INFECTION AND IMMUNITY, May 1993, p. 1909–1916
0019-9567/93/051909-08$02.00/0

Vol. 61, No. 5

Sequence Divergence in Two Tandemly Located Pilin Genes of *Eikenella corrodens*

TONE TØNJUM,[1]* SUSAN WEIR,[2] KJELL BØVRE,[1] ANN PROGULSKE-FOX,[3] AND CARL F. MARRS[2]

Kaptein W. Wilhelmsen og Frues Bakteriologiske Institutt, Rikshospitalet, N-0027 Oslo 1, Norway[1]; Department of Epidemiology, University of Michigan, Ann Arbor, Michigan 48109[2]; and Department of Oral Biology and Periodontal Disease Research Center, University of Florida, Gainesville, Florida 32610[3]

Received 20 November 1992/Accepted 16 February 1993

Eikenella corrodens **normally inhabits the human respiratory and gastrointestinal tracts but is frequently the cause of abscesses at various sites. Using the N-terminal portion of the *Moraxella nonliquefaciens* pilin gene as a hybridization probe, we cloned two tandemly located pilin genes of *E. corrodens* 31745, *ecpC* and *ecpD*, and expressed the two pilin genes separately in *Escherichia coli*. A comparison of the predicted amino acid sequences of *E. corrodens* 31745 EcpC and EcpD revealed considerable divergence between the sequences of these two pilins and even less similarity to EcpA and EcpB of *E. corrodens* type strain ATCC 23834. EcpC from *E. corrodens* 31745 displayed high degrees of homology to the pilins of *Neisseria gonorrhoeae* and *Pseudomonas aeruginosa*. EcpD from *E. corrodens* 31745 showed the highest homologies with the pilin of one of the three *P. aeruginosa* classes, whereas EcpA and EcpB of strain ATCC 23834 most closely resemble *Moraxella bovis* pilins. These findings raise interesting questions about potential genetic transfer between different bacterial species, as opposed to convergent evolution.**

Eikenella corrodens is part of the normal flora of the upper respiratory and gastrointestinal tracts of humans (7, 12). It is a fastidious, facultatively anaerobic, gram-negative, rod-shaped bacterium. This species has recently been found to belong to the family *Neisseriaceae*, and it is most closely related to the genera *Neisseria* and *Kingella* (6, 16, 36). *E. corrodens* is now being reported more frequently as a pathogen, primarily causing abscesses (2, 3, 14). It is also believed to play a role in dental pathogenesis by contributing to tissue destruction in periodontal disease (44, 46). In many cases, *E. corrodens* is found in coinfections with other organisms, such as *Streptococcus viridans* and anaerobes (44). It has been found as a sole isolate in blood, spinal and joint fluid, in cases of frontal sinusitis and orbital cellulitis, and in brain, abdominal, and neck abscesses as well as ascites (17, 18, 44).

E. corrodens displays a number of phenotypic traits associated with type IV (MePhe) pili, among them natural competence for genetic transformation (48), bacterial autoagglutination, agar pitting, twitching motility, and distinct spreading-corroding (SC) and smooth, non-spreading-corroding (N) colony morphology variants. These pili are proteinaceous surface appendages that have been shown to be virulence factors for *Neisseria gonorrhoeae* and many other species (45). Members of the type IV class of bacterial pili are found on a variety of gram-negative bacterial species (5), including *N. gonorrhoeae* (27), *Neisseria meningitidis* (13), the classical moraxellae (22, 37, 49), *Dichelobacter nodosus* (8, 14, 26), *Pseudomonas aeruginosa* (32, 40), and *Vibrio cholerae* (41). The main structural subunits of these pili, termed pilins, have a highly conserved amino-terminal domain of 25 to 30 amino acid residues (13), while the C-terminal part contains variable DNA regions that rearrange by different forms of recombination (20, 30, 52). *V. cholerae* (47) and enteropathogenic *Escherichia coli* (10) produce related pilins that have less amino-terminal sequence homology and contain a modified methionine as the first residue of the mature protein.

Considerable differences exist in the copy numbers and organization of type IV pilin genes and in the amount of amino acid sequence homology observed between pilins of the same and different species. *P. aeruginosa* strains have only a single copy of the pilin gene in each genome (40). Class I serotypes of *D. nodosus* have only a single gene, but class II serotypes have two pilin genes that are fairly divergent from each other and very different from class I pilin genes (23, 24). In *N. gonorrhoeae* and *N. meningitidis*, diversification of pilins results from homologous recombination between a single pilin gene expression locus and multiple partial pilin gene alleles (45, 52). *Moraxella bovis* and *Moraxella lacunata* each express one of two different pilin genes and can switch between them by an inversion of a 2.1-kb DNA fragment (9, 21, 37). Recently, two type IV pilin genes were cloned from the *E. corrodens* type strain ATCC 23834 (35).

The *E. corrodens* clinical strain 31745 (48) vividly expresses traits associated with piliation, such as distinct colony morphology variants. Because of its highly stable natural competence, we have routinely used this property for identification of new clinical isolates of *E. corrodens* by genetic transformation (48). The aim of this study was the cloning, sequencing, and expression of potential pilin genes of this strain as a basis for further studies on piliation and adherence.

MATERIALS AND METHODS

Bacterial strains, plasmids, and media. *E. corrodens* 31745 was isolated from spinal fluid (48). The strain was grown on 5% human blood agar (Difco, Detroit, Mich.) and identified as *E. corrodens* by genetic transformation (48). *E. coli*

* Corresponding author.

