JOURNAL OF BACTERIOLOGY, Feb. 2007, p. 761–771 Vol. 189, No. 3
0021-9193/07/$08.00+0 doi:10.1128/JB.01327-06

Novel Type IV Secretion System Involved in Propagation of Genomic Islands▽

Mario Juhas,[1]* Derrick W. Crook,[1] Ioanna D. Dimopoulou,[1] Gerton Lunter,[2] Rosalind M. Harding,[3] David J. P. Ferguson,[4] and Derek W. Hood[5]

Clinical Microbiology and Infectious Diseases, Nuffield Department of Clinical Laboratory Sciences, University of Oxford, OX3 9DU Oxford, United Kingdom[1]; MRC Functional Genetics Unit, Department of Physiology, Anatomy and Genetics, University of Oxford, OX1 3PT Oxford, United Kingdom[2]; Departments of Zoology and Statistics, University of Oxford, OX1 3TG Oxford, United Kingdom[3]; Department of Pathology, University of Oxford, OX3 9DU Oxford, United Kingdom[4]; and Molecular Infectious Diseases Group, The Weatherall Institute of Molecular Medicine, University of Oxford, OX3 9DS Oxford, United Kingdom[5]

Received 21 August 2006/Accepted 10 November 2006

Type IV secretion systems (T4SSs) mediate horizontal gene transfer, thus contributing to genome plasticity, evolution of infectious pathogens, and dissemination of antibiotic resistance and other virulence traits. A gene cluster of the *Haemophilus influenzae* genomic island ICE*Hin1056* has been identified as a T4SS involved in the propagation of genomic islands. This T4SS is novel and evolutionarily distant from the previously described systems. Mutation analysis showed that inactivation of key genes of this system resulted in a loss of phenotypic traits provided by a T4SS. Seven of 10 mutants with a mutation in this T4SS did not express the type IV secretion pilus. Correspondingly, disruption of the genes resulted in up to 100,000-fold reductions in conjugation frequencies compared to those of the parent strain. Moreover, the expression of this T4SS was found to be positively regulated by one of its components, the *tfc24* gene. We concluded that this gene cluster represents a novel family of T4SSs involved in propagation of genomic islands.

The horizontal gene pool contributes to the diversification and adaptation of microorganisms. A major part of the horizontal gene pool consists of genomic islands. Genomic islands are well-characterized segments of bacterial genomes, often inserted at tRNA genes, that contain homologues of genes encoding integrases and other genes associated with conjugative plasmids or phages (17).

Prior to the 1970s *Haemophilus influenzae* was universally susceptible to ampicillin. In 1972, the first ampicillin-resistant isolate was detected, and soon after this strains resistant to tetracycline, chloramphenicol, erythromycin, and multiple antibiotics were identified and spread rapidly around the globe. Work over the past few years has provided evidence that horizontal transfer of genes in bacteria, including transfer of antibiotic resistance, is facilitated by genomic islands.

Genomic islands of many bacterial plant and animal pathogens encode type IV secretion systems (T4SSs) which are preferentially used for delivery of bacterial effector proteins across the bacterial membrane and the plasmatic membrane into the eukaryotic host cells (4, 18, 24, 28). T4SSs also mediate horizontal gene transfer, thus contributing to genome plasticity, the evolution of infectious pathogens, and dissemination of antibiotic resistance and other virulence traits (9, 22). The structures of the genetic determinants of T4SSs vary and consist of multiple genes organized into a single functional unit. These structures have been classified into major types based on a combination of gene content and shared homology. Hitherto, two different grouping schemes and nomenclatures have been used by investigators, as described in recent reviews. In one classification there are three major types, referred to as types F, P, and I, and these types associate with model conjugation systems described for plasmids F, RP4, and R64, respectively. In the other classification, types F and P are grouped together as type IVA and type I is type IVB. A third group in this classification is composed of "other" T4SS representatives (9, 10, 13, 32).

Hitherto, a major unresolved feature of genomic islands was the mechanism by which they are transferred between bacteria. The most widely held view is that genomic islands represent mobile elements, such as phage or conjugative plasmids that have either lysogenized or cointegrated with the chromosome, and that their transfer functions have become degenerate (17). This hypothesis has been favored by a number of investigators; however, observations of a family of syntenic genomic islands with deep evolutionary relationships have challenged this hypothesis based on findings for two members of this family, ICE*Hin1056* and the *clc* element. It is recognized that both ICE*Hin1056* and the *clc* element are capable of integration into the chromosome of the host, excision, and self-transfer to a new host and reintegration (15, 16, 49).

Preliminary analysis of in silico data acquired from sequencing projects suggested that there is a highly conserved module of genes that is responsible for horizontal transfer of these genomic islands; however, no functional analysis of the components of this putative transfer module was performed, and there was no in-depth description. A better understanding of genes involved in conjugative transfer and their relationship to well-characterized conjugative systems should provide a better

* Corresponding author. Mailing address: Clinical Microbiology and Infectious Diseases, Nuffield Department of Clinical Laboratory Sciences, University of Oxford, Headington, OX3 9DU Oxford, United Kingdom. Phone: (0044) 01865 222194. Fax: (0044) 01865 222195. E-mail: mario.juhas@ndcls.ox.ac.uk.

▽ Published ahead of print on 22 November 2006.

TABLE 1. Bacterial strains and plasmids used in this work

Strain or plasmid	Characteristics	Reference or source
H. influenzae strains		
Rd	Wild type, standard laboratory strain	20
Rd11	Rd harboring ICE*Hin1056*	This study
Rd11*tfc1*	*tfc1* mutant of Rd11	This study
Rd11*tfc2*	*tfc2* mutant of Rd11	This study
Rd11*tfc4*	*tfc4* mutant of Rd11	This study
Rd11*tfc6*	*tfc6* mutant of Rd11	This study
Rd11*tfc12*	*tfc12* mutant of Rd11	This study
Rd11*tfc14*	*tfc14* mutant of Rd11	This study
Rd11*tfc16*	*tfc16* mutant of Rd11	This study
Rd11*tfc22*	*tfc22* mutant of Rd11	This study
Rd11*tfc23*	*tfc23* mutant of Rd11	This study
Rd11*tfc24*	*tfc24* mutant of Rd11	This study
E. coli DH5α		Lab collection
Plasmids		
pGEM-T Easy	Cloning vector, f1ori, *lacZ*, Amp^r	Promega
pGEM*tfc1*	pGEM-T Easy carrying disrupted *tfc1* gene	This study
pGEM*tfc2*	pGEM-T Easy carrying disrupted *tfc2* gene	This study
pGEM*tfc4*	pGEM-T Easy carrying disrupted *tfc4* gene	This study
pGEM*tfc6*	pGEM-T Easy carrying disrupted *tfc6* gene	This study
pGEM*tfc12*	pGEM-T Easy carrying disrupted *tfc12* gene	This study
pGEM*tfc14*	pGEM-T Easy carrying disrupted *tfc14* gene	This study
pGEM*tfc16*	pGEM-T Easy carrying disrupted *tfc16* gene	This study
pGEM*tfc22*	pGEM-T Easy carrying disrupted *tfc22* gene	This study
pGEM*tfc23*	pGEM-T Easy carrying disrupted *tfc23* gene	This study
pGEM*tfc24*	pGEM-T Easy carrying disrupted *tfc24* gene	This study

understanding of how at least one family of genomic islands is mobilized in bacteria and may suggest that many more genomic islands than currently recognized contain genes for self-mobilization.

Here we describe identification and functional analysis of a cluster of genes encoding an uncharacterized T4SS present in ICE*Hin1056*, an example of a diverse family of genomic islands. Within this gene cluster are homologues of genes in previously defined T4SSs. We show that this novel T4SS is responsible for formation of the conjugative pilus and the resulting conjugative transfer of ICE*Hin1056*. In this study we not only identified a new lineage of T4SSs but also demonstrated that a highly evolved and conserved cluster of genes is involved in the mobilization of genomic islands.

MATERIALS AND METHODS

Bacterial strains, plasmids, and growth conditions. All bacterial strains and plasmids used in this study are listed in Table 1. In most cases *H. influenzae* was grown on HIB medium (Columbia agar containing 15 μg/ml NAD and 15 μg/ml hemin). When required, this medium was supplemented with kanamycin (10 μg/ml), tetracycline (2 μg/ml), or ampicillin (4 μg/ml). All *Haemophilus* plate cultures were grown for 24 to 48 h at 37°C in an atmosphere containing 5% CO_2. Liquid cultures of *H. influenzae* were grown in brain heart infusion broth (BHI) supplemented with NAD (10 μg/ml), hemin (15 μg/ml), and, when necessary, antibiotics at the concentrations described above and incubated at 200 rpm on a rotatory shaker at 37°C. Luria-Bertani broth was routinely used for growing *Escherichia coli* strains. When appropriate, Luria-Bertani medium was supplemented with ampicillin (50 μg/ml), kanamycin (50 μg/ml), or gentamicin (5 μg/ml).

