Letter

Skewed genomic variability in strains of the toxigenic bacterial pathogen, *Clostridium perfringens*

Garry S.A. Myers,[1] David A. Rasko,[1,7] Jackie K. Cheung,[2] Jacques Ravel,[1]
Rekha Seshadri,[1] Robert T. DeBoy,[1] Qinghu Ren,[1] John Varga,[3] Milena M. Awad,[2]
Lauren M. Brinkac,[1] Sean C. Daugherty,[1] Daniel H. Haft,[1] Robert J. Dodson,[1]
Ramana Madupu,[1] William C. Nelson,[1] M.J. Rosovitz,[1] Steven A. Sullivan,[1]
Hoda Khouri,[1] George I. Dimitrov,[1] Kisha L. Watkins,[1] Stephanie Mulligan,[1]
Jonathan Benton,[1] Diana Radune,[1] Derek J. Fisher,[4] Helen S. Atkins,[5] Tom Hiscox,[2]
B. Helen Jost,[6] Stephen J. Billington,[6] J. Glenn Songer,[6] Bruce A. McClane,[4]
Richard W. Titball,[5] Julian I. Rood,[2] Stephen B. Melville,[3] and Ian T. Paulsen[1,8]

[1]*The Institute for Genomic Research, Rockville, Maryland 20850, USA;* [2]*Australian Bacterial Pathogenesis Program, Department of Microbiology, Monash University, Clayton 3800, Australia;* [3]*Department of Biological Sciences, Virginia Tech, Blacksburg, Virginia 24601, USA;* [4]*Department of Molecular Genetics and Biochemistry, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania 15261, USA;* [5]*Defence Science and Technology Laboratory, Porton Down, Salisbury SP4 0JQ, United Kingdom;* [6]*Department of Veterinary Science, University of Arizona, Tucson, Arizona 85721, USA*

Clostridium perfringens is a Gram-positive, anaerobic spore-forming bacterium commonly found in soil, sediments, and the human gastrointestinal tract. *C. perfringens* is responsible for a wide spectrum of disease, including food poisoning, gas gangrene (clostridial myonecrosis), enteritis necroticans, and non-foodborne gastrointestinal infections. The complete genome sequences of *Clostridium perfringens* strain ATCC 13124, a gas gangrene isolate and the species type strain, and the enterotoxin-producing food poisoning strain SM101, were determined and compared with the published *C. perfringens* strain 13 genome. Comparison of the three genomes revealed considerable genomic diversity with >300 unique "genomic islands" identified, with the majority of these islands unusually clustered on one replichore. PCR-based analysis indicated that the large genomic islands are widely variable across a large collection of *C. perfringens* strains. These islands encode genes that correlate to differences in virulence and phenotypic characteristics of these strains. Significant differences between the strains include numerous novel mobile elements and genes encoding metabolic capabilities, strain-specific extracellular polysaccharide capsule, sporulation factors, toxins, and other secreted enzymes, providing substantial insight into this medically important bacterial pathogen.

[Supplemental material is available online at www.genome.org. The sequence data from this study have been submitted to GenBank under accession numbers CP000246 (*C. perfringens* ATCC 13124 chromosome) and CP000312, CP000313, CP000314, and CP000315 (*C. perfringens* SM101 chromosome, pSM101A, pSM101B, and φSM101, respectively).]

Clostridium perfringens is a Gram-positive, anaerobic spore-forming bacterium responsible for a wide spectrum of disease in humans and animals. It is commonly found in the gastrointestinal (GI) tract of mammals, as well as in soil and freshwater sediments (Rood and Cole 1991). In humans, *C. perfringens* causes food poisoning, gas gangrene (clostridial myonecrosis), enteritis necroticans, and non-foodborne gastrointestinal infections. In addition, *C. perfringens* is a significant veterinary pathogen, causing a variety of enteric diseases in both domestic and wild animals (Songer 1997). As a species, *C. perfringens* is one of the most prolific producers of toxins (Rood 1998), with five biotypes (A-E) delineated on the basis of the differential production of α, β, ε,

and ι toxins. The different biotypes are associated with different diseases of humans and animals. Both the spores of the bacterium and the toxins have reportedly been of interest to several countries as possible biological weapons (Klietmann and Ruoff 2001), and the *C. perfringens* ε toxin is currently included on both the U.S. Centers for Disease Control (http://www.cdc.gov/od/sap/) and U.S. Department of Agriculture (http://www.aphis.usda.gov/programs/ag_selectagent/ag_bioterr_toxinslist.html) lists of select agents.

The genome sequence of *C. perfringens* strain 13, an enterotoxin-negative type A strain, has been reported (Shimizu et al. 2002). Strain 13 has been widely used as a laboratory model system for gangrene-related studies because of the ease of transformation relative to other strains (Rood and Cole 1991). However, compared with other gangrene isolates, the sequenced isolate of strain 13 can be considered atypical as it sporulates poorly in sporulation media, exhibits only moderate virulence in animal gangrene models (Awad et al. 1995, 2001; Stevens et al. 1997), and has a smaller genome size (3.03 Mb).

Table 1. General features of the *C. perfringens* genomes

Features	ATCC 13124	SM101[a]	Strain[b,c]
Size (bp)	3,256,682	2,897,392	3,031,430
G+C content	28.4	28.2	28.6
Protein-coding genes:			
Similar to proteins of known function	2105	1627	1774
Similar to proteins of unknown function	227	172	118
Conserved hypothetical proteins	382	296	246
Hypothetical proteins	157	331	407
Total	3040	2658	2841
Average CDS size (bp)	894	919	906
Coding (%)	83.5	82.0	82.9
Stable RNAs:			
rRNA	23	29	29
tRNA	93	95	96
sRNA	3	3	5

[a]Including SM101 plasmids (pSM101A and pSM101B) and SM101 episomal phage (φSM101).
[b]Including plasmid pCP13.
[c]Data from Shimizu et al. (2002).

We have sequenced the complete genomes of two additional *C. perfringens* type A strains, ATCC 13124 and SM101. ATCC 13124, the species type strain, was originally isolated from a human gas gangrene patient and produces large quantities of gangrene-associated toxins (Mollby and Holme 1976). *C. perfringens* SM101, a transformable derivative of *C. perfringens* food poisoning isolate NCTC 8798, produces *C. perfringens* enterotoxin (CPE). Although they represent less than 5% of all *C. perfringens* isolates, CPE-producing type A strains are major human GI pathogens, causing *C. perfringens* type A food poisoning (McClane 2001), responsible for ~250,000 reported cases in the United States annually (Mead et al. 1999), and other non-foodborne human GI diseases, such as antibiotic-associated diarrhea and sporadic diarrhea (Sarker et al. 1999; McClane et al. 2000).

Results and Discussion

General genome features

The *C. perfringens* ATCC 13124 and SM101 genomes each consists of a single circular chromosome of 3,256,682 bp and 2,897,392 bp, respectively (Table 1; Fig. 1). SM101 additionally contains two plasmids of 12,397 bp and 12,206 bp (pSM101A and pSM101B, respectively) and a complete episomal bacteriophage genome of 38,092 bp (*C. perfringens* φSM101). A total of 3040 and 2584 CDSs were identified in ATCC 13124 and SM101, respectively. SM101 encodes 10 rRNA operons as was previously described for *C. perfringens* strain 13 (Shimizu et al. 2002) and most *C. perfringens* strains (Rood and Cole 1991); however, ATCC 13124 encodes only eight rRNA operons. The absence of these two rRNA operons was confirmed by PCR amplification from flanking regions.

