Applied and Environmental Microbiology, July 2009, p. 4362–4373 Vol. 75, No. 13
0099-2240/09/$08.00+0 doi:10.1128/AEM.02521-08

Conjugative Type 4 Secretion System of a Novel Large Plasmid from the Chemoautotroph *Tetrathiobacter kashmirensis* and Construction of Shuttle Vectors for *Alcaligenaceae*▽†

Bomba Dam,[1]‡ Wriddhiman Ghosh,[1,2] and Sujoy K. Das Gupta[1]*

Department of Microbiology[1] *and Department of Biophysics,*[2] *Bose Institute, P-1/12, C. I. T. Scheme VII-M, Kolkata 700054, India*

Received 4 November 2008/Accepted 27 April 2009

Tetrathiobacter **spp. and other members of the** ***Alcaligenaceae*** **are metabolically versatile and environmentally significant. A novel, ~60-kb conjugative plasmid, pBTK445, from the sulfur chemolithoautotroph** ***Tetrathiobacter kashmirensis*****, was identified and characterized. This plasmid exists at a low copy number of 2 to 3 per host chromosome. The portion of pBTK445 sequenced so far (~25 kb) harbors genes putatively involved in replication, transfer functions, partition, and UV damage repair. A 1,373-bp region was identified as the minimal replicon. This region contains a** ***repA*** **gene encoding a protein belonging to the RPA (replication protein A) superfamily and an upstream, iteron-based** ***oriV*****. A contiguous 11-gene cluster homologous to various type 4 secretion systems (T4SSs) was identified. Insertional inactivation demonstrated that this cluster is involved in the conjugative transfer functions of pBTK445, and thus, it was named the** ***tagB*** **(transfer-associated gene homologous to** ***virB*****) locus. The core and peripheral TagB components show different phylogenetic affinities, suggesting that this system has evolved by assembling components from evolutionarily divergent T4SSs. A** ***virD4*** **homolog, putatively involved in nucleoprotein transfer, is also present downstream of the** ***tagB*** **locus. Although pBTK445 resembles IncP plasmids in terms of its genomic organization and the presence of an IncP-specific** ***trbM*** **homolog, it also shows several unique features. Unlike that of IncP, the** ***oriT*** **of pBTK445 is located in close proximity to the** ***oriV*****, and a** ***traL*** **homolog, which is generally present in the TraI locus of IncP, is present in pBTK445 in isolation, upstream of the** ***tagB*** **locus. A significant outcome of this study is the construction of conjugative shuttle vectors for** ***Tetrathiobacter*** **and related members of the** ***Alkaligenaceae*****.**

The genus *Tetrathiobacter* includes environmentally important betaproteobacteria belonging to the family *Alcaligenaceae*. Members of this family inhabit diverse habitats, ranging from animals and humans to soil, sewage, and sludge. They are also metabolically diverse and include facultative chemolithotrophs, versatile heterotrophs, xenobiotic degraders, fastidious parasites, and pathogens (15). While the type species, *Tetrathiobacter kashmirensis*, isolated from a temperate orchard soil, has been recognized as a thiosulfate- and tetrathionate-oxidizing facultative chemolithoautotroph (11, 15), *Tetrathiobacter mimigardefordensis*, isolated from compost, can utilize the organic disulfide 3,3′-dithiodipropionic acid for growth (42). More recently isolated soil-dwelling strains of *T. kashmirensis* can detoxify selenite by reducing it to insoluble elemental red selenium (18). Strains identified as *T. kashmirensis* on the basis of 16S rRNA gene sequence similarity (GenBank accession number EU523111) are allegedly involved in the biodegradation of thiodiglycol, the hydrolysis product of yperite, a highly hazardous derivative of mustard gas used in chemical weapons. In addition, bacteria isolated from a deep-sea environment and phylogenetically identified as *T. kashmirensis* (GenBank accession number EF619402) have been observed to degrade alkanes.

Species of *Alcaligenaceae* possess a wide repertoire of plasmids (21, 32), a feature pertinent to their biodegradative and biogeochemical roles in the environment. Many of these plasmids are well known for harboring genes involved in biodegradation (14, 39, 44). However, not many of them have been studied at the molecular level. In the present study, we have identified, partially sequenced, and characterized a large (~60-kb), low-copy-number, self-transmissible, novel plasmid, designated pBTK445, from *T. kashmirensis* strain WGT. We have characterized the minimal replication region of this plasmid and have subsequently constructed shuttle vectors that could be used for diverse members of the *Alcaligenaceae*, including *Tetrathiobacter*. A major part of the sequenced region was found to be occupied by genes homologous to constituents of various type 4 secretion systems (T4SSs) (5–7, 9). This locus was found to be involved in the conjugal transfer function of the new plasmid. Many features of pBTK445 resemble those of IncP plasmids, but the new plasmid also possesses several characteristics distinct from those of IncP plasmids. We discuss in detail those characteristics of pBTK445 that make it an interesting model for the study of the diversity and evolution of large plasmids.

MATERIALS AND METHODS

Bacterial strains and growth conditions. The bacterial strains used in this study and their sources are listed in Table 1. The growth conditions and minimal media used for growing and maintaining different strains of *Escherichia coli*, *T.*

* Corresponding author. Mailing address: Department of Microbiology, Bose Institute, P-1/12, C. I. T. Scheme VII-M, Kolkata 700 054, India. Phone: 91-033-2355-9416. Fax: 91-33-2355-3886. E-mail: sujoy@boseinst.ernet.in.

‡ Present address: Microbiology Laboratory, Department of Botany, Institute of Sciences, Visva-Bharati Central University, Santiniketan, West-Bengal 731 235, India.

† Supplemental material for this article may be found at http://aem.asm.org/.

▽ Published ahead of print on 1 May 2009.

TABLE 1. Bacterial strains and plasmids used in this study

Bacterial strain or plasmid	Relevant characteristic(s) and genetic marker(s)[a]	Source or reference
Strains		
Escherichia coli		
XL1-Blue	*recA1 lac endA1 gyrA46 thi hsdR17 supE44 relA1* [F′ *proAB lacI*^{q}ZΔM15 Tn*10* (Tet^r)]	Stratagene
SY327λ*pir*	Δ(*lac pro*) *argE*(Am) *recA* (Rif^r) *nalA* λ*pir*	27
SM10λ*pir*	*thi-1 thr leu tonA lacY supE recA*::RP4-2-Tc::Mu λ*pir*	27
Sulfur chemolithotrophs		
T. kashmirensis LMG 22696	Wild type harboring pBTK445; Ap^r Sox$^+$	15
T. kashmirensis SR	Spontaneous Sm^r Rif^r mutant of *T. kashmirensis* LMG 22696	11
T. kashmirensis C1	pBTK445-cured strain of *T. kashmirensis* SR	This study
BD 004	Tn*5-mob* insertion mutant of *T. kashmirensis* SR; Km^r Sox$^-$	11
BD 004 C1	pBTK445-cured strain of BD 004; Km^r	This study
Plasmids		
Large plasmids, cloning and expression vectors, and sources of antibiotic markers		
pBTK445	Large (~60-kb), low-copy-number plasmid harbored in *T. kashmirensis*	This study
pBTKCΩ*6*	*orf6*::ΩCm (pBTK445 with *orf6* inactivated)	This study
pBTKCΩ*B4-B5*	*tagB4-tagB5*::ΩCm (pBTK445 with *tagB4* and *tagB5* inactivated)	This study
pBluescript KS(+)	2.96-kb cloning vector; Ap^r	Stratagene
pKAS32	Cloning vector with dominant *rpsL* gene; Ap^r	36
pSD5B	Promoter probe vector (also used to probe the replication and transfer origins of pBTK445); Km^r	19
pQE30	T5 promoter with His_6 tag; Ap^r	Qiagen
pFH450	7.4-kb cloning vector containing P1 and ColE1 replicons; Cm^r	16
pUC4K	pUC vector carrying the kanamycin cassette from Tn*903*; Ap^r Km^r	Amersham
pUC4C	pUC vector carrying the chloramphenicol cassette from pFH450; Ap^r Cm^r	This study
Plasmids for shuttle vector construction, curing, mobilization, and complementation assays		
pBTKE5.7	5.7-kb EcoRI fragment of pBTK445 cloned into pBluescript; Ap^r	This study
pBTKE5.7R	5.7-kb EcoRI fragment of pBTK445 cloned into pBluescript in reverse direction; Ap^r	This study
pBTKS	1.9-kb Cm^r gene from pUC4C cloned into the PstI site of pBTKE5.7; Ap^r Cm^r	This study
pBTKSR	1.9-kb Cm^r gene from pUC4C cloned into the PstI site of pBTKE5.7R; Ap^r Cm^r	This study
pBTKAS	5.2-kb BamHI fragment of pBTKS cloned into pKAS32; Ap^r Cm^r	This study
pSUP5011	pBR325 (Bam$^-$)::Tn*5-mob*; Cm^r Ap^r Km^r	35
pBTKSM	3.6-kb PCR product from the *moco* locus of *T. kashmirensis* cloned into the XbaI site of pBTKS; Cm^r	This study
Constructs for in vivo determination of *oriV*, *oriT*, and P_{repA}		
pBTK1C	pBTKSR digested with SacII, with the larger fragment religated; Ap^r Cm^r	This study
pBTKHE3.1	pBTKE5.7 digested with HindIII and religated; Ap^r Cm^r	This study
pBTK2C	1.9-kb Cm^r gene from pUC4C cloned into the BamHI site of pBTKHE3.1; Ap^r Cm^r	This study
pBTK3C	pBTKS digested with ClaI and religated; Ap^r Cm^r	This study
pSD5MR	1,373-bp *repA*- and *oriV*-containing PCR product cloned into pSD5B; Km^r	This study
pSD5V	264-bp *oriV*- and *repA* promoter-containing PCR product cloned into the XbaI site of pSD5B; Km^r	This study
pSD5T	100-bp *oriT*-containing PCR product cloned into the XbaI site of pSD5B; Km^r	This study
Recombinant plasmids for knockout mutagenesis		
pBTKHE2.5	pBTKE5.7R digested with HindIII and religated; Ap^r	This study
pKORF6	pKAS32 containing *orf6* and *orf7* as a 1.97-kb BamHI and EcoRI fragment of pBTKHE2.5; Ap^r	This study
pKORF6C	pKORF6 with a Cm^r cassette inserted at the BamHI site of *orf6*; Ap^r Cm^r	This study
pBTKH1.9	1.9-kb HindIII fragment containing portions of *tagB4-tagB5* of pBTK445 cloned into pBluescript; Ap^r	This study
pKTAGB	pKAS32 containing partial *tagB4* and *tagB5* genes; Ap^r	This study
pKTAGBC	pKTAGB with *tagB4* and *tagB5* inactivated by a Cm^r cassette; Ap^r Cm^r	This study

[a] Sox$^+$ strains have the ability to oxidize reduced sulfur compounds, while Sox$^-$ strains do not.