PCR amplification and recombinant DNA methodology. Restriction endonucleases and DNA-modifying enzymes were obtained from New England Biolabs and were used according to the manufacturer's instructions. Recombinant DNA techniques and plasmid DNA preparation were performed as described by Sambrook et al. (45). Oligonucleotide primers were synthesized by Operon. Standard PCR amplifications were performed in 10-μl mixtures using *Taq* DNA polymerase or ProofStart DNA polymerase (QIAGEN) according to the supplier's instructions. Four-kilobase DNA fragments comprising the putative transfer region of ICE*Hin1056* were cloned into pGEM-T Easy (Promega). Recombinant plasmids were confirmed by PCR amplification and restriction endonuclease digestion. Genes of the putative type IV secretion system (*tfc1*, *tfc2*, *tfc4*, *tfc6*, *tfc12*, *tfc14*, *tfc16*, *tfc22*, *tfc23*, and *tfc24*) were inactivated by inserting a kanamycin resistance cassette and the *H. influenzae* DNA uptake signal sequence (47). These genes were chosen primarily because they exhibit homology to genes of other T4SSs; however, one of these genes, *tfc19*, a homologue of *traG*, lacked suitable restriction sites for insertional mutagenesis. An additional four genes in the module were inactivated because of well-placed unique restriction sites.

Generation of *H. influenzae* mutant strains. For construction of mutant strains by reciprocal recombination, 2 to 3 μg of the appropriate linearized plasmid DNA construct was used to transform *H. influenzae* strain Rd11 (Table 1), and transformants were selected with kanamycin (27). All transformants were checked by reculturing on HIB medium containing kanamycin and were subsequently confirmed to be mutants by PCR amplification. The PCR results were consistent with the hypothesis that only one copy of ICE*Hin1056* was present in each bacterium.

Conjugal transfer of the ICE*Hin1056* genomic island. The transfer efficiency of ICE*Hin1056* was determined using a modification of the method of Stuy (48). Donor or recipient cells were grown for 48 h on HIB agar, and approximately 10^8 cells were scraped off each plate and resuspended in 1 ml of BHI broth. Ten microliters of the suspension of donor cells and 100 μl of the suspension of recipient cells were gently mixed to obtain a ratio of donor cells to recipient cells of 1:10. Subsequently, this mixture was spread in the center of antibiotic-free HIB agar plates. The inoculum was allowed to dry prior to incubation for 6 h. The cells were then harvested by flooding the plate with 1 ml of BHI broth and resuspending them using a spreading paddle. Serial dilutions were plated to determine the viable counts and to determine the number of transconjugants, donors, and recipients using agar containing the appropriate selective antibiotic. Recipients and transconjugants were purified on agar plates containing tetracycline (2 μg/ml) and on agar plates containing kanamycin (10 μg/ml) plus tetracycline (2 μg/ml), respectively. Experiments were carried out in triplicate, and the mean and standard error for each strain were calculated.

Electron microscopy. To examine *H. influenzae* strains for the presence of the type IV secretion-dependent pilus, bacteria were grown overnight and then allowed to conjugate on antibiotic-free HIB agar plates for 6 h. The conjugating

TABLE 2. Oligonucleotide primers and TaqMan probes used in RT-PCR analysis of gene expression

Primer	Sequence (5′→3′)	Binding position within ICE*Hin1056* (bp)
RTTF1F	TTATGGTTTACCTGGGC	27426–27442
RTTF1R	CGACCACAACAGTTACAACAC	27590–27610
RTTF2F	GCTGATCGTGTACTGCAATATG	28983–29004
RTTF2R	GAAAGCAGCCCCTAATCCATATCC	30610–30633
RTTF3F	GTGGTGCGGGAATGCAGGTAAC	32057–32078
RTTF3R	GGTTGCTATATCTTGACTTGGCGC	32548–32571
RTTF4F	GGTGGCAATCTGCATTTTCTG	32522–32542
RTTF4R	GCACCCATAATGACAACATAGA	33742–33763
RTTF5F	TGGCGCTATAAGTGCTCAAG	33943–33962
RTTF5R	GATAACCCGCCATACACACT	34966–34985
RTTF6F	GAATTGCTATCGGTGCATTGC	35344–35364
RTTF6R	GCAGATGCCGTCATTGGTG	35980–35998
RTTF7F	CTGCATTGCCTGTGCGA	36499–36515
RTTF7R	ATTATGCTGAAATGTTGCTGC	36633–36653
RTTF8F	GGTTGTACAAGCTGGTAATC	37218–37237
RTTF8R	CCAGAGACAATTGCGCCTTC	37687–37706
RTTF9F	GCAGACTGGTTCCGTCAG	38053–38070
RTTF9R	AACCTGGTACAGGCATTTGCTCTG	38540–38563
RTTF10F	GATAATATTCGTAAGCAGCCAG	39646–39667
RTTF10R	GAATAGCCGAGCGGTCAG	40634–40651
RTTF11F	CATTTCCTCAGCGTGTATTTAC	41664–41685
RTTF11R	GTCGTAAATGGCGCATAGCC	42619–42638
RTTF12F	CCAGTGTTATATTCACGTAGCTC	45686–45708
RTTF12R	AGAAGATCCCGATGCTCAAGTG	45949–45970
RTTF13F	CAGCCGAATTTTGTACAAAATAAC	47761–47784
RTTF13R	GAGCTTGCAAGCCAGGTG	48254–48271
RTTF1F2	CCCGATGGCTTTACGTGATGC	28295–28315
RTTF1R2	GCAGCTCTCGCTCCATGCGG	28819–28839
RTTF2F2	GGATTAATGATGCGTGACCGTCG	33514–33536
RTTF2R2	CCCTGCCTGCACAGTAGTAATATC	34097–34120
RTTF3F2	CTCTGTGAGGTCAAGATGTC	34891–34910
RTTF4F2	CAAGGTCGATTCTATGCAGC	36618–36637
RTTF5F2	CCCGTACAGCAGAAAACGAGG	38447–38467
RTTF5R2	CGCTGCTTCTTCCCTCAGTTCC	38974–38995
RTTF6F2	GGGGATTGCAGTCAATGACG	41478–41497
RTTF7F2	GACATTAGCCCCTCTTATCCC	42516–42536
RTTF7R2	CATAAGCGTAAATAGAGATA AAAT	44556–44579
RTTF8F2	CTCCGGCCAAAATTGGACTC	44017–44036
RTTF8R2	CACAATAGTAACAGTATTGAAG AGCC	44675–44700
RTTF9F2	GCACTACCCGCTTTACAGCTCACC	44615–44638
RTTF9R2	CCAGGCATCTGCGCAATCAGG	45104–45124
RTTF10F2	GAAGCTGATACTGGCATGG	46916–46934
RTTF11F2	GCTGAGCTGCTACACTAGGATTCG	48175–48198
RTTF11R2	GCCCAAGTGTTTTCCCACACG	48844–48864
RTTF1F3	GGGGATTCATCAGCACTTTGATTG	26625–26648
RTTF1R3	ATGCGGTTGCTATCATCCAAGTG	27228–27250
RTTF2F3	CCTGGTGTCGGGTGATAGG	27131–27149
RTTF2R3	TTGTTGCTCATCGGCTGTTTG	27915–27935
RTTF3F3	GGCGGTGGTAATGGTATCAATGC	34581–34603
RTTF3R3	GCCTCATATTCTTCCGCTTGGC	48989–49010
QRTTF1F	GCCCGGAACGTGCATCT	40766–40782
QRTTF1R	CCATTTACCCCAGTGCAGAAA	40815–40835
QRTTF1P	TCTTGCAGATATGGCCGAAGCA ATGG	40785–40810
QRTTF2F	CGTGCTGCATTTCATTGCTAA	41605–41625
QRTTF2R	CGCTGAGGAAATGCGCTTA	41658–41676
QRTTF2P	ATTAAGACAAAAATTGGTTCAAAC TGCC	41627–41656

cells were harvested by flooding each plate with 1 ml of distilled water and subsequently resuspended them using a spreading paddle. A copper grid coated with Formvar and carbon was floated in the resulting suspension for 2 min. Excess fluid from the copper grid was removed prior to negative staining with 1% methyl tungstate. The cells were examined to determine the presence of a pilus by transmission electron microscopy.

Reverse transcriptase PCR (RT-PCR). Following growth for 12 h on solid medium and growth to an optical density at 600 nm of 0.6 in liquid medium, total RNA was prepared from harvested bacteria using the SV total RNA isolation system (Promega); the strains used were *H. influenzae* strains Rd and Rd11, as well as Rd11*tfc1*, Rd11*tfc2*, Rd11*tfc4*, Rd11*tfc6*, Rd11*tfc12*, Rd11*tfc14*, Rd11*tfc16*, Rd11*tfc22*, Rd11*tfc23*, Rd11*tfc24*, which represented mutants with mutations in the putative T4SS. Following RNA isolation, the yield was quantified spectrophotometrically. Subsequently, 550 ng of total RNA was reverse transcribed to produce cDNA using Superscript II reverse transcriptase (Invitrogen) and random primers (Promega). Coamplification of the cDNA transcripts for genes of the putative T4SS with an internal control allowed comparison of the levels of expression in the different RNA samples. PCR amplification of the cDNA of the putative T4SS genes was performed with oligonucleotides and probes listed in Table 2. Primers specific for the constitutively expressed housekeeping gene encoding beta-lactamase were also included in each PCR amplification as an internal control. To control for chromosomal DNA contamination, RNA samples were also directly used for PCR amplification.

