communications biology

Article

https://doi.org/10.1038/s42003-024-06617-4

Clostridium septicum manifests a bile salt germinant response mediated by *Clostridioides difficile csp* gene orthologs

Rongji Sum [1,2], Sylvester Jian Ming Lim [1,2], Ajitha Sundaresan [1,2], Sudipta Samanta[1], Muthukaruppan Swaminathan[1], Wayne Low[1], Madhumitha Ayyappan[1,3,4], Ting Wei Lim[1,3,4], Marvin Dragon Choo[1,3,4], Gabriel Junming Huang[1] & Ian Cheong [1,2]

Clostridium septicum infections are highly predictive of certain malignancies in human patients. To initiate infections, *C. septicum* spores must first germinate and regain vegetative growth. Yet, what triggers the germination of *C. septicum* spores is still unknown. Here, we observe that *C. septicum* germinates in response to specific bile salts. Putative bile salt recognition genes are identified in *C. septicum* based on their similarity in sequence and organization to bile salt-responsive *csp* genes in *Clostridioides difficile*. Inactivating two of these *csp* orthologs (*cspC-82* and *cspC-1718*) results in mutant spores that no longer germinate in the presence of their respective cognate bile salts. Additionally, inactivating the putative *cspBA* or *sleC* genes in *C. septicum* abrogates the germination response to all bile salt germinants, suggesting that both act at a convergent point downstream of *cspC-82* and *cspC-1718*. Molecular dynamics simulations show that both CspC-82 and CspC-1718 bear a strong structural congruence with *C. difficile*'s CspC. The existence of functional bile salt germination sensors in *C. septicum* may be relevant to the association between infection and malignancy.

Clostridium septicum infections are infrequent in human patients, accounting for only 1.3% of clostridial sepsis cases[1]. Yet, in a seminal review of 162 cases[1], the vast majority (80.8%) of *C. septicum*-infected patients were found to have an accompanying malignancy. Subsequent studies have also reported comparable figures (78-80%) for the strong association between *C. septicum* infections and cancer[2,3]. In contrast, only 4% of *Clostridium perfringens* infections are linked to an existing cancer. *C. septicum*-linked cancers are usually of colorectal (34%) or hematological (40.1%) origin[1]. In many cases (and especially for colorectal cancer), these cancers were unknown and only discovered after the infection. In short, *C. septicum* infections may be exceedingly rare, but its presence in a patient is highly predictive of a co-existing malignancy.

Besides its strong association with cancer, *C. septicum* also stands apart from the other pathogenic clostridia in how infections present and progress. *C. septicum* is the leading cause of spontaneous gas gangrene, which is highly lethal and arises without prior injury[3]. This lethality is due in large part to *C. septicum*'s α-toxin, which forms pores on plasma membranes, leading to death by pyroptosis[4,5]. Left untreated, infections progress swiftly into fulminant gas gangrene and myonecrosis, with mortality approaching 100%[1]. This is distinct from other clostridial infections, where injury is commonly the precursor and cause of infection[6]. Interestingly, when *C. septicum* is isolated from a patient, it tends to be the sole microorganism isolated, whereas most other clostridial infections result in mixed cultures[6]. This suggests that *C. septicum*'s host tropism may differ from other pathogenic clostridia.

There has been no explanation for why *C. septicum* infections are so different from the other pathogenic clostridia, and especially why they are so closely linked with malignancy. Since a clostridial spore can only colonize an environment that supports its germination, we reasoned that studying the germination triggers of *C. septicum* could provide clues as to why *C. septicum* preferentially infects cancer patients.

Many of the pathogenic clostridial species, which include *C. septicum, C. perfringens* and *C. botulinum*, share a common ancestry and form part of the 16 S rRNA-based phylogenetic group called cluster I[7]. Certain species (e.g. *Clostridioides difficile* and *Paraclostridium bifermentans*) which were formerly identified as *Clostridium* species were later renamed because they

[1]Temasek Life Sciences Laboratory, Singapore, Singapore. [2]Department of Biological Sciences, National University of Singapore, Singapore, Singapore. [3]NUS High School of Mathematics and Sciences, Singapore, Singapore. [4]These authors contributed equally: Madhumitha Ayyappan, Ting Wei Lim, Marvin Dragon Choo. e-mail: ian@tll.org.sg

were excluded from cluster I by phylogenetic distance[8]. Despite this, germination studies of *C. difficile* and *P. bifermentans* are as highly relevant as their cluster I neighbors for understanding the diversity of germination machinery among the sporulating anaerobic bacteria[7–12,11–14] From these studies, the clostridial germinant vocabulary spans a variety of L-amino acids[11–13,15], inorganic minerals[12,16–18] and organic molecules[11,15,19,20]. Cognate receptors that sense these germinants respond to their presence by activating the downstream germination signaling cascade[21,22]. For example, in *C. perfringens*, this germinant sensing function is played by GerKC, a germinant receptor which senses amino acid germinants and consequently activates the Csp (subtilisin-like pseudoprotease) gene products: CspA, CspB and CspC. These activated Csp proteases then proceed to activate pro-SleC by cleaving its N-terminal inhibitory pro-peptide, causing the release of catalytically active SleC[12,23–25]. Next, SleC degrades the cortex, leading to calcium-DPA release, core rehydration and eventual spore outgrowth[21,26].

Recently, bile salts have been implicated in shaping the composition of the gut microbiome[27–29]. Despite the evidence that the presence of specific bile acids can increase the germination frequency of many sporulating bacteria[30], *C. difficile* is currently the only one with a characterized bile salt-responsive germination sensor. This gene, *cspC*, was first implicated in bile salt signaling through a mutagenesis screen and further identified as a bile salt germinant receptor through an altered germinant specificity screen[31]. Guided by CspC's crystal structure, further functional studies identified mutations that diminished CspC's sensitivity for taurocholate and other co-germinants, hence establishing the role of CspC in taurocholate-mediated germination[32]. A subsequent study proposed that CspC and CspA inhibited CspB when associated with it, and that these inhibitory effects were released when germinants and co-germinants interacted with CspC and CspA respectively[33]. Unhindered CspB would then process pro-SleC to its active SleC form in a manner similar to *C. perfringens*, leading to spore outgrowth[22,33–35].

In this current study, an unbiased screen of single factor germinants for *C. septicum* revealed that sodium deoxycholate, sodium chenodeoxycholate and sodium glycocholate are potent germinants. To identify the genes responsible for *C. septicum*'s bile salt-germination, genes with high similarity to *C. difficile*'s *cspC* were first identified as candidate bile salt-sensors. Mutational inactivation of two of these *cspC* orthologs abolished *C. septicum*'s bile salt germination response to distinct sets of bile salts. Further, *cspBA* and *sleC* mutant spores were unresponsive to all three bile salt germinants, which is consistent with both proteins acting downstream of the *cspC* orthologs. Finally, structural modeling and molecular dynamics simulations showed high congruence between *C. difficile*'s CspC and both CspC orthologs in *C. septicum*. The striking similarity between the bile salt-germinant sensing mechanisms of *C. septicum* and *C. difficile* raises the possibility that other gut bacteria may use the same pathway.

Results

Characterization of C. septicum germination

To identify potential germinants, forty-eight compounds known to affect the growth, sporulation or germination of various clostridia (Supplementary Table T1) were individually tested for germinant activity as measured by a decrease in optical density of the spores over time. Of the screened compounds, only the bile salt sodium deoxycholate (DCA) triggered a significant OD drop (Supplementary Fig. S1a). The other two bile salts in the screen, sodium cholate (CA) and sodium taurocholate (TCA), did not elicit germination, suggesting that DCA's putative germinant activity was specific. In a follow-up experiment, dormant phase bright spores incubated with DCA turned phase dark, confirming that DCA was indeed a germinant (Supplementary Fig. S1b). Spores incubated with CA remained phase bright, supporting the initial screen results. Intriguingly, none of the L-amino acids in the screen were identified as germinants, even though L-amino acids are the dominant class of compounds acting either as germinants or co-germinants in other *Bacillus* and *Clostridium* spores[9,36]. Beyond single amino acids, the combination of L-asparagine, D-glucose, D-fructose, and K^+ (AGFK), which is a well-known potent germinant for *B. subtilis* and *C. perfringens* spores[12,37], also failed to germinate *C. septicum* spores (Supplementary Fig. S2).

Since the physical properties of bile salts are affected by pH, DCA's germinant activity was profiled from pH 6 to 8 (Supplementary Fig. S3a). *C. septicum* spores were observed to germinate most rapidly in the range of pH 6 to 7. However, a time-dependent OD increase after 100 min was also observed at pH 6 and 6.5. This artifact, possibly caused by complexation between spore-released Ca^{2+} and DCA[38], was unrelated to germination since all germinated spores remained phase dark after 100 min. This artifact was investigated by incubating Ca^{2+}, DPA and DCA in sodium phosphate buffer at various pHs (Supplementary Fig. S3b). A significant increase in OD was observed at pH 6.0 but not at pH 6.5 or pH 7.0. This pH-dependent effect was specific to DCA and not observed with either CA or GCA. To avoid this artifact in future experiments, all absorbance assays from this point were performed at pH 7.0.

We wondered if *C. septicum* responded to other bile salts besides DCA. When tested against a panel of bile salts, *C. septicum* spores germinated not just to DCA, but also to sodium chenodeoxycholate (CDCA) and sodium glycocholate (GCA) (Fig. 1a, Supplementary Fig. S4). GCA induced germination most rapidly, followed by DCA then CDCA. Sodium glycodeoxycholate (GDCA) induced weak partial germination at the highest concentrations, and hence we considered it a borderline germinant (Fig. 1a, Supplementary Fig. S4). No germination was observed with the other bile salts tested - sodium cholate (CA), sodium taurocholate (TCA) and sodium taurodeoxycholate (TDCA). Spore germination in these experiments was confirmed with phase contrast microscopy (Fig. 1b).

When spores were incubated with the various bile salts in the panel and plated on LB agar, the top germinators in liquid media (GCA, CDCA, DCA, GDCA) also produced the most colonies by rank order (Supplementary Fig. S5). Only GCA, CDCA and GDCA however were significantly higher than the negative control. Although CA, TCA and TDCA did not produce an OD decrease with the germination assay, they generated slightly higher numbers of colonies on LB agar than the control without bile salt, albeit without statistical significance. In this study, we focused on GCA, CDCA and DCA for subsequent experiments but we also note that solid growth agar may modulate the germinant effects of bile salts on *C. septicum*. All three germinants behaved somewhat differently in terms of germination extent and velocity as bile salt concentration was varied. Notably, the EC_{50} values for these germinants are within the physiological concentrations found in the intestinal lumen, and are also in the ballpark of taurocholate concentrations found to germinate *C. difficile* spores (Fig. 1c)[31,39]. Interestingly, germination extent was saturated at 80% for CDCA, compared to 100% for both DCA and GCA. All spores were observed to be either phase bright or phase dark, reflecting that CDCA germinated ~80% of spores in that experiment as opposed to a partial germination of all spores. Next, the maximal germination (V_{max}) velocity was plotted against concentration for each bile salt (Fig. 1d). All three V_{max} curves plateaued at different points in the concentration range. CDCA saturated early, with its V_{max} maxing out at a low value after 6.25 mM. In contrast, the V_{max} of GCA exceeded both DCA and CDCA after 12.5 mM, and had not plateaued by 50 mM, which was the highest concentration tested. DCA's V_{max} response straddled both GCA and CDCA and plateaued at around 12.5 to 50 mM. What is clear is that all three germinants have distinguishable germination kinetics. Notably, in comparison to GCA and DCA, CDCA exhibited slow kinetics, yet non-inferior potency.