TABLE 2. Primers used in this study

Primer	Sequence[a]	Construct(s)	Use
$OriV_R$	5′-gcTCTAGAAGCTTATTAAGGGTTTTG-3′	pSD5MR, pSD5V	Replication and *oriV* assays
$RepA_C$	5′-gcTCTAGATTAATTGAATAAGTTCGC-3′	pSD5MR	Replication assay
$OriV_F$	5′-gcTCTAGAGCTCGTTTCTTTAGTGC-3′	pSD5V	*oriV* assay
$OriT_F$	5′-gcTCTAGAGTTCTTGTATCACCAGCC-3′	pSD5T	*oriT* assay
$OriT_R$	5′-gcTCTAGACTTTTTTATTGCTGATGG-3′		
$RepA_N$	5′-ccgcGGATCCATGGAGTACCCTATTTTC-3′	pQERA	Promoter assay
$RepA_C$	5′-gccccAAGCTTAATTGAATAAGTTCGC-3′		
$ParS_R$	5′-AGAATGACCTTGTACCC-3′		Copy no. estimation
$ParA_N$	5′-ATGATAATTTCATTCC-3′		
P_F	5′-CAAAATCGCTAATAGAGGGTC-3′		pBTK445 detection
P_R	5′- GATACCGGAACCGCAACAC-3′		
$TB11_{F1}$	5′-CCAATCACCCTAACCATGTC-3′		Arbitrary PCR
$TB11_{F2}$	5′-AATCGCTCATTGCTGCTTGC-3′		
$Arbit_1$	5′-GGCCACGCGTCGACTAGTCANNNNNN-3′		Arbitrary PCR
$Arbit_2$	5′-GGCCACGCGTCGACTAGTCA-3′		
$SoxB_F$	5′-CAAGGGAAAAACACTGGTGAC-3′		Copy no. estimation
$SoxB_R$	5′-CCCTTTACGCCCCCTTTATC-3′		
$Moco_F$	5′-ctagTCTAGAGTGTACCTTCGGGTAC-3′	pBTKSM	Complementation
$Moco_R$	5′-ctagTCTAGATGAATCCCTTGCTGCG-3′		

[a] Underlined sequences indicate restriction sites; nucleotides shown in lowercase were added to the 5′ ends of the sites to ensure complete digestion.

kashmirensis, and other bacterial strains have been described previously (11, 15). When required, tetracycline, kanamycin, chloramphenicol, ampicillin, streptomycin, or rifampin (rifampicin) was added to the medium at a final concentration of 20, 50, 60, 100, 100, or 250 μg ml^{-1}, respectively.

Plasmid construction and DNA manipulations. The plasmids and oligonucleotides used in this study and their sources are listed in Tables 1 and 2, respectively. Large plasmids from strains of *T. kashmirensis* were isolated using a Qiagen large-construct plasmid isolation kit with an increased incubation time in certain steps. All regular DNA manipulations and hybridizations were carried out according to standard methods (33). Because *T. kashmirensis* is resistant to ampicillin (15), a kanamycin resistance (Km^r) or chloramphenicol resistance (Cm^r) cassette was incorporated into all constructs for their selection in the host. The Km^r cassette was derived from pUC4K, whereas the Cm^r cassette was obtained from a similar vector, pUC4C, in which a PstI-digested Cm^r gene of pFH450 replaces the Km^r gene of pUC4K. The *repA* overexpression plasmid pQERA was constructed by first amplifying the gene using primers $RepA_N$ and $RepA_C$ and then cloning it into the BamHI-HindIII sites of the expression vector pQE30 (Qiagen). The inserted fragment was sequenced in order to eliminate the possibility of any misincorporation of nucleotides. Expression of the recombinant RepA protein was checked by induction with isopropyl-β-D-thiogalactopyranoside (IPTG) (26, 33).

Sequencing of pBTK445. The nucleotide sequence of pBTK445 was determined either from the restriction fragment bank or by arbitrary primed PCR. For the construction of the bank, independent HindIII and EcoRI fragments were shotgun cloned into pBluescript KS(+) (Stratagene) and were sequenced starting with the T3 and T7 universal primers. To confirm the orientation of the fragments, two different restriction enzymes were used. To resolve ambiguities, close gaps, and cross restriction sites, a series of walking primers was constructed. The sequence downstream of the HindIII site, present within *tagB11*, was obtained by arbitrary primed PCR. Two sequential amplification steps were carried out with primers complementary to the sequenced regions of *tagB11* reading outward and with arbitrarily designed primers reading inward from the unknown sequence. In the first PCR, the arbitrary primer $Arbit_1$ and the *tagB11*-specific primer $TB11_{F1}$ were used against pBTK445 DNA as the template. The amplified PCR product was used as the template in a second PCR employing another *tagB11*-specific nested primer, $TB11_{F2}$, paired with a second arbitrary primer, $Arbit_2$, which is the constant region of primer $Arbit_1$. A specific 2.2-kb product obtained from the second reaction was gel purified and sequenced starting with the $TagB11_{F2}$ primer, followed by a series of walking primers. DNA sequencing was performed using an ABI dye terminator sequencing kit and an automated DNA sequencer, model ABI 377 (Applied Biosystems), at the Facility for Genomics and Proteomics at Bose Institute, Kolkata, India (Intensification of Research in High Priority Areas, Department of Science and Technology, Government of India).

Sequence analysis, Web servers, and homology predictions. Sequence analyses and translations were done with the online search engines at the ExPASy Proteomics Server (http://www.expasy.org), JustBio, and the Sequence Manipulation Suite, version 2 (http://www.bioinformatics.org/sms2/index.html). Annotation was performed using the GenBank CDS translation, PDB, SwissProt, PIR, and PRF protein databases and the DDBJ/EMBL/GenBank DNA databases with the BLAST programs (BLASTN, BLASTP, and BLASTX) at the National Center for Biotechnology Information website (http://www.ncbi.nlm.nih.gov/BLAST/) (2). Conserved domains were searched using the InterProScan (http://www.ebi.ac.uk/InterProScan/) and Pfam (http://pfam.sanger.ac.uk/) programs. Transmembrane helices were predicted for significant transmembrane segments using TMpred (http://www.ch.embnet.org/software/TMPRED_form.html). The Web-based software utilities CELLO, version 2.5 (http://cello.life.nctu.edu.tw/), and PSORTb, version 2.0.4 (http://www.psort.org/psortb/), were used to predict the cellular localization of the translated proteins. Plasmids and DNA fragments were drawn to scale using pDRAW32 software, version 1.0, from AcaClone.

Plasmid curing. The incompatibility characteristics of the pBTK445 partitioning locus (unpublished data) were utilized to develop plasmid-free strains of *T. kashmirensis* (38). pBTKAS, containing the partition site of pBTK445 on a pKAS32 backbone, was used as the curing agent (see the figure in the supplemental material). *E. coli* SM10λ*pir* harboring the plasmid was mated with *T. kashmirensis* SR (or BD 004) cells, and transconjugants were selected on chloramphenicol- and rifampin-containing Luria agar (LA) plates. Transconjugants were subcultured several times on LA plates containing chloramphenicol (specific for pBTKAS). Individual colonies were screened for the loss of pBTK445 by PCR using primers P_F and P_R. The primers were used to amplify a ~1.8-kb product (nucleotides [nt] 18380 to 20218) from a region of the plasmid (*tagB* locus) not incorporated into the shuttle vector. Southern hybridization with probes amplified using similar primers was also performed to confirm the loss of pBTK445. In the second step, derivatives of the transconjugants without pBTKAS were selected in the presence of streptomycin. Due to the presence of the *rpsL* gene, coding for ribosomal S12 protein, on the pKAS32 backbone, transconju-

gants with a single crossover were streptomycin sensitive and those with a double crossover were resistant to the antibiotic (36).

