S-Fimbria-Encoding Determinant *sfa*_I Is Located on Pathogenicity Island III₅₃₆ of Uropathogenic *Escherichia coli* Strain 536

ULRICH DOBRINDT,¹ GABRIELE BLUM-OEHLER,¹ THOMAS HARTSCH,² GERHARD GOTTSCHALK,²
ELIORA Z. RON,³ REINHARD FÜNFSTÜCK,⁴ AND JÖRG HACKER^{1*}

*Institut für Molekulare Infektionsbiologie der Universität Würzburg, 97070 Würzburg,¹ Institut für Mikrobiologie und Genetik,
Labor für Genomanalysen der Universität Göttingen, 37077 Göttingen,² and Klinik für Innere Medizin IV,
Klinikum der Universität Jena, 07740 Jena,⁴ Germany, and Department of Molecular Microbiology
and Biotechnology, Tel-Aviv University, Ramat Tel-Aviv, 69978 Tel Aviv, Israel³*

Received 18 September 2000/Returned for modification 11 January 2001/Accepted 10 April 2001

The *sfa*_I determinant encoding the S-fimbrial adhesin of uropathogenic *Escherichia coli* strains was found to be located on a pathogenicity island of uropathogenic *E. coli* strain 536. This pathogenicity island, designated PAI III₅₃₆, is located at 5.6 min of the *E. coli* chromosome and covers a region of at least 37 kb between the tRNA locus *thrW* and *yagU*. As far as it has been determined, PAI III₅₃₆ also contains genes which code for components of a putative enterochelin siderophore system of *E. coli* and *Salmonella* spp. as well as for colicin V immunity. Several intact or nonfunctional mobility genes of bacteriophages and insertion sequence elements such as transposases and integrases are present on PAI III₅₃₆. The presence of known PAI III₅₃₆ sequences has been investigated in several wild-type *E. coli* isolates. The results demonstrate that the determinants of the members of the S-family of fimbrial adhesins may be located on a common pathogenicity island which, in *E. coli* strain 536, replaces a 40-kb DNA region which represents an *E. coli* K-12-specific genomic island.

Extraintestinal *Escherichia coli* strains can be grouped into three different pathotypes: meningitis *E. coli* (MENEC) strains, which cause newborn meningitis; septicemia *E. coli* (SEPEC) strains, which cause septicemia infections; and uropathogenic *E. coli* (UPEC) strains, which are the most frequently isolated causative agents of infections of the bladder and the kidney in humans (45, 57). They are characterized by the expression of certain virulence factors, including adhesins, which contribute to the establishment of the infection and distinguish them from nonpathogenic *E. coli* strains (39). Members of the S-fimbrial family of adhesins are frequently expressed in extraintestinal *E. coli* strains isolated from men. This adhesion family consists of S-fimbriae (Sfa), with its subtypes SfaI and SfaII; F1C-fimbriae (Foc); and S/F1C-related fimbriae (Sfr) (25, 46). The AC/I-fimbriae (Fac) which are expressed by avian-pathogenic *E. coli* strains also belong to the S-family of adhesins (5, 6). All members are highly similar in the organization of their determinants and the sequence identities of the encoded proteins (45–47, 54). However, they differ in their receptor and therefore in their tissue specificity. S-fimbrial adhesins recognize α -sialyl-2-3- β -lactose-containing receptors and are predominantly expressed by strains which cause sepsis and meningitis but also by urinary tract infection (UTI) isolates (32, 33), whereas F1C-fimbrial adhesins bind to β -GalNac-1,4- β -Gal-containing structures (30) and are preferentially expressed by UTI isolates.

The uropathogenic *E. coli* strain 536 (O6:K15:H31) has previously been shown to produce various types of fimbrial adhesins, including type 1, P-related, and S-fimbriae (26). The presence of four pathogenicity islands (PAIs I₅₃₆ to IV₅₃₆) has

been described for *E. coli* strain 536 so far. All four PAIs have common characteristics which are typical of pathogenicity islands, i.e., association with a tRNA gene and the presence of often nonfunctional mobility genes (24). Earlier findings implied that the *sfa*_I determinant is part of a pathogenicity island of strain 536 designated PAI III₅₃₆ (23). Whereas the genetic organization of PAI I₅₃₆, PAI II₅₃₆, and PAI IV₅₃₆ has already been published, little information is currently available about the structure of PAI III₅₃₆. Therefore, one aim of this study was the structural characterization of pathogenicity island III₅₃₆. In order to improve our knowledge of the evolution and distribution of the members of the S-family of adhesins as a part of pathogenicity islands in *E. coli*, the chromosomal localization as well as the sequence context of operons coding for members of this adhesin family were investigated in different *E. coli* strains and compared with that of the PAI III₅₃₆-located *sfa*_I determinant of UPEC strain 536.

MATERIALS AND METHODS

Bacterial strains and culture conditions. A collection of 84 wild-type *E. coli* isolates and 28 strains of the *E. coli* Collection of Reference (ECOR) strain collection were used in this study. *E. coli* strains 536, 2980, 764, J96, E-B35, E351, E642, AD110, BK658, W1825, 4405/1, 7521/94-1, F18, 86-24, E2348/69, EDL933, and the ECOR strains have been described elsewhere (11, 13, 16, 17, 19, 27, 34, 36, 42, 43, 48, 49). Eighteen UPEC strains have been chosen from a collection of uropathogenic isolates described before (64, 65). The MENEC strains IHE3034, IHE3036, IHE3080, RS176, RS218, and RS226 were described earlier (1, 32, 41). The SEPEC and MENEC isolates B616, B10363, B13155, Ve239, and Ve1140 were kindly provided by H. Schrotten (Universitätsklinikum Düsseldorf, Düsseldorf, Germany). In addition, we used SEPEC isolates 269/93, 1939/93, 2656/93, 4549/93, 10413/93, and 10209. The enteroaggregative *E. coli* (EAEC) strain DPA065, the UPEC strain CFT073, and the necrotic *E. coli* (NTEC) strain S5 were provided by A. Giammanco (Dipartimento di Igiene e Microbiologica, University of Palermo, Palermo, Italy), by H. Mobley (University of Maryland School of Medicine, Baltimore, Md.), and by E. Oswald (Ecole Nationale Vétérinaire de Toulouse, Toulouse, France), respectively. The enteropathogenic *E. coli* (EPEC) strains 179/2, 156A, and 37-4, the enterohemorrhagic *E. coli* (EHEC) strains SF493/89, 3574/92, 2907/97, 5720/96, 3697/97, and ED142, the

* Corresponding author. Mailing address: Institut für Molekulare Infektionsbiologie, Röntgenring 11, D-97070 Würzburg, Germany. Phone: 49 (0)931 312575. Fax: 49 (0)931 312578. E-mail: j.hacker@mail.uni-wuerzburg.de.

enteroinvasive *E. coli* (EIEC) strain 76-5, and the EAEC 5477/94, as well as the enterotoxigenic (ETEC) strain C9221a (O6:K15:H16) and the reference strain (O149:K88), have been described before (40). The following strains were isolated during a long-term study of women with chronic UTIs: 13A1, 19A1, 16A2U, 20A1, 20A1U, 22A2, 22B2U, 3D5, 2E1U, 2E2U, 1G1, 1G1U, 1H1, and 1H1U. The *E. coli* K-12 strains MG1655, K-12, W3110, P678-54, 5K, C600, and XL-1 Blue have been described before (7, 15). All strains were grown in Luria-Bertani medium (53).

