Complete DNA sequence and gene analysis of the virulence plasmid pCP301 of *Shigella flexneri* 2a

ZHANG Jiyu (张继瑜)[1,2], LIU Hong (刘 红)[1], ZHANG Xiaobing (张笑兵)[1], YANG Jian (杨 剑)[3], YANG Fan (杨 帆)[1], YANG Guowei (杨国威)[1], SHEN Yan (沈 岩)[3], HOU Yunde (侯云德)[1] & JIN Qi (金 奇)[1]

1. State Key Laboratory for Molecular Virology and Genetic Engineering, Beijing 100052, China;
2. Lanzhou Institute of Animal Science and Veterinary Pharmaceutics, Chinese Academy of Agricultural Sciences, Lanzhou 730050, China;
3. Chinese National Human Genome Center, Beijing, Beijing 100176, China

Correspondence should be addressed to Jin Qi (email:zdsys@sina.com)

Received September 27, 2002

Abstract The complete nucleotide sequence and organization of the large virulence plasmid pCP301 (termed by us) of *Shigella flexneri* 2a strain 301 were determined and analyzed. The result showed that the entire DNA sequence of pCP301 is composed of 221618 bp which form a circular plasmid. Sequence analysis identified 272 open reading frames (ORFs), among which, 194 correspond to the proteins described previously, 61 have low identity (<60%) to known proteins and the rest 17 have no regions of significant homology with proteins in database. The genes of pCP301 mainly include the genes associated with bacterial virulence, the genes associated with regulation and the genes relating to plasmid maintenance, stability and DNA metabolism. Insertion sequence (IS) elements are 68 kb in length and account for 30 percent of complete sequence of the plasmid which indicates that gene multiple rearrangements of the pCP301 have taken place in *Shigella flexneri* evolution history. The research result is helpful for interpreting the pathogenesis of *Shigella*, as well as the genetics and evolution of the plasmid.

Keywords: *Shigella flexneri* 2a, virulence plasmid, pCP301, DNA, sequencing, analysis.

DOI: 10.1360/02yc0060

Bacteria *Shigella* spp. are highly contagious, severely harmful and gram-negative facultative intracellular pathogens. They may cause shigellosis characterized by fever, dehydration and hematochezia in clinic, and shigellosis has been remaining a leading cause of infant mortality in the world. *Shigella* belongs to the family *Enterobacteriaceae* and the group *Escherichiaeae*, which are divided into four species and at least 47 serotypes: *Shigella dysenteriae* (13 serotypes), *Shigella flexneri* (15 serotypes), *Shigella boydii* (18 serotypes), and *Shigella sonnei* (1 serotype)[1]. *Shigella flexneri*, the prevalent species in developing countries, kills more than 250000 lives annually[2]. In China, *Shigella flexneri* 2a has been found to be responsible for the major epidemics and pandemics of shigellosis, and about 1 million cases are identified every year, which makes a big threat to the public health.

The pathogenicity of *Shigella flexneri* is closely related to its schlepped large virulence plas-

mid which is about 220 kb and encodes a series of factors required for virulence[3,4]. The process of causing disease is mainly as follows: the bacteria reach the colonic mucusa and invade colonic epithelial cells, leading to intracellular bacterial multiplication. At the same time, they spread to adjacent cells, arousing cell death, and eventually inflammation and destruction of the colonic mucosa[5,6]. Expression of the virulence of the plasmid is regulated by several genes located both on the virulence plasmid and on the chromosome, as well as environmental stimuli[7,8].

Study of complete DNA sequence of the virulence plasmid not only offers the opportunity to understand the mechanism of pathogenicity of *Shigella flexneri*, but also provides comparisons of the virulence plasmid to those of different *Shigella* serotypes and closely related species. It is also important for the recognition of the relationship between pathogenicity and evolution of *Shigella flexneri*, shigellosis control, new drug discovery and vaccine development.

This paper reports the complete DNA sequence and gene analysis of the virulence plasmid of *Shigella flexneri* 2a strain 301, the most prevalent type in China, henceforth designated as pCP301.

