Distribution of *drb* Genes Coding for Dr Binding Adhesins among Uropathogenic and Fecal *Escherichia coli* Isolates and Identification of New Subtypes

Lixin Zhang,1 Betsy Foxman,1 Pat Tallman,1 Eduardo Cladera,1 Chantal Le Bouguenec,2 and Carl F. Marrs1*

Department of Epidemiology, University of Michigan School of Public Health, Ann Arbor, Michigan 48109,1 and Unité de Pathogénie Bactérienne des Muqueuses, Institut Pasteur, Paris, France2

Received 21 November 1996/Returned for modification 15 January 1997/Accepted 13 March 1997

The Dr family of related adherence structures, some fimbriated and others afimbriated, bind to decay-accelerating factor molecules on human cells. Dr is associated with recurring urinary tract infection (UTI), but the distribution of Dr subtypes among uropathogenic *Escherichia coli* causing UTI among otherwise healthy women has yet to be described. A total of 787 UTI and fecal *E. coli* isolates from college women were screened for the presence of Dr sequences (*drb*). Fifteen percent of UTI strains were *drb* positive, compared to 5% of fecal strains. The adhesin (E gene) subtype of each *drb*-positive strain was determined by type-specific PCR followed by restriction enzyme analysis. Among 78 *drb*-positive strains, we found 14 (18%) *afaE1*, 1 (1.3%) *afaE2*, 2 (1.3%) *afaE3*, 9 (12%) *draE*, 9 (12%) *draE-afaE3* hybrid, 1 (1.3%) *daaE*, 32 (41%) *afaE5*, 4 (5.1%) *F131* E gene-like, and 7 untypeable strains. All untypeable E genes were cloned and sequenced, revealing four additional new classes of E genes, including two similar to the previously identified nonfimbrial E series. While a great range of diversity exists among the E genes, restriction fragment length polymorphism analysis demonstrated that all of these *dra* operons share a highly conserved gene structure. The most common subtype, *afaE5*, occurred three times as often among UTI than fecal strains. Over half of the *drb*-positive strains and 80% of those positive for *afaE5* have the same virulence signature (positive for *aer*, *kpsMT*, *ompT*, and * fim*), suggesting an association of this profile with UTI pathogenesis.

Bacterial adherence to host cells is a crucial step in the initiation of various infectious diseases (9, 24). An array of adhesins including type 1, P, S, and G fimbriae as well as adhesins of the Dr family and nonfimbrial adhesins (NFAs) have been identified in *Escherichia coli* strains causing urinary tract infection (UTI) (17, 19, 22). Adherence factors facilitate the colonization of the urinary tract and promote *E. coli* colonization and persistence in the colon or vagina, which may serve as a reservoir for ascending infection in the urinary tract (10, 22). The variety of adhesins, as well as the antigenic diversity of certain adhesins, may extend bacterial tissue tropism to different sites, facilitating the multiple stages of the pathogenic process.

Several pieces of evidence suggest that Dr adhesins play an important role in UTI pathogenesis. Dr adhesins occur more frequently among *E. coli* UTI isolates than fecal isolates (10, 22), and the presence of Dr at the time of a first UTI has been associated with an increased risk of a second UTI (13). Dr receptors, the Dr blood group antigen component of decay-accelerating factor (31, 32), are widely distributed along the urinary tract, underlining the potential importance of Dr adhesin-producing *E. coli* in ascending colonization of the urinary tract. Furthermore, a pathogenicity study using a mouse model demonstrated that a Dr hemagglutinin-positive *E. coli* strain was much more virulent in terms of causing persistent and severe UTI than its isogenic mutant (10, 18).

Structural and assembly genes required to produce adhesins of the Dr family are present in operons, each consisting of at least five genes often named A to E (a representative *afa-I* operon is shown in Fig. 1) (26). This fimbrial gene organization is quite different from that for other operons such as type 1 and P-pilus operons but are similar to that reported for the adhesin gene cluster of enteroaggregative *E. coli* (34). Other potential regulatory genes located upstream of gene A have been identified in some adhesin-encoding gene clusters (5, 16). In all but the *dra* operon (coding for Dr hemagglutinin), genes A to D encoding mostly accessory functions among the family are very similar in terms of amino acid sequence homology and functional homology. The last gene in each operon, E, encodes a protein that functions both as the major structural protein and as an adhesin molecule. Most of the sequence heterogeneity between family members occurs in the E gene. Similar genes exist in the *dra* operon, but the *dra* genes were named in reverse alphabetical order compared to their counterparts in other operons; thus, *draA* encodes the adhesin/structural protein (33). Recently, Carnoy and Moseley (7b) renamed the *dra* genes to fit the order found in the *afa* genes. In this report, we are also following this latter naming system shared by the majority of *drb* operons, in which E is the last gene in each operon.

