

Bacteriophage 2851 Is a Prototype Phage for Dissemination of the Shiga Toxin Variant Gene 2c in *Escherichia coli* O157:H7[†]

Eckhard Strauch,¹ Jens Andre Hammerl,¹ Antje Konietzny,¹ Susanne Schneiker-Bekel,³ Walter Arnold,³ Alexander Goesmann,⁴ Alfred Pühler,³ and Lothar Beutin^{2*}

Molecular Diagnostics and Genetics, Department of Biological Safety, Federal Institute for Risk Assessment (BfR), 12277 Berlin, Germany¹; National Reference Laboratory for *Escherichia coli*, Federal Institute for Risk Assessment (BfR), 12277 Berlin, Germany²; Department of Genetics, Bielefeld University, 33594 Bielefeld, Germany³; and Center for Biotechnology (CeBiTec), Bielefeld University, D-33594 Bielefeld, Germany⁴

Received 16 July 2008/Returned for modification 19 August 2008/Accepted 19 September 2008

The production of Shiga toxin (Stx) (verocytotoxin) is a major virulence factor of *Escherichia coli* O157:H7 strains (Shiga toxin-producing *E. coli* [STEC] O157). Two types of Shiga toxins, designated Stx1 and Stx2, are produced in STEC O157. Variants of the Stx2 type (Stx2, Stx2c) are associated with high virulences of these strains for humans. A bacteriophage designated 2851 from a human STEC O157 encoding the Stx2c variant was described previously. Nucleotide sequence analysis of the phage 2851 genome revealed 75 predicted coding sequences and indicated a mosaic structure typical for lambdoid phages. Analyses of free phages and K-12 phage 2851 lysogens revealed that upon excision from the bacterial chromosome, the loss of a phage-encoded IS629 element leads to fusion of phage *antA* and *antB* genes, with the generation of a recombinant *antAB* gene encoding a strong antirepressor. In wild-type *E. coli* O157 as well as in K-12 strains, phage 2851 was found to be integrated in the *sbcB* locus. Additionally, phage 2851 carries an open reading frame which encodes an OspB-like type III effector similar to that found in *Shigella* spp. Investigation of 39 *stx*_{2c} *E. coli* O157 strains revealed that all except 1 were positive for most phage 2851-specific genes and possessed a prophage with the same border sequences integrated into the *sbcB* locus. Phage 2851-specific sequences were absent from most *stx*_{2c}-negative *E. coli* O157 strains, and we suggest that phage 2851-like phages contributed significantly to the dissemination of the Stx2c variant toxin within this group of *E. coli*.

The production of Shiga toxins (Stx) in *Escherichia coli* is associated with strains causing illness in humans. Two major toxin families, called Stx1 and Stx2, were established according to differences in their nucleotide sequences and their antigenicities. Stx-producing strains of serotype O157:H7 were most frequently associated with cases of hemorrhagic disease in humans, such as hemorrhagic colitis and hemolytic uremic syndrome (15, 40). Investigation of Shiga toxin-producing *E. coli* (STEC) O157:H7 for their Shiga toxins revealed that three types, namely, Stx1, Stx2, and Stx2c, occur alone or in combination in these strains (1, 6, 33, 51). They are encoded by *stx* genes on the genome of lysogenic lambdoid bacteriophages that are inserted into different loci in the chromosome of their bacterial host (24). Prototype Stx1- and Stx2-encoding phages were found to differ for their integration sites in the *E. coli* chromosome (45, 46). Accordingly, STEC O157:H7 strains positive for more than one Stx toxin are frequently found, and the spontaneous loss of inserted Stx phages may occasionally occur (9). Stx2 and Stx2c were significantly associated with hemorrhagic colitis and hemolytic uremic syndrome in human patients, indicating that the virulence of an STEC strain is

determined by its Stx type (18, 19). It was previously reported that O157:H7 strains encoding Stx2c produce lower amounts of toxin and might be less virulent than those producing Stx2 (6, 8, 38). The genes encoding Stx2 and Stx2c in *E. coli* O157 have been sequenced and show >99% identity in the A subunit and approximately 96% identity in the B subunit (25, 36). The prototype for the *stx*_{2c} gene was described for the O157:NM strain E32511, a reference strain which harbors two Shiga toxin-encoding genes: *stx*₂, composed of *stxA*₂ and *stxB*₂, and the *stx*_{2c} gene, composed of *stxA*₂ and *stxB*_{2c} (44).

For *stx* genotyping of larger numbers of O157:H7 strains, PCR amplification of *stx* genes followed by restriction enzyme digestion of PCR products giving characteristic patterns for the Stx subtypes was employed (6). Primers developed by Lin et al. (30) give in the case of the *stx*_{2c} allele a product of 903 bp, which can be distinguished from the *stx*₂ gene by a HincII digestion. Such PCR-restriction fragment length polymorphism (RFLP) investigations revealed that *stx*₁ and *stx*_{2c} as well as *stx*₂ and *stx*_{2c} are frequently associated in O157:H7 and strains encoding only Stx2c are more rarely isolated (6, 48, 52). While Stx1 and Stx2 are disseminated over a large variety of STEC serotypes, the Stx2c type was found to be restricted mostly to STEC serotype O157:H7 and, moreover, to certain clonal lineages of that serotype (54); however, the reasons for this finding are not known. Interestingly, only Stx2 and not Stx1 or Stx2c was found in isolates of sorbitol-fermenting O157:NM (nonmotile), which is an emerging highly virulent subtype of the STEC O157 group (9, 10).

In a previous study, we isolated an Stx2c-encoding bacterio-

* Corresponding author. Mailing address: National Reference Laboratory for *Escherichia coli*, Federal Institute for Risk Assessment (BfR), Diederdsdorfer Weg 1, D-12277 Berlin, Germany. Phone: 49 30 8412 2259. Fax: 49 30 8412 2983. E-mail: l.beutin@bfr.bund.de.

[†] Supplemental material for this article may be found at <http://iai.asm.org/>.

[‡] Published ahead of print on 29 September 2008.

TABLE 1. Bacterial strains and plasmids

<i>E. coli</i> strains and plasmids (source)	Description ^a	Origin or reference
<i>E. coli</i> strains		
Genehogs	F ⁻ <i>mcrA</i> Δ(<i>mrr-hsdRMS-mcrBC</i>) φ80Δ <i>lacZ</i> Δ <i>M15</i> Δ <i>lacX74</i> <i>recA1</i> <i>araD139</i> Δ(<i>ara-leu</i>)7697 <i>galU galK rpsL</i> (Sm ^r) <i>endA1 nupG fluA::IS2</i> (confers phage T1 resistance)	Invitrogen
C600	K-12, <i>stx</i> negative, <i>q</i> negative, O rough; Lac ⁻ Trp ⁺ His ⁺	48
CB2851	Wild-type O157:H7, <i>stx</i> _{2c}	48
TPE1874	C600, lysogenic for phage 2851	48
TPE1875	C600, lysogenic for phage 2851	48
Plasmids		
pLitmus38	2.2-kb cloning vector; Ap ^r	New England Biolabs
pMS470Δ8	6.9-kb cloning vector, pMB1 replicon, P _{<i>lacI</i>} , <i>lacI</i> ; T7 gene Shine-Dalgarno sequence; Ap ^r	4
pJH701 (CB2851)	pMS470Δ8 Δ[NdeI HindIII] Ω[phage <i>cl</i> , PCR NdeI-HindIII]	This work
pJH701X (CB2851)	pMS470Δ8 Δ[NdeI HindIII] Ω[phage 2851 <i>cl</i> negative, PCR NdeI-HindIII]	This work
pJH702 (CB2851)	pMS470Δ8 Δ[NdeI HindIII] Ω[phage 2851 <i>cII</i> , PCR NdeI-HindIII]	This work
pJH702X (CB2851)	pMS470Δ8 Δ[NdeI HindIII] Ω[phage 2851 <i>cII</i> negative, PCR NdeI-HindIII]	This work
pJH703 (CB2851)	pMS470Δ8 Δ[NdeI HindIII] Ω[phage 2851 <i>antA</i> , PCR NdeI-HindIII]	This work
pJH703X (CB2851)	pMS470Δ8 Δ[NdeI HindIII] Ω[phage 2851 <i>antA</i> negative, PCR NdeI-HindIII]	This work
pJH704 (CB2851)	pMS470Δ8 Δ[NdeI HindIII] Ω[phage 2851 <i>antB</i> , PCR NdeI-HindIII]	This work
pJH704X (CB2851)	pMS470Δ8 Δ[NdeI HindIII] Ω[phage 2851 <i>antB</i> negative, PCR NdeI-HindIII]	This work
pJH705 (TPE1874) ^b	pMS470Δ8 Δ[NdeI HindIII] Ω[phage 1874 <i>antAB</i> , PCR NdeI-HindIII]	This work
pJH705X (TPE1874)	pMS470Δ8 Δ[NdeI HindIII] Ω[phage 1874 <i>antAB</i> negative, PCR NdeI-HindIII]	This work
pJH707 (CB2851)	pMS470Δ8 Δ[NdeI HindIII] Ω[phage 2851 <i>cro</i> , PCR NdeI-HindIII]	This work
pJH707X (CB2851)	pMS470Δ8 Δ[NdeI HindIII] Ω[phage 2851 <i>cro</i> negative, PCR NdeI-HindIII]	This work
pJH709 (CB2851)	pMS470Δ8 Δ[NdeI HindIII] Ω[phage 2851 <i>ospB</i> , PCR NdeI-HindIII]	This work
pJH709X (CB2851)	pMS470Δ8 Δ[NdeI HindIII] Ω[phage 2851 <i>ospB</i> negative, PCR NdeI-HindIII]	This work

^a Ap^r, ampicillin resistance; Sm^r, streptomycin resistance.