1 Materials and methods

1.1 Bacterial strain and plasmid

Shigella flexneri 2a strain 301 was originally isolated from a typical patient of shigellosis at Changping County in China in 1984. Plasmid pCP301 was prepared from strain 301 grown overnight at 37℃ in L broth and purified with QIAGEN tip (QIAGEN Inc.)

1.2 Sub-cloning and sequencing determination

The abstracted DNA was fragmented by sonication. The fractions ranging from 1 to 5 kb were collected and purified and ligated with vector pBluescript to construct random shotgun libraries. Ligation mixtures were used to transform DH5α, and transformants were selected on a plate containing ampicillin, IPTG and X-gal. After 12 h growing at 37℃, the clones that did not express β-galactosidase activity were picked for plasmid amplification. Templets were abstracted with large-scale preparation protocol by using Vitagene kit (Vitagene Inc, China). Templates for sequencing were amplified in a thermocycler with cycle sequencing reactions (30 cycles, 96℃ for 2 min; 50℃ for 10 s; 60℃ for 4 min). About 2200 clones were sequenced from both ends (giving 10 times coverage) using BigDye terminator chemistry on ABI3700 (Perkin Elemer) automated sequencer.

1.3 DNA sequence assembly and annotation

The sequence assembly was carried out by using Phred/Phrap software, choosing optimized parameters and quality score (>= 20) in a computer work station. Fifty contigs based on 4400 sequences were constructed. Sequence gaps were filled by editing the ends of sequence traces with Consed or primer walking on plasmid clones selected according to the analysis of the forward and

reverse links between contigs. The predicting protein-coding regions were initially defined by searching for open reading frames (ORFs) of more than 50 amino acids in length with Glimmer2.01. The possible ORFs were decided by a combination of searching against the NR (the non-redundant NR proteins) and COG (Clusters of Orthologous Groups of proteins) database with Blastp through the internet.

1.4 Nucleotide sequence accession number

The sequence data were deposited in GenBank/EMBL/DDBJ under accession number AF386526.

2 Results and discussion

2.1 General review

The complete DNA sequence of pCP301 consists of 221618 bp, which forms a circular plasmid. Sequence analysis identified 272 potential ORFs, among which 194 have high identity (>60%) and 61 have low identity (<60%) to the subject proteins, another 17 have no regions of significant homology with proteins in the database which have been confirmed as new potential genes (plate I). About 68 kb and nearly 30% of the plasmid sequence appears to correspond to IS elements. The known genes on pCP301 include virulence associated genes, virulence regulation genes, and the genes responsible for plasmid maintenance, stability and nucleotide metabolism which makes up 100 kb in total. An almost 50 kb (24%) part is a non-encoding area.

2.2 Genes associated with virulence

2.2.1 *ipa-mxi-spa* genes. There was a 31 kb region in pCP301 (bp 104162—135422) which includes 34 ORFs (from ORF123 to ORF156) with high identity to the subject proteins in the database. This region was designated invasion region containing *ipa*, *mxi* and *spa* genes and was characterized by the following properties: (i) all of the genes clustered tightly, and there was no IS element existing between the genes; (ii) the region contains 10 genes, from *virB* to *icsB*, and the region contains 24 genes, from *ipgD* to *spa*40 transcribed in opposite directions on pCP301; (iii) the G+C content of this region is 33.4%, that is obviously lower than the average G+C content of pCP301 (45.2%) (fig. 1). This suggests that these genes have the same origin; (iv) two truncated IS600 elements were found to flank both sides of the region, suggesting that the gene rearrangement takes place at this site after acquisition of the invasion region.

The region of *ipa-mxi-spa* has remained a key object in research of the virulence plasmid of *Shigella flexneri*. *Ipa* proteins act as the invasins, while the Mxi-Spa proteins establish a type III secretion system. Both Ipa invasins and the type III secretion system proteins form the central effectors causing disease, leading to epithelial cell entry, phagosomal escape and induction of apoptosis as well in macrophages[9]. Synthesis of Ipa proteins is required for the *Shigella flexneri* invasive phenotype, as the transport of these invasins to the bacterial surface and the external medium via products of the *mxi* and *spa* operons[10]. It is clear that Ipa proteins contribute directly to shig-

ellosis. ipa proteins are rapidly secreted from the *Shigella flexneri* outer membrane and from intracellular stores during infection.

