ARTICLE

Cell Research (2004); 14(3): 208-216
http://www.cell-research.com

Characterization of a novel toxin-antitoxin module, VapBC, encoded by *Leptospira interrogans* chromosome

Yi Xuan ZHANG[1, 2], Xiao Kui GUO[3], Chuan WU[2], Bo BI[1], Shuang Xi REN[4], Chun Fu WU[1], Guo Ping ZHAO[2, 4*]

[1]*Pharmaceutical Department, Shenyang Pharmaceutical University, 103 Wenhua Road, Shenhe District, Shenyang 110016,China.*
[2]*Research Center of Biotechnology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, 500 Caobao Road, Shanghai 200233, China.*
[3]*Department of Microbiology and Parasitology, Shanghai Second Medical University, 280 Chongqingnan Road, Shanghai 200025, China.*
[4]*Chinese National Human Genome Center at Shanghai (CHGCS), 250 Bibo Road, ZhangJiang High Tech Park, Shanghai, 201203, China.*

ABSTRACT

Comparative genomic analysis of the coding sequences (CDSs) of *Leptospira interrogans* revealed a pair of closely linked genes homologous to the *vapBC* loci of many other bacteria with respect to both deduced amino acid sequences and operon organizations. Expression of single *vapC* gene in *Escherichia coli* resulted in inhibition of bacterial growth, whereas co-expression of *vapBC* restored the growth effectively. This phenotype is typical for three other characterized toxin-antitoxin systems of bacteria, *i.e.*, *mazEF*[1], *relBE*[2] and *chpIK*[3]. The VapC proteins of bacteria and a thermophilic archeae, *Solfolobus tokodaii*, form a structurally distinguished group of toxin different from the other known toxins of bacteria. Phylogenetic analysis of both toxins and antitoxins of all categories indicated that although toxins were evolved from divergent sources and may or may not follow their speciation paths (as indicated by their 16s RNA sequences), co-evolution with their antitoxins was obvious.

Keywords: *Leptospria,* toxin-antitoxin system, VapBC, co-evolution, speciation.

INTRODUCTION

Leptospira interrogans is the etiologic agent of leptospirosis, which is a worldwide zoonosis with a much greater incidence in tropical and subtropical regions[4]. The recently reported genomic sequence of the *L. interrogans* serogroup Icterohaemorrhagiae serovar lai [5] indicated that in contrast to the two strictly parasitic spirochetes, *Treponema pallidum*[6] and *Borrelia burg-dorferi* [7], *L. interrogans* possessed much more physiologically important genes for its facultative free-living metabolism, allowing it to adapt to its diverse environment. However, the regulation of the physiology is largely unknown.

Toxin-antitoxin (TA) modules are of importance among bacterial regulatory systems[8]. They were firstly found on plasmids and considered prevalently to be associated with stable plasmid inheritance at cell division[9-11]. The well-known plasmid-encoded TA loci contained CcdA/CcdB of F factors[12, 13], Kis/Kid of R1[14], Phd/Doc of P1[15], ParD/ParE of RK2[16], *pas* locus of pTF-CF2 [17], ω-ε-ξ operon of pSM19035[18], *stb* locus of pMYSH6000[19] and *relBE* locus of P307[20].

Recently, TA modules were also identified on chromosomes and considered to be associated with host stress response[21]. Only three TA loci families encoded by chromosomes have been identified so far, *relBE*[2], *mazEF* [1], and *chpIK*[3]. Generally, TA loci are organized into operons where the first cistron encodes a small instable antitoxin while the second cistron encodes a large stable antitoxin[21].

The presence of a TA module in *L. interrogans* was firstly reported by Picardeau *et al*[3] by characterizing the large chromosome (CI) encoded *chp* locus (*chpI-chpK*). It was proposed that the presence of this system might provide means for bacterial adaptation to a poor nutrition environment. Further analysis of the *L. interrogans* genome found a *mazEF* locus (LA1780/LA1781) belonging to the TA system. When a pair of *vapBC* (virulence-associated protein) genes were cloned and expressed in *Escherichia coli* for functional analysis, we noticed that the characteristics of these genes and their products were similar to that of the TA systems, but unlikely to be associated

*Correspondence: Guo Ping ZHAO, Tel: +86-21-50801919;
Fax: +86-21-50801922; E-mail: gpzhao@sibs.ac.cn.

Tab 1. *Leptospira* strains used in this study

Strains	Series code*	Source	Virulence
L. interrogans strains:			
serogroup Icterheamorrhagiae serovar lai	56601	CCDC	+
serogroup Ballum serovar ballum	56604	CCDC	+
serogroup Pyrogenes serovar pyrogenes	56605	CCDC	+
serogroup Autumnalis serovar autumnalis	56606	CCDC	+
serogroup Pomona serovar pomona	56608	CCDC	+
serogroup Grripotypphosa serovar lin 6	56609	CCDC	+
serogroup Hebdomadis serovar hebdomadis	56610	CCDC	+
serogroup Icterheamorrhagiae serovar lai	——	Inst. Pasteur	–
L. biflexa strains:			
serogroup Semaranga serovar monvalerio	57001	CCDC	–
serogroup Semaranga serovar patoc	651505	CCDC	–

*Code numbers of Chinese Center for Disease Control and Prevention

Tab 2. Oligonucleotide primers specific for *vapB* and *vapC* used in the PCR and RT-PCR reactions