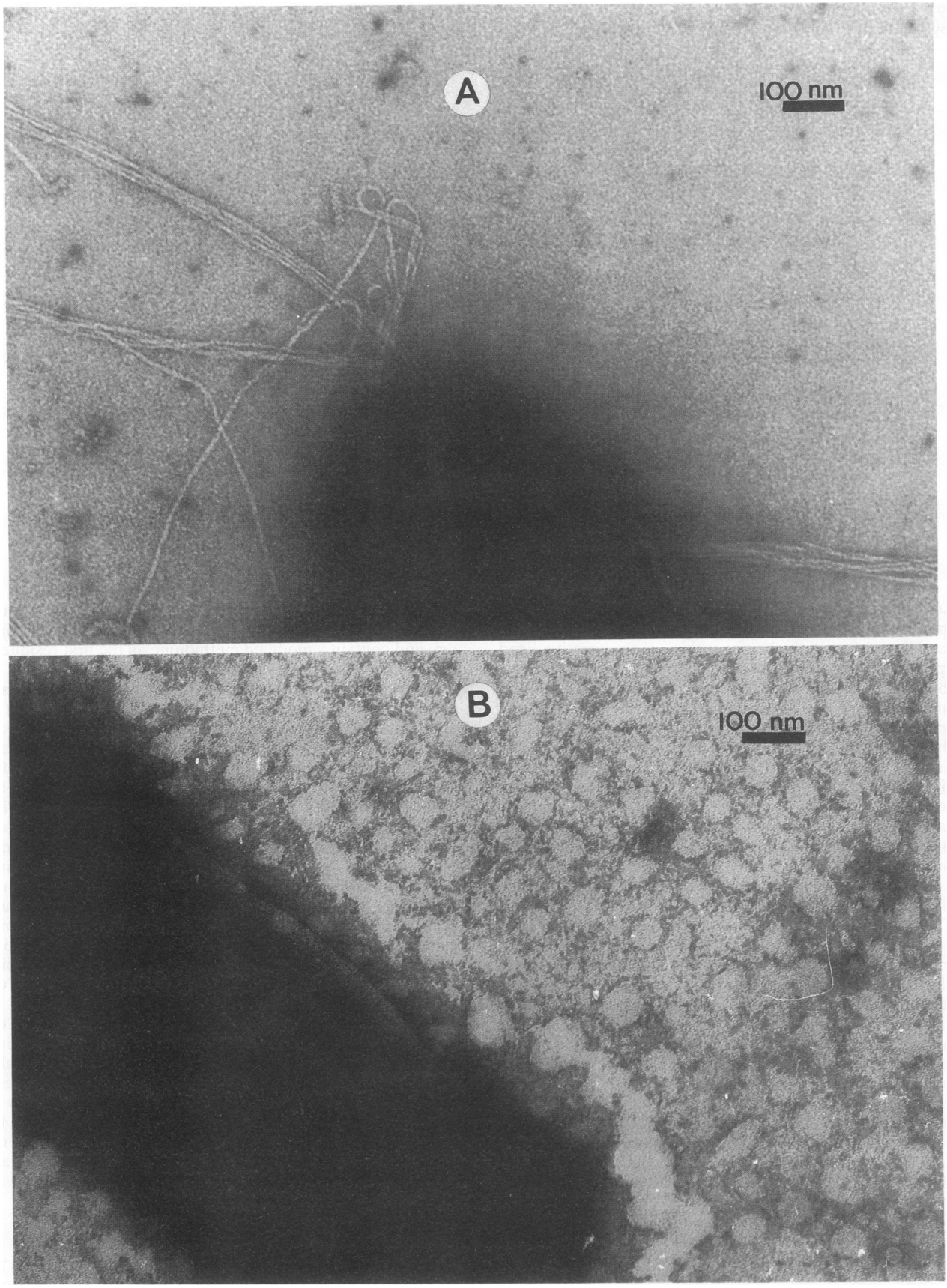

FIG. 1. Transmission electron micrographs of *E. corrodens* 31745 SC (A) and N (B) variants, grown for 3 days in static broth. Negative staining was performed with 1% uranyl acetate. Magnification, ×175,000.

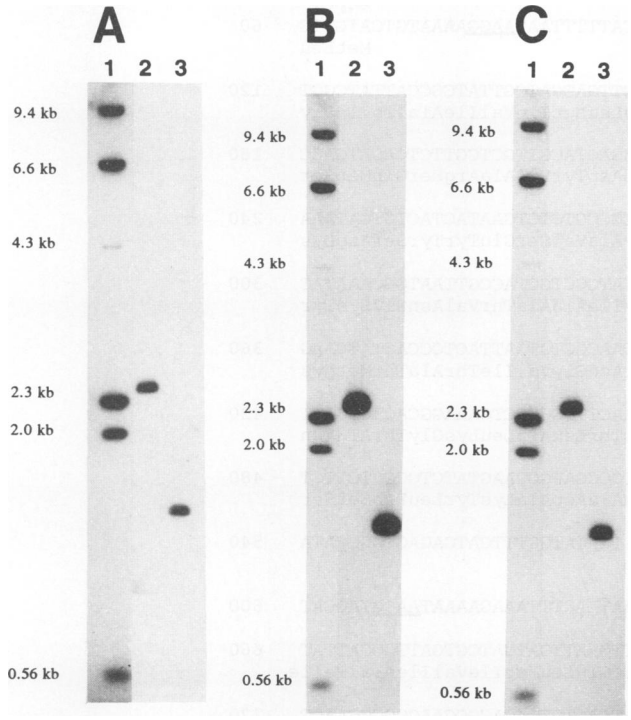

FIG. 2. Genomic Southern hybridization analysis of *Sma*I and *Kpn*I restriction endonuclease-digested *E. corrodens* 31745 total DNA with the N-terminal *M. nonliquefaciens* pilin gene (A), *E. corrodens* 31745 *ecpC* (B), and 31745 *ecpD* (C) PCR products as DNA probes. Lane 1, *Hin*dIII-digested lambda DNA; lane 2, 31745 digested with *Kpn*I; lane 3, 31745 digested with *Sma*I.

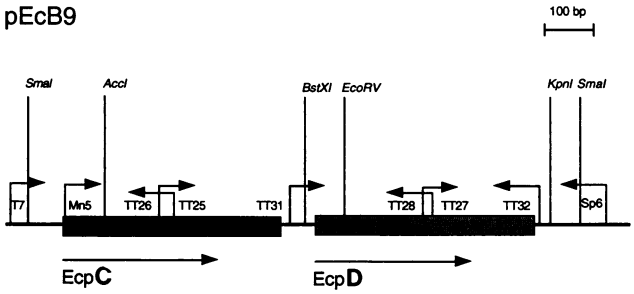

FIG. 3. Restriction map of the insert region of pEcB9, encoding the *E. corrodens* 31745 pilin genes *ecpC* and *ecpD*. The insert region is represented by the bold line between the *Sma*I sites. The positions and directions of the oligonucleotide primers used to determine the DNA sequence (T7, Sp6, TT25, TT26, TT27, and TT28) and for PCR (Mn5 plus TT26 and TT31 plus TT32 for *ecpC* and *ecpD*, respectively) are shown, as are the relative positions and directions of the genes within the insert.