Dr hemagglutinin, F1845, *AfaE-I*, and *AfaE-III*, have been studied in detail. Despite the similarities among them, interesting differences exist. Dr hemagglutinin, F1845, *AfaE-I*, and *AfaE-III* each bind to functionally distinct sites on decay-accelerating factor (32). Dr hemagglutinin and F1845 have fimbrial structures, while *AfaE-I* and *AfaE-III* are afimbrial. While Dr hemagglutinin and *AfaE-III* differ in only three amino acids, they differ in fimbriation, and Dr hemagglutinin recognizes type IV collagen as an additional receptor and is chloramphenicol sensitive; *AfaE-III* is not (38). The phenotypic differences among adhesin subtypes of the Dr family are probably due to sequence differences in their E genes (15, 28,
The differences in chloramphenicol sensitivity for hemagglutination between AfaE-III and Dr hemagglutinin were first shown to reside in the N-terminal 54 amino acids (36) and then specific localized to the presence of an asparagine residue and an aspartic acid residue at position 52, respectively (28). Other hybrid experiments have demonstrated that the N-terminal regions are critical for determination of fimbrial status (15). The existence of other adhesin variants such as AfaE-II and the newly discovered AfaE-V (14) and F131 (30) suggests even greater sequence heterogeneity in the genetic determinant encoding the structural adhesins of the Dr family.

Most previous studies describing the prevalence of *drb* among UTI isolates have used probes that cross-react with all members of the Dr family. Thus, the relative distribution of each subtype is unknown. Furthermore, many previous studies have had relatively small sample sizes and screened UTI isolates from laboratory collections rather than from defined populations (3, 8). In this study, we describe the frequency of the *drb* gene cluster and adhesin subtypes in a large collection of *E. coli* isolates from a well-defined population sample representative of otherwise healthy young women with first-time UTI and compare them with that of a fecal collection from healthy women of the same age group. We do similar comparisons with smaller collections of *E. coli* isolates from second UTI and from recurrent UTI (RUTI) (12). We also describe the antigenic diversity of adhesin-encoding genes within this population and four new adhesin subtypes. The co-occurrence of *drb* genes with other virulence factors is also investigated, and virulence profiles of *drb*-positive strains are presented.

### MATERIALS AND METHODS

#### Sample collection.
We used *E. coli* isolates from three different collections. Urinary *E. coli* isolates were collected from a cohort of 304 women 18 to 40 years of age with a first UTI followed up for recurrence during the next 6 months. A total of 315 first UTI isolates and 49 second UTI isolates were collected at the University of Michigan Health Service and the University of Texas at Austin Health Service during February and March 1996. Twenty-eight *E. coli* isolates were collected using the Cultureswab Transport System (Difco Laboratories, Detroit, Mich.) and streaked for isolation on both Trypticase-soy agar and *MacConkey* agar. Any phenotypically different colonies were identified and processed by using the same procedures as used for urinary *E. coli* (12).

#### Bacterial strains, plasmids, and culture conditions.
Strain TOP10F* (Invitrogen, San Diego, Calif.) was used as the host strain for recombinant plasmid studies. The following *E. coli* strains were used as controls for DNA hybridization and PCR studies: C248 (AfaE-I [afa-1]), C1845 (F1845 [daa]), and C131 (F131) were provided by S. L. Moseley of the University of Washington; and KS52 (AfaE-I [afa-1]), A22 (AfaE-II [afa-2] and AfaE-III [afa-3]), and AL851 (AfaE-V [afa-5]) (Invitrogen) was used as the cloning vector. Plasmid pLI14 (provided by A. Labigne-Roussel, Institut Pasteur, Paris, France), containing the *afa-1* operon, was used as the source for DNA probes. All clinical and control strains were grown in Luria-Bertani (LB) broth or on LB plates at 37°C. *E. coli* HB101 carrying plasmid pLI14 was grown in LB broth or on LB plates with ampicillin (100 μg/ml). Low-salt LB agar plates with isopropylthio-β-D-galactoside (50 μg/ml) and Zeocin (50 μg/ml) were used for the selection of transformants as instructed by the manufacturer (Invitrogen).

#### Hemagglutination.
We used the method of Nowicki et al. (32) for hemagglutinations. Briefly, normal human group O erythrocytes were washed in phosphate-buffered saline (136 mM NaCl, 1 mM NaH2PO4, 8 mM Na2HPO4) and resuspended to 3% (vol/vol) in phosphate-buffered saline with 2.5% α-methyl-d-mannoside. Hemagglutination was performed in microtiter plate wells at room temperature and at 4°C by mixing equal volumes (40 μl) of serial twofold bacterial dilutions and erythrocyte suspensions. Inhibition of mannose-resistant hemagglutination (MRHA) by chloramphenicol was tested by the addition of 10 μM chloramphenicol to an identical erythrocyte suspension prior to the assay for hemagglutination.