Genes	Reactions	Direction	Primer sequence*	Length (bp)
vapB	PCR	Sense	5'CGGAGATAA**CATATG**CAAACAGCCAAATT 3'	351
		Antisense	5'ACTCA**GGATCC**GCAACAGTTAAAGACGAAACAAA3'	
vapC	PCR	Sense	5'AAAAGCT**CATATG**TATCTTTTGGATA3'	443
		Antisense	5'CGC**GGATCC**GAAGTTAGTTGTAGTGG3'	
vapBC	PCR	Sense	5'CGGAGATAA**CATATG**CAAACAGCCAAATT3'	674
		Antisense	5'CGC**GGATCC**GAAGTTAGTTGTAGTGG3'	
vapBC	RT-PCR	Sense	5'AGTCAAGCGGTTCGACTACCA 3'	487
		Antisense	5'GTCTATTGAACCGATTACATTACCT 3'	

*The designed restriction sites incorporated into the primers, *Nde*I for the sense sequences and *Bam*HI for the antisense primers are bolded and underlined.

with virulence. Based on their sequence homology and operon organization, a novel group of TA system was characterized.

MATERIALS AND METHODS

Materials

Plasmid pET28b (+) from Novagen (Madison, Wisconsin, USA) was used for heterogeneous gene expression. Golden Taq DNA polymerase, BigDye Terminator Cycle Sequencing Ready Reaction Kit, restriction enzymes, T4 DNA ligase, and bacterial alkaline phosphatase were obtained from Takara (China). Trizol reagent and Freund's complete and incomplete adjuvant was obtained from GBICO-BRL (USA). Vector plasmid pUCm-T, AMV Reverse Transcription System, goat anti-rabbit immuno-globulin G (Fc)-Alkaline Phosphatase conjugated and BCIP/NBT color development substrate were purchased from Promega (USA). Nitrocellulose membranes were obtained from Amersham (USA). All other reagents were of the highest purity available.

Bacterial strains and culture conditions

Leptospira strains used in this work are shown in Tab 1. The avirulent *L. interrogans* strain of serovar lai was kindly supplied by Prof. Saint Girons Unité de Bactériologie Moléculaire et Médicale, Institut Pasteur, France. All of the other strains were maintained by the Institute for Infectious Disease Control and Prevention (IIDC), Chinese Center for Disease Control and Prevention (CCDC), Beijing, China. The strains were cultured in EMJH(Johnson and Marris modification of the Ellinghausen and McCullough medium[4]) at 28°C for 3-4 d under aerobic conditions until mid-log phase was reached (about 10^8 bacteria/ml). Bacterial counts were determined by dark-field microscopy. *E. coli* strains DH5α and BL21(DE3/pLysS) used as hosts for cloning or expression were cultured in LB (Luria-Bertani) medium.

Databases and softwares

The complete genomic sequence of the virulent *L. interrogans* serogroup Icterohaemorrhagiae serovar lai was obtained from http://www.chgc.sh.cn/lep/. BLAST searching the NCBI non-redundant

Tab 3. Vectors used in this study and plasmids with cloned genes

Plasmids	Construction	Source
pUB	ApUC19 derivative carrying the *vapB* gene	This work
pUC	ApUC19 derivative carrying the *vapC* gen	This work
pUBC	A pUC19 derivative carrying the *vapBC* gene	This work
pET28b (+)	Expression vector with strong T7 promotor, kan	Novagen (USA)
pTB	A pET28b derivative carrying the *vapB* gene, kan^r	This work
pTC	A pET28b derivative carrying the *vapC* gene, kan^r	This work
pTBC	A pET28b derivative carrying the *vapBC* gene, kan^r	This work

nucleotide database, or non-redundant protein database on SwissProt/TrEMBL were used to annotate the *L. interrogans vapBC* and *mazEF* genes. Bacterial protein sequences of *mazEF, vapBC, chpIK* and *relBE* were obtained from the genome database of KEGG by employing the TFASTA search (http://fasta.genome.ad.jp/ideas/fasta/tfasta_genome.html) with inputting VapB, VapC, MazE, MazF, ChpI and ChpK protein sequences of *L. interrogans*. Only those with similarity scores above e-8 were selected for further analysis. Multiple-sequence alignment was accomplished by using Bioedit. Phylogenetic trees were constructed using Mega 2.0.

Expressing *vapB* and *vapC* genes in *E. coli* and protein purification

PCR primers for *L. interrogans vapB*, *vapC* and *vapBC* genes were designed as shown in Tab 2. Forward primers incorporated an *Nde*I restriction site and reverse primers incorporated a *Bam*HI restriction site. The reverse primers were designed to be immediately behind the terminator "TTA" to ensure involvement of the terminator in the PCR product. All the genes were amplified using *Leptospira* spp. genomic DNA as templates, which were extracted from mid-log phase cultures of corresponding strains grown in EMJH medium[22]. PCR was performed using the following cycle; 95°C for 10 min, followed by 30 cycles at 94°C for 30 sec for denaturing, 54°C for 30 sec for *vapB* or 58°C for 1 min for *vapC* and *vapBC* for annealing, and 72°C for 1 min for reaction. The whole reaction was completed at 72°C for 10 min for elongation followed by 4°C for storage. Procedures for gene cloning, heterogeneous expression and protein purification were manipulated according to the manufacturer's manual. Plasmids constructed for this work are listed in Tab 3.

