CLINICAL MICROBIOLOGY REVIEWS, Oct. 1997, p. 569–584 Vol. 10, No. 4
0893-8512/97/$04.00+0

Occurrence, Distribution, and Associations of O and H Serogroups, Colonization Factor Antigens, and Toxins of Enterotoxigenic *Escherichia coli*

MARCIA K. WOLF*

Department of Gastroenterology, Division of Medicine, Walter Reed Army Institute of Research, Washington, DC 20307-5100

INTRODUCTION

Enterotoxigenic *Escherichia coli* (ETEC) is a common cause of infectious diarrhea (6), especially in tropical climates, where uncontaminated water is not readily available. Most illness, in terms of both numbers of cases and severity of symptoms, occurs in infants and young children after weaning. Another significant population at risk is travelers who lack recent exposure to ETEC, but mortality is not a problem in this group. Occasional outbreaks are reported in industrialized countries (4, 74).

Diarrhea caused by ETEC has much in common with cholera; both result from ingestion of rather large inocula of gram-negative bacteria that then colonize the small intestine and produce toxins that cause net secretion into the intestinal lumen. Typically, the diarrhea is not severe and resolves without treatment in about a week. Two enterotoxins are produced by ETEC, heat-labile toxin (LT) and heat-stable toxin (ST). One or both toxins may be expressed in ETEC. LT is very similar to cholera toxin; their amino acid sequences are approximately 80% homologous, and they have the same mode of action (66). Other than the toxins, ETEC and *Vibrio cholerae* are not closely related.

Besides the enterotoxins, three other factors have been used to identify and characterize ETEC: O serogroup, H serogroup, and colonization factor antigens (CFAs). These are based on lipopolysaccharide, flagella, and fimbrial adhesins, respectively. They are important since all are exposed on the surface of ETEC and are promising candidates for protective antigens in vaccines against ETEC.

This review examines the occurrence, distribution, and associations of O serogroup, H serogroup, CFAs, and toxins on ETEC from widespread locations. It is my hope that this will point to good candidates for vaccine components.

METHODS

The data included in this analysis are from reports that contain complete descriptions of ETEC strains that include O serogroup, H serogroup, CFA, toxin, location of isolation, and a measure of their frequency. These criteria were met in reports of ETEC from Argentina (5), Bangladesh (28, 42), Burma (39), Brazil (29), Central African Republic (7), Chile (1, 21), Egypt (73), India (7, 59), Japan (7), New Caledonia (3), Peru (39), Rhodesia/Zaire (39), Saudi Arabia (72, 73), Spain (7, 8, 27), Thailand (11, 42), United States (74), Latin America (7), and Southeast Asia (70). These were included in the database of 988 isolates, and the distribution is shown in Fig. 1. Most isolates were obtained from patients who visited clinics while suffering from diarrhea. Testing for antigens, especially for CFAs, varied; therefore, some ETEC listed as not expressing any CFA may have an antigen that was not tested in a given study. Some studies (5, 11, 28, 29, 42, 70) reported only O:H serotypes for CFA-bearing ETEC. These 711 CFA-negative ETEC are not included in the database and so are not represented in the tables and figures, but they are discussed where appropriate.

Data from these sites were entered into a database by using dBase III Plus (Ashton-Tate copyright 1985 and 1986). Fields included O serogroup, H serogroup, CFA, coli surface antigens (CS), toxin, location, and number of isolates. Some analyses were carried out with Microsoft Excel version 5.0 (copyright 1985 to 1993, Microsoft Corp.).

Interpretation of this database requires noting that it is not comprehensive. ETEC is known to occur in some locations that are not represented because the reports did not present the data in a way that allowed characterization of individual

* Mailing address: Department of Gastroenterology, Division of Medicine, Walter Reed Army Institute of Research, Washington, DC 20307-5100. Phone: (202) 782-7184. Fax: (202) 782-9147. E-mail: WolfM@wrsmtp-ccmail.army.mil.

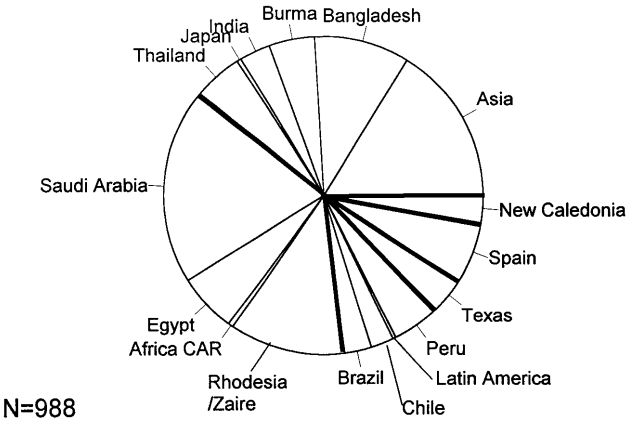

FIG. 1. Distribution of ETEC isolates in the database by location.

isolates (most notably Mexico) or did not include serotyping (13). Many regions surely have ETEC but have not been sampled. A second limitation is that the reported number of isolates is at best an estimate. The reports are generally from patients who presented with diarrhea at study sites and not from proactive sampling of all people with diarrhea. Some strains may be overrepresented if they caused an outbreak that occurred while sampling was in progress. Therefore, it is essential to consider both the occurrence and the distribution of phenotypes when weighing the relative contribution to diarrheal disease. A third caution is that any study only "takes a snapshot" of the etiology at a given time and the same location could have a different distribution of ETEC pathogens during another season. Fourth, some studies assayed for only a limited number of CFAs, so that those more recently described may have a much higher incidence than these numbers suggest.

O SEROGROUPS

Unlike the H antigens, CFA, and toxins, which are proteins, the O antigens are carbohydrates and are part of the lipopolysaccharide. Like the H antigens and CFA, the O antigen is composed of repeated subunits that extend from the surface of bacteria. For many years, the O serogroup was used to distinguish pathogenic from commensal *E. coli*. Since the pathogenesis of *E. coli* has become defined and pathogenic *E. coli* is classified by using probes of virulence factors, determining the O serogroup is less important. Still, O serogrouping provides worthwhile information; it is a good indication of the variety of strains that make up a group of *E. coli*. In ETEC, it was the category with the largest variety of antigens (Fig. 2).