#### DNA isolation and DNA probes.
Plasmid DNAs used for enzymatic manipulation were routinely prepared by an alkaline lysis procedure (21) and by using a plasmid preparation kit (Qiagen, Chatsworth, Calif.). Routine plasmid screening was done by the rapid lysis method of Kado and Liu (23). Larger amounts of plasmid DNA were prepared by an alkaline lysis procedure (20) followed by gel electrophoresis. Either a 2.6-kb *Pst*I fragment or three *Pst*I fragments (2.6, 1.1, and 0.4 kb) were cut out, and the DNA was eluted. The purified DNA fragments were used as templates to generate digoxigenin-labeled DNA probes (Genius System kit; Boehringer Mannheim, Indianapolis, Ind.).

#### Dot and Southern blots.
We screened for virulence factors by using dot blot hybridization as previously described (12). Both dot blot and Southern blots were hybridized to digoxigenin-labeled DNA probes, using conditions described previously (12).

#### Subtyping by PCR and restriction digestion.
The subtype of each *drb*-positive strain was determined by PCR using a conserved primer derived from a common 5' flanking region and a selective primer derived from the unique coding region within each known subtype gene. Table 1 lists the oligonucleotide primers used in PCR subtyping. Full-length adhesin-encoding genes of four Dr subtypes, AfaE-III, Dr hemagglutinin, AfaE-V, and F131, were amplified by a pair of primers designed to contain the unique coding region and a selective primer derived from the unique coding region. Amplification reactions were carried out with a MicroCycler (Eppendorf Inc., Cornelius, Calif.). To differentiate between AfaE-III and Dr hemagglutinin, restriction patterns were compared by digestion of PCR products with enzymes *Tp*922 and *Hpa*509I. To differentiate between AfaE-V and F131, restriction patterns were compared by digestion of PCR products with enzymes *Ban*HI and *Avr*II.

#### TABLE 1. Oligonucleotides used in PCR typing

<table>
<thead>
<tr>
<th>Name</th>
<th>Oligonucleotide</th>
<th>Strand</th>
<th>Size (bp)*</th>
<th>Reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conserved primer</td>
<td>5’-TTAGACCGTACGTGTTGTTACC-3’</td>
<td>+</td>
<td>487</td>
<td>This study</td>
</tr>
<tr>
<td>AfaE-I-type primer</td>
<td>5’-GTTGGTCTGGTTGTTACATTTCC-3’</td>
<td>–</td>
<td>375</td>
<td>28</td>
</tr>
<tr>
<td>AfaE-II-type primer</td>
<td>5’-GTTCCCGATAGACTGGAATGAAGC-3’</td>
<td>–</td>
<td>375</td>
<td>27p</td>
</tr>
<tr>
<td>AfaE-III/Dr-type primer</td>
<td>5’-ACCATTTGCTGCTTTAGCGCCGCGGAG-3’</td>
<td>–</td>
<td>408</td>
<td>28, 35</td>
</tr>
<tr>
<td>AfaE-V/F131-type primer</td>
<td>5’-AGCATCGCCGCGGTACGCT-3’</td>
<td>–</td>
<td>429</td>
<td>27, 30</td>
</tr>
<tr>
<td>F1845-type primer</td>
<td>5’-CATGTTGTCATTGGTCGCCCTG-3’</td>
<td>–</td>
<td>353</td>
<td>6</td>
</tr>
<tr>
<td>Degenerate primer</td>
<td>5’-TNAANTNNCCAGANNCCNCC-3’</td>
<td>–</td>
<td>530</td>
<td>This study</td>
</tr>
</tbody>
</table>

* Distance of PCR fragment generated by using the conserved primer plus each addition primer.
* Distance of PCR fragment generated by using the conserved primer plus each addition primer.

![Diagram of the afa-1 gene operon](http://iai.asm.org/)
Cloning and sequencing. Two approaches were used for cloning E genes of the new subtypes. A conserved primer and a degenerative primer (Table 1) were used in PCR to amplify E gene sequences from strains with untypable subtypes. The amplified DNA was cloned into plasmid vector pZErO-1. Two independent recombinant molecules constructed from the insertion of each PCR product were selected. Alternatively, DNA from these strains was isolated and digested with Smal, and the appropriate DNA fragment was identified by Southern hybridization with the DNA probe A (Fig. 1) and eluted from the gel after agarose gel electrophoresis. The DNA fragments were then purified and cloned into the plasmid vector. The recombinant plasmid DNAs were prepared by using a plasmid kit (Qiagen). The sequence of the double-stranded DNA was determined at the University of Michigan Molecular Biology Core Facility by an Applied Biosystems model 373A automated sequencer. When required, new internal oligonucleotide primers were synthesized on the basis of the determined sequence and used as new sequence primers. GenBank accession numbers for these sequences are U87540 and U87790.