^b *antAB* fusion gene generated from 5' region of *antA* and 3' region of *antB* (see text).

phage designated 2851 from an STEC O157:H7 strain which originated from a human patient (48). An 8.4-kb DNA region flanking the *stx*_{2c} gene in phage 2851 was sequenced and compared to the classical Stx2 phage 933W, revealing differences in regulatory genes, especially in the *q* gene, which is involved in the expression of late genes and the Shiga toxin genes. An investigation of 82 *E. coli* O157:H7 strains carrying *stx*₁, *stx*₂, and/or *stx*_{2c} genes revealed that these genetic differences are closely associated with strains carrying *stx*_{2c} genes. This suggests that *stx*_{2c} was disseminated in enterohemorrhagic *E. coli* (EHEC) O157 by a phage genetically similar to phage 2851. The phage genome sequencing results of the present study corroborate our previous suggestions that phage 2851 is the prototype for the global spread of the *stx*_{2c} gene in EHEC O157.

The identification of a specific chromosomal integration site for phage 2851 explains the coexistence of this determinant in strains carrying phage-encoded *stx*₁ or *stx*₂ genes. Comparison of phage *stx*_{2c} and neighboring sequences to sequences in non-O157 STEC strains indicates that phage 2851 is restricted to the O157 serotype and that Stx2c determinants harbored in non-O157 strains are encoded by other bacteriophages.

MATERIALS AND METHODS

Bacterial strains. The original *E. coli* O157:H7 strain, namely, CB2851, 2 phage 2851 lysogenic *E. coli* K-12 strains (TPE1874 and TPE1875), and 81 *E. coli* O157:H7 strains have been described previously (48). *E. coli* strain C600 was routinely used for the propagation of Stx phages (Table 1). All strains were incubated at 37°C in Luria-Bertani (LB) broth. Solid and soft agar media contained 1.8% and 0.7% (wt/vol) agar, respectively. *E. coli* K-12 strain Genehogs (Invitrogen, Karlsruhe, Germany) was used as the host for plasmid propagation

and standard cloning procedures. When appropriate, ampicillin was added to a final concentration of 100 μg/ml.

Induction and isolation of bacteriophages. The induction of temperate bacteriophages from *E. coli* strains CB2851, TPE1874, and TPE1875 was performed with mitomycin C (final concentration, 1 μg/ml) as described previously (48). After induction, the bacteria were sedimented by centrifugation, and the culture fluid was filtered through 0.45-μm-pore-size membranes (Schleicher and Schüll, Dassel, Germany). Phages from dilutions of supernatants were titrated on *E. coli* K-12 strain C600. Isolation of single plaques, preparation of high-titer phage stocks, and purification of phage DNA were performed as described previously (47, 48).

DNA library construction, nucleotide sequencing, and sequence analysis of phage 2851. DNA of phage 2851 was digested with the restriction enzymes HincII, BspI407I, EcoRI, and EcoRV. The resulting restriction fragments were cloned into the sequencing vector pLitmus38 (Fermentas, St. Leon-Rot, Germany). Three hundred forty-eight restriction library templates were sequenced by the dideoxy chain termination method with a Taq DyeDeoxy terminator cycle sequencing kit (Applied Biosystems, Darmstadt, Germany). Standard shotgun sequencing was done with the ABI 3730 (Applied Biosystems) DNA sequencer, resulting in 696 sequencing reads. Sequence quality control and assembly were performed applying PHRED (16, 17) and PHRAP (14). Genome finishing was done using the CONSED/AUTOFINISH software tool (21, 22). We decided to rely on Phred40 quality in the consensus sequence. For gap closure and polishing of the sequence, 139 and 40 primer walkings on phage 2851 DNA and DNA of restriction library clones as a template were performed, respectively, using the dye terminator chemistry on an ABI 377 sequencer (Applied Biosystems). The final assembly and editing of the DNA sequence data resulted in a single linear molecule with a total length of 57,248 bp.

Functional annotation and analysis. The finished phage 2851 sequence was annotated by using the GenDB annotation tool (34). Gene prediction was performed using the accurate gene-finding method REGANOR (31) implemented in GenDB. Additionally, BLAST analyses (2) of intergenic regions were performed. Prior to the manual annotation of each predicted gene, an automatic annotation was computed based on different tool results. Therefore, similarity searches were performed against different databases including SWISS-PROT and TrEMBL (11), KEGG (28), Pfam (5), TIGRFAM (23), and InterPro (3).

TABLE 2. Primers for detection of phage 2851-related sequences and restriction patterns of PCR products

Primer	Sequence (5'–3')	Product size (bp)	Genetic element ^a	Restriction analysis with:	Fragment sizes (bp) ^b
72	TCACTGCTGACTTGCGAAAGC	704	<i>int</i>	BglII	430, 274
73	CGGTTTGTATGTCTTCTCCAGGG				
74	GACAAAACAATCTCAATCCATCCC	797	<i>ospB</i>	PstI	505, 292
75	CCAAAGTGGCTCTGTAGTCGCTTC				
80	GTGCTTGGGTCTTTTCTCTG	1,212	<i>int</i> site, left	VspI	729, 483
81	TCTGCGGAACATCAATCAAC				
82	CCAGCGTGGGATAAAAAGAGAATC	1,235	<i>int</i> site, right	HindIII	1,020, 215
83	CATCAGTGAAGAACCACAAAACAC				
94	ACGGCTCTTCCAAGCACG	2,253	<i>q</i> , <i>stx</i> _{2c}	CfrI	1,169, 471
95	ATACACCAGTCGCCCTCCG				326, 287

^a *int* site, integration site.

^b Fragments generated after restriction endonuclease digestion of PCR products.

Putative tRNA genes were identified with tRNAscan-SE (32). Putative terminator sequences were predicted using Tranterm (26).

Phage comparison. For phage comparison, the annotated sequences of the following phages were downloaded and imported into the GenDB genome annotation system: enterobacteria phage lambda (J02459), bacteriophage 933W (AF125520), phage BP-4795 (AJ556162), enterobacteria phage VT1-Sakai (AP000400), enterobacteria phage VT2-Sakai (AP000422), Stx1-converting phage (AP005153), Stx2-converting phage (AP005154), Stx2-converting phage II (AP004402), bacteriophage CP-1639 (AJ304858), and enterobacteria phage HK97 (NC_002167). Similarity searches against the protein-encoding regions of the 10 phages were conducted by using BLASTP (2) with an E-value cutoff of 1E–30.

Preparation of total DNA from bacteria. Total DNA of bacteria was prepared from 1-ml overnight cultures (approximately 1×10^9 bacteria) with the ready-to-prepare spin bacteria DNA mini kit (Invitex, Berlin, Germany). DNA preparations were stored at 4°C for use. The amplification of DNA by PCR was performed in 35 cycles, as indicated in Table 2. Annealing temperatures were calculated with Mac Vector software (Oxford Molecular Group, Campbell, CA).

Plasmid construction. Genes of phage 2851 predicted to regulate phage propagation (*antA*, *antB*, *cro*, *ci*, and *chl*) or coding for a putative type III effector protein (*ospB*) were amplified by PCR. The sequences of the forward and reverse primers were modified to comprise sites for restriction endonucleases NdeI and HindIII, respectively (see Table S1 in the supplemental material). The PCR products were restricted and inserted in the corresponding sites of pMS470Δ8 (4), resulting in plasmids of the pJH700 series (Table 1). Plasmids of the pJH700X series generating a frameshift in the coding sequences (CDSs) of the corresponding gene encoded an additional guanine in the sequence of the forward primer (see Table S1 in the supplemental material). The putative antirepressor gene (*antAB*) of phage 1874 and a control construct were inserted in pMS470Δ8 in the same way (Table 1). All recombinant plasmids were introduced into *E. coli* C600. Propagation tests were carried out as follows. Cultures of recombinant plasmid carrying *E. coli* C600 strains were grown to mid-logarithmic phase (optical density at 600 nm of 0.3 to 0.5). Infections with phage lysates from strains CB2851 and TPE1874 were carried out as a standard procedure.

E. coli K-12 strains harboring the recombinant plasmids pJH709 and pJH709X, encoding the putative type III effector OspB, were grown to the mid-logarithmic phase in LB medium and the inducer IPTG (isopropyl-β-D-thiogalactopyranoside; final concentration, 1 mM) was added. Cell lysates were obtained by ultrasonication and analyzed by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis. A band corresponding to the putative OspB protein was cut out from the SDS gel, in gel digested by trypsin, and analyzed by high-performance liquid chromatography-electrospray ionization-tandem mass spectrometry (Agilent 1100 nano-high-performance liquid chromatograph coupled to Finnigan LTQ-FT). Data were converted and searched against the NCBI database.

Determination of the phage 2851 integration site. To obtain the nucleotide sequence of the phage integration site in the bacterial chromosome, total DNA of strains CB2851 and TPE1874 was isolated and digested with several restriction enzymes. Restriction fragments of each digest were ligated with T4 ligase and used as the template in a PCR using primer Stx2c70 out1 (5' TIGTCTGGC ACATCAGGAGC 3') in combination with primer Stx2c71 out2 (5' GACAAC TATTCCATGACAACG 3'); both primers bind within the CDS of the inte-

grase gene of phage 2851. A PCR product derived from amplification with these primers is obtained only if the DNA sequences adjacent to the 3' ends of the primers are joined to form a circular DNA molecule. PCR products obtained with ligation mixtures derived from EcoRI- and HindIII-digested total DNA were purified and analyzed for their nucleotide sequences. Primers for amplification were deduced from the *shcB* region (accession no. ECOHU43) and phage 2851 DNA (Table 2).

Preparation of RNA and cDNA from bacteria. For the preparation of RNA, bacteria were grown to late exponential phase (up to an optical density at 600 nm of 0.8 to 1.0) in LB without or with mitomycin C as described above. Total RNA was isolated from 5-ml bacterial cultures with the RNeasy mini kit (Qiagen, Hilden, Germany). RNA preparations were repeatedly digested with RNase-free DNase I (Roche Applied Systems, Mannheim, Germany) for 60 min at 37°C as described by the manufacturer. The absence of DNA from RNA samples was controlled by real-time PCR amplification of the *icdA*, *ospB*, and *stx*₂ genes. RNA preparations were stored at –20°C for use. The preparation of cDNA from the RNA samples (between 200 ng and 300 ng) was performed with the high-capacity cDNA reverse transcription kit (Applied Biosystems, Darmstadt, Germany) following the instructions supplied from the provider. cDNA reaction mixtures were diluted 10-fold for use in real-time PCR.