DNA techniques. Isolation of plasmid DNA and recombinant DNA was performed as previously published (53).

PCR. A description of the primers used in this study is available as supplementary material (<http://www.uni-wuerzburg.de/infektionsbiologie>). The differentiation between subtypes of the S-fimbrial family of adhesins was performed by PCR using primers which are specific for the individual major fimbrial subunits. The *sfa*₁-specific PCR product (161 bp) was obtained with the primers *sfa*AI.1 (5'-CGGGCATGCATCAATTATCTTTG-3') and *sfa*AI.2 (5'-TGTGTAGATG CAGTCACTCCG-3') using chromosomal DNA from *E. coli* 536 as a positive control. Detection of the *sfa*_{II} determinant using primer pair *sfa*AI.1 (5'-ACG AAAAGTTAGCTAATCTTGAT-3') and *sfa*AI.2 (5'-TATCTGCGCTTTG ATCCGAATT-3') and chromosomal DNA of *E. coli* strain IHE3034 as a positive control resulted in a 251-bp PCR product. PCR with the primer pair *foc*A.1 (5'-ATGGAGGAACCCAAACGCCA-3') and *foc*A.2 (5'-GCTCACTGTAA CCACTTTTGTG-3') and chromosomal DNA of *E. coli* strain 2980 as a reference strain should result in a 274-bp PCR product. The *sfr* determinant can be detected using primer pair *sfr*A.1 (5'-CTAAAAGCCGACGGGGATAAAA GTGCTGCT-3') and *sfr*A.2 (5'-ATAGCTAGCTGTTGGAGTATTTCCGTCG AAC-3') and chromosomal DNA of strain BK658 as a control. The corresponding PCR product should be 241 bp long. DNA primers were purchased from MWG Biotech (Ebersberg, Germany). The *Taq* DNA polymerase used for the detection of genes in different *E. coli* strains was purchased from Qiagen (Hilden, Germany).

DNA sequencing and sequence analysis. The sequence determination of the cosmid clones pANN1E6 and parts of pCos3grl3, which contain parts of PAI III₅₃₆ of *E. coli* strain 536, was performed as follows. One small insert library (2 to 2.5 kb) was generated by mechanical shearing of the cosmid DNA (44). After end repair with T4 polymerase, the fragments were ligated into the prepared pTZ19R vector. Isolated plasmids were sequenced from both ends using dye terminator chemistry and analyzed on ABI377 machines (Applied Biosystems, Munich, Germany). In the case of pANN1E6, after the assembly of 495 sequences, the remaining gaps were closed by primer walking on the plasmid clones. The Phrap software implemented in the Staden software package was used for assembly and editing of the sequence data (56). The sequence of the PAI III₅₃₆ fragment present on the cosmid 1E6 was submitted to the NCBI database.

Homology searches were performed with the BLASTN and BLASTX programs of the National Center for Biotechnology Information (3).

Southern hybridization. PCR results during the analysis of the chromosomal region downstream of *thrW* in different *E. coli* isolates were confirmed by Southern hybridization of *Eco*RI-digested chromosomal DNA of the investigated *E. coli* strains. After agarose gel electrophoresis, the *Eco*RI-digested *E. coli* genomic DNA was transferred to Biodyne B nylon membranes (PALL, Rosendorf, Germany). The PCR products of PCRs 19 to 46 (see also Fig. 1B) performed with chromosomal DNA of *E. coli* strain MG1655 as the template were used as probes for hybridization. Hybridization and detection were carried out using the ECL labeling and signal detection system (Amersham/Pharmacia Biotech, Freiburg, Germany) according to the manufacturer's recommendations.

RESULTS

Sequence analysis of left part of PAI III₅₃₆. The cosmid clone pANN1E6 was shown to contain the entire S-fimbrial determinant (*sfa*₁) together with the flanking regions of UPEC strain 536. The sequence determination of the 38-kb DNA insert revealed that about 33 kb of this DNA fragment represented a part of a pathogenicity island which was designated PAI III₅₃₆ previously (24). The right end of the cosmid insert of pANN1E6 and the left end of that of pCos3grl3 contain an overlapping region of about 4 kb covering a part of the *iro* determinant. The assembly of the sequences inserted into pANN1E6 and of the available sequences of pCos3grl3 re-

sulted in a 37-kb DNA sequence which covers the left-hand part of PAI III₅₃₆ (accession no. AF302690). It is estimated that the complete PAI III₅₃₆ is about 70 kb in size. As can be seen from the sequence analysis, PAI III₅₃₆ exhibits several features which are typical of pathogenicity islands. In addition to the *sfa*₁ determinant encoding the S-fimbrial adhesin, this fragment comprises the tRNA gene *thrW*, which serves as the chromosomal integration site of PAI III₅₃₆, several genes of mobile genetic elements such as bacteriophage genes and insertion sequence (IS) elements, as well as unknown open reading frames (Fig. 1A, Table 1). One typical characteristic of pathogenicity islands is the presence of a bacteriophage integrase gene downstream of the tRNA gene which serves as the chromosomal integration site of the PAI.

In PAI III₅₃₆, the tRNA gene *thrW* is followed by an integrase and excisionase gene of the O-antigen-modifying bacteriophage X of *Shigella flexneri*. PAI III₅₃₆ contains a 5-kb DNA region without sequence homology to known DNA sequences which is flanked by an IS100 element frequently found on *Yersinia pestis* plasmids and a nonfunctional IS3 element found on the SHI-2 pathogenicity island of *S. flexneri*. Five suppression subtractive hybridization (SSH) fragments (TspE4.K8, SauE15.C6, TspE4.B12, SauE15.G6, and SauE15.D4) obtained from a newborn meningitis (NBM)-causing *E. coli* isolate described by Bingen and coworkers (14) are homologous to this IS100-specific region. Two putative open reading frames, designated ORF 1 and ORF 2, can be determined in this region, which may represent so far unknown genes encoding a protein with homology to a hypothetical protein of *Thermotoga maritima* or a putative *pifA* gene product of *E. coli* and to a protein with homology to the 3-deoxy-D-arabinoheptulosonate-7-phosphate synthase (DAHPSynthase) of *Erwinia herbicola*, respectively. The fact that ORF 2 is preceded by a 150-bp DNA region that codes for a fragment similar to the transposase of the *Erwinia amylovora*-associated transposon Tn5393 implies that ORF 2 is a part of a former mobile genetic element. The IS3 element is followed by a 1.6-kb region which contains a fragment of the *mhcE* gene of *E. coli* strain H47, which encodes a component of the microcin H47 ABC transporter (4). About 1 kb of this DNA sequence is also homologous to an open reading frame which represents a truncated copy of the *cvaB* gene of plasmid pColV-K30, encoding a colicin V secretion protein. Another DNA region without any nucleotide sequence homology covers roughly 1.4 kb downstream of the *cvaB* fragment. This region seems to be linked with the *cvaB* gene, as it contains two small putative ORFs which code for peptides with homology to a microcin V immunity protein and the colicin V precursor, respectively. The gene product of another putative open reading frame (ORF 5) without any homology to known DNA sequences in this region exhibits similarity to a hypothetical protein of *Serratia marcescens*. This ORF is followed by a 2-kb DNA stretch containing ORFs 6 and 7. This part of PAI III₅₃₆ is homologous to sequences found in the enterohemorrhagic *E. coli* O157:H7 strain EDL933 in which the putative open reading frames Z1633 and Z1634 are located (50). Due to an internal stop codon, the putative gene product of Z1634 is 21 amino acids shorter in *E. coli* 536 than in strain EDL933. The 5' part of the next ORF (ORF 8) is identical to that of *E. coli* O157:H7 ORF Z1635, but the promoter-distal region shows no homology to

FIG. 1. Comparison of genetic organization of chromosomal downstream region of the tRNA gene *thrW* in *E. coli* strains 536 (A) and MG1655 (B); identification of different types of genomic islands. The structure of PAI III₅₃₆ of *E. coli* strain 536 and that of the genomic island (GEI) of *E. coli* K-12 strain MG1655 is shown in its chromosomal sequence context. Numbered lines indicate regions of PAI III₅₃₆ and of the GEI which have been detected by PCR. Solid arrows indicate ORFs representing the core chromosome, and open arrows indicate ORFs of mobile genetic elements. ORFs within regions without homology on the nucleotide level are indicated by hatched arrows.

known nucleotide sequences. The SSH fragment TspE4.H10 isolated from an NBM *E. coli* isolate (14) represents a DNA stretch homologous to a part of *sfaC*.