Quantification of the *tfc16* and *tfc17* transcripts in strains Rd11 and Rd11*tfc24* was performed using an ABI Prism 7000 sequence detection system (Applied Biosystems) according to the manufacturer's recommendations. Probes were labeled with 6-carboxyfluorescein. Plasmid standards were obtained by preparing serial 10-fold dilutions in 100 μl (final volume) of nuclease-free water such that 5 μl of standard added to a 50-μl real-time quantitative PCR mixture (1× TaqMan Universal PCR Master Mix [Applied Biosystems], each primer [Applied Biosystems] at a concentration of 500 nM, 100 nM TaqMan 6-carboxyfluorescein-labeled probe [Applied Biosystems]) contained from 3×10^5 to 3×10^1 copies. The reporter fluorescence was normalized to the fluorescence of the passive reference dye ROX. Amplification data, expressed as increases in reporter fluorescence, were collected in real time and were analyzed using the sequence detection system. Experiments were repeated in triplicate with two batches of cDNA made from RNA samples obtained from two independent preparations.

DNA sequence analysis. DNA sequence similarity searches using the BLASTN and BLASTX algorithms (1) and position-specific iterated BLAST (PSI-BLAST) (2) were performed by interrogating the National Center for Biotechnology Information (NCBI) website (http://ncbi.nlm.nih.gov/). The Smith-Waterman algorithm, using the Blosum65 score matrix, a gap opening penalty of 11, and a gap extension penalty of 1, was used to align low-homology open reading frames (ORFs) with the ORFs of known T4SSs and to compute *P* values for homology (29), using a Bonferroni correction to account for multiple testing. The Artemis comparison tool (ACT) (42) was used to visually compare T4SSs of various bacteria pairwise for homology by using the TBLASTX algorithm. The presence and locations of signal peptide cleavage sites in amino acid sequences were predicted with SignalP 3.0 Server (http://www.cbs.dtu.dk/services/SignalP/), using a combination of several artificial neural networks and hidden Markov models (6).

Evolutionary analysis. The evolutionary relationships of a set of genomic islands that carry putative T4SSs were investigated using three representative essential genes (*tfc6*, *tfc14*, and *tfc16*). Homologues of only these genes were present in all the genomic islands represented in the comparative analysis. It has been claimed elsewhere that *tfc14* and *tfc16* encode the archetypal T4SS proteins (32). The relationships were estimated using ClustalX based on either the amino acid sequences encoded by each gene individually or the amino acid sequences encoded by the concatenated sequences of *tfc14* and *tfc16*. For these analyses, the following settings were used: correction for multiple substitutions and exclusion of positions with gaps. The output from the phylogenetic analysis was visualized with the TreeView software (38).

RESULTS

Identification of the conserved type IV secretion system. Genomic island ICE*Hin1056*, which is an archetype of a genomic island that accounts for pandemic spread of antibiotic resistance among *H. influenzae*, is 59.4 kb long and consists of 64 ORFs. Based on preliminary bioinformatic analyses, 24 of the 64 ORFs of ICE*Hin1056*, designated *tfc1* to *tfc24* (Fig. 1a), were thought to encode an uncharacterized T4SS. Seven of the 24 gene products of the ICE*Hin1056* putative T4SS have transmembrane helices, suggesting membrane-bound localization of the relevant proteins (Fig. 1a). Moreover, sequence analysis with SignalP 3.0 Server (6) revealed that 10 of the 24 genes of this gene cluster have signal peptide sequences that are typical for genes involved in T4SSs (Fig. 1a) (34). Characteristically, T4SSs are encoded by many genes expressing transmembrane

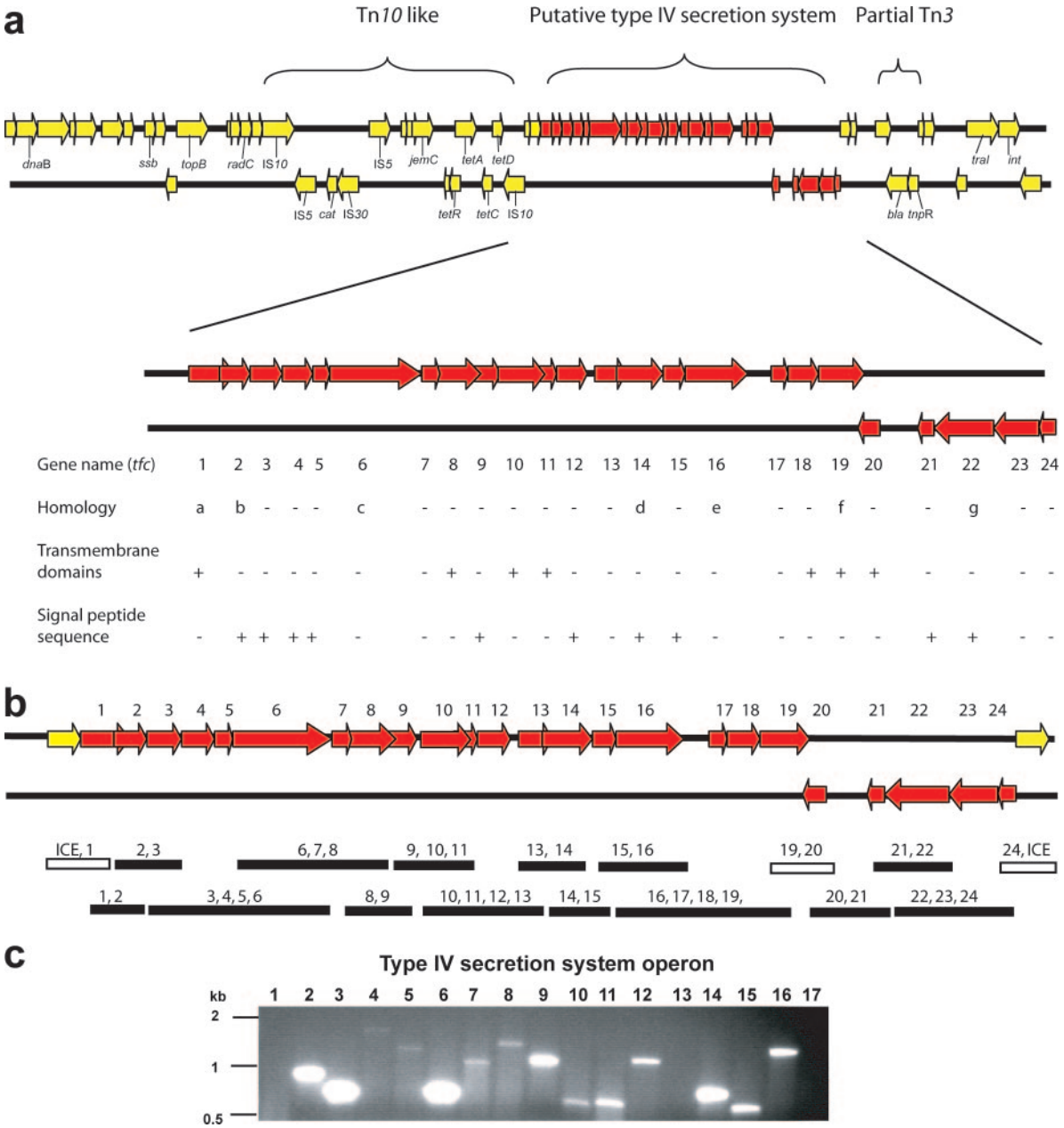


FIG. 1. (a) Schematic diagram of the ICE*Hin1056* genomic island and localization of the putative type IV secretion system (red). This T4SS is encoded by 24 genes, 7 of which exhibit homology to known T4SS components (components a to g). Seven of 24 gene products of the ICE*Hin1056* putative T4SS have transmembrane helices (indicated by plus signs), suggesting membrane-bound localization of the relevant proteins. Moreover, 10 of 24 genes of this gene cluster have signal peptide sequences (indicated by plus signs) that are typical for genes involved in T4SSs. a, TraX homologue (F plasmid); b, PilL homologue (plasmid pKLC102); c, TraD homologue (F plasmid); d, TraB homologue (F plasmid); e, TraC/VirB4 homologue (F plasmid); f, TraG homologue (F plasmid); g, PilT homologue (*L. pneumophila* str. Philadelphia 1). (b) Transcriptional mapping by RT-PCR of the ICE*Hin1056* T4SS gene cluster. The T4SS gene cluster is expressed in the form of two polycistronic transcripts. Cotranscription of neighboring genes is indicated by solid boxes. Neighboring genes which are not transcribed together are indicated by open boxes. (c) Gel showing that no mRNA transcripts were detected between *tfc19* and *tfc20* and between the first gene (*tfc1*) and the last gene (*tfc24*) of the ICE*Hin1056* T4SS and the neighboring genes in the rest of the ICE*Hin1056* sequence (ORF ICE*Hin1056*). Lane 1, ORF ICE*Hin1056-tfc1* transcript; lane 2, *tfc1-tfc2* transcript; lane 3, *tfc2-tfc3* transcript; lane 4, *tfc3-tfc4-tfc5-tfc6* transcript; lane 5, *tfc6-tfc7-tfc8* transcript; lane 6, *tfc8-tfc9* transcript; lane 7, *tfc9-tfc10-tfc11* transcript; lane 8, *tfc10-tfc11-tfc12-tfc13* transcript; lane 9, *tfc13-tfc14* transcript; lane 10, *tfc14-tfc15* transcript; lane 11, *tfc15-tfc16* transcript; lane 12, *tfc16-tfc17-tfc18-tfc19* transcript; lane 13, *tfc19-tfc20* transcript; lane 14, *tfc20-tfc21* transcript; lane 15, *tfc21-tfc22* transcript; lane 16, *tfc22-tfc23-tfc24* transcript; lane 17, *tfc24*-ORF ICE*Hin1056* transcript.

domains and signal peptide sequences (32). DNA sequence similarity searches using the BLASTN and BLASTX algorithms (1) and position-specific iterated BLAST (2) together with the Smith-Waterman algorithm (29) were employed to identify components of this putative T4SS. This bioinformatic analysis revealed homology between *tfc1* and the known T4SS gene *traX* ($E = 1 \times 10^{-60}$ in PSI-BLAST iteration 2), encoding a protein responsible for acetylation of F plasmid prepilin (35).