Transcriptional analysis. For transcriptional analysis of *ospB* and *stx*₂ genes by quantitative real-time reverse transcription-PCR (qRT-PCR) (7, 50), the qRT-PCR was performed with the Applied Biosystems 7500 real-time PCR system (Applied Biosystems) with cDNA samples from bacteria. Transcription rates of the *ospB* gene and the *stx*₂ gene were compared in parallel to those of the *icdA* housekeeping gene. Primers and TaqMan minor groove binder (MGB) probes were developed with ABI Prism primer software (Applied Biosystems) and produced by Applied Biosystems (Applied Biosystems, Darmstadt, Germany). Primers *icdA*-F and *icdA*-R (50) were used as described together with the 6-carboxyfluorescein-labeled *icdA* MGB probe (5' ACCCTGCAAAACGGCAA 3') (7).

The transcriptional analysis of Shiga toxin genes was performed with primers Stx2-f (5' CAGGCAGATACAGAGAGAAATTTCC 3') and Stx2-b (5' CCGGC GTCATCGTATACACA 3') together with VIC-labeled Stx2-specific TaqMan MGB probe (5' ACTGTCTGAAACTGCTC 3'). The primer system is specific for the gene of the A subunit of most *stx*₂ gene clusters (7). *ospB* gene expression was measured with primers 90 (5' ACACGTCTATGCCGAGTTTGAA 3') and 91 (5' CCAGTGCCATGGTAACCTGAT 3') by use of the VIC-labeled TaqMan MGB probe (5' TCGCGGAATTAAC 3'). Real-time PCR amplifications were performed for 35 cycles in 25-μl reaction volumes by use of TaqMan universal PCR master mix (Applied Biosystems). Relative quantification assays were performed with cDNA in an *icdA* and *ospB* multiplex assay. The cDNA was also analyzed in an *icdA* and *stx*₂ multiplex assay. qRT-PCR assays were analyzed with the 7500 system SDS software version 1.4 (Applied Biosystems).

Nucleotide sequence accession numbers. The complete nucleotide sequence of phage 2851 was submitted to EMBL under accession number FM180578. The following additional sequences were submitted (accession numbers are given in parentheses): the left integration sites of phage 2851 in strain CB2851 (AM982817) and in strain TPE1874 (AM982819) and the right integration sites of phage 2851 in strains CB2851 (AM982818) and TPE1874 (AM982820). The nucleotide sequence of the *stx*_{2c} gene of strain CB8028 was deposited under accession number AM982821, and those of the *stx*_{2c} genes of O177 strains

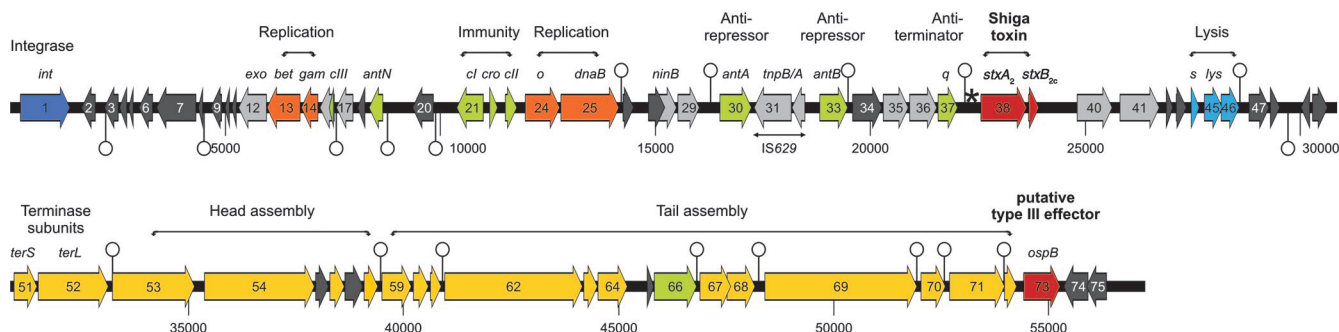


FIG. 1. Physical map of phage 2851 genome. Terminators are indicated by open circles (above terminator on plus strand; below terminator on minus strand); the black star in front of the *stxB*_{2c} gene depicts the position of tRNA genes. Box colors: dark blue, integrase gene; light blue, lysis genes; black, hypothetical genes; gray, conserved genes with or without known function; orange, replication; green, genes involved in lysogeny/lytic development; yellow, morphogenetic genes; dark red, virulence genes.

CB7126 and CB1748 were deposited under accession numbers FM177471 and FM177472, respectively.

RESULTS AND DISCUSSION

Complete nucleotide sequencing and annotation of phage 2851. The complete nucleotide sequence of phage 2851 was established by a terminal sequencing of inserts from four different restriction fragment libraries and the subsequent use of primer-walking strategies. Compilation of the complete phage 2851 nucleotide sequence revealed that the phage DNA is circularly permuted. The data were assembled into a sequence of 57,248 bp, including the left and right core sequences of the attachment sites. The mean G+C content of the genome is 51.05%. The annotation of the phage sequence revealed 75 predicted CDSs. Its genetic structure and modular organization are depicted in Fig. 1. The genome of phage 2851 contains 19 terminators (see Table S2 in the supplemental material) as indicated in Fig. 1. Phage 2851 is a lambdoid phage with a typical mosaic in genome structure, as lambdoid phages share a common gene pool from which DNA can be frequently exchanged (24, 27).

The phage 2851 sequence was arranged starting with the putative phage integrase gene *int*. The *int*₂₈₅₁ gene shows the highest similarity to integrase genes present in prophage sequences of EHEC O157 (strains EC4045 and EC4206; accession numbers ZP_02744119.1 and ZP_02752421.1, respectively) and of a *Shigella flexneri* strain (YP_689502.1), but so far it has not been found in viable Shiga toxin-converting bacteriophages. Downstream of the *int* gene are a number of phage genes which code for regulatory proteins, like repressors CI, CII, CIII, and Cro, antirepressors AntA and AntB, and anti-terminator N, replication proteins O and P, and proteins involved in nucleotide metabolism. The arrangement of regulatory genes *cI*, *cro*, *cII*, and *cIII* is very similar to what is seen for phage lambda (12). Analysis of the intergenic region between *cI* and *cro* shows the presence of three operator sequences with features conserved between it and the lambda system (data not shown). Phage 2851 carries one complete insertion sequence element, IS629, which is flanked by antirepressor genes *antA* and *antB* (see below). The Shiga toxin-encoding genes are part of a lysis cassette downstream of the gene (ORF37) for the Q antiterminator, which begins at nucleotide position 21,564 and

contains genes coding for a holin and lysozyme. The right half of the phage genome begins with the genes encoding a small terminase (ORF51 at position 31,500) and a large terminase (ORF52). Most of the right half of the genome consists of morphogenetic genes coding for head and tail structures. ORF73 (position 54,419) may encode an effector for a type III secretion system, as its gene product shows similarity to OspB of *Shigella*. In *Shigella* species, the gene for OspB is located on a virulence plasmid and secreted by a plasmid-encoded type III secretion system; however, a function for the OspB protein has yet to be determined (41, 55). The last putative gene of the phage genome (ORF75) encodes a hypothetical protein which shows similarity to hypothetical protein ECs2227 of *E. coli* EHEC strain O157 Sakai.

Comparison of phage 2851 with different phages. Overall, the gene content of the Stx1-converting phage BP-4795 (AJ556162) is most similar to that of phage 2851, with 48 related genes, and is followed in terms of similarity by an Stx2-converting phage (AP005154), enterobacteria phages VT2-Sakai (AP000422) and VT1-Sakai (AP000400), and another Stx1-converting bacteriophage (AP005153), with 40, 38, 37, and 33 related genes, respectively (Fig. 2). Enterobacteria phages HK97 (NC_002167) and lambda (J02459) share lower numbers of related genes with phage 2851 (14 and 14 similar genes, respectively). ORF46 of phage 2851 encodes a putative endopeptidase, which is predicted to be involved in the lysis of bacteria and is the only protein with related counterparts in all 10 compared phages. Surprisingly, the upstream lysozyme-like protein (gene product of ORF46) does not show similarity to any of the compared phages; thus, it seems to be specific for phage 2851 (Fig. 2). Further genes without any counterparts in one of the compared phages are as follows: ORF01, encoding putative prophage integrase; ORF47, encoding the KilA protein, related to phage lytic development; and ORF02, ORF48, ORF74, and ORF75, encoding hypothetical proteins. Interestingly, ORF24 of phage 2851, coding for bacteriophage replication protein O, has only one related counterpart with enterobacteria phage HK97 (NC_002167), which shares the lowest number of similar genes with phage 2851. ORF73, encoding an OspB-like protein which may act as a virulence factor, has only one counterpart that is present in the Stx1 phage BP-4795 (AJ556162) (13).

