J Mol Evol (2005) 60:546–556
DOI: 10.1007/s00239-004-0263-9

JOURNAL OF MOLECULAR EVOLUTION

© Springer Science+Business Media, Inc. 2005

Phage Associated Bacteriocins Reveal a Novel Mechanism for Bacteriocin Diversification in *Klebsiella*

Milind Chavan,[2] Hamid Rafi,[1] John Wertz,[1] Carla Goldstone,[1] Margaret A. Riley[2]

[1] Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT 06511, USA
[2] Department of Biology, University of Massachusetts Amherst, 611 North Pleasant Street, Amherst, MA 01003, USA

Received: 24 March 2004 / Accepted: 11 October 2004 [*Reviewing Editor:* Prof. David Guttman]

Abstract. Ninety-six isolates of *Klebsiella pneumoniae* and *K. oxytoca* were recovered from wild mammals in Australia. 14.6% of these bacteria produce killing phenotypes that suggest the production of bacteriocin toxins. Cloning and sequencing of the gene clusters encoding two of these killing phenotypes revealed two instances of a bacteriocin associated with a bacteriophage gene, the first such genetic organization described. The newly identified klebicin C gene cluster was discovered in both *K. pneumoniae* and *K. oxytoca*. The newly identified klebicin D gene cluster was detected in *K. oxytoca*. Protein sequence comparisons and phylogenetic inference suggest that klebicin C is most closely related to the rRNase group of colicins (such as colicin E4), while klebicin D is most closely related to the tRNase group of colicins (such as colicin D). The klebicin C and D gene clusters have similar genetic and regulatory organizations. In both cases, an operon structure is inferred consisting of a phage-associated open reading frame and klebicin activity and associated immunity genes. This novel bacteriophage/bacteriocin organization may provide a novel mechanism for the generation of bacteriocin diversity in *Klebsiella*.

Key words: Klebicin — Colicin — Diversification — Enteric bacteria — *Klebsiella*

Correspondence to: Margaret A. Riley; *email:* riley@bio.umass.edu

Introduction

Bacteriocins are a diverse family of protein toxins produced by bacteria, which kill or inhibit members of the same or closely related species (Tagg et al. 1976). The bacteriocins produced by *E. coli*, the colicins, were the first to be identified and have since served as a model system for investigations of bacteriocin structure and function, genetic organization, ecology and evolution (Braun et al. 1994; Riley 1993a,b, 1998; Riley and Wertz 2002; Roos et al. 1989; Tan and Riley 1997).

All colicins studied to date exhibit a unique, three-domain structure. The receptor-binding domain recognizes and binds to specific cell surface receptors on the target cells. The N-terminal translocation domain interacts with cell membrane proteins to gain access to the cell interior. The C-terminal domain specifies a killing activity, such as pore formation or nuclease activity. An immunity protein protects the producing cell by binding to the killing domain of the colicin protein. A lysis protein is often associated with colicin production and plays a role in the release of colicin from the producing cell (Pugsley and Schwartz 1983a,b).

Studies comparing colicin toxins and the genes associated with their production have proven useful in inferring the evolutionary mechanisms that contribute to the existing diversity of bacteriocins (Riley 1993a, 1998; Riley and Wertz 2002; Tan and Riley 1997). Such studies suggest that bacteriocins experience the powerful force of diversifying selection that continuously sifts through novel toxin

Table 1. *Klebsiella* strains used in this study, modified from Gordon and FitzGibbon (1999)

Strain	Host order–family	Host organsim	Location
		Klebsiella pneumoniae	
Misc40	Rodent–Muridae	*Mus musculus*	Stones Farm, Victoria
Misc138	Rodent–Muridae	*Notomys mitchellii*	Wyperfeld National Park, Victoria
		Klebsiella oxytoca	
Misc192	Rodent–Muridae	*Zyzomys argurus*	Kakadu National Park, Northern Territory
Misc712	Bat–Vespertilionidae	*Chalinolobus gouldii*	Gol Gol, New South Wales
Misc170	Marsupial–Phalangeridae	*Trichosurus vulpecula*	Batalling, Western Australia

variants created by recombination and mutation. An extraordinary array of colicins and their associated immunity proteins have resulted from these processes, making them a useful model for studying the evolution of protein diversity at the molecular level.

Colicins have served as the primary model for research on bacteriocin molecular evolution. Only recently have bacteriocins of other enteric bacteria been examined (Riley et al. 2001; Wertz and Riley 2004), such as cloacins of *Enterobacter cloacae* (Tieze et al. 1969), klebicins of *Klebsiella* species (James et al. 1987; Riley et al. 2001), alveicins of *Hafnia alvei* (Wertz and Riley 2004), marcescins of *Serratia marcescens* (Guasch et al. 1995; Nasu 1981; Traub et al. 1971; Viejo et al. 1992), and pesticins of *Yersinia pestis* (Rakin et al. 1996). However, as yet we know little about how this diversity of enteric toxins arose, diversified and is maintained in microbial populations.

In the present study the bacteriocin-encoding DNA of two klebicin-producing isolates of *Klebsiella* were cloned and their nucleotide sequences determined to further examine the evolutionary relationships between klebicins and other colicin and colicin-like toxins. DNA and protein sequence analysis reveals that these klebicins are the result of numerous recombination events generating novel combinations of receptor binding, translocation and killing domains. Further, these klebicins are the first bacteriocins found in close association with bacteriophage genes. This association suggests that *Klebsiella* has evolved a novel mechanism for the lateral transfer and diversification of its klebicins.

Materials and Methods

Microbial Strains and Media

The *Klebsiella* strains used in this study are from the Australian Enteric Collection, which comprises environmental enteric bacteria isolated from wild mammals in Australia (Gordon and FitzGibbon 1999). Information about the strains, including geographic origin and host, is listed in Table 1. The *E. coli* strain DH10B (F′–*mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) ϕ80*lac*ZΔM15 Δ*lac*X74 *deo*R *recA*1 *end*A1 *ara*D139 Δ(*ara*, *leu*)7697 *gal*U *gal*K λ$^-$*rps*L *nup*G) (Invitrogen) and NM522 (*supE*, *thi*, Δ(*lac-proAB*), Δ *hsd*5 (r-,m-), F′ (*proAB*, *lack*^{q}*Z* ΔM15) (Promega) were used as host strains for cloned bacteriocin genes. All *Klebsiella* and *E. coli* strains were cultured in Luria Broth (LB) (Sigma) at 37°C. LB medium was supplemented with ampicillin (100 μg/ml) for culturing the *E. coli* DH10B and *E. coli* NM522 containing *Klebsiella* genomic libraries.

Cloning and DNA Sequence Determination

Genomic and plasmid DNA were extracted using DNeasy Tissue kits (Qiagen) and Plasmid Miniprep kits (Qiagen). Cloning of the klebicin gene clusters was performed using standard protocols (Maniatis et al. 1982). Partially *Sau3a1* restricted genomic DNA from Misc40 was cloned into the *Bam*HI site of phagemid pBluescript II SK(+) (Strategene) and transformed into *E. coli* NM522 cells by electroporation. Partially restricted *Tsp509I* genomic DNA from Misc192 was cloned into the *Eco*RI site of phagemid pBluescript II SK(+) (Stratagene) and transformed into *E. coli* DH10B cells by electroporation. Transformants from both libraries were screened for bacteriocin production as described by Pugsley and Oudega (1987) using bacteriocin-sensitive strains Misc138 or Misc170, respectively, for screening the Misc40 or Misc192 libraries. Mitomycin C (0.5 μg/ml) was used to induce bacteriocin production in the transformants (Pugsley and Oudega 1987).