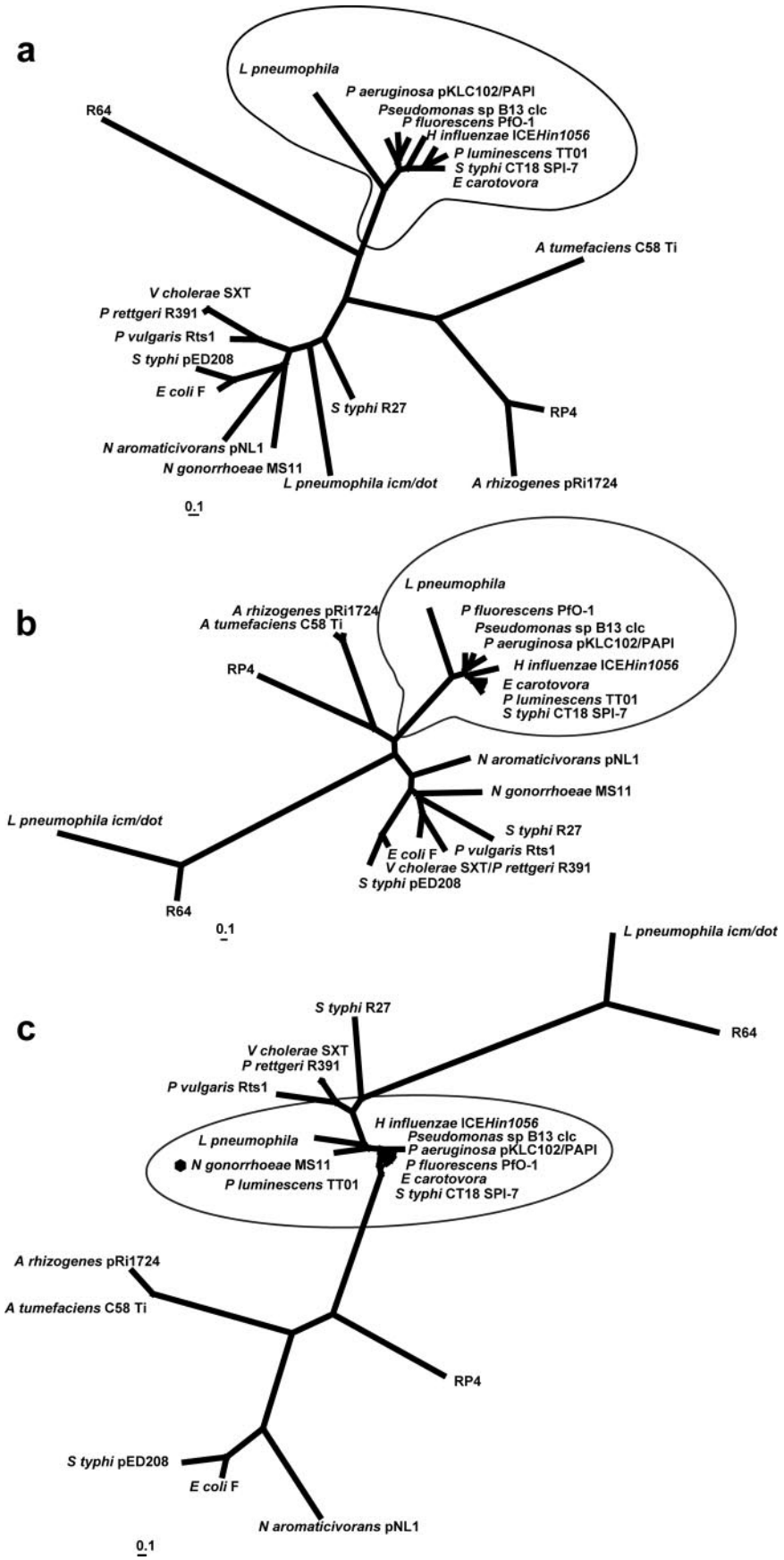
Moreover, *tfc6* was homologous to *traD* ($E = 2 \times 10^{-166}$ in PSI-BLAST iteration 1), a gene encoding a T4SS conjugative coupling protein involved in synchronization of mating pair formation with DNA transfer and "pumping" of DNA into recipient cells (22, 25). *tfc14*, *tfc16*, and *tfc19* are *traB* ($E = 9 \times 10^{-06}$ in PSI-BLAST iteration 2), *traC/virB4* ($E < 3 \times 10^{-159}$ in PSI-BLAST iteration 1), and *traG* ($E = 2 \times 10^{-42}$ in PSI-BLAST iteration 3) homologues, respectively, all of which are involved in pilus assembly (3, 32, 40). *tfc2* and *tfc22* exhibit homology with *pilL* ($E = 4 \times 10^{-19}$ in PSI-BLAST iteration 1) and *pilT* ($E = 1 \times 10^{-21}$ in PSI-BLAST iteration 1), respectively, whose products are required for type IV secretion pilus biogenesis (43, 44). Altogether, these in silico findings (Fig. 1a) suggest that the gene cluster which we investigated encodes components of a T4SS. Moreover, this gene cluster has been recognized in related genomic islands of other bacterial species, including *Legionella pneumophila*, *Pseudomonas aeruginosa*, *Pseudomonas fluorescens*, *Photorabdus luminescens*, *Erwinia carotovora*, and *Salmonella enterica* serovar Typhi.

ICE*Hin1056* type IV secretion system is evolutionarily distant from previously described T4SSs. Work over the past few years has led to classification of three distinct clusters of T4SSs: F-like (21, 22, 32, 33, 37), P-like (12, 13), and I-like (31, 46, 50). Relationships within and among clusters are based mainly on three genes, identified as either *tfc6*, *tfc14*, and *tfc16* or *traD*, *traB*, and *traC/virB4*, respectively (named for the *E. coli* F plasmid T4SS). The *tfc14* and *tfc16* genes encode the quintessential T4SS proteins (32) and as such are ideal for comparative amino acid alignment analysis. The trees for these three genes showed different patterns of clustering across genomic islands, but substantial compatibility between trees was evident based on homologies to *tfc14* and *tfc16* (Fig. 2a to c). Combining the data for both sets of homologues of *tfc14* and *tfc16* produced a tree with strong bootstrap support for four major clusters (Fig. 3). The fourth type of T4SS is carried by a wide variety of genomic islands, including the *clc* element, pKLC102, PAPI, and *Salmonella* pathogenicity island 7, as well as other genomic islands derived from the following bacterial strains: *L. pneumophila*, *P. aeruginosa*, *P. fluorescens*, *E. carotovora*, and *S. enterica* serovar Typhi (Fig. 3).

Role of the ICE*Hin1056* type IV secretion system in pilus biogenesis. T4SSs are multisubunit cell envelope-spanning structures comprising a secretion channel and often a pilus or other surface filament or protein (13, 19, 32, 41, 51, 52). Despite the presence of a vestigial pilus gene cluster in the genome, *H. influenzae* strain Rd produces neither fimbriae nor type IV pili (5, 20). However, electron microscopic examination of strain Rd harboring genomic island ICE*Hin1056* (strain Rd11) revealed long thin pili (Fig. 4a to d), which is consistent with the hypothesis that the genes responsible for formation of pili are located on ICE*Hin1056*. Only a single pilus approximately 7 to 7.5 nm in diameter and 1.5 μm long was attached to bacteria (Fig. 4a). In order to demonstrate the essential role of this putative T4SS for the formation of pili, 10 genes in this gene cluster, *tfc1*, *tfc2*, *tfc4*, *tfc6*, *tfc12*, *tfc14*, *tfc16*, *tfc22*, *tfc23*, and *tfc24*, were disrupted by insertion of a kanamycin resistance cassette, and mutant strains were examined for expression of pili by electron microscopy. These genes were chosen primarily because they exhibit homology to genes of other T4SSs or because of well-placed unique restriction sites. As shown in Table 3, disruption of these genes had a dramatic effect on the abilities of the mutants tested to produce pili. Pili were not detected in 7 of the 10 mutants analyzed, confirming the crucial role of this gene module in T4SS-dependent pilus expression. Not all genes associated with known T4SSs are involved in pilus biogenesis (13, 33). Disruption of the remaining three genes, *tfc1*, *tfc6* and *tfc24*, did not diminish the production of pili (Fig. 4b to d and Table 3). The pilus-positive phenotype of strains Rd11*tfc1* and Rd11*tfc24* provided an explanation for the observed slight changes in the conjugation frequencies described below and suggest that *tfc1* and *tfc24* do not play a significant role in the formation of the conjugative type IV pili under our in vitro conditions. By contrast, despite its massive reduction in conjugative transfer, the mutant with a mutation in the *tfc6* gene, encoding the conjugative coupling protein, produced pili (Table 3), confirming the hypothesis described by Panicker and Minkley (39) and verified by several other research groups that conjugative coupling proteins are essential for conjugation but are not required for T4SS pilus synthesis. However, most of the mutants examined were not able to express pili, revealing the essential role of this gene module in the expression of type IV secretion-dependent pili. None of the mutants has been complemented in *trans* as there is not a complementation system in *H. influenzae* that is easy to use; however, our experimental data show that the mutations had no polar effects on at least two of the downstream genes (*tfc6* and *tfc24* mutants produced pili in spite of pilus-negative phenotypes of upstream genes).