Although we had previously demonstrated that single amino acids were insufficient to trigger germination of *C. septicum* spores, amino acids can act as potent co-germinants for other sporulating bacteria[11]. To investigate this possibility for *C. septicum* spores, we screened an admixture of 20 amino acids for their ability to potentiate germination with suboptimal concentrations of CDCA, DCA or GCA. At these concentrations where germination was minimally induced, significant potentiation was observed with GCA, but not with CDCA or DCA (Supplementary Fig 6a). We further screened each of the 20 amino acids with the same suboptimal concentration of GCA (3.13 mM) and discovered that arginine produced the strongest

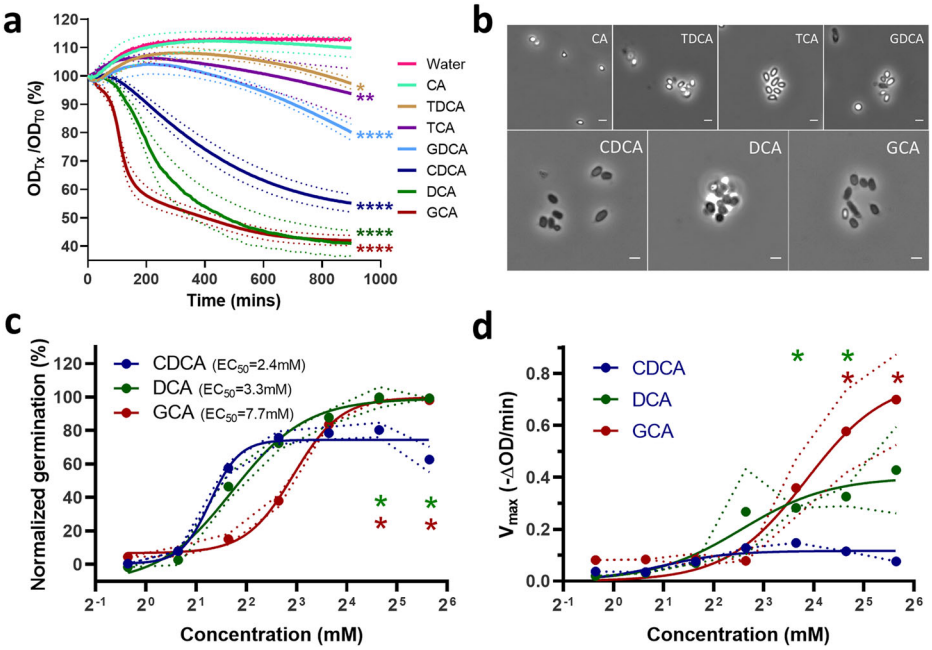

Fig. 1 | *C. septicum* germinates robustly in the presence of bile salts. a Germination measured by optical density (OD) is shown. OD values were normalized to the starting OD. Purified wild-type spores were incubated with 25 mM of each bile salt. Germination, as indicated by decreasing OD values. was observed with CDCA, DCA and GCA. **b** Representative phase contrast images of the spores at the end of the OD assay in **a** are shown. Spores treated with CDCA, DCA and GCA were phase dark, while spores treated with CA, GDCA, TCA and TDCA remained phase bright. Scale bar, 2 µm. **c** Germination response curves were plotted for the three bile salt germinants. Germination was measured at 900 min post-exposure to increasing bile salt concentrations. CDCA was the most potent germinant, followed by DCA and GCA. **d** Maximal germination velocity curves were plotted for the three bile salt germinants. For (**a**, **c**, **d**) graphed data represent the average at least 3 biologically independent experiments using at least two different spore preparations. For all bile salts, $n = 3$ except for GCA, where $n = 5$. Error envelopes (dashed lines) represent the standard error of the mean. For (**a**), statistical significance relative to water was determined by one-way ANOVA and Dunnett's test. For (**c**, **d**), statistical significance relative to CDCA was determined by unpaired *t* test between groups at each time point. ****$p < 0.0001$, ***$p < 0.001$, **$p < 0.01$ *$p < 0.05$.

potentiation while glutamic acid, asparagine and aspartic acid had relatively weaker effects (Supplementary Fig 6b).

The effect of heat activation on bile salt-induced germination was also investigated (Supplementary Fig. S7). Two trends were noted. Firstly, the kinetics of germination increased slightly when spores were heat activated at 60 °C or 68 °C as compared to no heat activation. This was reflected most noticeably by the steeper slopes for DCA and CDCA, and the increased ability of the spore to germinate to the weak germinant GDCA. Secondly, as activation temperature increased, the extent of germination induced by GCA decreased as spores went from no heat activation to heat activation at 68 °C. No germination was observed at ≥ 78 °C for all bile salts. With the exception of the weak germinant GDCA, heat activation below 78 °C did not enhance bile salt germination. Hence, heat activation was not incorporated into downstream germination assays in this study.

A correlation of bile acid structure with germinant activity revealed an intriguing observation. Taking GCA as the reference, the loss of either the R2 hydroxyl (as in GDCA) or the R4 conjugated glycine (as in CA) destroyed germination activity, implying that the presence of both features were important for the recognition of GCA (Supplementary Fig. S8). However, DCA lacked both these features but was still a germinant. This raised the possibility that there are at least two distinct proteins which separately respond to DCA and GCA. We next set out to identify the relevant sensors for the three bile salt germinants.

General features of the *C. septicum* ATCC 11424 genome

To identify the bile salt-responsive elements in *C. septicum*, we sequenced the *C. septicum* genome in search of genes with high similarity to *C. difficile*'s bile-salt germination genes, which are the only characterized bile salt germinant-responsive genes to date. De novo sequencing was necessary because the strain used in this paper (*C. septicum* ATCC 11424) was not among any of the strains sequenced before[40]. Genome assembly yielded a complete chromosome scaffold with three contigs and two plasmids (Supplementary Figs. S9–S11). The assembled chromosomal scaffold reveals features consistent with other sequenced *C. septicum* strains: a similar genome size of 3.41 million bp and a low G/C content of 27.79%. PGAP annotation uncovered a total of 3107 CDS, 115 RNA genes and one CRISPR array (Supplementary Table T2). Other features of the genome are described in Supplementary Tables T3-T4, Supplementary Data 2, and Supplementary Figs. S9–S11.

Identification of putative *csp* operons

Since *C. difficile's* subtilisin-like pseudo-protease (CspC) is currently the only known bile salt germinant sensor, we searched for orthologs of CspC in the *C. septicum* genome. Five *cspC* orthologs were identified and these were divided among three operons (Supplementary Figs. S12) as predicted by Operon-mapper[41]. For ease of reference, these orthologs are referred to in this study by their imputed orthologous gene name (*cspC*) hyphenated with a number denoting their gene order within the *C. septicum* genome. All orthologs contained the subtilase domain, which is the hallmark of Csp serine proteases[9]. The Asp-His-Ser (or DHS) catalytic triad within this domain was preserved in all but one ortholog (*cspC-1718*) where Ser was replaced by Cys, suggesting that *cspC-1718* may function as a cysteine protease instead of a serine protease (Supplementary Figs. S13, S14). One of the three identified operons was structured similarly to *C. perfringens* S40 and *C. difficile* 630 (Fig. 2a). Specifically, this operon contained the genes *cspBA-81*, *cspC-82* and *sleC-83*, which had similar counterparts in *C. difficile* and *C. perfringens* arranged in the same order (Fig. 2a). Two minor differences inconsequential to this work are however noted. Firstly, *cspB* and *cspA* are separate genes in *C. perfringens* whereas both genes are fused as *cspBA* in both *C. septicum* and *C. difficile*. Secondly, in *C. septicum* and *C. perfringens*, *sleC* follows immediately after *cspC*, whereas sleC is located further downstream (1.9 Mb) of *cspC* in *C. difficile*.

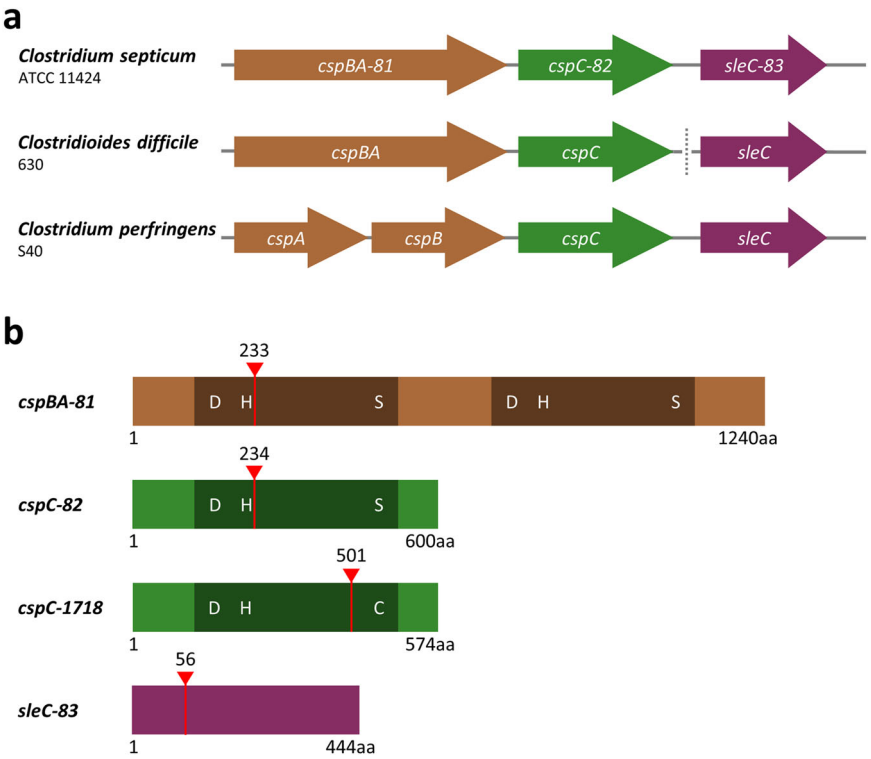

Fig. 2 | Identification of putative *cspC* genes and scheme for genetic inactivation. a Genetic layout of the *csp* operons for *C. septicum*, *C. difficile* and *C. perfringens*. *cspBA-81*, *cspC-82* and *sleC-83* is part of a pentacistronic operon as predicted by OperonMapper. The other two upstream genes were not predicted to be part of the germination mechanism and were hence omitted to avoid clutter. **b** Schematic of insertion sites of the TargeTron vector for generation of mutants in the *csp/sleC* operon and *cspC-1718*.