The cloned fragments encoding the klebicin killing activity (Misc40, klebicin C clone; and Misc192, klebicin D clone) were sequenced using Big Dye terminator in an ABI 377 DNA sequencer (Applied Biosystems) starting with the primers KS and SK located in the multiple cloning site of pBluescript SK(+) vector. Additional primers were designed to allow sequence walking through the cloned inserts. Double coverage was ensured by sequencing both strands of DNA. Sections of the bacteriocin gene cluster amplified by PCR from *Klebsiella* Misc712 were purified using QIA quick PCR purification Kit (Qiagen), and then sequenced as described above using the PCR primers.

PCR Screening

Klebsiella oxytoca Misc712 shared a similar killing spectrum as *K. pneumoniae* Misc40. This strain was screened for the presence of a klebicin gene cluster using PCR. Primers 5′-GCTCTGT AACCTTCAAGTTCTC-3′ and 5′-CAAGCAAGATTACGGTC TACTC-3′ yielded the expected 2.3-kb PCR product. PCR was performed using 1 unit of *Taq* polymerase (Promega) using standard PCR buffers. Annealing temperature of 54°C for 40 s and extension temperature of 72°C for 2 min were used for each of a total of 32 cycles.

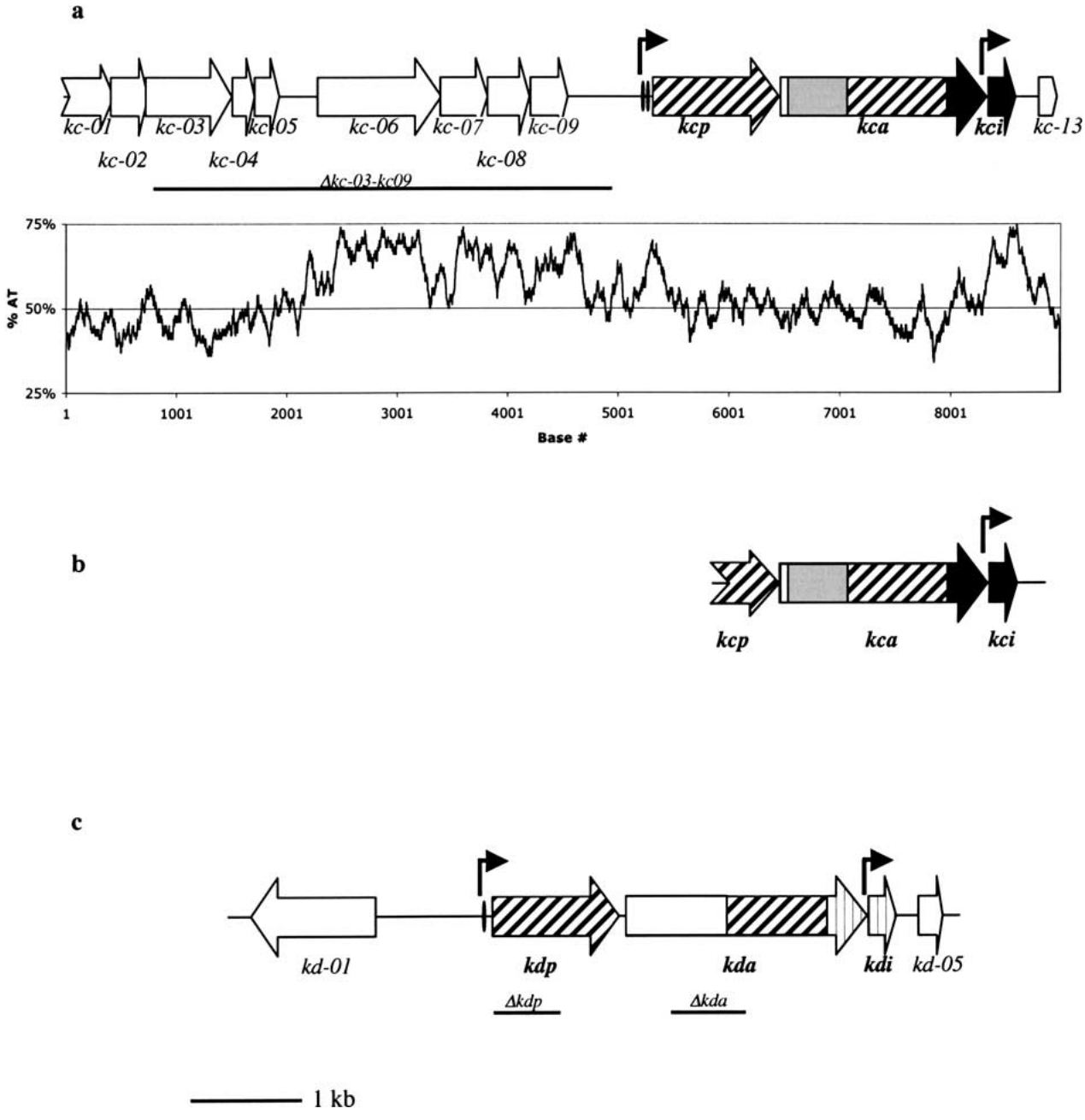

Fig. 1. Klebicin clone organization. (**a**) Klebicin C clone from *K. pneumoniae* Misc40. (**b**) Klebicin C gene cluster from *K. oxytoca* Misc712. (**c**) Klebicin D clone from *K. oxytoca* Misc192. Promoter regions (┌•), Lex A binding boxes (|), and similarity to each other (▨) and to known colicins E4 (■), E1 (▩), and D (▥) are indicated. The %AT composition over a 100-bp sliding window of klebicin C is represented below its genetic map. The Δ*kc-03–kc-09*, Δ*kdp*, and Δ*kda* deletions are indicated by bold lines.

Generation of Deletions in Klebicin Clones

A 4024-bp BsrGI fragment was deleted from the klebicin C clone to yield plasmid pBS-40r (Fig. 1a) using standard methods (Maniatis et al. 1982). A 768-bp *Eco*RI fragment, which includes 760 bases of the *kdp* gene, was deleted from the klebicin D clone to yield plasmid pBS-192 Δkdp (Fig. 1c). A 809-bp *Bsr*GI fragment was deleted from the *kda* gene of klebicin D clone to yield a plasmid pBS-192Δkda (Fig. 1c). The deletions were confirmed by sequencing (data not shown). The resulting plasmids were electroporated into *E. coli* NM522 cells, grown in LB at 37°C with shaking, and bacteriocin production was induced by the addition of 0.5 μg/ml mitomycin C (Maniatis et al. 1982; Pugsley and Oudega 1987). Killing activity was assayed against the bacteriocin-sensitive strain Misc 170 or Misc138.

