 14. RNA-seq analysis of sporulation genes dependent on  $\sigma^G$  or  $\sigma^K$ .** Expression patterns of individual  $\sigma^G$ -dependent (A-C) and  $\sigma^K$ -dependent (D-E) genes in a panel of *codY* mutants. Transcript levels (RPKMO values) are presented relative to the levels in the wild-type strain (set at 1.0). The different colored patterns within each part of the figure reflect the behavior of individual genes. Part E displays the results with a single gene (*cotB*), but all other parts of the figure show collections of unlinked genes that have similar levels of overexpression in the *codY* null mutant.

<https://doi.org/10.1371/journal.pone.0206896.g014>

## Discussion

### Impact of CodY on metabolism

More than 250 of the 552 genes overexpressed or underexpressed more than 3-fold in the *codY* null mutant strain during exponential growth are involved in metabolic processes, e.g., transport and metabolism of carbon and nitrogen sources, biosynthesis of amino acids, energy production, glycogen biosynthesis, sugar metabolism, and the Krebs Cycle. The results obtained with strains expressing mutant forms of CodY with different levels of residual activity imply that some metabolism genes (e.g., the *leu* and *glg* operons) are considerably derepressed when CodY activity is only slightly reduced, some genes (e.g., the *ntp* and *etf* operons) are only significantly derepressed when CodY activity is more reduced and other genes (e.g., the *xdhA* and *aksA* clusters) are only derepressed when CodY activity is eliminated. The implication of these findings is that the affinity of CodY for its binding sites varies considerably from target to target, a situation proven to be true on a genome-wide basis for *B. subtilis* CodY [30], but not yet tested for *C. difficile* CodY. An apparent implication of these results is that, as cells consume critical nutrients, expression of the operons cited above would increase first for the *leu* and *glg* operons and lastly for the *xdhA* and *aksA* operons. However, this implication may not be correct, because all of these operons are undoubtedly regulated by other factors as well, each of which responds to different changes in nutrient availability. Thus, the timing of expression of CodY-regulated genes and operons when cells transition from exponential to stationary phases is likely to be complex.

### Impact of CodY on virulence

The CodY proteins of multiple Gram-positive pathogens serve as virulence regulators. As shown in Fig 1, despite the high level of virulence of *C. difficile* strain UK1, inactivation of

CodY further increases its virulence significantly. Ribotype 027 strains have often been viewed as “hypervirulent” compared to other isolates, although not all members of the ribotype are highly virulent [85]. In *S. aureus* [39–42, 86, 87] and *Bacillus thuringiensis* [88], CodY is also a negative regulator of virulence. In fact, a hypervirulent *S. aureus* USA300 strain became even more virulent when the *codY* gene was inactivated [39]. Thus, the virulence of multiple species is restrained to a significant extent under conditions where CodY is active. In *Bacillus anthracis* [28], *Bacillus cereus* [35], *Listeria monocytogenes* [36], *Clostridium botulinum* [43, 89] and several Streptococcal species [33, 90, 91], however, CodY appears to activate virulence. Moreover, in *Clostridium perfringens*, the impact of CodY on virulence varies from strain to strain. In both the Type D strain CN3718 and the Type A strain SM101, CodY activates virulence gene expression, but the mechanisms of regulation are different. In CN3718, CodY directly activates expression of the *etx* gene, which encodes epsilon toxin [34], whereas, in SM101, CodY activates expression of the *cpe* (*C. perfringens* enterotoxin) gene indirectly by positively regulating certain key sporulation genes [92]. By contrast, CodY represses sporulation gene expression in strain CN3718 [34]. The variability in the impact of CodY on virulence and sporulation implies that different strains and species have evolved to use this global regulator in ways that suit their preferred conditions for sporulating and causing damage to the host.

### Mechanism of regulation of sporulation genes by CodY

Virtually all of the genes of known or predicted functions that are expected to be required for spore formation in *C. difficile* are expressed during exponential growth phase in mutant strains defective in CodY. The *spo0A* gene, which encodes the primary regulator of sporulation, is derepressed 4.5-fold in the *codY* null mutant strain (Fig 13). At first approximation, this result is surprising, since many other sporulation genes, all of which depend on direct or indirect activation by Spo0A, are derepressed to a much greater extent. Two factors may be responsible for this level of *spo0A* gene regulation. First, the *spo0A* gene in strain JIR8094 does not have a CodY-binding site [15]; the sequence upstream of *spo0A* in strain UK1 is very similar to that in JIR8094 and presumably also lacks a binding site. Thus, CodY regulation of *spo0A* is undoubtedly indirect. Second, in wild-type cells of strain UK1, *spo0A* mRNA is within the top 4% of mRNAs in terms of abundance (RPKMO value); in the *codY* null mutant, the *spo0A* mRNA moves to the top 3%. Thus, increased synthesis of *spo0A* mRNA may not be critical for spore formation. Instead, phosphorylation of already highly abundant Spo0A is likely to be the critical factor that initiates sporulation. The gene encoding a kinase known to phosphorylate Spo0A *in vitro* [80] has a CodY-binding site [17] and is derepressed in the *codY* null mutant (Fig 13), potentially serving as the primary factor in leading to sporulation gene expression. (The 4.5-fold increase in *spo0A* mRNA in the *codY* null mutant may be due to autoactivation of *spo0A* transcription by Spo0A~P.) In fact, very few sporulation genes have direct CodY-binding sites, suggesting that CodY regulates sporulation genes primarily through indirect effects.

Despite the limited direct effects, the impact of CodY on spore formation by *C. difficile* is much greater than for *B. subtilis*. Inactivation of CodY increases spore formation by *B. subtilis*, but the effect is relatively small and only occurs in particular media [21]. It is likely that the complex regulation of spore formation in *B. subtilis* minimizes the impact of inactivation of a single negative regulatory protein. *C. difficile*, by contrast, lacks several of the regulators of *B. subtilis* spore formation, including Spo0F and Spo0B (members of the Spo0A phosphorylation cascade), Rap proteins that dephosphorylate Spo0F~P, and Spo0E (a Spo0A~P phosphatase). The lack of such complex regulation may explain why a *codY* mutation has a stronger effect on sporulation in *C. difficile* than in *B. subtilis*.

## Co-regulation of sporulation and toxin synthesis

Toxin synthesis and spore formation are both major contributors to *C. difficile* infection. TcdA and TcdB are the primary factors causing damage to the host and spore formation is critical for resistance to antibiotics, for survival of the bacteria outside the GI tract and for transmission of the bacteria to other hosts. Thus, it makes sense that both processes are induced when nutrients become limiting and are subject to overlapping regulation, as shown here and in previous publications [16, 17]. Unlike sporulation genes, which appear to be mostly indirect targets of CodY, the *tcd* locus is a direct target of CodY [16]. Although CodY represses both processes, the results with the point mutants described here imply that many sporulation genes are turned on when CodY is only partially inactivated, whereas expression of the toxin genes requires complete inactivation of CodY.

A potential complication is the likely impact of toxin synthesis on spore formation and *vice versa*. Toxin activity is expected to increase the availability of nutrients due to damage to host cells. In cells that have not yet committed to sporulation, an increase in nutrient availability may permit renewed growth, thereby delaying or preventing sporulation. Since different RNA polymerase sigma factors are required for transcription of toxin genes ( $\sigma^A$ ,  $\sigma^D$  and TcdR) [9, 11–13] and sporulation genes ( $\sigma^A$ ,  $\sigma^H$ ,  $\sigma^F$ ,  $\sigma^E$ ,  $\sigma^G$  and  $\sigma^K$ ) [74, 76, 82], it seems unlikely that sporulating cells express the toxin genes, but no experiments to date have addressed this issue. How does the *C. difficile* population balance the production of toxins and spores? Is the population divided between the two responses to nutrient limitation? The relative behavior of the CodY point mutants described here implies that, as nutrients become limiting, many sporulation genes are turned on; the toxin genes by contrast are among the last genes to be turned on as nutrient availability decreases, implying that at least a portion of the population initiates sporulation before any of the cells produce the toxins. Since spore formation is irreversible after an early stage, the cells that choose to sporulate will complete the process, creating a mixed population that provides multiple advantages to the bacterium. The non-sporulating cells that produce toxin will generate nutrients that will allow them to multiply, while the sporulating members of the population will become prepared to survive outside the GI tract and initiate new infections.