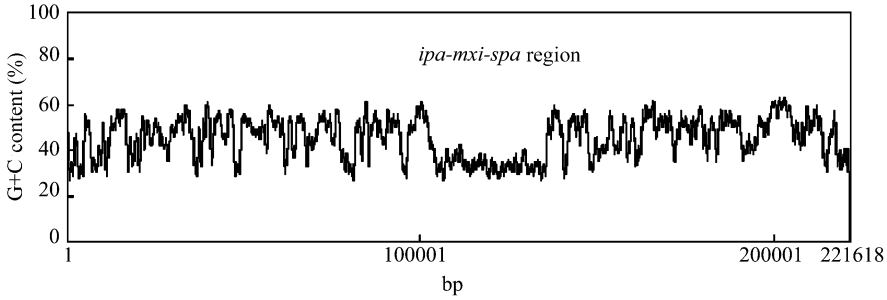

Fig. 1. Base composition of pCP301. The plots show that G+C content was derived by using the DNASIS program. The box indicates the position of the *ipa-mxi-spa* region on pCP301.

Type III secretory apparatus components have been identified in many mammalian and plant pathogens, including enterohemorphagic and enteropathogenic *Escherichia coli* and *Shigella, Salmonella*, *yersinia*, *Chamydia*, *Bordetella*, *Pseudomonas*, *Xanthomonas*, *Ralstonia*, and *Erwinia* spp[11]. In plasmid pCP301, several conservative components of type III secretory systems are also similar to proteins involved in flagellum biosynthesis, suggesting that the filiation of type III secretory systems is related with flagellar subunit secretion. These genes include *mxiJ*, *mxiA*, *spa*47, *spa24*, *spa29* and *spa40*, and their G+C contents are similar (table 1). The study on the genome of *Shigella flexneri* 2a found that mutation of some *Shigella* flagellar associated genes has taken place on the chromosome because the ORFs were truncated (unpublished), i.e. *flgF*, the cell-proximal portion of basal-body rod, *flgK* and *flgL*, the hook-filament junction protein, *flhA*, the possible export of flagellar proteins, *fliA*, the regulator of flagellar operons, and *fliF*, the basal-body-ring and collar protein. The inactivity of these genes may be directly related to the loss of motility in *Shigella* strains. On pCP301, the genes for flagellum biosynthesis may have no relationships with the formation and function of *Shigella* flagellar, so the invasive function could be included.

Table 1 Flagellar biosynthesis associated genes on pCP301

Gene	Site on pCP301/bp	Length (aa)	G+C content (%)	Accession no.
mxiJ	120503—121225	241	32.3	C45271
mxiA	126749—128806	688	33.4	D50601
spa47	129227—130516	430	37.6	C42284
spa24	132695—133342	216	31.4	F42284
spa29	133617—134387	256	30.5	I49846
spa40	134397—135422	342	30.2	A38908

2.2.2 *osp* genes. Six kinds of *osp* genes, including *ospB*, *ospC1—C3*, *ospD1—D3*, *ospE1-E2*, *ospF* and *ospG*, which encode the potential outer *Shigella* proteins, were identified on pCP301. These genes are distributed in different regions on the plasmid. For the *ospC* family, the sequence

of *ospC2* and *ospC3* exhibited 97% identity and showed the nearest relationship. The next closely related gene is *ospC4*, of which the sequence exhibits 84% and 85% identity to those of *ospC2* and *ospC3* respectively. *ospC1* was more distantly related to other genes in *osp* family, the sequence exhibited only 16% identity to that of *ospC4*. For other *osp* genes, the sequence of *ospE1* and *ospE2* exhibited 99% identity, while the sequence of *ospB*, *ospF*, *ospG* and three *ospD* genes showed much less identity to each other, suggesting different periods of duplication event. The G+C content of the *osp* genes ranged from 31% to 38%, which indicated that these genes have the same origin as the *isp-mxi-spa* region genes do. Comparison of pCP301 to pWR100 of *Shigella flexneri* 5a indicated that *ospC4* gene was frameshifted in pWR100 at position 98 and 560, while *osp*C4 gene was complete in pCP301, although whether the encoded protein was secreted and had some function remain unknown, and need to be further investigated in future.