Seventy-eight O serogroups were detected in 954 ETEC isolates (Table 1). The remaining 34 ETEC isolates either lacked side chains (were rough) or had unknown O serogroups. O6 was the most common, both in the number of isolates and in the number of locations where it was recovered; more than 16% (166) of the ETEC were O6, and it was isolated in 16 locations. O6, O78, O8, O128, and O153 accounted for over half the ETEC. Thirty-six O serogroups occurred only once and included some serogroups, such as O157, known to be associated with other pathogenic *E. coli* strains.

The data suggest that there is no incompatibility of a given O serogroup and H serogroup, CFA, or toxin, although often there is a "preferred" combination. O8 was a notable exception in that it was associated with more H serogroups (7) than any other O serogroup. O8:H9 was unusual because it was found with both CFA/II LTST and CS17 LT and both combinations were common. The significance of the large variety of antigens associated with O8, if any, is unknown.

It might be supposed that there are relationships among the O serogroups, but if such relationships exist, they are not evident from the data available at present. O serogroups are defined by reaction of the carbohydrate antigens with antibodies. O side chains with similar saccharides and linkages or constructed with common enzymes may not be detectable by using immunological assays. The carbohydrate structures of O6 (32, 33), O8 (35), O78 (34), and O153 (52) have been reported but do not reveal any commonality. Perhaps relationships would be apparent from genetic analysis of these antigens, but DNA sequences of biosynthetic enzymes are lacking.

H SEROGROUPS

The H serogroup is determined by the flagellar antigen. The role of flagella in the pathogenesis of ETEC has not been shown experimentally, and some authors suggest that the an-

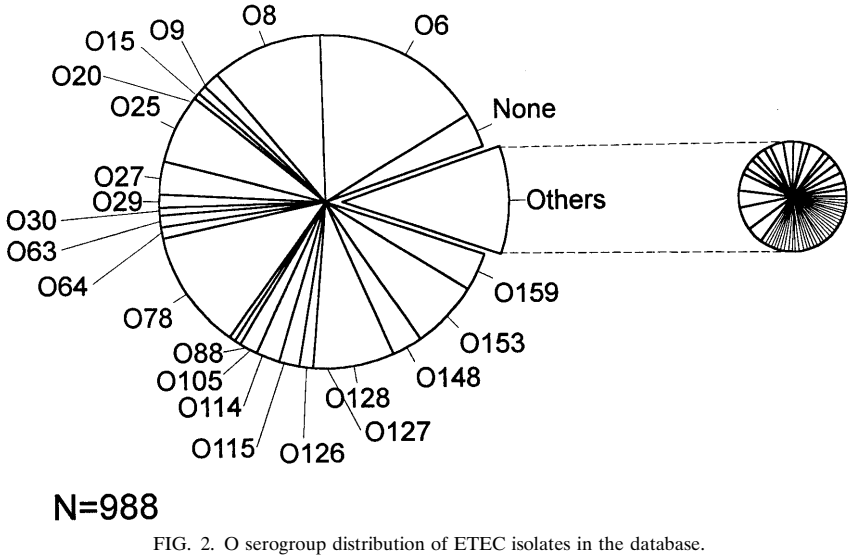

FIG. 2. O serogroup distribution of ETEC isolates in the database.

TABLE 1. Occurrence and association of O serogroups[a]

O serogroup	H serogroup			CFA			Toxin			Total	
	N	Identity	N_L	*N*	Identity	N_L	*N*	Identity	N_L	*N*	N_L
O1				1	ND	1	1	ST	1	1	1
O2	1	H7	1	2	ND	2	1	ST	1	2	2
							1	LT	1		
O3	1	H1	1	2	ND	1	2	LT	1	2	1
	1	H9	1								
O4				1	CFA/I	1				1	1
O6	143	H16	14	3	CFA/I	3	5	LT	3	166	16
	1	H40	1	151	CFA/II	15	121	LTST	15		
				3	CFA/IV	3					
				9	ND	3	4	ST	3		
O7	3	H18	2	1	CFA/I	1	2	LT	1	3	2
				2	ND	1	1	LTST	1		
O8	1	H1	1	3	CFA/I	1	46	LT	9	106	12
	60	H9	10	36	CFA/II	8	22	ST	4		
	1	H17	1	11	CFA/IV	2	16	LTST	6		
	4	H21	3	21	CS17	3					
	5	H32	1	35	ND	8					
	1	H33	1								
	1	H40	1								
O9	1	H2	1	1	CFA/II	1	4	LT	3	16	5
	11	H21	1	11	CFA/IV	1	11	ST	1		
				1	CS9	1					
				3	ND	2					
O11	1	H4	1	1	CFA/II	1	2	ST	2	3	2
	2	H33	1	2	CFAIV	2	1	LT	1		
O12				1	CFA/I	1	1	ST	1	1	1
O15	5	H11	2	1	CFA/I	1	3	ST	1	9	3
	1	H21	1	2	CS7	1	4	LTST	1		
				6	ND	2					
O17	2	H18	1	3	ND	1	1	LT	1	3	1
	1	H51	1				2	LTST	1		
O18				1	CFA/II	1	1	LTST	1	1	1
O20	1	H10	1	2	CFA/II	2	1	LT	1	6	3
				3	CFA/IV	1	2	ST	1		
				1	PCFO166	1	3	LTST	3		
O24	1	H15	1	1	ND	1	1	ST	1	1	1
O25	21	H42	4	1	CFA/II	1	43	LT	6	67	9
				1	CFA/III	1	2	ST	1		
				27	CFA/IV	6	5	LTST	3		
				38	ND	3					
O27	24	H7	6	20	CFA/IV	6	30	ST	8	30	8
	4	H20	2	10	ND	2					
O28	1	H9	1	1	ND	1	1	ST	1	1	1
O29	1	H4	1	12	CFA/IV	3	12	ST	3	12	3
	11	H21	2								
O30	7	H10	1	7	CFA/IV	1	7	ST	1	7	1
O32	1	H45	1	1	CFA/I	1	1	ST	1	1	1
O41	1	H30	1	1	ND	1	1	LT	1	1	1
O49				1	CFA/I	1	1	ST	1	1	1