Comparative analysis

The three sequenced strains of *C. perfringens* share a conserved syntenic core, based on whole-genome nucleotide comparisons (Delcher et al. 2002) and a three-way comparison of their predicted proteomes (Rasko et al. 2005). A total of 2170 genes from strain 13 were conserved in the other two sequenced *C. perfringens* isolates, with almost complete conservation of gene order within these regions (Fig. 2; Supplemental Figure S1). Outside of this conserved core though, considerable genomic diversity was discovered. Three hundred twenty-three divergent "islands" of at least 1 kb in size were identified that were unique to one strain or conserved in only two of the strains analyzed (Fig. 1; Supplemental Tables S1–S6 and Fig. S1). The largest of these genomic islands was a contiguous 242,969-bp island present in ATCC 13124, absent in strain 13, and fragmented in SM101. Conversely, almost no genomic synteny exists between the three sequenced *C. perfringens* isolates and the other published clostridial genomes, *C. tetani* (Bruggemann et al. 2003) and *C. acetobutylicum* (Nolling et al. 2001) (Supplemental Fig. S2).

While genomic diversity has been described in many bacterial lineages, two unusual aspects to the genomic variation have been observed in *C. perfringens*. Firstly, none of the unique genomic regions in any of the three strains is located in a region of atypical nucleotide composition. While a simplistic measure of potential lateral gene transfer at best, atypical nucleotide composition can point to recent gene acquisitions. Secondly, the genomic locations of the divergent regions in all three strains display a dramatic and unusual asymmetry, with an apparent bias for genomic variation on one replichore (Fig. 1). To our knowledge, no similar bias has been identified in any other sequenced genomes, although various genera have shown a tendency for genomic variation to be centered on the terminus of replication (Eisen et al. 2000; Tillier and Collins 2000; Suyama and Bork 2001). Despite the presence of numerous mobile elements in at least one of these genomes, the mechanistic basis for this striking bias in *C. perfringens* remains to be elucidated.

Mobile elements and genomic "islands"

The relative paucity of insertion (IS) elements in *C. perfringens* strain 13 and *C. acetobutylicum* genomes (Nolling et al. 2001) led to the suggestion that the clostridia have few mobile elements (Shimizu et al. 2002). Similarly the *C. tetani* genome (Bruggemann et al. 2003) also has few IS elements. Contradicting this apparent trend, *C. perfringens* SM101 contains 69 IS elements, compared with seven intact IS elements in strain 13 and only three in ATCC 13124 (Sarker et al. 1999). Moreover, 49/69 (71%) of the SM101 IS elements are members of previously undescribed IS families (ISCpe2–ISCpe7; Supplemental Table S7–S9). As observed for the genomic islands, the IS elements in SM101 are not evenly distributed and are located primarily on one replichore (Fig. 1). Only one of the elements (ISCpe4) is found in the same location between ATCC 13124 and strain 13; none of the other IS elements is conserved between the three sequenced strains. Two possible composite transposons were identified in SM101 (CPR_0681-CPR_0809 and CPR_0929-CPR_0946); both are flanked by inverted copies of ISCpe7, and both encode PTS sugar uptake systems; however, neither is flanked by short direct repeats indicative of transposition.

No complete integrated phage was observed; however, all three *C. perfringens* genomes show evidence of prior phage insertions at a variety of genomic locations (Fig. 1). Clostridial phages and their insertion sites are poorly characterized, with only two genome sequences reported: *C. perfringens* φSM101 (this study)
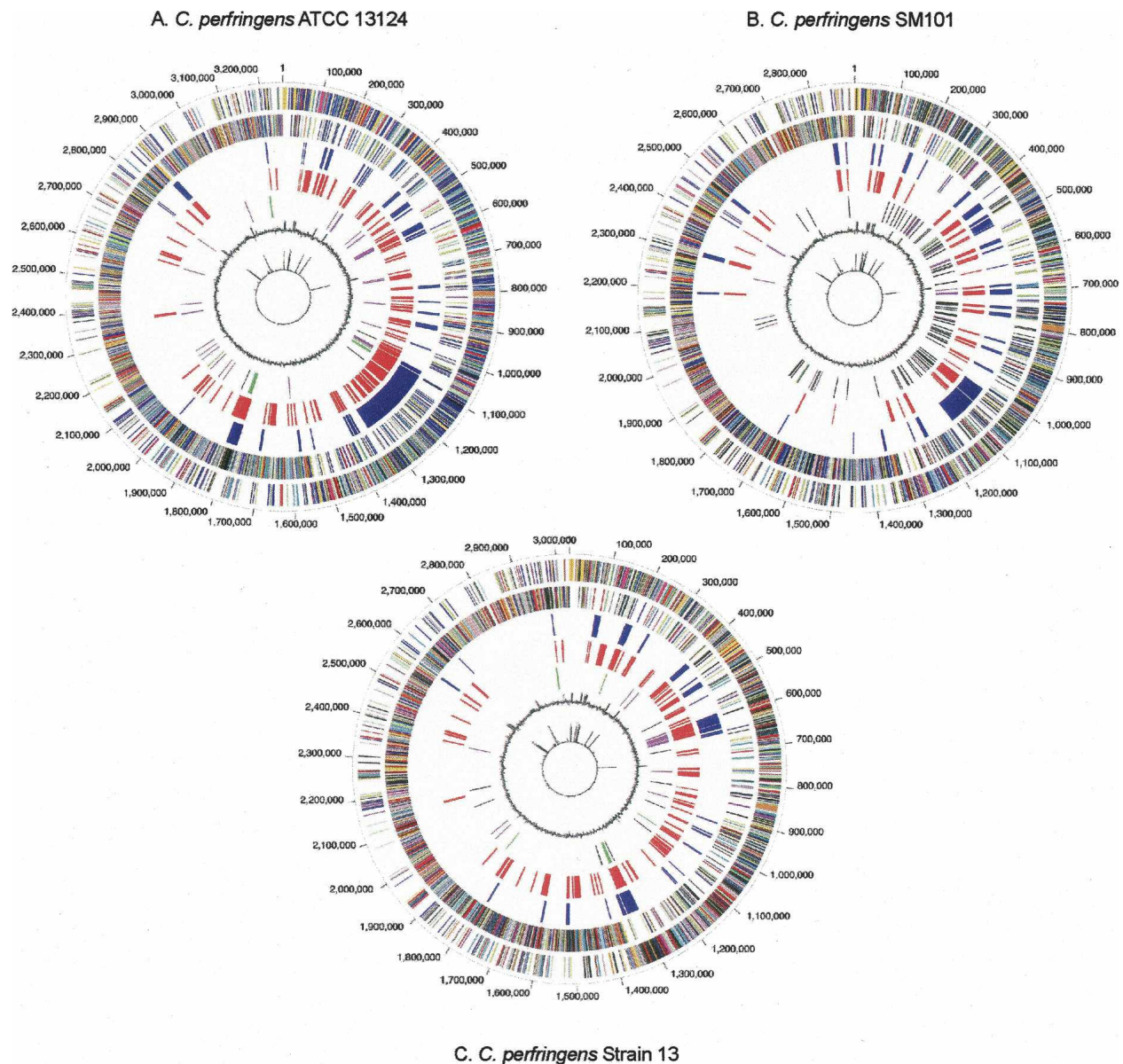
A. *C. perfringens* ATCC 13124