The *sfa*₁ determinant coding for the S-fimbrial adhesin is followed by a 9.3-kb DNA region which represents the *E. coli* homologue of the *iroBCDEN* operon of *Salmonella* spp. The *iroN* gene encoding a catechol siderophore receptor is followed by four putative ORFs with high homology to the *iroED* of *Salmonella enterica* and *iroBC* of *Salmonella enterica* serovar Typhi. The function of the *iroE* gene product is yet unknown. The gene products of *iroD* and *iroC* are homologues of a ferric enterochelin esterase of *S. enterica* and an ABC transport protein of serovar Typhi, respectively. The *iroB*-encoded protein is highly homologous to glucosyltransferases. The *iroNE* genes and an *iroD* fragment have recently been identified in another uropathogenic *E. coli* strain, CP9 (O4:K54:H5), closely linked to DNA sequences with homology to determinants of members of the S-adhesin family (52). One SSH fragment (TspE4.A8) obtained from an NMB-causing *E. coli* strain (14) is homologous to a region within *iroN*. The *iro* determinant is followed by an ORF which represents the second half of a gene encoding a hypothetical protein of 50.4 kDa in IS4.

Identification of right junction site of PAI III₅₃₆. The right junction site of PAI III₅₃₆ was determined by inverse PCR after ligation of *EcoRV* fragments of chromosomal DNA of *E. coli* strain 536, which contain the upstream region of *yagU* encoding a hypothetical protein. The sequence determination of the obtained DNA fragment (Fig. 2) revealed that 109 bp upstream of *yagU*, the *E. coli* MG1655-specific sequence representing the *E. coli* core chromosome is interrupted by sequences with homology to fragments of integrase genes of the *S. flexneri* bacteriophages SfII (AF021347) and SfV (SFU82619) as well as of the *E. coli* bacteriophage P22 (AF217253). This region is followed by DNA sequences with homology to fragments of the *insB* and *insA* genes of the *Iso-IS1* element present in *Shigella dysenteriae* (AF153317) and the EPEC plasmid pB171 (AB024946). These results demonstrate that PAI III₅₃₆ is flanked by sequences derived from *Shigella* bacteriophages, which integrate into highly homologous *attP* sites next to the tRNA gene *thrW*.

Identification of different members of S-family of adhesins by PCR. In order to be able to distinguish between the determinants that encode different subtypes of S-fimbriae (*sfa*₁ and *sfa*_{II}), F1C-fimbriae (*foc*), and S-fimbria-related fimbriae (*sfr*), PCRs were established with primer pairs which are specific for

TABLE 1. ORFs and mobile genetic elements within PAI III₅₃₆

ORF, IS, or bacteriophage	Position ^a	Similar sequences ^b	% Nucleotide (protein) identity
<i>int, xis</i> (SfX)	94–1478	Integrase and excisionase of the O-antigen-converting bacteriophage X (<i>S. flexneri</i>) (BXU82084)	98 (99)
IS100	2134–4011	IS100 (<i>Y. pestis</i> plasmids pCD1 and pMT1) (AF074612)	99
ORF 1	4062–5813	Hypothetical protein (<i>T. maritima</i>) (AE001775)	(23)
		Hypothetical <i>pifA</i> gene product (<i>E. coli</i>) (X04968)	(36)
ORF 2	7334–8368	DAHPSynthase (<i>E. herbicola</i>) (U93355)	(48)
IS3	9245–10458	IS3 (SHI-2, <i>S. flexneri</i>) (AF141323)	99
<i>ΔcvaB</i>	10491–11546	<i>E. coli mchE</i> gene for the microcin H47 exporter component protein (AJ278866)	97
		Colicin V secretion ATP-binding protein (pColV) (X57524)	(87–95)
ORF 3	11653–11874	Microcin V immunity protein (P22521)	(23)
ORF 4	11871–12149	Colicin V precursor (P22522)	(40)
ORF 5	12326–13012	Hypothetical protein (<i>S. marcescens</i>) (AF027768)	(39)
		YprB (<i>E. coli</i>) (P33774)	(40)
ORF 6	13107–13577	ORF Z1633 of <i>E. coli</i> O157:H7 (AE005311)	97
ORF 7	14023–14535	ORF Z1634 of <i>E. coli</i> O157:H7 (AE005311)	86
ORF 8	14987–15346	Nucleotides 1–160 of ORF Z1635 of <i>E. coli</i> O157:H7 (AE005311)	83
<i>sfaC</i>	15621–15842	<i>E. coli</i> 536 <i>sfaC</i> gene of the S-fimbrial adhesin determinant <i>sfa</i> (AF301153)	100 (56)
<i>sfaB</i>	16261–16590	<i>E. coli</i> 536 <i>sfaB</i> gene of the S-fimbrial adhesin determinant <i>sfa</i> (AF301153)	100
<i>sfaA</i>	16964–17509	<i>E. coli</i> 536 <i>sfaA</i> gene of the S-fimbrial adhesin determinant <i>sfa</i> (AF301153)	100
		<i>E. coli</i> 536 SfaA fimbrial subunit precursor (P12730)	(99)
<i>sfaD</i>	17580–18119	<i>E. coli</i> 536 <i>sfaD</i> gene of the S-fimbrial adhesin determinant <i>sfa</i> (AF301153)	100
<i>sfaE</i>	18160–18855	<i>E. coli</i> 536 <i>sfaE</i> gene of the S-fimbrial adhesin determinant <i>sfa</i> (AF301153),	100
		<i>E. coli</i> 536 sfa E fimbrial subunit (1713397D)	(99)
<i>sfaF</i>	18926–21530	<i>E. coli</i> 536 <i>sfaF</i> gene of the S-fimbrial adhesin determinant <i>sfa</i> (AF301153),	100
		<i>E. coli</i> 536 Sfa F fimbrial subunit (1713397E)	(99)
<i>sfaG</i>	21569–22096	<i>E. coli</i> 536 <i>sfaG</i> gene of the S-fimbrial adhesin determinant <i>sfa</i> (AF301153),	100
		<i>E. coli</i> 536 sfa G fimbrial subunit precursor (P13429)	(100)
<i>sfaS</i>	22118–22609	<i>E. coli</i> 536 <i>sfaS</i> gene of the S-fimbrial adhesin determinant <i>sfa</i> (AF301153)	100
<i>sfaH</i>	22671–23570	<i>E. coli</i> 536 <i>sfaH</i> gene of the S-fimbrial adhesin determinant <i>sfa</i> (AF301153)	100
		<i>E. coli</i> 536 SfaH fimbrial subunit precursor (P13431)	(99)
<i>sfaX</i> (17 kDa)	24811–25310	<i>E. coli</i> <i>prsX</i> gene (X62158)	96
		<i>E. coli</i> PrsX protein (P42192)	(98)
<i>iroN</i>	25804–27981	<i>E. coli</i> siderophore receptor IroN (AF135597)	99 (99)
<i>iroE</i>	28026–28982	<i>S. enterica</i> IroE (U97227)	73–76
<i>iroD</i>	29067–30296	<i>S. enterica</i> ferric enterochelin esterase homolog IroD (U97227)	72 (64)
<i>iroC</i>	30400–34185	Serovar Typhi ABC transport protein homolog IroC (T30293)	73–79 (68)
<i>iroB</i>	34199–35314	Serovar Typhi glucosyltransferase (T30292)	78 (80)
<i>Δyi4I</i> (IS4)	36162–36521	Plasmid pMM237, <i>tnpA</i> /IS4 right junction (MM7TN3IS4A)	99
		Insertion element IS4 hypothetical 50.4-kDa protein, amino acids 316–442 (P03835)	(99)