Quantitative measurement of plasmid copy number by real-time PCR. The plasmid copy number relative to the chromosome was determined by quantitative real-time PCR analysis. Whole cells lysed by heating at 95°C for 10 min were used as a template source for performing the quantitative PCR in order to avoid experimental errors arising from irreproducible DNA isolation. PCR amplicons of similar sizes (~100 bp) required for performing quantitative real-time PCR were obtained using primers $SoxB_F$ and $SoxB_R$ to fish out a 90-bp product from the genome of *T. kashmirensis* (unpublished data), whereas primers $ParS_R$ and $ParA_N$ were used to amplify a 101-bp product from the plasmid. The 7500 fast real-time PCR system (Applied Biosystems) was used for quantitative PCR amplification, and SYBR green was used for detecting the amplified DNA. Relative quantification, taking into account the different amplification efficiencies of amplicons for the chromosome and plasmids, was used in the plasmid copy number calculation according to standard methods (25, 37).

Promoter assays. Promoter assays were performed using the promoter probe vector pSD5B (19). The promoter sequence to be tested was cloned into an XbaI site upstream of the β-galactosidase reporter gene. Promoter assays were performed in *E. coli* XL1-Blue either alone or in combination with the RepA-expressing plasmid pQERA, as described previously (19, 26).

Conjugation and electroporation. Conjugative mating was performed by the plate mating method. Conjugation experiments, where the ability to transfer DNA to *T. kashmirensis* cells was studied, included a pBTK445-cured strain of the Km^r derivative of *T. kashmirensis* SR (BD 004 C1) as the recipient in order to avoid potential recombination events. Donors and recipients were grown in the presence of the appropriate antibiotic to an optical density at 600 nm of 0.7 and were mixed in a total volume of 1 ml at a ratio of 1:1. The mixture was then centrifuged, and the cells were washed twice with 0.9% NaCl, spread on an LA plate, and incubated overnight at 30°C (or 37°C, when *E. coli* was used as the recipient). The conjugate growth on the mating plate was suspended in 4 ml of 0.9% NaCl, and after appropriate dilutions it was spread on LA plates supplemented with the appropriate single or double antibiotic in order to determine the number of donors or transconjugants, respectively. The frequency of conjugation was expressed as the number of transconjugants per donor cell, and values are means for three independent experiments. Triparental mating was performed in a similar manner, with the donor, recipient, and helper strains mixed in a 1:1:2 ratio.

For electroporation, a mid-log-phase culture was harvested by centrifugation at 5,000 × *g* for 10 min. The cell pellet was first washed twice with chilled distilled water, then washed with chilled 15% (vol/vol) glycerol, and ultimately resuspended in a 1/20 volume of the same solution to give a cell density of 10^{12} to 10^{14} cells ml^{-1}. Forty microliters of this electroporation-competent cell suspension was mixed with 100 to 200 ng of plasmid DNA (1 to 4 μl) in the electroporation cuvette (Bio-Rad) and incubated for 10 min on ice to allow DNA-cell saturation. The chilled cell mixture was shocked using a Gene Pulser apparatus (Bio-Rad) set at 2.43 kV, 25 μF, and 200 Ω. The shocked cells were then resuspended in 1 ml of Luria broth, placed in sterile Eppendorf tubes, and incubated for 2 h at 30°C (or 37°C for *E. coli*). Appropriate dilutions were plated either on LA alone or on LA supplemented with a plasmid-specific antibiotic in order to determine the number of recipients or transformants, respectively. Transformation frequencies are expressed as the number of transformants per total number of competent cells per microgram of plasmid DNA, and values are means for three independent experiments. In both cases, plasmid DNA was isolated from the transconjugants/transformants and examined for its integrity by running it on an agarose gel and performing PCR using plasmid-specific primers.

Construction of a Cm^r derivative and a *tagB4 tagB5* knockout mutant of pBTK445. pBTK445 was tagged with a Cm^r cassette, in an unrelated open reading frame (ORF), *orf6*, via gene replacement using a homologous-recombination strategy. A 2.5-kb HindIII-EcoRI fragment containing *orf6* was cloned into pBluescript, creating pBTKHE2.5. A 1.97-kb fragment (harboring *orf6*) obtained by double digestion of the plasmid with BamHI and EcoRI was gel purified and cloned into the BglII-EcoRI site of the suicide vector pKAS32 (resulting in the construct pKORF6). A BamHI-digested Cm^r cartridge from pUC4C was then introduced into the BglII site (which is present once within *orf6*) of pKORF6. The resulting suicide construct, pKORF6C, was then conjugated from *E. coli* SM10λ*pir* cells to *T. kashmirensis* SR, and transconjugants in which a plasmid integration event had occurred were selected on plates containing chloramphenicol and rifampin. Transconjugants with a double crossover and harboring the mutated plasmid pBTKCΩ*6* were then selected in the presence of streptomycin as described above.

Two genes from the *tagB* locus (*tagB4* and *tagB5*) were disrupted by homologous recombination using a procedure similar to that described above. A 1.9-kb HindIII fragment of pBTK445 containing partial sequences from *tagB4* and *tagB5* was cloned into pBluescript to yield pBTKH1.9. A 1.81-kb fragment (containing the two *tagB* genes) obtained by digesting pBTKH1.9 with BamHI (derived from the vector) and EcoRI was cloned into the BglII-EcoRI site of pKAS32. A BamHI-digested Cm^r cartridge from pUC4C was then introduced into the two BglII sites (one in *tagB4* and the other in *tagB5*) of the resultant plasmid, pKTAGB. The recombinant suicide plasmid pKTAGBC was then conjugated from *E. coli* SM10λ*pir* to *T. kashmirensis* SR, and transconjugants with double crossovers were selected as described above. This ultimately resulted in the knockout construct pBTKCΩ*B4-B5*. In both cases, the strain(s) of *T. kashmirensis* harboring the mutated plasmid was confirmed both by PCR and by Southern blotting.

Nucleotide sequence accession number. The partial nucleotide sequence of pBTK445 has been deposited in the GenBank database and assigned accession number EU585932.

RESULTS

Isolation and characterization of a novel plasmid. The sulfur-chemolithoautotrophic betaproteobacterium *T. kashmirensis* strain WGT (LMG 22696) (15) was found to harbor a novel plasmid, which was designated pBTK445. Plasmids with similar sizes were also found in the other reported strains of *T. kashmirensis*, such as WT001^T (also referred to as MTCC 7002^T and LMG 22695^T), 445a, 445c, and WPT. The plasmid obtained from *T. kashmirensis* strain WGT was further characterized. Based on independent restriction mapping with EcoRI and HindIII, the size of the plasmid was found to be ~60 kb (Fig. 1). Quantitative real-time PCR experiments revealed that its copy number is low, in the range of 2 to 3 copies per host chromosome. The host cells could not be cured of the plasmid by treatment with chemical agents such as acridine orange, *para*-rosaniline, or ethidium bromide. However, curing was achieved by utilizing the partition-based incompatibility property of pBTK445 (data not shown). The cured strain was tested for antibiotic resistance and was found to have properties similar to those of the wild type. This indicates that the plasmid does not possess the tested antibiotic resistance markers, such as chloramphenicol, kanamycin, rifampin, streptomycin, and tetracycline resistance markers.

Analysis of the sequenced region of pBTK445. A ~24.5-kb contiguous region of pBTK445 (Fig. 2A) was sequenced (GenBank accession number EU585932) and found to have an overall G+C content of 45.49 mol%, which was distinct from the G+C content of the genomic DNA of *T. kashmirensis* (54 to 55.2 mol%) (15), indicating probably distinct lineages of the genome and at least the sequenced region of the plasmid. Bioinformatic analyses of the sequenced region revealed 27 putative ORFs, 22 of which could be annotated based on homology with protein sequences of either plasmid or chromosomal origin (Tables 3 and 4).

One of the ORFs (*orf5*) potentially encodes a 253-amino-acid product, showing homology to several genome- as well as plasmid-derived putative replication proteins. The highest similarity was observed with the RepA protein from the genome of *Rhodospirillum rubrum* (GenBank accession number YP_425444) (52% identity), followed by replication proteins of several betaproteobacterial plasmids, such as pPNAP06 of *Polaromonas naphthalenivorans* CJ2 (GenBank accession number YP_973978) (48% identity) and plasmid 2 of *Nitrosospira multiformis* (GenBank accession number YP_413485) (47% identity). pBTK445 RepA and its close relatives were found to

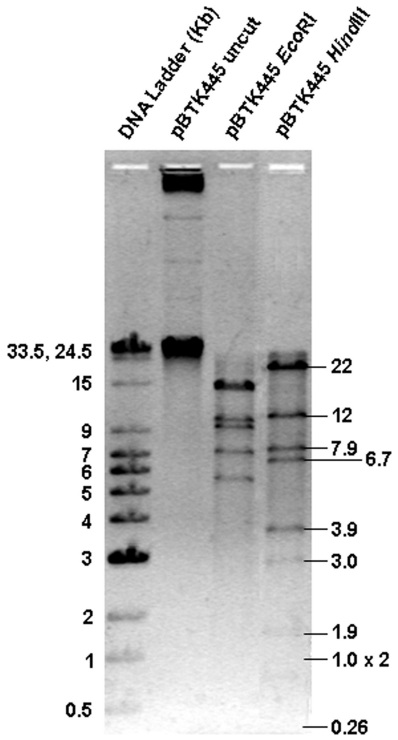

FIG. 1. Electropherogram of pBTK445 and its restriction profile for the determination of its size. Leftmost lane, DNA ladder; second lane, uncut plasmid; third and fourth lanes, plasmid DNA digested with EcoRI and HindIII, respectively. The sizes of the DNA bands (in kilobases) are given on the left for the DNA ladder and on the right for the restriction fragments of the plasmid digested with HindIII. The size of the plasmid determined from the summation of the restriction fragments is ~60 kb (the sizes of the HindIII fragments are shown). Although the lower bands are not clearly visible, they have been obtained as clones in the restriction fragment bank and sequenced.

belong to the RPA (replication protein A) superfamily (pfam10134) of proteins. Members of this family of proteins, characterized mainly in eukaryotic systems, including humans, are single-stranded DNA binding proteins involved in DNA replication, repair, and recombination (40, 43).