Multiple factors contribute to the regulation of toxin synthesis and sporulation, in some cases providing co-regulation and in other cases opposing regulation. Like CodY, the glucose-activated CcpA protein is a negative regulator of both the *tcd* gene cluster and the *spo0A* and *sigF* genes [14]. In addition, in strain R20291, a *tcdR* mutation causes a decrease in spore formation and heat resistance of the spores produced [10], implying that TcdR is a positive regulator of sporulation as well as toxin synthesis. Possible mechanisms include stimulation of sporulation by toxins (by an unknown mechanism) or a requirement for the TcdR sigma factor for maximal transcription of one or more sporulation genes. On the other hand, in strain 630 $\Delta$ erm, the RstA protein is an activator of sporulation, but an inhibitor of toxin synthesis [93]. In the same strain, one group found that a *spo0A* mutation causes overexpression of the *tcdA* gene [8]. Other researchers detected binding of Spo0A to the *tcdB* gene *in vitro*, but saw little or no effect of a *spo0A* mutation on toxin production *in vivo* [6]. A third group found evidence of overexpression of both *tcdA* and *tcdB* in a *spo0A* mutant of a ribotype 027 strain, but not in 630 $\Delta$ erm [7]. (It is likely that differences in the isolates and media used for these experiments contribute to their varying results.) Although the details remain to be resolved, it is likely that *C. difficile* has evolved to base its extent of toxin synthesis and sporulation (and potentially the decision as to which pathway to pursue at the single-cell level) on multiple environmental factors and regulatory proteins.

## Materials and methods

### Bacterial strains and growth conditions

The bacterial strains and plasmids used in this study are listed in Table 4. *C. difficile* strains were grown in tryptose-yeast extract (TY) medium [94], BHIS medium [95] or in defined CDMM [96], supplemented with 250 µg D-cycloserine per ml, 40 µg kanamycin per ml, 20 µg thiamphenicol per ml, and 20 µg lincomycin per ml or 5 µg erythromycin per ml, as needed. *C. difficile* strains were maintained at 37°C in an anaerobic chamber (Coy Laboratory Products) in an atmosphere of 10% H<sub>2</sub>, 5% CO<sub>2</sub> and 85% N<sub>2</sub>. *Escherichia coli* and *B. subtilis* strains were grown at 37°C in L broth [97] supplemented with 20 µg chloramphenicol per ml or 100 µg ampicillin per ml for *E. coli* and with 1 µg of erythromycin and 12.5 µg of lincomycin per ml for *B. subtilis*, as needed.

### Containment practices and biosafety precautions

Laboratory experiments using *C. difficile* strain UK1 and its derivatives were approved by the Tufts University Institutional Biosafety Committee as registration number 2013-BRIA16 from

**Table 4. Bacterial strains and plasmids.**

Strains and Plasmids	Description	Source and/or Reference
<b>Plasmids</b>		
pJS107	TargeTron vector	J. Sorg
pBL100	TargeTron vector	[98]
pBL92	pBL100 <i>codY::intron::ermB</i>	This work
pBL103	pJS107 <i>codY::intron::ermB</i>	This work
pSMB47	<i>E. coli</i> plasmid carrying part of Tn916	[99]
pBL26	Derivative of pSMB47 carrying the <i>cat</i> gene from pJIR1456	[17]
pND3	pBL26:: <i>codY</i> <sup>+</sup>	This work
pND5	pBL26:: <i>codY</i> (F74Y)	This work
pND6	pBL26:: <i>codY</i> (E99D)	This work
pND7	pBL26:: <i>codY</i> (E103D)	This work
pND9	pBL26:: <i>codY</i> (F74L)	This work
pND11	pBL26:: <i>codY</i> (F101W)	This work
pND12	pBL26:: <i>codY</i> (P102G)	This work
<b><i>Escherichia coli</i></b>		
HB101 (pRK24)	F <sup>+</sup> <i>supE44 hsdS20(rB<sup>+</sup>mB<sup>+</sup>) recA13 ara-14 proA2 lacY1 galK2 rpsL20 xyl-5 mtl-1 (Tra<sup>+</sup> Mob<sup>+</sup> Amp<sup>r</sup> Tc<sup>r</sup>)</i>	[95]
<b><i>Bacillus subtilis</i></b>		
BS49	Contains integrated Tn916	[100]
<b><i>Clostridoides difficile</i></b>		
UK1	Ribotype 027	D. Gerding
LB-CD16	UK1 <i>codY::intron::ermB</i>	[44]
ND-CD5	UK1 <i>codY::intron::ermB</i> Tn916	This work
ND-CD6	UK1 <i>codY::intron::ermB</i> Tn916:: <i>codY</i> (E103D)	This work
ND-CD10	UK1 <i>codY::intron::ermB</i> Tn916:: <i>codY</i> <sup>+</sup>	This work
ND-CD12	UK1 <i>codY::intron::ermB</i> Tn916:: <i>codY</i> (F74Y)	This work
ND-CD13	UK1 <i>codY::intron::ermB</i> Tn916:: <i>codY</i> (E99D)	This work
ND-CD17	UK1 <i>codY::intron::ermB</i> Tn916:: <i>codY</i> (F101W)	This work
ND-CD18	UK1 <i>codY::intron::ermB</i> Tn916:: <i>codY</i> (P102G)	This work

<https://doi.org/10.1371/journal.pone.0206896.t004>

April 25, 2013 to April 1, 2016 and as number 2016-BR04 from March 31, 2016 to March 31, 2019. All researchers were required to pass annual laboratory safety training tests. In accordance with the registration, live samples of *C. difficile* strains were grown and stored in an anaerobic chamber and killed by treatment with bleach or by autoclaving before disposal. To extract DNA, RNA or protein from *C. difficile* cultures, the cells were killed in order to allow the molecules to be purified outside the chamber. Researchers handled the material outside the chamber in sealed containers and were protected by gloves and lab coats.

### Storage and sharing plan

Individual strains of *C. difficile* were stored on labeled Petri plates inside the anaerobic chamber or at -80°C in tubes containing labels that reveal their origin. The sites of storage were only accessible by researchers who had passed the laboratory safety training tests. If strains of *C. difficile* were sent to other labs as part of this project or upon requests by other scientific groups, the bacteria were only sent if the proposed recipient was experienced in working with *C. difficile* and dealing with its potential hazards. Such shipments met the requirements of UN 3373 (Biological Substance, Category B, as defined by IATA DGR 6.2 –Infectious Substances). Infection of mice by *C. difficile* was carried out under the prescribed, careful conditions described below.

### Strain and plasmid construction

Oligonucleotides used in this study are listed in [S8 Table](#). Creation by TargeTron methodology of an insertional mutation in the *codY* gene of *C. difficile* strain UK1 was described by Mooyottu et al. [44]. To complement the *codY* gene disruption, a 983-bp fragment containing the wild-type *codY* gene and its upstream region was amplified using primers oLB275 and oLB276 ([S10 Table](#)) and cloned between the *Bam*HI and *Hind*III sites of pBL26, a region homologous to the transposon Tn916. The resulting plasmid was named pND3. To create each desired point mutation in the *codY* GAF domain, two sets of primers were designed ([S10 Table](#)) to amplify overlapping portions of the *codY* gene, each of which would have the desired mutation. After amplification by Hi-Fidelity Phusion polymerase (New England Biolabs), the two PCR products were purified, mixed and amplified using the primers oLB275 and oLB276, which yielded the entire *codY* gene with the intended point mutation. Each of the mutated genes was then cloned in plasmid pBL26, generating plasmids pND5, pND6, pND7, pND11 and pND12.

Introduction into *E. coli* strain JM107 caused the plasmids derived from pBL26 to create concatemers that facilitated subsequent transformation of *B. subtilis* strain BS49, which carries Tn916 within its chromosome; the *codY* gene inserted within the chromosomal Tn916 by homologous recombination. The various *B. subtilis* strains were then mated with *C. difficile* strain LB-CD16 (*codY::intron::erm*), resulting in strains ND-CD10 (*codY::intron::erm* Tn916::*codY*<sup>+</sup>), ND-CD12 (*codY::intron::erm* Tn916::*codY* F74Y), ND-CD13 (*codY::intron::erm* Tn916::*codY* E99D), ND-CD17 (*codY::intron::erm* Tn916::*codY* F101W), and ND-CD18 (*codY::intron::erm* Tn916::*codY* P102G), each of which has a full-length copy of the *codY* gene (with or without a point mutation) and a copy interrupted by the intron. The creation of mutant strains of *C. difficile* was approved by the Institutional Biosafety Committee for Tufts University and Tufts Medical Center and was assigned the registration number 2016-BR04 for the period 03/31/16-03/31/19.

To determine the number of *codY*-containing Tn916 insertions per *C. difficile* genome, chromosomal DNA was extracted as previously described (72). As a model, DNA extracted from strain UK1 was 5-fold serially diluted (from 5 to 0.016 ng/μl) and used as a template for

quantitative PCR of *codY* (primers oND34/oND35) using the Roche SYBR Green I PCR mix and a Roche LightCycler 480 II thermocycler. Reactions were performed in triplicate in a final volume of 20  $\mu$ l using 1  $\mu$ l of the serially diluted DNA and each primer at 0.5  $\mu$ M. The threshold cycle value of each dilution was determined. In a base 10 logarithmic graph, the threshold cycle was plotted versus the dilution factor and the data were fitted to a straight line. The correlation coefficient ( $R^2$ ) for the line was 0.99 or greater. This plot was then used as a standard curve for extrapolating the relative concentration levels of the *codY* gene in strain LB-CD16 carrying Tn916 embedded with various versions of the *codY* gene. The relative concentration value was then used to generate the number of *codY* copies using the website <http://cels.uri.edu/gsc/cndna.html>. A separate standard curve was generated for the housekeeping gene *rpoA* using oligos oLB273 and oLB274. The number of *codY* copies in each sample was then normalized to that of *rpoA*. Clones that appeared to have a single full-length *codY* copy per chromosome, indicated with an asterisk in S2 Fig, were used in subsequent studies. However, strain ND-CD13 was later found to have two copies of the *codY* gene (S2 Table).