Targeted insertional mutagenesis using a group II intron vector[42,43] was performed to probe the function of all seven orthologs (five *cspCs*, one *cspBA* and one *sleC*). For technical reasons, we were unable to inactivate three of the orthologs. The remaining four orthologs shown in Fig. 2b (*cspBA-81*, *cspC-82*, *sleC-83* and *cspC-1718*) were successfully inactivated and verified by Southern blot analysis to have single intron insertions (Supplementary Fig. S15). These four mutants were sufficient to provide an insight into the bile salt germination mechanism of *C. septicum*.

Inactivation of *cspC* orthologs abrogates germination to specific bile salts

Spores with inactivated *cspC-82* or *cspC-1718* were evaluated for germination in the presence of bile salts in three ways: (1) drop in optical density, (2) release of calcium-dipicolinate (Ca-DPA) from the spore core[44,45] and (3) phase contrast microscopy.

Inactivation of *cspC-1718* attenuated the germination response to GCA and inactivation of *cspC-82* did the same for CDCA and DCA (Fig. 3, Supplementary Figs. S16). While intron insertional methods can lead to polar effects on the downstream genes, this was not observed in our mutants. Complementation of the mutants with their respective inactivated gene products restored the original germination response. Similar results were also observed with an independent spore batch, further supporting the previous observations (Supplementary Figs. S17). These results, demonstrating that two gene products were responsible for sensing non-coincidental bile salt germinants, was consistent with our earlier hypothesis that there were at least two bile salt germination sensors in *C. septicum*. As expected, there was no appreciable change in germination phenotype with the weakly-germinant and non-germinant bile salts - CA, GDCA, TCA and TDCA (Fig. 3, Supplementary Figs. S16, S17). Interestingly, the OD decrease and Ca-DPA release triggered by GDCA was attenuated by *cspC-82* inactivation but restored by complementation. This suggests that GDCA, which is structurally similar to GCA, may be weakly recognized by the cognate receptor for GCA.

Interestingly, *cspC-1718* mutant had approximately one log lower Ca-DPA content compared to its WT counterpart (Supplementary Fig. S18). Although complementation of the *cspC-1718* mutant did not reverse this phenotype, germination of the *cspC-1718* mutant and its complement still released Ca-DPA, accompanied by the usual OD drop and phase dark spores (Fig. 3, Supplementary Figs. S16, S17).

Inactivation of *cspBA-81* and *sleC-83* abrogates germination to all bile salts

Both *cspBA-81* and *sleC-83* are potentially important for bile salt germination because they reside in the same operon as *cspC-82*, the gene shown by previous experiments to mediate CDCA/DCA-triggered germination. The *cspBA*, *cspC*, and *sleC* genes in *C. difficile* play a pivotal role in its bile salt germination pathway. Specifically, in *C. difficile*, CspBA is cleaved by the YabG protease into CspB and CspA[46]. CspB then undergoes autoprocessing to remove its own inhibitory prodomain, resulting in an active form which converts pro-SleC to SleC. The activated SleC proceeds to hydrolyze the spore peptidoglycan to initiate spore germination[34,47,48].

To study if *cspBA-81* and *sleC-83* in *C. septicum* played a similar role to their counterparts in *C. difficile*, spores with these genes inactivated were studied in the same way as the *cspC* mutants. Here, we observed that inactivation of either gene negated the germination response of *C. septicum* spores to all three bile salt germinants, and complementation with their respective gene product restored germinant sensitivity, suggesting that no polar effects were induced by Targetron insertional mutagenesis (Fig. 4, Supplementary Fig. S19, S20). Not surprisingly, and similar to the *cspC* mutants, there was no appreciable change in germination phenotype with the borderline or non-germinant bile salts - CA, GDCA, TCA and TDCA (Fig. 4, Supplementary Figs. S19). These results were also replicated using an independent spore batch (Supplementary Figs. S20).

FInally, we tested the viability of the mutant and complemented spores on BHI-S plates (Supplementary Fig. S21). All *cspC* mutant spores were viable, but *cspC-1718* spores exhibited approximately 10-fold fewer colonies compared to WT. This lower germination efficiency seems to be consistent with the one log lower Ca-DPA content previously observed for *cspC-1718* (Supplementary Fig. S18). Complementing cspC-1718 restored viability to WT levels, suggesting that the cspC-1718 may somehow influence spore Ca-DPA content. Compared to the cspC mutants, *cspBA* and *sleC* mutant spores were not viable, further corroborating our observations in the

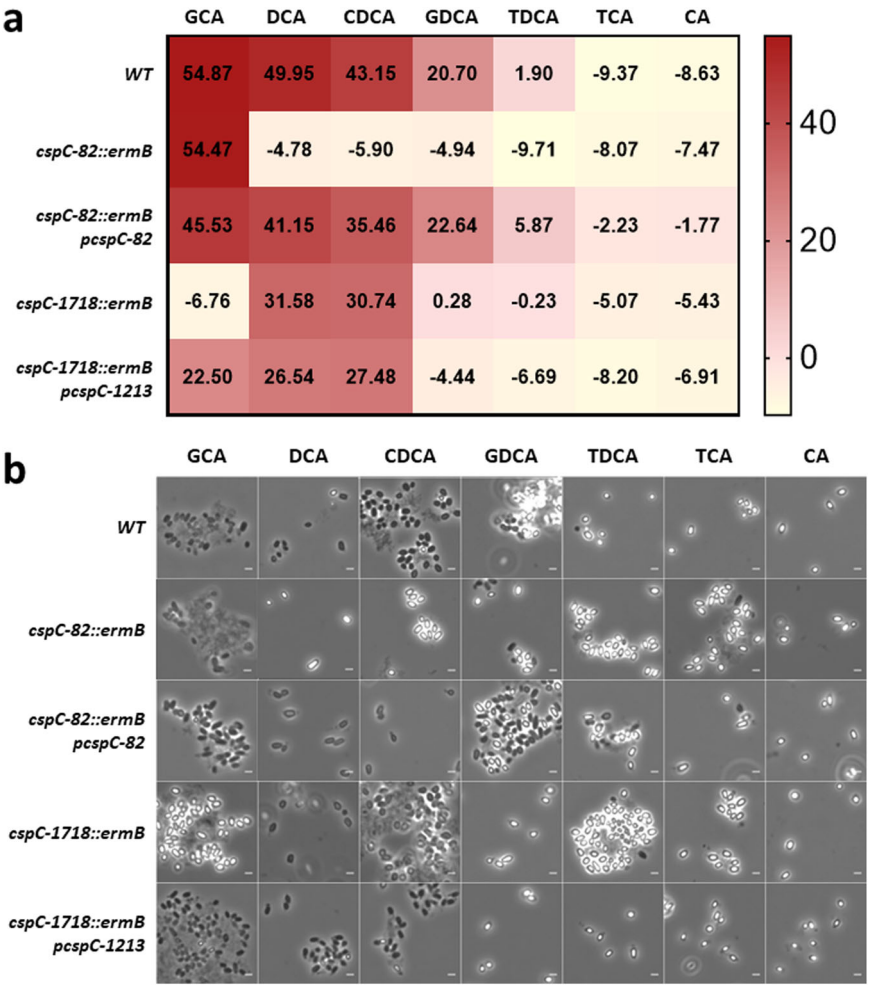

Fig. 3 | Both *cspC-82* and *cspC-1718* mutants show a differential response to bile salt germinants. Purified spores of wild-type (*WT*), *cspC* mutants (*cspC-82::ermB* and *cspC-1718::ermB*) and complemented *cspC* mutants (*cspC-82::ermB pcspC-82* and *cspC-1718::ermB pcspC-1718*) were incubated with 10 mM bile salts in the presence of oxyrase. **a** Heat map depicting %OD drop values at the end of the assay. Colors indicate extent of germination (red: germination, yellow: no germination). The *cspC-82* mutant was able to germinate only in the presence of GCA while the *cspC-1718* mutant was able to germinate in the presence of CDCA and DCA. Complementation of the mutant with their respective gene rescued the mutant phenotype. Data represent the average of $n = 2$ biologically independent experiments. **b** Representative phase contrast images of *WT*, *cspC* mutants and their respective complemented *cspC* mutants after 900 minutes of incubation with various bile-salts. Scale bar, 2 µm.

previous germination assays. As expected, complementing *cspBA* and *sleC* restored colony counts to *WT* levels.

Structural comparisons of CspC-82 and CspC-1718 with *C. difficile* CspC

Having established that CspC-82 and CspC-1718 were similar to *C. difficile*'s CspC in biological function, we wondered what structural similarities existed between the three proteins. A Molecular Dynamics (MD) simulation approach was used to study this question. These MD experiments were not intended to provide functional insight, just to cross-compare the overarching physical attributes of all three proteins.

Homology models for CspC-82 and CspC-1718 were generated using i-TASSER and AlphaFold since crystal structures for both of these proteins were unavailable (Supplementary Data 1). The structural analog identified with the closest similarity to both proteins was *C. difficile*'s CspC, with a high TM score of 0.899 and 0.936 respectively when both proteins were aligned to it. MD simulations of both proteins were then performed to test and refine the accuracy of their homology-modeled structures (Supplementary Table T5). The same simulation was also performed for *C. difficile*'s CspC using its known crystal structure[32].

First the structural stability of the three proteins were verified through the time evolution of the structure over the simulation runs. Specifically, the RMSD with respect to the initial energy minimized structure was calculated. These RMSD values provide microscopic insight on local dynamics of different segments of the protein structure. The RMSD values were calculated by considering only the non-hydrogen atoms of the protein backbone and amino acid side chains. For all three proteins, three independent simulations were performed using three different initial structures to check the observed simulation results are independent of initial protein configuration (Supplementary Fig. S22). Figure 5a shows the change in RMSD with respect to time for three independent protein simulations of *C. difficile*'s CspC. For all three simulations, RMSD showed similar time evolution, ensuring that the relative protein dynamics were independent of the initial structures. Similar protein dynamics were also observed for CspC-82 and CspC-1718. Figure 5b shows the time evolution of the RMSD values for all three proteins. For all three proteins, RMSD quickly increased with time and reached a plateau within 20 ns, demonstrating that for all cases, the structures of the proteins stabilized rapidly. The RMSD profiles exhibited stable fluctuations around mean values of 4.07 Å, 4.22 Å and 1.92 Å for CspC-82, CspC-1718 and *C. difficile's* CspC respectively. The lower RMSD value of *C. difficile*'s CspC indicates higher structural rigidity compared to the other two proteins. Based on the observed RMSDs, the first 20 ns was used as the equilibration run and the next 30 ns as the production run.