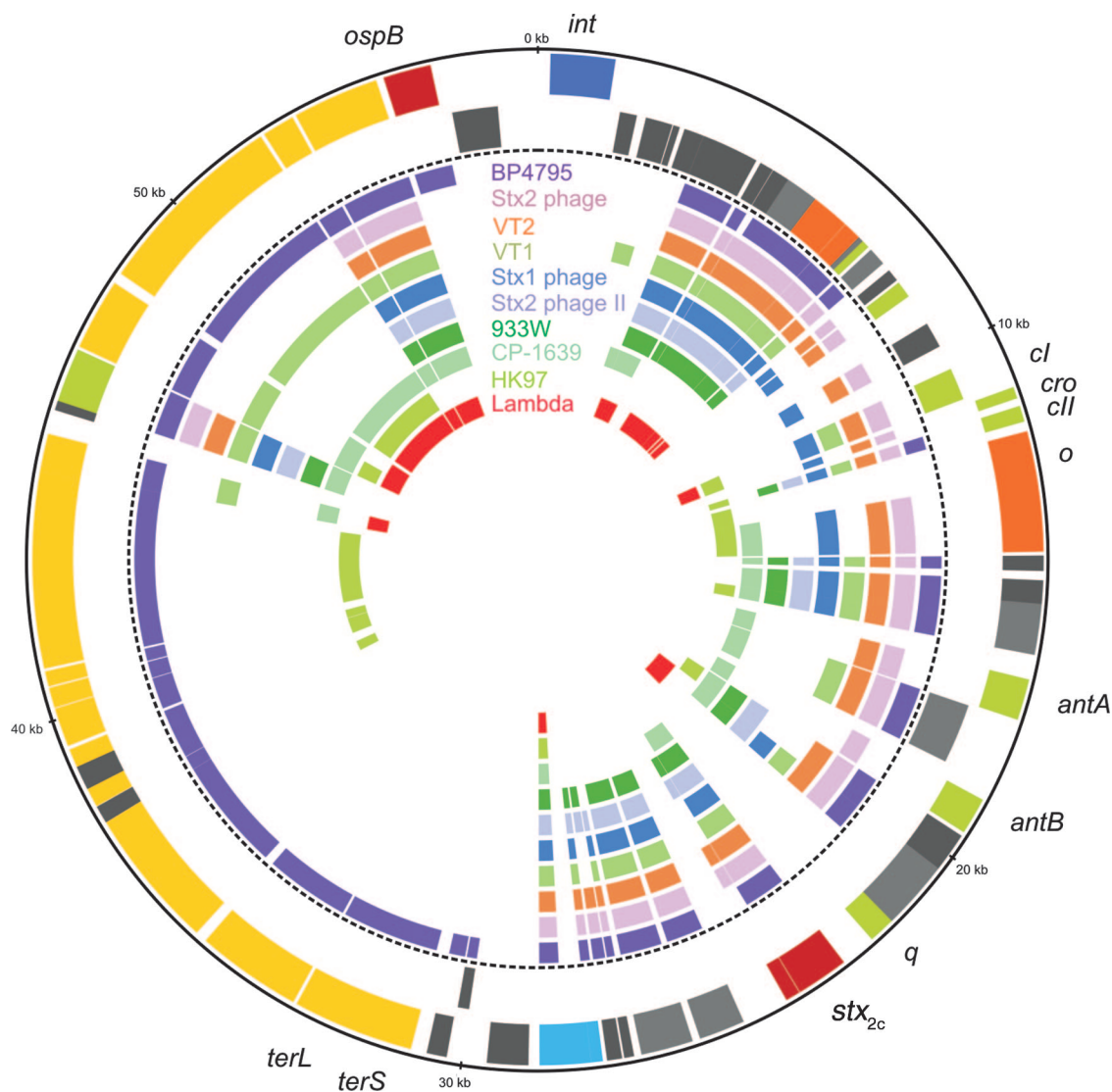


FIG. 2. Circular synthetic plot of phage 2851: circular representation of the CDS of phage 2851 and related proteins in the other phages. From the outermost to the innermost concentric circle, the circles indicate the following: 1, genomic position in kbp; 2 and 3, CDSs of phage 2851 on the forward (circle 2) and the reverse (circle 3) strands colored according to the modules in Fig. 1; 4 to 13, CDSs encoding similar proteins in phage BP4795 (lilac; AJ556162), Stx2-converting phage (pink; AP005154), enterobacteria phage VT2-Sakai (brown; AP000422), enterobacteria phage VT1-Sakai (light green; AP000400), Stx1-converting phage (blue; AP005153), Stx2-converting phage II (light blue; AP004402), bacteriophage 933W (green; AF125520), bacteriophage CP-1639 (turquoise; AJ304858), HK97 (olive; NC_002167), and lambda (red; J02459). The related proteins were identified by BlastP matches of amino acid sequences (E-value of $<1E-30$).

IS629-mediated deletion of phage 2851 DNA between two antirepressor genes. The determination of the phage nucleotide sequence revealed two types of phage genomes in released phages. In the genomic region encompassing the two putative antirepressor genes *antA* (ORF30) and *antB* (ORF33), differences were observed. When preparations of phages from strain CB2851 were analyzed by use of primers 46B and 49 (48), two PCR products with different sizes were found. One PCR product possessed a size of 2,334 bp, corresponding to the region of the putative antirepressor genes *antA* and *antB* (ORF30, ORF33) and the IS629 element of the phage sequence as depicted in Fig. 1. The second PCR product, however, showed a truncated sequence (714 bp) of this region.

The sequence analysis of *antA* and *antB* revealed that in the

coding regions of the two genes, identical direct repeats with lengths of 107 bp are present. Upon phage release, a recombination between these repeats takes place frequently, leading to a deletion of the 3' coding region of *antA*, the IS629, and the 5' region of the *antB* region. A new fusion gene, *antAB*, is formed, containing only one copy of the repeat sequence in the middle of the CDS (Fig. 3). The start of the repeat sequence is in the +1 frame of the CDSs of the *antA*, *antB*, and *antAB* genes; thus, in all putative gene products, the same amino acid sequence is encoded by the repeat.

The viability of both types of phages was confirmed, as we obtained *E. coli* C600 strains lysogenic for the original prophage (designated TPE1875) and the truncated phage (strain TPE1874). By growing the lysogenic C600 strains in LB me-

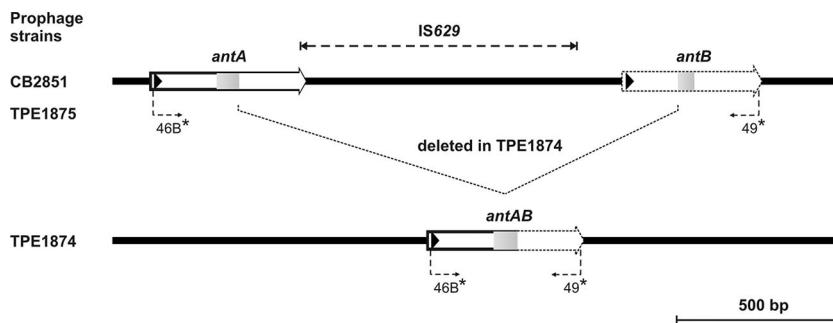


FIG. 3. *IS629*-mediated deletion of phage 2851 sequence. (Top) Structures of *antA*, *IS629*, and *antB* regions in prophages of strains CB2851 and TPE1875. (Bottom) Structure of fusion gene *antAB* after loss of 2,242 bp. The positions of primers 46B and 49 are indicated.

dium and adding mitomycin C to the medium, we observed that the phage titer after overnight growth was generally approximately 10-fold higher in supernatants of TPE1874 cultures than for TPE1875 cultures. These results indicate that the fusion gene *antAB* might act as an efficient antirepressor promoting the release of the integrated prophages. Strains carrying phages with the *IS629* element were shown to spontaneously release numbers of *stx* phages that were 1 log unit lower than those seen for phages without *IS629*. This could indicate that the *IS629* insertion between *antA* and *antB* contributes to the stability of the lysogenic state and the survival of its *E. coli* host.

Antirepressor activity encoded by fusion gene *antAB*. To confirm the functionality of the putative antirepressor *antAB* gene, we investigated a number of genes from phage 2851 which are supposed to influence the propagation of the phage. The genes chosen as controls were the repressor genes *cI* and *cII* and the genes *cro*, *antA*, and *antB* of prophage CB2851, as well as the *antAB* fusion gene of strain TPE1874, which encode regulatory proteins involved in lytic development. The experiments were carried out by cloning the genes of interest into vector pMS470Δ8 (4). The plasmid constructs were introduced into strain C600. All strains were grown to the mid-logarithmic phase and infected with phages isolated from TPE1874 and CB2851. In cases where high intracellular levels of repressor molecules (CI or CII) were present in the recombinant strains,

the formation of phages was reduced, as the repressor proteins interact with incoming phages, a phenomenon designated hyperinfection immunity. In cases where there were elevated levels of intracellular regulatory proteins, like Cro and antirepressors AntA and AntB, the opposite effect is observed after phage infection, as a high number of phages enter the lytic pathway compared to what is seen for the control. As controls to measure phage development, we made a control construct for each investigated gene containing the gene mutated by introducing a frameshift in the CDS. Phage development was measured by comparing phage titers obtained after the infection of the recombinant C600 strains with an intact gene to what was seen for the corresponding control.

Figure 4 shows the results of these experiments. The expression of cloned repressor genes *cI* and *cII* reduced the phage titers drastically compared to what was seen for the controls. For the cloned *cI* gene, a more-than-100-fold reduction of phage development was observed, and for the cloned *cII* gene, lytic development was completely inhibited. By introducing the cloned genes *cro*, *antA*, and *antB* and thereby initiating the lytic pathway, we found an increase of the phage titers compared to what was found in the control experiments. In the case in which the fusion gene *antAB* was created by spontaneous recombination, the increase of phage release was more than 10-fold, indicating that the encoded antirepressor was functional and enhanced phage development. This effect on phage develop-

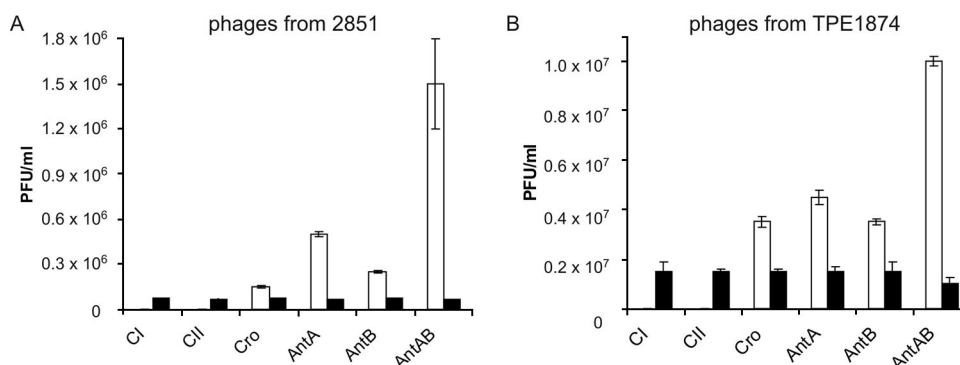


FIG. 4. PFU per ml on C600 strains harboring recombinant plasmids with genes influencing phage propagation. White bars indicate strains with functional regulatory genes; black bars indicate control strains harboring genes with frameshift mutations. In C600 strains with functional CI and CII repressors, phage titers were reduced to below 1% (white bars missing). (A) Phage lysate from strain CB2851 had a titer of 7.5×10^4 phages (PFU/ml) on C600; the standard deviations for control strains were too low to be visible. (B) Phage lysate from strain TPE1874 had a titer of 1.5×10^6 (PFU/ml) on C600. Means and standard deviations are derived from three different experiments.