To map the sites of Tn916 insertion, two approaches were taken. First, applying the method of Hava and Camilli [101] chromosomal DNA of each mutant strain was amplified using a Tn916-specific primer (oLS9) and a partially randomized primer (ARB1); after 30 cycles, the products were purified using a QiaQuick PCR Purification Kit (Qiagen) and then subjected to a second round of PCR using a second Tn916-specific primer (oLS17) and a non-random primer (ARB2) that anneals to products produced by ARB1. After purification, the products of the second round were sequenced using the Tn916-specific primer oLS17. The second approach was a verification based on the RNA-seq results. That is, in most cases, the RNA-seq data revealed a junction between the mRNA for a chromosomal gene and one end of Tn916, thereby indicating where the transposon had inserted. In all strains except ND-CD13, only one site of Tn916 insertion was detected. In strain ND-CD13, however, three sites were found. Insertions between R20291\_0447 and \_0448 and between R20291\_1900 and \_1901 contained pSMB47 carrying the mutant version of the *codY* gene. The insertion between R20291\_0467 and \_0468 did not contain pSMB47.

### Determination of CodY stability and abundance by immunoblotting

*C. difficile* strains grown to mid-exponential phase in TY medium (6 ml) were collected by centrifugation and stored at  $-80^{\circ}\text{C}$  overnight. Cells were washed with 0.5 mL of solution A (50 mM Tris-HCl, 2 mM EDTA, 1 mM DTT), resuspended in 0.5 mL of solution A and disrupted using a Mini BeadBeater (two beatings of 30 sec each at speed 48.) The samples were stored to  $-80^{\circ}\text{C}$  for 1 hour, thawed and subjected to a second round of disruption. Unbroken cells and cell debris were removed by centrifugation at 13,000 rpm for 15 min at  $4^{\circ}\text{C}$ . The total protein concentration of the supernatant was determined using the Bio-Rad protein assay reagent; for each sample, 4  $\mu$ g of total protein was then analyzed by Western blotting as follows. Protein samples were mixed with an equal volume of 2 $\times$  SDS-PAGE-loading buffer, boiled for 3 min, loaded on 12% polyacrylamide-SDS gels, and subjected to electrophoresis at a constant voltage of 100V for 2 h. Proteins were then electro-transferred to Immobilon-P membranes (Millipore) and immunoblotted using rabbit antibody to CodY (prepared by Biodesign International (21)) diluted 5000-fold in 5% skim milk in 50 mM Tris, 150 mM NaCl, 0.1% Tween 20. Immune reactions were visualized using an enzyme-based chemiluminescence (ECL) kit (ThermoFisher).

### Animal infection conditions

*C. difficile* spores were prepared as described (45). Briefly, an overnight culture in BHIS medium was diluted in fresh medium to an optical density at 600 nm of 0.2. A 150- $\mu$ l portion

of this suspension was spread onto 5 ml of BHIS agar in each well of a six-well tissue culture dish. The dish was incubated anaerobically for 4 to 7 days to induce sporulation. The spores were washed off the plates with phosphate-buffered saline (PBS). The spore suspension was then heated at 60°C for 20 min to kill vegetative cells. The spore suspension was stored at 4°C, and the spore concentration was determined by serial dilution and plating on BHIS agar supplemented with 1% sodium taurocholate.

Four groups of 10 C57Bl/6 mice (Jackson Laboratories) were pretreated with a cocktail of five antibiotics designed to generate the following approximate daily doses: kanamycin (40 mg/kg), gentamicin (3.5 mg/kg), colistin (4.2 mg/kg), metronidazole (21.5 mg/kg), and vancomycin (4.5 mg/kg). Antibiotics were administered in sterile drinking water for 3 days, followed by 2 days of sterile drinking water without antibiotics. Mice were then given one dose of clindamycin (20 mg/kg weight) intraperitoneally one day prior to infection by oral gavage with  $10^4$  or  $10^5$  spores of *C. difficile* strains UK1 or UK1 *codY* (LB-CD16). After infection, mice were monitored daily for weight loss, diarrhea, and mortality. At a dose of  $10^5$  spores, one mouse died after exposure to strain UK1 and two mice died after exposure to LB-CD16. At a dose of  $10^4$  spores, no mice died after exposure to either strain. A weight loss of 20% required that the mice be euthanized by carbon dioxide asphyxiation followed by cervical dislocation. At the end of each experiment (7–9 days after infection), mice were euthanized using the same methods.

### Ethics statement

Animal studies followed the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health and were approved by the Tufts University Institutional Animal Care and Use Committee under the protocol #G2012-70. All researchers doing animal studies were required to pass Mandatory Animal Care and Use (MACU) training at Tufts University.

### Quantification of toxin by ELISA

*C. difficile* strains were grown in BHIS medium in an anaerobic chamber. Samples (1 ml) were collected after 24 hrs of incubation, and were adjusted to an identical  $OD_{650}$  to ensure equal concentrations of cells in each sample. Samples were then filtered and culture fluid was added to Costar 96-well microplates that had been coated with 100  $\mu$ l of anti-TcdA antibody (PCG 4.1; Novus Biologicals) (1  $\mu$ g/ml) or anti-TcdB antibody (5A8-E11; GeneTex) (10  $\mu$ g/ml) and incubated overnight in phosphate-buffered saline (PBS) at 4°C. The samples were then blocked for 1 hr with 5% (wt/wt) skim milk diluted in PBS. Standards (TcdA and TcdB purified as previously described [102]) and samples (100  $\mu$ l) were added to each well in duplicate and the microplates were incubated for 90 min at 25°C. After another set of washes, HRP-chicken anti-*C. difficile* toxin A or B (1:5,000 dilution in PBS, Gallus Immunotech) was added to the wells for 30 min at 25°C. A final set of three washes preceded the addition of TMB Microwell Peroxidase Substrate and incubation for 20 min at 25°C in the dark. The reaction was stopped with 2 N  $H_2SO_4$ . Absorbance was measured using a plate reader at 450 nm, and the ELISA was analyzed spectrophotometrically utilizing BioTek Gen5 Version 2.0 Data Analysis Software.

### Determination of the crystal structure of *C. difficile* CodY

See Supplementary Materials and Methods.



### Quantitative reverse transcription-PCR (qRT-PCR) analysis

The levels of *tcdA* and *tcdB* RNA displayed in Fig 3 were assayed in cultures of *C. difficile* grown in CDMM medium and harvested after 8 and 24 hrs (i.e., in stationary phase). DNA-free RNA (500 ng), prepared as previously described [16,103], was quantitated by absorbance ( $A_{260}$  and  $A_{260}/A_{280}$  ratio) using a NanoDrop ND-1000 spectrophotometer (Thermo Scientific) and subjected to cDNA synthesis using a QuantiTect Reverse Transcription Kit (Qiagen) following the manufacturer's recommendation. To control for contamination by DNA, mock cDNA synthesis reactions containing no reverse transcriptase were used as negative controls in subsequent amplifications. Primers for qRT-PCR were designed using the online PrimerQuest tool from Integrated DNA Technologies (<http://www.idtdna.com/Scitools/Applications/Primerquest>), and amplification efficiencies for each primer set were determined prior to use. cDNA samples were used as templates for quantitative PCR of *rpoA* (defined as that of R20291\_0096) (primers oLB273/oLB274), *tcdA* (primers oLB131/oLB132) and *tcdB* (primers oND32/oND33) using Roche SYBR Green I PCR mix and a Roche LightCycler 480 II thermocycler. Reactions were performed in a final volume of 20  $\mu$ l using 4  $\mu$ l of cDNA (25 ng) and 1  $\mu$ M of each primer. Amplification included 45 cycles of the following steps: 10 s at 95°C, 10 s at 53°C, 15 s at 72°C. Results were calculated using the comparative cycle threshold method [104], in which the amount of target mRNA is normalized to that of an internal control transcript (*rpoA*). Reactions were performed in triplicate using cDNA extracted from three biological replicates, and results are presented as the means and standard deviations of the data obtained.

To assess the impact of *codY* mutations on gene expression during growth, three independent cultures of the wild-type, the null mutant and various *codY* point mutant strains were grown in TY medium to a culture density of  $A_{600} = 0.4$ – $0.6$ . Purified RNA was analyzed by qRT-PCR using primers (S10 Table) specific for the genes encoding a peptidase (R20291\_2712), chloromuconate cycloisomerase (R20291\_1235), *ilvC* (R20291\_1413), a cell surface protein (R20291\_1698), and *glgC* (R20291\_0812), with *rpoA* as a standard for comparison.