ICE*Hin1056* type IV secretion system is important for conjugation. Conjugation systems, assembled by a subfamily of the T4SSs, are usually used by bacteria in the process of conjugative transfer of DNA from donor cells to recipient cells (11). Genomic island ICE*Hin1056* in the parent strain was transferred from the donor to the recipient strain of *H. influenzae* at a frequency of 3×10^{-2}. To obtain evidence for the function of the ICE*Hin1056* T4SS, all T4SS mutant strains were tested to determine their abilities to transfer the genomic island by conjugation to the same recipient. As shown in Table 3, the conjugation frequencies of the mutants were significantly lower than the conjugation frequency of the parent strain and ranged from 6×10^{-3} to 1×10^{-7}. In agreement with our expectations, the lowest conjugation frequency, 1×10^{-7}, was observed for the mutant with the disrupted *tfc6* gene, encoding a putative conjugative coupling protein (8). Disruption of three genes, *tfc1*, *tfc4*, and *tfc24*, reduced the conjugal transfer frequencies to approximately 10^{-3}, whereas disruption of the remaining seven genes, *tfc2*, *tfc6*, *tfc12*, *tfc14*, *tfc16*, *tfc22*, and *tfc23*, led to reduction of the conjugal transfer frequencies to approximately 10^{-7}. In summary, the abilities of the majority of the mutants tested to conjugate were critically impaired, and the frequencies of transfer were reduced up to 100,000-fold compared to the frequencies of transfer observed for the parent strain, providing functional evidence of the importance of the T4SS for conjugal transfer of ICE*Hin1056* DNA; hence, we used the designation *tfc* (*t*ype *f*our *c*onjugation) for all components of this gene cluster.

Transcriptional analysis of the ICE*Hin1056* type IV secretion system. RT-PCR analysis of gene expression revealed that the ICE*Hin1056* T4SS gene cluster is expressed in the form of polycistronic transcripts (Fig. 1b and c). No mRNA transcripts



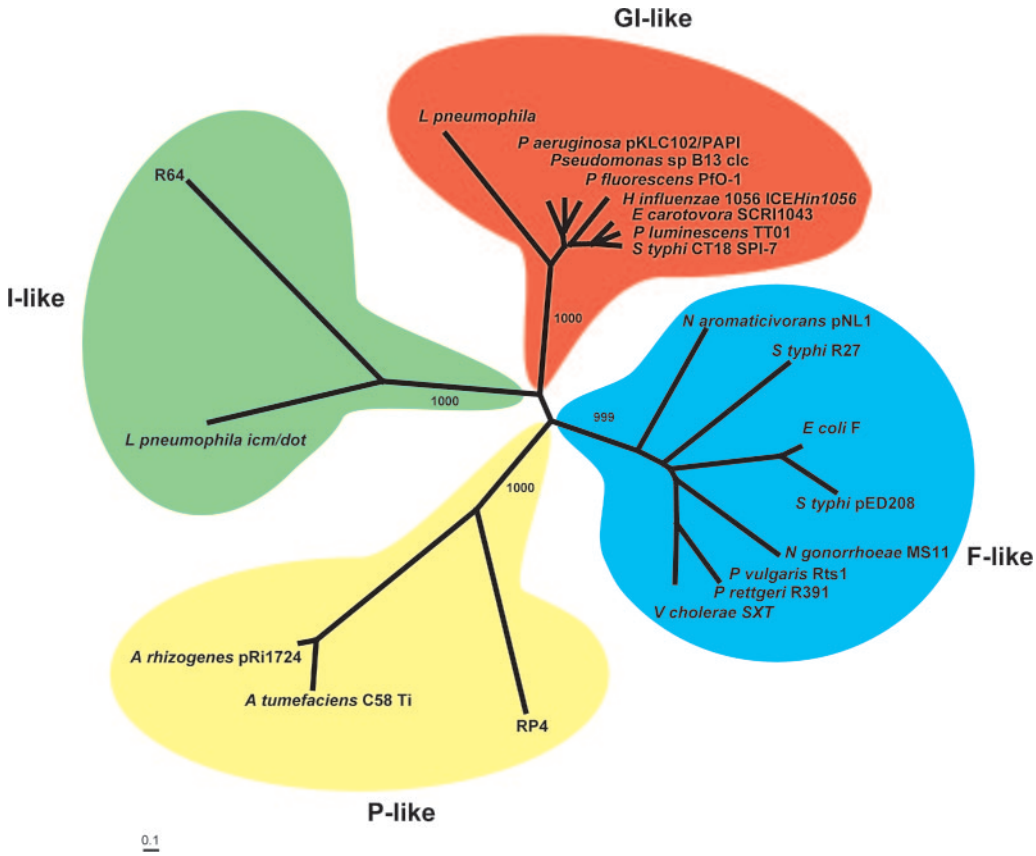


FIG. 3. Phylogenetic tree based on ClustalX analysis, showing the evolutionary relationship between concatenated amino acid sequences of *tfc14* and *tfc16* of the ICE*Hin1056* T4SS and the previously described T4SSs. The ICE*Hin1056* T4SS does not belong to any of the previously described groups and is evolutionarily distant. Moreover, syntenic T4SSs located on a broad spectrum of genomic islands present in a wide variety of bacterial species belong to this evolutionarily distant group. For an explanation of abbreviations see the legend to Fig. 2.

were detected between *tfc19* and *tfc20* and between the first gene (*tfc1*) and the last gene (*tfc24*) of the ICE*Hin1056* type IV secretion cluster analyzed and the neighboring genes in the rest of the ICE*Hin1056* sequence (Fig. 1b and c). As Fig. 1b and c show, the RT-PCR analysis results suggest that the module consists of two transcripts. Transcriptional analysis by RT-PCR (Fig. 1b and c) revealed amplicons for the gene downstream of the gene interrupted by insertion of an antibiotic resistance cassette in each of the 10 T4SS mutant strains tested (data not shown). This suggests that there are no polar effects on genes downstream of insertional mutations.

Interestingly, RT-PCR analysis of gene expression of the ICE*Hin1056* T4SS indicated that there were no substantial changes in the pattern of expression in the mutants tested compared to the wild-type pattern, except for mutant *tfc24*. As shown in Fig. 5a, mutation of *tfc24* greatly reduced expression of many of the type IV secretion genes. Quantification of *tfc16* and *tfc17* mRNA transcripts in strains Rd11 and Rd11*tfc24* using the ABI Prism 7000 sequence detection system revealed that the number of copies of the *tfc16* gene transcript was reduced from 530×10^2 in strain Rd11 to 263×10^2 in mutant Rd11*tfc24* (Fig. 5b). Moreover, the number of copies of the *tfc17* gene transcript was reduced from 537×10^2 in strain Rd11 to 411×10^2 in mutant Rd11*tfc24* (Fig. 5b). The *tfc24* product exhibits strong homology with conserved proteins having unknown functions found in a wide variety of bacterial species, including *Haemophilus somnus*, *Haemophilus ducreyi*, *P. aeruginosa*, *P. fluorescens*, *Pseudomonas syringae*, *P. luminescens*, *E. carotovora*, *S. enterica* serovar Typhi, and *Yersinia pseudotuberculosis*, and it contains a protein kinase C phosphorylation site. This suggests that *tfc24* might encode a protein which plays a crucial role in regulation of the T4SS, whose function might be conserved between species. The role of this gene in regulation remains to be elucidated.

FIG. 2. Phylogenetic trees based on ClustalX analysis, showing the evolutionary relationships between amino acid sequences encoded by *tfc14* (a), *tfc16* (b), and *tfc6* (c) of the ICE*Hin1056* T4SS and homologues from a set including the previously described T4SSs and homologues identified by bioinformatic analysis from a wide range of proteobacteria. The topology of the trees resulting from alignment of amino acid sequences encoded by *tfc6*, *tfc14*, and *tfc16* and their homologues revealed that the topology was conserved only for *tfc14* and *tfc16*. *S Typhi*, *Salmonella enterica* serovar Typhi; *V cholerae*, *Vibrio cholerae*; *P rettgeri*, *Proteus rettgeri*; *P vulgaris*, *Proteus vulgaris*; *N aromaticivorans*, *Neisseria aromaticivorans*; *N gonorrhoeae*, *Neisseria gonorrhoeae*; *A rhizogenes*, *Agrobacterium rhizogenes*.