^a The location is given as nucleotide positions, numbering from the first nucleotide downstream of the tRNA gene *thrW*.

^b Homology based on BLASTN and BLASTX database searches.

*sfaA*_I, *sfaA*_{II}, *focA*, and *sfrA*. These primers can be used for specific amplification of the major subunit-encoding gene of the abovementioned members of the S-family of adhesins. *E. coli* strains 536, IHE3034, 2980, and BK658 were used as positive controls during the *sfaA*_I-, *sfaA*_{II}-, *focA*-, and *sfrA*-specific PCRs, respectively. According to the results of the PCRs, the strains used in this study were grouped as indicated in Table 2. Of the 66 *E. coli* strains expressing members of the S-family of adhesins (including strain 536), 13 strains were shown to be *sfaA*_I positive. In 12 *E. coli* strains, the *sfaA*_{II} determinant was detectable. *E. coli* strain AC/1 harbors the *foc* determinant, and 34 strains in this study contain the *foc* operon. With primers

specific for conserved regions of the *sfaA*_I determinant, the presence of corresponding sequences was detectable in six *E. coli* strains. However, further identification using primer pairs specific for *sfaA*_I, *sfaA*_{II}, or *focA* was not possible. It has been reported that strain BK658 expresses an S-fimbria-related adhesin (Sfr) which is related but not identical to S- and F1C-fimbriae (49). In strains BK658 and 19A1, the *sfrA*-specific PCR was positive. The fimbrial determinants of the other four strains which could not be identified as *sfaA*_I, *sfaA*_{II}, *foc*, or *sfr* have been designated *sfx*, as they also seem to encode an unknown subtype of S-fimbrial adhesin. The corresponding strains do not agglutinate bovine erythrocytes but contain se-

FIG. 2. Junction sites between the *E. coli* core chromosome and PAI III₅₃₆ in *E. coli* strain 536. Solid arrows represent ORFs located on the core chromosome. PAI III₅₃₆ is flanked by the tRNA gene *thrW* and by *yagU*, representing the adjacent genes of the core chromosome. Both ends of PAI III₅₃₆ show homology to sequences of *Shigella* O-antigen-modifying bacteriophages which integrate into the tRNA gene *thrW*. Solid arrows indicate ORFs representing the core chromosome, and open arrows indicate ORFs of mobile genetic elements.

quences homologous to *sfa* and *foc* which are detectable by Southern hybridization.

Analysis of presence of PAI III₅₃₆-like genetic elements in *E. coli* strains expressing members of S-family of fimbrial adhesins. Based on the DNA sequence representing the left-hand part of PAI III₅₃₆, we selected DNA-primer combinations in order to detect the known PAI III₅₃₆ fragment in 112 different *E. coli* strains, including many extraintestinal isolates which are known to frequently express fimbrial adhesins belonging to the S-family of *E. coli* adhesins. In 35 extraintestinal and fecal strains as well as in several isolates of different intestinal *E. coli* pathotypes, no PAI III₅₃₆-specific sequences were detectable. Among them, the presence of the *iro* determinant without PAI III₅₃₆-specific sequence context was demonstrated in 11 extraintestinal and fecal strains. As shown in Table 2, the entire *E. coli* 536-specific PAI III₅₃₆ fragment was not detectable by PCR in 65 *E. coli* strains containing genes which encode S-family fimbrial adhesins. Interestingly, in eight *E. coli* strains harboring the *sfa*_I determinant, almost the entire left-hand part of PAI III₅₃₆ (covered by PCRs 3 to 18) was detectable. With primers generated from conserved regions of the *E. coli* 536-specific *sfa*_I gene cluster (region 4), the different S-fimbria determinants, including *foc* of the avian-pathogenic strain AC/I, were completely detectable in all S-fimbria-positive *E. coli* strains used in this study. In all cases, the different determinants coding for members of the S-adhesin family are

associated with the 3'-flanking region (region 5, *iro* gene cluster) of the *E. coli* 536-specific *sfa*_I determinant (Fig. 1 and Table 2). Some differences can be detected with respect to the 3' region of the *iro* determinant and the following sequence context (PCR 18). It is very likely that these differences are mainly due to the presence of additional IS4-related sequences located downstream of *iro* which are absent in *E. coli* strain 536. With the exception of the *sfa*_{II}-positive *E. coli* strains, the 5'-flanking region 3 (*E. coli* O157:H7-related sequences) of the *sfa*_I₅₃₆ determinant was also detectable upstream of the corresponding determinants of all strains expressing an adhesin of the S-adhesin family (Fig. 1 and Table 2). Sequence determination of the upstream region of the *E. coli* J96 *foc* determinant confirmed these findings, as the obtained sequences were identical with 750 bp of the upstream region of the PAI III₅₃₆-located *sfa*_I operon (data not shown). Other parts of PAI III₅₃₆ (regions 1 to 3) could not be amplified by PCR in *E. coli* isolates harboring other subtypes of S-fimbrial adhesin determinants (Fig. 1, Table 2). However, the IS3 transposase gene embedded in a PAI III₅₃₆-specific sequence context (PCR 5) could be amplified in most of the S-fimbria-positive strains tested. These results demonstrate that the different determinants of the S-fimbrial adhesin family are located within a common DNA region. They are associated with the *iro* gene cluster and, except for the *sfa*_{II} determinant, with sequences found within the genome of *E. coli* strain EDL933. This com-

TABLE 2. Presence of PAI III₅₃₆-specific DNA regions in *E. coli* isolates expressing different types of S-fimbrial adhesin^a

Adhesin type	No. of isolates	Region 1 (mobility) (PCRs 1 and 2)	Region 2 (unknown ORFs) (PCRs 3 and 4)	Region 3 (colicin) (PCR 6)	Region 4 (O157:H7 related) (PCRs 7 and 8)	Region 5 (S-adhesin) (PCRs 9–13)	Region 6 (<i>iro</i>) (PCRs 14–17)
<i>sfa</i> _I	12	—	+	++	++	++	++
<i>sfa</i> _{II}	12	—	—	—	—	++	++
<i>foc</i>	34	—	—	—	++	++	++
<i>foc</i>	1	—	—	—	++	++	++
<i>sfr</i>	2	—	—	—	++	++	++
<i>sfx</i>	4	—	—	—	++	++	++

^a ++, PAI III₅₃₆-specific region detectable in 100% of *E. coli* strains expressing a certain type of S-adhesin; +, PAI III₅₃₆-specific region detectable in less than 100% of the *E. coli* strains expressing a certain type of S-adhesin; —, PAI III₅₃₆-specific region not detectable in *E. coli* strains expressing a certain type of S-adhesin.