Total RNA extraction, reverse transcription-PCR (RT-PCR)

Total RNA was extracted from *Leptospira* strains cultured in EMJH medium for 24, 48 (log-phase), or144 h using Trizol Reagent (GBICO-BRL). The AMV Reverse Transcription System (Promega) was used for reverse transcription experiments. RT-PCR primers were designed to simultaneously detect the 3' portion of the *vapB* gene together with the 5' portion of the corresponding *vapC* gene (as shown in Tab 2). Reverse transcription synthesized first-strand cDNA was used as the PCR template. The amount of template cDNA used was 10-15 ng per assay. The PCR protocol was as described above, except for the annealing step which was done at 54°C for 1 min.

Bacterial growth curve detection

Growth curves of the *E. coli* strains BL21(DE3/pLysS) containing pTB, pTC, pTBC were determined by measuring the $O.D_{600nm}$. Cells were initially grown in LB overnight with appropriate antibiotics (kanamycin, 50 μg/ml) and then, transferred into fresh LB at 1% inoculums in the presence of IPTG (1 mM).

Plasmid stability test

Single colonies of *E. coli* DH5α carrying plasmid pUC19, pUB, pUC, or pUBC were picked from LB plates containing ampicillin (50 μg/ml) and innoculated into 5 ml of LB without ampicillin. After shaking at 37°C, 220 rpm for 12 h, cells were reinnoculated into 5 ml of fresh LB without ampicillin, but containing 1 mM IPTG (1% innoculum) for a further 12 h. This kind of reinnoculation and cultivation was continued for up to 336 h. 100 μl samples were collected every 24 h, diluted and plated onto LB-plates, with or without ampicillin (50 μg/ml). Viable colonies were counted after 12 h incubation at 37°C. The ratio of the Amp^R colonies versus the total viable colony counts was used to estimate the percentage of plasmid maintained in the population.

RESULTS

The expression level of *vapC* in *E. coli* without or with the co-expression of *vapB* differed significantly

Among the 4,769 predicted genes of *L. interrogans* (based on its genomic sequence – http://www.chgc.sh.cn/lep/), a pair of closely linked CDS (LA1002/LA1001) was shown to be strikingly similar to the bacterial virulence association proteins (VapB/VapC) [23], as well as a plasmid maintenance stability determinant locus, (*stb*) from Shigella flexneri[19] .

The leptospira *vapBC* locus consists of an upstream 231 bp CDS (*vapB*) that encodes a putative protein of 77 amino acids and a downstream 399 bp CDS (*vapC*) that encodes a putative protein of 133 amino acids (Fig 1). The translational start codon for *vapC* overlaps with the last base of the translational stop codon of *vapB*, which is a strong indication of translational coupling. This *vapBC* operon structure is strikingly similar to that of the three known TA operons, *relBE*[1, 24], *mazEF*[1] and *chp*[3, 25]. A putative promoter region was revealed by the presence

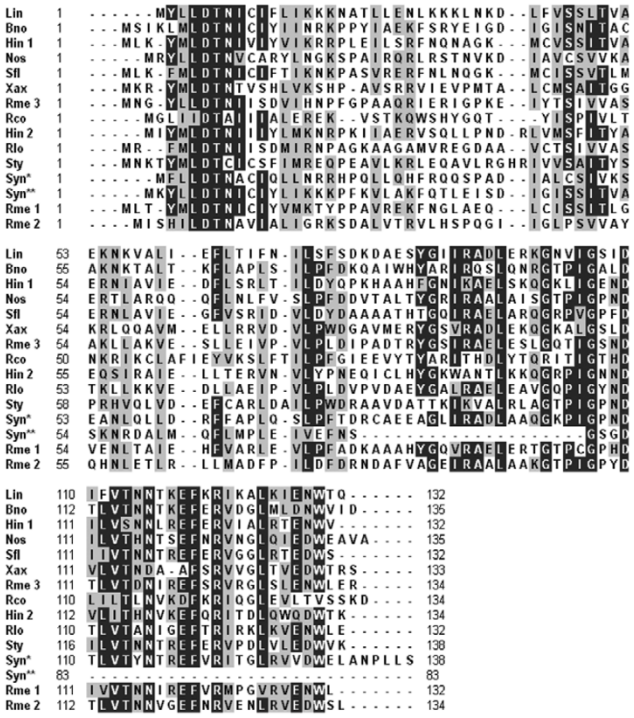

Fig 1. Multiple-sequence alignment of 18 VapC-homolog proteins from *L. interrogans* and other bacteria with similarity scores above e-8. Only the VapC proteins with obvious upstream VapB partners were included in the alignment analysis. Identical amino acids are shown in black, while conserved amino acids are shown in grey. The alignment was accomplished using Bioedit software. Gaps introduced to maximize the similarity alignment are indicated by dashes. The bacterial abbreviations used are listed in Tab 4.

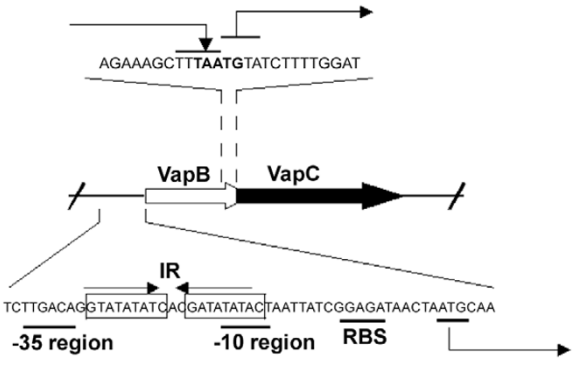

Fig 2. Schematic representation of the putative promoter region of the chromosome-encoded *vapBC* of L. *interrogans*. Single arrows show all of the translation directions. The overlap of the stop codon of *vapB* with the start codon of *vapC* is enlarged. The 9-bp inverted repeats (IRs) are boxed and highlighted by two head-to-head arrows. Typical -10 and -35 regions as well as the ribosome-binding site (RBS) are marked and the *vapB* start coden is bolded.