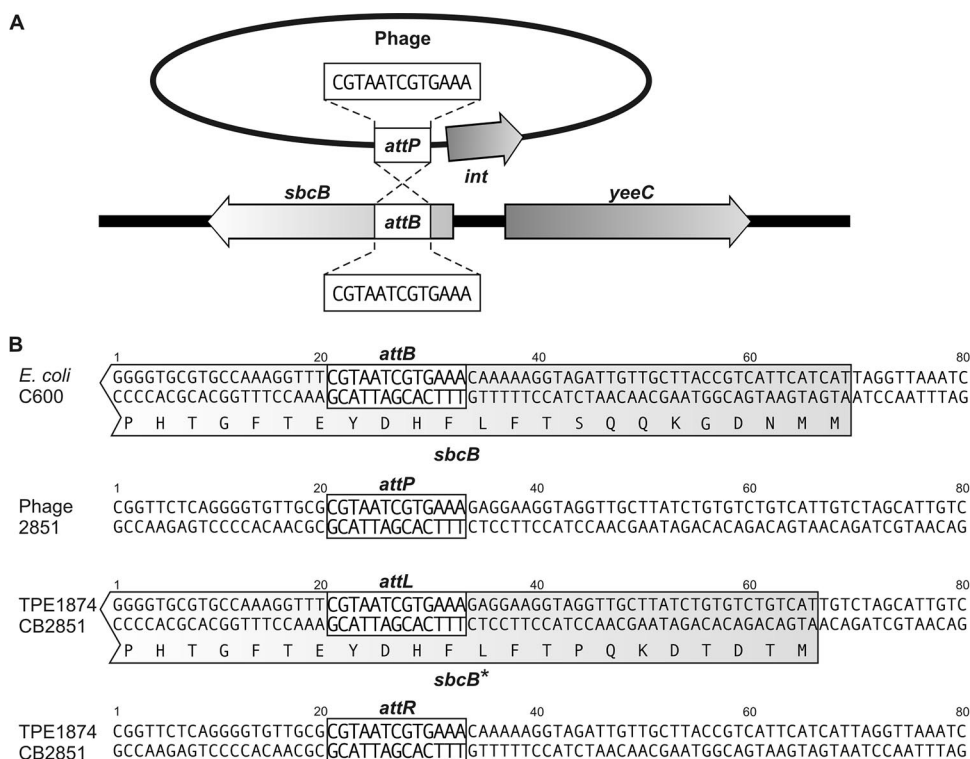


FIG. 5. Integration site of phage 2851. (A) Integration of phage 2851 into the chromosome of *E. coli* C600, the *attP* attachment site of the phage, and the *attB* bacterial attachment site. (B) Sequences flanking the attachment sites in, from top to bottom, C600, phage 2851, and the left (*attL*) and right (*attR*) core sequences in lysogenic strains TPE1874 and CB2851. *sbcB** indicates a putative modified *sbcB* gene.

ment was greater for the AntAB antirepressor than for Cro, AntA, or AntB.

Occurrence of *antAB* fusion genes in other STEC O157 strains. To find out if the fusion gene *antAB* is present in more STEC O157 strains, we performed PCR with primers located in the 5' region of the CDS of *antA* and the 3' region of the coding region of *antB* (primers 46-B and 49) (Fig. 3) (48). We found the fusion gene *antAB* in two Stx2c strains from our previous investigation (CB7917, CB7890). In those Stx2c strains, which had been positive for a PCR amplifying an intergenic region between *antA* and *antB* (48), the two genes were arranged head to tail, which was shown by partially sequencing some of the PCR products (data not shown). A search of the GenBank database for *antA*- and *antB*-related sequences revealed a number of related sequences in coliphages and *E. coli* strains. An *antAB* gene identical to the gene present in prophage TPE1874 was found in the VT2-Sakai phage (accession AP000363) and in another very similar phage (Stx2-converting phage AP005154). The head-to-tail arrangement of two putative antirepressor genes, *antA* and *antB*, is present in the cryptic Stx2c prophage Nil2 (AJ413274); however, no IS629 element is inserted into the intergenic region. It is likely that antirepressor genes are located in phage modules that are subject to frequent events of recombinations and deletions, leading to a great sequence variety in this phage module. Thus, PCRs targeting this genomic region of phages are not suited for diagnostic purposes.

***E. coli sbcB* locus as the integration site of phage 2851.** The integration site of phage 2851 into the chromosome of the host

bacteria was determined by amplifying DNA adjacent to the phage integrase gene (ORF1, *int*) by use of primers binding within the CDS of the integrase gene (see Materials and Methods). The integration of lambdoid phages takes place by site-specific recombination between an attachment site in the bacterial chromosome (*attB*) and one in the phage genome (*attP*). The *attP* site is usually immediately upstream or downstream of the phage integration gene. The analysis revealed that the integration of phage 2851 occurs in the 5' end of the coding region of the *E. coli sbcB* gene, which encodes an exonuclease (Fig. 5). The core region of the attachment sites is a 13-bp-long sequence, CGTAATCGTGAAA. The *sbcB* locus was already identified as an integration site for Stx phages; however, the integrated phages have not been further characterized (39, 45). In the two C600 strains lysogenic for phage 2851, designated TPE1874 and TPE1875, the *sbcB* locus was also confirmed as integration site, while the locus was unoccupied in the parental C600 strain. The nucleotide sequences of the left and right integration sites in strain CB2851 and TPE1874 have been deposited in the EMBL database (see Materials and Methods).

The analysis of the left integration site revealed that the N terminus of the exonuclease might be modified through the integration of phage 2851 in TPE1874 (Fig. 5) and TPE1875 (data not shown). The integration of the phage DNA into the CDS of *sbcB* creates a slightly altered 5' coding region of the *sbcB* gene upstream of the core attachment site (*sbcB* of strains C600 and TPE1874) (Fig. 5B). The same modification of the *sbcB* gene has probably occurred in strain CB2851; however, no phage 2851-free derivative of the strain is available.

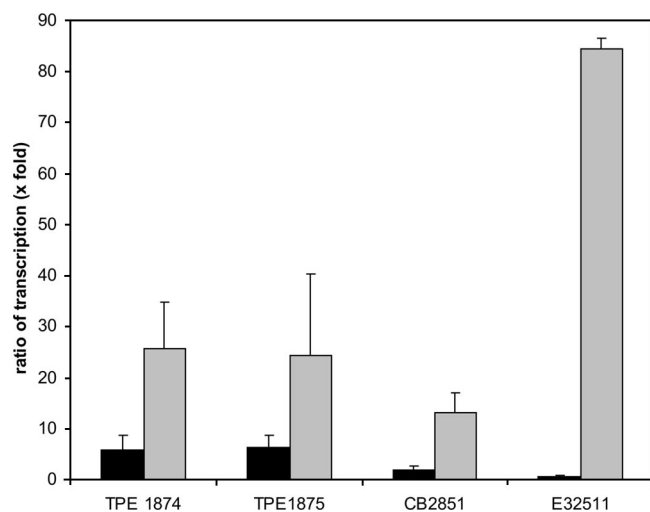


FIG. 6. Variation of *ospB* and *stxA*₂ relative expression in relation to the expression of the housekeeping gene *icdA* (*x*-fold) after induction with mitomycin C. Black bars indicate the ratios of transcription of *ospB* to *icdA*; gray bars indicate the ratios of transcription of *stx2* to *icdA*. Strain E32511 contains two *stxA*₂ genes. Means and standard deviations are from two separate sets of experiments, with measurements performed in triplicate.

Analysis of *ospB* expression. The sequence annotation revealed that phage 2851 may encode an OspB-like type III effector (ORF73) which has not been detected in other *stx* phages so far. The putative OspB protein-encoding region was cloned in vector pMS470Δ8, and a control construct with a frameshift mutation was constructed (see Materials and Methods). SDS-polyacrylamide gel electrophoresis analysis of the *E. coli* K-12 strains harboring the recombinant plasmids pJH709 and pJH709X were grown to mid-logarithmic phase with or without IPTG. An additional protein with the expected size of 32 kDa was visible only in cell extracts from a strain harboring pJH709 after induction with IPTG. Mass spectrometry of tryptic peptides obtained from this protein demonstrated that the cloned open reading frame indeed encodes the predicted protein. The total coverage of the protein sequence by peptides was 45% (data not shown).

To find out if the *ospB* gene is expressed together with the phage-encoded Shiga toxin genes, a transcriptional analysis of Stx2c strains TPE1874, TPE1875, CB2851, and E32511 was performed. Reference strain E32511 possesses two Shiga toxin-encoding genes; besides *stx*_{2c}, the prototypical *stx*₂ gene is also present (44). As a negative control, K-12 strain C600 was also analyzed, and it did not show *ospB* and *stxA*₂ transcripts. The strains were grown with or without mitomycin C, as this antibiotic stimulates phage release and induces Shiga toxin production. Bacteria were harvested at the late logarithmic growth phase, RNA was isolated, and the relative expressions of the Shiga toxin A gene and the *ospB* gene in relation to that of the housekeeping gene *icdA* (isocitrate dehydrogenase) were determined (Fig. 6) (53). We found that *ospB* transcription was increased between two- and sixfold in the phage-releasing strains CB2851, TPE1874, and TPE1875 compared to what was seen for the housekeeping gene upon phage induction. When *stxA*₂ transcription was measured, the effect of mitomy-

cin C induction was clearly more pronounced. The increase of *stxA*₂ gene transcription was approximately 25-fold in TPE1875 and TPE1874 and half of that in CB2851. In strain E32511, *ospB* expression was nearly unchanged by mitomycin C; however, *stxA*₂ transcription was dramatically increased (Fig. 6). This might be explained by the fact that strain E32511 harbors two *stxA*₂ genes, one gene belonging to *stx*_{2c} and one belonging to the prototypical *stx*₂ gene group (44).