TABLE 3. Comparison of *thrW* downstream region in different *E. coli* strains

<i>E. coli</i> strain (no. tested)	No. (%) of strains with <i>E. coli</i> K-12-specific ORF ^a				
	<i>yafW</i> (PCR 19)	<i>ykfC</i> (PCR 28)	<i>yagF</i> (PCR 32)	<i>yagT</i> (PCR 43)	<i>ykgJ</i> (PCR 45)
K-12 (7)	7 (100)	7 (100)	7 (100)	7 (100)	7 (100)
Fecal (19)	0 (0)	0 (0)	0 (0)	0 (0)	18 (95)
UPEC (19)	0 (0)	0 (0)	0 (0)	0 (0)	19 (100)
NBM, sepsis (10)	0 (0)	0 (0)	0 (0)	0 (0)	9 (90)

^a Values in parentheses indicate the percentage of strains with a positive PCR result within every subgroup analyzed in this study.

mon sequence context may be part of a widespread pathogenicity island or may represent a part of a commonly transferable genetic element. The *iro* determinant was only detectable in extraintestinal and fecal *E. coli* isolates. In all *E. coli* isolates used in this study which represent different *E. coli* intestinal pathotypes, no PAI III₅₃₆-specific sequences were detectable.

Analysis of *thrW* downstream region in different *E. coli* strains: identification of *E. coli* K-12-specific genomic island. The PCR-based analysis of the *thrW* downstream region, covering the chromosomal region between *yafW* and *yagU* of *E. coli* strain MG1655 (Fig. 1B), revealed that this strain-specific DNA sequence is not present in this form in the following eight pathogenic *E. coli* isolates: 536 (UPEC), J96 (UPEC), IHE3034 (MENEC), C9221a (ETEC), E2348/69 (EPEC), EDL1284 (EIEC), DPA065 (EAEC), and EDL933 (EHEC). Although not every ORF or gene in this region has been analyzed by PCR (PCRs 19 to 46; see Fig. 1B), our results imply that this ~40-kb fragment is almost completely missing in the different *E. coli* strains tested. ORFs located in regions A, B, and C as well as that amplified by PCR 40 are absent in all eight strains (Fig. 1B). ORFs *ykfC* (PCR 28), *yagI* (PCR 34), and *yagM* (PCR 39) as well as *yagP*, *yagR*, and *yagT* (PCRs 41 to 43) are present in some of these pathogenic *E. coli* strains. With the exception of the EIEC strain EDL1284, the ORFs *yagU*, *ykgJ*, and *yagV* (PCRs 44 to 46) are present in all other pathogenic isolates (Fig. 1B). The PCR results have been confirmed by Southern blot analysis, underlining that, with a few exceptions, the ORFs investigated are absent in the eight pathogenic *E. coli* isolates used in this study.

In order to find out whether the DNA region between *thrW* and *yagU* is specific for *E. coli* K-12 strains, we expanded our study and screened several *E. coli* K-12 strains, several strains representing different *E. coli* pathotypes, and members of every subgroup of the ECOR strain collection for the presence of five representative ORFs which are located in the DNA region between *thrW* and *yagV*. The results presented in Table 3 clearly demonstrate that in contrast to the analyzed *E. coli* K-12 strains, the ORFs located between *thrW* and *yagU* were not detectable in the different *E. coli* isolates and ECOR strains studied. On the basis of the PCR results presented in Table 3, it can be concluded that the DNA fragment covering a region of ~40 kb downstream of the tRNA gene *thrW* in *E. coli* MG1655 represents in this form a genomic island specific for *E. coli* K-12 strains and is absent in all *E. coli* isolates investigated in this study.

DISCUSSION

Fimbrial adhesins of the S-adhesin family are virulence factors frequently produced by most extraintestinal *E. coli* isolates that cause a variety of clinical symptoms in mammals and poultry. Although the genetic organization of their determinants and the sequences of many of their encoded gene products are highly homologous, their receptor specificities vary, thus changing the host and tissue specificity. Therefore, from an evolutionary point of view, the operons coding for the members of the S-adhesin family can serve as a model system to study combinatorial gene shuffling as an adaptive process in host-pathogen interactions. The frequent occurrence of S-fimbria-encoding operons in different extraintestinal *E. coli* strains implies that they are located on a pathogenicity island which may have been acquired by horizontal gene transfer. The localization of other fimbrial adhesin determinants on pathogenicity islands of extraintestinal *E. coli*, i.e., those coding for members of the P-adhesin family, has been shown before (12, 24, 58).

The sequence analysis of the cosmid clone pANN1E6 proved that the *sfa*_I determinant of *E. coli* strain 536 is located on a pathogenicity island which has been designated PAI III₅₃₆ before (23). Although PAI III₅₃₆ has not been completely characterized, the known sequences exhibit several characteristics which are typical of pathogenicity islands. PAI III₅₃₆ is integrated at the tRNA gene *thrW* into the bacterial chromosome. tRNA genes frequently serve as chromosomal integration sites for bacteriophages. From our point of view, the fact that PAIs may have originated from temperate bacteriophages which have been "immobilized" within the chromosome is also reflected by the presence of mobility genes coding for integrases or transposases of mobile genetic elements on these PAIs (23). In the case of PAI III₅₃₆, the tRNA gene *thrW* is immediately followed by an attachment site for different bacteriophages, including P22 and DLP12, and the complete integrase and excisionase gene of the O-antigen-converting bacteriophage X of *Shigella flexneri*. The *attP* sites of *S. flexneri* bacteriophage X and other serotype-converting *Shigella* phages are identical to the core *attP* sequence of P22 and DLP12. Additionally, the proteins encoded by the integrase genes of the O-antigen-converting *Shigella* phages are very similar to the corresponding proteins of P22 and DLP12 (2). The presence of sequences at the right junction site of PAI III₅₃₆ with homology to integrase gene fragments of the *Shigella* O-antigen-converting bacteriophages SfII and SfV or to that of P22 demonstrates that this pathogenicity island may originally have evolved from integrated *Shigella* bacteriophages. Additionally, three regions with homology to IS elements (IS100, I3S, IS4, and iso-IS1) are located on PAI III₅₃₆, thereby increasing the number of mobility genes on this pathogenicity island. Whereas the sequence of the IS100 element found on PAI III₅₃₆ is identical to those frequently found on *Yersinia pestis* and EPEC plasmids (37, 59) the IS3 element on PAI III₅₃₆ is highly homologous to a nonfunctional IS3 element found on *S. flexneri* pathogenicity island SHI-2 (61). These data underline the frequent presence of horizontally acquired genetic elements within pathogenicity islands.

Another typical feature of PAIs is the presence of unidentified ORFs. On PAI III₅₃₆, there are two regions which show

no homology to known DNA sequences but which contain ORFs that, after translation, encode proteins with similarity to hypothetical proteins of other bacterial species or plasmid-encoded proteins. Whether these ORFs are actually functional will have to be investigated in the future. Determinants coding for siderophore systems are frequently located on pathogenicity islands (55, 61, 63). The *iroBCDEN* operon, which is homologous to the *iro* gene cluster of *Salmonella* species, has been detected downstream of the *sfa*_I determinant on PAI III₅₃₆. In *E. coli*, the siderophore which is recognized by *Iron* is unknown, as is the function of the other genes of this operon. *iroN* and parts of the accompanying *iroE* and *iroD* genes have recently been demonstrated to be induced in extraintestinal *E. coli* isolate CP9 upon incubation in urine. Interestingly, it has been shown that the *iroN* gene of strain CP9 is preceded by a fragment of an IS element and by DNA sequences with homology to the *prs* or *foc* gene cluster (52). These data support our findings that the *iro* genes are associated with determinants coding for members of the S-adhesin family of *E. coli*.