Continued on following page

TABLE 1—*Continued*

O serogroup	H serogroup			CFA			Toxin			Total	
	N	Identity	N_L	*N*	Identity	N_L	*N*	Identity	N_L	*N*	N_L
O52	1	H10	1	1	CFA/I	1	1	ST	1	1	1
O56				1	ND	1	1	LT	1	1	1
O63	2	H10	1	10	CFA/I	3	3	ST	2	11	3
	6	H12	3	1	CFA/II	1	3	LTST	1		
O64	3	H5	2	10	ND	4	9	LT	4	10	4
							1	LTST	1		
O70				1	ND	1	1	ST	1	1	1
O71				2	ND	1	2	LT	1	2	1
O73	1	H17	1	1	ND	1	1	LT	1	1	1
O76	5	H21	1	5	ND	1	5	ST	1	5	1
O77	1	H7	1	1	CFA/IV	1	1	ST	1	1	1
O78	9	H11	2	112	CFA/I	7	36	ST	6	116	8
	79	H12	6	2	CFA/II	1	26	LTST	4		
	2	H18	1	2	PCFO166	1					
O80	1	H10	1	1	ND	1	1	LT	1	1	1
O81	1	H33	1	1	CFA/II	1	2	LTST	1	2	1
	1	H45	1	1	ND	1					
O85	1	H4	1	1	CFA/I	1	1	ST	1	1	1
O86	1	H18	1	1	CFA/II	1	1	ST	1	1	1
O88	4	H25	2	6	ND	3	5	LT	2	6	3
							1	ST	1		
O89	1	H15	1	1	ND	1	1	ST	1	1	1
O90	1	H12	1	1	CFA/I	1				1	1
O91	1	H10	1	1	ND	1	1	LTST	1	1	1
O92	1	H10	1	3	CFAIV	1	3	ST	1	3	1
	1	H12	1								
	1	H29	1								
O99	1	H6	1	1	ND	1	1	LT	1	1	1
O101				3	ND	1	1	LT	1	3	1
							2	ST	1		
O103	1	H49	1	1	CS7	1	1	LT	1	1	1
O104				1	CFA/IV	1	1	LT	1	1	1
O105	6	H18	1	6	ND	1	6	LT	1	6	1
O109	4	H45	1	4	ND	1	4	LTST	1	4	1
O110	1	H12	1	1	CFA/I	1				1	1
O113				1	ND	1	1	LTST	1	1	1
O114	11	H21	2	6	CS7	2	18	LT	5	18	5
	2	H49	1	11	CS17	2					
				1	ND	1					
O115	1	H1	1	2	CFA/I	2	10	ST	5	23	6
	1	H5	1	9	CFA/II	2	2	LTST	1		
	5	H12	1	12	CFA/IV	5					
	1	H35	1								
	5	H40	3								
	8	H51	2								

Continued on following page

TABLE 1—*Continued*

O serogroup	H serogroup			CFA			Toxin			Total	
	N	Identity	N_L	*N*	Identity	N_L	*N*	Identity	N_L	*N*	N_L
O126	13	H12	4	18	CFA/I	5	11	ST	4	19	5
	1	H30	1	1	ND	1	1	LTST	1		
O127	13	H21	1	13	CFA/I	1	13	ST	1	13	1
O128	2	H2	1	51	CFA/I	5	33	ST	5	76	7
	48	H12	5	19	CFA/IV	3	3	LT	2		
	13	H21	1	6	ND	2	19	LTST	2		
O133	1	H2	1	2	ND	2	1	LT	1	2	2
	1	H33	1				1	LTST	1		
O136	1	H9	1	1	ND	1	1	LT	1	1	1
O139	1	H28	1	1	CFA/II	1				1	1
O140	1	H20	1	1	CFA/II	1	1	ST	1	1	1
O141	2	H27	1	3	ND	1	3	LT	1	3	1
O143	1	H19	1	1	CS9	1	1	LT	1	1	1
O146				1	ND	1	1	LT	1	1	1
O148	26	H28	6	8	CFA/II	1	8	LT	1	29	6
				20	CFA/IV	5	19	ST	4		
				1	ND	1	2	LTST	2		
O149				1	ND	1	1	LT	1	1	1
O151	2	H17	1	1	CFA/II	1	1	ST	1	2	1
				1	CFA/IV	1	1	LTST	1		
O153	4	H10	3	64	CFA/I	7	63	ST	7	67	8
	1	H12	1	2	CFA/IV	1	2	LTST	2		
	58	H45	5	1	ND	1					
O154				1	CFA/II	1	1	LT	1	1	1
O155	1	H45	1	1	CFA/I	1	1	ST	1	1	1
O157				1	CFA/II	1	1	ST	1	1	1
O158	1	H10	1	1	CFA/II	1	2	LT	2	3	2
	1	H45	1	1	CFA/IV	1	1	LTST	1		
				1	ND	1					
O159	2	H4	1	1	CFAI	1	4	LT	3	37	5
	1	H20	1	31	CFAIV	1	32	ST	2		
				1	PCFO159	1	1	LTST	1		
				4	ND	3					
O160	1	H10	1	1	CFA/II	1	1	LTST	1	1	1
O162	1	H27	1	2	ND	2	1	LT	1	2	2
							1	ST	1		
O163	2	H5	1	1	CFA/I	1	4	ST	3	5	3
	1	H33	1	2	CFA/II	1	1	LTST	1		
	1	H34	1	2	CFA/IV	1					
O165	1	H9	1	1	CFA/I	1	1	ST	1	1	1
O166	1	H21	1	2	CFA/IV	2	1	LT	1	5	4
	2	H27	1	2	PCFO166	1	3	ST	2		
	1	H30	1	1	ND	1	1	LTST	1		
	1	H33	1								
O167	4	H5	2	4	CFA/IV	2	2	LT	2	5	3
				1	CS17	1					
O169	1	H20	1	9	CFA/IV	5	9	ST	5	9	5

[a] *N*, number of isolates; N_L, number of locations; ND, none detected.