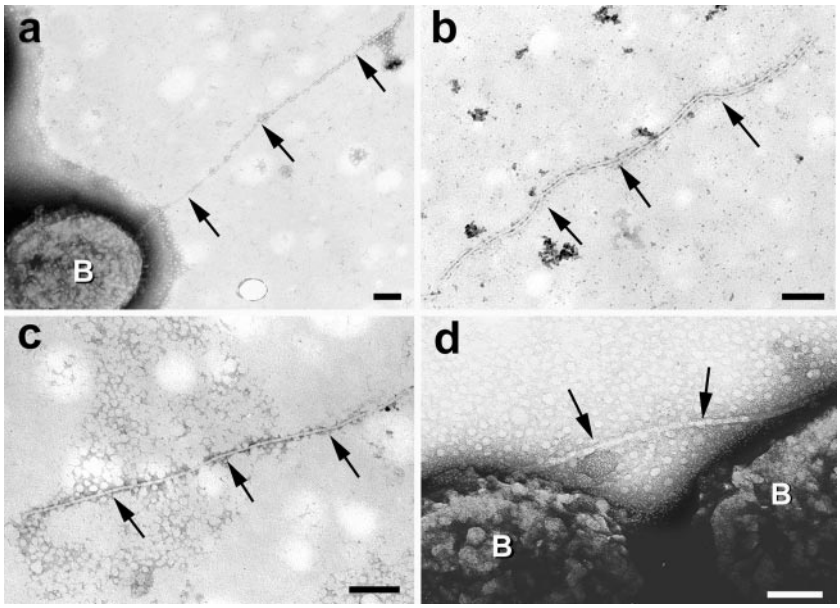


FIG. 4. Transmission electron micrographs of negatively stained samples, showing the presence of pili dependent on the T4SS (arrows) in *H. influenzae* strain Rd harboring ICE*Hin1056* (a), the *tfc1* mutant (b), the *tfc6* mutant (c), and the *tfc24* mutant (d), B, bacterium. Bars = 100 nm.

DISCUSSION

A set of genes present in ICE*Hin1056*, a vector of antibiotic resistance in *H. influenzae*, was characterized, and we concluded that this vector represents a new family of T4SSs involved in DNA translocation among gram-negative bacteria. This interpretation was based on bioinformatic, evolutionary, and functional analyses. First, in the bioinformatic analysis, both protein motifs typical of T4SS gene clusters were found, and homologies to genes known to be associated with T4SSs were identified. Second, the evolutionary analysis of the two archetypal T4SS proteins revealed high levels of similarity between amino acid sequences for the three previously identified T4SS types plus a novel fourth type. Third, mutagenesis of genes indicated that the genes are involved in pilus formation and the resulting transfer of DNA by conjugation. Lastly, a transcriptional analysis suggested that this gene cluster is probably part of a regulated operon typical of other T4SSs. The putative origin of transfer has been localized between ORFs 60 and 61 of ICE*Hin1056*, a position that is distinct from the position of the putative T4SS module investigated.

TABLE 3. Conjugation frequencies and pilus formation for mutants with mutations in the ICE*Hin1056* type IV secretion system

Donor strain	Conjugation frequency (mean ± SE)[a]	Pilus formation
Parent strain	$3 \times 10^{-2} \pm 0.7 \times 10^{-2}$	+
Rd11*tfc1*	$1 \times 10^{-3} \pm 0.1 \times 10^{-3}$	+
Rd11*tfc2*	$5 \times 10^{-7} \pm 0.8 \times 10^{-7}$	−
Rd11*tfc4*	$2 \times 10^{-3} \pm 0.8 \times 10^{-3}$	−
Rd11*tfc6*	$1 \times 10^{-7} \pm 0.4 \times 10^{-7}$	+
Rd11*tfc12*	$7 \times 10^{-7} \pm 1.2 \times 10^{-7}$	−
Rd11*tfc14*	$3 \times 10^{-7} \pm 2.9 \times 10^{-7}$	−
Rd11*tfc16*	$4 \times 10^{-7} \pm 1.3 \times 10^{-7}$	−
Rd11*tfc22*	$7 \times 10^{-7} \pm 3.3 \times 10^{-7}$	−
Rd11*tfc23*	$6 \times 10^{-7} \pm 2.6 \times 10^{-7}$	−
Rd11*tfc24*	$6 \times 10^{-3} \pm 2 \times 10^{-3}$	+

[a] The values are the values from three independent experiments.

The genes that we analyzed are members of a set of core genes found in a diverse family of syntenic genomic islands (36). Well-recognized members of this family of genomic islands are widely distributed among proteobacteria, and additional variants have been identified among completed genome sequences present in NCBI databases.

Analysis of mRNA by RT-PCR suggested that transcription of the T4SS genes is polycistronic. The gene disruptions in the T4SS mutants described here have not been complemented in *trans* as no reliable complementation system is available for *H. influenzae*; however, our experimental data show that the mutations did not have polar effects on the downstream genes. Both *tfc6* and *tfc24* mutants produce pili in spite of the pilus-negative phenotype of upstream genes. Similarly, the conjugation frequencies of *tfc4* and *tfc24* mutants are only slightly reduced compared to those of mutants with mutations in upstream genes. Furthermore, transcriptional analysis suggested that there was no clear polar effect of gene disruption on downstream transcripts. This might be explained by the fact that the kanamycin resistance cassette used harbors an endogenous promoter which, when oriented in the same direction as the target gene, does not interfere with downstream gene expression.

Recent reviews of T4SSs have described the relationship between the recognized T4SSs (9, 13, 32). A simple consensus view is that T4SSs share two features: they encode bacterial protein and nucleoprotein export systems, and members of each type exhibit evidence of common ancestry. This definition is broad, and its interpretation is flexible; therefore, it is not surprising that the T4SS gene clusters described previously are diverse in a number of respects. They vary in gene content, the number of homologues that they share, and the gene order. This heterogeneity is a challenge to a simple classification.

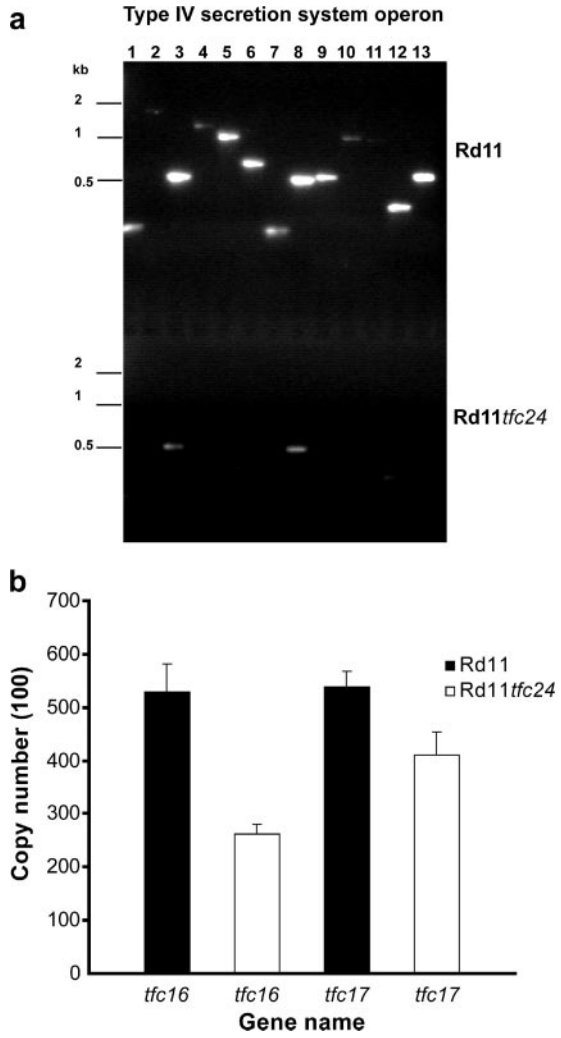


FIG. 5. RT-PCR showing transcription levels of genes of the ICE*Hin1056* T4SS in *H. influenzae* strain Rd harboring ICE*Hin1056* (Rd11) and in the *tfc24* mutant (Rd11*tfc24*). (a) Expression of the majority of the genes of the T4SS examined is downregulated in the *tfc24* mutant. The results also demonstrate that there is cotranscription of T4SS genes. Lane 1, *tfc1* transcript; lane 2, *tfc3-tfc4-tfc5-tfc6* transcript; lane 3, *tfc6* transcript; lane 4, *tfc6-tfc7-tfc8* transcript; lane 5, *tfc9-tfc10-tfc11* transcript; lane 6, *tfc12-tfc13* transcript; lane 7, *tfc13* transcript; lane 8, *tfc14* transcript; lane 9, *tfc14-tfc15* transcript; lane 10, *tfc16* transcript; lane 11, *tfc17-tfc18-tfc19* transcript; lane 12, *tfc22* transcript; lane 13, *tfc23-tfc24* transcript. (b) Quantification of the *tfc16* and *tfc17* mRNA transcripts in strains Rd11 and Rd11*tfc24* using the ABI Prism 7000 sequence detection system. Interpolating from the standard curve, the number of copies of the *tfc16* gene transcript was reduced from 530×10^2 in strain Rd11 to 263×10^2 in mutant Rd11*tfc24* and the number of copies of the *tfc17* gene transcript was reduced from 537×10^2 in strain Rd11 to 411×10^2 in mutant Rd11*tfc24*.

Therefore, it is not surprising that two approaches have been used for classification. Both approaches were founded on a combination of the genetic organization of the cluster and the evolutionary relationships among homologous genes. The first classification recognized three types based on the incompatibility group of the representative and archetypal DNA conjugative plasmid, namely, type F for IncF (plasmid F), type P for IncP (plasmid RP4), and type I for IncI (plasmid R64) (32). There is greater conservation of gene content and order within each type than there is between types. Alignment of amino acid sequences shared by proteins encoded by homologous genes supports the subdivision into these three groups. The alternative classification recognizes two major groups, type IVA and type IVB, and a third, "other" group (9, 10, 13). Type IVA includes the F and P types referred to above. Type IVB consists of type I, whose component modules are considered to be highly divergent from those of members of both type F and type P. The third group, labeled the "other" group, has not been well characterized.