The comparison of the flanking regions of S-fimbrial gene clusters in several extraintestinal *E. coli* isolates shows that indeed a colocalization of the corresponding determinants and common flanking regions exists. Interestingly, differences between the upstream regions of the different fimbrial determinants can be seen. The upstream region of the *sfa*_{II} determinant seems to be different from that of the other determinants coding for members of the S-family of adhesins. Our results demonstrate that all determinants coding for the different members of the S-adhesin family are associated with *iro* genes and, with the exception of the *sfa*_{II} genes, are localized in a common sequence context. This may be indicative of their localization within a common pathogenicity island or of their former acquisition within a block of commonly transferable DNA. Our results contradict the recently published finding that the *iroN*_{*E. coli*} sequences are not strictly associated with operons coding for adhesins of the S-family (28). Furthermore, the *iro*-encoded siderophore system seems to be characteristic of fecal and extraintestinal *E. coli* isolates.

Mobile genetic elements such as plasmids, bacteriophages, and pathogenicity islands belong to a group of genomic additions that contribute in particular to the evolution of bacterial pathogens and to the evolution of bacterial species in general (18). The increasing number of genetic elements with structural similarities to pathogenicity islands in nonpathogenic bacteria led to an extension of the pathogenicity island concept to the so-called concept of genomic islands (21, 22). In recent years, much attention has been paid to the detection and characterization of pathogenicity islands and other horizontally acquired DNA fragments that have been added to the former core chromosome and confer pathogenicity or fitness on certain bacterial strains (29, 35). Few examples of deletions of large DNA fragments from the chromosome have been reported with respect to their impact on virulence properties (20, 31, 38). In most cases, these unstable DNA fragments turned out to be pathogenicity islands. The comparison of the size of the completely sequenced chromosome of *E. coli* strain MG1655 (4,639 kb) (10) with estimates of the genome sizes of natural *E. coli* isolates (ranging from 4.5 to 5.52 kb) shows that substantial differences in genome size exist among various *E. coli* strains (8, 9). The analysis of type-specific genome size

differences among different *E. coli* pathotypes and *E. coli* MG1655 by macrorestriction demonstrated that not only the addition of DNA fragments to but also the deletion of chromosomal regions from the chromosome contribute to different genome sizes in comparison to *E. coli* MG1655 and consequently to different phenotypes or pathotypes (51). We identified a region of the *E. coli* MG1655 chromosome which is specific for *E. coli* K-12 strains (Fig. 1B, Table 3). This DNA fragment covers approximately 40 kb downstream of the *thrW* tRNA gene in strain MG1655 and is not present in this form in several extraintestinal *E. coli* isolates as well as in representatives of all four subgroups of the ECOR strain collection, including the A subgroup, which also includes *E. coli* strain K-12.

This 40-kb *E. coli* K-12-specific DNA region downstream of the tRNA-encoding gene *thrW* (b0245 to b0286) exhibits a GC content of 54%, contains several repeat elements as well as IS elements or at least parts of them (IS911A, IS30, IS5, IS1, and IS30), and consists mainly of genes whose gene products are hypothetical proteins. The presence of many mobile genetic elements in this region could explain the loss of the 40-kb fragment in many *E. coli* strains due to high recombinational activity in this DNA region. The central part of this region contains the *argF* gene, flanked by IS1 elements. This structure, only present in *E. coli* K-12 strains, together with the high *argF* GC content (59%), argues for its acquisition by horizontal gene transfer (60, 62). Interestingly, a 53-kb deletion compared to *E. coli* MG1655 has been identified in uropathogenic *E. coli* strain J96 in the vicinity of the *foc* determinant (51). On the other hand, the fact that the presence of this region is restricted to *E. coli* K-12 strains and the occurrence of several mobility genes are features which are characteristic of genomic islands. Therefore, this *E. coli* K-12-specific DNA fragment could represent a genomic island. Our study presents for the first time in detail the occurrence of an *E. coli* K-12-specific DNA region and improves our knowledge on the genome variability of *E. coli* strains.

ACKNOWLEDGMENTS

The work of the Würzburg group was supported by the Deutsche Forschungsgemeinschaft (Sonderforschungsbereich 479 and by the "Fonds der Chemischen Industrie." The Würzburg group and the Tel Aviv group cooperated in a joint research project (contract no. I-502-171-01/96) supported by the German Israeli Foundation. The Göttingen Genomics Laboratory received support from "Forschungsmittel des Landes Niedersachsen."

We thank H. Karch (Würzburg, Germany) and H. Schrotten (Düsseldorf, Germany) for the gift of several MENEC and SEPEC strains used in this study. We are grateful to A. Giammanco (Palermo, Italy) and E. Oswald (Toulouse, France) for providing the EAEC strain DPA065 and the NTEC strain S5, respectively. We thank A. Sauer (Würzburg) for technical assistance.

REFERENCES

1. Achtman, M., A. Mercer, B. Kusecek, A. Pohl, M. Heuzenroeder, W. Aaronson, A. Sutton, and R. P. Silver. 1983. Six widespread bacterial clones among *Escherichia coli* K1 isolates. *Infect. Immun.* **39**:315-335.
2. Allison, G. E., and N. K. Verma. 2000. Serotype-converting bacteriophages and O-antigen modification in *Shigella flexneri*. *Trends Microbiol.* **8**:17-23.
3. 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.
4. Azpiroz, M. F., E. Rodriguez, and M. Lavina. 2001. Structure, function, and origin of the Microcin H47 ATP-binding cassette exporter are similar to those of colicin V. *Antimicrob. Agents Chemother.* **45**:969-972.