Characteristics and structures of the Osp proteins are significantly different in various microorganisms. The role of Osp proteins during the pathogenesis process of *Shigella flexneri* is still unclear. They may participate in infection and immunity and thus be helpful to disease prevention and vaccine development[12].

2.2.3 *ipa*H genes. Family of *ipaH* genes is a genus encoding invasive plasmid antigens, existing on both the chromosome and virulence plasmid of *Shigella flexneri*. pCP301 contains five *ipaH* copies, which are designated *ipaH1.4*, *ipaH2.5*, *ipaH4.5*, *ipaH7.8* and *ipaH9.8*. We found that of the five *ipaH* copies, *ipaH4.5*, *ipaH7.8* and *ipaH9.8* exist on the chromosome of *Shigella flexneri* 2a strain 301. The previous report indicated that all of the *ipaH* genes in *Shigella flexneri* 5a are characterized by a 5′ end 600—760 bp variable region and a 3′ end 839 bp constant region. The conclusion was confirmed in pCP301. Additionally, the constant region of *ipaH4.5*, ipa*H7.8* and *ipaH9.8* extended 83 bp towards 5′ end to 922 bp, and the constant region of *ipaH7.8* and *ipa*H9.8 extended 39 bp towards 3′ end to 961 bp in pCP301. The sequence of *ipaH1.4* and *ipaH2.5* exhibited the highest identity of 99% in all the *ipaH* genes.

The G+C contents of the constant and variable regions were significantly different. The G+C content of the variable region ranged from 35.0% (*ipaH7.8*) to 39.2% (*ipaH4.1*), while that of the constant region was 52.4%, which was similar to the average chromosome G+C content (50.8%) of *Shigella flexneri* 2a. This suggested that the constant and variable regions of *ipaH* genes have different origins. The former probably came from *Shigella* chromosome or other organisms and was integrated into these genes by an unknown mechanism.

The origin and function of the IpaH proteins are still unknown. They might be secreted by the type III secretion apparatus and related to invasion of *Shigella flexneri*.

2.2.4 *vir* genes and virulence regulation of the plasmid. Virulence gene expression in the plasmid was subject to strict control of the plasmid harbouring key regulatory genes, and regulatory genes located on the chromosome. In pCP301, the sequences of four ORFs including ORF46, ORF123, ORF181 and ORF182 exhibited 100% identity to *virF*, *virB*, *virA* and *virG* respectively.

These genes were located at different sites except *virA* and *virG* on pCP301. The analysis indicated that the sequence of the four genes exhibited very low identity to each other. Moreover, the G+C content of these genes ranged from 30.1% (*virF*) to 41.6% (*virK*), which indicated that they might come from different origins.

ORF237, showing 99% identity to *virK*, were previously identified with *virG* on plasmid pMYSH6000 from *Shigella flexneri* 2a. The *virG*-encoded out membrane proteins play an important role in the activity of *Shigella flexneri* in infection cells, while the function of *virK* was unclear except that *virK*-encoded product is needed for the expression and localization of VirG protein[13]. *virG* and *virK* are located in two different regions with similar G+C content on pCP301, indicating that they may come from the same origin at the same time. On pCP301, the plasmid linked *virB* was located at a distance from the main virulence gene operons, but *ipa* and *mxi-spa* genes were regulated collectively at the transcription level, with regulation being exerted by the products of *virB* genes[14]. *VirA*, located upstream and transcribed divergently from *virG*, was involved in invasion and spreading. The two genes were firstly identified in pMYSH6000 from *Shigella flexneri*.