The RMSD of the different segments of the protein were further analyzed at the domain level to explore the microscopic details of the local configurational fluctuations within the protein. For the purpose of this analysis, we defined four domains within each protein (Supplementary Table T6) and their relative position within the protein structures are shown in Fig. 5c. The domain-level analyses show that the RMSD values for domains one and two were significantly larger for CspC-82 and CspC-1718 in comparison to *C. difficile's* CspC, indicating that both domains had higher flexibility (Fig. 5d). On the other hand, the similar RMSDs for domains three and four suggest that all three proteins exhibit similar rigidity in these regions. In summary, the previous observation of higher RMSD values for CspC-82 and CspC-1718 were due to the higher flexibility of domains one and two.

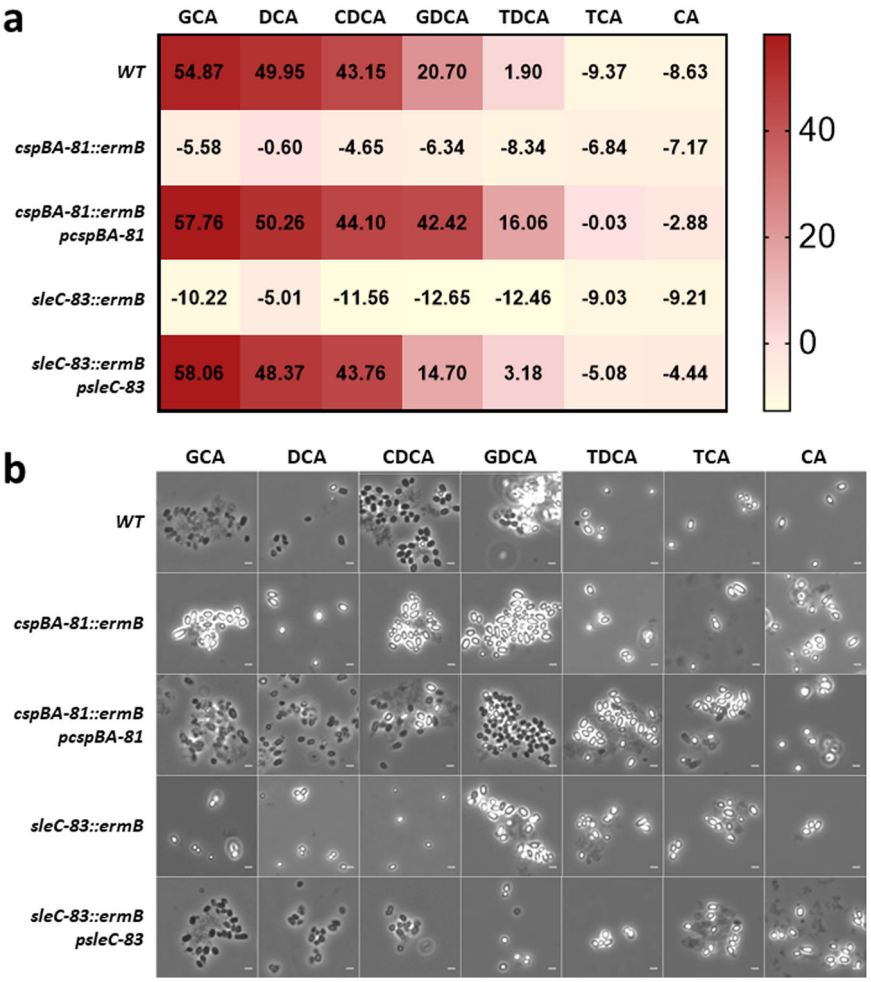

Fig. 4 | Inactivating *cspBA* or *sleC* removes the germination response to all bile salt germinants. Purified spores of *cspBA* or *sleC* mutants and their complemented counterparts were incubated with 10 mM bile salts in the presence of oxyrase. **a** Heat map depicting %OD drop values at the end of the assay. Colors indicate extent of germination (red: germination, yellow: no germination). *cspBA-81* and *sleC-83* mutants were unable to germinate with any bile salts. Complementation of the mutant with their respective gene rescued the mutant phenotype. Data represent the average of $n = 2$ biologically independent experiments. **b** Representative phase contrast images of *WT*, *cspBA* and *sleC* mutants and their respective complements after 900 minutes of incubation with various bile-salts. Scale bar, 2μm.

The interface solvent accessible surface area (interSASA) for each protein domain was also analyzed at the domain level. A larger interSASA value implies a greater chance for interaction between that domain and the rest of the protein. Conversely, a smaller interSASA value indicates a lower surface area for interactions and potential for more flexibility. Here, we observed that all three proteins internally showed the same relative pattern of interSASA values across their four domains, with domain 2 having the highest interSASA and domains 1 and 3 having the lowest interSASA (Supplementary Table T7). Within each domain, C. difficile's CspC showed the highest interSASA while the *C. septicum* CspC orthologs had similar and smaller interSASA values. This result supports the RMSD analysis showing the comparatively greater structural flexibility of *C. septicum's* CspC orthologs at both the protein and domain levels.

To summarize, there was a high degree of general structural congruence between the 3 proteins, providing further corroboration that *C. septicum* CspC orthologs play a similar bile salt-sensing role to *C. difficile's* CspC. Yet, there were also microscopic differences in domain flexibility and interSASA which remain to be explored and could shed light on differences in bile salt specificity or signaling.

Discussion

In this study, we investigated the germinant vocabulary of *C. septicum*. The most noteworthy finding was the strong germination response to bile salts germinants, and the identification of two *cspC* genes mediating this effect. While *C. difficile* has hitherto provided the only example of a *cspC* gene playing a bile salt germinant-sensing role, the two new examples now identified in *C. septicum* create more opportunities for understanding the nature of bile salt-sensing by CspC. Further, we speculate that bile salt dysregulation may plausibly link *C. septicum* with cancer.

Up till now, *C. difficile*'s *cspC* has been the only known gene known so far to mediate bacterial spore germination through bile salt sensing[11,31,32]. Hence, *C. septicum* was sequenced with the aim of identifying *cspC* orthologs which could play the same germinant-sensing function. While *cspC* orthologs were our best lead, the mere presence of *cspC* genes does not guarantee bile salt germination. For example, the well-characterized Csp proteases in *C. perfringens* are not germinant sensors, but instead act downstream of germinant receptor GerKC to activate SleC for cortex hydrolysis[23,24]. In another example, *P. bifermentans* does not germinate in response to bile salts despite the presence of *cspC*[49]. On the flip side, many microbes (besides *C. difficile*) which are known to exhibit bile salt-enhanced germination are also predicted in a Blastp search to possess at least one *cspC* ortholog. These microbes include *Clostridium perfringens, Paeniclostridium sordellii* (previously *Clostridium. sordellii*), *Clostridium baratii*, *Clostridium innocuum*, *Flavonifractor plautii* (previously *Eubacterium plautii*) and *Hungatella hathewayi* (previously *Clostridium hathewayi*)[30,50,51]. Now that *C. septicum* has presented us with a second example of *cspC*'s role in bile salt-sensing, it would be relevant to investigate if the *cspC* orthologs in the abovementioned microbes play a similar bile salt-sensing role. All four deposited *C. septicum* genomes in the NCBI database submitted prior to our study contain paralogs identical in sequence to CspC-82 and CspC-1718. The bile salt-sensing function of *C. septicum* CspCs demonstrated in this study is hence unlikely to be strain-specific. We postulate that CspC's bile sensing properties extend far beyond *C. difficile* and *C. septicum*. Recent work with *C. perfringens* has reported that taurocholate, glycocholate, taurochenodeoxycholate and taurodeoxycholate can potently induce spore germination when combined with L-alanine[51]. If Ger proteins are the canonical sensors for amino acid germinants, it is

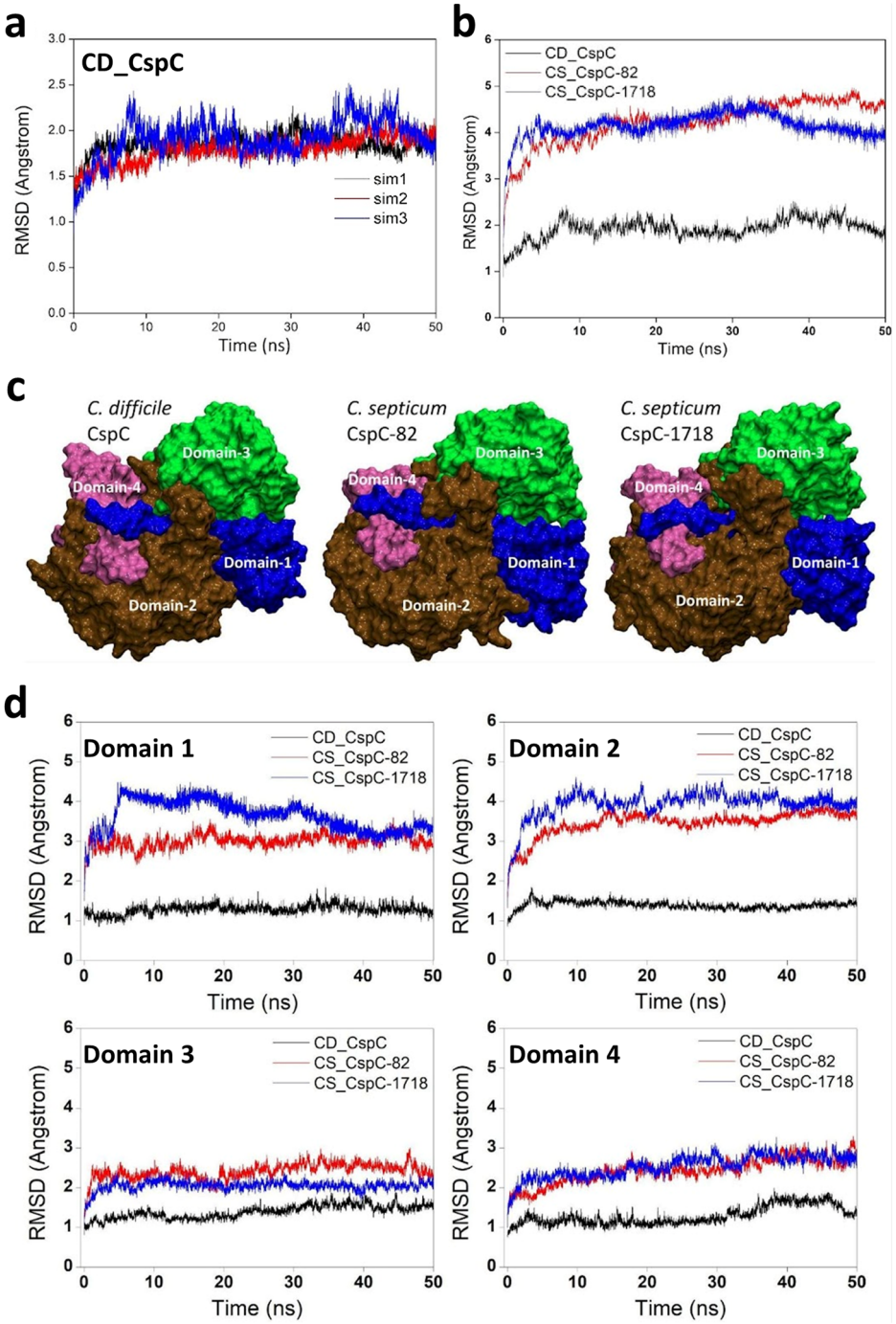

Fig. 5 | Molecular Dynamics Simulation. The structure and stability of *C. septicum*'s (CS) CspC-82 and CspC-1718 were compared with *C. difficile's* (CD) CspC. **a** The change in RMSD with respect to time for three independent protein simulations of *C. difficile*'s CspC is shown. Only non-hydrogen atoms were considered for the calculation of RMSDs. **b** The time evolution of RMSD from the initial energy minimized structure is shown for the three proteins. **c** Space-filled representations of CD_CspC, CS_CspC-82 and CS_CspC-1718, each with four domains, are shown. **d** The time evolution of RMSD measured in (**a**), but resolved for the individual domains shown in **b**, were calculated to understand the local structural fluctuations of the three studied proteins.

conceivable that Csp proteins may likewise form the molecular basis for bile acid germinant sensing.