B. *C. perfringens* SM101

C. *C. perfringens* Strain 13

Figure 1. Comparative analysis of each *C. perfringens* genome. (*A*) ATCC 13124; (*B*) SM101; (*C*) Strain 13. For each genome, data are from outermost circle to innermost. (Circles 1 and 2) Tick marks represent predicted coding sequences on the plus strand (circle 1) and minus strand (circle 2) colored by cellular role. Role categories and colors are as follows: amino acid biosynthesis, violet; biosynthesis of cofactors, prosthetic groups and carriers, light blue; cell envelope, light green; cellular processes, red; central intermediary metabolism, brown; DNA metabolism, gold; energy metabolism, light gray; fatty acid and phospholipid metabolism, magenta; protein synthesis and fate, pink; biosynthesis of purines, pyrimidines, nucleosides, and nucleotides, orange; regulatory functions and signal transduction, olive; transcription, dark green; transport and binding proteins, blue–green; other categories, salmon; unknown function, gray; conserved hypothetical proteins, blue; hypothetical proteins, black. (Circles 3 [blue] and 4 [red]) Genes identified by MUMmer analysis to be present in one genome but absent at that position in one other genome: (*A*) Circle 3: present in ATCC 13124 and absent in Strain 13; Circle 4: present in ATCC 13124 and absent in SM101. (*B*) Circle 3: present in SM101 but absent in Strain 13; Circle 4: present in SM101 but absent in ATCC 13124. (*C*) Circle 3: present in Strain 13 but absent in ATCC 13124; Circle 4: present in Strain 13 but absent in SM101. Circle 5: Phage genes (green), capsule-associated genes (purple) and IS elements (black). Circle 6: %G+C. Circle 7: χ^2 of trinucleotide composition.

and *C. perfringens* φ3626 (Zimmer et al. 2002). Using the preliminary public release sequence of the ATCC 13124 genome, a 12-bp bacterial attachment site (*attBB′*) was reported for φ3626 within the 3′ end of *guaA* (CPF_2557) (Zimmer et al. 2002). This attachment site in the *guaA* region is identical in ATCC 13124, SM101, and strain 13. φSM101 does not contain the equivalent φ3626 *attPP′* site and exhibits minimal similarity at the nucleo-

tide level to φ3626, suggesting that if φSM101 does integrate, it will do so elsewhere in the genome.

In all three strains, the unique genomic islands often encode genes likely to confer specific virulence, metabolic, or catabolic capabilities to the host strain. Some of these islands are associated with mobile elements and putative bacteriophage-related regions. The ATCC 13124 prophage remnant regions are associ-
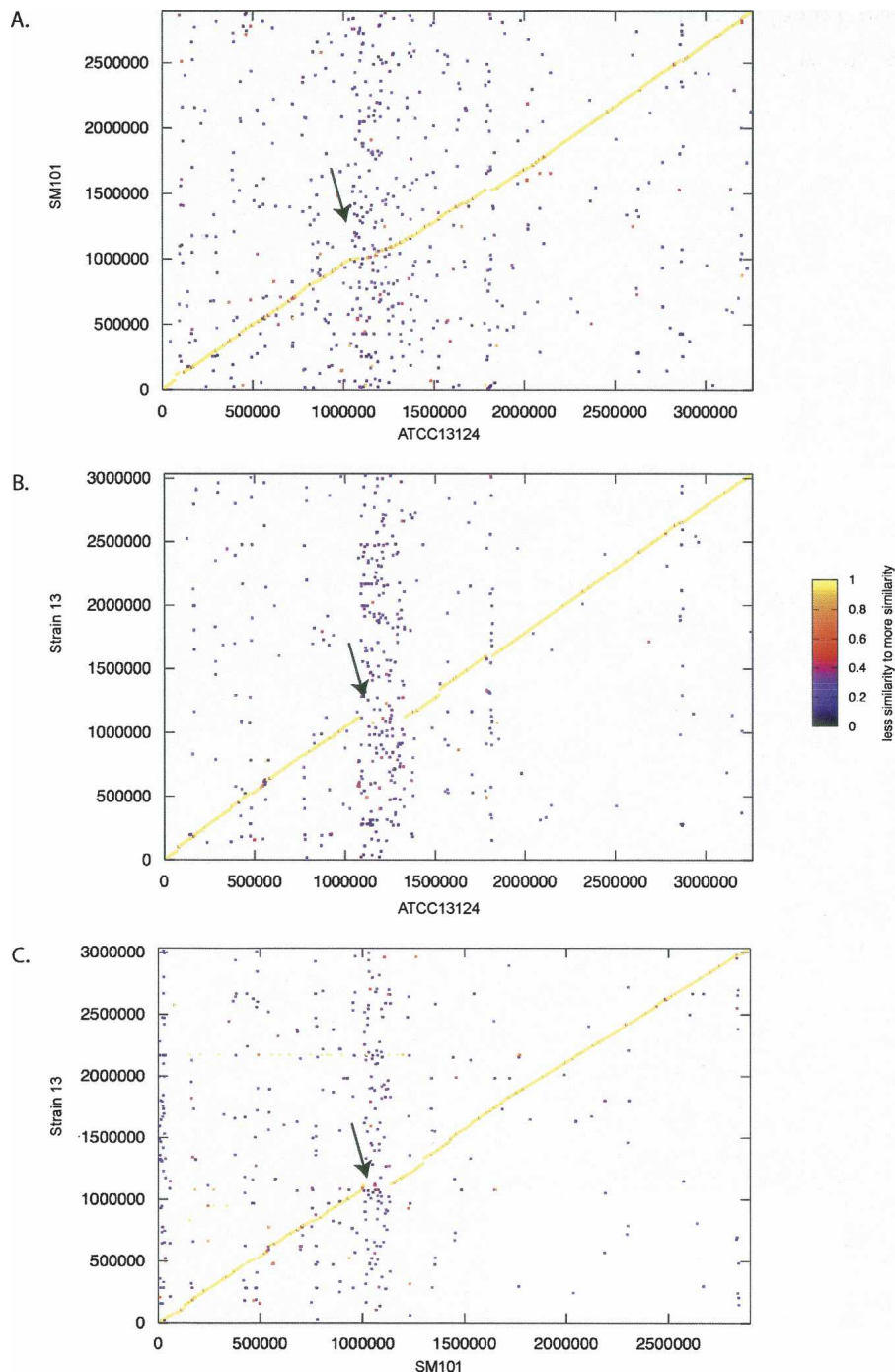
Figure 2. Synteny plots using the BLAST Score Ratio method (Rasko et al. 2005). (*A*) ATCC 13124 vs. SM101. (*B*) ATCC 13124 vs. Strain 13. (*C*) SM101 vs. Strain 13. Each plot point represents a single peptide in the indicated *C. perfringens* genome pair. The color of each point indicates the level of similarity of that peptide in each genome pair, according to Rasko et al. (2005). Arrows indicate the 242,969-bp island.

port genes, restriction-modification systems, a fucose utilization operon, and numerous genes encoding various glycosyl hydrolases (Supplemental Tables S1–S6).

Iron is a key element for bacterial growth and survival; consequently, iron limitation is a key host defense strategy against pathogens. Bacterial iron acquisition genes are widely reported in association with pathogenicity islands (Carniel 1999). All three *C. perfringens* strains possess siderophore-based systems for the binding, uptake, and utilization of biologically bound ferric iron (III). Under anaerobic conditions, however, soluble ferrous iron (II) is the predominant form (Andrews et al. 2003). All three strains encode homologs of the iron (II) uptake *feoAB* system (Kammler et al. 1993). The *feoAB* operon found in strain 13 is conserved in both ATCC 13124 and SM101; however, the latter strains contain additional copies of this operon. ATCC 13124 possesses two extra *feoAB* clusters, both located with the 243-kb island, whereas SM101 contains one of these clusters within the equivalent fragmented parts of this island in SM101. These additional *feoAB* operons may contribute to virulence by increasing the capability of these strains to cope with anaerobic iron limitation in the host.

Unlike *C. perfringens* strains 13 and SM101, *C. perfringens* strain ATCC 13124 is refractory to genetic transformation by electroporation. ATCC 13124 contains a putative type I restriction-modification system (CPF_2599–CPF_2596), three type II restriction endonucleases and associated methyltransferases (CPF_0140, CPF_1468, and CPF_1016), and a putative type III restriction modification system (CPF_0362–CPF_0361), all located within genomic regions unique to ATCC 13124. There is no evidence for any restriction-modification systems in SM101 or strain 13. The presence or absence of restriction-modification systems within these strains correlates with the observed differences in genetic malleability.