DNA and Protein Sequence Analysis

The sequenced contigs were assembled using Sequencher 4.1 software (Gene Codes Corp.). The DNA sequences were queried against NCBI databases (NCBI: http://www.ncbi.nlm.nih.gov/BLAST/) using both nucleotide-BLAST (blastn) and translated-BLAST (tblastx) programs. Open reading frames (ORFs) were detected with software MapDraw (DNASTAR, Inc.) and the

Table 2. Klebicin C clone open reading frames

ORF	Name	Amino acids	ORF location	Blast similarity	Accession No.	e value
1	*kc-01*	>158	<1–475	Replication protein P–bacteriophage lambda	NP_040632.1	1e-60
2	*kc-02*	111	468–803	Putative DNA-binding protein–Gipsy2 prophage in *Salmonella typhimurium* LT2	NP_459991 AAL19950	2e-12
3	*kc-03*	261	796–1581	Phage-related conserved hypothetical protein–*Bordetella bronchiseptica*	NP_888219.1	1e-70
4	*kc-04*	67	1578–1781	Hypothetical protein–*E. coli*	NP 755480	1e-05
5	*kc-05*	74	1774–1998	Hypothetical protein–*S. typhi*	NP_058299.1	0.011
6	*kc-06*	373	2304–3425	Ser/Thr protein kinase–*Plasmodium falciparum*	NP_705451.1	0.011
7	*kc-07*	143	3438–3869	putative vaccinia A22R homolog–*Melanoplus sanguinipes entomopoxvirus*	AAC97652	0.79
8	*kc-08*	125	3869–4246	Hypothetical protein–*Ustilago maydis*	UM_05432.1	1
9	*kc-09*	113	4179–4520	Cytidine deaminase–*Streptococcus pneumoniae*	NP_358781.1	5.3
10	*kcp*	380	5307–6449	Tail fiber distal subunit–bacteriophage KVP40	NP_899545.1	1e-16
11	*kca*	619	6449–8308	Colicin E1–*E. fergusonii*	AAN76838.1	4e-31
				Colicin D–*E. coli*	P17998	4e-62
				Colicin E4–*E. coli*	S49176	1e-48
12	*kci*	84	8318–8572	E4 immunity–*E. coli*	S49177	3e-31
13	*kc-13*	>57	8808–	Protein Q–prophage Fels-1	NP_459880	3e-20

translated ORFs were searched using the blastp program against the NCBI databases. Bacteriocin activity proteins were aligned using the ClustalX algorithm in Megalign (DNASTAR, Inc.) and then converted back into DNA for subsequent phylogenetic analysis. A maximum parsimony tree was inferred using PAUP version 4.0b10 (Swofford 2002) and evaluated using 10,000 bootstrap replications. Kyte–Doolittle hydropathy plots of proteins were constructed using the software Protean 5.52 (DNASTAR, Inc.) using a sliding window of nine amino acids. The AT composition of the cloned sequences was evaluated using a sliding window of 100 bp.

Results

A previous survey of bacteriocin production in isolates of *K. pneumoniae* and *K. oxytoca* revealed two novel bacteriocin-like killing phenotypes. Genomic libraries of the two producing strains (Misc40 and 192) were created and screened for killing phenotypes. A single klebicin-producing clone containing a 9.0-kb insert (GenBank accession AY578793) was recovered from Misc40 and was referred to as the klebicin C clone. A single klebicin-producing clone containing a 6.5-kb insert (GenBank accession AY578792) was recovered from Miscl92 and was referred to as the klebicin D clone.

Klebicin C Clone

A 9.0-kb fragment containing the klebicin C killing activity was cloned from *K. pneumoniae* Misc40. The nucleotide sequence of the clone contains thirteen open reading frames (ORFs), all in the same transcriptional orientation (Fig. 1a and Table 2). Five of the ORFs are similar in sequence to bacteriophage genes (Table 2) including the ORFs; *kc*-01, which encodes a protein 72% identical to replication protein P of bacteriophage lambda (Sanger et al. 1982); *kc*-13, which encodes a protein 84% identical to prophage Fels1 antiterminator; and *kcp*, which encodes a protein 45% similar to tail fiber proteins of the Myoviridae bacteriophage family (Table 2). A deletion of ~4 kb from the klebicin C clone, which includes *kc-03* to *kc-09* (Fig. 1a), did not affect the resulting killing phenotype. This deletion reveals that ORFs *kc-03* to *kc*-09 do not play a role in bacteriocin-mediated killing.

Two colicin-like genes were identified in the klebicin C clone. The 1860-bp *kca* and 255-bp *kci* genes encode putative bacteriocin activity and immunity proteins, respectively (Fig. 1a and Table 2). These genes are found in the same transcriptional orientation, an arrangement typical of nuclease colicins. The encoded klebicin C activity protein (klebicin C) contains 619 amino acids with a deduced molecular mass of ~66 kDa. The C-terminal 95 amino acids are 90% identical to the killing domain of colicin E4 (Fig. 1a), a member of the rRNase group of colicins, which kills by inhibiting protein synthesis through hydrolysis of 16S rRNA (Smarda et al. 1988). The remainder of the protein shows more limited sequence similarity with known bacteriocins. Residues 22–239 are 45% identical to the translocation domain of E1, which is a pore-forming colicin produced by *Escherichia fergusonii* that utilizes the TolCAQ translocation pathway (Lazdunski et al. 1998). The central region (residues 201–521) is 45% identical to the colicin D receptor-binding domain, which is a colicin produced by *E. coli* and known to bind to the FepA cell surface receptor (Braun 1995). However, this central region is most similar (63% sequence identity) to the corresponding region in klebicin D described below.

Table 3. Klebicin D clone open reading frames

ORF	Name	Amino acids	ORF	Blast similarity	Accession No.	e value
1	*kd-01*	206	821–201	Putative resolvase–*Y. pestis*	NP_395416	1e-93
2	*kdp*	377	2381–3514	Protein GP37–bacteriophage K3	Q38394	1e-15
3	*kda*	716	3595–5745	Colicin D-157–*E. coli*	CAA71433.1	e-131
4	*kdi*	88	5742–6008	D immunity–*E. coli*	S09255	3e-25
5	*kd-05*	75	6188–6415	No significant similarity found	—	-

The klebicin C immunity protein contains 84 amino acids (molecular mass, ~10 kDa) and shares 78% protein sequence identity with the colicin E4 immunity protein (Table 2). Two regions of the immunity protein (residues 6–20 and 51–73) harbor 17 of the 19 residues that differ from colicin E4. This same region segregates a significant number of differences when the immunity proteins of other rRNase colicins are compared. The elevated levels of divergence in these two restricted regions suggest they play a role in creating immunity specificity.

Colicin gene clusters are characteristically regulated by the SOS response and are thus repressed by the LexA repressor (Parker 1986; Riley et al. 2001; van den Elzen et al. 1982, 1983a). Riley et al. (2001) published an alignment of colicin promoters showing a dual overlapping LexA binding box between the Pribnow box and the start codon of most colicin genes. A similarly structured presumptive promoter containing the perfect palindrome, 5′ ATAACCACTGTATAAATACACAGCTGTGTATTTATACAGTGGTTAT 3′, that includes two predicted LexA binding sites (underlined bases) was identified ~1200 bp upstream of *kca* (Fig. 1a). This promoter appears to regulate the expression of an operon consisting of the phage-like *kcp* gene and the colicin-like *kca* and *kci* genes (Fig. 1a). A putative stem-loop transcriptional terminator (5′ AAATATAGCCAGCCTTCGTAGTTGGCTATATTT 3′) was identified downstream of *kci*.