Other genetic methods. Restriction enzyme digestion, nucleotides treatment, and ligation were performed by standard procedures according to the methods of Mamiatis et al. (29) and the manufacturer's instructions. Transformation was accomplished by calcium chloride preparation of competent cells. DNA was recovered from agarose gels by using an Elu-Quick DNA purification kit (Schlecher & Schuell, Keene, N.H.). PCR-amplified DNA was purified by using a QIAquick Spin PCR purification kit (Qiagen). The afaD genes were amplified by PCR using a pair of primers (5'-CGTAGCCTCAGCGGAGGAC-3'; 5'-TTCCCTGGGCACCACAGG-3') derived from afaD3 internal sequences (15).

Data analysis. Differences between groups were tested by using y2, a test for equality of two proportions, or, for smaller sample sizes, the Fisher exact test. Excel software (Microsoft, Redmond, Wash.) was used for data entry. Software packages from DNAStar (Madison, Wis.) were used for DNA sequence and amino acid sequence analysis and comparison. For amino acid comparisons, a Lipman-Pearson protein alignment (Ktuple 2; gap penalty 5) was defined as 5%

RESULTS

Detection of drb sequences among E. coli isolates. A total of 787 E. coli strains from otherwise healthy college women with a first UTI, a second UTI, or RUTI and from fecal specimens of college women without UTI were screened by dot blot hybridization for the presence of a sequence homologous to that of drb with a conserved DNA probe derived from the afa-1 operon (probe A in Fig. 1). Strains positive by dot blot hybridization were confirmed by Southern hybridization using the same probe. Sequence homology to drb was found in 49 of the 315 first UTI strains (16%), 7 of the 49 second UTI strains (14%), 2 of the 28 RUTI strains (7.1%), and 20 of the 395 fecal strains (5%). 3.6% of the 315 first UTI strains (16%), 2 of the 49 second UTI strains (14%), 2 of the 28 RUTI strains (7.1%), and 20 of the 395 fecal strains (5%) had the nucleotides A, C, and T at these positions (1). Hybridization with the drb gene probe was 3.5 (odds ratio) (95% confidence interval, 2.0 to 5.9) times more common among E. coli isolates from patients with first cystitis than from control fecal isolates.

Adhesin subtype classification. Because of the homology between the adhesin subtypes, DNA hybridization cannot distinguish one subtype from another. We used PCR to differentiate between subtypes for each adhesin primer from the sequence immediately upstream of the adhesin-encoding gene that was conserved among all subtypes, and type primers derived from various regions of the E gene sequence where each given subtype had a sequence uniquely different from those of all other subtypes (Table 1). However, there were two exceptions: since afaE3 and draE are 99.4% identical, and afaE5 and f131E are 97% identical, at the DNA level, the same type primer was used within each pair. AfaE-III and Dr, and AfaE-V and F131, were thus classified into the same groups by PCR. We confirmed that the correct subtype was amplified in each PCR by showing the different-size PCR fragments were produced by the different types of primers (353 to 487 bp [Table 1]), a result verified by agarose gel electrophoresis (Fig. 2).

Discovery of afaE3-draE and draE-afaE3 hybrids. To differentiate between AfaE-III and Dr hemagglutinins as well as between AfaE-V and F131, we used the minor changes in endonuclease restriction sites that caused the difference between the respective pairs of sequences. A 534-bp DNA fragment containing the entire E gene from all four subtypes was PCR amplified. Of the PCR-amplified DNA fragments, AfaE-V was distinct from F131 by a BamHI restriction site (none in AfaE-V and one in F131) and by a AvaII restriction site (two in AfaE-V and one in F131). AfaE-III was distinct from Dr hemagglutinin by Tsp509I restriction sites (one in AfaE-III and two in Dr hemagglutinin) and by a Hsp92I restriction site (none in AfaE-III and one in Dr hemagglutinin). By the restriction pattern, the F131–AfaE-V group could be clearly broken down into two expected subtypes: 4 F131 types and 32 AfaE-V types. This was not true for the AfaE-III–Dr hemagglutinin group; of 19 strains, 1 and 9 could be classified as AfaE-III and Dr, respectively, by both restriction sites. Eight strains were Dr by Tsp509I restriction pattern but AfaE-III by Hsp92I restriction pattern (afaE3-draE hybrids). One was AfaE-III by Tsp509I restriction pattern but Hsp92I restriction pattern (draE-afaE3 hybrid).