5. Babai, R., G. Blum-Oehler, B. E. Stern, J. Hacker, and E. Z. Ron. 1997. Virulence patterns from septicemic *Escherichia coli* O78 strains. *FEMS Microbiol. Lett.* **149**:99–105.
6. Babai, R., B. E. Stern, J. Hacker, and E. Z. Ron. 2000. A new fimbrial gene cluster of the S-fimbrial adhesin family. *Infect. Immun.* **68**:5901–5907.
7. Bachmann, B. J. 1987. Derivatives and genotypes of some mutant derivatives of *Escherichia coli* K-12, p. 1190–1219. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology, 1st ed. ASM Press, Washington, D.C.
8. Bergthorsson, U., and H. Ochman. 1998. Distribution of chromosomal length variations in natural isolates of *Escherichia coli*. *Mol. Biol. Evol.* **15**:6–16.
9. Bergthorsson, U., and H. Ochman. 1995. Heterogeneity of genome sizes among natural isolates of *Escherichia coli*. *J. Bacteriol.* **177**:5784–5789.
10. Blattner, F. R., G. Plunkett III, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. Collado-Villes, J. D. Glasner, C. K. Rode, G. F. Mayhew, J. Gregor, N. W. Davis, H. A. Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau, and Y. Shao. 1997. The complete genome sequence of *Escherichia coli* K-12. *Science* **277**:1453–1462.
11. Blum, G., M. Ott, A. Cross, and J. Hacker. 1991. Virulence determinants of *Escherichia coli* O6 extraintestinal isolates analysed by Southern hybridizations and DNA long range mapping techniques. *Microb. Pathog.* **10**:127–136.
12. Blum, G., M. Ott, A. Lischewski, A. Ritter, H. Imrich, H. Tschäpe, and J. Hacker. 1994. Excision of large DNA regions termed pathogenicity islands from tRNA-specific loci in the chromosome of an *Escherichia coli* wild-type pathogen. *Infect. Immun.* **62**:606–614.
13. Blum, G., J. Heesemann, D. Kranzfelder, F. Scheutz, and J. Hacker. 1996. Abortion by a woman carrying a pathogenic *Escherichia coli* of serotype O12:K1:H7. *Eur. J. Clin. Microbiol. Infect. Dis.* **16**:153–155.
14. Bonacorsi, S. P., O. Clermont, C. Tinsley, I. Le Gall, J. C. Beaudoin, J. Elion, X. Nassif, and E. Bingen. 2000. Identification of regions of the *Escherichia coli* chromosome specific for neonatal meningitis-associated strains. *Infect. Immun.* **68**:2096–2101.
15. Boyd, E. F., and D. L. Hartl. 1998. Chromosomal regions specific to pathogenic isolates of *Escherichia coli* have a phylogenetically clustered distribution. *J. Bacteriol.* **180**:1159–1165.
16. Burghoff, R., L. Pallesen, K. A. Krogfelt, J. V. Newman, M. Richardson, J. L. Bliss, D. C. Laux, and P. S. Cohen. 1993. Utilization of the mouse large intestine to select an *Escherichia coli* F-18 DNA sequence that enhances colonizing ability and stimulates synthesis of type 1 fimbriae. *Infect. Immun.* **61**:1293–1300.
17. Caprioli, A., V. Falbo, F. M. Ruggeri, F. Minelli, I. Ørskov, and G. Donelli. 1989. Relationship between cytotoxic necrotizing factor production and serotype in hemolytic *Escherichia coli*. *J. Clin. Microbiol.* **27**:758–761.
18. Dobrindt, U., and J. Hacker. 1999. Plasmids, phages and pathogenicity islands: lessons on the evolution of bacterial toxins, p. 3–23. *In* J. E. Allouf and J. Freer (ed.), *The comprehensive sourcebook of bacterial protein toxins*. Academic Press, London, England.
19. Donnenberg, M. S., S. Tzipori, M. L. McKee, A. D. O'Brien, J. Alroy, and J. B. Kaper. 1993. The role of the *eae* gene of enterohemorrhagic *Escherichia coli* in intimate attachment in vitro and in a porcine model. *J. Clin. Invest.* **3**:1418–1424.
20. Fetherston, J. D., P. Schuetze, and R. D. Perry. 1992. Loss of the pigmentation phenotype in *Yersinia pestis* is due to the spontaneous deletions of 102 kb of chromosomal DNA, which is flanked by a repetitive element. *Mol. Microbiol.* **6**:2693–2704.
21. Hacker, J., and J. B. Kaper. 1999. The concept of pathogenicity islands, p. 1–11. *In* J. B. Kaper and J. Hacker (ed.), *Pathogenicity islands and other mobile virulence elements*. ASM Press, Washington, D.C.
22. Hacker, J., and J. B. Kaper. 2000. Pathogenicity islands and the evolution of microbes. *Annu. Rev. Microbiol.* **54**:641–679.
23. Hacker, J., G. Blum-Oehler, B. Janke, G. Nagy, and W. Goebel. 1999. Pathogenicity islands of extraintestinal *Escherichia coli*, p. 59–76. *In* J. B. Kaper and J. Hacker (ed.), *Pathogenicity islands and other mobile virulence elements*. ASM Press, Washington, D.C.
24. Hacker, J., G. Blum-Oehler, I. Mühldorfer, and H. Tschäpe. 1997. Pathogenicity islands of virulent bacteria: structure, function and impact on microbial evolution. *Mol. Microbiol.* **23**:1089–1097.
25. Hacker, J., H. Kestler, H. Hoshützky, K. Jann, F. Lottspeich, and T. K. Korhonen. 1993. Cloning and characterization of the S fimbrial adhesin II complex of an *Escherichia coli* O18:K1 meningitis isolate. *Infect. Immun.* **61**:544–550.
26. Hacker, J., L. Bender, M. Ott, J. Wingender, B. Lund, R. Marre, and W. Goebel. 1990. Deletions of chromosomal regions coding for fimbriae and hemolysins occur in vitro and in vivo in various extraintestinal *Escherichia coli* isolates. *Microb. Pathog.* **8**:213–225.
27. Hull, R. A., E. R. 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.
28. Johnson, J. R., T. A. Russo, P. I. Tarr, U. Carlino, S. S. Bilge, J. C. Vary, Jr., and A. L. Stell. 2000. Molecular epidemiological and phylogenetic associations of two novel putative virulence genes, *iha* and *iroN_{E. coli}*, among *Escherichia coli* isolates from patients with urosepsis. *Infect. Immun.* **68**:3040–3047.
29. Kaper, J. B., and J. Hacker (ed.). 1999. Pathogenicity islands and other mobile virulence elements. ASM Press, Washington, D.C.
30. Khan, A. S., B. Knip, T. A. Ölschlager, I. van Die, T. K. Korhonen, and J. Hacker. 2000. Receptor structure for F1C fimbriae of uropathogenic *Escherichia coli*. *Infect. Immun.* **68**:3541–3547.
31. Knapp, S., J. Hacker, T. Jarchau, and W. Goebel. 1986. Large, unstable inserts in the chromosome affect the virulence properties of uropathogenic *Escherichia coli* O6 strain 536. *J. Bacteriol.* **168**:22–30.
32. Korhonen, T. K., M. V. Valtonen, J. Parkkinen, V. Väisänen-Rhen, J. Finne, I. Ørskov, F. Ørskov, S. B. Svenson, and P. H. Mäkelä. 1985. Serotype, hemolysin production, and receptor recognition of *Escherichia coli* strains associated with neonatal sepsis and meningitis. *Infect. Immun.* **48**:486–491.
33. Korhonen, T. K., V. Väisänen-Rhen, M. Rhen, A. Pere, J. Parkkinen, and J. Finne. 1984. *Escherichia coli* recognizing sialyl galactosides. *J. Bacteriol.* **159**:762–766.
34. Krallmann-Wenzel, U., M. Ott, J. Hacker, and G. Schmidt. 1989. Chromosomal mapping of genes encoding mannose-sensitive (type I) and mannose-resistant F (P) fimbriae of *Escherichia coli* O18:K5:H5. *FEMS Microbiol. Lett.* **49**:315–321.
35. Lawrence, J. G., and H. Ochman. 1998. Molecular archaeology of the *Escherichia coli* genome. *Proc. Natl. Acad. Sci. USA* **95**:9413–9417.
36. Levine, M. M., E. J. Bergquist, D. R. Nalin, D. A. Waterman, R. B. Hornick, C. R. Young, S. Scotman, and B. Row. 1978. *Escherichia coli* strains that cause diarrhea but do not produce heat labile or heat stable enterotoxins are not invasive. *Lancet* **i**:1119–1122.
37. Lindler, L. E., G. V. Plano, V. Burland, G. F. Mayhew, and F. R. Blattner. 1998. Complete DNA sequence and detailed analysis of the *Yersinia pestis* KIM5 plasmid encoding murine toxin and capsular antigen. *Infect. Immun.* **66**:5731–5742.
38. Maurelli, A. T., R. E. Fernandez, C. A. Bloch, C. K. Rode, and A. Fasano. 1998. Black holes and bacterial pathogenicity: a large genomic deletion that enhances the virulence of *Shigella* spp. and enteroinvasive *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **95**:3943–3949.
39. Mühldorfer, I., and J. Hacker. 1994. Genetic aspects of *Escherichia coli* virulence. *Microb. Pathog.* **16**:171–181.
40. Nagy, G., U. Dobrindt, M. Kupfer, L. Emödy, H. Karch, and J. Hacker. 2001. Expression of hemin receptor molecule ChuA is influenced by RfaH in uropathogenic *Escherichia coli* strain 536. *Infect. Immun.* **69**:1924–1928.
41. Nowicki, B., J. Vuopio-Varkila, P. Viljanen, T. K. Korhonen, and P. H. Mäkelä. 1986. Fimbrial phase variation and systemic *Escherichia coli* infection studied in the mouse peritonitis model. *Microb. Pathog.* **1**:335–347.
42. O'Brien, A. D., T. A. Lively, T. W. Chang, and S. L. Gorbach. 1983. Purification of *Shigella dysenteriae* 1 (Shiga)-like toxin from *Escherichia coli* O157:H7 strain associated with haemorrhagic colitis. *Lancet* **ii**:573.
43. Ochman, H., and R. K. Selander. 1984. Standard reference strains of *Escherichia coli* from natural populations. *J. Bacteriol.* **157**:690–693.
44. Oefner, P., S. P. Hunicke-Smith, L. Chiang, F. Dietrich, J. Mulligan, and R. W. Davis. 1996. Efficient random subcloning of DNA sheared in a recirculating point-sink flow system. *Nucleic Acids Res.* **24**:3879–3886.
45. Ørskov, I., and F. Ørskov. 1985. *Escherichia coli* in extraintestinal infections. *J. Hyg.* **95**:551–575.
46. Ott, M., H. Hoshützky, K. Jann, I. van Die, and J. Hacker. 1988. Gene clusters for S fimbrial adhesin (*sfa*) and F1C fimbriae (*foc*) of *Escherichia coli*: comparative aspects of structure and function. *J. Bacteriol.* **170**:3983–3990.
47. Ott, M., J. Hacker, T. Schmoll, T. Jarchau, T. K. Korhonen, and W. Goebel. 1986. Analysis of the genetic determinants coding for the S-fimbrial adhesin (*sfa*) in different *Escherichia coli* strains causing meningitis or urinary tract infections. *Infect. Immun.* **54**:646–653.
48. Ott, M., L. Bender, G. Blum, M. Schmittroth, M. Achtman, H. Tschäpe, and J. Hacker. 1991. Virulence patterns and long-range genetic mapping of extraintestinal *Escherichia coli* K1, K5, and K100 isolates: use of pulsed-field gel electrophoresis. *Infect. Immun.* **59**:2664–2672.
49. Pawelzik, M., J. Heesemann, J. Hacker, and W. Opferkuch. 1988. Cloning and characterization of a new type of fimbriae (S/FIC related fimbria) expressed by an *Escherichia coli* O75:K1:H7 blood culture isolate. *Infect. Immun.* **56**:2918–2924.
50. Perna, N. T., G. Plunkett III, V. Burland, B. Mau, J. D. Glasner, D. J. Rose, G. F. Mayhew, P. S. Evans, J. Gregor, H. A. Kirkpatrick, G. Pósfai, J. Hackett, S. Klink, A. Boutin, Y. Shao, L. Miller, E. J. Grotbeck, N. W. Davis, A. Lim, E. T. Dimalanta, K. D. Potamouisis, J. Apodaca, T. S. Anantharaman, J. Lin, G. Yen, D. C. Schwartz, R. A. Welch, and F. R. Blattner. 2001. Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature* **409**:529–533.
51. Rode, C. K., L. J. Melkerson-Watson, A. T. Johnson, and C. A. Bloch. 1999. Type-specific contributions to chromosome size differences in *Escherichia coli*. *Infect. Immun.* **19**:230–236.
52. Russo, T. A., U. B. Carlino, A. Mong, and S. T. Jodush. 1999. Identification