A killing phenotype similar to that produced by *K. pneumoniae* Misc40 was detected in *K. oxytoca* Misc712. This prompted a PCR screen of Misc712 using primers designed from the klebicin C sequence. PCR products were sequenced and assembled to obtain a 2992-bp sequence (Fig. 1b; GenBank accession AY578794). The klebicin C gene cluster is highly conserved (99.7% identity) between these two producers, isolated from two different species, *K. pneumoniae* and *K. oxytoca*. Two nonsynonymous changes (amino acid replacements 255-Y to F and 366-K to N) were detected in the *kcp* gene. Three differences were detected in the *kca* gene: two nonsynonymous substitutions (116-L to Q and 260-K to M) and one synonymous substitution (position 1359). The immunity gene, *kci*, is identical between these strains. Four changes were detected in the 3′ flanking region (35,134, 224, and 255 bp downstream of the *kci* stop codon).

Klebicin D

A 6510-bp fragment containing the klebicin D killing activity was cloned from *K. oxytoca* Misc192. The nucleotide sequence of the clone revealed five ORFs, three of which form a putative bacteriocin operon (Fig. 1c and Table 3). The klebicin D operon consists of *kdp* (1134-bp homologue of the bacteriophage-like *kcp* gene identified in the klebicin C operon), *kda* (2151-bp putative bacteriocin activity gene), and *kdi* (267-bp putative immunity gene). *kdp* encodes a protein 70% identical to that encoded by *kcp* and 43% similar to the C-terminal domain of the phage tail protein of bacteriophage K3 (Fig. 1 and Table 3). Deletion of 760 bp of the *kdp* gene (Fig. 1c) results in loss of killing activity, suggesting that this protein is required for *kda*-encoded toxin activity.

The encoded klebicin D protein contains 716 amino acids with a molecular mass of ~76 kDa. The C-terminal 91 amino acids are 85% identical to the corresponding region in colicin D, which is produced by *E. coli* and functions by hydrolyzing arginine tRNA to inhibit protein synthesis (Masaki and Ogawa 2002; Tomita et al. 2000) (Fig. 1c). This suggests that klebicin D is a member of the tRNase group of bacteriocins that includes colicins D and E5 (Hofinger et al. 1998; Masaki et al. 1997; Roos et al. 1989). The receptor-binding domain of klebicin D (residues 301–613) is most similar to the corresponding region of klebicin C (63% sequence identity) (Fig. 1). More limited similarity (45%) is detected to the colicin D receptor region. The N-terminal translocation domain is unique. Deletion of the 809 bp of the *kda* gene (Fig. 1c) results in loss of killing activity, suggesting that *kda* is responsible for the killing phenotype.

The immunity gene, *kdi*, is located downstream of *kda* in the same transcriptional orientation, as observed in klebicin C and other nuclease colicin operons. It encodes an 88-amino acid protein (molecular mass, ~10 kDa) that is 59% identical to the colicin D

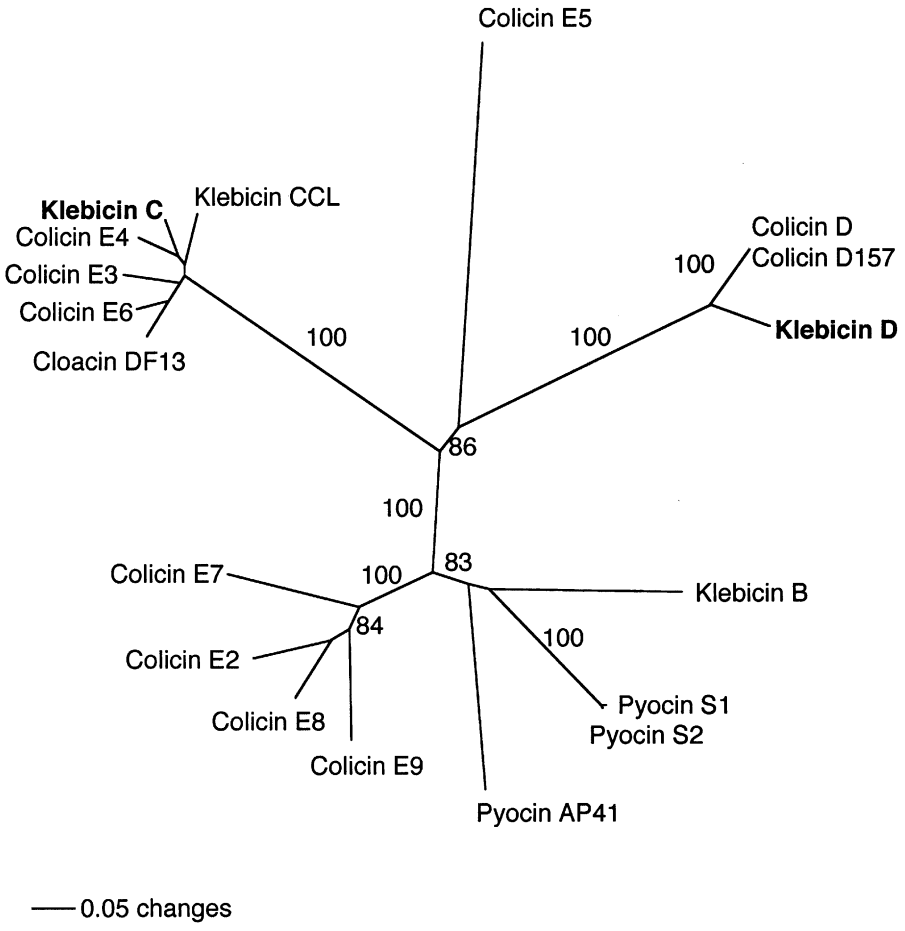

Fig. 2. Phylogenetic tree inferred for killing domains of nuclease colicin and colicin-like proteins; 10,000 bootstrap replicates were performed and values greater than 65% are reported. Klebicins C and D, described in this article, are in boldface.

immunity protein. A total of 35 amino acid replacements distributed over the length of protein were observed between the immunity proteins of klebicin D and colicin D.

The operon structure of klebicin D is similar to that determined for klebicin C (Fig. 1c). The *kda* start codon is 80 bases downstream of the *kdp* stop codon, while *kdi* overlaps *kda* by 4 bp. Expression of this gene cluster appears to be controlled by an SOS regulated promoter identified upstream of *kdp*. Unlike the dual overlapping LexA sites found in klebicin C and most colicin operons, the klebicin D operator appears to contain only a single LexA binding site (5′ TACTGTATGTATATACAGTA 3′) that matches the bacterial LexA binding site consensus sequence 5′ TACTG(TA)$_5$CAGTA 3′ (Walker 1984) more closely than any other colicin promoters (Riley et al. 2001). The observation that klebicin D expression is induced by mitomycin C (data not shown) further suggests that the LexA binding site controls klebicin expression. A second putative promoter was identified upstream of *kdi* and is presumed to be a constitutive promoter that ensures low-level-immunity protein production to protect the cell from klebicin D activity (Chak and James 1984, 1985; Cole et al. 1985). A putative stem-loop transcription terminator sequence (5′ TAGCCAACCTCTCGGGTTGGCTA 3′) was identified 26 bp downstream of the *kdi* stop codon.