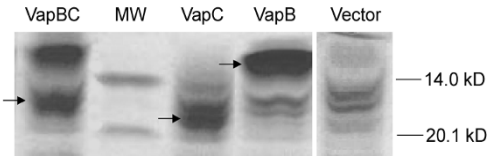

Fig 3. Proteins encoded by *vapB* and *vapC* clones expressed in *E. coli* BL21 (DE3/pLysS). The proteins were expressed as described in Materials and Methods. The arrows indicate each expressed His-tag recombinant protein.

of a pair of 9 bp inverted repeats (IR). This shares the same characteristics with IRs of the *E. coli chp* promoter, known for specific DNA binding of the Chp/Pem protein [26]. Both the -10 and -35 regions for the promoter and ribosomal binding site (RBS) are typical for bacteria (Fig 2).

In order to characterize this pair of predicted proteins, we cloned the *vapB* and *vapC* genes into pET expression vectors, either individually or together (Tab 3) and attempted to express the genes in *E. coli* BL21(DE3 pLysS) (Novagen). The recombinant proteins were expressed after IPTG induction and shown to have a molecular weight (MW) of 10.7 kD for VapB with his-tag and 17.1 kD for VapC with his-tag (Fig 3). The expression level of VapC was significantly lower than that of its partner VapB if expressed alone. However, when these two proteins were expressed simultaneously and coordinately from a construct containing their original operon with the characteristic coupled cistronic structure (Fig 2), the expression levels of both were high and essentially the same (Fig 3).

The growth of the *E. coli* strain carrying the *vapC* expression plasmid was restrained but was relieved with the co-expression of the *vapB* gene

The growth of the *E. coli* carrying plasmid pTC (*vapC* expression plasmid) sharply decreased in the presence of IPTG compared to that of the hosts with plasmids pTB (*vapB* expression plasmid), pTBC (*vapBC* coupled expression plasmid) or pET28b (the negative control vector) (Fig 4). Quantitatively, there was no difference in their doubling times (44 min for pTC, 47 min for pTB, 45 min for pTBC and 46 min for pET28b). However, the difference was significant for the length of the lag phase (248 min for pTC but 28 min for pTB, 40 min for pTBC, and 34 min for pET28b). This difference directly leads to the difference in the speed of accumulation of biomass. At the 14^{th} h of cultivation, the average OD_{600} reading of *E. coli* harboring pTC (1.15) was much lower than that of all the other 3 strains (1.43 for pTB, 1.49 for pTBC and 1.51 for pET28b).

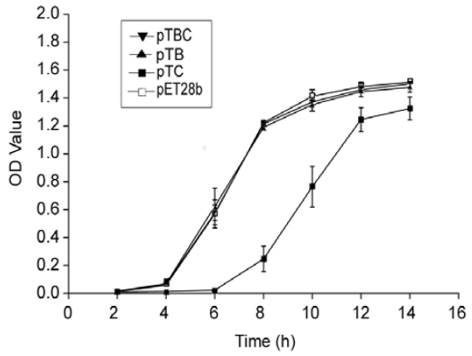

Fig 4. Effect of the over-expression of the *L. interrogans vapC* in the presence or absence of *vapB* on growth of *E. coli* cells harboring the expression plasmids. Symbols used for the plasmids transformed to the *E. coli* hosts: pET28b (□), pTB (▲), pTC (■), pTBC (▼). Growth was determined by measuring the optical density at 600 nm.