One representative strain from each of the afaE3-draE and draE-afaE3 hybrid groups was chosen for sequence analysis. The three nucleotides that differentiate afaE3 and draE are at positions +217 (an A in afaE3 or a G in draE), +326 (T or C, respectively), and +395 (C or T, respectively). The sequenced afaE3-draE hybrid had the nucleotides A, C, and T at these three sites and also had one nucleotide different from either afaE3 or draE at position +22. This difference is predicted to change one amino acid in the signal sequence from Ala to Thr. The sequenced draE-afaE3 hybrid had two silent changes at positions +24 and +411. Since position +217 (amino acid 52 in the mature adhesin) has been shown to be critical for both fimbriation and receptor binding characteristics (15), we predicted that afaE3-draE hybrids would be phenotypically AfaE-III-like and that the draE-afaE3 hybrid would be phenotypically Dr-like. A hemagglutination assay (see Materials and Methods) was used to test for inhibition of MRHA by chloramphenicol. As predicted, the four hemagglutination-positive afaE3-draE hybrids were resistant to inhibition by chloramphenicol (similar to afaE3-positive strains), while the draE-afaE3 hybrid had its hemagglutination inhibited by chloramphenicol (similar to draE-positive strains).
**Subtype distribution by strain source.** Every known Dr adhesin subtype was found in our urinary E. coli collection, but the subtypes occurred at different frequencies. Among the *drb*-positive UTI strains, 41% were AfaE-V, 18% were the AfaE-I subtype, 13% were Dr-like (nine *draE* plus one *afaE3-* *draE* hybrid), 12% were AfaE-III-like (one *afaE3* plus eight *afaE3-* *draE* hybrids), 5% were F131, 1% were AfaE-II, and 1% were the F1845 subtype. While the predominant subtypes were the same in both UTI and fecal isolates, the relative frequencies were very different (Table 2). Every subtype was found more frequently among urinary than fecal isolates. The most common subtype, AfaE-I, was over three times more frequent in first UTI than fecal isolates (*P* = 0.01), as were the Dr-like subtypes (*P* = 0.03).

**Cloning and sequencing of new adhesin-encoding genes.** Seven of the 78 total *drb* probe-positive strains (four UTI strains and three fecal strains) could not be classified as any known subtype by PCR. These strains could represent additional adhesin variants or deletions in the gene clusters. To address this question, we used both PCR cloning and genomic cloning to identify potential DNA segments that carry alternative or deleted E genes. The cloned DNA segments were sequenced to identify the relatedness to known Dr subtypes.

All known subtypes have a conserved DNA sequence immediately upstream of the E gene and share two conserved amino acids at the C-terminal end. Thus, we hypothesized that additional adhesin variants could have similar sequence conservation. A pair of primers was used to amplify the full length of the E gene region in the seven strains with a possible new adhesin subtype: the same conserved primer used in PCR typing and a downstream degenerate primer (Table 1) including all DNA sequences which could encode the conserved C-terminal amino acids and stop codon. Various amplification conditions were tried, and a DNA fragment with the expected size (~580 bp) was obtained from four of the seven strains. These DNA fragments were cloned and sequenced. Sequence results identified two new classes of E genes, *drb*121 and *drb*122. The *drb*121 gene gives a predicted amino acid sequence for Dr121 that matches most closely that of F1845 (48% identical, 69% similar), while *drb*122 gives a predicted amino acid sequence for Dr122 that matches most closely those of F131 and AfaE-V (78% identical, 88% similar).

The remaining three isolates did not produce a band by PCR. Based on Southern hybridization and PCR analysis, these three strains differ in the 5’ flanking regions; thus, the conserved primer does not work in these three strains (data not shown). We then used genomic cloning to isolate the E gene from the three remaining strains. Total DNA was digested with *Sma*I. DNA fragments of about 3 and 3.5 kb, which should include the E gene region, were identified by Southern hybridization using an ~200-bp DNA probe that hybridized to sequences downstream of the *Sma*I site (Fig. 1). DNA fragments were purified from the agarose gel and cloned into the plasmid vector pZErO-1. Upon sequencing the cloned fragments, we identified two additional classes of E genes, which we named *nfa*111 and *nfa*116. While *Drb*121 and *Drb*122 appear to belong to the previously defined Dr family, NFA111 and NFA116 have no amino acid sequence similarity with any Dr adhesin subtype. NFA111 and NFA116 do have sequences similar to the NFA series (Table 3). Although NFAs have not previously

### TABLE 2. Percent *drb* subtype distribution among all strains screened by strain source

<table>
<thead>
<tr>
<th>Group (no. of isolates)</th>
<th>No. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AfaE-I</td>
</tr>
<tr>
<td>1st UTI (315)</td>
<td>8 (2.5)</td>
</tr>
<tr>
<td>Fecal (395)</td>
<td>4 (1.0)</td>
</tr>
<tr>
<td>2nd UTI (49)</td>
<td>2 (4.1)</td>
</tr>
<tr>
<td>RUTI (28)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>All sources (787)</td>
<td>14 (1.8)</td>
</tr>
</tbody>
</table>

* Includes eight *afaE3-* *draE* hybrids that are predicted to have phenotypes similar to that of AfaE-III (see the text).