SURE (*recB recJ sbc-201 uvrC umuC*::Tn*5* [Kan^r] *mcrA mcrB mrr lacD* [*hsdRMS*] *endA1 gyrA96 thi relA1 supE44* [F′ *proAB lacI^{q}ZΔM15* Tn*10* [Tet^r]) (Stratagene, La Jolla, Calif.) was used for amplification and screening of the genomic library, as it has the *mcrAB mrr* mutant background suitable for cloning methylated DNA (52). *E. coli* XL-1Blue [*recA1 endA1 gyrA96 thi hsdR17* (r_K^- m_K^{q+}) *supE44 relA1* 1_R- *Δlac* (F′ *proAB lacI^{q}ZΔM15* Tn*10* (Tet^r)] (Stratagene) was used as a host for subcloning. *E. coli* containing drug-resistant plasmids was grown on LB-agar containing carbenicillin (100 μg/ml; Beecham Pharmaceuticals, Surrey, U.K.). Plasmid pGEM7 (Promega, Madison, Wis.) was used for subcloning.

PCR. Synthetic oligonucleotides (Genetic Designs Inc., Houston, Tex.) Mn5 (5′-TGAACGCTCAAAAAGGTTTTA CCC-3′) and Mn6 (5′-GCTTCTGATACCTGTGCACGGGC -3′) (based on the N-terminal sequence of the *M. nonliquefaciens* pilin gene [49]) were used as primers for polymerase chain reaction (PCR) amplification of a pilin-specific probe with *M. nonliquefaciens* NCTC 7784 DNA as the template. The resulting amplification product was 120 bp long. Primers Mn5 and TT26 (5′-AATCACACCGTTAGTACC-3′) were used for amplification of the first *E. corrodens* pilin gene, and TT31 (5′-GATGAGAGTGGCCATACC-3′) and TT32 (5′-CT TGCTCTCTGGTTAAGC-3′) were used for the second pilin gene. Samples were amplified in 50 μl of a reaction mixture containing 10 ng of template DNA, 10 mM Tris-HCl, 50 mM KCl, 2 mM $MgCl_2$, 200 μM each deoxynucleotide triphosphate (Pharmacia), 0.001% (wt/vol) gelatin, 0.5 μM each primer, and 0.5 U of Ampli-Taq polymerase (Cetus, Norwalk, Conn.). Each sample was amplified for 25 cycles of 30 s at 95°C, 1 min at 65°C, and 1 min at 72°C in a thermal cycler (Perkin Elmer Cetus, Norwalk, Conn.).

DNA isolation and manipulation. Total DNA was prepared by the method of Hull et al. (15). A gene bank of *E. corrodens* 31745 (Stratagene, La Jolla, Calif.) was screened with an *M. nonliquefaciens* pilin-specific PCR probe by a plaque blotting technique (38). The phagemid Ec1A was chosen for in vivo excision of plasmid pBluescript with an insert, termed pEc2A. Size-fractionated *Sma*I fragments of pEc2A hybridizing to the pilin-specific PCR probe were obtained by electroelution from 1% agarose (SeaKem; FMC Bioproducts, Rockland, Maine). These fragments were ligated into *Sma*I-cleaved pGEM7, transformed into competent *E. coli* XL-1Blue cells, and selected on LB plates with carbenicillin. Cb^r transformants were screened for hybridization to the *M. nonliquefaciens* pilin-specific PCR probe by a colony blot method (38). The *Sma*I subclones pEcA4 and pEcB9, containing the same insert in opposite directions, were chosen for detailed analysis.

Plasmid DNA was isolated by the method of Birnboim and Doly (1) in Qiagen columns according to the manufacturer's recommendations. Restriction endonucleases were purchased from Bethesda Research Laboratories Inc.; New England Biolabs Inc., Beverly, Mass.; Boehringer Mannheim Biochemicals, Indianapolis, Ind.; or Promega, Madison, Wis. Restriction endonuclease digestions and agarose gel electrophoresis were performed as described previously (21). DNA transfer from agarose gels to nylon membranes (Hybond-N; Amersham, Arlington Heights, Ill.) was done by a modification of the method of Southern (43), with the Vacu-Gene System (Pharmacia-LKB, Piscataway, N.J.). Hybridizations were performed as described previously (22). The PCR pilin probes were labeled with [α-^{32}P]dCTP (Amersham) by use of a Megaprime kit (Amersham).