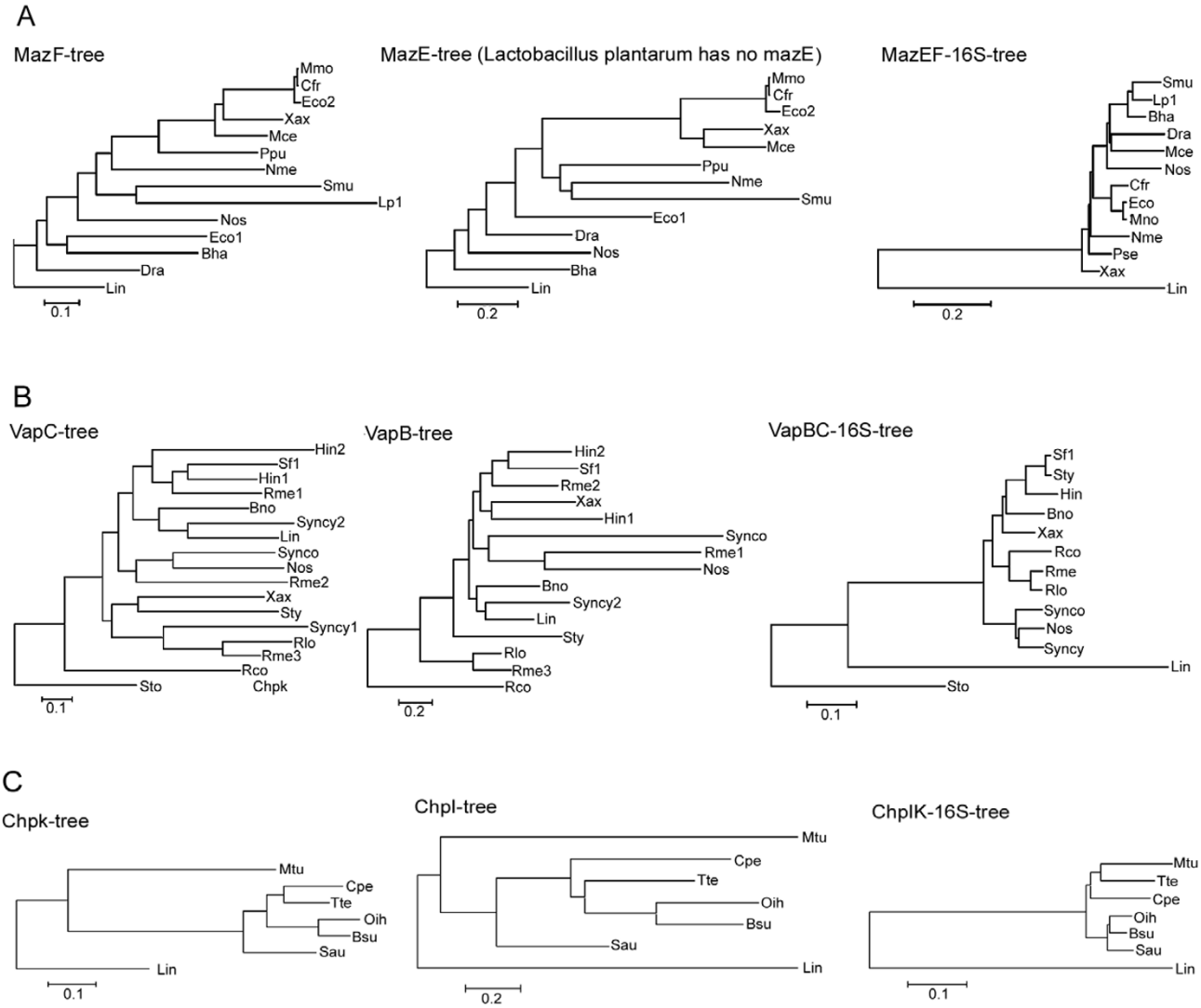

Fig 5. Unrooted phylogenetic trees for the amino acid sequences of bacterial antitoxin-toxin MazEF (panel A), VapBC (panel B) and ChpIK (panel C) and 16s rRNA sequences corresponding to each kind of TA system. In panel B, *S. tokodaii* was drawn as the root in the VapC tree because this archeae strain is the most primitive among all the bacteria in this group. As *S. tokodaii* does not have a *vapB* gene, *R. conorii*, the strain closet to *S. tokodaii* in the tree for VapC, was drawn as the root in the VapB tree. All the other trees were unrooted. The abbreviations for the bacterial species are listed in Tab 4.

Tab 4. The chromosome-encoded TA systems distributed in microorganisms

Abbreviation	*Bacterial species/strains*	*vapB*	*vapC*	*mazE*	*mazF*	*relB*	*relE*	*chpK*	*chpI*
Lin	*Leptospira interrogans*	LA1002	LA1001	LA1780	LA1781			LA2843	LA2844
Eco	*Escherichia coli*			P18534	P33645	P07007	P07008		
				P13975	P13976	Q8VU39	Q8VU40		
Sfl	*Shigella flexneri*	O06663	O06662						
Hin	*Haemophilus influenzae*	Q57534	P71363						
		Q57120	Q57122						
Rme	*Rhizobium meliloti*	Q92KY4	Q92KY5			NO	Q92Q75		
		Q92ME1	Q92ME0						
		Q9L381	Q9L380						
Xax	*Xanthomonas axonopodis*	Q8PRH2	Q8PRH3	Q8PRN0	Q8PRN1				
Sty	*Salmonella typhimurium*	Q8L236	Q8L237			Q8ZPG0	Q8ZPF9		
						Q8XES7	Q8XGL8		
Rlo	*Rhizobium loti*	Q98K21	Q98K22						
Rco	*Rickettsia conorii*	Q92IP0	Q92IP1						
Bno	*Bacteroides nodosus*	Q46558	Q46557						
Synco	*Synechococcus sp.*	Q9Z3G8	Q9Z3G9						
Nos	*Nostoc sp.*	Q8YK90	Q8YK91	Q8YS79	Q8YS80				
Syncy	*Synechocystis sp.*	NO	Q55189						
		P72994	P72993						
Bsu	*Bacillus subtilis*							P96622	P96621
Dra	*Deinococcus radiodurans*			Q9RX99	Q9RX98				
Nme	*Neisseria meningitides*			Q9JZS7	Q9JZS8				
Ppu	*Pseudomonas putida*			AAN66395	AAN66396				
Smu	*Streptococcus mutans*			Q8DW96	Q8DW95				
	UA159								
Mce	*Mycobacterium celatum*			Q93S65	Q93S64				
Mtu	*Mycobacterium tuberculosis*							Q10867	Q10868
Cfr	*Citrobacter freundi*			Q8GFY0	Q8GFY1				
Mmo	*Morganella morganii*			O52204	O52205	Q9ZH43	Q9ZH42		
Lpl	*Lactobacillus plantarum*			NO	gp|AL935260				
Oih	*Oceanobacillus iheyensis*							BAC12579	Q8ESW4
Tte	*Thermoanaerobacter*							Q8R861	Q8R5Q2
	tengcongens								
Cpe	*Clostridium perfringens*							Q8XNN7	Q8XNN8
Sau	*Staphylococcus aureus*							O05341	Q99SI6
Asa	*Aeromonas salmonicida*					CAD48430	Q8GMN9		
Vch	*Vibrio cholerae*					Q08800	Q08799		
Bha	*Bacillus halodurans*			Q9K6K9	Q9K6K8				
Sto	*Solfolobus tokodaii*	NO	Q96Z88						

The protein sequences codes are from Swiss-Prot or TrEMBL.