The T4SS gene cluster reported here is distinct from previously recognized types whichever typing scheme is used. It represents a fourth lineage with a genetic distance as great as that between the F, P, and I lineages. The component genes of this module exhibit homology to a limited number of genes in the other lineages; only four homologues, *tfc6* (*traD*), *tfc14* (*traB*), *tfc16* (*traC*/*virB4*), and *tfc19* (*traG*) (Fig. 1a), are shared with the 12-gene VirB/D4 system found in the Ti plasmid of *Agrobacterium tumefaciens* and with the more numerous genes contributing to the T4SS gene clusters of the F plasmid and plasmid R64. Thus, using the alternative classification system, this new T4SS would be classified as a member of neither type IVA or type IVB but as a member of the "other" group. Despite the involvement in DNA conjugation, there is no recognized incompatibility group for the elements containing the new T4SS; consequently, it does not readily fit into the typing scheme consisting of types F, P, and I. We designated this lineage the GI type, as this T4SS is associated exclusively with genomic islands, as explained further below. The GI type T4SS gene cluster is a core component of a highly diverse family of syntenic genomic islands that presumably evolved in proteobacteria. Remarkably, the G+C contents of the GI type gene clusters found in this family of genomic islands range from 40% to approximately 70%, indicating that there is substantial divergence and evolutionary separation; however, the gene content and order have remained largely intact (36). This suggests that there is a particular fitness advantage to this coherent set of genes acting together, and intriguingly, so far these genes appear to be uniquely associated with a specific family of genomic islands.

The observations reported here describing the GI type of T4SS module from the archetypal genomic island ICE*Hin1056* provide direct evidence of how genomic islands may conjugally transfer. Hitherto, there was doubt about how such widely distributed structures could rapidly mobilize and spread through bacterial populations. Here we present evidence that one related, diverse, and widely distributed family of genomic islands possesses the genetic wherewithal to conjugate. Only two examples of this family of genomic islands have been shown to be conjugally transferable, namely, ICE*Hin1056* and the *clc* element (23, 36). It can be inferred that other members of this family of genomic islands mobilize intact from the distribution of the following genomic islands among their host bacteria: *Salmonella* pathogenicity island 7 (*S. enterica* serovar Typhi) (7), the *Yersinia* adhesion pathogenicity island (*Y. pseudotuberculosis*) (14), and PAP1 and pKLC102-like genomic islands (*P. aeruginosa*) (26, 30). The presence of a highly evolved and efficient conjugation system for mobilizing genomic islands provides a better explanation of how

one family of genomic islands is able to propagate and efficiently enable bacterial populations to adapt to rapidly changing environments. What remains to be seen is whether the many diverse genomic islands not related to ICE*Hin1056* have dedicated conjugation systems that enable them to propagate through conjugation. Only recently have bioinformatic techniques become powerful enough to recognize and make predictions allowing recognition of the ICE*Hin1056* family of genomic islands and their related T4SS gene clusters. It is quite plausible that other distinct families of genomic islands will be identified and that new and divergent lineages (families) of T4SS gene clusters will be recognized.

Potentially, 24 genes were found in the T4SS module of ICE*Hin1056*, and 7 of these genes are homologous to genes directly or indirectly associated with other T4SS gene clusters. Four of these seven genes are homologues to VirB/D4 genes found in the Ti plasmid of *A. tumefaciens*, a gene cluster considered by many workers to be the minimal essential set of T4SS genes; therefore, there are at least 18 genes with no homology to any of the remaining genes in the other type F, P, and I T4SS lineages, as determined by currently available techniques. Similarly, there are genes in types F, P, and I that are unique to each type, and the number may be more than 10 for type I and for type F. This heterogeneity between types for genes forming a T4SS module is consistent with the hypothesis that these genes were assembled through recruitment by a mixture of convergent (horizontal transfer) and divergent evolutionary processes. Curiously, within each type, there is marked preservation of coherence in each T4SS gene cluster. Once assembled, despite opportunities for recombination and reordering of component genes, the cluster apparently remains coherent due to a particular fitness advantage for a coordinated set of genes. Intriguingly, at least two genes, *traC*/*virB4* and *traB*, appear to be sufficient to define membership in their respective lineages and may be essential for a cluster to function as a T4SS export system. If these two genes provide the key signature identifying other poorly related T4SSs among genomes and genomic islands in particular, searching for homologues of these genes may be the way to identify other distinct T4SS lineages.

ACKNOWLEDGMENTS

We cordially thank E. Richard Moxon and Sarah L. Turner for critical reviews and comments on the manuscript and Kate Dingle and Gaynor Jenkins for their technical assistance with RT-PCR expression analysis.

This work was supported by a grant from the Medical Research Council. D.J.P.F. was supported by an equipment grant from the Wellcome Trust.

REFERENCES

1. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403–410.
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. **Anthony, K. G., W. A. Klimke, J. Manchak, and L. S. Frost.** 1999. Comparison of proteins involved in pilus synthesis and mating pair stabilization from the related plasmids F and R100-1: insights into the mechanism of conjugation. J. Bacteriol. **181:**5149–5159.
4. **Baar, C., M. Eppinger, G. Raddatz, J. Simon, C. Lanz, O. Klimmek, R. Nandakumar, R. Gross, A. Rosinus, H. Keller, P. Jagtap, B. Linke, F. Meyer, H. Lederer, and S. C. Schuster.** 2003. Complete genome sequence and analysis of *Wolinella succinogenes*. Proc. Natl. Acad. Sci. USA **100:**11690–11695.
5. **Bakaletz, L. O., B. D. Baker, J. A. Jurcisek, A. Harrison, L. A. Novotny, J. E. Bookwalter, R. Mungur, and R. S. Munson, Jr.** 2005. Demonstration of type IV pilus expression and twitching phenotype by *Haemophilus influenzae*. Infect. Immun. **73:**1635–1643.
6. **Bendtsen, J. D., H. Nielsen, G. von Heijne, and S. Brunek.** 2004. Improved prediction of signal peptides: SignalP 3.0. J. Mol. Biol. **340:**783–795.
7. **Bueno, S. M., C. A. Santiviago, A. A. Murillo, J. A. Fuentes, A. N. Trombert, P. I. Rodas, P. Youderian, and G. C. Mora.** 2004. Precise excision of the large pathogenicity island, SPI7, in *Salmonella enterica* serovar Typhi. J. Bacteriol. **186:**3202–3213.
8. **Cabezon, E., J. I. Sastre, and F. de la Cruz.** 1997. Genetic evidence of a coupling role for the TraG protein family in bacterial conjugation. Mol. Gen. Genet. **254:**400–406.
9. **Cascales, E., and P. J. Christie.** 2004. Definition of a bacterial type IV secretion pathway for a DNA substrate. Science **304:**1170–1173.
10. **Christie, P. J., and J. P. Vogel.** 2000. Bacterial type IV secretion: conjugation systems adapted to deliver effector molecules to host cells. Trends Microbiol. **8:**354–360.
11. **Christie, P. J.** 2001. Type IV secretion: intercellular transfer of macromolecules by systems ancestrally related to conjugation machines. Mol. Microbiol. **40:**294–305.
12. **Christie, P. J.** 2004. Bacterial type IV secretion: the *Agrobacterium* VirB/D4 and related conjugation systems. Biochim. Biophys. Acta **1694:**219–234.
13. **Christie, P. J., K. Atmakuri, V. Krishnamoorthy, S. Jakubowski, and E. Cascales.** 2005. Biogenesis, architecture and function of bacterial type IV secretion systems. Annu. Rev. Microbiol. **59:**451–485.
14. **Collyn, F., A. Billault, C. Mullet, M. Simonet, and M. Marceau.** 2004. YAPI, a new *Yersinia pseudotuberculosis* pathogenicity island. Infect. Immun. **72:**4784–4790.
15. **Dimopoulou, I. D., J. Z. Jordens, N. J. Legakis, and D. W. Crook.** 1997. A molecular analysis of Greek and UK *Haemophilus influenzae* conjugative resistance plasmids. J. Antimicrob. Chemother. **39:**303–307.
16. **Dimopoulou, I. D., J. E. Russell, Z. Mohd-Zain, R. Herbert, and D. W. Crook.** 2002. Site-specific recombination with the chromosomal $tRNA^{Leu}$ gene by the large conjugative *Haemophilus* resistance plasmid. Antimicrob. Agents Chemother. **46:**1602–1603.
17. **Dobrindt, U., B. Hochhut, U. Hentschel, and J. Hacker.** 2004. Genomic islands in pathogenic and environmental microorganisms. Nat. Rev. Microbiol. **2:**414–424.
18. **Duchaud, E., C. Rusniok, L. Frangeul, C. Buchrieser, A. Givaudan, S. Taourit, S. Bocs, C. Boursaux-Eude, M. Chandler, J. F. Charles, E. Dassa, R. Derose, S. Derzelle, G. Freyssinet, S. Gaudriault, C. Medigue, A. Lanois, K. Powell, P. Siguier, R. Vincent, V. Wingate, M. Zouine, P. Glaser, N. Boemare, A. Danchin, and F. Kunst.** 2003. The genome sequence of the entomopathogenic bacterium *Photorhabdus luminescens*. Nat. Biotechnol. **21:**1294–1295.
19. **Eisenbrandt, R., M. Kalkum, L. Lurz, and E. Lanka.** 2000. Maturation of IncP pilin precursors resembles the catalytic dyad-like mechanism of leader peptidases. J. Bacteriol. **182:**6751–6761.
20. **Fleischmann, R. D., M. D. Adams, O. White, R. A. Clayton, E. F. Kirkness, A. R. Kerlavage, C. J. Bult, J. F. Tomb, B. A. Dougherty, J. M. Merrick, K. McKenney, G. Sutton, W. Fitzhugh, C. Fields, J. D. Gocayne, J. Scott, R. Shirley, L. I. Liu, A. Glodek, J. M. Kelley, J. F. Weidman, C. A. Phillips, T. Spriggs, E. Hedblom, M. D. Cotton, T. R. Utterback, M. C. Hanna, D. T. Nguyen, D. M. Saudek, R. C. Brandon, L. D. Fine, J. L. Fritchman, J. L. Fuhrmann, N. S. M. Geoghagen, C. L. Gnehm, L. A. McDonald, K. V. Small, C. M. Fraser, H. O. Smith, and J. C. Venter.** 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science **269:**496–512.
21. **Frost, L. S., K. Ippen-Ihler, and R. A. Skurray.** 1994. Analysis of the sequence and gene products of the transfer region of the F sex factor. Microbiol. Rev. **58:**162–210.
22. **Frost, L. S., R. Leplae, A. O. Summers, and A. Toussaint.** 2005. Mobile genetic elements: the agents of open source evolution. Nat. Rev. Microbiol. **3:**722–732.
23. **Gaillard, M., T. Vallaeys, F. J. Vorholter, M. Minoia, C. Werlen, V. Sentchillo, A. Puhler, and J. R. van der Meer.** 2006. The *clc* element of *Pseudomonas* sp. strain B13, a genomic island with various catabolic properties. J. Bacteriol. **188:**1999–2013.
24. **Galibert, F., T. M. Finan, S. R. Long, A. Puhler, P. Abola, F. Ampe, F. Barloy-Hubler, M. J. Barnett, A. Becker, P. Boistard, G. Bothe, M. Boutry, L. Bowser, J. Buhrmester, E. Cadieu, D. Capela, P. Chain, A. Cowie, R. W. Davis, S. Dreano, N. A. Federspiel, R. F. Fisher, S. Gloux, T. Godrie, A. Goffeau, B. Golding, J. Gouzy, M. Gurjal, I. Hernandez-Lucas, A. Hong, L. Huizar, R. W. Hyman, T. Jones, D. Kahn, M. L. Kahn, S. Kalman, D. H. Keating, E. Kiss, C. Komp, V. Lelaure, D. Masuy, C. Palm, M. C. Peck, T. M. Pohl, D. Portetelle, B. Purnelle, U. Ramsperger, R. Surzycki, P. Thebault, M. Vandenbol, F. J. Vorholter, S. Weidner, D. H. Wells, K. Wong, K. C. Yeh, and J. Batut.** 2001. The composite genome of the legume symbiont *Sinorhizobium meliloti*. Science **293:**668–672.
25. **Gomis-Ruth, F. X., M. Sola, F. de la Cruz, and M. Coll.** 2004. Coupling