All proteolytically functional Csp proteins belong to the PA clan (Proteases of mixed nucleophile, superfamily A) and have a catalytic triad which is either DHS (aspartate-histidine-serine) or DHC (aspartate-histidine-cysteine). In contrast, *C. difficile*'s CspC triad has been rendered non-functional through mutation to DTG. Recently, it was speculated that this loss of its catalytic triad was required for *C. difficile* to evolve a bile salt sensing function independent of its original proteolytic one[52]. The evidence in this study seems to suggest otherwise. Both of *C. septicum*'s bile salt-sensing orthologs (CspC-1718 and CspC-82) possess catalytic triads, showing that a catalytic triad is not a barrier to gaining a new bile salt-sensing function. Further, this appears to be true regardless of whether the catalytic triad is DHC (as for CspC-1718) or DHS (as for CspC-82). It remains to be seen whether these triads in *C. septicum* are truly functional. If so, it would imply that CspC's bile salt sensing and proteolytic activity are independent of each other.

While the gene inactivation results suggest that CspC-1718 and CspC-82 are bona fide bile salt germination sensors, the structural modeling experiments show that there is a strong similarity between these proteins and *C. difficile*'s CspC (Fig. 5). A cross-comparison between CspC-82, CspC-1718 and *C. difficile*'s CspC showed a high level of structural as well as dynamical congruence between the three proteins. The major structural domains of *C. difficile*'s CspC were conserved in *C. septicum*'s orthologs — the jellyroll domain (domain 3), which imparts structural rigidity to CspC, as well as the prodomain (domain 1), which functions as an intramolecular chaperone (IMC) for protein folding. The jellyroll domain is thought to play an important role in bile salt germinant sensing. In a recent study, a single

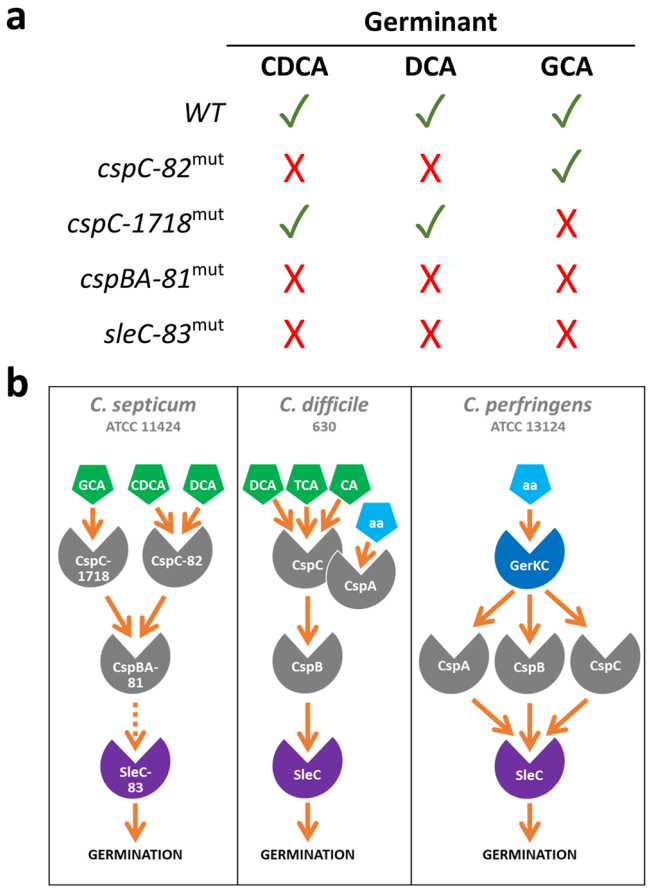

Fig. 6 | Summary of bile-salt response and proposed germination pathway. **a** Summary of bile-salts and their germinant effects on the various mutants. **b** The proposed germination pathway of *C. septicum* is compared with the known pathway elements for *C. difficile* and *C. perfringens*. CspC-1718 and CspC-82 are bile salt sensors (similar to *C. difficile's* CspC) which recognize their cognate bile-salt germinants, leading to downstream activation events which trigger germination. While *C. septicum* does not appear to have a germinant response to amino acids (unlike *C. difficile* and *C. perfringens*), several amino acids can act as co-germinants with sub-optimal levels of GCA (Supplementary Fig. S6A). We have not included potential mediators of amino acid sensing in this model. A dotted arrow joins CspBA-81 and SleC-83 because it is unclear if SleC is activated solely by CspB (like in *C. difficile*), redundantly by multiple Csp proteins (like in *C. perfringens*), or a novel alternative configuration.

amino acid substitution (R358A) in the jellyroll domain of *C. difficile*'s CspC dramatically reduced *C. difficile*'s germination response to taurocholate[32]. One major limitation of these simulations is that they are limited only to bile salt-responsive CspCs. A full molecular dynamical comparison of bile salt responsive CspC's from *C. septicum* and *C. difficile*, against bile salt non-responsive CspC's from other clostridia is warranted. This may reveal structural features within CspC which are essential for bile salt sensing.

Altogether, this study provides sufficient data to begin piecing together a rudimentary germination signaling framework for *C. septicum*. The four bile salt-responsive elements identified in this study corresponded to similar elements in *C. difficile*. Among them, CspC-82 and CspC-1718 responded to non-overlapping subsets of the bile salt germinants while CspBA-81 and SleC-83 responded to all of them (Fig. 6a). This suggests that CspC-82 and CspC-1718 may be individual bile sensing elements, both signaling through downstream CspBA-81 and SleC-83 to cause germination. Here, there is a strong parallel with *C. difficile*, where CspC is the bile salt element which signals through CspB (cleaved from CspBA) to activate SleC's cortex degrading activity (Fig. 6b)[34]. Since CspB also signals downstream to SleC in *C. perfringens*[24], we hypothesize that information flow may be similar in *C. septicum*, proceeding from both CspC elements to CspBA-82 and then SleC-83.

We do not know if CspBA-81 is processed to yield CspB and CspA, but we speculate this is likely because *C. difficile's* orthologous CspBA fusion gene product is non-functional until proteolytically cleaved into functional CspB and CspA by the YabG protease[46]. This suggests that cleavage is necessary for function. In comparison, the need for processing is removed in *C. perfringens* because both CspB and CspA are already expressed as distinct functional proteins.

Although the 20 canonical amino acids have been shown to trigger spore germination in most clostridia[12,13,53,54], it turns out that *C. septicum* is one of the rare cases which only uses these amino acids as co-germinants (Supplementary Figs. S1a, S6). Arginine exhibited the strongest co-germinant effect, triggering spore germination in the presence of sub-germinant concentrations of GCA. This co-germination effect was not observed with the other bile salt germinants (DCA and CDCA), suggesting that amino acid co-germination is specific for the GCA sensing pathway. Based on our current understanding, there are two likely mechanisms through which arginine sensing could occur. In one mechanism, *C. septicum*'s Csp proteins sense arginine similarly to how *C. difficile*'s CspA[33] and CspC[31] are involved in sensing glycine as a co-germinant. Since there is already a strong parallel in how both bacteria use CspC to sense bile salt germinants, *C. septicum* might also use the same co-germinant sensing paradigm as *C. difficile*. On the other hand, there is one big difference between both bacteria; Ger family orthologs are present in the genome of *C. septicum* (Supplementary Table T3) but not *C. difficile*. Since Ger orthologs are the canonical sensor for amino acid germinants in other sporulating bacteria (e.g. *C. perfringens*[12]), it is conceivable that Ger orthologs mediate co-germinant sensing in *C. septicum*. Only further genetic studies with CspA and GerK can shed further light on this and hence we have not speculated on the signaling framework for co-germinant sensing in Fig. 6.

While there are contrasting reports debating the presence of *C. septicum* in humans[55–58], our results now suggest that *C. septicum* is arguably a native inhabitant of the vertebrate gut, especially since bile acids are found exclusively in the vertebrate gut[27]. Physiological concentrations of bile salts within the intestinal lumen[39] fall within the range of EC_{50} values of the bile germinants in this study (Fig. 1c).

Since the bile salt germinants for *C. septicum* (CDCA, GCA, DCA) are found to varying degrees in healthy human individuals, it is perhaps surprising that *C. septicum* is not more commonly detected in the gut microbiome. However, we speculate that its presence in a healthy colon may be suppressed by growth-inhibitory bile salts as well as mucosal immunity. A shift to a germination and growth-friendly bile salt profile, accompanied by a suitable necrotic or ischemic environment for colonization, are likely necessary conditions for *C. septicum* to replicate efficiently within the colon. This is consistent with the current thought that *C. septicum* infections begin in the gut and enter the bloodstream through gastric lesions, which are often but not exclusively caused or facilitated by malignancies. Once in the bloodstream, *C. septicum* can establish itself at distant anatomical sites, leading to spontaneous myonecrosis[59]. Compromised mucosal immunity, a common feature of certain malignancies[60] could further release the brakes on such infections which could explain the high association with malignancy.

In light of our study, it is timely to revisit *C. septicum*'s unique association with colorectal and hematological malignancies[1]. Recent studies have now established a link between certain bile acids and cancer. For example, up to 90% of colorectal cancer cases are influenced by high lipid diets that upregulate bile acid secretion in the gut, with DCA being one such bile acid[61]. Further, DCA and CDCA are known carcinogens implicated in colorectal carcinogenesis, while GCA has been proposed to be a biomarker for hepatocellular carcinoma[62–67]. If the same bile acids which promote carcinogenesis also trigger *C. septicum* spore germination, then bile acid dysbiosis could very well be the common cause linking *C. septicum* infection with malignancy.