The DNA sequences of the insert regions of pEcA4 and pEcB9 were determined by the dideoxy chain termination method of Sanger et al. (39). The primers T7 (5′-TAATAC GACTCACTATA-3′) and SP6 (5′-TAAATCCACTGTGAT AT-3′) were sequences in the pGEM7 vector flanking the insert. The primers TT25 (5′-AACGGTGTGATTACTGCC-3′), TT26, TT27 (5′-GTGTAGCGACTGACCCGA-3′), and TT28 (5′-ATTCGGGTCAGTCGCTAC-3′) were also used for sequencing, and their positions are shown in Fig. 3. Some sequence was obtained by using primers T7 and Sp6 on

```
                                                                ecpC>
       CCCGGGTCGGAAGATTCGGTTTATTTAAGTTCCATTTTTAACAAGGAAAATGTCATGTTG   60
                                                                MetLeu

  61   AAACAAGTACAAAAAGGTTTTACCCTGATCGAGTTGATGATCGTTATCGCCATTATCGGT  120
       LysGlnValGlnLysGlyPheThrLeuIleGluLeuMetIleValIleAlaIleIleGly

 121   ATCTTGGCTGCTATCGCTCTGCCGGCTTACCAAGACTACGTTGCTCGTTCTCAGATGAGC  180
       IleLeuAlaAlaIleAlaLeuProAlaTyrGlnAspTyrValAlaArgSerGlnMetSer

 181   GAAGCCTTCAACTTGGCCGGTGGTCAAAAAGGTGCCGTCTCTGAATACTACTCTGATAAA  240
       GluAlaPheAsnLeuAlaGlyGlyGlnLysGlyAlaValSerGluTyrTyrSerAspLys

 241   GGTGTGTGGCCGGCTGATAATGCTGCTGCCGGTATCGCTGCTACCGTTAATGGCAAATAC  300
       GlyValTrpProAlaAspAsnAlaAlaAlaGlyIleAlaAlaThrValAsnGlyLysTyr

 301   GTTAATTCTGTTGTAGTAAGTGCTGCTGGTACTAACGGTGTGATTACTGCCACTATGAAG  360
       ValAsnSerValValValSerAlaAlaGlyThrAsnGlyValIleThrAlaThrMetLys

 361   TCTACTGGTGTAGCCAAAGGTGTACAAGGCAAAACTTTGGCCCTGAAAGGCACTGCGAAT  420
       SerThrGlyValAlaLysGlyValGlnGlyLysThrLeuAlaLeuLysGlyThrAlaAsn

 421   GATGGCTCTTTCTCTTGGGAGTGCTCATCTAATGCCGATGCGAAGTATCTGCCCTCTTCT  480
       AspGlySerPheSerTrpGluCysSerSerAsnAlaAspAlaLysTyrLeuProSerSer

 481   TGCCGTAATGCTGCTACTCCGACTCCGACTCCGTAATATGGTTTGATGAGAGTGGCCATA  540
       CysArgAsnAlaAlaThrProThrProThrPro

 541   CCATTGGGTGTGGCCATTTTTACATTAAAAGGAATAACTTTAAAGAAAATAAGGTAGCAT  600
                ecpD>
 601   TATGCAAGCTAAGATATCAGGTTTTACCTTGATTGAATTGATGATCGTGATTGCCATTAT  660
       MetGlnAlaLysIleSerGlyPheThrLeuIleGluLeuMetIleValIleAlaIleIle

 661   TGGGATCTTGGCAGCCATTGCTTTGCCGGCATATCAGTCTTACACGGCAAGGGCGCAAGT  720
       GlyIleLeuAlaAlaIleAlaLeuProAlaTyrGlnSerTyrThrAlaArgAlaGlnVal

 721   GTCGGAAGCTATCTCTTTGATGGATGGTTTAAAAAGCTCTGTGGCGGATAATTATTTTAA  780
       SerGluAlaIleSerLeuMetAspGlyLeuLysSerSerValAlaAspAsnTyrPheAsn

 781   TAGTATGATTTGTGCCGATAACCAAACAGTTAATCATTTTGGAATTGCGCAGAGAAGCCG  840
       SerMetIleCysAlaAspAsnGlnThrValAsnHisPheGlyIleAlaGlnArgSerArg

 841   TATTTCAGGCAATTATGTAGAATCAATACATACTCGTGTAGCGACTGACCCGAATTATAA  900
       IleSerGlyAsnTyrValGluSerIleHisThrArgValAlaThrAspProAsnTyrAsn

 901   TTGCGAAATGTTGGCTACGTTTAAATCAACAGATGCTGCCGCCCCTATACGGGGGAAGAC  960
       CysGluMetLeuAlaThrPheLysSerThrAspAlaAlaAlaProIleArgGlyLysThr

 961   TGTACTGCTTAGCATGAAAATAGTTGATGGCGGAACGGTTTGGAATTGTTCTTCTTCTGA 1020
       ValLeuLeuSerMetLysIleValAspGlyGlyThrValTrpAsnCysSerSerSerAsp

1021   TTTAGCGAATGAGTTTTTGCCAACAGCTTGCCGTCATTAGTATAAAGAGTTGAGATAGTT 1080
       LeuAlaAsnGluPheLeuProThrAlaCysArgHis

1081   TAGTATGTTGAAATATATTAAAAACTTAGGTTAATGGAATGTTGCTTAACCAGAGAGCAA 1140

1141   GATTTGCCGCCAAGTTGGCCGGCTGCCATTTAATCATCAATATCATTGCTGGCTGTTTTG 1200

1201   GTCGCAATTTTAGTATTTTATATATGGTACCCTTACCCTTACCGGCAAATGCTAGGAGGA 1260

1261   TTGGGTTTGTTTTCTTTGGTTATCGGTATTGATGTGGTGTGTGGACCAGTATTAACTGCG 1320

1321   GTATTGGCCAATCCGAAAAAATCGACCCGGG
```

FIG. 4. Nucleotide sequence and translation of the *Sma*I fragment containing the *E. corrodens* 31745 *ecpC* and *ecpD* pilin genes. The predicted amino acid sequence of each open reading frame containing the coding sequence of the pilin is shown. The AAGG sequences (underlined) upstream of the starting methionines are putative ribosome-binding sites.

*Kpn*I-, *Bst*XI-, and *Eco*RV-deleted subclones of pEcA4 and pEcB9.

DNA sequence analysis. DNA sequence analysis was performed with software from DNASTAR (Madison, Wis.). The optimized scores used to compare pilin protein sequences in Fig. 6 were obtained by using the AAlign program, which uses the amino acid sequence similarity algorithms of Needleman and Wunsch (28) and the scoring system of Lipman and Pearson (19).

Nucleotide sequence accession number. The nucleotide sequence presented here has been submitted to the GenBank data base under accession number L12049.

Western immunoblotting. Proteins of whole-cell extracts of bacteria were boiled in loading buffer, separated by molecular weight by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), and transferred to nitrocellulose filter membranes (Schleicher & Schuell, Keene, N.H.) (50). Rabbit antiserum 130D, raised against purified

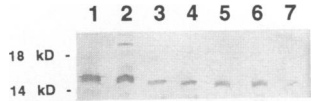

FIG. 5. Immunoblot analysis of crude pili and whole-cell extracts. Lane 1, purified pili from *M. nonliquefaciens* NCTC 7784; lane 2, *E. corrodens* 31745; lanes 3 and 4, *E. coli* XL-1Blue containing the cloned *E. corrodens* 31745 *Sma*I fragment homologous to the *M. nonliquefaciens* pilin-specific PCR probe in pEcB9 and pEcA4, respectively; lane 5, *E. coli* XL-1Blue containing the cloned *E. corrodens* 31745 *ecpC*; lane 6, *E. coli* XL-1Blue containing the cloned *E. corrodens ecpD*; lane 7, *E. coli* XL-1Blue with pBluescript without insert. Rabbit antiserum 130D, raised against purified pili of *M. nonliquefaciens* NCTC 7784, was used at a dilution of 1:1,000. Molecular size markers noted on the left side are approximately 14 and 18 kDa.

pili of *M. nonliquefaciens* NCTC 7784, was kindly provided by L. O. Frøholm, Folkehelsa, Oslo, Norway, and was diluted 1:1,000 when reacted with the filters.