The 1,071-bp noncoding region upstream of *repA* (nt 2982 to 4052) shows similarity to well-characterized iteron-based plasmid replication origin (*oriV*) sequences. This region is relatively AT rich (55%) and contains four 22-bp direct repeats (having the consensus sequence 5′-GTGCAGGTGAAATGC CGATAGT-3′) in addition to other direct- and inverted-repeat sequences (Fig. 2B). One of the three consecutive ORFs (*orf4*), present immediately downstream of *repA* and transcribed in the same direction, encodes a 214-amino-acid product, which shows homology with members of the ParA family of ATPases, involved in plasmid partition (17, 28). Moreover, a stretch of 11 genes that shows homology with members of the diverse T4SS family was identified (Tables 3 and 4). Two other genes, *orf6* and *orf7*, show homology with the *umuC* and *umuD* genes, respectively (Table 3), products of which are involved in UV tolerance and mutagenic DNA repair (41). Three contiguous putative transcriptional repressors (*orf10*, *orf11*, and *orf12*) were also identified upstream of the T4SS locus.

Mapping the minimal replication region. A 5.7-kb EcoRI fragment (nt 460 to 6145) (Fig. 2A) includes the putative *repA* gene and the *cis*-acting elements and thus could contain the minimal replicon. To verify this conjecture, the fragment was cloned into the ColE1 vector pBluescript (resulting in pBTKE5.7), and subsequently a Cm^r cassette was added (because *T. kashmirensis* is resistant to ampicillin) to produce plasmid pBTKS (see the figure in the supplemental material). This plasmid, but not the vector, could be stably maintained in *T. kashmirensis* (transformation frequency, 4×10^{-5}). Thus, pBTKS comprises origin regions for both *E. coli* and *T. kashmirensis* and could be used as an efficient shuttle vector.

For mapping of the minimal replication region, constructs (pBTKAS, pBTK1C, pBTK2C, and pBTK3C) with deletion derivatives of the 5.7-kb fragment were prepared, and their transforming abilities in *T. kashmirensis* C1 were tested by electroporation (Fig. 2A). All the constructs except pBTK3C were successfully transformed at an average frequency of 2×10^{-5} to 6×10^{-5}. These deletions indicate that a 1,373-bp region (nt 2220 to 3592) spanning the *repA* gene and the upstream 611-bp intergenic region could serve as the minimal replicon. To confirm this, the region was first amplified using primers $OriV_R$ and $RepA_C$ and then cloned into the P15A *ori*-based vector pSD5B, containing a Km^r marker that could be used for its selection in *T. kashmirensis*. The resulting construct (pSD5MR), but not the vector alone, gave positive transformants in *T. kashmirensis* C1 (frequency, 2×10^{-5}) (Fig. 2A). In all the cases mentioned above, plasmid DNA was isolated from positive transformants and checked for its integrity. The fact that the plasmids suffered no distortion or recombination inside *T. kashmirensis* was proved by their similar restriction profiles and positive PCR amplification by insert-specific primers, before and after transformation (data not shown).

Localization of the replication origin, *oriV*. To localize the *cis*-acting *oriV*, primers $OriV_F$ and $OriV_R$ were used to amplify a 264-bp region (nt 3329 to 3592) spanning the iteron-like repeats; the product was subcloned into pSD5B (to give pSD5V), and its ability to replicate in *T. kashmirensis* C1 was tested. This construct failed to give any transformant of its own (frequency, $<10^{-9}$), but when it was electroporated into *T. kashmirensis* cells containing pBTKS (supplying the replication-related protein in *trans*), double-resistant (Km^r Cm^r) transformants were obtained at a frequency of 5×10^{-5} (Fig. 2A). An analysis of the plasmid DNA isolated from the transformants showed that both plasmids exist as self-replicating entities (data not shown). To avoid any incidence of plasmid loss due to incompatibility (since pSD5V and pBTKS share the same *oriV*), transformants were always selected and maintained in the presence of both the antibiotics. This result shows that the 264-bp fragment harbors the *oriV* and is activated when factors encoded by the replication region are supplied in *trans*.

Shuttle vector for diverse members of the *Alkaligenaceae*. To determine the range of hosts in which the shuttle vector could be used, it was electroporated into both phylogenetically related and phylogenetically distant hosts. pBTKS could be successfully transformed into the phylogenetically close members of the *Alkaligenaceae* (15), such as *Alcaligenes defragrans* (LMG 18538^T), *Alcaligenes faecalis* (ATCC 8750), *Achromobacter*

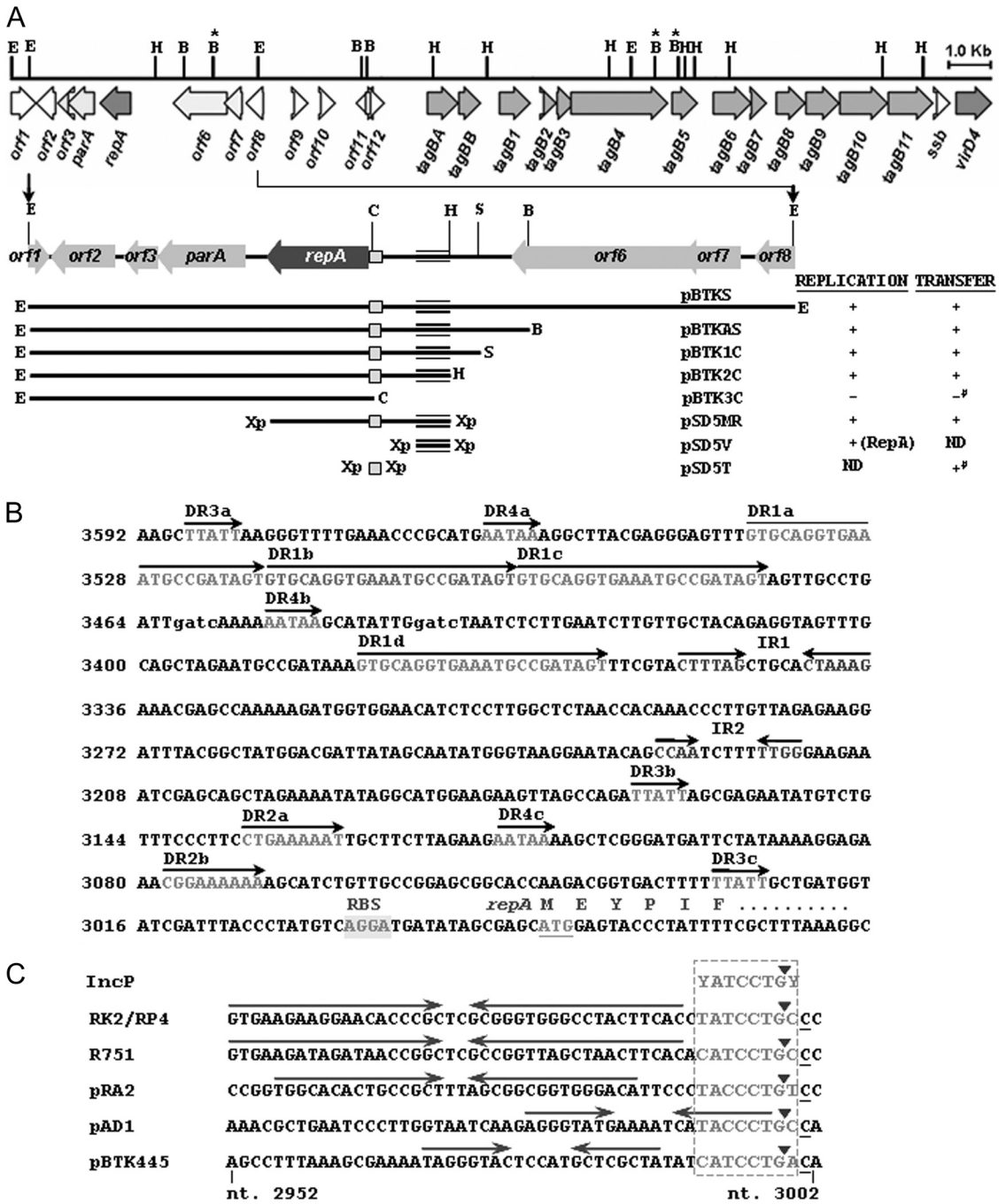

FIG. 2. (A) Identification of the minimal-replicon, *oriV*, and *oriT* regions of pBTK445. The 24,440-bp sequenced region of pBTK445 is drawn to scale as a thick horizontal line. ORFs are represented by arrows, with arrowheads indicating the transcriptional orientation. Vertical lines above the fragment represent the restriction sites (B, BamHI; B marked with an asterisk, BglII; C, ClaI; E, EcoRI; H, HindIII; S, SacII; Xp, XbaI), derived from the corresponding primers. The EcoRI and HindIII sites are used for the construction of banks, and the BglII sites are used for knockout mutants (see Materials and Methods). The 5.7-kb EcoRI fragment used for the characterization of the replication region is diagramed below. Restriction fragments incorporated into different constructs are indicated as horizontal lines, with the enzyme abbreviations at their two ends. In the "Replication" column, a plus sign indicates the ability of a plasmid to replicate in *T. kashmirensis*, and a minus sign indicates the inability to do so. RepA protein, when supplied in *trans*, is shown in parentheses. ND, not determined. The 264-bp minimal *oriV* is represented as triple horizontal lines. The mobilization abilities of the derivatives are given under the "Transfer" column. A plus sign indicates a mobilization frequency of $>10^{-5}$, and a minus sign indicates a mobilization frequency of $<10^{-7}$. Conjugative mating was performed with *T. kashmirensis* strains as the recipients. Experiments using *E. coli* as the recipient are marked with a pound sign. The 100-bp minimal *oriT* is represented by squares in the plasmid diagrams. Inferences regarding the replication and conjugation abilities of the constructs are drawn from three different experiments. (B) Organization of the *repA* gene and the corresponding *cis*-acting elements. The reverse complementary sequence of the region is shown (nt 2953 to 3592). DR and IR, direct and inverted repeats, respectively. Dam methylase recognition sites are represented by lowercase letters (gatc). (C) DNA sequence comparison of the *oriT* regions of RP4/RK2, R751, pRA2, pAD1, and pBTK445. The IncP-like *oriT* consensus sequence is boxed. Inverted repeats are marked by arrows. Solid arrowheads indicate the known or predicted *nic* (nick) sites.