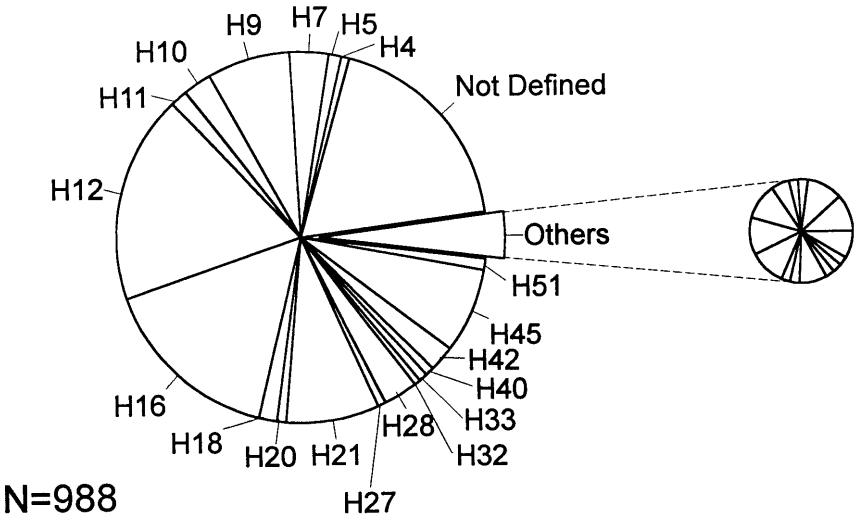

FIG. 3. H serogroup distribution of ETEC isolates in the database.

tigen is not a virulence factor (43, 51). The fact that most ETEC possess a flagellar antigen and analogy with *Vibrio cholerae* infection (43) suggest that it may be important in pathogenesis. The contribution of the H antigen to immunity to ETEC is unknown. At any rate, H serogroups serve as useful antigenic markers and as potential components of an ETEC vaccine.

Considerably fewer H serogroups than O serogroups are associated with ETEC (Fig. 3; Table 2). This reflects the smaller number of H serogroups in the *E. coli* typing scheme. A total of 34 H serogroups have been identified from 730 ETEC isolates (the remaining 258 of the ETEC strains either were nonmotile, lacked a defined H type, or were not tested). H12, H16, H21, H45, and H9 accounted for over half the ETEC in the database and were widespread, and so they are the most common H serogroups of ETEC. H7, H10, and H28 were widespread but were found less often. H7 has achieved some notoriety in recent years in combination with O157 and SLT in *E. coli* that cause a hemorrhagic colitis and hemolytic uremic syndrome (2). H7 was never associated with O157 in ETEC.

Some H serogroups were strongly associated with an O serogroup (that is, the O serogroup was present on at least 70% of the ETEC strains of a given H serogroup and the serotypes occurred at least seven times); those are O27:H7, O8:H9, O78 or O128:H12, O6:H16, O148:H28, O25:H42, and O153:H45. Some H types were strongly associated with a CFA (H11, H12, and H45 with CFA/I; H16 and H51 with CFA/II; and H20 and H42 with CFA/IV) or with a toxin (H20 and H45 with ST, and H11 and H16 with LTST). The significance of these combinations of antigens regarding enhanced virulence or genetic linkages, if any, is not known.

CFAs

CFAs, CS, and putative colonization factors (PCF) are proteins exposed on the surface of ETEC and are fimbrial (or fibrillar if they are especially thin). They promote attachment of the ETEC to epithelial cells of the small intestine and therefore serve as virulence factors (10). Both epidemiological (14) and challenge (10, 18, 38, 67) experiments in humans suggest that CFA are protective antigens such that immunity to a CFA protects against challenge by other ETEC strains expressing the same CFA. Because of technical difficulties and practical considerations, there are no hard data to suggest that all of the CFAs are colonization factors or protective antigens in humans, but it seems likely.

CFA/I was the first to be described (20), and there are convincing data that it is a colonization factor and a protective antigen in humans (18, 19, 22). Discovery of CFA/II followed a few years later (17), and CFA/II was later shown to be composed of CS3 alone or with CS1 or CS2 (63). These data led investigators to focus on them as a small group of proteins that might be included in vaccines to protect against ETEC. Most surveys of ETEC assyed for the presence of these two CFAs, and many also assayed for CFA/IV (68). CFA/IV was later shown to be composed of CS6 alone or with CS4 or CS5 (69). Over the years, additional fimbrial antigens have been discovered, but most studies did not include these.

CFA/I, CFA/II, and CFA/IV have been extensively characterized both biochemically and genetically (10). Genes for the structural subunits occur in operons with genes for synthesis, transport, and assembly of the subunits on the bacterial surface. CFA/I, CS1, CS2, CS4, CS17, and PCFO166 have much homology, while CS3, CS5, and CS6 are distinct. The genes for all are commonly carried on plasmids, and pieces of insertion sequences and transposons flank the operons.

By any measure, CFA/I, CFA/II, and CFA/IV are common and widespread (Fig. 4; Table 3). Together, these have been found in 23 to 94% of the ETEC isolates at the locations tested, depending on the methods used to detect CFAs (see the discussion below).

CFA/I

Almost one-third (a total of 299) of the ETEC isolates in the database expressed CFA/I, and it was detected from 13 locations. CFA/I occurred with 23 O types, and 14 of these were found only once. O78, O153, O128, O126, O127, and O63 accounted for almost 90% of the ETEC isolates bearing CFA/I. All were widespread except for O127 and O63. Remarkably, none of the ETEC isolates in the database expressing CFA/I expressed LT as the sole toxin. Of those with the toxin indicated, two-thirds were ST and the remainder were LTST.