Discussion

The focus of this study is the molecular characterization of two novel klebicin gene clusters and an investigation into the evolutionary mechanisms responsible for generating klebicin diversity. To date, only two bacteriocins from *Klebsiella* have been sequenced, klebicins B (Riley et al. 2001) and CCL (NCBI accession AF190857). The klebicin B operon is a chimera, composed of short regions each with a different evolutionary origin (Riley et al. 2001). The encoded killing domain is most similar to the corresponding region in pyocin S1, a DNase bacteriocin produced by *Pseudomonas aeruginosa*. The receptor-binding and translocation regions are unique. The lysis gene and flanking regions are most similar to those found in colicin A, a pore-forming bacteriocin produced by *Citrobacter freundii* (Riley et al. 2001). The resulting chimeric klebicin B operon encodes two functions (killing and lysis) found in multiple species

(killing, *P. aeruginosa* and *K. pneumoniae*; lysis, *C. freundii and K. pneumoniae*) and two functions found only in *Klebsiella* (unique receptor recognition and translocation functions), which presumably serve to restrict bacteriocin activity to close relatives of the producing strains.

In sharp contrast, the entire klebicin CCL operon is nearly identical (99%) to that of cloacin DF13, which encodes a nuclease bacteriocin produced by *Enterobacter cloacae*. Cloacin DF13 utilizes the TolABQR pathway for translocation and employs IutA as a cell surface receptor (Cooper and James 1985; James et al. 1987). The near-identity of the translocation and receptor binding domains in DF13 and klebicin CCL suggests that the Tol pathway and the IutA receptor are shared between these species. Indeed, the partially completed *K. pneumoniae* genome sequence (Washington University in St. Louis Genome Sequencing Center; http://genome.wustl.edu) encodes an IutA protein that is 99% identical in sequence to that identified from *E. coli*. Further, this same receptor is present in most enteric species, suggesting that a bacteriocin targeting such a receptor would have a relatively broad killing breadth. A partial sequence of a putative tolABQR gene cluster was identified in the *K. pneumoniae* partial genome sequence, which is 81% identical to the *tol*A gene identified from *E. coli*. The tolABQR genes are found in most species of enteric bacteria, suggesting the presence of a common mode of entry for colicins and colicin-like bacteriocins.

We describe here the sequence and genetic structure of two novel klebicins, C and D. Klebicin C is predicted to be a member of the rRNase bacteriocin group (based on protein sequence similarity) and is produced by both *K. pneumoniae* and *K. oxytoca*. Klebicin D has been found only in *K. pneumoniae* and is predicted to be a member of the tRNase bacteriocin group. Sequence analysis reveals that both klebicins are expressed as SOS-regulated transcriptional units consisting of three genes: *kcp*, *kca*, and *kci* in klebicin C and *kdp*, *kda*, and *kdi* in klebicin D. The *kca*/*kda* and *kci*/*kdi* genes encode the bacteriocin activity and immunity proteins, respectively. The *kcp*/*kdp* genes are similar in sequence (70% identity) and encode a protein similar (~45%) to a family of phage tail fiber-like proteins (Miller et al. 2003; Riede et al. 1986). Although the specific function of the *kcp*/*kdp* genes is unknown, our preliminary evidence (consisting of deletion analysis) suggests that this gene or gene product is required for klebicin production and/or activity. Such an operon structure, with an additional bacteriophage-related gene inserted between the bacteriocin promoter and the toxin-related genes, has not previously been described in the literature.

The four characterized klebicin proteins (B, C, D, and CCL) all appear to share a three-domain structure typical of colicin proteins, with an N-terminal translocation domain, a central receptor binding domain, and a C-terminal killing domain. In the case of klebicin C, the three domains appear to be a composite of sequences found in at least three other bacteriocins (Fig. 1a). The N-terminal translocation domain is similar to the corresponding region of colicin E1, which employs the Tol system to translocate into the target cell (Lazdunski et al. 1998). The highest level of similarity is found in the TolC domain of colicins E1, 5, and 10, suggesting that TolC may be required for translocation of klebicin C (Pilsl and Braun 1995). A protein with 84% identity to the TolC protein of *E. coli* was found in the partially complete *K. pneumoniae* genome sequence (Washington University in St. Louis Genome Sequencing Center; http://genome.wustl.edu), which further supports the proposed role of TolC in klebicin C translocation. The klebicin C receptor recognition region is most similar to the corresponding region in klebicin D, which is unique to these two klebicins. The killing domain of klebicin C is homologous to the corresponding region in rRNase colicins, which function by inhibiting protein synthesis by hydrolyzing 16S rRNA (Smarda et al. 1988). Bacteriocins belonging to rRNase group include E3, E4, and E6 (produced by *E. coli*), cloacin DF13 (produced by *E. cloacae*), and klebicins CCL and C (produced by *K. pneumoniae*). In summary, klebicin C appears to have evolved from multiple recombination events resulting in a chimeric protein consisting of a novel translocation domain fused to a colicin E1-like TolC box, a klebicin D-like receptor binding domain, and a colicin E4-like rRNase killing domain.

The klebicin C operon is highly conserved in two *Klebsiella* isolates (*K. pneumoniae* Misc40 and *K. oxytoca* Misc712), which were identified with identical killing profiles (Riley et al. 2003). These strains were isolated from different hosts (*Mus musculus* and *Chalinolobus gouldii*) and from widely separated geographic regions (Victoria and New South Wales) (Table 1). The near-identity of these two bacteriocin operons (only 9 substitutions in 2977 bp) recovered from two species of *Klebsiella*, and from such disparate mammalian hosts and location, suggests both a recent common ancestry and rapid spread of the bacteriocin.

The partial ORFs flanking the klebicin C operon are similar to the P and Q genes of lambdoid bacteriophage. In phage the region between P and Q can be deleted or replaced while maintaining phage viability. For example, lambdoid phages harboring shiga toxins in this segment have been described (Karch et al. 1999). The klebicin C operon (composed of ~50% G and C) appears to have been inserted into the P

and Q regions of an existing prophage (~50% GC). In addition, an AT-rich region (35% GC) extending from *kc-06* to *kc-09* (Fig. 1a) suggests that this segment may have been acquired via a separate horizontal transfer event. One gene in this segment, *kc-07*, is similar to RTX toxins (Table 2).

Klebicin D belongs to a small family of tRNase bacteriocins, which kill by hydrolyzing specific tRNAs. This family consists of only two additional members, colicins D and E5, which differ with respect to the tRNA species hydrolyzed. Klebicin D is more closely related to colicin D, which targets arginine tRNA (Masaki and Ogawa 2002). Colicin D is processed within the target cell by peptidase to release its tRNase activity (de Zamaroczy et al. 2001). The colicin D processing site is proposed to reside between residues 571 and 606 (de Zamaroczy et al. 2001). This region is highly conserved (92% similarity) between colicin D and klebicin D, suggesting that klebicin D undergoes similar processing. The receptor-binding domain of klebicin D is most closely related to that of klebicin C and its translocation domain is unique. Thus, klebicin D is also a chimeric protein, with a killing domain derived from a nuclease colicin ancestor (such as colicin D), a receptor recognition domain derived from a klebicin C-like ancestor and a unique translocation domain.

Most colicin and colicin-like gene clusters contain a closely linked lysis gene, which encodes a protein involved in the release of colicins from the producer cell (Pugsley and Schwartz 1983b; Riley and Wertz 2002). It is absent from some pore-forming colicin gene clusters, such as colicin B (Pressler et al. 1986) and Ib (Mankovich et al. 1986), but is always found within 100 bp 3′ to the immunity gene in nuclease bacteriocin gene clusters, including klebicins B and CCL of *K. pneumoniae*. No identifiable lysis gene was detected within 400 bp downstream of the immunity genes in the klebicin C and D operons. The identification of putative transcriptional terminators immediately downstream of *kci* and *kdi* and the killing activity of the cloned insert suggests that there is no lysis gene associated with this operon. If so, klebicins C and D are the first colicin-like nuclease bacteriocin operons that lack an identifiable lysis gene.