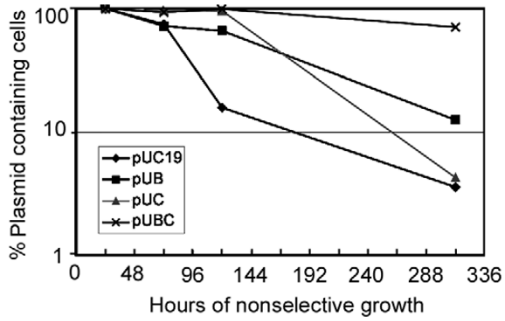

Fig 6. Genetic stabilization effect of the *vapBC* system upon the high-copy unstable plasmid, pUC19. Plasmids used were derivatives of pUC19: pUB (carrying *vapB*), pUC (carrying *vapC*), pUBC (carrying *vapBC*) and pUC19 (blank control). Plasmid construction and experimental details are described in Materials and Methods.

The *vapBC* locus is conserved in different species or serovars of *Leptospira*

The presence of the *vapBC* locus and the expression of this operon was tested in various kinds of pathogenic or saprophytic leptospires (Tab 1). PCR amplification of *vapBC* genes from genomic DNA templates of the experimental strains showed that both the 231 bp *vapB* fragment and the 399 bp *vapC* were present in all the strains tested (data not shown). Subsequent DNA sequencing results showed that the strains tested had identical gene fragments in *vapB* and *vapC* (data not shown). RT-PCR experiments employing total RNA isolated from cells grown on EMJH medium with gene specific primers (see Materials and Methods) showed that these two closely linked genes were transcribed simultaneously in all the *Leptospira* strains tested (data not shown).

The *L. interrogans vapBC* system is effective in stabilizing pUC plasmid in *E. coli*

The stability of a series of pUC19 derived plasmids containing *vapB* (pUB), *vapC* (pUC), or *vapBC* (pUBC) genes with the cloned genes expressed by the *lac* promotor of pUC19 were measured in *E. coli* DH5α (see Materials and Methods). As shown in Fig 6, the percentage of ampicilin resistant cells in the population of *E. coli* carrying pUC19 decreased. On the other hand, the majority of *E. coli* cells carrying pUBC with *vapBC* were relatively stable after prelonged incubation without any selection pressure. Neither pUC nor pUB was stable under the same conditions, although the rate of plasmid lost was slightly improved.

DISCUSSION

The *vap* module was first found in the genome of the strictly anaerobic bacterium, *Dichelobacter nodosus* (*Bacteroides nodosus*)[18], a major pathogen implicated in foot rot (a highly contagious disease of sheep). Hybridization experiments showed that the *vap* module was present in all the genomes of virulent isolates, while it was absent in 67% of the benign isolates. Therefore, it was named the virulent-associated region and the hypothetical proteins encoded in this region were designated as virulent-associated proteins (*vap*), although their definitive functions were unknown.

Subsequently, many *vapBC* loci were found in other bacteria (Fig 1, Tab 4), encoded by either chromosomes or plasmids and in particular, the *stb* locus encoded on the large virulent plasmid pMYSH6000 of *S. flexneri,* are both structurally and functionally similar to the *vap* system. By studying the physiology of the overexpression of *L. interrogans vapBC* genes in *E. coli*, we recognized that, similar to the case of *L. interrogans chpIK* operon[3], *vapBC* is a novel type of toxin-antitoxin system. In contrast to *B. nodosus*, the *vapBC* operon is widely present and expressed in leptospires. Therefore, it is more likely related to physiological regulation rather than to the virulence of the bacterium.

Deduced amino acid sequence alignment of the VapC proteins indicated that besides the general similarity among the family members, highly conserved specific sites were observed (Fig 1). All the VapC proteins contain a common PIN domain (IPR002716, InterPro) putatively responsible for nucleic acid binding, which might infer the possibility of a common cellular target for the Vap TA system.

Further similarity searches and comparisons in the microbial genomic database with respect to TA systems revealed a total of four types encoded by bacterial chromosomes. They are the *relBE, mazEF, chpIK,* and *vapBC* loci (Tab 4). Except for the *chpIK* locus, all the others are also found in plasmids (Tab 4).

It is interesting to note that several TA systems, or multiple copies of one TA system, may present in one genome (Tab 4). In this connection, *L. interrogans* stands out as the only bacterium that possesses three TA systems, the chpIK previously identified[3], the *vapBC* characterized in this study, and the *mazEF* (LA1780/LA1781) recognized by homology searches. It is known[21] that TA systems encoded by plasmids prevent plasmid loss at cell division, whereas TA loci found on chromosomes are associated with species stability or cell programmed death (bacterial apoptosis). Further studies have revealed the specific molecular targets of some of the toxins, *i.e.*, CcdB inhibits DNA gyrase, while PemK/Kid inhibits DNA replication, presumably *via* interaction with the DnaB helicase[27]. The RelE proteins encoded by the *E. coli* K-12 chromosome presumably inhibit translation[2, 20], but

their specific target(s) within the translational machinery is not yet known. Recent studies have shown that the ChpK proteins of *E.coli* inhibit cell growth by cleavage of mRNA [28]. Similarly, RelE toxin inhibits protein synthesis by cleavage of ribosome bound mRNA in response to nutritional stress[29]. These findings further support the hypothesis proposed by Gerdes[21]. Toxins from the TA gene families, the RelE, VapC, MazF and ChpK, may mediate inhibition of translation or DNA replication[21] as part of the global cellular response to environmental stress, rather than being simply cell-killing. The presence of multiple TA systems in one particular genome, as shown in the case of *L. interrogans*, may imply the complexity of regulation during the transition of metabolism.