TABLE 3. Predicted ORFs found in the partially sequenced genome of pBTK445

Position (start-stop)[a]	ORF name	No. of amino acids	Nearest homologous protein[b]	Functional notes	% Amino acid identity with nearest phylogenetic relative	Nearest phylogenetic relative (GenBank accession no.)
1–613	*orf1*	Partial				
1,101–631	*orf2*	156				
1416–1159	*orf3*	85				
2064–1420	*orf4*/*parA*	214	ParA	ATPase involved in plasmid partitioning	46	*Chlorobium limicola* plasmid pCL1 (NP_052163)
2981–2220	*orf5*/*repA*	253	RepA (belonging to RPA superfamily)	Plasmid replication initiator protein	52	*Rhodospirillum rubrum* genome (YP_425444)
5369–4053	*orf6*	438	DNA-directed DNA polymerase; subunit UmuC	Error-prone DNA repair	49	*Methylibium petroleiphilum* PM1 plasmid RPME01 (YP_001023066)
5749–5321	*orf7*	142	DNA-directed DNA polymerase; subunit UmuD	Error-prone DNA repair	57	*Methylibium petroleiphilum* PM1 plasmid RPME01 (YP_001023062)
6283–5858	*orf8*	141				
7046–7399	*orf9*	117				
7711–8076	*orf10*	121	Putative transcriptional regulator		32	*Shigella boydii* BS512 (ZP_00699160)
8838–8599	*orf11*	79	Prophage-related conserved transcriptional repressor protein		47	*Xylella fastidiosa* 9a5c (NP_297791)
8951–9298	*orf12*	115	Putative phage repressor		30	*Bordetella avium* (CAJ48906)
10371–22972	*tagB* locus (*orf13* to *orf25*)		T4SS homologs	Conjugal transfer and type 4 secretion	See Table 4	
23070–23393	*orf26*/*ssb*	107	ssDNA-binding protein	Controls activity of RecBCD nuclease	47	*Salmonella enterica* subsp. *enterica* (ZP_02353510)
23548–24537	*virD4*	Partial	TraN protein	VirD4-like protein	47	Plasmid pIPO2T (NP_444530)

[a] Ascending or descending nucleotide positions indicate that the gene is located on the positive or negative strand, respectively.
[b] ssDNA, single-stranded DNA.

xylosoxidans (ATCC 27061), *Pelistega europaea* (LMG 10982^T), *Taylorella equigenitalis* (LMG 6222^T), and *Pigmentiphaga kullae* (LMG 21665), at an average frequency of 2×10^{-6} to 4×10^{-6}. The plasmid, however, gave no transformants in the distantly related hosts tested, including the alphaproteobacteria *Pseudaminobacter salicylatoxidans* KCT001 and *Mesorhizobium thiogangeticum* (MTCC 7001^T) (frequency, $<10^{-9}$). These results suggest that the 5.7-kb fragment supports the replication of the plasmid only in related hosts and that the shuttle vector could be used in diverse members of the *Alkaligenaceae*.

To demonstrate a practical application for the constructed shuttle vector, an attempt was made to perform a complementation experiment. A sulfur oxidation mutant of *T. kashmirensis* (BD 004) with a single copy of Tn*5-mob* inserted into the *moeA* gene (a component of the molybdopterin cofactor biosynthesis [*moco*] locus) was used for this purpose (11). A 3.6-kb fragment (nt 442 to 4045; GenBank accession number EF017216) from the *moco* locus of the wild-type genome was introduced into the mutant strain using pBTKSM (which has the fragment cloned into pBTKS). The sulfur oxidation property of this complemented mutant was found to be similar to that of the wild type (data not shown). The shuttle vector could thus be used as a tool for the genetic engineering of *T. kashmirensis*.

Localization of the transfer origin, *oriT*. Based on comparison with the *oriT* regions of plasmids RK2/RP4, R751, pRA2, and pAD1, a potential IncP-like *oriT* consensus sequence (Y ATCCTGY) (3) was detected in pBTK445 (Fig. 2C) in close proximity to *oriV*, upstream of the *repA* gene (Fig. 2A). This region had a specific cleavage site called the "*nic*" site, preceded by an inverted-repeat sequence (Fig. 2C). The shuttle vector includes this *oriT*-like element and thus could support conjugative transfer, provided the transfer functions are supplied in *trans*. To test this, conjugation experiments were performed with *E. coli* SM10λ*pir* cells (providing RP4 "*tra*" functions from its genome in *trans*) harboring pBTKS as the donor and with the pBTK445-cured cells of the Km^r derivative of *T. kashmirensis* (BD 004 C1) as the recipient. Transconjugants were selected on kanamycin- and chloramphenicol-containing LA plates. The results show that the shuttle vector (but not pBluescript) was successfully transferred at a frequency of 3×10^{-5} (Fig. 2A). Deletion derivatives of the shuttle vector (pBTKAS, pBTK1C, pBTK2C, and pSD5MR) were also transferred with approximately the same efficiency, while the pBTK3C construct was not (Fig. 2A). Because this construct lacks the *oriV* in addition to the *oriT*-like region, its ability to be transferred to another *E. coli* strain (XL1-Blue) was tested. The plasmid was not transferred to the *E. coli* host, either (Fig. 2A), suggesting that the additional deleted region (as in pBTK2C) does contain the *cis*-acting element for transfer in addition to the *oriV*.

To prove that the *oriT*-like region is functional, it was amplified as a 100-bp (nt 2938 to 3037) PCR product using the primer pair $OriT_F$–$OriT_R$ and was cloned into pSD5B (which contains no *oriT*). Attempts were then made to mobilize the resulting construct, pSD5T, from *E. coli* SM10λ*pir* to *E. coli* XL1-Blue. The result shows that pSD5T (and not the empty vector) could be mobilized at a frequency of 2×10^{-4} (Fig. 2A), proving that the 100-bp fragment harbors the *oriT*.

The *repA* promoter. Bioinformatic screening revealed the presence of a promoter-like region, localized within the 264-bp

TABLE 4. Properties of pBTK445 *tagB* ORFs and their deduced products

Position (start-stop)	TagB ORF (no. of amino acid residues)	Location[a]	G+C content (%)	Homolog		% Amino acid identity with nearest phylogenetic relative	Nearest phylogenetic relative (GenBank accession no.)
				IncP	*Agrobacterium tumefaciens* or *Brucella suis*		
10371–11135	*tagBA* (254)	CP/PP	42	TraL		43	*Xylella fastidiosa* Ann-1 genome (ZP_00682673)
						42	*Verminephrobacter eiseniae* EF01-2, plasmid pVEIS01 (YP_980149)
11149–11697	*tagB* (182)	PP	48	TraR/TrbM		38	Conjugative broad-host-range plasmid pSB102 (NP_361032)
						35	Plasmid pKJK5 (YP_709155)
12162–12929	*tagB1* (255)	PP	52	TrbN	VirB1	55	*Xanthomonas axonopodis* pv. citri plasmid pXAC64 (NP_644764)
						44	*Xanthomonas campestris* pv. vesicatoria plasmid pXCV38 (YP_361544)
13265–13585	*tagB2* (106)	IM	50	TrbC	VirB2	32	*Aeromonas veronii* plasmid pAc3249A (ABI83638)
13604–14035	*tagB3* (143)	IM	45	TrbD	VirB3	34	*V. eiseniae* EF01-2 plasmid pVEIS01 (YP_980124)
13944–16358	*tagB4* (804)	OM	48	TrbE	VirB4	43	*Moraxella bovis* plasmid pMBO-2 (BAD83765)
						42	*V. eiseniae* EF01-2 plasmid pVEIS01 (YP_980123)
16462–17103	*tagB5* (213)	PP	45	TrbJ	VirB5	33	*X. axonopodis* pv. glycines plasmid pXAG81 (AAX12232)
17502–18443	*tagB6* (313)	IM	41	TrbL	VirB6	45	*M. bovis* plasmid pMBO-2 (BAD83763)
						34	*X. axonopodis* pv. citri plasmid pXAC64 (NP_644769)
18449–18820	*tagB7* (123)	IM	46	TrbK	VirB7		
19072–19791	*tagB8* (239)	PP/IM	50	TrbF	VirB8	42	*X. fastidiosa* Ann-1 ctg_247 (ZP_00682689)
19810–20634	*tagB9* (274)	OM	53	TrbG	VirB9	42	*X. fastidiosa* Ann-1 ctg_247 (ZP_00682690)
20645–21850	*tagB10* (401)	OM	55	TrbI	VirB10	42	*Bartonella henselae* (YP_034271)
21860–22972	*tagB11* (370)	CP	51	TrbB	VirB11	43	*X. fastidiosa* Ann-1 ctg_247 (ZP_00682692)

[a] IM, inner membrane; OM, outer membrane; CP, cytoplasmic; PP, periplasmic.

oriV-containing region, upstream of *repA*. To confirm that this promoter is functional, the recombinant plasmid pSD5V, mentioned above, was used, because in this construct the 264-bp segment is transcriptionally fused to the β-*gal* reporter gene. β-Galactosidase activity due to this element was then measured in *E. coli* XL1-Blue. While the promoterless vector pSD5B showed negligible activity, pSD5V exhibited a 75-fold increase in β-galactosidase activity (Fig. 3). This indicates that the 264-bp DNA fragment does contain a transcriptional promoter, which was designated *PrepA*. Contrary to expectation, however, the activity of *PrepA* remained unchanged when pBTK445 RepA was supplied in *trans* from the cotransformed inducible expression vector pQERA (Fig. 3). The results suggest that pBTK445 RepA has no regulatory effect on its own promoters, at least in *E. coli*.