TABLE 2. Occurrence and association of H serogroups[a]

H serogroup	O serogroup			CFA			Toxin			Total	
	N	Identity	N_L	*N*	Identity	N_L	*N*	Identity	N_L	*N*	N_L
H1	1	O3	1	2	CFA/IV	2	1	LT	1	4	2
	1	O8	1	2	ND	1	3	ST	2		
	1	O115	1								
H2	1	O9	1	1	CFA/II	1	1	LT	1	4	3
	2	O128	1	1	CFA/IV	1	1	ST	1		
	1	O133	1	1	ND	1					
H4	1	O11	1	1	CFA/I	1	3	LT	2	6	3
	1	O29	1	3	CFA/IV	1	3	ST	2		
	1	O85	1	2	ND	1					
	2	O159	1								
H5	3	O64	2	2	CFA/II	2	3	LT	3	10	5
	1	O115	1	4	CFA/IV	2	2	LTST	2		
	2	O163	1	1	CS17	1	2	ST	2		
	4	O167	2	3	ND	2					
H6	1	O99	1	1	ND	1	1	LT	1	1	1
H7	1	O2	1	16	CFA/IV	4	4	LT	4	29	8
	24	O27	6	13	ND	4	25	ST	6		
	1	O77	1								
H9	1	O3	1	2	CFA/I	2	37	LT	5	65	12
	60	O8	10	28	CFA/II	7	9	LTST	5		
	1	O28	1	21	CS17	3	3	ST	2		
	1	O136	1	14	ND	5					
	1	O165	1								
H10	1	O20	1	5	CFA/I	3	1	LT	1	23	8
	7	O30	1	2	CFA/II	1	4	LTST	3		
	1	O52	1	13	CFA/IV	3	15	ST	5		
	2	O63	1	3	ND	3					
	1	O80	1								
	1	O91	1								
	1	O92	1								
	4	O153	3								
	1	O158	1								
	1	O160	1								
H11	5	O15	2	10	CFA/I	3	9	LTST	1	14	2
	9	O78	2	2	CS7	?					
				2	ND	1					
H12	6	O63	3	148	CFA/I	9	3	LT	2	162	9
	79	O78	6	2	CFA/II	1	41	LTST	4		
	1	O90	1	6	CFA/IV	1	62	ST	8		
	1	O92	1	6	ND	2					
	1	O110	1								
	5	O115	1								
	13	O126	4								
	48	O128	5								
	1	O153	1								
H14		(O−)		1	CFA/II	1	1	LTST	1	1	1
H15	1	O24	1	2	ND	1	2	ST	1	2	1
	1	O89	1								
H16	143	O6	14	3	CFA/I	3	5	LT	3	145	14
				130	CFA/II	13	105	LTST	13		
				3	CFA/IV	3	4	ST	3		
				9	ND	3					
H17	1	O8	1	1	CFA/I	1	1	LT	1	4	1
	1	O73	1	1	CFA/II	1	1	LTST	1		
	2	O151	1	1	CFA/IV	1	2	ST	1		
				1	ND	1					
H18	3	O7	2	1	CFA/I	1	8	LT	2	14	5
	2	O17	1	1	CFA/II	1	3	LTST	2		

Continued on following page

TABLE 2—*Continued*

H serogroup	O serogroup			CFA			Toxin			Total	
	N	Identity	N_L	*N*	Identity	N_L	*N*	Identity	N_L	*N*	N_L
	2	O78	1	2	PCFO166	1	3	ST	2		
	1	O86	1	10	ND	2					
	6	O105	1								
H19	1	O143	1	1	CS9	1	4	LT	1	4	1
				3	ND	1					
H20	4	O27	2	1	CFA/II	1	1	LTST	1	7	4
	1	O140	1	5	CFA/IV	3	6	ST	3		
	1	O159	1	1	PCFO159	1					
	1	O169	1								
H21	4	O8	3	13	CFA/I	1	13	LT	3	71	6
	11	O9	1	1	CFA/II	1	2	LTST	2		
	1	O15	1	38	CFA/IV	3	55	ST	3		
	11	O29	2	11	CS17	2					
	5	O76	1	8	ND	3					
	11	O114	2								
	13	O127	1								
	13	O128	1								
	1	O166	1								
H25	4	O88	2	4	ND	2	3	LT	1	4	2
							1	ST	1		
H26		(O−)		1	ND	1	1	LT	1	1	1
H27	2	O141	1	2	PCFO166	1	2	LT	1	6	3
	1	O162	1	4	ND	2	1	LTST	1		
	2	O166	1				3	ST	2		
H28	1	O139	1	9	CFA/II	2	8	LT	1	27	7
	26	O148	6	17	CFA/IV	5	2	LTST	2		
				1	ND	1	16	ST	4		
H29	1	O92	1	1	CFA/IV	1	1	ST	1	1	1
H30	1	O41	1	3	ND	3	1	LT	1	3	3
	1	O126	1				2	ST	2		
	1	O166	1								
H32	5	O8	1	1	CFA/IV	1	3	LT	1	5	1
				4	ND	1	1	LTST	1		
							1	ST	1		
H33	1	O8	1	1	CFA/I	1	2	LT	1	7	3
	2	O11	1	2	CFA/II	2	2	LTST	2		
	1	O81	1	2	CFA/IV	1	3	ST	2		
	1	O133	1	2	ND	1					
	1	O163	1								
	1	O166	1								
H34	1	O163	1	1	CFA/IV	1	1	ST	1	1	1
H35	1	O151	1	1	CFA/IV	1	1	ST	1	1	1
H40	1	O6	1	1	CFA/I	1	1	LT	1	7	5
	1	O8	1	1	CFA/II	1	1	LTST	1		
	5	O115	3	4	CFA/IV	2	2	ST	2		
				1	ND	1					
H42	21	O25	4	21	CFA/IV	4	4	LTST	2	21	4
							1	ST	1		
H45	1	O32	1	61	CFA/I	5	1	LT	1	67	6
	1	O81	1	6	ND	2	6	LTST	2		
	4	O109	1				60	ST	5		
	58	O153	5								
H48		(O−)		1	CFA/IV	1	1	ST	1	1	1
H49	1	O103	1	3	CS7	1	3	LT	1	3	1
	2	O114	1								
H51	1	O17	1	8	CFA/II	2	1	LT	1	9	3
	8	O115	2	1	ND	1	2	LTST	1		

[a] See Table 1, footnote *a*.