Previous studies on virulence determinants of *Shigella* species and EIEC strains encoded by plasmid indicated that loci linked to *kcpA*, *ompR-envZ* and *glpK* appeared to regulate expression of the virulence plasmid[15]. This result showed that *kcpA* gene is located on *Shigella flexneri* 2a 301 chromosome between 1302968 bp and 1303982 bp (the complete DNA sequence data of *Shigella flexneri* strain 301 have been deposited in GenBank with accession number AE005674, unreleased) and is associated with expression of *virG* encoded protein. *Shigella flexneri* strains that were defective in either *kcpA* or *virG* expression were fully invasive in the tissue culture model, but they multiplied within localized areas of the cytoplasm of infected cells and did not spread to contiguous cells. The *envZ* gene product was a transmembrane osmolarity sensor that phosphorylated the product of the *ompR* gene, and the latter protein modulates transcription of porin proteins *ompF* and *ompC*. The products of the two-component *ompR-envZ* regulatory locus modulate the expression of the plasmid-encoded invasive phenotype when *Shigella* was exposed to the hypertonic colonic contents of the primate intestine.

The regulation of the *virulence genes* expression was also dependent on temperature. The virulence of bacteria strain would be induced and expressed at 37℃, while it would be repressed at 30℃. The instability of the virulent phenotype of *Shigella flexneri* 2a may be beneficial to the bacteria for survival outside the host, by allowing conservation of energy that would be directed towards virulence maintenance.

2.2.5 Genes for plasmid maintenance, stability and partitioning. Two small contiguous ORF256 (198515—198120 bp) and ORF257 (198742—198518 bp) showed 100% sequence identity to gene *mvpA* and *mvpT* respectively in pCP301. The third contiguous ORF255 (197932—198205 bp) also showed 99% identity to an unknown function protein from *Shigella* in the data-

base, which might be related to *mvpA* and *mvpT* too. The general organization of *mvp* resembles the post-segregational killing systems which encode a toxin and an unstable antidote, when the plasmid is lost from the cell, the antidote decays but the toxin persists, eventually killing most of the progeny of the plasmid free cells. This promotes the maintenance of the plasmid in the growing population[16].

ORF202 (162813—161857 bp) and ORF201 (161854—161462 bp) in pCP301 displayed 99% and 100% identity to protein sequence of *stbA* and *stbB*, synchronously, presented 43% and 29% identity to *parA* and *parB* of R100 respectively, and also exhibited 43% and 27% identity to *parA* and *parB* in *E.coli* respectively. There was an ORF200 which was contiguous to ORF202 and ORF201. An unknown protein in the database might also be related with *stb* function. *stb* system might stabilize a partition-defective P1 mini-plasmid in *E. coli*. It appeared to promote better-than-random distribution of plasmid copies, being a determinant factor of plasmid incompatibility, and causing the coexisting plasmid containing *stb* to produce incompatibility[17].

The pCP301 also contained a *parAB* loci (26955—29131 bp) including *parA* and *parB* genes, which encoded a 399-aa ParA and a 326-aa ParB proteins and exhibited 47% and 58% identity to the partitioning protein ParA and ParB in *E. coli* bacteriphage P1 respectively. The G+C content of *parA* and *parB* was 42.3% and 51.9%, respectively, suggesting that the two genes have different origins: *parB* might come from the chromosome, while *parA* from other plasmids or bacteriphages. The partitioning system was confirmed in the large virulence plasmid from *Shigella flexneri* 5a. It played an important role in partition and segregation of low copy number plasmids in descendant cells[18].

2.3 Evolution and pathogenicity of pCP301

S*higella* was first named *Bacillus dysenteriae* historically, for it was the cause of bacilliary dysentery and closely related to *Bacillus* (*Escherichia*). Both *Shigella* and *E. coli* had always been considered to be very closely related, and some time, they were considered to be so similar that they could be placed in the same species. In the 1940s, four species of the new genus *Shigella* were recognized. Recent work showed that *Shigella* should belong to the *E. coli* taxonomic group in taxology, and it might be divided into at least three clusters as being categorized with some housekeeping genes of *Shigella* strains in genetics, the estimated schedule engendering these three clusters should be 35000—270000 years ago[19]. It appeared that the *Shigella* species designations currently used could not even accurately reflect common evolutionary history, as each of the three major clusters of *Shigella* strains containing sero-types had been assigned to different species by category.