Electron micrographs. Transmission electron micrographs of *E. corrodens* 31745 SC and N variants grown for 3 days in broth without shaking were obtained by negative staining with 1% uranyl acetate (34).

RESULTS

Imaging *E. corrodens* surface proteins by electron microscopy. Electron micrographs demonstrating surface appendages consistent with pili on an *E. corrodens* 31745 SC variant (Fig. 1A) and no pili on a 31745 N variant (Fig. 1B) are shown. Similar to other type IV pili, these structures are 5 to 6 nm in diameter. These pili were very long and many times the length of the cell; we were unable to determine their exact lengths. Piliation did not seem to be peritrichous; however, there appeared to be multiple origins of bundles of pili on a single cell.

Cloning of the two *E. corrodens* 31745 pilin genes. Since type IV pilin genes are known to have a conserved N-terminal sequence, we used the pilin-specific PCR product of the N-terminal portion of the *M. nonliquefaciens* pilin gene as a genomic Southern hybridization probe to detect the equivalent *E. corrodens* pilin gene(s) (Fig. 2A). Strong hybridization to an approximately 1.4-kb band in the *Sma*I-digested lane and to an approximately 2.5-kb band in the *Kpn*I-digested lane was observed. To our surprise, *ecpA* and *ecpB* gene probes from strain ATCC 23834 did not give nearly as strong a hybridization signal as did the *M. nonliquefaciens* PCR probe (data not shown). Therefore, we used the *M. nonliquefaciens* PCR probe to screen the *E. corrodens* 31745 λZap gene bank and detected multiple hybridizing plaques. After rescreening and plaque purification, one of the hybridizing plaques was chosen for in vivo excision with an M13 helper phage (42), resulting in plasmid pEc2A. The 1.4-kb *Sma*I fragment which hybridized to the PCR probe was then further subcloned in each orientation into vector pGEM7, producing plasmids pEcA4 and pEcB9. A map of pEcB9 is presented in Fig. 3.

Sequencing of the *E. corrodens* pilin gene clone. The sequence of the entire *Sma*I insert in both directions was obtained by using subclones and six oligonucleotides, whose positions are shown in Fig. 3. The nucleotide sequence of the 1,351-bp *Sma*I fragment is presented in Fig. 4, together with the predicted amino acid sequence encoded by the two tandemly arranged pilin genes contained on the fragment. Using previously established nomenclature, we named these pilin genes of strain 31745 *ecpC* and *ecpD*. The AAGG sequence present on the 5′ side of the starting ATGs was (partly) homologous to known ribosome-binding sites in prokaryotes (11, 31). The *ecpC* and *ecpD* genes were predicted to produce mature pilins of 145 amino acids with leader sequences of 8 and 7 amino acids, respectively.

Southern blot analysis with *ecpA* and *ecpB* as probes. PCR products of the cloned genes *ecpC* and *ecpD* of strain 31745 were used as hybridization probes for Southern analysis of strain 31745 DNA (Fig. 2B and C). The resulting hybridization pattern was identical to that observed with the original *M. nonliquefaciens* PCR probe.

Western blot analysis. A *Bst*XI deletion derivative of pEcB9 (pEc1) contains only *ecpC*, and a *Bst*XI deletion derivative of pEcA4 (pEc2) contains only *ecpD* (Fig. 3). As demonstrated in Fig. 5, rabbit antiserum 130D, raised against purified pili of *M. nonliquefaciens*, reacted with a protein band of approximately 15 kDa in whole-cell extracts of *E. coli* containing pEcB9, pEc1, or pEc2. These 15-kDa bands comigrated with purified pili from *M. nonliquefaciens*. In contrast, *E. coli* containing the vector alone consistently made only very small amounts of a cross-reactive protein of this size (Fig. 5, lane 7).

DISCUSSION

Phenotypic traits associated with pili were found in *E. corrodens* 31745, and structures that resembled type IV pili were detected by electron microscopy (Fig. 1). Two tandemly duplicated pilin genes of *E. corrodens* 31745 were cloned and characterized. The two pilins appeared to be complete and were transcribed in the same orientation. The predicted amino acid sequence of both pilin subunits consisted of 145 amino acids of mature protein, with leader sequences of 8 and 7 amino acids for EcpC and EcpD, respectively. Cysteine loops were located at the extreme ends of the C termini, with EcpB possessing three cysteine residues. Western blot analysis showed that the *E. corrodens* 31745 pilin genes were expressed as proteins of approximately 15 kDa in *E. coli*, whether present in tandem (pEcB9 and pEcA4) or in single copy (pEc1 and pEc2), and comigrated with the *M. nonliquefaciens* pilin protein (Fig. 5). The expression of pilins from pEc1 and pEc2 is probably due to promoters provided by the vector. Future studies will clarify whether both genes are normally expressed in *E. corrodens* and whether the resulting pilus structures are homopolymers or heteropolymers.

Comparison of the amino acid sequences of the pilins of *E. corrodens* 31745 showed greatly conserved N-terminal sequences as well as major differences in the hypervariable regions (Fig. 6). The *E. corrodens* 31745 pilin protein EcpC had features of both *N. gonorrhoeae* (27, 29) and *P. aeruginosa* K122-4 (33) pilins, displayed moderate homology with *E. corrodens* 31745 EcpD (this study) and *M. nonliquefaciens* TfpA (49), and differed substantially from *E. corrodens* ATCC 23834 EcpA and EcpB (35) and pilins from *D. nodosus* serogroups A through H (8, 14, 24), *P. aeruginosa* PAK, PAO, and P1 (33, 40), *M. bovis* (11, 22), and *M. lacunata* (37). In contrast, *E. corrodens* 31745 EcpD diverged significantly from all other pilins investigated except that of *P. aeruginosa* K122-4 (33).