### RNA-sequencing library construction, sequencing and analysis

RNA was extracted from two independent cultures of the wild-type strain and two independent cultures of each of the mutant *C. difficile* strains grown in TY medium to  $OD_{600} = 0.4$  to  $0.6$  (mid-exponential phase) using the RNeasy kit (Qiagen) and treated with TURBO DNA-free DNase (Ambion). Removal of DNA was considered successful if no PCR product was detected after 30 cycles of amplification using oND52/53 primers specific for genes encoding 16s rRNA. DNA-free RNA quality was assessed using a Bioanalyzer and RNA Pico Chips (both from Agilent Technologies). Only RNA samples with an RNA integrity number  $>8$  were used for library construction. rRNA was depleted from DNA-free RNA preparations using the RiboZero Magnetic kit (Gram-positive kit; Epicentre). mRNAs were fragmented using the NEB Next RNA Fragmentation Module (New England Biolabs) and further assessed using the Bioanalyzer and RNA Pico Chip. Fragmented RNA was purified and concentrated using the RNA Clean & Concentrator-5 kit (Zymo Research Corporation). First-strand cDNA was synthesized using SuperScript III reverse transcriptase (Life Technologies, Inc.); actinomycin D (8  $\mu$ g) was included in each reaction to prevent spurious second-strand synthesis. The first-strand cDNAs were purified using the RNA Clean & Concentrator-5 kit and subjected to second-strand cDNA synthesis with 13.3 units of DNA polymerase I and 3.3 units each of *E. coli* DNA ligase and RNase H using dUTP in place of dTTP in the reaction mixture. The double-stranded cDNAs were blunted using the Quick Blunting Kit (New England Biolabs) and A-tailed using Klenow fragment (3'-to-5' Exo minus), to which the universal adaptor Olj331/

Olj543 (S10 Table) was ligated in a 10:1 (adaptor/fragment) molar ratio. The libraries were purified and size-selected using AMPure XP SPRI beads (Agencourt), as directed, with elution in 1× low TE buffer [10 mM Tris·Cl (pH 8), 0.1 mM EDTA]. Quantifluor dsDNA dye (Promega) was used to quantify double-stranded cDNA and DNA. The second strand was selectively degraded using 1 unit of USER enzyme (New England Biolabs) before library enrichment and barcoding using Phusion HiFi polymerase (New England Biolabs) and oligonucleotides containing unique 6-bp barcodes. Final library size distributions were determined using a fragment analyzer before pooling samples. Samples were loaded into single lanes of a HiSeq 2500 instrument (Illumina) in the Tufts University Genomics Core and sequenced in multiplex (single-end 50-bp reads) using v3 chemistry. Reads were aligned to the reference genome of the *C. difficile* ribotype 027 strain R20291 using Burrows–Wheeler Aligner version 5.9 [105]; the genome sequence and gene annotations for this strain were obtained from RefSeq ([www.ncbi.nlm.nih.gov/refseq/](http://www.ncbi.nlm.nih.gov/refseq/)). The overall read coverage of genomic regions corresponding to features such as ORFs and rRNAs was conducted as described [51, 106]. Differential expression analysis was conducted using DESeq (50). The primary data from the RNA-seq experiments have been uploaded by the NIH Sequence Read Archive for public access as BioProject PRJNA438155, accessible at [https://www.ncbi.nlm.nih.gov/sra?linkname=bioproject\\_sra\\_all&from\\_uid=438155](https://www.ncbi.nlm.nih.gov/sra?linkname=bioproject_sra_all&from_uid=438155) with accession numbers SRX4158084–SRX4158093.

## Metabolomic analysis

Cultures of *C. difficile* strain UK1 and mutant derivatives were grown anaerobically in TY medium to mid-exponential phase ( $A_{600} = 0.5–0.6$ ). A sample of 13 ml was rapidly collected under vacuum on a nitrocellulose membrane (2 m; Millipore). The membrane was quickly washed with 5 ml of PBS, pH 7.5, to remove excess medium and non-cellular material, and then immersed in 1 ml of a mixture of acetonitrile, methanol and water (40:40:20) supplemented with 0.1 M formic acid pre-cooled at  $-20^{\circ}\text{C}$  and stored at  $-80^{\circ}\text{C}$ .

Metabolites were extracted from duplicate samples and analyzed by LC-MS as described in de Carvalho et al. [107], Pesek et al. [108] and Brinsmade et al. [30]. Data analysis was as outlined in Brinsmade et al. [30].

## Supporting information

### S1 File. Supplementary materials and methods.

(DOCX)

**S1 Fig. The quaternary structure of the GAF domain of *C. difficile* CodY.** **A.** The hexamer formed by the six molecules of the asymmetric unit of the CdCodY(1–156) crystals. The view is down the three-fold symmetry axis with the three intersecting 2-fold symmetry axes in the plane of the page. The chains are coloured A (ice blue) B (gold) C (coral) D (blue) E (pink) and F (red). The isoleucine ligands are shown as spheres with carbon, nitrogen and oxygen atoms coloured green, blue and red respectively. **B.** The dimer formed by the GAF domains of CodY from *B. subtilis*. The two chains are coloured gold and blue respectively and the isoleucine effector is shown as spheres. **C.** The A (ice blue) and B (gold) subunits from the CdCodY GAF domain hexamer shown in **A**. It is evident that these two molecules are juxtaposed in a very similar manner to the subunits in *B. subtilis* CodY, the obvious difference being that helices  $\alpha 1$  have been displaced from the AB dimer interface so that they instead pack with neighbouring dimer interfaces in the hexamer. **D.** Superposition of GAF domain dimer of *B. subtilis* CodY (white) with a ‘hybrid’ GAF domain dimer of CdCodY (ice blue) formed by substituting the  $\alpha 1$  helices of chains A and B with those from chains F (red) and C (coral) respectively.

Following least squares superposition of 238 C $\alpha$  atoms from the GAF domain dimers from the two species, the positional rms $\Delta$  is 1.6 Å. The carbon atoms of the effectors are colored green and grey for CdCodY and BsCodY respectively.

(TIF)

**S2 Fig. Quantification of *codY* copy number in the different mutant strains.** DNA was extracted from *C. difficile* wild-type and *codY* mutant strains harboring *codY* variants with single amino acid substitutions and quantified by real time PCR (qPCR). The numbers indicate different isolates of the same *codY* variant. Only clones having a single, full-length, uninterrupted copy of *codY*, indicated with an asterisk (\*), were used in subsequent studies.

(TIF)

**S3 Fig. Verification of stability of mutant CodY proteins.** Crude lysates of *C. difficile* strains carrying both a *codY* null mutation and a version of the *codY* gene with a point mutation were assayed by Western blotting using rabbit anti-CodY antibodies. Proteins of each lysate (4  $\mu$ g) were separated by SDS-PAGE. The proteins were electrotransferred and immunoblotted with a polyclonal CodY antibody. Lane 1 contains purified *B. subtilis* CodY protein. Lanes 2–10 display extracts of various *codY* point mutants. (Each one is a derivative of strain LB-CD16 (*codY::erm*) in which a point mutant form of *codY* has been integrated into the chromosome.) Lanes 11 and 12 display lysates from the strain LB-CD6 with (lane 11) or without (lane 12) the empty vector pBL26. Lane 13 displays the lysate from wild-type cells.

(TIF)

**S1 Table. X-ray data collection and refinement statistics.**

(PDF)

**S2 Table. Sites of Tn916 insertion in *codY* mutant strains.** The chromosomal sites of Tn916-*codY* insertion were determined by sequencing and by analysis of RNA-seq data. See [Materials and Methods](#) for details.

(DOCX)

**S3 Table. RPKMO values for all genes of the parental strain UK1, the *codY* null mutant and three point mutants.** Two samples were assayed and averaged for each strain.

(XLSX)

**S4 Table. Genes overexpressed (A) or underexpressed (B) >3-fold in the *codY* null mutant strain.** The average RPKMO values for two samples of strains UK1 (*codY*<sup>+</sup>) and LB-CD16 (*codY::intron::erm*) were determined (columns D and G). The ratios of the LB-CD16/UK1 averages are presented in column H.

(XLSX)

**S5 Table. Genes overexpressed (A) or underexpressed (B) >3-fold in the *codY* mutant strain ND-CD13.** The average RPKMO values for two samples of strains UK1 (*codY*<sup>+</sup>) and ND-CD13 (*codY::intron::erm* Tn916::*codY* (E99D)) were determined (columns D and G). The ratios of the ND-CD13/UK1 averages are presented in column H.