A T4SS locus in pBTK445. Eleven genes located within a 10.8-kb region of pBTK445 were found to be homologous to the constituents of different T4SSs, including the TraII core of typical IncP plasmids (Fig. 4), and were designated *tagB* genes

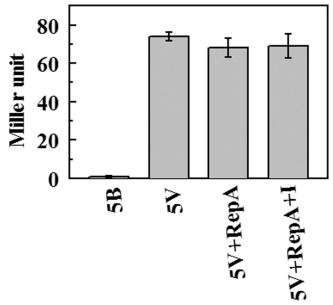

FIG. 3. Promoter assays. *E. coli* XL1-Blue transformed with the promoter probe vector pSD5B or the promoter construct pSD5V, either alone or in combination with the RepA expression vector pQERA, was assayed for the level of β-galactosidase activity under noninducing or inducing (+I) conditions. pSD5B alone (5B), pSD5V alone (5V), pSD5V plus pQERA (5V+RepA), and pSD5V plus pQERA induced by IPTG (5V+RepA+I) were tested. Activities are given in Miller units. Data are means for three different experiments. Error bars indicate standard errors.

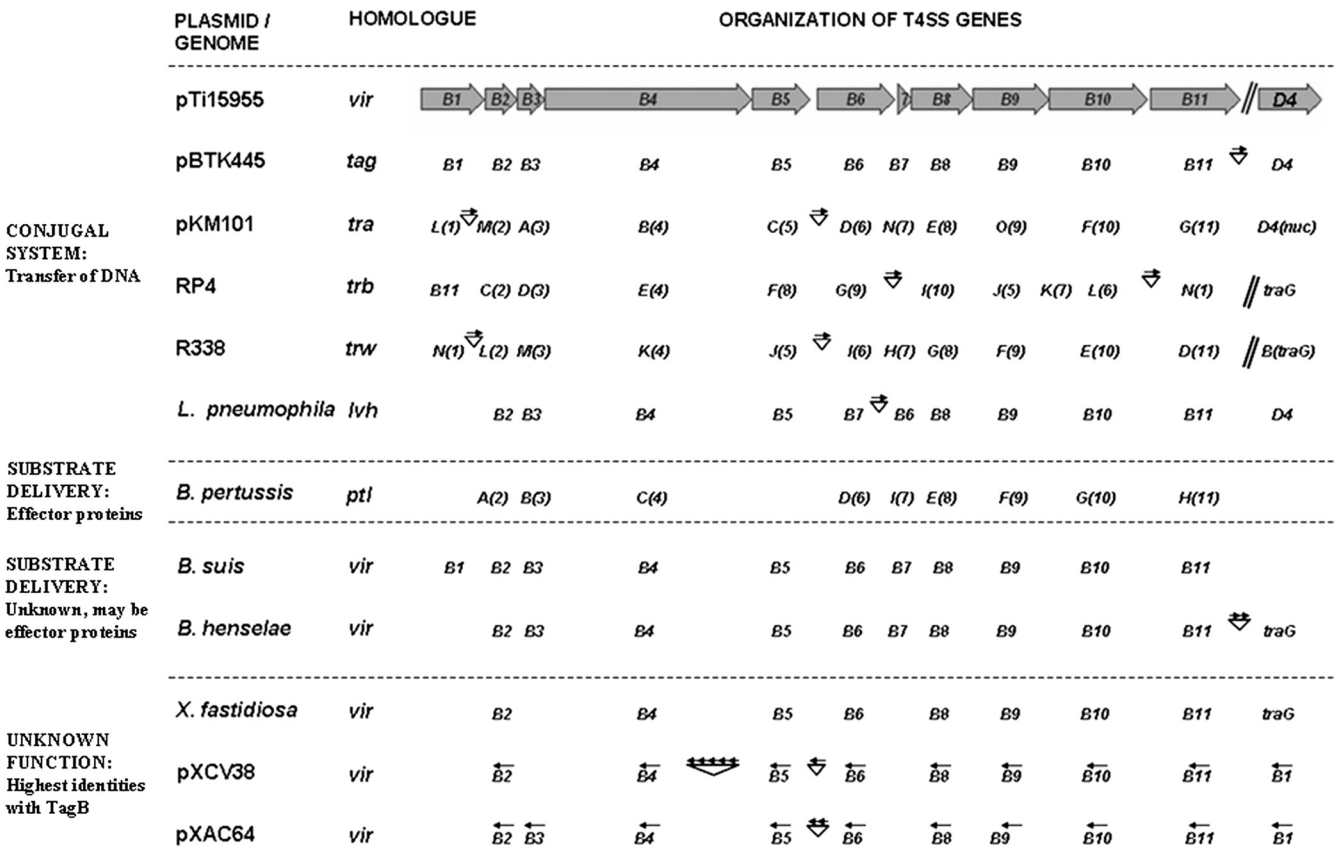

FIG. 4. Comparison of the T4SS homologs. The *virB* locus of the *A. tumefaciens* plasmid pTi15955 (X06826; numbers in parentheses herein refer to GenBank accession numbers), used as the reference, is diagramed at the top. Filled arrows represent genes, with arrowheads indicating their transcriptional orientations. The T4SS homologs are divided into four groups, separated by dashed lines. The presence of a particular homolog in a system is indicated by the name of the gene below the arrow for the corresponding reference gene. Unless otherwise indicated, the genes are transcribed in the same direction as those in pTi15955. Some systems are disrupted by one or many unrelated ORFs (represented by arrows above inverted triangles, with each arrow representing one ORF and the arrowheads indicating their orientations). The *tagB* genes of pBTK445 are included in the first group along with homologs from plasmids RP4 (M93696 and CAA38334) and R388 (X81123 and X63150), in addition to the genomic counterpart of *Legionella pneumophila* (Y19029). This group corresponds to systems known to transfer DNA between bacteria. The *B. pertussis* (BX640422) *ptl* system, known to deliver the pertussis toxin (substrate) to mammalian cells, is in the second group. The third group consists of systems whose substrates are presently unknown but are postulated to be effector proteins. These systems are represented by the genomic homologs of *B. suis* (AF141604) and *B. henselae* (BX897699). The last group consists of systems that have the highest identities with TagB and are available in the database without any functional information. These include the *virB* loci from the genome of *X. fastidiosa* (NZ_AAAM0300005), plasmid pXCV38 (AM039950) of *Xanthomonas campestris* pv. vesicatoria, and plasmid pXAC64 (AE008925) of *Xanthomonas axonopodis* pv. citri.

(transfer-associated genes homologous to *virB*). The predicted function and cellular localization of these gene products (Table 4) suggest that they might form a multicomponent channel extending from the bacterial membranes to the external environment (6, 9). The TagB system potentially encodes all the essential components that make up a T4SS. This includes the putative membrane proteins TagB8 through TagB10, homologs of which constitute the "core region" of the transporter (12, 23). The other putative core constituent, TagB7, shows insufficient sequence similarity with any known protein in the database and was annotated primarily on the basis of its genomic location between *tagB6* and *tagB8*. The predicted protein is larger (123 amino acids) than the established VirB7 members and shows a different transmembrane topology model in comparison to these VirB7 members (6). The system also encodes the predicted integral membrane protein TagB6, homologs of which are required for T4SS assembly and the passage of substrates to the cell exterior (20). The predicted energetic components include the two highly conserved nucleoside triphosphatases TagB4 and TagB11, homologs of which are known to energize the early substrate transfer reactions by an ATP-dependent mechanism (4). The putative peripheral pilus components include the TagB2 protein, which is probably associated with the inner membrane, the periplasmic TagB5 protein, and a homolog of the minor pilus subunit, TagB3 (6, 23, 24, 34). In addition, the system encodes the rarely found putative periplasmic protein TagB1, which possesses the characteristic lytic transglycosylase and goose egg white lysozyme domains (22), spanning 57 to 189 amino acids of its sequence. Homologs of this protein are capable of locally enlarging gaps in the peptidoglycan meshwork to allow the efficient assembly and anchoring of supramolecular transport complexes in the cell envelope (22). An ORF putatively encoding a single-stranded DNA binding protein (Ssb) is located downstream of *tagB11*,

followed by a *virD4* homolog, potentially encoding another energetic component of the system (4). Homologs of VirD4 proteins are also known to be necessary for the transfer of nucleoprotein particles by pathogenic T4SSs (10).