Deletion analysis revealed that *kdp*, and presumably also *kcp*, is required for klebicin activity. Hydropathy plots and predicted secondary structures were compared between the *kcp* and *kdp* encoded proteins and those of representative colicin lysis proteins. No similarities, in terms of predicted structure or even hydrophilic tendencies, were detected. This suggests that, if *kcp* and *kdp* encode proteins involved in klebicin release, they do so in a manner quite different from traditional colicin lysis proteins. In *Serratia marcescens*, a phage holin-like protein has been proposed to release marcesin 28B, a chromosomally encoded bacteriocin (Ferrer et al. 1996).

Klebicin gene clusters show several similarities to colicin operons. As is consistently found for nuclease colicins, klebicin C and D gene clusters have activity and immunity genes encoded on the same strand of DNA (Braun et al. 1994). This arrangement differs from that found in the pore-forming colicins, which encode the colicin and immunity genes in the opposite transcriptional orientation (Braun et al. 1994). The regulatory organization of the klebicin C and D operons is similar to that observed in colicins. Both klebicin operons are expressed from an SOS-regulated promoter with LexA binding sites located downstream of a −10 promoter. A sequence downstream of the immunity genes that forms a potential stem-loop may serve as the transcriptional terminator for both of these operons. An additional promoter detected upstream of the immunity genes appears to ensure constitutive expression of the immunity protein to protect the cells from their cognate klebicins. Such SOS-independent expression of immunity genes is a general feature of nuclease colicin operons (Chak and James 1984, 1985; Cole et al. 1985). The klebicin C operator has two overlapping LexA binding sites, typical of most colicin operators, including colicins A, B, D, E1, E2, E3, E6, K, N, S4, U, Y, 5, 10, klebicin B, and pesticin (Riley et al. 2001; van den Elzen et al. 1982). Klebicin D has a single LexA binding site, as is found in cloacin DF13 (van den Elzen et al. 1983b), klebicin CCL (NCBI AF190857), marcescin A (unpublished result), and colicins Ia and Ib (Fernandez De Henestrosa et al. 2000).

Klebicins B, C, D, and CCL are predicted to function as nucleases. Figure 2 provides a phylogenetic tree inferred for nuclease killing domains. Three highly divergent lineages corresponding to the DNase, rRNase, and tRNase functional groups of nuclease bacteriocins are distinguished. Klebicins are represented in all three groups: the DNase family (klebicin B), the tRNase family (klebicin D), and the rRNase family (klebicins C and CCL).

It has been demonstrated that many colicins and colicin-like bacteriocins evolve by recombination between existing colicin gene clusters, resulting in novel arrangements of colicin functional domains (such as found in colicins B and D) or in different combinations of colicin, immunity, and lysis genes (such as found in colicins K, 5, and 10) (Braun et al. 1994; Riley 1998; Riley and Wertz 2002). Not surprisingly, klebicins C and D follow this general pattern of recombination-mediated diversification. Each bacteriocin toxin and bacteriocin operon is composed of regions derived from several different putative ancestral sequences.

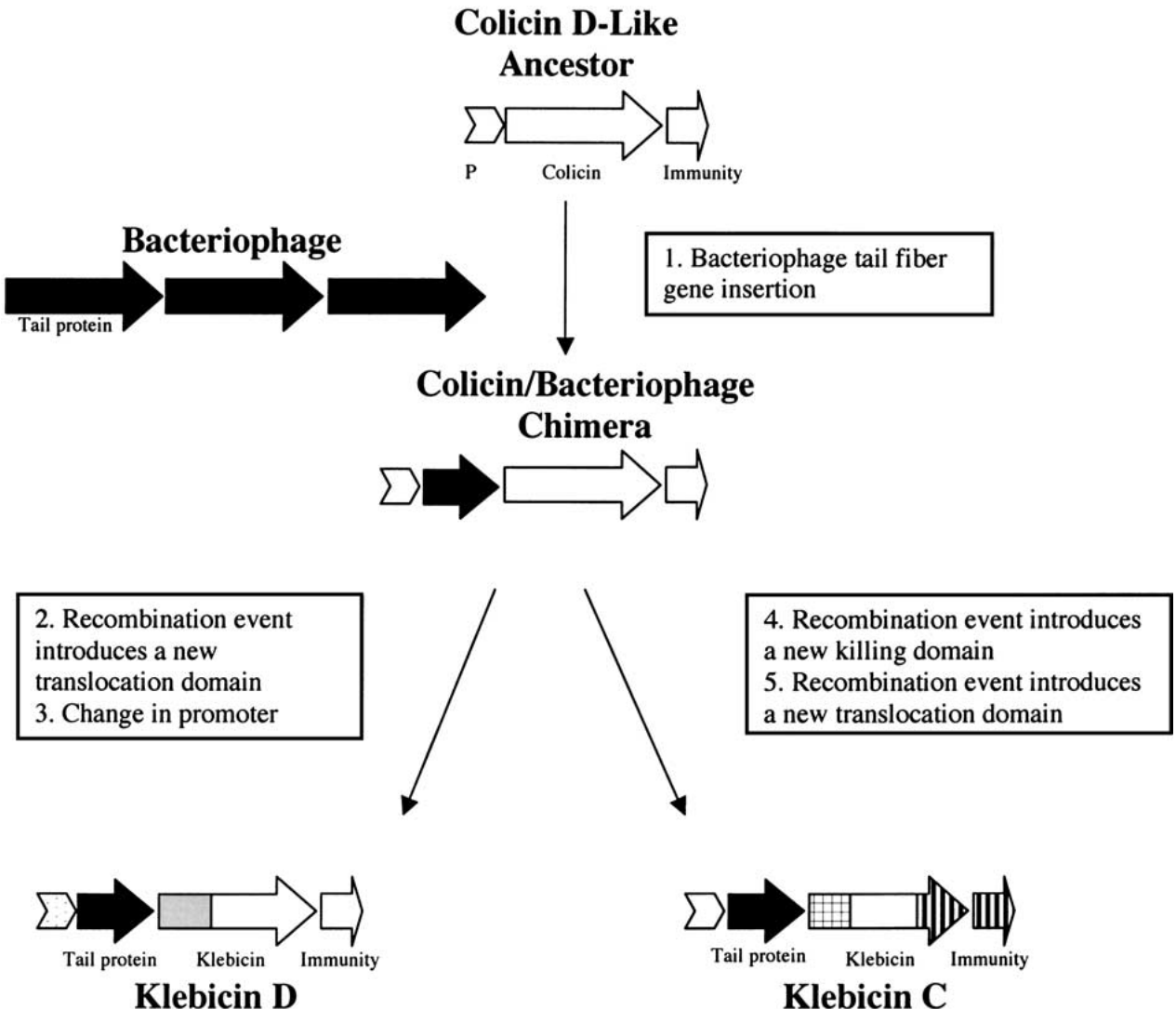

Fig. 3. Hypothetical model of bacteriophage-mediated klebicin diversification, bacteriocin (➪) and bacteriophage (➡) ancestors are indicated and hypothetical recombination events are shaded differentially. LexA binding promoter region is indicated (∑).