* Includes one *draE-* *afaE3* hybrid that is predicted to have a Dr phenotype (see the text).

### TABLE 3. Pairwise percent identity* of deduced protein sequence among various *drb* subtypes

<table>
<thead>
<tr>
<th>Subtype</th>
<th>AfaE-I</th>
<th>AfaE-II</th>
<th>AfaE-III</th>
<th>Dr*</th>
<th>F1845</th>
<th>AfaE-V</th>
<th>F131</th>
<th>Drb121</th>
<th>Drb122</th>
<th>NFAI</th>
<th>NFA111</th>
<th>NFA116</th>
</tr>
</thead>
<tbody>
<tr>
<td>AfaE-I</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AfaE-II</td>
<td>27</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AfaE-III</td>
<td>37</td>
<td>31</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dr*</td>
<td>37</td>
<td>31</td>
<td>98</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F1845</td>
<td>43</td>
<td>26</td>
<td>57</td>
<td>57</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AfaE-V</td>
<td>32</td>
<td>28</td>
<td>52</td>
<td>52</td>
<td>48</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F131</td>
<td>32</td>
<td>28</td>
<td>51</td>
<td>51</td>
<td>47</td>
<td>93</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drb121</td>
<td>43</td>
<td>25</td>
<td>45</td>
<td>46</td>
<td>48</td>
<td>41</td>
<td>41</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drb122</td>
<td>32</td>
<td>31</td>
<td>47</td>
<td>47</td>
<td>43</td>
<td>78</td>
<td>78</td>
<td>38</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NFAI*</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NFA111</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>38</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>NFA116</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>98</td>
<td>39</td>
<td>100</td>
</tr>
</tbody>
</table>

* As defined in Materials and Methods.

* Boldface numbers indicate the highest identity percentage between newly identified subtypes and a known subtype.

* DraE protein sequence used in comparisons.

* NfaA protein sequence (4) used in comparisons.

* —, Too different to obtain an identity comparison.
been considered part of the Dr family, they are very similar in genetic organization (1, 16). Since these NFA series can be detected by the same Dr probe and have upstream flanking sequences similar to those of the Dr family (Fig. 3), they may well belong to the same group. However, it remains to be shown that they bind to decay-accelerating factor.

**Genetic comparison of various adhesin carrying operons.**

Genetic conservation has been demonstrated with a common 4.1-kb internal fragment shared by all afa-1-related members (25). Here we extend the comparison to all E gene classes, including our four newly identified ones. We compared restriction fragment length polymorphism (RFLP) patterns among all adhesin classes by Southern hybridization using the accessory gene probe (probe B in Fig. 1), with a 4.1-kb DNA fragment composed of three PstI fragments [2.6, 1.1, and 0.4 kb] from the afa-1 operon and the E gene probe (three PCR-amplified DNA fragments in Fig. 2, lanes 2, 4, and 6). All Drb subtypes including the NFA series share a very conserved 4.1-kb (three PstI fragments [2.6, 1.1, and 0.4 kb]) segment (Fig. 4), while E genes and their 3’ flanking regions are highly variable (Fig. 5).

It has been suggested that afaD3 is the second adhesin gene in the afa-3 operon (15). To determine if this sequence was also conserved among the D genes in other subtypes, we used PCR analysis with a pair of primers internal to the afaD3 gene.

The D gene sequence appears to be conserved in most Drb subtypes strains tested (Fig. 6). NFA111 shows no PCR fragment, and Drb122 shows a PCR fragment with smaller size.