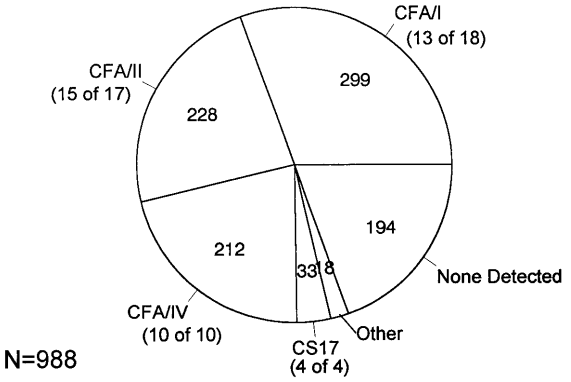

FIG. 4. CFA distribution of ETEC isolates in the database. The numbers in parentheses are the number of locations where a CFA was detected and the number of locations where that CFA was tested for.

CFA/II

Of the ETEC isolates in the database, 23% expressed CFA/II. This was the most widespread type of ETEC and was found in 15 locations. Two-thirds were O6:H16 LTST, but 20 additional O serogroups and 14 more H serogroups were expressed with CFA/II. Besides O6:H16, only strains with O8:H9 were common, so that strains with CFA/II have a variety of O and H types but only a few are common. This is different from strains with other CFAs that have a number of common serotypes. Of the ETEC isolates expressing CFA/II, 85% also expressed LTST, making this a very strong association.

Only half of the CFA/II isolates were tested for CS1 and CS2, and so the incidence is probably more than that reflected in this database. CS1 and CS2 were never found together and were never found in the absence of CS3, but CS3 occurred alone in 15% of the isolates that were also tested for CS1 and CS2. The distribution and occurrence of CS1 and CS2 are similar. Both are associated with O6:H16 LTST.

CFA/IV

Of the ETEC isolates in the database, 21% expressed CFA/IV, and it was found in strains from 10 locations. The number of O serogroups and H serogroups is more varied with strains expressing CFA/IV than with strains expressing CFA/I and CFA/II, and no O or H types predominate. Of the ETEC expressing CFA/IV, 80% expressed ST alone.

CS4 and CS5 were never found together and were never found in the absence of CS6, but CS6 occurred alone in 92% of the isolates that were tested for CS4 and occurred alone in 72% of the isolates tested for CS5, showing that CS6 is usually not expressed with CS4 or CS5. There are reports of CS6 being expressed with other fimbriae (37, 40), so that CS6 may be expressed with fimbrial antigens other than CS4 or CS5. CS4 and CS5 were not common but were widespread. Neither was associated with one O serogroup or H serogroup. CS5 expression was strongly associated with ST.

Others

CS17 has been tested only at four locations but has been found at all four, suggesting that it may be common. CS17 has been found most often with serotypes O8:H9 and O114:H21. One instance of O167:H5 with CS17 has been reported. O8:H9 with CFA/II as well as CS17 is common. CS17 has been expressed only with LT. Since LT isolates tend to lack defined CFAs (see below), CS17 is an important candidate for identifying fimbriae on these isolates.

Testing for CS7, CS9, PCFO159, and PCFO166 is limited, and they have not been documented to be common. Only O114 CS7 ETEC strains have been reported from more than one location, and that phenotype was found at two locations. Perhaps additional data will show that these antigens are significant, but for now this has not been documented.

CFA Not Detected

Of the ETEC isolates in the database, 20% lacked a defined CFA. Unfortunately, studies did not report the O:H serotypes of an additional 711 ETEC isolates lacking CFAs, and so they were not included in the database (5, 11, 28, 29, 42, 70). The lack of CFAs was not associated with a lack of testing for CFA/IV but was related to the sensitivity of the assays. The proportion of ETEC with defined CFAs was lowest in the studies in which immunodiffusion (33% of 458 ETEC strains) (70) and slide agglutination (43% of 613) (5, 8, 11, 28, 29, 59) were used as the assays for CFA expression and highest (69% of 573) in studies involving enzyme-linked immunosorbent assays or dot blot assays (39, 42, 72, 73). Assays with enzyme-linked antibodies should be considered the standard assay based on their sensitivity. Although it was not documented in this database, polyclonal antibodies are preferred over monoclonal antibodies.

There was a reproducibly low proportion of ETEC isolates with defined CFAs from Thailand (40% [42], 29% [70], and 28% [11]), even if an enzyme-linked immunosorbent assay was used (42), but the proportion of ETEC isolates expressing CFAs from Bangladesh (75% [28], 64% [42], and 38% [42]) reflected the sensitivity of the assay. At present, those strains tested by slide agglutination or immunodiffusion should be suspected of underrepresenting the true proportion of ETEC strains bearing CFAs.

Admittedly, the absence of 711 ETEC isolates from the database because of lack of reporting of O and H serogroups for ETEC without detected CFAs decreases the accuracy of the conclusions about these ETEC strains. However, since 646 of the 711 were from studies in which the less sensitive methods for detection of CFAs were used, it is likely that some of these do in fact express known CFAs, so that their inclusion in the database would be misleading. At any rate, analysis of the database is instructive, if flawed. The 20% of the ETEC isolates in the database lacking defined CFAs were recovered from 13 locations. The occurrence of a prominent phenotype might lead to a new common CFA, but instead there was a large variety of O and H serogroups associated with CFA-negative ETEC in the database: 44 O serogroups, 26 H serogroups, and 41 O:H serotypes. Half the O serogroups occurred only once. In contrast to the large variety of rare O and H serogroups associated with CFA-negative ETEC isolates, a majority (61%) expressed LT alone.

TOXINS

The distribution of toxins in ETEC is the simplest to analyze because there are only two toxins, LT and ST. *E. coli* must express one or both toxins to be classified as ETEC.