C. perfringens exploits many complex sugars from its environment (Sneath 1986); accordingly, a wide variety of genes encoding glycolytic activities are found in these genomes. Many of these genes are found within strain-specific genomic islands, indicating that each strain has both conserved and unique glycolytic repertoires, reflecting the different capabilities of each strain to exploit different environmental conditions. For example, the gas gangrene strain ATCC 13124 possesses a putative fucose utilization operon

ated with two of the large genomic islands in that strain. A 17-kb phage region, which includes a phage integrase gene, is located at the 5′ end of the 243-kb island in ATCC 13124. This 236-gene island includes a variety of genes that are not found in one or both of the other two *C. perfringens* strains and that may confer ATCC 13124-specific phenotypes, including ferrous iron trans-

within the 243-kb island, and a gene encoding α-1,3–4-fucosidase. These genes are also present in the histotoxic strain 13 but absent in the food poisoning pathogen SM101. L-fucose is typically a terminal α-linked component of cell-surface glycoconjugates, which are abundant in mucin and on host epithelial cells. For both the tissue-degrading *C. perfringens* ATCC 13124 and strain 13 isolates, fucosylated glycoconjugates are likely to be an abundant energy source both in the soil environment and in vivo, deriving from the necrotic effects of infection.

It has been recognized that *C. perfringens* α-toxin, as a known phospholipase C, has the capability to hydrolyze membrane phospholipids, such as sphingomyelin and various phosphoglycerides, including phospatidylcholine and phospatidylethanolamine (Bryant and Stevens 1997). Both ATCC 13124 and strain 13 possess the capability to utilize ethanolamine through the *eut* operon. SM101 lacks this ethanolamine utilization operon. Ethanolamine utilization has been proposed as an important survival strategy against nutritional deprivation (Stojiljkovic et al. 1995). As noted above with fucose utilization, ethanolamine will likely be in abundance in vivo following the necrotic effects of infection, with both gangrene-capable isolates able to make use of ethanolamine as a source of carbon and nitrogen.

Toxins

Members of the clostridia produce more toxins and secreted virulence factors than any other bacterial genus (Johnson 1997; Rood 1998); consequently, these toxins and virulence factors have been intensively studied, with the pattern of toxin production used as a key typing system for *C. perfringens*. While the strains examined in this study differ in the quantity of toxins produced (data not shown), the genome sequences of ATCC 13124 and SM101 show that both possess the expected chromosomally encoded toxins and secreted enzymes characteristic of type A myonecrotic and food poisoning isolates (Supplemental Table S10). Specifically, all three strains encode α-toxin (phospholipase C), whereas only ATCC 13124 and strain 13 encode θ-toxin (perfringolysin O). SM101 does not have the perfringolysin O structural gene, *pfoA*, and only SM101 encodes CPE.

One notable secreted enzyme, sialidase (neuraminidase), has long been recognized in isolates of *C. perfringens*, with three different sialidases identified: NanH (non-secreted, 43 kDa), NanI (secreted, 77 kDa) and NanJ (secreted, 122 kDa). These enzymes could act on the abundant sialic acid conjugates present in the soil and mammalian intestinal tracts, providing an important nutritional source. Additionally, these enzymes could contribute to the pathogenesis of gas gangrene, for example, by enhancing the sensitivity of host cells to the action of α-toxin (Flores-Diaz et al. 2005). ATCC 13124 possesses all three sialidases, while strain 13 has NanI and NanJ but lacks the non-secreted NanH. The food poisoning strain, SM101, possesses only the non-secreted sialidase NanH, which is thought to act on sialic acids conjugated to short oligosaccharides that are able to be transported across the bacterial membrane (Roggentin and Schauer 1997; Walters et al. 1999). This lack of secreted sialidase activity in SM101 further highlights the differences between the myonecrotic and food poisoning strains noted earlier.

Bacteriocins are ribosomally synthesized antimicrobial toxins, best known for a wide distribution amongst lactic acid bacteria (LAB), with a spectrum of activity against a variety of Gram-positive bacteria including other LAB, *Listeria,* and *Clostridium* (Nes and Johnsborg 2004). In addition, many *C. perfringens* iso-

lates produce bacteriocins (Rood and Cole 1991), such that bacteriocin typing was an early typing scheme of *C. perfringens*. Shimizu et al. (2002) reported a single bacteriocin gene (*bcn5*) on the strain 13 chromosome in association with the prophage remnant. No bacteriocin or accessory genes are found in the ATCC 13124 genome. By contrast, several bacteriocin loci are found in the SM101 genome and plasmid sequences. The two SM101 plasmids each bear a complete UV-inducible bacteriocin operon (*uviA-uviB-bcn5*), similar to the archetypal clostridial bacteriocin-bearing plasmid pIP404 (Rood and Cole 1991).

In addition to the plasmid-borne bacteriocins, the SM101 chromosome encodes several gene clusters, each flanked by IS1470 elements. One cluster of genes is similar to lactococcin A (*lcnA*) production, immunity and signal sequence-independent secretion genes (Supplemental Fig. S3A). Other clusters include genes with homology with *lagD*, an *L. lactis* ABC transporter with both proteolytic and translocation functions typically associated with lactococcin G, another *L. lactis* bacteriocin. Two of these *lagD*-containing clusters encode at least one putative novel bacteriocin each (CPR_0317 and CPR_0658) and a bacteriocin ABC transporter accessory gene. SM101 also encodes a truncated *bcn5*-type bacteriocin (CPR_0631) on the chromosome. A frameshifted plasmid replication protein (CPR_0633) lies downstream of this gene. This diversity of bacteriocins within SM101, a food-poisoning isolate, likely represents an adaptation for competition and survival in the complex microbial communities that *C. perfringens* commonly inhabits, such as sewage or soil. In addition, an IS*1470* element is also found immediately upstream of a lactocepin precursor gene, *prtP* (Supplemental Fig. S3A) that encodes a cell wall-associated serine protease.

The association of all chromosomally encoded bacteriocin gene clusters in SM101 with directly repeated IS*1470* elements is intriguing, particularly as the location of the CPE gene is linked to disease type. Isolates with a chromosomally encoded CPE gene will usually cause food poisoning, whereas isolates containing a CPE-bearing plasmid are typically associated with non-foodborne GI disease (Collie and McClane 1998). IS*1470* was originally described in *C. perfringens* food poisoning isolates in association with the gene encoding CPE (Brynestad et al. 1997). The chromosomal CPE locus in SM101, and probably most or all other type A food poisoning isolates carrying a chromosomal *cpe* gene, consists of directly repeated IS*1470* elements flanking the *cpe* gene (CPR_0381) and IS*1469*, as previously reported (Brynestad et al. 1997; Miyamoto et al. 2002)—this configuration is similar to that observed in the bacteriocin loci discussed above (Supplemental Fig. S3B). Interestingly, considerably different arrangements of IS elements are associated with plasmid *cpe* loci of type A isolates. Thus, IS*1470* elements appear to be directly linked to the dissemination of toxins and other putative virulence factors within food-poisoning isolates of *C. perfringens*, with many of these genes appearing to originate from LAB and other gastrointestinal inhabitants.

Sporulation

The production of heat-resistant endospores plays an important role in food poisoning and gangrene infections caused by *C. perfringens*. Gangrene is often initiated from contamination of a wound with spores that subsequently germinate in vivo, leading to the characteristic rapid infection and associated gangrenous tissue damage (Stevens 1997). Both strain 13 and ATCC 13124 have been shown to cause gas gangrene in a mouse myonecrosis

model (Stevens et al. 1997), but SM101 did not cause this disease when tested in the same model (data not shown). In cases of acute food poisoning, multiple studies have shown that only sporulating forms of *C. perfringens* bearing the *cpe* gene will produce the enterotoxin (CPE) responsible for the associated diarrhea (Melville et al. 1997).