Shigella strains are characterized by the presence of a large virulence plasmid, lack of flagellar motility and inability to ferment lactose compared with *E. coli*. It was commonly considered that *Shigella* could have independently arisen convergent evolution for 7—8 times in family *E. coli*. However, during the evolution, the change of those functional genes led *Shigella* strains to

lose the pathogenic phenotype characteristics under circumstance press[20], but it obtained a large virulence plasmid. The obtain of a large virulence plasmid should be the most important characteristic for *Shigella* during convergent evolution, for it is a proficient vehicle for lateral gene transfer.

Shigella flexneri causes disease by specially invading into mucosal epithelium cells of the large intestine. The gene characterized by coding invasion is located at pCP301in *Shigella flexneri* 2a 301. This was similar to other results on large virulence plasmids of *Shigella* strains. It was found that the nucleotide sequence of pCP301 is 8124 bp longer than that of pWR100 compared with large plasmid pWR 100 of *Shigella flexneri* 5a. All of the known genes corresponding to invasive phenotypes are identical in both plasmids. The different sequences in pCP301 come from IS elements and other foreign plasmids, of which the encoded products and functions are waiting for further investigation.

In addition, analysis of the genome of *Shigella flexneri* 2a 301 showed that the lactose degradation-associated *lacZ* gene and flagellar biosynthesis-associated *fliN*, *fliF* and *flhD* genes were ruptured and lost activity by mutation in the chromosome (unpublished). This phenomenon might be caused by restricting intracellular movement of *Shigella flexneri* in the cell of host, meanwhile, the energy source needed for movement was also limited. Moreover, loss of lysine decarboxylase through genetic deletion appeared to be favorable, as the product of the lysine decarboxylase reaction inhibited enterotoxin activity.

Nearly one third of the sequence of pCP301 is composed of IS elements. Such a high ratio of IS has never been discovered in any microorganism. The complex components and sources of IS elements suggest a heterogeneity of DNA acquisition in plasmid evolution. The G+C contents of both the *ipa-mxi-spa* and *osp* region genes are similar, indicating that these genes have the same or similar origin and time of genes acquisition, but the G+C contents of several other regions of the gene were significantly different. Insertion of DNA and rearrangement of gene caused the phenotype and pathogenicity of pCP301 change in the convergent evolution of plasmid. The functions of 17 newly identified ORFs in pCP301 are still unclear. In the mean time, the functions of many genes and products remained to be entirely clarified, especially those regulation mechanisms during the pathogenesis period of *Shigella flexneri* 2a 301. Further research will further supply evidence for interpreting the essence of descendibility evolution and pathogenesis of the plasmid.

Acknowledgments We thank Zhu Junping, Sun Lilian, Wu Hongtao, Wang Jing and Chen Shuxia for their expert technical assistance and the sequencing group of Chinese National Human Genome Center, Beijing for technical work. This work was supported by the State Key Basic Research Program and High Technology Project of the Ministry of Science and Technology of China.

References

1. Fernando, R. N., Liao, F. M., David, R. M. et al., Strategy for cross-protection among *Shigella flexneri* serotypes, Infect. Immun., 1999, 67: 782—788.
2. Sansonetti, P. J., *Shigella* plays dangerous games, ASM News, 1999, 65: 612—616.
3. Marquart, M. E., Picking, W. L., Picking, W. D., Soluble invasion plasmid antigen C (IpaC) from *Shigella flexneri* elicits

epithelial cell responses related to pathogen invasion, Infect. Immun., 1996, 64: 4182—4187.