Co-evolution is usually represented by a change in genetic composition of one gene in response to the genetic change in another[30]. The translational coupling of two closely linked genes with a strongly interactive functional relationship to the proteins they encoded, such as the case of TA system, seems to enforce the strongest selection pressure for co-evolution. Amino acid sequence similarity and clustering analysis of each group of the TA system were carried out and the results were compared to their corresponding 16s RNA phylogenetic trees (Fig 5). In general, the co-evolution between toxins and their corresponding antitoxins is obvious, while the development of these TA systems along with their speciation route varies from case to case. It is also interesting to note that homologous analysis could spontaneously put the toxin sequences into four groups, but not for their antitoxins (data not shown). These observations indicate that quite a few of the TA systems did not originally exist in the species, *i.e.*, horizontal gene transfer might account for significant portions of the event. Therefore, co-evolution of the toxin genes, as well as the antitoxin genes with the 16s RNA, is only observed in part of the species as individual groups, but not along the whole family. This phenomenon is clearly illustrated in the case of the MazEF family (Fig 5A) where most of the Gram-negative bacteria except *Nostoc*, are grouped together as that of their speciation. However, Gram-positive bacteria are grouped separately into different Gram-negative bacterial clusters. Furthermore, even for each TA system, the toxin and antitoxin may not have originated from the same ancestor, or even been acquired simultaneously. This phenomenon is particularly obvious in case where multiple copies of one TA system existed in one strain. For instance, *Rhizobium meliloti* has two sets of VapBC systems encoded by its genome, but the phylogenetic position VapB1 only matches with that of VapC2 and vice versa. This implies not only an independent origination, but also a possible trans-acting mechanism between two TA modules of the same category (Fig 5B). In addition, it seems that if the evolving of a TA system among bacteria was a relatively homogeneous process, *e.g.*, without gene duplication so that only a single copy of a TA module existed in one species (as is the case of *chpIK*) parallel evolution between the TA module and the 16s RNA could be observed (Fig 5C). A similar situation was observed in the family of RelBE, which consists of Gram-negative bacteria only (data not shown). This fact implies that the evolvement of *relBE* module might be late in the evolution history and it might have no chance to be spread into bacteria other than the Gram-negative ones. Finally, we noticed that in certain cases, even the strong co-evolutionary relationship between toxin and antitoxin was distorted. For instance, antitoxin was not found in the only archeae that possessed a toxin gene (*vapC*), *Solfolobus tokodaii*. This is interesting because it suggests that the function of the toxin-like proteins found in archeae might not necessarily be the same as originally thought. Alternatively, it is equally possible that although these proteins are toxic to the host, they might not be functionally expressed under the physiological conditions of these hosts. In either case, it suggests that antitoxins are likely introduced into the system later in the evolutionary path, and that the genetic signature of this "double evolving" event is still embedded in the coding sequence that is seen today.

CONCLUSIONS

Our data have demonstrated that the *vapBC* locus of *L. interrogans* belongs to the TA family rather than encoding virulence-associated proteins. The product of the *vapC* gene is toxic to *E. coli* cells, while the product of the *vapB* gene counteracts VapC toxicity. The presence of the *vapBC* operon on an unstable plasmid under the control of a *lac* promoter can effectively prevent plasmid loss under non-selective conditions. A pair of 9 bp IRs in the promoter region of the *L. interrogans vapBC* operon share the same characterization with IRs of the *E. coli chpB* promoter region, indicating it might be a specific DNA binding site for VapC proteins. At present, the cellular target of VapC protein remains unknown. Further study on *L. interrogans* VapB and VapC could help in the understanding of the cellular target of these toxins and their physiological role in bacteria.

ACKNOWLEDGEMENTS

We thank Xiu Gao JIANG (The Institute of Epidemiology and Microbiology, Chinese Academy of Preventive Medicine) and Bao Yu HU (Department of Microbiology and Parasitology, Shanghai Second Medical University) for kindly providing *Leptospira* strains and cultures. We also thank You Gang MIAO (Shanghai Information Center for Life Sciences, Chinese Academy of Sciences), Hai XU

(Institute of Plant Physiology & Ecology, Chinese Academy of Sciences), Hui Feng XI and Lu Ling FENG (CHGCS) for their support in the bioinformatics analysis. This work was supported by the National High Technology Research and Development Program of China (Program No. 2003AA223031).