LT is very similar to cholera toxin; their amino acid sequences are approximately 80% homologous, and they have the same mode of action (57, 66). LT is composed of an A subunit that carries the enzymatic activity and five B subunits that bind to the receptor. Aside from this toxic activity, both LT and cholera toxin are potent antigens and are strong adju-

TABLE 3. Occurrence and associations of CFA[a]

CFA	CS	O serogroup			H serogroup			Serotype			Toxin			Total	
		N	Identity	N_L	*N*	Identity	N_L	*N*	Identity	N_L	*N*	Identity	N_L	*N*	N_L
CFA/I	NA	1	O4	1	1	H4	1	3	O6:H16	3	0	LT	0	299	13
		3	O6	3	2	H9	2	1	O7:H18	1	149	ST	12		
		1	O7	1	5	H10	3	1	O8:H9	1	55	LTST	6		
		3	O8	1	10	H11	2	1	O8:H17	1					
		1	O12	1	156	H12	9	1	O15:H11	1					
		1	O15	1	3	H16	3	1	O32:H45	1					
		1	O32	1	1	H17	1	1	O52:H10	1					
		1	O49	1	1	H18	1	2	O63:H10	1					
		1	O52	1	13	H21	1	6	O63:H12	3					
		10	O63	3	1	H33	1	9	O78:H11	2					
		112	O78	7	1	H40	1	79	O78:H12	6					
		1	O85	1	61	H45	5	1	O85:H4	1					
		1	O90	1				1	O90:H12	1					
		1	O110	1				1	O110:H12	1					
		2	O115	2				1	O115:H40	1					
		18	O126	5				13	O126:H12	4					
		13	O127	1				13	O127:H21	1					
		51	O128	5				42	O128:H12	5					
		64	O153	7				2	O153:H10	2					
		1	O155	1				1	O153:H12	1					
		1	O159	1				58	O153:H45	6					
		1	O163	1				1	O163:H33	1					
		1	O165	1				1	O165:H9	1					
CFA/II	CS3	151	O6	15	1	H2	1	128	O6:H16	13	16	LT	4	228	15
		36	O8	8	2	H5	1	27	O8:H9	7	8	ST	3		
		1	O9	1	28	H9	7	1	O8:H21	1	132	LTST	14		
		1	O11	1	2	H10	1	1	O9:H2	1					
		1	O18	1	2	H12	1	1	O11:H33	1					
		2	O20	2	1	H14	1	2	O78:H12	1					
		1	O25	1	130	H16	13	1	O81:H33	1					
		1	O63	1	1	H17	1	1	O86:H18	1					
		2	O78	1	1	H18	1	8	O115:H51	2					
		1	O81	1	1	H20	1	1	O139:H28	1					
		1	O86	1	1	H21	1	1	O140:H20	1					
		9	O115	2	9	H28	2	8	O148:H28	1					
		1	O139	1	2	H33	2	1	O151:H17	1					
		1	O140	1	1	H40	1	2	O158:H10	1					
		8	O148	1	8	H51	2	1	O160:H10	1					
		1	O151	1				2	O163:H5	1					
		1	O154	1											
		1	O158	1											
		1	O160	1											
		2	O163	1											
	CS1	34	O6	6	1	H10	1	34	O6:H16	6	3	LT	2	41	6
		2	O20	2	34	H16	6	1	O81:H33	1	2	ST	2		
		1	O81	1	1	H18	1	1	O86:H18	1	36	LTST	6		
		1	O86	1	1	H33	1	1	O160:H10	1					
		1	O154	1											
		1	O157	1											
		1	O160	1											
	CS2	47	O6	7	2	H5	1	40	O6:H16	5	0	LT	0	56	7
		2	O8	1	1	H10	1	1	O11:H33	1	6	ST	2		
		1	O11	1	40	H16	5	1	O140:H20	1	50	LTST	7		
		1	O18	1	1	H17	1	1	O151:H17	1					
		1	O140	1	1	H20	1	1	O158:H10	1					
		1	O151	1	1	H33	1	2	O163:H5	1					
		1	O158	1											
		2	O163	1											
CFA/IV	CS6	3	O6	3	2	H1	2	3	O6:H16	3	10	LT	4	212	10
		11	O8	2	2	H2	1	1	O8:H21	1	169	ST	9		

Continued on following page

TABLE 3—*Continued*

CFA	CS	O serogroup			H serogroup			Serotype			Toxin			Total	
		N	Identity	N_L	*N*	Identity	N_L	*N*	Identity	N_L	*N*	Identity	N_L	*N*	N_L
		11	O9	1	3	H4	1	1	O8:H32	1	11	LTST	5		
		2	O11	2	4	H5	2	11	O9:H21	1					
		3	O20	1	16	H7	4	1	O11:H4	1					
		27	O25	6	13	H10	3	1	O11:H33	1					
		20	O27	6	6	H12	1	1	O20:H10	1					
		12	O29	3	3	H16	3	21	O25:H42	3					
		7	O30	1	1	H17	1	14	O27:H7	4					
		1	O77	1	5	H20	3	4	O27:H20	2					
		3	O92	1	38	H21	3	1	O29:H4	1					
		1	O104	1	17	H28	5	11	O29:H21	2					
		12	O115	5	1	H29	1	7	O30:H10	1					
		19	O128	3	1	H32	1	1	O77:H7	1					
		20	O148	5	2	H33	1	1	O92:H10	1					
		1	O151	1	1	H34	1	1	O92:H12	1					
		2	O153	1	1	H35	1	1	O92:H29	1					
					4	H40	2								
		31	O159	1	21	H42	3	1	O115:H5	1					
		2	O163	1	1	H48	1	5	O115:H12	1					
		2	O166	2				1	O115:H35	1					
		4	O167	2				4	O115:H40	2					
		9	O169	5				2	O128:H2	1					
								13	O128:H21	1					
								17	O148:H28	5					
								1	O151:H17	1					
								2	O153:H10	1					
								1	O163:H34	1					
								1	O166:H21	1					
								1	O166:H33	1					
								3	O167:H5	1					
								1	O169:H20	1					
	CS4	2	O8	2	1	H7	1	1	O8:H21	1	0	LT	0	16	5
		1	O20	1	1	H10	1	1	O20:H10	1	10	ST	2		
		7	O25	4	2	H20	2	5	O25:H42	2	5	LTST	3		
		3	O27	2	1	H21	1	1	O27:H7	1					
		1	O148	1	1	H28	1	2	O27:H20	1					
		1	O136	1	1	H34	1	1	O148:H28	1					
					5	H42	2	1	O163:H34	1					
					1	H48	1								
	CS5	1	O6	1	2	H2	1	1	O6:H16	1	2	LT	1	39	6
		12	O29	3	2	H4	1	1	O29:H4	1	37	ST	6		
		1	O92	1	1	H5	1	11	O29:H21	2	0	LTST	0		
		6	O115	2	1	H7	1	1	O92:H21	1					
		15	O128	2	6	H12	1	1	O115:H5	1					
					1	H16	1	5	O115:H12	1					
					25	H21	2	2	O128:H2	1					
								13	O128:H21	1					
NA	CS7[b]	2	O15	?	2	H11	1	2	O15:H11	1	7	LT	2	9	2
		1	O103	1	3	H49	1	1	O103:H49	1	0	ST	0		
		6	O114	2				2	O114:H49	1	2	LTST	?		
	CS9[b]	1	O9	1	1	H19	1	1	O143:H19	1	3	LT	1	3	1
		1	O143	1							0	ST	0		
											0	LTST	0		
	CS17[b]	21	O8	3	1	H5	1	21	O8:H9	3	33	LT	4	33	4
		11	O114	2	21	H9	3	11	O114:H21	2	0	ST	0		
		1	O167	1	11	H21	2	1	O167:H5	1	0	LTST	0		
	O159[b]	1	O159	1	1	H20	1	1	O159:H20	1	0	LT	0	1	1
											0	ST	0		
											1	LTST	1		