The observed differences in transcription patterns of the *ospB* and *stxA*₂ genes indicate that neither gene is coregulated. Phage induction by mitomycin C is known to result in higher transcriptional activity of late phage genes (35, 49); however, the differential increase between the *ospB* expression and *stxA*₂ expression indicates that in the case of *stxA*₂, additional expression mechanisms exist. Analysis of the upstream regions of the *ospB* gene as well as the *stxA*₂ gene by use of bioinformatic tools (http://www.fruitfly.org/seq_tools/promoter.html) indicates putative promoters for the two genes. In the case of Shiga toxin expression, the increase of transcription is frequently found to be strongly enhanced, even if no viable phages are produced by the *E. coli* host strain (7). Remarkably, in the case of strain E32511, a very strong expression of the *stxA*₂ gene was observed upon induction, whereas *ospB* expression remained more or less unchanged. The reference strain contains two *stxA*₂ genes that cannot be discerned by the transcriptional analysis. We analyzed released phages from strain E32511 by PCR targeting the B subunit of the Shiga toxin gene (primers GK3 and GK4) (29) and *ospB* (primers 74 and 75) (Table 2). The sequence analysis revealed that these phages carry the *stx*₂ gene (and not *stx*_{2c}) and do not possess the *ospB* gene.

Spread of phage 2851 in *E. coli* O157 strains. To determine if phage 2851 may have contributed to the dissemination of the *stx*_{2c} gene in STEC O157 strains, we performed diagnostic PCRs targeting phage 2851-related sequences. Total DNAs of 82 *E. coli* O157 strains, among which were 39 strains with *stx*_{2c} genes, were used as templates for this PCR approach (48). A previously reported association between the *q* gene of phage 2851 and the *stx*_{2c} gene was confirmed, as 38 of 39 Stx2c strains were positive for the *q*₂₈₅₁ gene. A similar close association was found between the presence of the *o*₂₈₅₁ gene, which encodes a replication protein, and *stx*_{2c} (36/39 strains). In order to investigate more deeply the association of the *stx*_{2c} gene with more phage 2851-like phages, we developed PCRs targeting other phage-specific genes such as *ospB*, *int*, and the left and right integration sites of phage 2851, as determined from strains CB2851 and TPE1874 (see primers in Table 2). Integration and integration sites of phage 2851 were negative in all 43 O157 strains that did not harbor an *stx*_{2c} gene. On the other hand, these genes, as well as *ospB*₂₈₅₁, were present in 38 of the 39 investigated O157 Stx2c strains. The only exception was O157 Stx2c strain CB8028, which was negative for all tested phage 2851 sequences. The entire nucleotide sequence of the *stx*_{2c} gene of CB8028 (AM982821) was found more related to that of a previously published O177 strain (EU086525) (20) compared to phage 2851-associated *stx*_{2c} sequences. Remarkably, strain CB8028 was the only urease-positive isolate among all O157 strains, and it represents a different subclone which is not commonly found among O157 isolates.

In our previous study, we had found a *q*₂₈₅₁ gene in four O157 strains that tested negative for *stx*_{2c} but positive for *stx*₂

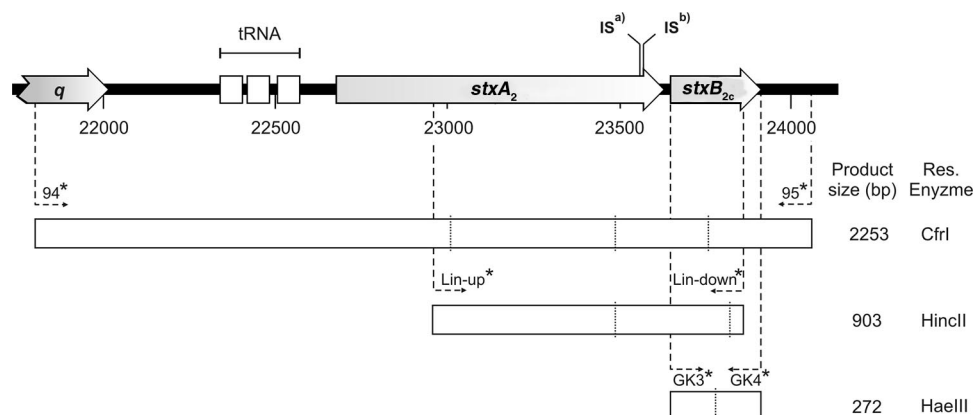


FIG. 7. Partial genomic sequence of phage 2851. (Top) Shiga toxin-encoding region. IS^a) indicates IS629 insertion sites in strains CB7209, CB7323, and CB7735, while IS^b) indicates the insertion site in strain CB7709. Open boxes below depict PCR fragments amplified to analyze *stx*_{2c} genes and the sizes of the products. Primers are indicated by arrows. Vertical lines in open boxes indicate restriction sites for enzymes (shown on the right sight of each box).

(strains CB7209, CB7323, and CB7735) and *stx*₁ (strain CB7709) by the PCR-RFLP method (48). In order to explore the origin of the *q*₂₈₅₁ gene, we analyzed these strains more deeply by PCR using the primers GK3 and GK4, which amplify only the B subunits of various *stx*₂ genes (Fig. 7) (29). Sequencing of the GK3/GK4 products revealed sequence ambiguities in the three Stx2-producing strains CB7209, CB7323, and CB7735 that indicated the presence of two *stx*₂ genes. In the fourth strain, CB7709, which had been classified as an Stx1 producer, a PCR product was generated which contained the CDS of the *stxB*_{2c} subunit.

In order to clarify the *stx* genes present in the four strains, we selected primers binding in the *q*₂₈₅₁ gene and downstream of the *stxB*_{2c} gene (primers 94 and 95) (Table 2). These primers enclose a 2,253-bp region of phage 2851 encompassing the 3' region of the *q* gene, three tRNA genes, and *stxA*₂ and *stxB*_{2c} (Fig. 7). PCR with primers 94 and 95 produced amplicons in strains CB7209, CB7323, CB7735, and CB7709 with sizes of 3.1 kbp. We partially sequenced these products and found that in all four strains a copy of insertion element IS629 inserted in the 3' region of the CDS of *stxA*₂ (Fig. 7). As IS629 prevented the formation of a PCR product with primers used for typing before, this finding could explain why these four strains were not identified as Stx2c producers in our previous study.

As a possible tool for the PCR identification of O157 Stx2c strains, we used primers 94 and 95 PCR (Table 2) on all 82 O157 strains. These primers were suitable for diagnostic purposes, as they specifically detected all O157 strains carrying *stx*_{2c} genes in our study. The digestion of the 94 and 95 PCR products with CfrI was used for verification of the specific product, since one CfrI restriction site is present in the nucleotide sequence of the *stxB*_{2c} gene but not in the sequence of the *stxB*₂ gene.

By using this diagnostic PCR targeting phage 2851 sequences, we found that 38 out of 39 strains harboring *stx*_{2c} genes were positive for most of the phage 2851-associated sequences (Table 3). Interestingly, all strains negative for the phage 2851 sequences also possessed an unoccupied *sbcB* locus. Our results suggest that phage 2851 is the prototype phage for the

dissemination of the Stx2c variant in pathogenic STEC O157 strains that emerged in different places and time periods.

Presence of phage 2851-related sequences in Stx2c strains not belonging to the O157 serotype. The O157:NM strain E32511 is the prototype strain for the *stx*_{2c} gene (42, 44). The *stx*_{2c} gene can be discerned from *stx*₂ by nucleotide sequence alterations in the CDS of the B subunit of the toxin, which can be identified by PCR-RFLP typing. *stx*_{2c} genes were rarely identified in non-O157 STEC (48). We investigated if phage 2851 may have contributed to the dissemination of the *stx*_{2c} genes in STEC strains belonging to non-O157 serotypes. Two STEC O177:H strains from our collection (strain CB7126 of bovine and CB7148 of human origin) possess the *q*₂₈₅₁ gene as well as the *stxB*_{2c} gene. We performed PCR using primers 94 and 95 with DNA from these strains and obtained PCR products from both strains which were approximately 1.3 kbp in size. Partial sequencing of these products revealed that the *q*₂₈₅₁ gene is not linked to the *stx*_{2c} gene but is located upstream of an open reading frame encoding a hypothetical phage-related protein not found for phage 2851. In one of the two strains, the *int*₂₈₅₁-related gene and the left integration site (primers 80 and 81) could be amplified. PCR with primers 82 and 83 for the right integration site failed in both strains, which, in contrast, were positive for the *ospB* gene (PCR with primers 74 and 75) (Table 2). In order to analyze the possible origin of the *stx*_{2c} gene in STEC O177 strains, we sequenced a 1,536-bp stretch starting 276 nucleotides upstream of the *stxA*₂ and ending 19 nucleotides downstream of the *stxB*_{2c} gene in strains CB7126 (GenBank accession number FM177471) and CB7148 (GenBank accession number FM177472). The genes for the Stx2A subunits of the O177:NM strains share 99% identity at the nucleotide level with the phage 2851 *stxA*₂ and 100% identity with its *stxB*_{2c} gene (AJ605767) but were slightly different from the O177 *stx*_{2c} described earlier (EU086525) (20). The 276-bp region located upstream of the *stxA*₂ gene in the O177 strains from our collection was most similar (98% identity, 100% query coverage) to that located upstream of the *E. coli* O91:H21 *stxA*_{2d2} subunit (accession number AF479829). It showed less similarity

TABLE 3. Association of phage 2851 genes in wild-type *E. coli* O157: H7 strains

<i>stx</i> _{2c} gene ^a	<i>stx</i> gene	Total no. of strains	Origin (no. of strains from indicated origin)	No. of strains positive in PCR for ^b :					
				<i>q</i> ₂₈₅₁	<i>o</i> ₂₈₅₁	<i>int</i> ₂₈₅₁	<i>attL</i> ₂₈₅₁	<i>attR</i> ₂₈₅₁	<i>ospB</i> ₂₈₅₁
Negative	Negative	8	Germany (6), The Netherlands (2)	0	0	0	0	0	0
	<i>stx</i> ₂	21	Germany (20), United States (1)	0	0	0	0	0	1
	<i>stx</i> ₁ + <i>stx</i> ₂	14	United States (9), Canada (3), United Kingdom (1), Germany (1)	0	1	0	0	0	2
Total negative		43		0	1	0	0	0	3
Positive	<i>stx</i> _{2c}	4	Germany (4) ^c	3	3	3	3	3	3
	<i>stx</i> ₁ + <i>stx</i> _{2c}	27	Germany (26), United States (1)	27	25	27	27	27	27
	<i>stx</i> ₂ + <i>stx</i> _{2c}	8	United States (1), Denmark (2), United Kingdom (1), Germany (4)	8	7	8	8	8	8
Total positive		39		38	36	38	38	38	38

^a *stx*_{2c} gene determined by PCR using primers 94 and 95.