(XLSX)

**S6 Table. Genes overexpressed (A) or underexpressed (B) >3-fold in the *codY* mutant strain ND-CD17.** The average RPKMO values for two samples of strains UK1 (*codY*<sup>+</sup>) and ND-CD17 (*codY::intron::erm* Tn916::*codY* (F101W)) were determined (columns D and G). The ratios of the ND-CD17/UK1 averages are presented in column H.

(XLSX)

**S7 Table. Genes overexpressed (A) or underexpressed (B) >3-fold in the *codY* mutant strain ND-CD12.** The average RPKMO values for two samples of strains UK1 (*codY*<sup>+</sup>) and ND-CD12 (*codY::intron::erm Tn916::codY* (F74Y)) were determined (columns D and G). The ratios of the ND-CD12/UK1 averages are presented in column H.  
(XLSX)

**S8 Table. Expression of ethanolamine metabolism genes in *codY* mutant strains.** The average RPKMO values for each of the genes is shown for the wild-type and the *codY* mutant strains.  
(XLSX)

**S9 Table. Sporulation genes regulated by CodY.** CodY-repressed genes (>3-fold) in strain UK1 (annotated as R20291 genes) that were found to be Spo0A-dependent in strain JIR8074, a derivative of strain 630 (A. Shen, personal communication) are listed with their 630 analogs.  
(XLSX)

**S10 Table. List of oligonucleotides.**  
(XLS)

## Acknowledgments

We are very grateful to Drs. Aimee Shen, Rita Tamayo, Lars Barquist and Robert Fagan for helpful suggestions and for making unpublished or pre-publication information available.

## Author Contributions

**Conceptualization:** Nadine Daou, Vladimir M. Levdikov, Laurent Bouillaut, Kyu Rhee, Anthony J. Wilkinson, Xingmin Sun, Abraham L. Sonenshein.

**Data curation:** Yuanguo Wang, Vladimir M. Levdikov, Madhumitha Nandakumar, Jonathan Livny, Laurent Bouillaut, Elena Blagova, Boris R. Belitsky, Kyu Rhee, Anthony J. Wilkinson.

**Formal analysis:** Nadine Daou, Vladimir M. Levdikov, Jonathan Livny, Laurent Bouillaut, Elena Blagova, Boris R. Belitsky, Kyu Rhee, Anthony J. Wilkinson, Xingmin Sun, Abraham L. Sonenshein.

**Funding acquisition:** Kyu Rhee, Anthony J. Wilkinson, Abraham L. Sonenshein.

**Investigation:** Nadine Daou, Yuanguo Wang, Vladimir M. Levdikov, Madhumitha Nandakumar, Laurent Bouillaut, Elena Blagova, Keshan Zhang, Boris R. Belitsky, Kyu Rhee, Anthony J. Wilkinson, Xingmin Sun, Abraham L. Sonenshein.

**Methodology:** Nadine Daou, Vladimir M. Levdikov, Laurent Bouillaut, Kyu Rhee, Anthony J. Wilkinson, Xingmin Sun, Abraham L. Sonenshein.

**Project administration:** Anthony J. Wilkinson, Xingmin Sun, Abraham L. Sonenshein.

**Resources:** Abraham L. Sonenshein.

**Supervision:** Laurent Bouillaut, Kyu Rhee, Anthony J. Wilkinson, Xingmin Sun, Abraham L. Sonenshein.

**Validation:** Jonathan Livny, Anthony J. Wilkinson.

**Writing – original draft:** Laurent Bouillaut, Kyu Rhee, Anthony J. Wilkinson, Xingmin Sun, Abraham L. Sonenshein.

**Writing – review & editing:** Nadine Daou, Jonathan Livny, Laurent Bouillaut, Kyu Rhee, Anthony J. Wilkinson, Xingmin Sun, Abraham L. Sonenshein.

## References

1. Lessa FC, Mu Y, Bamberg WM, Beldavs ZG, Dumyati GK, Dunn JR, et al. Burden of Clostridium difficile infection in the United States. *N Engl J Med*. 2015; 372(9):825–34. <https://doi.org/10.1056/NEJMoa1408913> PMID: 25714160
2. Shields K, Araujo-Castillo RV, Theethira TG, Alonso CD, Kelly CP. Recurrent Clostridium difficile infection: From colonization to cure. *Anaerobe*. 2015; 34:59–73. <https://doi.org/10.1016/j.anaerobe.2015.04.012> PMID: 25930686
3. Prevention CfDca. Antibiotic resistance threats in the United State, 2013. <http://www.cdc.gov/drugresistance/threat-report-2013/pdf/ar-threats-2013-508.pdf>. 2013.
4. Bouza E. Consequences of Clostridium difficile infection: understanding the healthcare burden. *Clin Microbiol Infect*. 2012; 18 Suppl 6:5–12.
5. Lyerly DM, Lockwood DE, Richardson SH, Wilkins TD. Biological activities of toxins A and B of Clostridium difficile. *Infect Immun*. 1982; 35(3):1147–50. PMID: 7068215
6. Rosenbusch KE, Bakker D, Kuijper EJ, Smits WK. C. difficile 630Deltaerm Spo0A regulates sporulation, but does not contribute to toxin production, by direct high-affinity binding to target DNA. *PLoS One*. 2012; 7(10):e48608. <https://doi.org/10.1371/journal.pone.0048608> PMID: 23119071
7. Mackin KE, Carter GP, Howarth P, Rood JI, Lyras D. Spo0A differentially regulates toxin production in evolutionarily diverse strains of Clostridium difficile. *PLoS One*. 2013; 8(11):e79666. <https://doi.org/10.1371/journal.pone.0079666> PMID: 24236153
8. Pettit LJ, Browne HP, Yu L, Smits WK, Fagan RP, Barquist L, et al. Functional genomics reveals that Clostridium difficile Spo0A coordinates sporulation, virulence and metabolism. *BMC Genomics*. 2014; 15:160. <https://doi.org/10.1186/1471-2164-15-160> PMID: 24568651
9. Mani N, Dupuy B. Regulation of toxin synthesis in Clostridium difficile by an alternative RNA polymerase sigma factor. *Proc Natl Acad Sci U S A*. 2001; 98(10):5844–9. <https://doi.org/10.1073/pnas.101126598> PMID: 11320220
10. Girinathan BP, Monot M, Boyle D, McAllister KN, Sorg JA, Dupuy B, et al. Effect of tcdR Mutation on Sporulation in the Epidemic Clostridium difficile Strain R20291. *mSphere*. 2017; 2(1).
11. Mani N, Lyras D, Barroso L, Howarth P, Wilkins T, Rood JI, et al. Environmental response and autoregulation of Clostridium difficile TxeR, a sigma factor for toxin gene expression. *J Bacteriol*. 2002; 184(21):5971–8. <https://doi.org/10.1128/JB.184.21.5971-5978.2002> PMID: 12374831
12. Bouillaut L, Dubois T, Sonenshein AL, Dupuy B. Integration of metabolism and virulence in Clostridium difficile. *Res Microbiol*. 2015; 166(4):375–83. <https://doi.org/10.1016/j.resmic.2014.10.002> PMID: 25445566
13. El Meouche I, Peltier J, Monot M, Soutourina O, Pestel-Caron M, Dupuy B, et al. Characterization of the SigD regulon of C. difficile and its positive control of toxin production through the regulation of tcdR. *PLoS One*. 2013; 8(12):e83748. <https://doi.org/10.1371/journal.pone.0083748> PMID: 24358307
14. Antunes A, Camiade E, Monot M, Courtois E, Barbut F, Sernova NV, et al. Global transcriptional control by glucose and carbon regulator CcpA in Clostridium difficile. *Nucleic Acids Res*. 2012; 40(21):10701–18. <https://doi.org/10.1093/nar/gks864> PMID: 22989714
15. Dineen SS, McBride SM, Sonenshein AL. Integration of metabolism and virulence by Clostridium difficile CodY. *J Bacteriol*. 2010; 192(20):5350–62. <https://doi.org/10.1128/JB.00341-10> PMID: 20709897
16. Dineen SS, Villapakkam AC, Nordman JT, Sonenshein AL. Repression of Clostridium difficile toxin gene expression by CodY. *Mol Microbiol*. 2007; 66(1):206–19. <https://doi.org/10.1111/j.1365-2958.2007.05906.x> PMID: 17725558
17. Nawrocki KL, Edwards AN, Daou N, Bouillaut L, McBride SM. CodY-Dependent Regulation of Sporulation in Clostridium difficile. *J Bacteriol*. 2016; 198(15):2113–30. <https://doi.org/10.1128/JB.00220-16> PMID: 27246573
18. Guedon E, Serron P, Ehrlich SD, Renault P, Delorme C. Pleiotropic transcriptional repressor CodY senses the intracellular pool of branched-chain amino acids in Lactococcus lactis. *Mol Microbiol*. 2001; 40(5):1227–39. PMID: 11401725
19. Richardson AR, Somerville GA, Sonenshein AL. Regulating the Intersection of Metabolism and Pathogenesis in Gram-positive Bacteria. *Microbiol Spectr*. 2015; 3(3).
20. Handke LD, Shivers RP, Sonenshein AL. Interaction of Bacillus subtilis CodY with GTP. *J Bacteriol*. 2008; 190(3):798–806. <https://doi.org/10.1128/JB.01115-07> PMID: 17993518