The T4SS constituents are known to coevolve and function in a highly concerted manner (6). In contrast, the TagB proteins show differential phylogenetic affinities. The putative core components (TagB8, TagB9, and TagB10) show maximum identities with homologs from chromosomally located T4SSs that have putative or proven pathogenic functions, including those from the genomes of *Xylella fastidiosa* (GenBank accession number NZ_AAAM03000051) and *Bartonella henselae* (GenBank accession number BX897699). On the other hand, the predicted peripheral TagB proteins were found to be more closely related to homologs from plasmid-encoded systems, plausibly having conjugative transfer functions, including those from the xanthomonad plasmids pXCV38 (GenBank accession number AM039950) and pXAC64 (GenBank accession number AE008925) (Table 4; Fig. 4). Thus, the TagB system appears to be an assembled locus comprising a mosaic of genes of discrete evolutionary lineages.

A homolog of *trbM*, a characteristic component associated with the TraII core of IncP plasmids (30, 31), is present upstream of the *tagB* locus and hence is designated *tagBB*. Although the TraI core present in typical IncP plasmids was not detected in the sequenced region of pBTK445, a solitary homolog of *traL*, an inherent component of the TraI core (31), was found upstream of *tagBB* and is designated *tagBA*.

Transfer functions conferred by the *tagB* locus. Since pBTK445 did not carry any selectable marker, essential for studying conjugative transfer, a Cm^r derivative of the plasmid (pBTKCΩ*6*) was constructed by homologous recombination. The plasmid was found to be efficiently and stably transferred from *T. kashmirensis* SR to BD 004 C1 cells at a frequency of 2×10^{-5}. However, another construct (pBTKCΩ*B4-B5*), in which the Cm^r cassette had been inserted into the *tagB4-tagB5* genes, resulting in the deletion of a 589-bp region from the terminal sequence of *tagB4* and the beginning of *tagB5* (nt 16033 to 16621), failed to conjugate (frequency, $<10^{-8}$) to and from designated *T. kashmirensis* cells. In addition, pBTKCΩ*6*, when used as a helper plasmid in triparental mating experiments, helped the transfer of the non-self-transmissible ColE1 plasmid pSUP5011 from *E. coli* strain DH5α to XL1-Blue at a frequency of 4×10^{-6}. However, pBTKCΩ*B4-B5*, when used as the helper plasmid, failed to mobilize pSUP5011 under similar experimental conditions. All these observations confirmed the involvement of the products of *tagB4* and *tagB5*, and consequently that of the *tagB* locus, in the DNA transfer process.

DISCUSSION

This paper presents the molecular characteristics of the replication and conjugation system of a new large plasmid, pBTK445, indigenous to *T. kashmirensis*. About 50% of the ~60-kb plasmid was sequenced, revealing that in terms of gene organization there exists a considerable degree of similarity between pBTK445 and typical IncP plasmids, such as RP4 and R721 (31). However, the proteins encoded by the ORFs of the new plasmid possess an overall identity level of only ~30% with those harbored by known IncP plasmids.

The replication region of this novel plasmid comprises a *repA* gene and an upstream iterated segment that functions as the *oriV*. Although this replication region has organizational similarities with those of other IncP plasmids, pBTK445 RepA shows no significant homology with its IncP counterparts. Rather, it possesses the highest sequence identity with certain functionally uncharacterized genome- and plasmid-derived replication proteins. In silico analysis suggested that pBTK445 RepA and its relatives belong to the RPA superfamily of proteins, which are thought to function by binding to single-stranded DNA (pfam10134). Although there is no significant information on the prokaryotic members of this family, substantial information is available for several eukaryotic versions, including those from humans (40, 43). These proteins have been implicated in the unwinding of double-stranded DNA, which is a prerequisite for the initiation of DNA replication. It is not unlikely that pBTK445 RepA and its related proteins function by destabilizing helices prior to replication in a manner similar to that of the eukaryotic RPAs. It would be interesting to study this possibility using the recombinant version of RepA reported here.

The *oriV* region also includes the promoter for *repA* transcription. This arrangement may suggest that RepA, apart from providing a replication function, also transcriptionally controls its own synthesis (1, 8). However, we could not detect any effect of RepA on the activity of its promoter, which may be due to the fact that a heterologous host (*E. coli*) was used. However, in its native host, *T. kashmirensis*, where all ancillary factors would be available, pBTK445 RepA-mediated regulation of promoter activity could well be possible, albeit not necessarily in the same manner as that generally accepted for IncP family members. Again, preliminary binding experiments using electrophoretic mobility shift assays showed that pBTK445 RepA does not bind to the origin/promoter region (unpublished data). This is not surprising in view of the fact that the class of proteins in question does not bind efficiently to double-stranded DNA. Whatever the mechanism of pBTK445 replication may be, its minimal replication region could be used to construct shuttle vectors for diverse members of the *Alcaligenaceae*.

Although the sequence signatures of the identified *oriV* and *oriT* resemble those of IncP, their genomic localizations are totally different. These two *ori* sequences of the plasmid occur in close proximity to each other, upstream of *repA*. In other well-known conjugative systems, the *oriT* is found in close association with the genes encoding the DNA-processing enzymes, such as the Mob (RK2, pRA2), TraI (RP4), or Dtr (pCC31) system, whose products are required for the nicking and subsequent transfer of DNA (13, 29, 45). The location of *oriT* in isolation from genes known to be involved in its activation is intriguing and suggests that the transfer origin of this plasmid may function in a manner that is at least partly different from that demonstrated earlier. Since origins require DNA melting, it is possible that the binding of RepA to *oriV* may also promote nicking at *oriT*, either directly or indirectly. However, there must be some mechanism by which the functioning of the two origins is regulated, and this needs to be investigated. Again, in view of the fact that the new plasmid is considerably

divergent from typical IncP members, it was interesting that RP4 transfer functions (supplied in *trans*) could mobilize otherwise nonmobilizable plasmids, such as pSD5B, when the pBTK445 *oriT* was cloned into them.

Apart from the functions described above, the presence of three consecutive genes encoding transcriptional repressor-like proteins (*orf10* to *orf12*) upstream of the *tagB* locus is noteworthy. The UV repair-related locus might confer survival advantages on *T. kashmirensis*.

The TagB system shares different phylogenomic features with different T4SSs that have diverse functional attributes. Although the new system is involved in conjugation, the *tagB* genes are more similar to the *virB* homologs of *Agrobacterium tumefaciens* and *Brucella suis* in terms of their genomic organization than to the components of the TraII core of IncP plasmids (Fig. 4). The presence of a *virD4* homolog in close association with the *tagB* locus suggests that transfer of nucleoprotein particles to eukaryotic cells might also be a function of this system (10). The fact that *T. kashmirensis* is a close relative of *Bordetella pertussis*, which has a well-characterized T4SS (the *ptl* system) involved in the transfer of the pertussis toxin, supports such a possibility.

The predicted products of the *tagB* genes have maximum identity with the putative products of T4SSs that are discontinuously organized and are frequently flanked by signatures of lateral gene transfer, such as insertion elements, transposases, and tRNA sequences (Fig. 4, last group). Thus, it appears that although the TagB system has acquired many of its genes, especially the peripheral constituents, from sources similar to those of the T4SSs in the xanthomonad plasmids pXCV38 and pXAC64, only the pBTK445-borne locus has been able to put together a full *virB*-like operon, while the xanthomonad plasmids have apparently failed to do so. Hence, a comprehensive study involving TagB and these fragmented systems would help us understand the phylogenomics and evolution of T4SSs better. The system is likely to have evolved to its present state in fairly recent times, as indicated by the presence of a VirB11 homolog, which is generally present in more recently evolved systems (6). The T4SSs that have been characterized in detail have mostly come from parasitic or symbiotic organisms. The present study of the pBTK445-borne T4SS, the conjugative transfer functions of which are described here in detail, is perhaps the first molecular study of such a system from a free-living, autotrophic soil bacterium. Moreover, the *tagB* locus of pBTK445 can carry out transfer functions in *trans* and facilitate the conjugation of other, unrelated plasmids, an ability that endows such conjugative plasmids with immense potential to effect widespread horizontal gene transfer, genome reorganization, and genome innovation in nature.

ACKNOWLEDGMENTS

B.D. and W.G. are grateful to the CSIR, Government of India, and the DST (Bose Institute), Government of India, respectively, for their fellowships.

This work is dedicated to the memory of the late Pradosh Roy.

Mrinal Das is acknowledged for assistance during DNA sequencing. We thank Anne O. Summers, Max Mergeay, and P. K. Chakraborty for valuable suggestions and encouragement from time to time. Jaydeep Chakraborty extended unconditional help throughout the work. Plasmid pFH450 was a generous gift from Finbarr Hayes.