A second process of diversification has been implicated in nuclease bacteriocin evolution, which involves diversifying selection acting on the immunity protein and the immunity binding region of the cognate bacteriocin (Riley 1993a, 1998; Riley and Wertz 2002). Studies of the levels and patterns of DNA sequence polymorphism and divergence among nuclease colicin operons have suggested that immunity proteins and the immunity binding regions of their cognate bacteriocins segregate similarly high levels of nucleotide diversity, relative to the remainder of the gene cluster (Riley 1993a, b). These data, combined with experimental evolution studies (Tan and Riley 1997), have been interpreted to suggest that nuclease colicins experience strong diversifying selection acting on novel immunity specificities created by point mutations (Riley and Wertz 2002).

Table 4 provides an estimate of the nucleotide divergence observed between klebicin C and its closest rRNase bacteriocins (cloacin DF13, klebicin CCL, colicins E3, E4, and E6) and klebicin D and its closest tRNase bacteriocin (colicin D). Levels of synonymous substitutions are nearly equal when compared between the regions encoding the bacte-

Table 4. Nucleotide divergence in rRNase and tRNase bacteriocins

Group	K_t	K_s	K_a
rRNase			
Killing domain	0.213	0.66	0.09
Immunity	0.229	0.58	0.15
tRNase			
Killing domain	0.151	0.75	0.09
Immunity	0.220	0.73	0.22

Note. K_t, total sites; K_s, synonymous sites; K_a, nonsynonymous sites. Based on Korber (2000).

riocin killing and immunity functions. In contrast, twice as many nonsynonymous substitutions have accumulated in the region encoding immunity function compared to the region encoding killing function. This pattern of substitution is quite different from that observed among nuclease colicins experiencing diversifying selections, which suggests that a different sort of evolutionary force may be acting on klebicins. Unfortunately, the high levels of recombination detected in klebicin operons preclude analyses aimed at determined whether these operons are experiencing selection.

The klebicin operons described here suggest a third mechanism for bacteriocin diversification, which involves recruiting an additional protein, a phage tail fiber-like protein, within the bacteriocin operon. Phage tail fiber proteins function in contact initiation and binding of the bacteriophage to its target bacterial host (Hendrix et al. 1999, 2003; Tetart et al. 1998). Host specificity is attributed to the variable C-terminal domain of the tail fiber, which can be swapped between different phages (Tetart et al. 1998).

The proteins encoded by *kcp* and *kdp* are similar to the C-terminal domain of several tail fiber proteins, whose functions are equivalent to those provided by the receptor binding domains of colicins. It is possible that klebicins have recruited binding proteins from bacteriophage, in addition to tapping into the existing pool of shared colicin functional domains. Figure 3 illustrates a hypothetical process, involving bacteriocin and bacteriophage interactions, which would result in the observed klebicin C and D operons. The first step in this process proposes a recombination event between a hypothetical colicin D-like ancestral bacteriocin and a Myoviridae-like bacteriophage ancestor resulting in insertion of a phage tail fiber-like protein between the bacteriocin promoter and the activity gene. Subsequent acquisition of a novel translocation domain results in a klebicin D-like bacteriocin. An alternative path involves duplication of the LexA binding site in the bacteriocin promoter region and acquisition of novel killing and translocation domains, resulting in a klebicin C-like bacteriocin.

Not shown in this scenario is the insertion of the entire bacteriocin gene cluster into an existing prophage. The presence of phage-like P and Q genes in the klebicin C operon further supports this hypothesized unique role bacteriophage may play in the evolution of klebicins. The klebicin C operon is straddled by sequences known to reside in prophage. Such genetic elements are mobile and could carry the encoded bacteriocin to new bacterial hosts, as is observed for klebicin C, which was found in Misc40 and Misc712 belonging to *K. pneumoniae* and *K. oxytoca*, respectively.

Acknowledgments. We are grateful to David Gordon (Department of Botany and Zoology, Australia National University, Canberra) for providing all of the natural isolates used in this study. We acknowledge Lisa Nigro and Cynthia Winkworth for their contributions to the sequencing of the klebicin clones. This material is based on work supported by the National Institutes of Health (M.A.R., Grant GM58433-02) as well as the National Science Foundation (J.E.W).

References

Braun V (1995) Energy-coupled transport and signal transduction through the gram-negative outer membrane via TonB-ExbB-ExbD-dependent receptor proteins. FEMS Microbiol Rev 16:295–307

Braun V, Pilsl H, Gross P (1994) Colicins: structures, modes of action, transfer through membranes, and evolution. Arch Microbiol 161:199–206

Chak KF, James R (1984) Localization and characterization of a gene on the ColE3-CA38 plasmid that confers immunity to colicin E8. J Gen Microbiol 130:701–710

Chak KF, James R (1985) Analysis of the promoters for the two immunity genes present in the ColE3-CA38 plasmid using two new promoter probe vectors. Nucleic Acids Res 13:2519–2531

Cole ST, Saint-Joanis B, Pugsley AP (1985) Molecular characterisation of the colicin E2 operon and identification of its products. Mol Gen Genet 198:465–472

Cooper PC, James R (1985) Three immunity types of klebicins which use the cloacin DF13 receptor of *Klebsiella pneumoniae*. J Gen Microbiol 131:2313–2318

De Fernandez Henestrosa AR, Ogi T, Aoyagi S, Chafin D, Hayes JJ, Ohmori H, Woodgate R (2000) Identification of additional genes belonging to the LexA regulon in *Escherichia coli*. Mol Microbiol 35:1560–1572

de Zamaroczy M, Mora L, Lecuyer A, Geli V, Buckingham RH (2001) Cleavage of colicin D is necessary for cell killing and requires the inner membrane peptidase. LepB. Mol Cell 8:159–168

Ferrer S, Viejo MB, Guasch JF, Enfedaque J, Regue M (1996) Genetic evidence for an activator required for induction of colicin-like bacteriocin 28b production in *Serratia marcescens* by DNA-damaging agents. J Bacteriol 178:951–960

Gordon DM, FitzGibbon F (1999) The distribution of enteric bacteria from Australian mammals: host and geographical effects. Microbiology 145:2663–2671

Guasch JF, Enfedaque J, Ferrer S, Gargallo D, Regue M (1995) Bacteriocin 28b, a chromosomally encoded bacteriocin produced by most *Serratia marcescens* biotypes. Res Microbiol 146:477–483

Hendrix RW, Smith MC, Burns RN, Ford ME, Hatfull GF (1999) Evolutionary relationships among diverse bacteriophages and prophages: all the world's a phage. Proc Natl Acad Sci USA 96:2192–2197

Hendrix RW, Hatfull GF, Smith MC (2003) Bacteriophages with tails: chasing their origins and evolution. Res Microbiol 154:253–257

Hofinger C, Karch H, Schmidt H (1998) Structure and function of plasmid pCo1D157 of enterohemorrhagic *Escherichia coli* O157 and its distribution among strains from patients with diarrhea and hemolytic-uremic syndrome. J Clin Microbiol 36:24–29

James R, Schneider J, Cooper PC (1987) Characterization of three group A klebicin plasmids: localization of their E colicin immunity genes. J Gen Microbiol 133:2253–2262

Karch H, Schmidt H, Janetzki-Mittmann C, Scheef J, Kroger M (1999) Shiga toxins even when different are encoded at identical positions in the genomes of related temperate bacteriophages. Mol Gen Genet 262:600–607