21. Ratnayake-Lecamwasam M, Serror P, Wong KW, Sonenshein AL. *Bacillus subtilis* CodY represses early-stationary-phase genes by sensing GTP levels. *Genes Dev.* 2001; 15(9):1093–103. <https://doi.org/10.1101/gad.874201> PMID: 11331605
22. Brinsmade SR, Sonenshein AL. Dissecting complex metabolic integration provides direct genetic evidence for CodY activation by guanine nucleotides. *J Bacteriol.* 2011; 193(20):5637–48. <https://doi.org/10.1128/JB.05510-11> PMID: 21856856
23. Han AR, Kang HR, Son J, Kwon DH, Kim S, Lee WC, et al. The structure of the pleiotropic transcription regulator CodY provides insight into its GTP-sensing mechanism. *Nucleic Acids Res.* 2016; 44(19):9483–93. <https://doi.org/10.1093/nar/gkw775> PMID: 27596595
24. Kaneda T. Fatty acids in the genus *Bacillus*. I. Iso- and anteiso-fatty acids as characteristic constituents of lipids in 10 species. *J Bacteriol.* 1967; 93(3):894–903. PMID: 4960925
25. Kaneda T. Incorporation of branched-chain C6-fatty acid isomers into the related long-chain fatty acids by growing cells of *Bacillus subtilis*. *Biochemistry.* 1971; 10(2):340–7. PMID: 4992629
26. Ochi K, Kandala J, Freese E. Evidence that *Bacillus subtilis* sporulation induced by the stringent response is caused by the decrease in GTP or GDP. *J Bacteriol.* 1982; 151(2):1062–5. PMID: 6807955
27. Ochi K, Kandala JC, Freese E. Initiation of *Bacillus subtilis* sporulation by the stringent response to partial amino acid deprivation. *J Biol Chem.* 1981; 256(13):6866–75. PMID: 6113248
28. van Schaik W, Chateau A, Dillies MA, Coppee JY, Sonenshein AL, Fouet A. The global regulator CodY regulates toxin gene expression in *Bacillus anthracis* and is required for full virulence. *Infect Immun.* 2009; 77(10):4437–45. <https://doi.org/10.1128/IAI.00716-09> PMID: 19651859
29. Bennett HJ, Pearce DM, Glenn S, Taylor CM, Kuhn M, Sonenshein AL, et al. Characterization of *relA* and *codY* mutants of *Listeria monocytogenes*: identification of the CodY regulon and its role in virulence. *Mol Microbiol.* 2007; 63(5):1453–67. <https://doi.org/10.1111/j.1365-2958.2007.05597.x> PMID: 17302820
30. Brinsmade SR, Alexander EL, Livny J, Stettner AI, Segre D, Rhee KY, et al. Hierarchical expression of genes controlled by the *Bacillus subtilis* global regulatory protein CodY. *Proc Natl Acad Sci U S A.* 2014; 111(22):8227–32. <https://doi.org/10.1073/pnas.1321308111> PMID: 24843172
31. Chateau A, van Schaik W, Joseph P, Handke LD, McBride SM, Smeets FM, et al. Identification of CodY targets in *Bacillus anthracis* by genome-wide in vitro binding analysis. *J Bacteriol.* 2013; 195(6):1204–13. <https://doi.org/10.1128/JB.02041-12> PMID: 23292769
32. den Hengst CD, van Hijum SA, Geurts JM, Nauta A, Kok J, Kuipers OP. The *Lactococcus lactis* CodY regulon: identification of a conserved cis-regulatory element. *J Biol Chem.* 2005; 280(40):34332–42. <https://doi.org/10.1074/jbc.M502349200> PMID: 16040604
33. Lemos JA, Nascimento MM, Lin VK, Abranches J, Burne RA. Global regulation by (p)ppGpp and CodY in *Streptococcus* mutants. *J Bacteriol.* 2008; 190(15):5291–9. <https://doi.org/10.1128/JB.00288-08> PMID: 18539745
34. Li J, Ma M, Sarker MR, McClane BA. CodY is a global regulator of virulence-associated properties for *Clostridium perfringens* type D strain CN3718. *MBio.* 2013; 4(5):e00770–13. <https://doi.org/10.1128/mBio.00770-13> PMID: 24105766
35. Lindback T, Mols M, Basset C, Granum PE, Kuipers OP, Kovacs AT. CodY, a pleiotropic regulator, influences multicellular behaviour and efficient production of virulence factors in *Bacillus cereus*. *Environ Microbiol.* 2012; 14(8):2233–46. <https://doi.org/10.1111/j.1462-2920.2012.02766.x> PMID: 22540344
36. Lobel L, Sigal N, Borovok I, Ruppin E, Herskovits AA. Integrative genomic analysis identifies isoleucine and CodY as regulators of *Listeria monocytogenes* virulence. *PLoS Genet.* 2012; 8(9):e1002887. <https://doi.org/10.1371/journal.pgen.1002887> PMID: 22969433
37. Majerczyk CD, Dunman PM, Luong TT, Lee CY, Sadykov MR, Somerville GA, et al. Direct targets of CodY in *Staphylococcus aureus*. *J Bacteriol.* 2010; 192(11):2861–77. <https://doi.org/10.1128/JB.00220-10> PMID: 20363936
38. Malke H, Steiner K, McShan WM, Ferretti JJ. Linking the nutritional status of *Streptococcus pyogenes* to alteration of transcriptional gene expression: the action of CodY and RelA. *Int J Med Microbiol.* 2006; 296(4–5):259–75. <https://doi.org/10.1016/j.ijmm.2005.11.008> PMID: 16531115
39. Montgomery CP, Boyle-Vavra S, Roux A, Ebine K, Sonenshein AL, Daum RS. CodY deletion enhances in vivo virulence of community-associated methicillin-resistant *Staphylococcus aureus* clone USA300. *Infect Immun.* 2012; 80(7):2382–9. <https://doi.org/10.1128/IAI.06172-11> PMID: 22526672