factors in macromolecular type-IV secretion machineries. Curr. Pharm. Des. **10:**1551–1565.
26. **He, J., R. L. Baldini, E. Deziel, M. Saucier, Q. Zhang, N. T. Liberati, D. Lee, J. Urbach, H. M. Goodman, and L. G. Rahme.** 2004. The broad host range pathogen *Pseudomonas aeruginosa* strain PA14 carries two pathogenicity islands harboring plant and animal virulence genes. Proc. Natl. Acad. Sci. USA **101:**2530–2535.
27. **Herriott, R. M., E. M. Meyer, and M. J. Vogt.** 1970. Defined non-growth media for stage II development of competence in *Haemophilus influenzae*. J. Bacteriol. **101:**517–524.
28. **Kaneko, T., Y. Nakamura, S. Sato, K. Minamisawa, T. Uchiumi, S. Sasamoto, A. Watanabe, K. Idesawa, M. Iriguchi, K. Kawashima, M. Kohara, M. Matsumoto, S. Shimpo, H. Tsuruoka, T. Wada, M. Yamada, and S. Tabata.** 2002. Complete genomic sequence of nitrogen-fixing symbiotic bacterium *Bradyrhizobium japonicum* USDA110. DNA Res. **9:**189–197.
29. **Kann, M. G., and R. A. Goldstein.** 2002. Performance evaluation of a new algorithm for the detection of remote homologs with sequence comparison. Proteins **48:**367–376.
30. **Klockgether, J., O. Reva, K. Larbig, and B. Tummler.** 2004. Sequence analysis of the mobile genome island pKLC102 of *Pseudomonas aeruginosa* C. J. Bacteriol. **186:**518–534.
31. **Komano, T., T. Yoshida, K. Narahara, and N. Furuya.** 2000. The transfer region of IncI1 plasmid R64: similarities between R64 *tra* and *Legionella icm*/*dot* genes. Mol. Microbiol. **35:**1348–1359.
32. **Lawley, T. D., W. A. Klimke, M. J. Gubbins, and L. S. Frost.** 2003. F factor conjugation is a true type IV secretion system. FEMS Microbiol. Lett. **224:** 1–15.
33. **Lawley, T. D., B. M. Wilkins, and L. S. Frost.** 2003. Bacterial conjugation in gram-negative bacteria, p. 203–226. *In* G. Phillips and B. Funnel (ed.), Biology of plasmids. ASM Press, Washington, DC.
34. **Lory, S., and M. S. Strom.** 1997. Structure-function relationship of type IV prepilin peptidase of *Pseudomonas aeruginosa*—a review. Gene **192:**117–121.
35. **Maneewannakul, K., S. Maneewannakul, and K. Ippen-Ihler.** 1993. Synthesis of F pilin. J. Bacteriol. **175:**1384–1391.
36. **Mohd-Zain, Z., S. L. Turner, A. M. Cerdeno-Tarraga, A. K. Lilley, T. J. Inzana, A. J. Duncan, R. M. Harding, D. W. Hood, T. E. Peto, and D. W. Crook.** 2004. Transferable antibiotic resistance elements in *Haemophilus influenzae* share a common evolutionary origin with a diverse family of syntenic genomic islands. J. Bacteriol. **186:**8114–8122.
37. **Moore, D., K. Maneewannakul, S. Maneewannakul, J. H. Wu, K. Ippen-Ihler, and D. E. Bradley.** 1990. Characterization of the F-plasmid conjugative transfer gene *traU*. J. Bacteriol. **172:**4263–4270.
38. **Page, R. D.** 1996. TreeView; an application to display phylogenetic trees on personal computers. Comput. Appl. Biosci. **12:**357–358.
39. **Panicker, M. M., and E. G. Minkley, Jr.** 1985. DNA transfer occurs during a cell surface contact stage of F sex factor-mediated bacterial conjugation. J. Bacteriol. **162:**584–590.
40. **Rabel, C., A. M. Grahn, R. Lurz, and E. Lanka.** 2003. The VirB4 family of proposed traffic nucleoside triphosphatases: common motifs in plasmid RP4 TrbE are essential for conjugation and phage adsorption. J. Bacteriol. **185:** 1045–1058.
41. **Rohde, M., J. Puls, R. Buhrdorf, W. Fischer, and R. Haas.** 2003. A novel sheathed surface organelle of the *Helicobacter pylori cag* type IV secretion system. Mol. Microbiol. **49:**219–234.
42. **Rutherford, K. J., J. Parkhill, J. Crook, T. Horsnell, P. Rice, M. A. Rajandream, and B. Barrell.** 2000. Artemis: sequence visualization and annotation. Bioinformatics **16:**944–945.
43. **Sakai, D., and T. Komano.** 2000. The *pilL* and *pilN* genes of IncI1 plasmids R64 and Collb-P9 encode outer membrane lipoproteins responsible for thin pilus biosynthesis. Plasmid **43:**149–152.
44. **Sakai, D., and T. Komano.** 2002. Genes required for plasmid R64 thin-pilus biogenesis: identification and localization of products of the *pilK*, pilM, *pilO*, *pilP*, *pilR*, and *pilT* genes. J. Bacteriol. **184:**444–451.
45. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
46. **Segal, G., M. Feldman, and T. Zusman.** 2005. The Icm/Dot type-IV secretion systems of *Legionella pneumophila* and *Coxiella burnetii*. FEMS Microbiol. Rev. **29:**65–81.
47. **Smith, H. O., J. F. Tomb, B. A. Dougherty, R. D. Fleischmann, and J. C. Venter.** 1995. Frequency and distribution of DNA uptake signal sequences in the *Haemophilus influenzae* Rd genome. Science **269:**468–470.
48. **Stuy, J. H.** 1979. Plasmid transfer in *Haemophilus influenzae*. J. Bacteriol. **139:**520–529.
49. **van der Meer, J. R., and V. Sentchilo.** 2003. Genomic islands and the evolution of catabolic pathways in bacteria. Curr. Opin. Biotechnol. **14:**248–254.
50. **Vogel, J. P., H. L. Andrews, S. K. Wong, and R. R. Isberg.** 1998. Conjugative transfer by the virulence system of *Legionella pneumophila*. Science **279:**873–876.
51. **Watarai, M., H. L. Andrews, and R. Isberg.** 2000. Formation of fibrous structure on the surface of *Legionella pneumophila* associated with exposure of DotH and DotO proteins after intracellular growth. Mol. Microbiol. **39:**313–329.
52. **Waters, C. M., H. Hirt, J. K. McCormick, P. M. Schlievert, C. L. Wells, and G. M. Dunny.** 2004. An amino-terminal domain of *Enterococcus faecalis* aggregation substance is required for aggregation, bacterial internalization by epithelial cells and binding to lipoteichoic acid. Mol. Microbiol. **52:**1159–1171.