Materials and methods

Bacterial strains and growth conditions

C. septicum strains were grown on BHIS (Brain Heart Infusion supplemented with 0.5 g/l L-cysteine) in an anaerobic chamber (Plas lab) at 37 °C (85% N_2, 10% H_2, and 5% CO_2). Antibiotics were added as needed (10 µg/mL thiamphenicol, polymyxin B 60 U/ml or 2.5 µg/mL erythromycin). The *E. coli* strains, XL1-Blue and S17-1 (Biomedal S.L) were cultured aerobically at 37 °C on either LB plates or 2xYT broth and supplemented with 10 µg/ml chloramphenicol. XL1-Blue was used as the host for plasmid construction while S17-1 was used for conjugal transfer of plasmids to *C. septicum*.

Generation of *csp* and *sleC* mutants

Oligonucleotides and synthesized DNA fragments were listed in Supplementary Materials 1, 2. The Targetron plasmid pJIR750ai (Sigma) was used as the starting plasmid for generation of mutants and was modified in several ways. First, pRJ1 was created by cloning the oriT site into pJIR750ai. Next, BtsMutI restriction sites were knocked out and re-assembled with NEBuilder HiFi DNA assembly (NEB) to yield the plasmid pRJ2. The original promoter for the transcription of the group II intron machinery, beta-2p, was replaced with the promoter for the alpha toxin gene in *C. septicum* (P*csa*) to yield pRJ3. The *ErmB* retrotransposition activated selectable marker (RAM), which confers erythromycin resistance upon insertion of the intron into the genome, was inserted into the MluI site of PFGv2-Pcsa to finally yield the base plasmid pRJ4.

Insertion sites in each gene were scored and predicted per gene using an algorithm[68]. Our choices of insertion sites were chosen according to two criteria, 1) high scores 2) position of insertion site relative to the N-terminus of the protein so that the protein is truncated early and abolishment of the function can be determined. The identified retargeting regions were synthesized by Gene Universal and ligated into pRJ4 at the HindIII and BsrGI restriction site. These plasmids were then sequenced and used for electroporation of the transfer strain S17-1

csp and *sleC* mutants were generated in *C. septicum* ATCC 11424 with the TargeTron mutagenesis system. Conjugal transfer of the Targetron plasmid (Supplementary Material 3) to the recipient *C. septicum* strains was performed in the anaerobic chamber by dripping 200 µl of the conjugation mixture on BHIS plates for six hours. The mix was then harvested with 1 ml of PBS and 100 µl of the slurry was spreaded on a BHIS plate supplemented with 10 µg/ml thiamphenicol and 60U/ml polymyxin B for 16–24 h. Replica plating was then performed on erythromycin containing BHIS plates (2.5 µg/ml) and grown for a further 16–24 h before screening with PCR to confirm insertion of introns. Mutants were then further streaked on BHIS plates to shed the TargeTron plasmid.

Complementation of *C.septicum* mutants

The modular plasmid pMTL83153[69] was modified by replacing the promoter P_{fdx} with P_{csa} at the SacI and AscI restriction site to yield pRJ5. Finally, the base constitutive expression plasmid pRJ6 was constructed by inverse PCR with primers complementary to the FLAG tag and the end of the alpha toxin promoter sequence. The respective genes for complementation were amplified by the Q5 polymerase (NEB) with their respective primers and assembled into pRJ6 with NEBuilder HiFi DNA assembly (NEB). These plasmids (Supplementary Material 3) were then sequenced and used for electroporation of the transfer strain S17-1. Conjugal transfer of the respective plasmids to the respective mutant recipient strains was performed as described above. The recipient strains carrying their respective plasmids were verified by PCR.

Genomic DNA isolation

Genomic DNA (gDNA) was isolated with a few modifications from the original protocol described in ref. 70. RNAseA (Sigma) and Proteinase K (Roche) treatment was only done following crude DNA extraction for 2 h each at 37 °C. Lastly, gDNA was purified by phenol/chloroform extraction and quantified with Picogreen (Life Technologies) using lambda DNA (NEB) as standards.

Southern blot analysis

0.5 µg of gDNA was digested with 10 U of AseI, separated by gel electrophoresis in a 0.8% agarose gel and transferred to a nylon membrane (Amersham Hybond-N +) with a vacuum blotting system (GE, VacuGeneXL). Hybridization of the blot with the intron probe was done following manufacturer instructions (Roche) The blot was hybridized with a 378 bp DIG-labeled intron probe (Roche) generated from the plasmid pRJ3.

DNA sequencing, assembly and annotation

High molecular genomic DNA was sheared, and the 10kB fragments were isolated for end-repair and universal hairpin adapters ligation. Subsequent steps were followed as per PacBio protocol to prepare the SMRTbell library. The library was sequenced in PacBio Sequel and the resulting sequencing reads were assembled with Flye v2.8[71]. Genome annotation was carried out using NCBI's prokaryotic genome annotation pipeline (PGAP) build 6021[72]. Schematic representations of the contigs and plasmids were generated with Proksee (https://proksee.ca/). Prophage analysis was predicted with the PHASTER server[73]. Analysis of Insertion elements were carried out with ISEscan[74]. Evaluation of genome completeness was executed with CheckM v1.0.18[75]. Genome assembly statistics was reported by Quast 4.4 and PGAP[72,76]. Both Quast and CheckM were hosted on Kbase[77]. The assembled genome can be found on NCBI's GenBank through accession number JARRAV000000000 under Bioproject ID PRJNA872817.

Bioinformatic identification and comparison of Ger/Csp orthologs

To identify putative CspC orthologs in *C. septicum*, CspC from *C. difficile* was used to query the *C. septicum* CDS sequences with the Protein Blast (Blastp) algorithm. Protein Blast (Blastp) was performed over nucleotide blast (Blastn) to account for synonymous codons. GerK orthologs in *C. septicum* were identified by querying the sequences of *C. perfringens* GerAA, GerKA, GerKB and GerKC. Identification of putative *cspC* orthologs of *C. baratii* (taxid: 1561), *C. innocuum* (taxid: 1522), *F. plautii* (taxid: 292800) *H. hathewayi* (taxid: 154046) and *P. bifermentans* (taxid: 1490) was performed with Blastp using CspC (accession number: WP_003433821) from *C. difficile*. CspBA and CspC orthologs of *C. perfringens*, *C. difficile* and *C. septicum* were analyzed in MegAlign Pro v13 (DNASTAR) using the Clustal Omega algorithm.

Sporulation and purification of *C. septicum* spores

C. septicum spores were purified as described previously[78], with some modifications. Briefly, overnight cultures of *C. septicum* were diluted 50X into 100 ml of BHI-S and incubated until the OD600 is 1.5 to 3 before the entire culture was added into 900 ml of sporulation media (0.5% L-cysteine, 3% Bacto polypeptone, 5% dehydrated cooked meat medium and 10% fetal bovine serum). Thiamphenicol (5 µg/ml) or erythromycin (2.5 µg/ml) were added when necessary. This sporulation media was then incubated for a further 5 days. Spores were purified from vegetative cells on a 80% discontinuous Percoll gradient, washed twice, resuspended in water and stored at 4 °C until use. Purified spores were observed to be >99% phase bright.

Germination of *C. septicum* spores

All germination assays were conducted at 37 °C in the presence of 0.2x Oxyrase for broth (Oxyrase®). Single factor screening experiments of WT spores (Supplementary Table T1) were performed (Oxyrase®) in a 384 well plate (Greiner) and covered with a plastic film (Excel Scientific). Determination of the optimum pH for WT spore germination was performed as above except WT spores were germinated in a germination assay buffer containing 10 mM sodium phosphate buffer with a range of pH from 6 to 8, 20 mM DCA and bacterial spores adjusted to an initial OD_{600} of 1.0±0.2. Data from the above experiments were collected from the infinite 200 Microplate reader (Tecan).

Subsequently all data was collected from the Spark Multimode Microplate reader (Tecan) To investigate whether Ca-DPA release from the spores was responsible for the high OD artifact observed with DCA induced spore germination, we performed germination assays as described above, except that Ca^{2+} and DPA was added instead of spores. Briefly, a germination solution consisting of 10 mM sodium phosphate buffer at the tested pH of 6, 6.5 or 7, 20 mM DCA, CA, or GCA, 50 uM Ca^{2+} and DPA was incubated at 37 °C and the OD was measured every 5 mins over 15 hours.

Germination assays of WT spores in serial dilutions of bile salts was performed in a germination assay media containing 10 mM sodium phosphate buffer pH 7.0, 0.2X Oxyrase, serial dilutions of bile salts (final concentrations 50 mM–0.78 mM) and lastly bacterial spores adjusted to an initial OD_{600} of 1.0 ± 0.2 in a 384 well UV transparentplate (Greiner UV Star plate) covered with an optically clear qPCR film (ABI technologies) to maintain anaerobicity. All measurements were taken in intervals of five minutes for 15 hours.

Amino acid co-germination assays were performed as above with sub-optimal concentrations of CDCA (1.56 mM), DCA (1.56 mM) and GCA (3.13 mM), except that a 20- amino acid mastermix was added to the germination buffer at a final concentration of 5 mM for each amino acid. Individual amino acids (5 mM) were also tested with GCA (3.13 mM) as described above. The final percentage of OD decrease relative to its initial OD ($OD_{final}/OD_{initial}$) of each individual amino acid with GCA is then subtracted by the GCA only control to establish the effect of the amino acid.

For heat activation of WT spores, WT spores were incubated in a thermocycler (AIT Biotech) in their respective temperatures for 15 minutes before being cooled on ice for five minutes prior to performing the germination assay as described above with a few modifications. The respective bile salts and $TbCl_3$ were added to a final concentration of 10 mM and 0.1 mM respectively. Released DPA complexes with terbium to form a fluorescence complex and is measured at an excitation of 272 nm with emission of 545 nm at z = 18216 µm and a time lag of 20 µs before measurement. All measurements were taken in kinetic mode in 5 minute steps for 15 hours. Germination assays of WT, *csp*, *sleC* mutants and complemented mutants were conducted in the same way as heat activation assays except spores were not heat activated. All values below zero were considered to be zero. Due to low Ca-DPA content in *cspC-1718* mutants and its complemented mutant spores, spores were added to an initial OD600 of 1.3 ± 0.2 and the acquisition parameters were optimized as follows: number of flashes (50), lag and integration time (2000 us).

Total DPA release assay

To assess the total DPA content in the spore preparation, spores suspensions (OD 0.5 in 100 ul) were heated at 100 °C for an hour. The spores were then cooled on ice for five minutes. Next, the spore suspension is spun at 17,000 g for 5 minutes and 20ul of the supernatant was assayed for DPA content in the same buffer as described above for the spore germination assay for the mutant, except that bile salts are excluded.

Recovery of colonies from bile salt exposed *C. septicum* spores on agar

2×10^7 *C. septicum* spores were incubated in 0.1% solutions of various bile acids for 10 min in the anaerobic chamber before they were plated on LB overnight. The colonies were counted the next day and compared to GCA exposed spores.

Comparison of the viability of *C. septicum* mutant and complemented spores

Working concentrations of *C. septicum* WT, mutant and complemented spores (2×10^7 spores/ml) were serially diluted and 5ul was plated on BHI-S plates and left to dry for 30 minutes. After drying, plates were incubated in a GasPak jar (BD, 260628) overnight.