^b Diagnostic PCRs were for the following genes of phage 2851 (Table 2 shows primers and RFLP): *q*₂₈₅₁, *q* antiterminator gene of phage 2851; *o*₂₈₅₁, *o* replication gene; *int*₂₈₅₁, integrase gene; *attL*₂₈₅₁ and *attR*₂₈₅₁, left and right integration loci (attachment sites), respectively, of phage 2851; and *ospB*₂₈₅₁, gene for putative type III effector.

^c Strain CB8028 (urease positive) was negative for all phage 2851 sequences.

(96% identity, 94% query coverage) with the corresponding region in phage 2851. Taken together, these results indicate that the origin of the *stx*_{2c} genes in STEC O177 strains is not linked to a 2851-like prophage.

We also conducted a database search for published *stx*_{2c} genes in non-O157 *E. coli* strains. Three additional *stx*_{2c} sequences are published that stem from one O91:H21 strain (AF479828), one O174:H21 strain (AF500198), and one O:untypeable strain (AF500193). All three strains possess an activatable Shiga toxin 2A subunit, which can be discerned by RFLP of PCR products of the Shiga toxin-encoding region by use of restriction enzyme PstI from the gene for the Stx₂A subunit of phage 2851. In contrast, no Stx_{2c} serotype O157 strains from our investigation possess an activatable A subunit (including the untypical urease-positive strain CB8028). Only the sequence of an O91:H21 strain (AF479828) contains sequence information of the *stx*_{2c} gene-flanking DNA region. In that case, the *q*₂₈₅₁ gene is not present.

Conclusion. In the present study, we have characterized phage 2851 for its complete nucleotide sequence and present data suggesting that it is a prototype phage for the global spread of the Stx_{2c} determinant in STEC O157:H7 strains. An interesting feature of the phage is the presence of an IS629 element in its genome, which is frequently lost upon the excision of the prophage, leading to a deletion of phage sequences. It is conceivable that deletions or rearrangement frequently occur in O157 strains, with IS629 being involved. IS629 seems to be an important genetic element in the evolution of Stx_{2c} strains, since it contributes to phage stability and inactivates the expression of the toxin gene. The fusion gene *antAB*, which is formed by homologous recombination, was shown to encode a functional antirepressor and was found in more O157 strains and Stx phages, thus indicating that this recombination event occurs in nature.

The integration site of phage 2851 was shown to be the *sbcB* locus in *E. coli* O157, as well as in *E. coli* K-12 strains which

were successfully converted to Shiga toxin production after the integration of phage 2851 (48). So far, five different insertion sites have been described for lambdoid phages in STEC strains (45). The *sbcB* locus was first described as an integration locus for Stx1 and Stx2 phages in a whole-genome PCR scanning approach for O157 strains (39). In our strain collection, only strains harboring phage 2851 sequences had an occupied *sbcB* locus, suggesting that the integration of this phage normally occurs into this locus, as was also demonstrated for the K-12 strains. It is conceivable that phage 2851 could integrate at another locus of the host strain chromosome if this site were occupied by a different prophage (45).

The putative type III effector *OspB* has not been found in other Stx phages so far, and its possible role in virulence needs to be investigated. Interestingly, another type III effector called NleA was found encoded by the *stx*₁ phage BP-4795, indicating that Stx phages play a role in the dissemination of type III effectors (13).

ospB was shown to be transcriptionally upregulated upon the induction of the phage by mitomycin C. However, the transcriptional increase was not as dramatic as for the *stx*_{2c} genes of strains TPE1874, TPE1875, and CB2851. It is conceivable that the increased transcription of *ospB* is an effect of increased copies of phage genomes which result from phage replication. In the case of increased *stx*_{2c} expression, it can be assumed that additionally the Q protein allows readthrough through a transcription promoter, thus enabling RNA polymerase to efficiently transcribe the *stx*_{2c} gene (37, 43). This would explain why in strain E32511 the *ospB* transcription rate remains unchanged, as it is located on a silent prophage that is not induced by mitomycin C. However, the second *stx*₂ gene of E32511, located on a viable phage, is efficiently induced, and the *stx*₂ expression is high.

A comparison of phage *stx*_{2c} and neighboring sequences to sequences in non-O157 STEC strains indicates that phage 2851 is restricted to the O157 serotype and that Stx_{2c} de-

terminants of non-O157 strains are encoded by other lambdoid phages.

ACKNOWLEDGMENTS

We thank students Sandra Lindenberg and Marc Bohmer for help with PCR analyses; Dorothea Knabner, Sabine Haby, and Katja Steege for technical help; and Ulf Bergmann for mass spectrometric analysis of the OspB protein.