The sequence divergence between the two pilins of *E. corrodens* 31745 contrasts strongly with the much higher homology found between pilins EcpA and EcpB of *E. corrodens* ATCC 23834 (35). The discrepancy is so evident that it can be compared with that between the two classes of

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
(A) 31745 EcpC	678																		
(B) 31745 EcpD	280	700																	
(C) 23834 EcpA	212	176	718																
(D) 23834 EcpB	180	148	322	752															
(E) N. gonorr.(P9)	381	242	172	165	758														
(F) N. gonorr.(MS11)	348	238	185	160	574	753													
(G) M. nonliq TfpA	263	235	190	171	214	220	698												
(H) D. nodosus A	208	203	173	147	206	206	248	712											
(I) D. nodosus C	214	194	191	158	215	210	264	468	708										
(J) D. nodosus D	190	183	258	216	180	183	186	188	193	734									
(K) D. nodosus H	199	187	243	220	182	187	181	190	197	562	732								
(L) D. nodosus FimZ	173	168	221	216	154	154	167	160	182	340	335	738							
(M) P. aerug. PAK	217	206	226	195	168	172	217	193	192	240	242	225	682						
(N) P. aerug. PAO	196	203	232	206	166	166	203	204	191	272	264	229	442	653					
(O) P. aerug. P1	167	216	314	253	169	169	188	177	180	287	244	209	268	270	683				
(P) P. aerug. K122-4	387	352	178	166	266	279	262	239	249	223	209	180	251	235	184	693			
(Q) M. lacunata Q	180	171	319	261	165	168	179	157	169	258	244	239	245	218	301	165	715		
(R) M. bovis Q	188	188	395	282	171	170	186	158	168	260	266	224	252	250	351	182	457	694	
(S) M. bovis I	175	167	382	302	159	159	179	153	163	261	249	215	235	249	322	161	448	501	688

FIG. 6. Values are the optimized scores for two-by-two protein sequence comparisons as described in the text. The sequences compared are as follows: A, *E. corrodens* 31745 EcpC (this study); B, *E. corrodens* 31745 EcpD (this study); C, *E. corrodens* ATCC 23834 EcpA (35); D, *E. corrodens* ATCC 23834 EcpB (35); E, *N. gonorrhoeae* P9 pilin (29); F, *N. gonorrhoeae* MS11 pilin (27); G, *M. nonliquefaciens* NCTC 7784 TfpA (49); H, *D. nodosus* serogroup A pilin (8); I, *D. nodosus* serogroup C pilin; J, *D. nodosus* serogroup D pilin (23); K, *D. nodosus* serogroup H pilin (8); L, *D. nodosus* FimZ (14); M, *P. aeruginosa* PAK pilin (32, 40); N, *P. aeruginosa* PAO pilin (40); O, *P. aeruginosa* P1 pilin (33); P, *P. aeruginosa* K122-4 pilin (33); Q, *M. lacunata* ATCC 17956 TfpQ (37); R, *M. bovis* Epp63 TfpQ (21); S, *M. bovis* Epp63 TfpI (9). For practical purposes, 300 was chosen as an arbitrary level of significance.

pilins in *D. nodosus*, which also demonstrate low sequence homology (8, 14, 24) (Fig. 6).

While most type IV pilin proteins from different strains of the same species are quite homologous (i.e., have high optimized scores), this was not the case for the *E. corrodens* 31745 EcpC and EcpD pilin proteins when compared with those of *E. corrodens* ATCC 23834. The only similar situation known to date exists between class I (serogroups A and C in Fig. 4) and class II (serogroups D and H) *D. nodosus* strain pilins (8, 14, 24). In this case, while serogroups within the same class give high homology scores (468 for serogroup A versus C and 562 for serogroup D versus H), comparisons between classes give relatively low scores, i.e., 188 for A versus D, 190 for A versus H, 193 for C versus D, and 197 for C versus H. Therefore, it could be that some strains of *E. corrodens* are similar to *D. nodosus* in having two or more different major groupings, or perhaps they have even greater variability.

P. aeruginosa P1 and K122-4 pilins (with two cysteine loops) are divergent from those of *P. aeruginosa* PAK and PAO (33, 40). Interestingly, the pilins from these two *P. aeruginosa* strains have only limited homology (with a score of 184), each somewhat more homologous to the other than to other *P. aeruginosa* pilins. These are thus additional examples of strains of the same species with very divergent pilins, but in the case of *P. aeruginosa*, there are at least three classes of pilins. It is also notable that *P. aeruginosa* K122-4 pilin has large regions of homology with the strain 31745 EcpC and EcpD pilins, with optimized scores of 387 and 352, respectively. While *P. aeruginosa* P1 pilin matched *E. corrodens* ATCC 23834 EcpA with a score of 314, the ATCC 23834 EcpA matched *M. bovis* and *M. lacunata* pilins, all with scores ranging from 319 to 382.

In considering the significance of the relative similarities and differences seen in the pairwise combinations of pilin proteins (including leader sequences) listed in Fig. 6, the following features should be noted. A minimal baseline similarity exists in all true type IV pilins because of the highly conserved amino termini. The lowest common score in Fig. 6 is the 148 given the match between EcpD of *E. corrodens* 31745 and EcpB of *E. corrodens* ATCC 23834. This score is still highly significant compared with the scores between any of the listed true type IV pilins and the related TcpA pilin from *V. cholerae* (41), which range from a low of 38 compared with *D. nodosus* FimZ to a high of 100 compared with EcpC of strain 31745.

Sequence differences, both within the same species and potentially between different species, arise from selective pressure to avoid host immune responses. Since in type IV pili, unlike many other described bacterial pili, the main pilin structural subunit appears to contain the receptor-binding domain, there is potential selection for similarity in some pilin sequences based on the common ability of being able to bind to the same or similar receptors. If this point is valid, one would predict that pilins of different genera which have high optimization match scores may also have similar receptor-binding features. It is interesting to speculate that the relatively low homology observed between the pilins of *E. corrodens* 31745 and ATCC 23834 is because one strain was isolated from the oral environment and the other from spinal fluid (although from the same region of Europe). It may be that the ability to bind to different receptors in these different host environments determines which strains can be maintained in each of these different host locations.

Regardless of the exact selective pressures for specific pilin sequence similarities found between distantly related type IV pilus-containing bacteria, a second question is how these similarities have arisen. One possibility is that there is a finite range of pilin amino acid sequences that allow correct assembly of pilins into functional pili. In this case, the strong similarities between genera could be due to either chance matches of limited choices during evolution or convergent evolution toward the ability to optimally bind to similar receptors. However, since type IV piliation is also correlated with competence for transformation of naked DNA, it is possible that horizontal gene transfer may be occurring occasionally even between bacteria of different genera and families (25). Once enough is known about codon usage in the different species harboring type IV pili, comparison of pilin gene codon usage to regular codon usage within the other genes of the same species may make it possible to determine whether independent evolution or horizontal gene transfer is the predominant mode of generating similar pilins in widely different bacteria.