Analysis of sporulation genes has identified 63 homologs of sporulation-associated genes in SM101 and ATCC 13124. Strain 13 possessed the same complement with the exception that it is missing spore coat composition proteins CotJB and CotJC. Growth of these strains in sporulation medium resulted in strain SM101, strain 13, and ATCC 13124 producing sporulation efficiencies of 70% ($\sim 1 \times 10^7$ heat-resistant spores/mL), 0.1% (1×10^5 heat-resistant spores/mL), and 0% (<10 spores/mL, Supplemental Fig. S4), respectively. Similar results were obtained when the spores were tested for resistance to UV irradiation and desiccation (data not shown). Whether the absence of CotJB and CotJC is involved in the lower sporulation efficiency of strain 13 is unknown, nor is there an obvious genomic explanation for the inability of ATCC 13124 to sporulate effectively.

The sequence of the master sporulation regulatory gene *spo0A* in the poorly sporulating strain 13 isolate analyzed by Shimizu et al. (2002) indicated a premature stop codon was present near the center of the coding region. One of our laboratories (that of S.B.M.) had another version of strain 13, obtained from a different source than the sequenced strain that sporulates much better than the Shimizu strain (albeit still poorly compared with SM101). Analysis of the *spo0A* gene from this strain demonstrated that the premature stop codon observed by Shimizu et al. was not present; instead a codon encoding serine was found. The increased level of sporulation and presence of an intact *spo0A* gene in this version of strain 13 suggests the stop codon in the sequenced strain may be the result of a laboratory artifact resulting in a loss of function.

Environmental interactions

The production of several *C. perfringens* toxins and extracellular enzymes, including α-toxin (phospholipase C), perfringolysin O (θ-toxin), collagenase (κ-toxin), sialidase, and α-clostripain, are regulated by the VirS/VirR two-component signal transduction system (Rood 1998). The VirR binding site upstream of the perfringolysin O gene (*pfoA*) has been characterized in vitro (Cheung and Rood 2000; Cheung et al. 2004). Analysis of the strain 13 genome identified four additional predicted genes with putative VirR sites (Shimizu et al. 2002), which have also subsequently been shown to be functional (Cheung et al. 2004).

In both the ATCC 13124 and SM101 genomes, putative VirR boxes were identified upstream of several additional genes that were absent in strain 13 (data not shown). By use of a combination of a *pfoA* reporter construct (Cheung et al. 2004), perfringolysin O assays and gel mobility shift experiments (Fig. 3A) we were able to confirm that the VirR boxes upstream of CPF_1074 and CPF_0461 in ATCC 13124 are functional (Table 2). Both of these genes encode membrane proteins, with CPF_1074 also likely to be secreted. However, as no function can be attributed to these genes through bioinformatic annotation processes, their precise role in virulence remains to be determined. Numerous VirR boxes were also identified in association with novel genes in the SM101 genome, two of which were chosen for further analysis. The results showed that the VirR boxes upstream of CPR_0761 (Table 2) were active but that the VirR boxes upstream of CPR_1650 were not functional (Fig. 3B). The latter VirR boxes had less similarity to the consensus VirR box sequence. Since SM101 does not have the *pfoA* gene it was important to determine whether its VirR protein ($VirR_{101}$) was able to recognize and bind to the selected VirR boxes. Therefore, the *virR* gene from SM101 was cloned into an expression vector and purified (Cheung et al. 2004). Gel mobility shift analysis showed that this protein could bind to VirR boxes from both strain 13 and SM101 (Fig. 3B), providing evidence that SM101 has a functional VirR response regulator.

In addition to the key VirS/VirR system, phylogenetic analysis of two-component signal transduction systems in *C. perfringens* indicated that there were 23 sensor kinases and 17 response regulators conserved in all three strains (Supplemental Fig. S5). ATCC 13124 and strain 13 share a number of two-component systems that are absent in SM101, including systems predicted to be involved in sugar metabolism, ethanolamine metabolism, and osmoregulation.
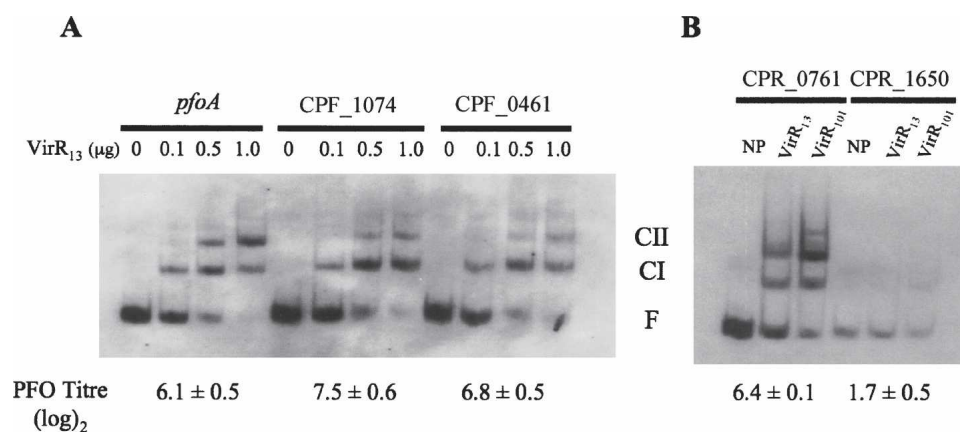

Figure 3. Gel mobility shift assay of VirR binding to novel putative VirR boxes from *C. perfringens* ATCC 13124 and SM101. Varying amounts of VirR protein from strain 13 ($VirR_{13}$) was used in *A*, whereas either no protein (NP) or 1.0 µg of either $VirR_{13}$ or the VirR protein from SM101 ($VirR_{101}$) was added in *B*. The genes downstream of the putative VirR boxes are shown *above* each gel section. The *pfoA* gene from strain 13 was used as the positive control. The respective perfringolysin O (PFO) titers obtained when the various VirR box regions fused to a *pfoA* gene were introduced into a *pfoA* mutant (Cheung and Rood 2000) are indicated *below* the gel.

Table 2. VirR binding of *C. perfringens* ATCC 13124 and SM101 genes associated with VirR boxes

Genome	Position	Strand	Match Sequence[a]	Description	VirR function?
ATCC 13124	2274124	+[b]	**CCA**A**TT**TCGTTAACTTTTCGA**CCAGTT**GCG**CAC**	CPF_1074: hypothetical secreted membrane protein	Yes
ATCC 13124	2951186	+[b]	**CCA**T**TT**ATGTTTAAAATTTGA**CCAGTT**TTG**CAC**	CPF_0461: hypothetical membrane protein	Yes
SM101	890287	+[b]	**CCA**G**TT**ATGTTTAAAATTTGA**CCA**a**TT**TTG**CAC**	CPR_0761: hypothetical membrane protein	Yes
SM101	1844335	+[c]	**C**t**ATTT**CCTTGATGAGTTATT**CCA**G**TT**ATAa**AC**	CPR_1650: aspartokinase	No

[a]Using the motif **CCA**N(1)**TT**N(15)**CCAGTT**N(3)**CAC**, as defined by Shimizu et al. (2002), with up to three mismatches.
[b]Upstream, intergenic.
[c]Intragenic.