40. Pohl K, Francois P, Stenz L, Schlink F, Geiger T, Herbert S, et al. CodY in *Staphylococcus aureus*: a regulatory link between metabolism and virulence gene expression. *J Bacteriol.* 2009; 191(9):2953–63. <https://doi.org/10.1128/JB.01492-08> PMID: 19251851
41. Rivera FE, Miller HK, Kolar SL, Stevens SM Jr., Shaw LN. The impact of CodY on virulence determinant production in community-associated methicillin-resistant *Staphylococcus aureus*. *Proteomics.* 2012; 12(2):263–8. <https://doi.org/10.1002/pmic.201100298> PMID: 22106056
42. Waters NR, Samuels DJ, Behera RK, Livny J, Rhee KY, Sadykov MR, et al. A spectrum of CodY activities drives metabolic reorganization and virulence gene expression in *Staphylococcus aureus*. *Mol Microbiol.* 2016; 101(3):495–514. <https://doi.org/10.1111/mmi.13404> PMID: 27116338
43. Zhang Z, Dahlsten E, Korkeala H, Lindstrom M. Positive regulation of botulinum neurotoxin gene expression by CodY in *Clostridium botulinum* ATCC 3502. *Appl Environ Microbiol.* 2014; 80(24):7651–8. <https://doi.org/10.1128/AEM.02838-14> PMID: 25281376
44. Mooyottu S, Kollanoor-Johny A, Flock G, Bouillaut L, Upadhyay A, Sonenshein AL, et al. Carvacrol and trans-cinnamaldehyde reduce *Clostridium difficile* toxin production and cytotoxicity in vitro. *Int J Mol Sci.* 2014; 15(3):4415–30. <https://doi.org/10.3390/ijms15034415> PMID: 24625665
45. Sorg JA, Sonenshein AL. Inhibiting the initiation of *Clostridium difficile* spore germination using analogs of chenodeoxycholic acid, a bile acid. *J Bacteriol.* 2010; 192(19):4983–90. <https://doi.org/10.1128/JB.00610-10> PMID: 20675492
46. Chen X, Katchar K, Goldsmith JD, Nanthakumar N, Cheknis A, Gerding DN, et al. A mouse model of *Clostridium difficile*-associated disease. *Gastroenterology.* 2008; 135(6):1984–92. <https://doi.org/10.1053/j.gastro.2008.09.002> PMID: 18848941
47. Levdivikov VM, Blagova E, Colledge VL, Lebedev AA, Williamson DC, Sonenshein AL, et al. Structural rearrangement accompanying ligand binding in the GAF domain of CodY from *Bacillus subtilis*. *J Mol Biol.* 2009; 390(5):1007–18. <https://doi.org/10.1016/j.jmb.2009.05.077> PMID: 19500589
48. Levdivikov VM, Blagova E, Young VL, Belitsky BR, Lebedev A, Sonenshein AL, et al. Structure of the Branched-chain Amino Acid and GTP-sensing Global Regulator, CodY, from *Bacillus subtilis*. *J Biol Chem.* 2017; 292(7):2714–28. <https://doi.org/10.1074/jbc.M116.754309> PMID: 28011634
49. Levdivikov VM, Blagova E, Joseph P, Sonenshein AL, Wilkinson AJ. The structure of CodY, a GTP- and isoleucine-responsive regulator of stationary phase and virulence in gram-positive bacteria. *J Biol Chem.* 2006; 281(16):11366–73. <https://doi.org/10.1074/jbc.M513015200> PMID: 16488888
50. Anders S, Huber W. Differential expression analysis for sequence count data. *Genome Biol.* 2010; 11(10):R106. <https://doi.org/10.1186/gb-2010-11-10-r106> PMID: 20979621
51. Mandlik A, Livny J, Robins WP, Ritchie JM, Mekalanos JJ, Waldor MK. RNA-Seq-based monitoring of infection-linked changes in *Vibrio cholerae* gene expression. *Cell Host Microbe.* 2011; 10(2):165–74. <https://doi.org/10.1016/j.chom.2011.07.007> PMID: 21843873
52. Janoir C. Virulence factors of *Clostridium difficile* and their role during infection. *Anaerobe.* 2016; 37:13–24. <https://doi.org/10.1016/j.anaerobe.2015.10.009> PMID: 26596863
53. Carter GP, Lyras D, Allen DL, Mackin KE, Howarth PM, O'Connor JR, et al. Binary toxin production in *Clostridium difficile* is regulated by CdtR, a LytTR family response regulator. *J Bacteriol.* 2007; 189(20):7290–301. <https://doi.org/10.1128/JB.00731-07> PMID: 17693517
54. Bilverstone TW, Kinsmore NL, Minton NP, Kuehne SA. Development of *Clostridium difficile* R20291DeltaPaLoc model strains and in vitro methodologies reveals CdtR is required for the production of CDT to cytotoxic levels. *Anaerobe.* 2017.
55. Lyon SA, Hutton ML, Rood JI, Cheung JK, Lyras D. CdtR Regulates TcdA and TcdB Production in *Clostridium difficile*. *PLoS Pathog.* 2016; 12(7):e1005758. <https://doi.org/10.1371/journal.ppat.1005758> PMID: 27414650
56. McBride SM, Sonenshein AL. The *dlt* operon confers resistance to cationic antimicrobial peptides in *Clostridium difficile*. *Microbiology.* 2011; 157(Pt 5):1457–65. <https://doi.org/10.1099/mic.0.045997-0> PMID: 21330441
57. Barketi-Klai A, Hoys S, Lambert-Bordes S, Collignon A, Kansau I. Role of fibronectin-binding protein A in *Clostridium difficile* intestinal colonization. *J Med Microbiol.* 2011; 60(Pt 8):1155–61. <https://doi.org/10.1099/jmm.0.029553-0> PMID: 21349990
58. Hennequin C, Porcheray F, Waligora-Dupriet A, Collignon A, Barc M, Bourlioux P, et al. GroEL (Hsp60) of *Clostridium difficile* is involved in cell adherence. *Microbiology.* 2001; 147(Pt 1):87–96. <https://doi.org/10.1099/00221287-147-1-87> PMID: 11160803
59. Tasteyre A, Barc MC, Collignon A, Boureau H, Karjalainen T. Role of FliC and FliD flagellar proteins of *Clostridium difficile* in adherence and gut colonization. *Infect Immun.* 2001; 69(12):7937–40. <https://doi.org/10.1128/IAI.69.12.7937-7940.2001> PMID: 11705981

60. Ho TD, Williams KB, Chen Y, Helm RF, Popham DL, Ellermeier CD. Clostridium difficile extracytoplasmic function sigma factor sigmaV regulates lysozyme resistance and is necessary for pathogenesis in the hamster model of infection. *Infect Immun*. 2014; 82(6):2345–55. <https://doi.org/10.1128/IAI.01483-13> PMID: 24664503
61. Cafardi V, Biagini M, Martinelli M, Leuzzi R, Rubino JT, Cantini F, et al. Identification of a novel zinc metalloprotease through a global analysis of Clostridium difficile extracellular proteins. *PLoS One*. 2013; 8(11):e81306. <https://doi.org/10.1371/journal.pone.0081306> PMID: 24303041
62. Oliveira Paiva AM, Friggen AH, Hossein-Javaheri S, Smits WK. The Signal Sequence of the Abundant Extracellular Metalloprotease PPEP-1 Can Be Used to Secrete Synthetic Reporter Proteins in Clostridium difficile. *ACS Synth Biol*. 2016.
63. Hensbergen PJ, Klychnikov OI, Bakker D, Dragan I, Kelly ML, Minton NP, et al. Clostridium difficile secreted Pro-Pro endopeptidase PPEP-1 (ZMP1/CD2830) modulates adhesion through cleavage of the collagen binding protein CD2831. *FEBS Lett*. 2015; 589(24 Pt B):3952–8.
64. Kovacs-Simon A, Leuzzi R, Kasendra M, Minton N, Titball RW, Michell SL. Lipoprotein CD0873 is a novel adhesin of Clostridium difficile. *J Infect Dis*. 2014; 210(2):274–84. <https://doi.org/10.1093/infdis/jiu070> PMID: 24482399
65. Fagan RP, Janoir C, Collignon A, Mastrantonio P, Poxton IR, Fairweather NF. A proposed nomenclature for cell wall proteins of Clostridium difficile. *J Med Microbiol*. 2011; 60(Pt 8):1225–8. <https://doi.org/10.1099/jmm.0.028472-0> PMID: 21252271
66. Willing SE, Candela T, Shaw HA, Seager Z, Mesnage S, Fagan RP, et al. Clostridium difficile surface proteins are anchored to the cell wall using CWB2 motifs that recognise the anionic polymer PSII. *Mol Microbiol*. 2015; 96(3):596–608. <https://doi.org/10.1111/mmi.12958> PMID: 25649385
67. Chu M, Mallozzi MJ, Roxas BP, Bertolo L, Monteiro MA, Agellon A, et al. A Clostridium difficile Cell Wall Glycopolymer Locus Influences Bacterial Shape, Polysaccharide Production and Virulence. *PLoS Pathog*. 2016; 12(10):e1005946. <https://doi.org/10.1371/journal.ppat.1005946> PMID: 27741317
68. Purcell EB, McKee RW, McBride SM, Waters CM, Tamayo R. Cyclic diguanylate inversely regulates motility and aggregation in Clostridium difficile. *J Bacteriol*. 2012; 194(13):3307–16. <https://doi.org/10.1128/JB.00100-12> PMID: 22522894
69. McKee RW, Mangalea MR, Purcell EB, Borchardt EK, Tamayo R. The second messenger cyclic Di-GMP regulates Clostridium difficile toxin production by controlling expression of sigD. *J Bacteriol*. 2013; 195(22):5174–85. <https://doi.org/10.1128/JB.00501-13> PMID: 24039264
70. Bordeleau E, Fortier LC, Malouin F, Burrus V. c-di-GMP turn-over in Clostridium difficile is controlled by a plethora of diguanylate cyclases and phosphodiesterases. *PLoS Genet*. 2011; 7(3):e1002039. <https://doi.org/10.1371/journal.pgen.1002039> PMID: 21483756
71. Purcell EB, McKee RW, Courson DS, Garrett EM, McBride SM, Cheney RE, et al. A nutrient-regulated cyclic diguanylate phosphodiesterase controls Clostridium difficile biofilm and toxin production during stationary phase. *Infect Immun*. 2017.
72. Nawrocki KL, Wetzel D, Jones JB, Woods EC, McBride SM. Ethanolamine is a valuable nutrient source that impacts Clostridium difficile pathogenesis. *Environ Microbiol*. 2018; 20(4):1419–35. <https://doi.org/10.1111/1462-2920.14048> PMID: 29349925
73. Belitsky BR, Brinsmade SR, Sonenshein AL. Intermediate Levels of Bacillus subtilis CodY Activity Are Required for Derepression of the Branched-Chain Amino Acid Permease, BraB. *PLoS Genet*. 2015; 11(10):e1005600. <https://doi.org/10.1371/journal.pgen.1005600> PMID: 26473603
74. Stragier P, Losick R. Cascades of sigma factors revisited. *Mol Microbiol*. 1990; 4(11):1801–6. PMID: 2127951
75. Piggot PJ, Hilbert DW. Sporulation of Bacillus subtilis. *Curr Opin Microbiol*. 2004; 7(6):579–86. <https://doi.org/10.1016/j.mib.2004.10.001> PMID: 15556029
76. Saujet L, Monot M, Dupuy B, Soutourina O, Martin-Verstraete I. The key sigma factor of transition phase, SigH, controls sporulation, metabolism, and virulence factor expression in Clostridium difficile. *J Bacteriol*. 2011; 193(13):3186–96. <https://doi.org/10.1128/JB.00272-11> PMID: 21572003
77. Predich M, Nair G, Smith I. Bacillus subtilis early sporulation genes kinA, spo0F, and spo0A are transcribed by the RNA polymerase containing sigma H. *J Bacteriol*. 1992; 174(9):2771–8. PMID: 1569009
78. Siranosian KJ, Grossman AD. Activation of spo0A transcription by sigma H is necessary for sporulation but not for competence in Bacillus subtilis. *J Bacteriol*. 1994; 176(12):3812–5. PMID: 8206860
79. Wu JJ, Piggot PJ, Tatti KM, Moran CP Jr. Transcription of the Bacillus subtilis spoIIA locus. *Gene*. 1991; 101(1):113–6. PMID: 1905664