ACKNOWLEDGMENTS

This work was supported by grants from Anders Jahres Foundation and Valborg Aschoug's Legat to T.T., from the Norwegian-America Association (the Marshall Fund) to S.W., and from the NIH (grant EY-07125) to C.F.M.

We thank N. Freitag for critical reading of the manuscript.

REFERENCES

1. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1513–1523.
2. **Brooks, G. F., J. M. O'Donoghue, J. P. Rissing, K. Soapes, and J. W. Smith.** 1974. *Eikenella corrodens*, a recently recognized pathogen: infections in medical-surgical patients and in association with methylphenidate abuse. Medicine **53**:325–342.
3. **Chen, C. K. C., R. G. Dunford, H. S. Reynolds, and J. J. Zambon.** 1989. *Eikenella corrodens* in the human oral cavity. J. Periodontol. **60**:611–616.
4. **Chen, C. K. C., T. V. Potts, and M. E. Wilson.** 1990. DNA homologies shared among *E. corrodens* isolates and other corroding bacilli from the oral cavity. J. Periodont. Res. **25**:106–112.
5. **Dalrymple, B., and J. S. Mattick.** 1987. An analysis of the organization and evolution of type 4 fimbrial (MePhe) subunit proteins. J. Mol. Evol. **25**:261–269.
6. **Dewhirst, F. E., B. J. Paster, and P. L. Bright.** 1989. *Chromobacterium*, *Eikenella*, *Kingella*, *Neisseria*, *Simonsiella*, and *Vitreoscilla* species comprise a major branch of the beta group *Proteobacteria* by 16S ribosomal ribonucleic acid sequence comparison: transfer of *Eikenella* and *Simonsiella* to the family *Neisseriaceae* (emend.). Int. J. Syst. Bacteriol. **39**:258–266.
7. **Eiken, M.** 1958. Studies on an anaerobic, rod-shaped Gram-negative microorganism: *Bacteroides corrodens* n.sp. Acta Pathol. Microbiol. Scand. **43**:404–416.
8. **Elleman, T. C., and P. A. Hoyne.** 1984. Nucleotide sequence of the gene encoding pilin of *Bacteroides nodosus*, the causal organism of ovine footrot. J. Bacteriol. **160**:1184–1187.
9. **Fulks, K. A., S. P. Stephens, M. R. Green, and C. F. Marrs.** 1990. Sequence analysis of the inversion region containing the pilin genes of *Moraxella bovis*. J. Bacteriol. **172**:310–316.
10. **Giron, J., A. S. Yue, and G. K. Schoolnik.** 1991. An inducible bundle-forming pilus of enteropathogenic *Escherichia coli*. Science **254**:710–713.
11. **Gold, L., D. Pribnow, T. Schneider, S. Shinedling, B. Swebilius-Singer, and G. Stormo.** 1981. Translational initiation in prokaryotes. Annu. Rev. Microbiol. **35**:365–403.
12. **Henriksen, S. D.** 1969. Corroding bacteria from the respiratory tract. II. *Bacteroides corrodens*. Acta Pathol. Microbiol. Scand. **75**:91–96.
13. **Hermodson, M. A., K. C. S. Chen, and T. M. Buchanan.** 1978. *Neisseria* pilin proteins: amino-terminal amino-acid sequences

and identification of an unusual amino acid. Biochemistry **17:**442–445.