4. High, N., Mounier, J., Prevost, M. C. et al., IpaB of *Shigella flexneri* causes entry into epithelial cells and escape from the phagocytic vacuole, EMBO J., 1992, 11: 1991—1999.
5. LaBrec, E. H., Schneider, H., Magnani, T. J. et al., Epithelial cell penetration as an essential step in the pathogenesis of bacillary dysentery, J. Bacteriol., 1988, 88: 1503—1518.
6. Bernardini, M. L., Mounier, J., Hauteville, H. D. et al., Identification of *icsA*, a plasmid locus of *Shigella flexneri* that governs intra- and intercellular spread through interaction with F-actin, Proc. Natl. Acad. Sci. USA, 1989, 86: 3867—3871.
7. Maurelli, A. T., Sansonetti, P. J., Identification of a chromosomal gene controlling tempreture-regulated expression of *Shigella* virulence, Proc. Natl. Acad. Sci. USA, 1988, 85: 2820—2824.
8. Tobe, T., Yoshikawa, M., Mizuno, T. et al., Transcriptional control of the invasion regulatory gene *virB* of S*higella flexneri*: Activation by *virF* and repression by H-NS, J. Bacteriol., 1993, 175: 6142—6149.
9. Zychlinsky, A., Prevost, M. C., Sansonetti, P. J., *Shigella flexneri* induces apoptosis in infected macrophages, Nature, 1992, 358: 167—169.
10. Watarai, M., Tobe, T., Yoshikawa, M. et al., Contact of *Shigella* with host cells triggers release of Ipa invasins and is an essential function of invasiveness, EMBO J., 1995, 14: 2461—2470.
11. Van, G. F., Genin, S., Evolutionary conservation of pathogenicity determinants among plant and animal pathogenic bacteria, Trends Microbiol., 1993, 1: 175—180.
12. Schutzer, S. E., Cole, P. K., Krupp, L. B. et al., Simultaneous expression of Borrelia OspA and OspC and IgM response in cerebrospinal fluid in early neurologic lyme disease, J. Clin. Invest., 1997, 100: 763—767.
13. Porter, M. E., Smith, S. G., Dorman, C. J., Two highly related regulatory proteins, *Shigella flexneri* VirF and enterotoxigenic *Escherichia coli* Rns, have common and distinct regulatory properties, FEMS Microbiol. Lett., 1998, 162: 303—309.
14. Berlutti, F., Casalino, M., Zagaglia, C. et al., Expression of the virulence plasmid-carried apyrase gene *(apy*) of enteroinvasive *Escherichia coli* and *Shigella flexneri* is under the control of H-NS and the VirF and VirB regulatory cascade, Infect. Immun., 1998, 66: 4957—4964.
15. Watanabe, H., Genetics of virulence of *Shigella* species, Microbiol. Sci., 1988, 5: 307—310.
16. Sayeed, S., Reaves, L., Radnedge, L. et al., The stability region of the large virulence plasmid of *Shigella flexneri* encodes an efficient postsegregational killing system, J. Bacteriol., 2000, 182: 2416—2421.
17. Radnedge, L., Davis, M. A., Youngren, B. et al., Plasmid maintenance functions of the large virulence plasmid of *Shigella flexneri*, J. Bacteriol., 1997, 179: 3670—3675.
18. Kwong, S. M., Yeo, C. C., Poh, C. L., Molecular analysis of the pRA2 partitioning region: ParB autoregulates *parAB* transcription and forms a nucleoprotein complex with the plasmid partition site, *pars*, Mol. Microbiol., 2001, 40: 621—633.
19. Pupo, G. M., Lan, R., Reeves, P. R., Multiple independent origins of *Shigella* clones of *Escherichia coli* and convergent evolution of many of their characteristics, Proc. Natl. Acad. Sci. USA, 1998, 97: 10567—10572.
20. Mamun, A. A., Tominaga, A., Enomoto, M., Cloning and characterization of the region III flagellar operons of the four *Shigella* subgroups: Genetic defects that cause loss of flagella of *Shigella boydii* and *Shigella sonnei*, J. Bacteriol., 1997, 179: 4493—4500.