Karlin S, Mrazek J, Campbell AM (1998) Codon usages in different gene classes of the *Escherichia coli* genome. Mol Microbiol 29:1341–1355

Korber B (2000) HIV signature and sequence variation analysis. In: Rodrigo AG, Learn GH (eds) Computational analysis of HIV moleoular sequences. Kluwer Academic, Dordrecht, Netherlands, pp 55–72

Lazdunski CJ, Bouveret E, Rigal A, Journet L, Lloubes R, Benedetti H (1998) Colicin import into *Escherichia coli* cells. J Bacteriol 180:4993–5002

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: A laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

Mankovich JA, Hsu CH, Konisky J (1986) DNA and amino acid sequence analysis of structural and immunity genes of colicins Ia and Ib. J Bacteriol 168:228–236

Masaki H, Ogawa T (2002) The modes of action of colicins E5 and D, and related cytotoxic tRNases. Biochimie 84:433–438

Masaki H, Ogawa T, Tomita K, Ueda T, Watanabe K, Uozumi T (1997) Colicin E5 as a new type of cytotoxin, which cleaves a specific group of tRNAs. Nucleic Acids Symp Ser 287–288

Miller ES, Heidelberg JF, Eisen JA, Nelson WC, Durkin AS, Ciecko A, Feldblyum TV, White O, Paulsen IT, Nierman WC, Lee J, Szczypinski B, Fraser CM (2003) Complete genome sequence of the broad-host-range vibriophage KVP40: comparative genomics of a T4-related bacteriophage. J Bacteriol 185:5220–5233

Nasu M (1981) Epidemiological studies of *Serratia marcescens* infections by bacteriocin (marcescin) typing. Tohoku J Exp Med 133:45–52

Nei M, Gojobori T (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol Biol Evol 3:418–426

Parker RC (1986) Mitomycin C-induced bidirectional transcription from the colicin E1 promoter region in plasmid ColE1. Biochim Biophys Acta 868:39–44

Pilsl H, Braun V (1995) Novel colicin 10: assignment of four domains to TonB- and TolC-dependent uptake via the Tsx receptor and to pore formation. Mol Microbiol 16:57–67

Pressler U, Braun V, Wittmann-Liebold B, Benz R (1986) Structural and functional properties of colicin B. J Biol Chem 261:2654–2659

Pugsley AP, Oudega B (1987) Methods for studying colicins and their plasmids. IRL Press, Oxford

Pugsley AP, Schwartz M (1983a) Expression of a gene in a 400-base-pair fragment of colicin plasmid ColE2-P9 is sufficient to cause host cell lysis. J Bacteriol 156:109–114

Pugsley AP, Schwartz M (1983b) A genetic approach to the study of mitomycin-induced lysis of *Escherichia coli* K-12 strains which produce colicin E2. Mol Gen Genet 190:366–372

Rakin A, Boolgakowa E, Heesemann J (1996) Structural and functional organization of the *Yersinia pestis* bacteriocin pesticin gene cluster. Microbiology 142:3415–3424

Riede I, Drexler K, Eschbach ML, Henning U (1986) DNA sequence of the tail fiber genes 37, encoding the receptor recognizing part of the fiber, of bacteriophages T2 and K3. J Mol Biol 191:255–266

Riley MA (1993a) Molecular mechanisms of colicin evolution. Mol Biol Evol 10:1380–1395

Riley MA (1993b) Positive selection for colicin diversity in bacteria. Mol Biol Evol 10:1048–1059

Riley MA (1998) Molecular mechanisms of bacteriocin evolution. Annu Rev Genet 32:255–278

Riley MA, Wertz JE (2002) Bacteriocins: evolution, ecology, and application. Annu Rev Microbiol 56:117–137

Riley MA, Pinou T, Wertz JE, Tan Y, Valletta CM (2001) Molecular characterization of the klebicin B plasmid of *Klebsiella pneumoniae*. Plasmid 45:209–221

Riley MA, Goldstone CM, Wertz JE, Gordon D (2003) A phylogenetic approach to assessing the targets of microbial warfare. J Evol Biol 16:690–697

Roos U, Harkness RE, Braun V (1989) Assembly of colicin genes from a few DNA fragments. Nucleotide sequence of colicin D. Mol Microbiol 3:891–902

Sanger F, Coulson AR, Hong GF, Hill DF, Petersen GB (1982) Nucleotide sequence of bacteriophage lambda DNA. J Mol Biol 162:729–773

Smarda J, Uher P, Osecky P (1988) Modes of action of colicins E4-E7: rates of basic biosyntheses inhibition. Zentralbl Bakteriol Mikrobiol Hyg [A] 269:7–14

Swofford DL (2002) PAUP*: Phylogenetic Analysis Using Parsimony (and Other Methods), version 4.0 Beta

Tagg JR, Dajani AS, Wannamaker LW (1976) Bacteriocins of gram-positive bacteria. Bacteriol Rev 40:722–756

Tan Y, Riley MA (1997) Nucleotide polymorphism in colicin E2 gene clusters: evidence for nonneutral evolution. Mol Biol Evol 14:666–673

Tetart F, Desplats C, Krisch HM (1998) Genome plasticity in the distal tail fiber locus of the T-even bacteriophage: recombination between conserved motifs swaps adhesin specificity. J Mol Biol 282:543–556

Tieze GA, Stouthamer AH, Jansz HS, Zandberg J, van Bruggen EFJ (1969) A bacteriocinogenic factor of *Enterobacter cloacae*. Mol Gen Genet 106:48–65

Tomita K, Ogawa T, Uozumi T, Watanabe K, Masaki H (2000) A cytotoxic ribonuclease which specifically cleaves four isoaccepting arginine tRNAs at their anticodon loops. Proc Natl Acad Sci USA 97:8278–8283

Traub WH, Raymond EA, Startsmar TS (1971) Bacteriocin (Marcescin) typing of clinical isolates of *Serratia marcescens*. Appl Microbiol 21:837–840

van den Elzen PJ, Hakkaart MJ, van Putten AJ, Walters HH, Veltkamp E, Nijkamp HJ (1983a) Structure and regulation of gene expression of a Clo DF13 plasmid DNA region involved in plasmid segregation and incompatibility. Nucleic Acids Res 11:8791–8808

van den Elzen PJ, Maat J, Walters HH, Veltkamp E, Nijkamp HJ (1982) The nucleotide sequence of the bacteriocin promoters of plasmids Clo DF13 and Col E1: role of lexA repressor and cAMP in the regulation of promoter activity. Nucleic Acids Res 10:1913–1928

van den Elzen PJ, Walters HH, Veltkamp E, Nijkamp HJ (1983b) Molecular structure and function of the bacteriocin gene and bacteriocin protein of plasmid Clo DF13. Nucleic Acids Res 11:2465–2477

Viejo MB, Gargallo D, Ferrer S, Enfedaque J, Regue M (1992) Cloning and DNA sequence analysis of a bacteriocin gene of *Serratia marcescens*. J Gen Microbiol 138: 1737–1743

Walker GC (1984) Mutagenesis and inducible responses to deoxyribonucleic acid damage in *Escherichia coli*. Microbiol Rev 48:60–93

Wertz JE, Riley MA (2004) Chimeric Nature of two plasmids of *Hafinia alvei* encoding the bacteriocins alveicins A and B. J Bacteriol 186:1598–1605