C. perfringens may encounter many distinct environments from soil and water-saturated sediments to intestinal and other human tissues; consequently, several mechanisms that counteract oxidative and acid stresses are evident within the genomes. *C. perfringens* is an aerotolerant anaerobe, with growth being arrested in the presence of oxygen and associated oxidative compounds (Jean et al. 2004). All three genomes encode putative factors thought to contribute to an adaptive stress response, including a Mn/Fe superoxide dismutase, two methionine sulfoxide reductases (MsrA), and two glutathione peroxidases. In addition, ATCC 13124 and strain 13 encode a putative alkyl hydroperoxide reductase (AhpC) that is absent from SM101. Furthermore, all three genomes encode genes necessary for selenoprotein synthesis (*selADB*). Prokaryotic selenoproteins typically catalyze redox reactions and formation of selenoethers in stress-induced metabolism and energy production. The only additional identifiable selenoproteins in ATCC 13124 and SM101 are three members of the HesB family, which are predicted to be involved in Fe-S cluster assembly and are thus putative redox factors.

As with oxidative stress, acid stress is likely to be a key feature of the environments frequented by *C. perfringens*, particularly the mammalian GI tract, where pH will vary from the high acidity of the stomach to the neutral or slightly alkaline regions of the colon. All three genomes utilize the arginine deiminase pathway to counter acidification by production of ammonia from arginine. All three genomes also encode glutamate decarboxylase, which together with arginine deiminase plays a major role in lactococcal acid resistance. This similarity of clostridial acid stress response mechanisms to those of the lactococci is not surprising given that both inhabit or transit the mammalian GI tract. Extending this similarity, ATCC 13124 lacks an operon found in both SM101 and strain 13 that encodes factors for citrate fermentation. Many bacteria will use citrate under fermentative conditions but, notably, lactococci utilize the citrate fermentation pathway to improve the cometabolism of glucose and citrate at low pH and to alleviate the inhibitory effects of acid stress (Martin et al. 2004). It is possible the presence of this citrate fermentation pathway in SM101 and strain 13 represents a specific adaptation to acidic conditions.

Analysis of the strain 13 genome identified genes potentially involved in type IV pilin biogenesis (Kanehisa and Goto 2000; Shimizu et al. 2002), which are rarely seen in Gram-positive bacteria. While the N-terminal regions of the putative genes encoding the major component of type IV pili, PilA (or pilin), were highly conserved compared to *pilA* genes from Gram-negative bacteria, the rest of the gene sequences exhibit little similarity. This C-terminal variability is characteristic of PilA proteins from Gram-negative bacteria (Mattick 2002). Strain 13 and SM101 each contain two putative *pilA* genes, whereas strain ATCC 13124 has four putative *pilA* genes.

Capsule

A principal virulence determinant of many bacterial pathogens is the possession of an extracellular capsule. A striking feature of the three *C. perfringens* strains is the correlation of unique capsule-associated loci in each strain with genomic islands (Fig. 1). Comparative analysis of these loci indicates each strain is producing different extracellular carbohydrate structures. Strain 13 possesses one large capsular cluster, encoding 49 genes. This cluster is flanked on the 3′ end by a tRNA (tRNA–Met-7), suggesting that it may have been acquired and inserted into this genomic location, as is typical of pathogenicity islands (Morschhauser et al. 2000). Within this cluster, two regions exhibit a significant level of similarity and synteny to dispersed capsular gene clusters in the other two *C. perfringens* genomes, responsible for the synthesis and export of capsular structures. These genes, having capsular synthesis and export functions, may constitute a core set of clostridial capsule genes. The peptides encoded by CPE0494 to CPE0496 are potentially responsible for the activation and addition of mannose to a growing polysaccharide structure. This conclusion is supported by the recent experimental description of mannose residues in extracellular polysaccharide associated with *C. perfringens* strain 13 (O'Brien and Melville 2003); indeed, it has been suggested that the mannose-containing capsule of strain 13 is not antiphagocytic as observed in other Gram-positive bacteria but rather enables phagocytosis and subsequent escape from the phagosome of macrophages, leading to persistence within the macrophage cytoplasm (O'Brien and Melville 2003).

There are two primary capsular clusters in *C. perfringens* ATCC 13124, one of which is the putative core set of genes and the other is unique. This region is flanked on the 3′ end by tRNA–Met-3, as observed for the equivalent region in strain 13. However, this region encodes genes for the activation and addition of rhamnose, which is structurally and antigenically different from the mannose that is present in *C. perfringens* strain 13 capsule.

By contrast to strain 13 and ATCC 13124, there does not appear to be a single large capsule biosynthesis operon in the SM101 genome. However, there are three smaller gene clusters, each with a distinct putative function. The first cluster contains the conserved clostridial capsule genes already discussed. The second cluster does not include the genes necessary for producing an extracellular capsule but instead may be involved in the decoration of surface exposed structures with phosphocholine. Choline is thought to decorate outer membrane proteins and carbohydrate structures (Campbell and Kent 2001) and can be directly linked to an increase in pathogenesis. Additionally, this region contains two homologs of the PspA protein from *Staphy-*

lococcus aureus, which has been shown to be involved in antigenic variation. This combination of glycosyltransferases, choline decoration proteins, and potential antigenically variable peptides in a single genomic location suggests a role in immune evasion. However, the transient nature of food poisoning isolates within the mammalian GI tract, where immune responses are generally weaker than in other mammalian tissues, suggests other uses than immune evasion alone. The third region, CPR_0254–CPR_0259, encodes proteins likely to be involved in the synthesis of teichoic acid, and is conserved in the other two strains.

Most strains of *C. perfringens* possess serologically distinct and complex capsular structures, with several hundred serological types reported (Kalelkar et al. 1997). However, experimental data describing the makeup of the *C. perfringens* capsule or its role in virulence are limited and have not been an area of research interest historically. The association of significant genomic variability of *C. perfringens* isolates with capsule gene loci reported in this work indicates that expansion of research into the clostridial extracellular capsule is warranted.

Diversity of *C. perfringens* isolates

The presence of over 300 genomic islands located within *C. perfringens* SM101, ATCC 13124, and strain 13 revealed an unexpected degree of genomic variation. To investigate the degree of diversity among a broader range of *C. perfringens* isolates, we undertook a PCR-based survey of a collection of 46 *C. perfringens* strains from both human and animal sources, different geographical regions, and including representatives of biotypes A through E (Supplemental Table S11). We designed primers internal to, and flanking, the four largest regions unique to ATCC 13124 (Supplemental Table S12). Two of these regions (the 17-kb and 37-kb loci) are absent in almost all strains tested. The large 243-kb island, which is present in a fragmented form in SM101, appears to be present but highly diverse amongst many of the strains tested. None of the strains tested appeared to have the entire island as seen in ATCC 13124; however, all but one of the strains contained at least one portion, although the strains differ with respect to the specific portions conserved. The remaining island (48-kb phage) was present in whole or in part in 16/48 strains, especially among food-poisoning isolates. There is no obvious pattern of conservation of these four regions that appears to be specific to any of the biotypes, or to the host species of origin. It should be recognized that this strain survey is limited in terms of the number of regions investigated and by the use of PCR, which is subject to experimental variability and does not take into account genomic location or variation at the primer binding sites. Nevertheless, these data reinforce the findings from the genome sequencing that there is a high degree of genomic diversity between *C. perfringens* strains and within the genomic islands found in these strains.

Conclusion

Completion and comparison of the three *C. perfringens* complete genome sequences has revealed an unexpected degree of genome variability between isolates of this species. Many of the gene content differences observed are contained within discrete islands unusually clustered on one replichore. In addition, PCR-based surveys of selected islands across additional *C. perfringens* isolates indicate that this observed diversity is even more varied when other strains are considered. Analysis of these islands has enabled the investigation of numerous genomic features that

may contribute to the different disease phenotypes and virulence characteristics of these pathogenic strains. These include mobile elements, metabolic capabilities, extracellular capsules, toxins and other secreted enzymes that act on the surroundings and that enable *C. perfringens* to adapt to its local environment, including human tissue.