**Virulence profiles of strains producing Dr family adhesins.**

All of our E. coli isolates were screened by DNA hybridization for the presence of nine virulence factors: type I (fim), Pap and Prs pili (prf), S fimbria (sfa), aerobactin (aer), group II capsule (kpsMT), outer membrane protease T (ompT), α-hemolysin (hly), cytotoxic necrotizing factor 1 (cnf1), and adhesins of the Dr family (drb) as previously described (12). This genotyping generated a virulence signature for each strain based on the presence or absence of these nine virulence factors. Most of the Drb-positive strains were also positive for aer (83%). Only one strain in our collection that was Drb-positive was positive for the toxin hly or cnf1. The Dr adhesions occurred only rarely with other MRHA adhesins such as S fimbria and P pili: 99 and 80% of Drb-positive strains were negative for sfa and prf, respectively. Of greatest interest, half (51.3%) of the Drb-positive strains belonged to a single virulence signature: 100111001 respectively. Of greatest interest, half (51.3%) of the Drb-positive strains belonged to a single virulence signature: 100111001 (Table 4), positive for fim, aer, kpsMT, ompT, and Drb and lacking the other genes tested. This virulence signature was present in 7.6% of first UTI isolates, 8.2% of second UTI isolates, and 2.8% of fecal isolates. We carried out RFLP analysis using variation in the fim locus on the 25 isolates with the most common virulence signature and the afaE5 gene. Figure 7 shows the hybridization patterns seen for seven strains with this predominant pattern, plus two strains with a second pattern and a third with yet a different pattern. In total, four...
were discovered when restriction enzyme analysis of the PCR isolates from healthy college women without UTI (15% versus wise healthy college women to the frequency among 395 fecal isolates, suggesting a possible specific role in adherence to the bladder.

Chromosomal and plasmid location of \textit{drb} genes. The presence of the Dr family of adhesins is highly correlated with the presence of aerobactin (encoded by \textit{aer}). Since \textit{aer} can also be plasmid borne, the correlation might be caused by physical linkage. To address this question, the locations of both \textit{drb} and \textit{aer} determinants among \textit{drb}-positive strains were analyzed by Southern hybridization using DNA probes (Fig. 8). Only 12% of the \textit{drb} genes were found on large plasmids; all of the remaining \textit{drb} genes were located on the chromosome. By subtype, 30% of \textit{draE}, 16% of \textit{afaE3}, 11% of \textit{afaE3}, and 0% of all other subtypes were found on plasmids. Of the 65 \textit{drb}-positive strains that were also positive for \textit{aer}, at most five (6%) had both genes on the same plasmid. Thus, the correlation between \textit{drb} and \textit{aer} genes can be only partially explained by colocalization on a plasmid.

**DISCUSSION**

Previous studies have failed to agree on the frequencies of \textit{drb} in UTI and fecal \textit{E. coli} isolates (2, 3, 8, 25, 33, 35). Several studies have small sample size and/or variable or poorly defined study groups probably contributing to the discrepancies. Frequencies may also differ by age, geographic location, or disease severity. Data pooled across several studies suggests that Dr family adhesins occur twice as frequently among UTI compared to fecal isolates (10). We observed a threefold difference in sequences homologous to \textit{drb} when we compared the frequency among 315 first-time UTI isolates from otherwise healthy college women to the frequency among 395 fecal isolates from healthy college women without UTI (15% versus 5%). This finding strongly supports an association between \textit{drb} and first-time UTI in otherwise healthy young women. AfaE-V was the most common subtype in each \textit{E. coli} collection but occurred proportionally more frequently in first-time UTI than fecal isolates, suggesting a possible specific role in adherence to the bladder.

In subtyping our collection of \textit{E. coli} strains, we discovered genes with a sequence intermediate between the sequence found in subtypes \textit{afaE3} and \textit{draE}. These intermediate strains were discovered when restriction enzyme analysis of the PCR products revealed that only 10 of 19 strains amplified with the subtype AfaE-III/Dra-type primer (Table 1) gave PCR products which corresponded to either \textit{afaE3} or \textit{draE} sequences. Eight products appear to be a hybrid containing the N-terminal half of \textit{afaE3} and the C-terminal half of \textit{draE}. One product is the opposite hybrid: it contains the N-terminal half of \textit{draE} and the C-terminal half of \textit{afaE3}. N-terminal regions are critical for the differences seen between AfaE-III and Dr relative to fimbrial behavior. Since the phenotypic differences are the ones that we expect to play a role in any pathogenic differences, we did our distribution analysis with the \textit{afaE3-draE} hybrids listed as AfaE-III-like and the \textit{draE-afaE3} hybrid listed as Dr-like in Tables 2 and 4.