- of genes in an extraintestinal isolate of *Escherichia coli* with increased expression after exposure to human urine. *Infect. Immun.* **67**:5306–5314.
53. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
 54. Schmoll, T., H. Hoshützky, J. Morschhäuser, F. Lottspeich, K. Jann, and J. Hacker. 1989. Analysis of genes coding for the sialic acid-binding adhesin and two other minor fimbrial subunits of the S-fimbrial adhesin determinant of *Escherichia coli*. *Mol. Microbiol.* **3**:1735–1744.
 55. Schubert, S., A. Rakin, H. Karch, E. Carniel, and J. Heesemann. 1998. Prevalence of the “high-pathogenicity island” of *Yersinia* species among *Escherichia coli* strains that are pathogenic to humans. *Infect. Immun.* **66**:480–485.
 56. Staden, R., K. F. Beal, and J. K. Bonfield. 2000. The Staden package 1998. *Methods Mol. Biol.* **132**:115–130.
 57. Sussman, M. 1997. *Escherichia coli* and human disease, p. 3–48. In M. Sussman (ed.), *Escherichia coli*: mechanisms of virulence. Cambridge University Press, Cambridge, England.
 58. Swenson, D. L., N. O. Bukanov, D. E. Berg, and R. A. Welch. 1996. Two pathogenicity islands in uropathogenic *Escherichia coli* J96: cosmid cloning and sample sequencing. *Infect. Immun.* **64**:3736–3743.
 59. Tobe, T., T. Hayashi, C. G. Han, G. K. Schoolnik, E. Ohtsubo, and C. Sasakawa. 1999. Complete DNA sequence and structural analysis of the enteropathogenic *Escherichia coli* adherence factor plasmid. *Infect. Immun.* **67**:5455–5462.
 60. Van Vliet, F., A. Boyen, and N. Glansdorff. 1988. On interspecies gene transfer: the case of the *argF* gene of *Escherichia coli*. *Ann. Inst. Pasteur Microbiol.* **139**:493–496.
 61. Vokes, S. A., S. A. Reeves, A. G. Torres, and S. M. Payne. 1999. The aerobactin iron transport system genes in *Shigella flexneri* are present within a pathogenicity island. *Mol. Microbiol.* **33**:63–73.
 62. York, H. K., and M. Stodolsky. 1981. Characterization of *P1argF* derivatives from *Escherichia coli* K12 transduction. I. *ISI* elements flank the *argF* gene segment. *Mol. Gen. Genet.* **181**:230–240.
 63. Zhou, D., W.-D. Hardt, and J. Galán. 1999. *Salmonella typhimurium* encodes a putative iron transport system within the centisome 63 pathogenicity island. *Infect. Immun.* **67**:1974–1981.
 64. Zingler, G., M. Ott, G. Blum, U. Falkenhagen, G. Naumann, W. Sokolowska-Köhler, and J. Hacker. 1992. Clonal analysis of *Escherichia coli* serotype O6 strains from urinary tract infections. *Microb. Pathog.* **12**:299–310.
 65. Zingler, G., G. Blum, U. Falkenhagen, I. Ørskov, F. Ørskov, J. Hacker, and M. Ott. 1993. Clonal differentiation of uropathogenic *Escherichia coli* isolates of serotype O6:K5 by fimbrial antigen typing and DNA long-range mapping techniques. *Med. Microbiol. Immunol.* **182**:13–24.

Editor: V. J. DiRita