Methods

Genome sequencing and annotation

The complete genome sequences of *C. perfringens* ATCC 13124 and SM101 were determined using the whole-genome shotgun method (Fraser et al. 1997). Physical and sequencing gaps were closed using a combination of primer walking, generation and sequencing of transposon-tagged libraries of large-insert clones, and multiplex PCR (Tettelin et al. 1999). Identification of putative protein-encoding genes and annotation of the genome were performed as previously described (Fraser et al. 1997). An initial set of open reading frames (ORFs) predicted to encode proteins was identified with GLIMMER (Delcher et al. 1999). ORFs consisting of fewer than 30 codons and those containing overlaps were eliminated. Frame shifts and point mutations were corrected or designated "authentic." Functional assignment, identification of membrane-spanning domains, determination of paralogous gene families, and identification of regions of unusual nucleotide composition were performed as previously described (Fraser et al. 1997). Sequence alignments and phylogenetic trees were generated using the methods described previously (Fraser et al. 1997).

Trinucleotide composition

Distribution of all 64 trinucleotides (3mers) was determined, and the 3mer distribution in 2000-bp windows that overlapped by half their length (1000 bp) across the genome was computed. For each window, we computed the χ^2 statistic on the difference between its 3mer content and that of the whole chromosome. A large value for χ^2 indicates the 3mer composition in this window is different from the rest of the chromosome. Probability values for this analysis are based on assumptions that the DNA composition is relatively uniform throughout the genome and that 3mer composition is independent. Because these assumptions may be incorrect, we prefer to interpret high χ^2 values as indicators of regions on the chromosome that appear unusual and demand further scrutiny.

Comparative genomics

C. perfringens ATCC 13124, SM101, and other sequenced clostridial genomes were compared at the nucleotide level by suffix tree analysis using MUMmer (Delcher et al. 2002) and the data parsed by custom Perl scripts. Predicted *C. perfringens* genes were compared by BLAST against the complete set of genes from other clostridial genomes using an E-value cutoff of 10^{-5}. Synteny and BLAST Score Ratio analyses were performed as previously described (Rasko et al. 2005).

Validation of VirR binding sites

Expression and purification of His-tagged VirR protein, perfringolysin O (PFO) assays, and the construction of recombinant plasmids for determining the function of putative VirR boxes were performed as previously described (Cheung and Rood 2000). These constructs were also used as templates in PCR amplification with oligonucleotides #4565 and #5126 to generate the probes used in gel mobility shift assays (Cheung and Rood 2000). Each separate 183-bp DNA probe consisted of the putative

VirR boxes being tested and the VirR-dependent promoter region upstream of the *pfoA* gene. Probes were end-labeled with digoxigenin-11-ddUTP and used in binding reactions with varying amounts of VirR protein, 1 μg of poly[d(I-C)], and 0.1 μg of poly-lysine in binding buffer [20 mM HEPES at pH 7.6, 1 mM EDTA, 10 mM $(NH_4)_2SO_4$, 1 mM DTT, 0.2% (w/v) Tween$_{20}$, 30 mM KCl]. Following incubation at room temperature for 15 min, binding reactions were separated by electrophoresis in a 6% native $0.5\times$ TBE polyacrylamide gel. Shifted DNA transferred onto Hybond Nylon+ membranes (Amersham) and detected using CSPD (Roche) as per the manufacturer's instructions. Relevant strains and plasmids are listed in Supplemental Table S13 and oligonucleotide primers in Supplemental Table S14.

Acknowledgments

References

Andrews, S.C., Robinson, A.K., and Rodriguez-Quinones, F. 2003. Bacterial iron homeostasis. *FEMS Microbiol. Rev.* **27:** 215–237.

Awad, M.M., Bryant, A.E., Stevens, D.L., and Rood, J.I. 1995. Virulence studies on chromosomal α-toxin and θ-toxin mutants constructed by allelic exchange provide genetic evidence for the essential role of α-toxin in *Clostridium perfringens*-mediated gas gangrene. *Mol. Microbiol.* **15:** 191–202.

Awad, M.M., Ellemor, D.M., Boyd, R.L., Emmins, J.J., and Rood, J.I. 2001. Synergistic effects of α-toxin and perfringolysin O in *Clostridium perfringens*-mediated gas gangrene. *Infect. Immun.* **69:** 7904–7910.

Bruggemann, H., Baumer, S., Fricke, W.F., Wiezer, A., Liesegang, H., Decker, I., Herzberg, C., Martinez-Arias, R., Merkl, R., Henne, A., et al. 2003. The genome sequence of *Clostridium tetani*, the causative agent of tetanus disease. *Proc. Natl. Acad. Sci.* **100:** 1316–1321.

Bryant, A.E. and Stevens, D.L. 1997. The pathogenesis of gas gangrene. In *The Clostridia. Molecular biology, and pathogenesis* (eds. J.I. Rood, et al.), pp. 185–196. Academic Press, London.

Brynestad, S., Synstad, B., and Granum, P.E. 1997. The *Clostridium perfringens* enterotoxin gene is on a transposable element in type A human food poisoning strains. *Microbiol.* **143:** 2109–2115.

Campbell, H.A. and Kent, C. 2001. The CTP:phosphocholine cytidylyltransferase encoded by the licC gene of *Streptococcus pneumoniae*: Cloning, expression, purification, and characterization. *Biochim. Biophys. Acta* **1534:** 85–95.

Carniel, E. 1999. The *Yersinia* high-pathogenicity island. *Int. Microbiol.* **2:** 161–167.

Cheung, J.K. and Rood, J.I. 2000. The VirR response regulator from *Clostridium perfringens* binds independently to two imperfect direct repeats located upstream of the pfoA promoter. *J. Bacteriol.* **182:** 57–66.

Cheung, J.K., Dupuy, B., Deveson, D.S., and Rood, J.I. 2004. The spatial organization of the VirR boxes is critical for VirR-mediated expression of the perfringolysin O gene, pfoA, from *Clostridium perfringens*. *J. Bacteriol.* **186:** 3321–3330.

Collie, R.E. and McClane, B.A.. 1998. Evidence that the enterotoxin gene can be episomal in *Clostridium perfringens* isolates associated with non-food-borne human gastrointestinal diseases. *J. Clin. Microbiol.* **36:** 30–36.

Delcher, A.L., Harmon, D., Kasif, S., White, O., and Salzberg, S.L. 1999. Improved microbial gene identification with GLIMMER. *Nucleic Acids Res.* **27:** 4636–4641.

Delcher, A.L., Phillippy, A., Carlton, J., and Salzberg, S.L. 2002. Fast algorithms for large-scale genome alignment and comparison. *Nucleic Acids Res.* **30:** 2478–2483.

Eisen, J.A., Heidelberg, J.F., White, O., and Salzberg, S.L. 2000. Evidence for symmetric chromosomal inversions around the replication origin in bacteria. *Genome Biol.* **1:** research0011.

Flores-Diaz, M., Alape-Giron, A., Clark, G., Catimel, B., Hirabayashi, Y., Nice, E., Gutierrez, J.M., Titball, R., and Thelestam, M. 2005. A cellular deficiency of gangliosides causes hypersensitivity to *Clostridium perfringens* phospholipase C. *J. Biol. Chem.* **280:** 26680–26689.

Fraser, C.M., Casjens, S., Huang, W.M., Sutton, G.G., Clayton, R., Lathigra, R., White, O., Ketchum, K.A., Dodson, R., Hickey, E.K., et al. 1997. Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*. *Nature* **390:** 580–586.

Jean, D., Briolat, V., and Reysset, G. 2004. Oxidative stress response in *Clostridium perfringens*. *Microbiol.* **150:** 1649–1659.

Johnson, E.A. 1997. Extrachromosomal virulence determinants in the Clostridia. In *The Clostridia. Molecular biology and pathogenesis* (eds. J.I. Rood, et al.), pp. 35–48. Academic Press, London.

Kalelkar, S., Glushka, J., van Halbeek, H., Morris, L.C., and Cherniak, R. 1997. Structure of the capsular polysaccharide of *Clostridium perfringens* Hobbs 5 as determined by NMR spectroscopy. *Carbohydr. Res.* **299:** 119–128.