Received, Oct 21, 2003
Revised, Apr 25, 2004
Accepted, Apr 28, 2004

REFERENCES

1 Aizenman E, Engelberg-Kulka H, Glaser G. An *Escherichia coli* chromosomal "addiction module" regulated by guanosine 3', 5'-bispyrophosphate: a model for programmed bacterial cell death. Proc Natl Acad Sci 1996; **93**:6059-63.
2 Gotfredsen M, Gerdes K. The Escherichia coli relBE genes belong to a new toxin-antitoxin gene family. Mol Microbiol 1998; **29**:1065-76.
3 Picardeau M, Ren S, Saint Girons I. Killing effect and antitoxic activity of the Leptospira interrogans toxin-antitoxin system in Escherichia coli. J Bacteriol 2001; **183**:6494-7.
4 Faine S, Adler B, Bolin C, Perolat P. *Leptospira and Leptospirosis.* 1999; Medisci Melbourne.
5 Ren SX, Fu G, Jiang XG, et al. Unique physiological and pathogenic features of Leptospira interrogans revealed by whole-genome sequencing. Nature 2003; **422**:888-93.
6 Fraser CM, Norris SJ, Weinstock GM et al. Complete genome sequence of *Treponema pallidum*, the syphilis spirochete. Science 1998; **281**:375-88.
7 Fraser CM, Casjens S, Huang WM et al. Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi.* Nature 1997; **390**:580-6.
8 Gerdes K. Toxin-antitoxin modules may regulate synthesis of macromolecules during nutritional stress. J Bacteriol 2000; **182**: 561-72.
9 Holcik M. Lyer VN. Conditionally lethal genes associated with bacterial plasmids. Microbiology 1997; **143**:3403-16.
10 Jensen RB, Gerdes K. Programmed cell death in bacteria: proteic killer gene systems. Mol Microbiol 1995; **17**:205-10.
11 Kobayashi I. Selfishness and death: raison d'etre of restriction, recombination and mitochondria. Trends Genet 1998; **14**:368-74.
12 Bahassi EM, O'Dea MH, Allali N, Messens J, Gellert M, Couturier M. Interactions of CcdB with DNA gyrase. Inactivation of GyrA, poisoning of the gyrase-DNA complex, and the antidote action of CcdA. J Biol Chem 1999; **274**:10936-44.
13 Tam JE, Kline BC. The F plasmid *ccd* autorepressor is a complex of CcdA and CcdB proteins. Mol Gen Genet 1989; **219**:26-32.
14 Ruiz-Echevarria MJ, Torre MA, Diaz-Orejas RA mutation that decreases the efficiency of plasmid R1 replication leads to the activation of *parD*, a killer stability system of the plasmid. FEMS Microbiol Lett 1995; **130**:129-5.
15 Gazit E, Sauer RT. The Doc toxin and Phd antidote proteins of the bacteriophage P1 plasmid addiction system form a heterotrimeric complex. J Biol Chem 1999; **274**:16813-8.
16 Johnson EP, Ström AR, Helinski DR. Plasmid RK2 toxin protein ParE: purification and interaction with the ParD antitoxin protein. J Bacteriol 1996; **178**:1420-9.
17 Smith AS, Rawlings DE. The poison-antidote stability system of the broad-host-range *Thiobacillus ferrooxidans* plasmid pTF-FC2. Mol Microbiol 1997; **26**:961-970.
18 Ceglowski P, Boitsov A, Chai S, Alonso JC. Analysis of the stabilization system of pSM19035-derived plasmid pBT233 in Bacillus subtilis. Gene 1993; **136**:1-12.
19 Radnedge L, Davis MA, Youngren B, Austin SJ. Plasmid maintenance functions of the large virulence plasmid of *Shigella flexneri.* J Bacteriol 1997; **179**:3670-5.
20 Grønlund H, Gerdes K. Toxin-antitoxin systems homologous to *relBE* of *Escherichia coli* plasmid P307 are ubiquitous in prokaryotes. J Mol Biol 1999; **285**:1401-15.
21 Kenn Gerdes. Toxin-Antitoxin Modules May Regulate Synthesis of Macromolecules during Nutritional Stress. J Bacteriol 2000; **182**:561-72.
22 Segers RP, van der Drift A, de Nijs A, Corcione P et al. Molecular analysis of a sphingomyelinase C gene from *Leptospira interrogans* serovar hardjo. Infect Immun 1990; **58**:2177-85.
23 Katz ME, Howarth PM, Yong WK et al. Identification of three gene regions associated with virulence in Dichelobacter nodosus, the causative agent of ovine footrot. J Gen Microbiol 1991; **137**:2117-24.
24 Gotfredsen M, Gerdes K. The *Escherichia coli relBE* genes belong to a new toxin-antitoxin gene family. Mol Microbiol 1998; **29**:1065-76.
25 Masuda Y, Miyakawa K, Nishimura Y et al. *chpA* and *chpB*, *Escherichia coli* chromosomal homologs of the *pem* locus responsible for stable maintenance of plasmid R100. J Bacteriol 1993; **175**:6850-6.
26 Bravo A, de Torrontegui G, Diaz R. Identification of components of a new stability system of plasmid R1, ParD, that is close to the origin of replication of this plasmid. Mol Gen Genet 1987; **210**:101-10.
27 Ruiz-Echevarria MJ, Gimenez-Gallego G, Sabariegos-Jareno R, Diaz-Orejas R. Kid, a small protein of the *parD* stability system of plasmid R1, is an inhibitor of DNA replication acting at the initiation of DNA synthesis. J Mol Biol 1995; **247**:568-77.
28 Christensen SK, Pedersen K, Hansen FG et al. Toxin-antitoxin loci as stress-response-elements: ChpAK/MazF and ChpBK cleave translated RNAs and are counteracted by tmRNA. J Mol Biol 2003; **332**:809-19.
29 Pedersen K, Zavialov AV, Pavlov MY et al. The bacterial toxin RelE displays codon-specific cleavage of mRNAs in the ribosomal A site. Cell 2003; **112**:131-40.
30 Pal C, Hurst LD. Evidence for co-evolution of gene order and recombination rate. Nat Genet 2003; **33**:392-5.