Protein structure prediction and simulation details

The all-atom molecular dynamics (MD) simulation was performed for three systems: (i) *C. difficile* CspC, (ii) *C. septicum* CspC-82, and (iii) *C. septicum* CspC-1718 to explore the structural stability and similarities of the three proteins. NAMD simulation program v2.14 with CHARMM36 force field was used to perform all simulations[79,80]. The initial configuration for the simulation of *C. difficile* CspC was obtained from Protein Data Bank (PDB entry 6MW4; 1.1 Å resolution), whereas the initial configurations of *C. septicum* CspC-82 and *C. septicum* CspC-1718 were modeled using i-TASSER[81] and AlphaFold[82]. The end residues for all proteins were properly capped (Acetyl group at the N-terminal and N-methyl group at the C-terminal). The side-chain protonation states of asparagine, glutamine, and histidine were considered and resolved through optimization of local hydrogen bonding. The three proteins were solvated by cubic TIP3 water boxes, and the overall charge of each system was neutralized by adding counter ions. Each of the three systems were energy minimized by performing 50,000 steps of Steepest Descent to eliminate the close van der Waals contacts. After that, the temperature of the systems was gradually increased to room temperature (300 K), followed by 50 ps equilibration using the NPT algorithm, where temperature was controlled by the Langevin algorithm. The three systems were then run for 50 ns under a constant pressure of 1 atm, and a constant temperature of 303 K (NPT ensemble). The time steps for each simulation was 2 fs, and trajectories were stored every 2 ps for analysis. The long-range electrostatic forces were calculated using Particle Mesh Ewald method. Periodic boundary condition and a 10 Å cutoff were applied for non-bonded short range interactions.The interface Solvent-Accessible Surface Area (inter*SASA*) for a specific domain *i* among the set of all domains *P* is defined as: $interSASA_{\{i\}} = SASA_{\{i\}} + SASA_{P\text{-}\{i\}} - SASA_P$ where $SASA_{\{i\}}$, $SASA_{P\text{-}\{i\}}$ and $SASA_P$ are the *SASA* values for all domains, domain *i* and all domains excluding *i* respectively.

Statistics and Reproducibility

Spores of the order of 10^9 CFU/ml were used for all germination assays. All experiments were independently replicated at least twice with two embedded technical replicates. All attempts at replicating our findings were successful. Where applicable, spores from a different batch were also assayed. For viability and recovery of spores on agar, at least 2×10^7 spores/ml were used and this amount produced colonies which were serially diluted for accurate counting. Figures were generated on Graphpad Prism 9 for Windows (Graphpad software). Error bars represent the standard error of the mean. Statistical analysis of germination curves, CFU recovery, amino acid co-germination, heat activation and total DPA content was done with one way ANOVA and Dunnett with the final endpoint data. Comparison of OD increase of different bile acids in varying pH was performed with two way ANOVA and Dunnett with the final endpoint data.

To screen for potential germinants, OD drop over the moving average of 6 timepoints (25 min) was calculated and the maximum OD drop velocity value over the duration of the experiment was used. For bile salt concentration response, 100% germination response was defined as the maximal germination response with DCA since it showed the highest OD drop while 0% germination response was defined as the germination response in water instead of bile salts. The standard curve of each bile salt germinant was generated using a four-parameter logistic curve and EC_{50} values were interpolated from the generated curves. To fit V_{max} curves, V_{max} was calculated after the assay started for 30 min over a moving window of three time points (10 min) and the maximum velocity of each concentration was used. The curve of each bile salt was generated using the built-in allosteric sigmoidal curve fit. An unpaired *t*-test was conducted to assess if there were differences in EC_{50} and V_{max} values between the germinants.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All data is available in the main text or the Supplementary Information. The assembled genome can be found on NCBI's GenBank through accession number JARRAV000000000. The source data behind the graphs in the paper can be found in Supplementary Data 3.

Received: 9 May 2023; Accepted: 23 July 2024;
Published online: 06 August 2024

References

1. Kornbluth, A. A., Danzig, J. B. & Bernstein, L. H. Clostridium septicum infection and associated malignancy. Report of 2 cases and review of the literature. *Medicine* **68**, 30–37 (1989).
2. Hermsen, J. L., Schurr, M. J., Kudsk, K. A. & Faucher, L. D. Phenotyping Clostridium septicum infection: a surgeon's infectious disease. *J. Surg. Res.* **148**, 67–76 (2008).
3. Srivastava, I., Aldape, M. J., Bryant, A. E., Stevens, D. L. & Spontaneous, C. septicum gas gangrene: A literature review. *Anaerobe* **48**, 165–171 (2017).
4. Kennedy, C. L. et al. Pore-Forming Activity of Alpha-Toxin Is Essential for Clostridium septicum-Mediated Myonecrosis. *Infect. Immun.* **77**, 943–951 (2009).
5. Jing, W. et al. *Clostridium septicum* α-toxin activates the NLRP3 inflammasome by engaging GPI-anchored proteins. *Sci. Immunol.* **7**, eabm1803 (2022).
6. Larson, C. M., Bubrick, M. P., Jacobs, D. M. & West, M. A. Malignancy, mortality, and medicosurgical management of Clostridium septicum infection. *Surgery* **118**, 592–597 (1995). discussion 597–8.
7. Prescott, J. F., MacInnes, J.I. & Wu, A. K. K. Taxonomic relationships among the Clostridia. in *Clostridial Diseases of Animals* 1–5 (John Wiley & Sons, Inc, Hoboken, NJ, 2016).
8. Lawson, P. A., Citron, D. M., Tyrrell, K. L. & Finegold, S. M. Reclassification of Clostridium difficile as Clostridioides difficile (Hall and O'Toole 1935) Prévot 1938. *Anaerobe* **40**, 95–99 (2016).
9. Bhattacharjee, D., McAllister, K. N. & Sorg, J. A. Germinants and Their Receptors in Clostridia. *J. Bacteriol.* **198**, 2767–2775 (2016).
10. Shen, A., Edwards, A. N., Sarker, M. R. & Paredes-Sabja, D. Sporulation and Germination in Clostridial Pathogens. *Microbiol Spectr* **7**, 1–30 (2019).
11. Sorg, J. A. & Sonenshein, A. L. Bile salts and glycine as cogerminants for Clostridium difficile spores. *J. Bacteriol.* **190**, 2505–2512 (2008).
12. Paredes-Sabja, D., Torres, J. A., Setlow, P. & Sarker, M. R. Clostridium perfringens spore germination: characterization of germinants and their receptors. *J. Bacteriol.* **190**, 1190–1201 (2008).
13. Brunt, J. et al. Functional characterisation of germinant receptors in Clostridium botulinum and Clostridium sporogenes presents novel insights into spore germination systems. *PLoS Pathog.* **10**, e1004382 (2014).
14. Waites, W. M. & Wyatt, L. R. The effect of pH, germinants and temperature on the germination of spores of Clostridium bifermentans. *J. Gen. Microbiol.* **80**, 253–258 (1974).
15. Gibbs, P. A. The activation of spores of Clostridium bifermentans. *J. Gen. Microbiol.* **46**, 285–291 (1967).
16. Waites, W. M. & Wyatt, L. R. Germination of spores of Clostridium bifermentans by certain amino acids, lactate and pyruvate in the presence of sodium or potassium ions. *J. Gen. Microbiol.* **67**, 215–222 (1971).
17. Kochan, T. J. et al. Intestinal calcium and bile salts facilitate germination of Clostridium difficile spores. *PLoS Pathog*. **13**, 1–21 (2017).
18. Paredes-Sabja, D., Udompijitkul, P. & Sarker, M. R. Inorganic phosphate and sodium ions are cogerminants for spores of Clostridium perfringens type A food poisoning-related isolates. *Appl. Environ. Microbiol.* **75**, 6299–6305 (2009).
19. Alberto, F., Broussolle, V., Mason, D. R., Carlin, F. & Peck, M. W. Variability in spore germination response by strains of proteolytic Clostridium botulinum types A, B and F. *Lett. Appl. Microbiol.* **36**, 41–45 (2003).
20. Plowman, J. & Peck, M. W. Use of a novel method to characterize the response of spores of non-proteolytic Clostridium botulinum types B, E and F to a wide range of germinants and conditions. *J. Appl. Microbiol.* **92**, 681–694 (2002).
21. Setlow, P., Wang, S. & Li, Y.-Q. Germination of Spores of the Orders Bacillales and Clostridiales. *Annu. Rev. Microbiol.* **71**, 459–477 (2017).
22. Shen, A. Clostridioides difficile Spore Formation and Germination: New Insights and Opportunities for Intervention. *Annu. Rev. Microbiol.* **74**, 545–566 (2020).
23. Talukdar, P. K. & Sarker, M. R. The serine proteases CspA and CspC are essential for germination of spores of Clostridium perfringens SM101 through activating SleC and cortex hydrolysis. *Food Microbiol.* **86**, 103325 (2020).
24. Paredes-Sabja, D., Setlow, P. & Sarker, M. R. The protease CspB is essential for initiation of cortex hydrolysis and dipicolinic acid (DPA) release during germination of spores of Clostridium perfringens type A food poisoning isolates. *Microbiology* **155**, 3464–3472 (2009).
25. Shimamoto, S., Moriyama, R., Sugimoto, K., Miyata, S. & Makino, S. Partial characterization of an enzyme fraction with protease activity which converts the spore peptidoglycan hydrolase (SleC) precursor to an active enzyme during germination of Clostridium perfringens S40 spores and analysis of a gene cluster involved in the activity. *J. Bacteriol.* **183**, 3742–3751 (2001).
26. Paredes-Sabja, D., Setlow, P. & Sarker, M. R. SleC is essential for cortex peptidoglycan hydrolysis during germination of spores of the pathogenic bacterium Clostridium perfringens. *J. Bacteriol.* **191**, 2711–2720 (2009).
27. Ridlon, J. M., Kang, D. J., Hylemon, P. B. & Bajaj, J. S. Bile acids and the gut microbiome. *Curr. Opin. Gastroenterol.* **30**, 332–338 (2014).
28. Tian, Y. et al. The microbiome modulating activity of bile acids. *Gut Microbes* **11**, 979–996 (2020).
29. Larabi, A. B., Masson, H. L. P. & Bäumler, A. J. Bile acids as modulators of gut microbiota composition and function. *Gut Microbes* **15**, 2172671 (2023).
30. Browne, H. P. et al. Culturing of 'unculturable' human microbiota reveals novel taxa and extensive sporulation. *Nature* **533**, 543–546 (2016).
31. Francis, M. B., Allen, C. A., Shrestha, R. & Sorg, J. A. Bile acid recognition by the Clostridium difficile germinant receptor, CspC, is important for establishing infection. *PLoS Pathog.* **9**, e1003356 (2013).
32. Rohlfing, A. E. et al. The CspC pseudoprotease regulates germination of Clostridioides difficile spores in response to multiple environmental signals. *PLoS Genet* **15**, e1008224 (2019).
33. Shrestha, R., Cochran, A. M. & Sorg, J. A. The requirement for co-germinants during Clostridium difficile spore germination is influenced by mutations in yabG and cspA. *PLoS Pathog.* **15**, e1007681 (2019).
34. Adams, C. M., Eckenroth, B. E., Putnam, E. E., Doublié, S. & Shen, A. Structural and functional analysis of the CspB protease required for Clostridium spore germination. *PLoS Pathog.* **9**, e1003165 (2013).
35. Lawler, A. J., Lambert, P. A. & Worthington, T. A Revised Understanding of Clostridioides difficile Spore Germination. *Trends Microbiol* **28**, 744–752 (2020).
36. Setlow, P. Spore germination. *Curr. Opin. Microbiol.* **6**, 550–556 (2003).
37. Zhu, Y. et al. Rapid determination of spore germinability of Clostridium perfringens based on microscopic hyperspectral imaging technology and chemometrics. *J. Food Eng.* **280**, 109896 (2020).