Kammler, M., Schon, C., and Hantke, K. 1993. Characterization of the ferrous iron uptake system of *Escherichia coli*. *J. Bacteriol.* **175:** 6212–6219.

Kanehisa, M. and Goto, S. 2000. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28:** 27–30.

Klietmann, W.F. and Ruoff, K.L. 2001. Bioterrorism: Implications for the clinical microbiologist. *Clin. Microbiol. Rev.* **14:** 364–381.

Martin, M.G., Sender, P.D., Peiru, S., de Mendoza, D., and Magni, C. 2004. Acid-inducible transcription of the operon encoding the citrate lyase complex of *Lactococcus lactis* Biovar diacetylactis CRL264. *J. Bacteriol.* **186:** 5649–5660.

Mattick, J.S. 2002. Type IV pili and twitching motility. *Annu. Rev. Microbiol.* **56:** 289–314.

McClane, B. 2001. *Clostridium perfringens*. American Society for Microbiology Press, Washington, DC.

McClane, B., Lyerly, D., Moncrief, J.S., and Wilkins, T. 2000. Enterotoxic Clostridia: *Clostridium perfringens* and *Clostridium difficile*. In *Gram Positive Pathogens* (eds. V.A. Fischetti, et al.), pp. 551–562. ASM Press, Washington, DC.

Mead, P.S., Slutsker, L., Dietz, V., McCaig, L.F., Bresee, J.S., Shapiro, C., Griffin, P.M., and Tauxe, R.V. 1999. Food-related illness and death in the United States. *Emerg. Infect. Dis.* **5:** 607–625.

Melville, S.B., Collie, R.E., and McClane, B.A. 1997. Regulation of enterotoxin production in *Clostridium perfringens*. In *The Clostridia: Molecular biology and pathogenesis* (eds. J.I. Rood, et al.), pp. 471–490. Academic Press, London.

Miyamoto, K., Chakrabarti, G., Morino, Y., and McClane, B.A. 2002. Organization of the plasmid cpe locus in *Clostridium perfringens* type A isolates. *Infect. Immun.* **70:** 4261–4272.

Mollby, R. and Holme, T. 1976. Production of phospholipase C (α-toxin), haemolysins and lethal toxins by *Clostridium perfringens* types A to D. *J. Gen. Microbiol.* **96:** 137–144.

Morschhauser, J., Kohler, G., Ziebuhr, W., Blum-Oehler, G., Dobrindt, U., and Hacker, J. 2000. Evolution of microbial pathogens. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **355:** 695–704.

Nes, I.F. and Johnsborg, O. 2004. Exploration of antimicrobial potential in LAB by genomics. *Curr. Opin. Biotechnol.* **15:** 100–104.

Nolling, J., Breton, G., Omelchenko, M.V., Makarova, K.S., Zeng, Q., Gibson, R., Lee, H.M., Dubois, J., Qiu, D., Hitti, J., et al. 2001. Genome sequence and comparative analysis of the solvent-producing bacterium *Clostridium acetobutylicum*. *J. Bacteriol.* **183:** 4823–4838.

O'Brien, D.K. and Melville, S.B. 2003. Multiple effects on *Clostridium perfringens* binding, uptake and trafficking to lysosomes by inhibitors of macrophage phagocytosis receptors. *Microbiol.* **149:** 1377–1386.

Rasko, D.A., Myers, G.S., and Ravel, J. 2005. Visualization of comparative genomic analyses by BLAST score ratio. *BMC Bioinformatics* **6:** 2.

Roggentin, P. and Schauer, R. 1997. *Clostridial sialidases*. In *The Clostridia. Molecular biology and pathogenesis* (eds. J.I. Rood, et al.), pp. 421–437. Academic Press, London.

Rood, J.I. 1998. Virulence genes of *Clostridium perfringens*. *Annu. Rev. Microbiol.* **52:** 333–360.

Rood, J.I. and Cole, S.T. 1991. Molecular genetics and pathogenesis of *Clostridium perfringens*. *Microbiol. Rev.* **55:** 621–648.

Sarker, M.R., Carman, R.J., and McClane, B.A. 1999. Inactivation of the gene (cpe) encoding *Clostridium perfringens* enterotoxin eliminates the ability of two cpe-positive *C. perfringens* type A human gastrointestinal disease isolates to affect rabbit ileal loops. *Mol. Microbiol.* **33:** 946–958.

Shimizu, T., Ohtani, K., Hirakawa, H., Ohshima, K., Yamashita, A., Shiba, T., Ogasawara, N., Hattori, M., Kuhara, S., and Hayashi, H. 2002. Complete genome sequence of *Clostridium perfringens*, an anaerobic flesh-eater. *Proc. Natl. Acad. Sci.* **99:** 996–1001.

Sneath, P.H.A. 1986. Endospore-forming Gram positive rods and cocci. In *Bergey's Manual of Systematic Bacteriology* (eds. P.H.A. Sneath, et al.), pp. 1104–1207. Williams & Wilkins, Baltimore, MD.

Songer, J.G. 1997. Clostridial disease of animals. In *The Clostridia. Molecular biology and pathogenesis* (eds. J.I. Rood, et al.), pp. 154–182. Academic Press, London.

Stevens, D.L. 1997. Necrotizing Clostridial soft tissue infections. In *The Clostridia: Molecular biology and pathognesis*. (eds. J.I. Rood, et al.), pp. 141–152. Academic Press, London.

Stevens, D.L., Tweten, R.K., Awad, M.M., Rood, J.I., and Bryant, A.E. 1997. Clostridial gas gangrene: Evidence that α and θ toxins differentially modulate the immune response and induce acute tissue necrosis. *J. Infect. Dis.* **176:** 189–195.

Stojiljkovic, I., Baumler, A.J., and Heffron, F. 1995. Ethanolamine utilization in *Salmonella typhimurium*: Nucleotide sequence, protein expression, and mutational analysis of the cchA cchB eutE eutJ eutG eutH gene cluster. *J. Bacteriol.* **177:** 1357–1366.

Suyama, M. and Bork, P. 2001. Evolution of prokaryotic gene order: Genome rearrangements in closely related species. *Trends Genet.* **17:** 10–13.

Tettelin, H., Radune, D., Kasif, S., Khouri, H., and Salzberg, S.L. 1999. Optimized multiplex PCR: Efficiently closing a whole-genome shotgun sequencing project. *Genomics* **62:** 500–507.

Tillier, E.R. and Collins, R.A. 2000. Genome rearrangement by replication-directed translocation. *Nat. Genet.* **26:** 195–197.

Walters, D.M., Stirewalt, V.L., and Melville, S.B. 1999. Cloning, sequence, and transcriptional regulation of the operon encoding a putative N-acetylmannosamine-6-phosphate epimerase (nanE) and sialic acid lyase (nanA) in *Clostridium perfringens*. *J. Bacteriol.* **181:** 4526–4532.

Zimmer, M., Scherer, S., and Loessner, M.J. 2002. Genomic analysis of *Clostridium perfringens* bacteriophage phi3626, which integrates into guaA and possibly affects sporulation. *J. Bacteriol.* **184:** 4359–4368.

Skewed genomic variability in strains of the toxigenic bacterial pathogen, *Clostridium perfringens*

Garry S.A. Myers, David A. Rasko, Jackie K. Cheung, et al.

Supplemental Material	http://genome.cshlp.org/content/suppl/2006/07/07/gr.5238106.DC1
References	This article cites 39 articles, 16 of which can be accessed free at: http://genome.cshlp.org/content/16/8/1031.full.html#ref-list-1
License	
Email Alerting Service	Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here.**

To subscribe to *Genome Research* go to:
https://genome.cshlp.org/subscriptions