The \textit{drb} sequences in our \textit{E. coli} collection included the whole spectrum of Dr subtypes previously identified as well as our four completely new ones, for a total of at least 10 E gene subtypes (depending on how \textit{draE} genes with intermediate sequence between \textit{afaE3} and \textit{draE} are categorized). Despite significant sequence differences among E genes, all previously identified Dr subtypes and as well as our newly identified Dr subtypes (including NFA-like sequences) share highly conserved genes within the same gene cluster. They have the same

![](image)

**FIG. 7.** Southern hybridization of total DNA from 10 of the 25 strains with the 100111001 virulence signature, using \textit{fim} as the probe. DNAs were digested with \textit{HindIII} and \textit{EcoRV}. Lane 1, \textit{HindIII} DNA marker; lane 2, 144-11; lane 3, 155-11; lane 4, 164-11; lane 5, 275-11; lane 6, 339-11; lane 7, 356-11; lane 8, F14-61; lane 9, 1188-11; lane 10, F1-62; lane 11, F200-61.
DNA PstI fragment of about 4.1 kb which has been demonstrated among afa operons (25). The ~150-bp segments between PstI (p1 to p3 in Fig. 1) and the beginning of the E gene are almost identical among the classical Dr subtypes. Recent studies suggest that afaD3 of the afa-3 gene cluster may act as a second adhesin molecule (15). We found the D gene to be highly conserved.

Over half of the drb-positive strains in our collection had the same virulence signature: 100111001, which is positive for fim, aer, kpsMT, ompT, and drb and negative for pf, sfa, hly, and cnf1 (Table 4). This virulence signature group accounted for 78% of the most common Dr subtype, AfaE-V, 80% of the Dr-like strains, and 75% of the F131 strains. Whether virulence signature 100111001 corresponds to serotype O75, which has been associated with the Dr family of adhesins (33, 39), is unknown.

There are at least three possible explanations for the observed association of the Dr adhesin with a set of virulence genes: (i) physical co-localization, (ii) clonality, and (iii) codependence in particular pathogenic disease pathways. Pathogenicity islands containing multiple virulence factors such as P pili, hemolysin, and CNF1 found in uropathogenic E. coli isolates (7, 37) are a prime example of physical colocalization. fim, kpsMT, ompT, and, when on the chromosome, aer and drb map to different chromosomal locations. Further, only 9% of drb genes in our strains are on plasmids, and only half of those plasmids also contain aer. Therefore colocalization of the genes does not seem to be a plausible explanation of the association of drb with this particular virulence signature.

If a single clone were responsible for a significant proportion of UTIs caused by Dr-adherent strains, then the copresence of the other virulence genes might be a chance occurrence. However, the majority of isolates from three different drb subtypes have the same virulence signature. RFLP analysis using variation in the fim locus showed that 20 of 25 isolates with the most common virulence signature and the afaE5 gene fall into a single RFLP group. Therefore, half of the 40 strains with the predominant virulence signature are part of a single predominant clone, but the other half fit into at least eight other clonal types with no more than four members each, based on subtype and RFLP analysis.

Finally, codependence of these virulence factors in particular pathogenic disease pathways may explain the association of drb with a particular virulence signature. Our data support the possibility that strains with the 100111001 virulence signature use a particular pathogenic pathway to cause UTI. This pathway may differ from that used by UTI E. coli expressing P pili, hemolysin, and CNF1. Different pathogenic pathways may involve differences in mode of transmission. Epidemiologic studies exploring the latter possibility are ongoing.

Dr is associated with risk of a second UTI (13); thus, a better understanding of the role of Dr subtypes in pathogenesis is warranted. In a large population-based sample of UTI isolates, we found a particular Dr-associated virulence profile to predominate, suggesting an association of this profile with UTI pathogenesis. In between 20 and 30% of all women with a first UTI, the infection recurs within 6 months (11); recurrences account for the vast majority of UTIs treated annually. It will therefore be important to determine if the predominant Dr-associated virulence profile found in this study is also present in significant numbers of strains from women with recurrences.

ACKNOWLEDGMENTS

We thank Steve L. Moseley (University of Washington) for kindly providing unpublished F131 sequences; Charlotte Williams and other staff at the University of Michigan Student Health Service and Scott Spear, Shirley Arldt, Nell Curtis, Barbara Locke, and other staff at the University of Texas at Austin Student Health Service for bacterial samples; and Bonnie Andree, M. Joan Marty, Angela Spindler, and Divya Agrawal for assistance in data collection, data management, and storage and processing of study isolates.

This work was supported by Public Health Service grants DK-35568 and DK-47519.

REFERENCES


FIG. 8. (a) Agarose gel electrophoresis of purified plasmid DNAs. Lane 1, AmpIII DNA marker; lane 2, 245-11 (AfaE-III); lane 3, 260-11 (AfaE-III); lane 4, 1175-11 (AfaE-III); lane 5, 1157-11 (AfaE-III); lane 6, 1072-11 (AfaE-V); lane 7, F1-62 (AfaE-V); lane 8, F34-62 (AfaE-V); lane 9, 365-11. (b) Southern hybridization of plasmid DNAs by using drb probe A (Fig. 1). (c) Southern hybridization of plasmid DNAs by using an aer probe.
27. Le Bouguenec, C. Unpublished results.
30. Moseley, S. L. Personal communication.