38. Hofmann, A. F. & Mysels, K. J. Bile acid solubility and precipitation in vitro and in vivo: the role of conjugation, pH, and Ca2+ ions. *J. Lipid Res.* **33**, 617–626 (1992).
39. Northfield, T. C. & McColl, I. Postprandial concentrations of free and conjugated bile acids down the length of the normal human small intestine. *Gut* **14**, 513–518 (1973).
40. Thomas, P. et al. First Comparative Analysis of Clostridium septicum Genomes Provides Insights Into the Taxonomy, Species Genetic Diversity, and Virulence Related to Gas Gangrene. *Front. Microbiol.* **12**, 771945 (2021).
41. Taboada, B., Estrada, K., Ciria, R. & Merino, E. Operon-mapper: a web server for precise operon identification in bacterial and archaeal genomes. *Bioinformatics* **34**, 4118–4120 (2018).
42. Kuehne, S. A., Heap, J. T., Cooksley, C. M., Cartman, S. T. & Minton, N. P. ClosTron-mediated engineering of Clostridium. *Methods Mol. Biol.* **765**, 389–407 (2011).
43. Shrestha, R., Lockless, S. W. & Sorg, J. A. A Clostridium difficile alanine racemase affects spore germination and accommodates serine as a substrate. *J. Biol. Chem.* **292**, 10735–10742 (2017).
44. Ragkousi, K., Eichenberger, P., van Ooij, C. & Setlow, P. Identification of a new gene essential for germination of Bacillus subtilis spores with Ca2+-dipicolinate. *J. Bacteriol.* **185**, 2315–2329 (2003).
45. Tehri, N., Kumar, N., Raghu, H. V. & Vashishth, A. Biomarkers of bacterial spore germination. *Ann. Microbiol.* **68**, 513–523 (2018).
46. Kevorkian, Y., Shirley, D. J. & Shen, A. Regulation of Clostridium difficile spore germination by the CspA pseudoprotease domain. *Biochimie* **122**, 243–254 (2016).
47. Burns, D. A., Heap, J. T. & Minton, N. P. SleC is essential for germination of Clostridium difficile spores in nutrient-rich medium supplemented with the bile salt taurocholate. *J. Bacteriol.* **192**, 657–664 (2010).
48. Gutelius, D., Hokeness, K., Logan, S. M. & Reid, C. W. Functional analysis of SleC from Clostridium difficile: an essential lytic transglycosylase involved in spore germination. *Microbiology* **160**, 209–216 (2014).
49. Bhattacharjee, D. & Sorg, J. A. Conservation of the 'Outside-in' Germination Pathway in Paraclostridium bifermentans. *Front. Microbiol.* **9**, 2487 (2018).
50. Liggins, M., Ramirez, N., Magnuson, N. & Abel-Santos, E. Progesterone analogs influence germination of Clostridium sordellii and Clostridium difficile spores in vitro. *J. Bacteriol.* **193**, 2776–2783 (2011).
51. Liggins, M., Ramírez Ramírez, N. & Abel-Santos, E. Comparison of sporulation and germination conditions for Clostridium perfringens type A and G strains. *Front. Microbiol.* **14**, 1143399 (2023).
52. Donnelly, M. L., Forster, E. R., Rohlfing, A. E. & Shen, A. Differential effects of 'resurrecting' Csp pseudoproteases during Clostridioides difficile spore germination. *Biochem. J.* **477**, 1459–1478 (2020).
53. Ramirez, N. & Abel-Santos, E. Requirements for germination of Clostridium sordellii spores in vitro. *J. Bacteriol.* **192**, 418–425 (2010).
54. Sundaresan, A. et al. A design of experiments screen reveals that Clostridium novyi-NT spore germinant sensing is stereoflexible for valine and its analogs. *Commun. Biol.* **6**, 118 (2023).
55. Finegold, S. M., Attebery, H. R. & Sutter, V. L. Effect of diet on human fecal flora: comparison of Japanese and American diets. *Am. J. Clin. Nutr.* **27**, 1456–1469 (1974).
56. Marchesi, J. R. et al. Towards the human colorectal cancer microbiome. *PLoS One* **6**, e20447 (2011).
57. Kopliku, F. A. et al. Low prevalence of Clostridium septicum fecal carriage in an adult population. *Anaerobe* **32**, 34–36 (2015).
58. Ely, J., Ruder, C. A. & Vogt, J. C. A survey of the microbes in human bile. *Trans. Kans. Acad. Sci.* **88**, 20–28 (1985).
59. Mallozzi, M. J. G. & Clark, A. E. Trusting Your Gut: Diagnosis and Management of Clostridium septicum Infections. *Clin. Microbiol. Newsl.* **38**, 187–191 (2016).
60. Hibberd, P. L. Immunizations in adults with cancer. in *UpToDate* (eds. Boeckh, M. & Bogorodskaya, M.) (UpToDate, Waltham, MA, 2021).
61. Ocvirk, S. & O'Keefe, S. J. Influence of Bile Acids on Colorectal Cancer Risk: Potential Mechanisms Mediated by Diet - Gut Microbiota Interactions. *Curr. Nutr. Rep.* **6**, 315–322 (2017).
62. Wu, Y.-C., Chiu, C.-F., Hsueh, C.-T. & Hsueh, C.-T. The role of bile acids in cellular invasiveness of gastric cancer. *Cancer Cell Int* **18**, 75 (2018).
63. Wu, M. et al. Diagnosis of hepatocellular carcinoma using a novel anti-glycocholic acid monoclonal antibody-based method. *Oncol. Lett.* **17**, 3103–3112 (2019).
64. Phelan, J. P., Reen, F. J., Caparros-Martin, J. A., O'Connor, R. & O'Gara, F. Rethinking the bile acid/gut microbiome axis in cancer. *Oncotarget* **8**, 115736–115747 (2017).
65. Yoshimoto, S. et al. Obesity-induced gut microbial metabolite promotes liver cancer through senescence secretome. *Nature* **499**, 97–101 (2013).
66. Di Ciaula, A. et al. Bile Acids and Cancer: Direct and Environmental-Dependent Effects. *Ann. Hepatol.* **16**, S87–S105 (2017).
67. Zhang, A. et al. Urinary metabolic profiling identifies a key role for glycocholic acid in human liver cancer by ultra-performance liquid-chromatography coupled with high-definition mass spectrometry. *Clin. Chim. Acta* **418**, 86–90 (2013).
68. Perutka, J., Wang, W., Goerlitz, D. & Lambowitz, A. M. Use of computer-designed group II introns to disrupt Escherichia coli DExH/D-box protein and DNA helicase genes. *J. Mol. Biol.* **336**, 421–439 (2004).
69. Heap, J. T., Pennington, O. J., Cartman, S. T. & Minton, N. P. A modular system for Clostridium shuttle plasmids. *J. Microbiol. Methods* **78**, 79–85 (2009).
70. Pospiech, A. & Neumann, B. A versatile quick-prep of genomic DNA from gram-positive bacteria. *Trends Genet* **11**, 217–218 (1995).
71. Kolmogorov, M., Yuan, J., Lin, Y. & Pevzner, P. A. Assembly of long, error-prone reads using repeat graphs. *Nat. Biotechnol.* **37**, 540–546 (2019).
72. Tatusova, T. et al. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res* **44**, 6614–6624 (2016).
73. Arndt, D. et al. PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res* **44**, W16–W21 (2016).
74. Xie, Z. & Tang, H. ISEScan: automated identification of insertion sequence elements in prokaryotic genomes. *Bioinformatics* **33**, 3340–3347 (2017).
75. Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P. & Tyson, G. W. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* **25**, 1043–1055 (2015).
76. Gurevich, A., Saveliev, V., Vyahhi, N. & Tesler, G. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* **29**, 1072–1075 (2013).
77. Arkin, A. P. et al. KBase: The United States Department of Energy Systems Biology Knowledgebase. *Nat. Biotechnol.* **36**, 566–569 (2018).
78. Dang, L. H., Bettegowda, C., Huso, D. L., Kinzler, K. W. & Vogelstein, B. Combination bacteriolytic therapy for the treatment of experimental tumors. *Proc. Natl Acad. Sci. Usa.* **98**, 15155–15160 (2001).
79. Best, R. B. et al. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone ϕ, ψ and Side-Chain χ1 and χ2 Dihedral Angles. *J. Chem. Theory Comput.* **8**, 3257–3273 (2012).
80. Phillips, J. C. et al. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802 (2005).
81. Roy, A., Kucukural, A. & Zhang, Y. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.* **5**, 725–738 (2010).
82. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).

Acknowledgements
This work was supported by Temasek Life Sciences Laboratory Core Funding (http://www.tll.org.sg) (to I.C.). The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. We thank colleagues in Temasek Life Sciences Laboratory. Specifically, Ding Yichen for guidance on the assembly of the C. septicum genome; Liu Yanbin, Cai Lin and Ji Liang Hui for equipment and technical advice. This study was supported by Temasek Life Science Laboratories.

Author contributions
Conceptualization (R.S., A.S., I.C). Data curation (R.S., S.J.M.L, S.S., M.S., I.C). Formal analysis (R.S., S.J.M.L, S.S., M.S., I.C.). Funding acquisition (I.C.). Investigation (R.S., S.J.M.L, A.S., S.S. M.S. W.L., A.M., T.W.L., M.D.C., G.J.H). Methodology (R.S., S.S., M.S., I.C.). Project administration (R.S). Software (S.S., M.S). Supervision (R.S., I.C.). Validation (R.S). Visualization (R.S., S.J.M.L, S.S., M.S., I.C.). Writing – original draft (R.S., S.J.M.L, A.S., I.C.). Writing – review & editing (R.S., S.J.M.L, A.S., S.S., M.S., W.L., A.M., T.W.L., M.D.C., G.J.H., I.C.).

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42003-024-06617-4.

Correspondence and requests for materials should be addressed to Ian Cheong.

Peer review information *Communications Biology* thanks Ernesto Abel-Santos and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Tobias Goris. A peer review file is available.