14. **Hobbs, M., B. P. Dalrymple, P. T. Cox, S. P. Livingstone, S. F. Delaney, and J. S. Mattick.** 1991. Organization of the fimbrial gene region of *Bacteroides nodosus*: class I and class II strains. Mol. Microbiol. **5:**543–560.
15. **Hull, R., R. E. Gill, P. Hsu, B. H. Minshew, and S. Falkow.** 1981. Construction and expression of recombinant plasmids encoding type 1 or D-mannose-resistant pili from a urinary tract infection *Escherichia coli* isolate. Infect. Immun. **33:**933–938.
16. **Jackson, F. L., and Y. E. Goodman.** 1972. Transfer of the facultatively anaerobic organism *Bacteroides corrodens* Eiken to a new genus, *Eikenella*. Int. J. Syst. Bacteriol. **22:**73–77.
17. **Janda, D. H., D. H. Ringler, J. K. Hilliard, R. C. Hankin, and F. M. Hankin.** 1990. Nonhuman primate bites. J. Orthop. Res. **8:**146–150.
18. **Klein, B., J. Couch, and J. Thompson.** 1990. Ocular infections associated with *Eikenella corrodens*. Am. J. Ophthalmol. **109:** 127–131.
19. **Lipman, D. J., and W. R. Pearson.** 1985. Rapid and sensitive protein similarity searches. Science **227:**1435–1441.
20. **Marrs, C. F., F. W. Rozsa, M. Hackel, S. P. Stephens, and A. C. Glasgow.** 1990. Identification, cloning, and sequencing of *piv*, a new gene involved in inverting the pilin genes of *Moraxella lacunata*. J. Bacteriol. **172:**4370–4377.
21. **Marrs, C. F., W. W. Ruehl, G. K. Schoolnik, and S. Falkow.** 1988. Pilin gene phase variation of *Moraxella bovis* is caused by an inversion of the pilin genes. J. Bacteriol. **170:**3032–3039.
22. **Marrs, C. F., G. K. Schoolnik, J. M. Koomey, J. Hardy, J. Rothbard, and S. Falkow.** 1985. Cloning and sequencing of a *Moraxella bovis* pilin gene. J. Bacteriol. **163:**132–139.
23. **Mattick, J. S.** 1989. The molecular biology of the fimbriae of *Bacteroides nodosus* and the development of a recombinant DNA-based vaccine against footrot, p. 195–218. *In* J. R. Egerton, W. K. Yang, and G. G. Rifkin (ed.), Footrot and foot abscess of ruminants. CRC Press, Boca Raton, Fla.
24. **Mattick, J. S., B. J. Anderson, P. T. Cox, B. P. Dalrymple, M. M. Bills, M. Hobbs, and J. R. Egerton.** 1991. Gene sequences and comparison of the fimbrial subunits representative of *Bacteroides nodosus* serotypes A to I: class I and class II strains. Mol. Microbiol. **5:**561–573.
25. **Mazodier, P., and J. Davies.** 1991. Gene transfer between distantly related bacteria. Annu. Rev. Genet. **25:**147–171.
26. **McKern, N. M., I. J. O'Donnel, A. S. Inglis, D. J. Stewart, and B. L. Clark.** 1983. Amino acid sequence of pilin from *Bacteroides nodosus* (strain 198), the causative organism of ovine footrot. FEBS Lett. **164:**149–153.
27. **Meyer, T. F., E. Billyard, R. Haas, S. Storzbach, and M. So.** 1984. Pilus genes of *Neisseria gonorrhoeae*: chromosomal organization and DNA sequence. Proc. Natl. Acad. Sci. USA **81:**6110–6114.
28. **Needleman, S. B., and C. D. Wunsch.** 1970. A general method applicable to the search for similarities in the amino acid sequence of two proteins. J. Mol. Biol. **48:**443–453.
29. **Nicolson, I. J., A. C. F. Perry, M. Virji, J. E. Heckels, and J. R. Saunders.** 1987. Localization of antibody-binding sites by sequence analysis of cloned pilin genes from *Neisseria gonorrhoeae*. J. Gen. Microbiol. **133:**825–833.
30. **Nunn, D., S. Bergman, and S. Lory.** 1990. Products of three accessory genes, *pilB*, *pilC*, and *pilD*, are required for biogenesis of *Pseudomonas aeruginosa* pili. J. Bacteriol. **172:**2911–2919.
31. **Pasloske, B. L., D. S. Drummond, L. S. Frost, and W. Paranchych.** 1989. The activity of the *Pseudomonas aeruginosa* pilin promoter is enhanced by an upstream regulatory site. Gene **81:**25–34.
32. **Pasloske, B. L., B. B. Finlay, and W. Paranchych.** 1985. Cloning and sequencing of the *Pseudomonas aeruginosa* PAK pilin gene. FEBS Lett. **183:**408–412.
33. **Pasloske, B. L., P. A. Sastry, B. B. Finlay, and W. Paranchych.** 1988. Two unusual pilin sequences from different isolates of *Pseudomonas aeruginosa*. J. Bacteriol. **170:**1691–1697.
34. **Progulske, A., and S. C. Holt.** 1980. Transmission-scanning electron microscopic observations of selected *Eikenella corrodens* strains. J. Bacteriol. **143:**1003–1018.
35. **Rao, V., and A. Progulske-Fox.** 1993. Identification of a novel type 4 pilin gene arrangement: cloning and sequencing of two *N*-methylphenylalanine pilin genes from *Eikenella corrodens*. J. Gen. Microbiol. **139:**651–660.
36. **Rossau, R., G. Vandenbussche, S. Thielemans, P. Segers, H. Grosch, E. Göthe, W. Mannheim, and J. De Ley.** 1989. Ribosomal ribonucleic acid cistron similarities and deoxyribonucleic acid homologies of *Neisseria*, *Kingella*, *Eikenella*, *Simonsiella*, *Alysiella*, and Centers for Disease Control groups EF-4 and M-5 in the emend family *Neisseriaceae*. Int. J. Syst. Bacteriol. **39:**185–198.
37. **Rozsa, F. W., and C. F. Marrs.** 1991. Interesting sequence differences between the pilin gene inversion regions of *Moraxella lacunata* ATCC 17956 and *Moraxella bovis* EPP63. J. Bacteriol. **173:**4000–4006.
38. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed., p. 1.90–1.104. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
39. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
40. **Sastry, P. A., B. B. Finlay, B. L. Pasloske, W. Paranchych, J. R. Pearlstone, and L. B. Smillie.** 1985. Comparative studies of the amino acid and nucleotide sequences of pilin derived from *Pseudomonas aeruginosa* PAK and PAO. J. Bacteriol. **164:**571–577.
41. **Shaw, C. E., and R. K. Taylor.** 1990. *Vibrio cholerae* O395 *tcpA* pilin gene sequence and comparison of predicted protein structural features to those of type 4 pilins. Infect. Immun. **58:**3042–3049.
42. **Short, J. M., J. M. Fernandez, J. A. Sorge, and W. D. Huse.** 1988. λ Zap: a bacteriophage λ expression vector with *in vivo* expression properties. Nucleic Acids Res. **16:**7583–7600.
43. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98:**503–517.
44. **Stoloff, A. L., and M. L. Miller.** 1986. Infections with *Eikenella corrodens* in a general hospital: a report of 33 cases. Rev. Infect. Dis. **8:**50–53.
45. **Swanson, J., and J. M. Koomey.** 1989. Mechanisms for variation of pili and outer membrane protein II in *Neisseria gonorrhoeae*, p. 743–761. *In* D. Berg and M. Howe (ed.), Mobile DNA. American Society for Microbiology, Washington, D.C.
46. **Tanner, A. C. R., C. Haffer, G. T. Brathal, R. A. Visconti, and S. S. Socransky.** 1979. A study of the bacteria associated with advancing periodontitis in man. J. Clin. Periodontol. **6:**278–307.
47. **Taylor, R. K., V. L. Miller, D. B. Furlong, and J. J. Mekalanos.** 1987. Use of *phoA* gene fusions to identify a pilus colonization factor coordinately regulated with cholera toxin. Proc. Natl. Acad. Sci. USA **81:**2833–2837.
48. **Tønjum, T., N. Hagen, and K. Bøvre.** 1985. Identification of *Eikenella corrodens* and *Cardiobacterium hominis* by genetic transformation. Acta Pathol. Microbiol. Immunol. Scand. Sect. B **93:**389–394.
49. **Tønjum, T., C. F. Marrs, F. Rozsa, and K. Bøvre.** 1991. The type 4 pilin of *Moraxella nonliquefaciens* exhibits unique similarities with the pilins of *Neisseria gonorrhoeae* and *Dichelobacter* (*Bacteroides*) *nodosus*. J. Gen. Microbiol. **137:**2483–2490.
50. **Towbin, H., T. Staehelin, and J. Gordon.** 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc. Natl. Acad. Sci. USA **76:**4350–4354.
51. **West, S. E. H., and V. L. Clark.** 1989. Genetic loci and linkage associations in *Neisseria gonorrhoeae* and *Neisseria meningitidis*. Clin. Microbiol. Rev. **2:**S92–S103.
52. **Zhang, Q. Y., D. DeRyckere, P. Lauer, and M. Koomey.** 1992. Gene conversion in *Neisseria gonorrhoeae*: evidence for its role in pilus antigenic variation. Proc. Natl. Acad. Sci. USA **89:** 5366–5370.