80. Underwood S, Guan S, Vijayasubhash V, Baines SD, Graham L, Lewis RJ, et al. Characterization of the sporulation initiation pathway of *Clostridium difficile* and its role in toxin production. *J Bacteriol*. 2009; 191(23):7296–305. <https://doi.org/10.1128/JB.00882-09> PMID: 19783633
81. Darkoh C, DuPont HL, Norris SJ, Kaplan HB. Toxin synthesis by *Clostridium difficile* is regulated through quorum signaling. *MBio*. 2015; 6(2):e02569. <https://doi.org/10.1128/mBio.02569-14> PMID: 25714717
82. Fimlaid KA, Bond JP, Schutz KC, Putnam EE, Leung JM, Lawley TD, et al. Global analysis of the sporulation pathway of *Clostridium difficile*. *PLoS Genet*. 2013; 9(8):e1003660. <https://doi.org/10.1371/journal.pgen.1003660> PMID: 23950727
83. Pishdadian K, Fimlaid KA, Shen A. SpoIIID-mediated regulation of sigmaK function during *Clostridium difficile* sporulation. *Mol Microbiol*. 2015; 95(2):189–208. <https://doi.org/10.1111/mmi.12856> PMID: 25393584
84. Dembek M, Barquist L, Boinett CJ, Cain AK, Mayho M, Lawley TD, et al. High-throughput analysis of gene essentiality and sporulation in *Clostridium difficile*. *MBio*. 2015; 6(2):e02383. <https://doi.org/10.1128/mBio.02383-14> PMID: 25714712
85. Smits WK. Hype or hypervirulence: a reflection on problematic *C. difficile* strains. *Virulence*. 2013; 4(7):592–6. <https://doi.org/10.4161/viru.26297> PMID: 24060961
86. Majerczyk CD, Sadykov MR, Luong TT, Lee C, Somerville GA, Sonenshein AL. *Staphylococcus aureus* CodY negatively regulates virulence gene expression. *J Bacteriol*. 2008; 190(7):2257–65. <https://doi.org/10.1128/JB.01545-07> PMID: 18156263
87. Ibberson CB, Jones CL, Singh S, Wise MC, Hart ME, Zurawski DV, et al. *Staphylococcus aureus* hyaluronidase is a CodY-regulated virulence factor. *Infect Immun*. 2014; 82(10):4253–64. <https://doi.org/10.1128/IAI.01710-14> PMID: 25069977
88. Dubois T, Perchat S, Verplaetse E, Gominet M, Lemy C, Aumont-Nicaise M, et al. Activity of the *Bacillus thuringiensis* NprR-NprX cell-cell communication system is co-ordinated to the physiological stage through a complex transcriptional regulation. *Mol Microbiol*. 2013; 88(1):48–63. <https://doi.org/10.1111/mmi.12168> PMID: 23388036
89. Ihekweba AE, Mura I, Walshaw J, Peck MW, Barker GC. An Integrative Approach to Computational Modelling of the Gene Regulatory Network Controlling *Clostridium botulinum* Type A1 Toxin Production. *PLoS Comput Biol*. 2016; 12(11):e1005205. <https://doi.org/10.1371/journal.pcbi.1005205> PMID: 27855161
90. Kreth J, Chen Z, Ferretti J, Malke H. Counteractive balancing of transcriptome expression involving CodY and CovRS in *Streptococcus pyogenes*. *J Bacteriol*. 2011; 193(16):4153–65. <https://doi.org/10.1128/JB.00061-11> PMID: 21705595
91. Feng L, Zhu J, Chang H, Gao X, Gao C, Wei X, et al. The CodY regulator is essential for virulence in *Streptococcus suis* serotype 2. *Sci Rep*. 2016; 6:21241. <https://doi.org/10.1038/srep21241> PMID: 26883762
92. Li J, Freedman JC, Evans DR, McClane BA. CodY Promotes Sporulation and Enterotoxin Production by *Clostridium perfringens* Type A Strain SM101. *Infect Immun*. 2017; 85(3).
93. Edwards AN, Tamayo R, McBride SM. A novel regulator controls *Clostridium difficile* sporulation, motility and toxin production. *Mol Microbiol*. 2016; 100(6):954–71. <https://doi.org/10.1111/mmi.13361> PMID: 26915493
94. Dupuy B, Sonenshein AL. Regulated transcription of *Clostridium difficile* toxin genes. *Mol Microbiol*. 1998; 27(1):107–20. PMID: 9466260
95. Bouillaut L, McBride SM, Sorg JA. Genetic manipulation of *Clostridium difficile*. *Curr Protoc Microbiol*. 2011; Chapter 9:Unit 9A 2.
96. Cartman ST, Minton NP. A mariner-based transposon system for in vivo random mutagenesis of *Clostridium difficile*. *Appl Environ Microbiol*. 2010; 76(4):1103–9. <https://doi.org/10.1128/AEM.02525-09> PMID: 20023081
97. Miller JH. *Experiments in molecular genetics*. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory; 1972. xvi, 466 p. p.
98. Bouillaut L, Self WT, Sonenshein AL. Proline-dependent regulation of *Clostridium difficile* Stickland metabolism. *J Bacteriol*. 2013; 195(4):844–54. <https://doi.org/10.1128/JB.01492-12> PMID: 23222730
99. Manganelli R, Provvedi R, Berneri C, Oggioni MR, Pozzi G. Insertion vectors for construction of recombinant conjugative transposons in *Bacillus subtilis* and *Enterococcus faecalis*. *FEMS Microbiol Lett*. 1998; 168(2):259–68. <https://doi.org/10.1111/j.1574-6968.1998.tb13282.x> PMID: 9835037
100. Mullany P, Wilks M, Tabaqchali S. Transfer of Tn916 and Tn916 delta E into *Clostridium difficile*: demonstration of a hot-spot for these elements in the *C. difficile* genome. *FEMS Microbiol Lett*. 1991; 63(2–3):191–4. PMID: 1647998

101. Hava DL, Camilli A. Large-scale identification of serotype 4 *Streptococcus pneumoniae* virulence factors. *Mol Microbiol.* 2002; 45(5):1389–406. PMID: [12207705](#)
102. Yang G, Zhou B, Wang J, He X, Sun X, Nie W, et al. Expression of recombinant *Clostridium difficile* toxin A and B in *Bacillus megaterium*. *BMC Microbiol.* 2008; 8:192. <https://doi.org/10.1186/1471-2180-8-192> PMID: [18990232](#)
103. McBride SM, Sonenshein AL. Identification of a genetic locus responsible for antimicrobial peptide resistance in *Clostridium difficile*. *Infect Immun.* 2011; 79(1):167–76. <https://doi.org/10.1128/IAI.00731-10> PMID: [20974818](#)
104. Schmittgen TD, Livak KJ. Analyzing real-time PCR data by the comparative C(T) method. *Nat Protoc.* 2008; 3(6):1101–8. PMID: [18546601](#)
105. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2010; 26(5):589–95. <https://doi.org/10.1093/bioinformatics/btp698> PMID: [20080505](#)
106. Livny J, Zhou X, Mandlik A, Hubbard T, Davis BM, Waldor MK. Comparative RNA-Seq based dissection of the regulatory networks and environmental stimuli underlying *Vibrio parahaemolyticus* gene expression during infection. *Nucleic Acids Res.* 2014; 42(19):12212–23. <https://doi.org/10.1093/nar/gku891> PMID: [25262354](#)
107. de Carvalho LP, Fischer SM, Marrero J, Nathan C, Ehrt S, Rhee KY. Metabolomics of *Mycobacterium tuberculosis* reveals compartmentalized co-catabolism of carbon substrates. *Chem Biol.* 2010; 17(10):1122–31. <https://doi.org/10.1016/j.chembiol.2010.08.009> PMID: [21035735](#)
108. Pesek JJ, Matyska MT, Loo JA, Fischer SM, Sana TR. Analysis of hydrophilic metabolites in physiological fluids by HPLC-MS using a silica hydride-based stationary phase. *J Sep Sci.* 2009; 32(13):2200–8. <https://doi.org/10.1002/jssc.200900270> PMID: [19569099](#)