REFERENCES

1. **Abeles, A. L., K. M. Snyder, and D. K. Chattoraj.** 1984. P1 plasmid replication: replicon structure. J. Mol. Biol. **173:**307–324.
2. **Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:**3389–3402.
3. **An, F. Y., and D. B. Clewell.** 1997. The origin of transfer (*oriT*) of the enterococcal, pheromone-responding, cytolysin plasmid pAD1 is located within the *repA* determinant. Plasmid **37:**87–94.
4. **Atmakuri, K., E. Cascales, and P. J. Christie.** 2004. Energetic components VirD4, VirB11 and VirB4 mediate early DNA transfer reactions required for bacterial type IV secretion. Mol. Microbiol. **54:**1199–1211.
5. **Burns, D. L.** 1999. Biochemistry of type IV secretion. Curr. Opin. Microbiol. **2:**25–29.
6. **Cao, T. B., and M. H. Saier, Jr.** 2001. Conjugal type IV macromolecular transfer systems of Gram-negative bacteria: organismal distribution, structural constraints and evolutionary conclusions. Microbiology **147:**3201–3214.
7. **Cascales, E., and P. J. Christie.** 2003. The versatile bacterial type IV secretion systems. Nat. Rev. Microbiol. **1:**137–149.
8. **Chattoraj, D. K., K. M. Snyder, and A. L. Abeles.** 1985. P1 plasmid replication: multiple functions of RepA protein at the origin. Proc. Natl. Acad. Sci. USA **82:**2588–2592.
9. **Christie, P. J.** 2001. Type IV secretion: intercellular transfer of macromolecules by systems ancestrally related to conjugation machines. Mol. Microbiol. **40:**294–305.
10. **Covacci, A., J. L. Telford, G. Del Giudice, J. Parsonnet, and R. Rappuoli.** 1999. *Helicobacter pylori* virulence and genetic geography. Science **284:**1328–1333.
11. **Dam, B., S. Mandal, W. Ghosh, S. K. Das Gupta, and P. Roy.** 2007. The S4-intermediate pathway for the oxidation of thiosulfate by the chemolithoautotroph *Tetrathiobacter kashmirensis* and inhibition of tetrathionate oxidation by sulfite. Res. Microbiol. **158:**330–338.
12. **de Paz, H. D., F. J. Sangari, S. Bolland, J. M. Garcia-Lobo, C. Dehio, F. de la Cruz, and M. Llosa.** 2005. Functional interactions between type IV secretion systems involved in DNA transfer and virulence. Microbiology **151:**3505–3516.
13. **Fürste, J. P., W. Pansegrau, G. Ziegelin, M. Kroger, and E. Lanka.** 1989. Conjugative transfer of promiscuous IncP plasmids: interaction of plasmid-encoded products with the transfer origin. Proc. Natl. Acad. Sci. USA **86:**1771–1775.
14. **Ghosal, D., I. S. You, D. K. Chatterjee, and A. M. Chakrabarty.** 1985. Genes specifying degradation of 3-chlorobenzoic acid in plasmids pAC27 and pJP4. Proc. Natl. Acad. Sci. USA **82:**1638–1642.
15. **Ghosh, W., A. Bagchi, S. Mandal, B. Dam, and P. Roy.** 2005. *Tetrathiobacter kashmirensis* gen. nov., sp. nov., a novel mesophilic, neutrophilic, tetrathionate-oxidizing, facultatively chemolithotrophic betaproteobacterium isolated from soil from a temperate orchard in Jammu and Kashmir, India. Int. J. Syst. Evol. Microbiol. **55:**1779–1787.
16. **Hayes, F.** 1998. A family of stability determinants in pathogenic bacteria. J. Bacteriol. **180:**6415–6418.
17. **Hayes, F.** 2000. The partition system of multidrug resistance plasmid TP228 includes a novel protein that epitomizes an evolutionarily distinct subgroup of the ParA superfamily. Mol. Microbiol. **37:**528–541.
18. **Hunter, W. J., and D. K. Manter.** 2008. Bio-reduction of selenite to elemental red selenium by *Tetrathiobacter kashmirensis*. Curr. Microbiol. **57:**83–88.
19. **Jain, S., D. Kaushal, S. K. DasGupta, and A. K. Tyagi.** 1997. Construction of shuttle vectors for genetic manipulation and molecular analysis of mycobacteria. Gene **190:**37–44.
20. **Jakubowski, S. J., V. Krishnamoorthy, E. Cascales, and P. J. Christie.** 2004. *Agrobacterium tumefaciens* VirB6 domains direct the ordered export of a DNA substrate through a type IV secretion system. J. Mol. Biol. **341:**961–977.
21. **Kamachi, K., M. Sota, Y. Tamai, N. Nagata, T. Konda, T. Inoue, E. M. Top, and Y. Arakawa.** 2006. Plasmid pBP136 from *Bordetella pertussis* represents an ancestral form of IncP-1β plasmids without accessory mobile elements. Microbiology **152:**3477–3484.
22. **Koraimann, G.** 2003. Lytic transglycosylases in macromolecular transport systems of Gram-negative bacteria. Cell. Mol. Life Sci. **60:**2371–2388.
23. **Krall, L., U. Wiedemann, G. Unsin, S. Weiss, N. Domke, and C. Baron.** 2002. Detergent extraction identifies different VirB protein subassemblies of the type IV secretion machinery in the membranes of *Agrobacterium tumefaciens*. Proc. Natl. Acad. Sci. USA **99:**11405–11410.
24. **Lai, E. M., and C. I. Kado.** 1998. Processed VirB2 is the major subunit of the promiscuous pilus of *Agrobacterium tumefaciens*. J. Bacteriol. **180:**2711–2717.
25. **Lee, C. L., D. S. Ow, and S. K. Oh.** 2006. Quantitative real-time polymerase chain reaction for determination of plasmid copy number in bacteria. J. Microbiol. Methods **65:**258–267.
26. **Mandal, S., S. Chatterjee, B. Dam, P. Roy, and S. K. Das Gupta.** 2007. The dimeric repressor SoxR binds cooperatively to the promoter(s) regulating expression of the sulfur oxidation (*sox*) operon of *Pseudaminobacter salicylatoxidans* KCT001. Microbiology **153:**80–91.

27. **Miller, V. L., and J. J. Mekalanos.** 1988. A novel suicide vector and its use in construction of insertion mutations: osmoregulation of outer membrane proteins and virulence determinants in *Vibrio cholerae* requires *toxR*. J. Bacteriol. **170:**2575–2583.
28. **Motallebi-Veshareh, M., D. A. Rouch, and C. M. Thomas.** 1990. A family of ATPases involved in active partitioning of diverse bacterial plasmids. Mol. Microbiol. **4:**1455–1463.
29. **Pansegrau, W., and E. Lanka.** 1991. Common sequence motifs in DNA relaxases and nick regions from a variety of DNA transfer systems. Nucleic Acids Res. **19:**3455.
30. **Pansegrau, W., and E. Lanka.** 1996. Enzymology of DNA transfer by conjugative mechanisms. Prog. Nucleic Acid Res. Mol. Biol. **54:**197–251.
31. **Pansegrau, W., E. Lanka, P. T. Barth, D. H. Figurski, D. G. Guiney, D. Haas, D. R. Helinski, H. Schwab, V. A. Stanisich, and C. M. Thomas.** 1994. Complete nucleotide sequence of Birmingham IncPα plasmids: compilation and comparative analysis. J. Mol. Biol. **239:**623–663.
32. **Riccio, M. L., L. Pallecchi, R. Fontana, and G. M. Rossolini.** 2001. In*70* of plasmid pAX22, a bla_{VIM-1}-containing integron carrying a new aminoglycoside phosphotransferase gene cassette. Antimicrob. Agents Chemother. **45:** 1249–1253.
33. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
34. **Schmidt-Eisenlohr, H., N. Domke, and C. Baron.** 1999. TraC of IncN plasmid pKM101 associates with membranes and extracellular high-molecular-weight structures in *Escherichia coli*. J. Bacteriol. **181:**5563–5571.
35. **Simon, R.** 1984. High frequency mobilization of gram-negative bacterial replicons by the in vitro constructed Tn*5*-*mob* transposon. Mol. Gen. Genet. **196:**413–420.
36. **Skorupski, K., and R. K. Taylor.** 1996. Positive selection vectors for allelic exchange. Gene **169:**47–52.
37. **Skulj, M., V. Okrslar, S. Jalen, S. Jevsevar, P. Slanc, B. Strukelj, and V. Menart.** 2008. Improved determination of plasmid copy number using quantitative real-time PCR for monitoring fermentation processes. Microb. Cell Fact. **7:**6.
38. **Uraji, M., K. Suzuki, and K. Yoshida.** 2002. A novel plasmid curing method using incompatibility of plant pathogenic Ti plasmids in *Agrobacterium tumefaciens*. Genes Genet. Syst. **77:**1–9.
39. **Vedler, E., M. Vahter, and A. Heinaru.** 2004. The completely sequenced plasmid pEST4011 contains a novel IncP1 backbone and a catabolic transposon harboring *tfd* genes for 2,4-dichlorophenoxyacetic acid degradation. J. Bacteriol. **186:**7161–7174.
40. **Wold, M. S.** 1997. Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. Annu. Rev. Biochem. **66:**61–92.
41. **Woodgate, R., and S. G. Sedgwick.** 1992. Mutagenesis induced by bacterial UmuDC proteins and their plasmid homologues. Mol. Microbiol. **6:**2213–2218.
42. **Wübbeler, J. H., T. Lutke-Eversloh, S. Van Trappen, P. Vandamme, and A. Steinbuchel.** 2006. *Tetrathiobacter mimigardefordensis* sp. nov., isolated from compost, a betaproteobacterium capable of utilizing the organic disulfide 3,3′-dithiodipropionic acid. Int. J. Syst. Evol. Microbiol. **56:** 1305–1310.
43. **Wyka, I. M., K. Dhar, S. K. Binz, and M. S. Wold.** 2003. Replication protein A interactions with DNA: differential binding of the core domains and analysis of the DNA interaction surface. Biochemistry **42:**12909–12918.
44. **Wyndham, R. C., R. K. Singh, and N. A. Straus.** 1988. Catabolic instability, plasmid gene deletion and recombination in *Alcaligenes* sp. BR60. Arch. Microbiol. **150:**237–243.
45. **Ziegelin, G., J. P. Fürste, and E. Lanka.** 1989. TraJ protein of plasmid RP4 binds to a 19-base pair invert sequence repetition within the transfer origin. J. Biol. Chem. **264:**11989–11994.