REFERENCES

- Allison, L. J., P. E. Carter, and F. M. Thomson-Carter. 2000. Characterization of a recurrent clonal type of *Escherichia coli* O157:H7 causing major outbreaks of infection in Scotland. *J. Clin. Microbiol.* **38**:1632–1635.
- Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**:3389–3402.
- Apweiler, R., T. K. Attwood, A. Bairoch, A. Bateman, E. Birney, M. Biswas, P. Bucher, L. Cerutti, F. Corpet, M. D. Croning, R. Durbin, L. Falquet, W. Fleischmann, J. Gouzy, H. Hermjakob, N. Hulo, I. Jonassen, D. Kahn, A. Kanapin, Y. Karavidopoulou, R. Lopez, B. Marx, N. J. Mulder, T. M. Oinn, M. Pagni, F. Servant, C. J. Sigrist, and E. M. Zdobnov. 2001. The InterPro database, an integrated documentation resource for protein families, domains and functional sites. *Nucleic Acids Res.* **29**:37–40.
- Balzer, D., G. Ziegelin, W. Pansegrau, V. Kruff, and E. Lanka. 1992. KorB protein of promiscuous plasmid RP4 recognizes inverted sequence repetitions in regions essential for conjugative plasmid transfer. *Nucleic Acids Res.* **20**:1851–1858.
- Bateman, A., E. Birney, L. Cerruti, R. Durbin, L. Etmiller, S. R. Eddy, S. Griffiths-Jones, K. L. Howe, M. Marshall, and E. L. Sonnhammer. 2002. The Pfam protein families database. *Nucleic Acids Res.* **30**:276–280.
- Beutin, L., S. Kaulfuss, T. Cheasty, B. Brandenburg, S. Zimmermann, K. Gleier, G. A. Willshaw, and H. R. Smith. 2002. Characteristics and association with disease of two major subclones of Shiga toxin (verocytotoxin)-producing strains of *Escherichia coli* (STEC) O157 that are present among isolates from patients in Germany. *Diagn. Microbiol. Infect. Dis.* **44**:337–346.
- Beutin, L., U. Kruger, G. Krause, A. Miko, A. Martin, and E. Strauch. 2008. Evaluation of major types of Shiga toxin 2e-producing *Escherichia coli* present in food, pigs, and the environment as potential pathogens for humans. *Appl. Environ. Microbiol.* **74**:4806–4816.
- Beutin, L., S. Zimmermann, and K. Gleier. 1996. Rapid detection and isolation of Shiga-like toxin (verocytotoxin)-producing *Escherichia coli* by direct testing of individual enterohemolytic colonies from washed sheep blood agar plates in the VTEC-RPLA assay. *J. Clin. Microbiol.* **34**:2812–2814.
- Bielaszewska, M., R. Kock, A. W. Friedrich, C. von Eiff, L. B. Zimmerhackl, H. Karch, and A. Mellmann. 2007. Shiga toxin-mediated hemolytic uremic syndrome: time to change the diagnostic paradigm? *PLoS ONE* **2**:e1024.
- Bielaszewska, M., H. Schmidt, M. A. Karmali, R. Khakhria, J. Janda, K. Blahova, and H. Karch. 1998. Isolation and characterization of sorbitol-fermenting Shiga toxin (verocytotoxin)-producing *Escherichia coli* O157:H⁻ strains in the Czech Republic. *J. Clin. Microbiol.* **36**:2135–2137.
- Boeckmann, B., A. Bairoch, R. Apweiler, M. C. Blatter, A. Estreicher, E. Gasteiger, M. J. Martin, K. Michoud, C. O'Donovan, I. Phan, S. Pilbout, and M. Schneider. 2003. The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Res.* **31**:365–370.
- Casjens, S., and R. Hendrix. 1988. Control mechanisms in dsDNA bacteriophage assembly, p. 15–91. *In* R. Calendar (ed.), *The bacteriophages*, vol. 1. Plenum Press, New York, NY.
- Creuzburg, K., J. Recktenwald, V. Kuhle, S. Herold, M. Hensel, and H. Schmidt. 2005. The Shiga toxin 1-converting bacteriophage BP-4795 encodes an NleA-like type III effector protein. *J. Bacteriol.* **187**:8494–8498.
- de la Bastide, M., and W. R. McCombie. 2007. Assembling genomic DNA sequences with PHRAP. *Curr. Protoc. Bioinformatics* **Chapter 11**:Unit11.4.
- EFSA. 2007. Scientific opinion of the panel on biological hazards on a request from EFSA on monitoring of verotoxigenic *Escherichia coli* (VTEC) and identification of human pathogenic types. *EFSA J.* **5**:79–161.
- Ewing, B., and P. Green. 1998. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res.* **8**:186–194.
- Ewing, B., L. Hillier, M. C. Wendt, and P. Green. 1998. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res.* **8**:175–185.
- Friedrich, A. W., M. Bielaszewska, W. L. Zhang, M. Pulz, T. Kuczus, A. Ammon, and H. Karch. 2002. *Escherichia coli* harboring Shiga toxin 2 gene variants: frequency and association with clinical symptoms. *J. Infect. Dis.* **185**:74–84.
- Friedrich, A. W., J. Borell, M. Bielaszewska, A. Fruth, H. Tschape, and H. Karch. 2003. Shiga toxin 1c-producing *Escherichia coli* strains: phenotypic and genetic characterization and association with human disease. *J. Clin. Microbiol.* **41**:2448–2453.
- Gilmour, M. W., H. Tabor, G. Wang, C. G. Clark, D. M. Tracz, A. B. Olson, M. Mascarenhas, M. A. Karmali, T. Mailman, and L. K. Ng. 2007. Isolation and genetic characterization of a coinfection of non-O157 Shiga toxin-producing *Escherichia coli*. *J. Clin. Microbiol.* **45**:3771–3773.
- Gordon, D., C. Abajian, and P. Green. 1998. Consed: a graphical tool for sequence finishing. *Genome Res.* **8**:195–202.
- Gordon, D., C. Desmarais, and P. Green. 2001. Automated finishing with autofinish. *Genome Res.* **11**:614–625.
- Haft, D. H., B. J. Loftus, D. L. Richardson, F. Yang, J. A. Eisen, I. T. Paulsen, and O. White. 2001. TIGRFAMs: a protein family resource for the functional identification of proteins. *Nucleic Acids Res.* **29**:41–43.
- Herold, S., H. Karch, and H. Schmidt. 2004. Shiga toxin-encoding bacteriophages—genes in motion. *Int. J. Med. Microbiol.* **294**:115–121.
- Jackson, M. P., J. W. Newland, R. K. Holmes, and A. D. O'Brien. 1987. Nucleotide sequence analysis of the structural genes for Shiga-like toxin I encoded by bacteriophage 933J from *Escherichia coli*. *Microb. Pathog.* **2**:147–153.
- Jacobs, G. H., P. A. Stockwell, W. P. Tate, and C. M. Brown. 2006. Transterm-extended search facilities and improved integration with other databases. *Nucleic Acids Res.* **34**:37–40.
- Johansen, B. K., Y. Wasteson, P. E. Granum, and S. Brynstad. 2001. Mosaic structure of Shiga-toxin-2-encoding phages isolated from *Escherichia coli* O157:H7 indicates frequent gene exchange between lambdoid phage genomes. *Microbiology* **147**:1929–1936.
- Kanehisa, M., and S. Goto. 2000. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**:27–30.
- Karch, H., H. I. Huppertz, J. Bockemuhl, H. Schmidt, A. Schwarzkopf, and R. Lissner. 1997. Shiga toxin-producing *Escherichia coli* infections in Germany. *J. Food Prot.* **60**:1454–1457.
- Lin, Z., H. Kurazono, S. Yamasaki, and Y. Takeda. 1993. Detection of various variant verotoxin genes in *Escherichia coli* by polymerase chain reaction. *Microbiol. Immunol.* **37**:543–548.
- Linke, B., A. C. McHardy, H. Neuweger, L. Krause, and F. Meyer. 2006. REGANOR: a gene prediction server for prokaryotic genomes and a database of high quality gene predictions for prokaryotes. *Appl. Bioinformatics* **5**:193–198.
- Lowe, T. M., and S. R. Eddy. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* **25**:955–964.
- Mead, P. S., L. Slutsker, V. Dietz, L. F. McCaig, J. S. Bresee, C. Shapiro, P. M. Griffin, and R. V. Tauxe. 1999. Food-related illness and death in the United States. *Emerg. Infect. Dis.* **5**:607–625.
- Meyer, F., A. Goesmann, A. C. McHardy, D. Bartels, T. Bekel, J. Clausen, J. Kalinowski, B. Linke, O. Rupp, R. Giegerich, and A. Puhler. 2003. GenDB—an open source genome annotation system for prokaryote genomes. *Nucleic Acids Res.* **31**:2187–2195.
- Muniesa, M., M. de Simon, G. Prats, D. Ferrer, H. Panella, and J. Jofre. 2003. Shiga toxin 2-converting bacteriophages associated with clonal variability in *Escherichia coli* O157:H7 strains of human origin isolated from a single outbreak. *Infect. Immun.* **71**:4554–4562.
- Nakao, H., K. Kimura, H. Murakami, T. Maruyama, and T. Takeda. 2002. Subtyping of Shiga toxin 2 variants in human-derived Shiga toxin-producing *Escherichia coli* strains isolated in Japan. *FEMS Immunol. Med. Microbiol.* **34**:289–297.
- Neely, M. N., and D. I. Friedman. 1998. Arrangement and functional identification of genes in the regulatory region of lambdoid phage H-19B, a carrier of a Shiga-like toxin. *Gene* **223**:105–113.
- Nishikawa, Y., Z. Zhou, A. Hase, J. Ogasawara, T. Cheasty, and K. Haruki. 2000. Relationship of genetic type of Shiga toxin to manifestation of bloody diarrhea due to enterohemorrhagic *Escherichia coli* serogroup O157 isolates in Osaka City, Japan. *J. Clin. Microbiol.* **38**:2440–2442.
- Ohnishi, M., J. Terajima, K. Kurokawa, K. Nakayama, T. Murata, K. Tamura, Y. Ogura, H. Watanabe, and T. Hayashi. 2002. Genomic diversity of enterohemorrhagic *Escherichia coli* O157 revealed by whole genome PCR scanning. *Proc. Natl. Acad. Sci. USA* **99**:17043–17048.
- Rangel, J. M., P. H. Sparling, C. Crowe, P. M. Griffin, and D. L. Swerdlow. 2005. Epidemiology of *Escherichia coli* O157:H7 outbreaks, United States, 1982–2002. *Emerg. Infect. Dis.* **11**:603–609.
- Santapaola, D., F. Del Chierico, A. Petrucca, S. Uzzau, M. Casalino, B. Colonna, R. Sessa, F. Berlutti, and M. Nicoletti. 2006. Apyrase, the product of the virulence plasmid-encoded *phoN2* (*apy*) gene of *Shigella flexneri*, is necessary for proper unipolar IcsA localization and for efficient intercellular spread. *J. Bacteriol.* **188**:1620–1627.
- Scheutz, F., and B. Strockbine. 2005. Genus I. *Escherichia*, p. 607–624. *In* G. M. Garrity, D. J. Brenner, N. R. Krieg, and J. T. Staley (ed.), *Bergey's manual of systematic bacteriology*. Springer, New York, NY.
- Schmidt, H. 2001. Shiga-toxin-converting bacteriophages. *Res. Microbiol.* **152**:687–695.
- Schmitt, C. K., M. L. McKee, and A. D. O'Brien. 1991. Two copies of Shiga-like toxin II-related genes common in enterohemorrhagic *Escherichia coli* strains are responsible for the antigenic heterogeneity of the O157:H⁻ strain E32511. *Infect. Immun.* **59**:1065–1073.

45. Serra-Moreno, R., J. Jofre, and M. Muniesa. 2007. Insertion site occupancy by *stx*₂ bacteriophages depends on the locus availability of the host strain chromosome. *J. Bacteriol.* **189**:6645–6654.
46. Shaikh, N., and P. I. Tarr. 2003. *Escherichia coli* O157:H7 Shiga toxin-encoding bacteriophages: integrations, excisions, truncations, and evolutionary implications. *J. Bacteriol.* **185**:3596–3605.
47. Strauch, E., R. Lurz, and L. Beutin. 2001. Characterization of a Shiga toxin-encoding temperate bacteriophage of *Shigella sonnei*. *Infect. Immun.* **69**:7588–7595.
48. Strauch, E., C. Schaudinn, and L. Beutin. 2004. First-time isolation and characterization of a bacteriophage encoding the Shiga toxin 2c variant, which is globally spread in strains of *Escherichia coli* O157. *Infect. Immun.* **72**:7030–7039.
49. Wagner, P. L., J. Livny, M. N. Neely, D. W. Acheson, D. I. Friedman, and M. K. Waldor. 2002. Bacteriophage control of Shiga toxin 1 production and release by *Escherichia coli*. *Mol. Microbiol.* **44**:957–970.
50. Wei, Y., J. M. Lee, D. R. Smulski, and R. A. LaRossa. 2001. Global impact of *sdhA* amplification revealed by comprehensive gene expression profiling of *Escherichia coli*. *J. Bacteriol.* **183**:2265–2272.
51. Willshaw, G. A., T. Cheasty, H. R. Smith, S. J. O'Brien, and G. K. Adak. 2001. Verocytotoxin-producing *Escherichia coli* (VTEC) O157 and other VTEC from human infections in England and Wales: 1995–1998. *J. Med. Microbiol.* **50**:135–142.
52. Willshaw, G. A., H. R. Smith, T. Cheasty, P. G. Wall, and B. Rowe. 1997. Vero cytotoxin-producing *Escherichia coli* O157 outbreaks in England and Wales, 1995: phenotypic methods and genotypic subtyping. *Emerg. Infect. Dis.* **3**:561–565.
53. Zhang, W., M. Bielaszewska, A. W. Friedrich, T. Kuczius, and H. Karch. 2005. Transcriptional analysis of genes encoding Shiga toxin 2 and its variants in *Escherichia coli*. *Appl. Environ. Microbiol.* **71**:558–561.
54. Ziebell, K., M. Steele, Y. Zhang, A. Benson, E. N. Taboada, C. Laing, S. McEwen, B. Ciebin, R. Johnson, and V. Gannon. 2008. Genotypic characterization and prevalence of virulence factors among Canadian *Escherichia coli* O157:H7 strains. *Appl. Environ. Microbiol.* **74**:4314–4323.
55. Zurawski, D. V., C. Mitsuhashi, K. L. Mumy, B. A. McCormick, and A. T. Maurelli. 2006. OspF and OspC1 are *Shigella flexneri* type III secretion system effectors that are required for postinvasion aspects of virulence. *Infect. Immun.* **74**:5964–5976.

Editor: B. A. McCormick