Continued on following page

TABLE 3—*Continued*

CFA	CS	O serogroup			H serogroup			Serotype			Toxin			Total	
		N	Identity	N_L	*N*	Identity	N_L	*N*	Identity	N_L	*N*	Identity	N_L	*N*	N_L
	O166[b]	1	O20	1	2	H18	1	2	O78:H18	1	0	LT	0	5	2
		2	O78	1	2	H27	1	2	O166:H27	1	4	ST	1		
		2	O166	1							1	LTST	1		
ND[b]		1	O1	1	2	H1	1	1	O2:H7	1	126	LT	13	197	13
		2	O2	2	1	H2	1	1	O3:H1	1	42	ST	8		
		2	O3	1	2	H4	1	1	O3:H9	1	29	LTST	8		
		9	O6	3	3	H5	2	9	O6:H16	3					
		2	O7	1	1	H6	1	2	O7:H18	1					
		35	O8	8	13	H7	4	1	O8:H1	1					
		3	O9	2	14	H9	5	11	O8:H9	3					
		6	O15	2	3	H10	3	2	O8:H21	2					
		3	O17	1	2	H11	1	4	O8:H32	1					
		1	O24	1	6	H12	2	1	O8:H33	1					
		38	O25	3	2	H15	1	1	O8:H40	1					
		10	O27	2	9	H16	3	2	O15:H11	1					
		1	O28	1	1	H17	1	1	O15:H21	1					
		1	O41	1	10	H18	2	2	O17:H18	1					
		1	O56	1	3	H19	1	1	O17:H51	1					
		10	O64	4	8	H21	3	1	O24:H15	1					
		1	O70	1	4	H25	2	10	O27:H7	2					
					1	H26	10								
		1	O73	1	4	H27	2	1	O41:H30	1					
		5	O76	1	1	H28	1	3	O64:H5	2					
		1	O80	1	3	H30	3	1	O73:H17	1					
		1	O81	1	4	H32	1	5	O76:H21	1					
		6	O88	3	2	H33	1	1	O80:H10	1					
		1	O89	1	1	H40	1	1	O81:H45	1					
		1	O91	1	6	H45	2	3	O88:H25	1					
		1	O99	1	1	H51	1	1	O89:H15	1					
		3	O101	1				1	O91:H10	1					
		6	O105	1				1	O99:H6	1					
		4	O109	1				6	O105:H18	1					
		1	O113	1				4	O109:H45	1					
		1	O114	1				1	O126:H30	1					
		1	O126	1				6	O128:H12	2					
		6	O128	2				1	O133:H2	1					
		2	O133	2				1	O133:H33	1					
		1	O136	1				1	O136:H9	1					
		3	O141	1				2	O141:H27	1					
		1	O146	1				1	O148:H28	1					
		1	O148	1				1	O158:H45	1					
		1	O149	1				2	O159:H4	1					
		1	O153	1				1	O162:H27	1					
		1	O158	1				1	O166:H30	1					
		4	O159	3											
		2	O162	2											
		1	O166	1											

[a] See Table 1, footnote *a*. NA, not applicable; ND, none detected.
[b] Testing for these antigens was limited.

vants. This activity is being investigated to enhance antigen delivery to the mucosal immune system (31, 64). There is evidence that the B subunit of cholera toxin is protective against ETEC strains that express LT (12, 48), but multiple episodes of diarrhea caused by ETEC expressing LT are common (14, 75).

ST is present in 75% of the ETEC strains, either alone or with LT (Fig. 5). Its importance as a virulence factor is suggested in case-control or cohort studies, which find a significant correlation of ETEC expressing ST and diarrhea (13, 16, 26, 75). The molecular nature of ST and its receptor has been extensively studied in recent years (24, 45, 46, 56, 57, 60). ST is a family of peptides of less than 20 amino acids related to the mammalian hormone guanylin (60). (There is a second family of ST, called STII, STb, ST_B, or ST2, to distinguish it from the other family, more properly called STI, ST1, STa, or ST_A. STII peptides are not homologous to STI peptides, are not active in the suckling mouse model [15] that is the traditional assay for STI, and do not cause diarrhea by the same mechanism as STI [23, 24, 30]). Since ST was assayed by the suckling mouse assay in most of the ETEC isolates in the database, ST should be taken to mean STI in this review. This is appropriate because STII has not been proven to be a significant human virulence factor [57].) There are a number of variants of ST that cause secretory diarrhea by activating guanylyl cyclase (24, 57). Although ST is poorly immunogenic, it has been fused to LT (9, 54, 55) and antibodies have been detected that recognize native ST and partially neutralize toxicity in vitro (9, 55). It remains to be determined if these antibodies are protective in vivo. Since ST is an analog of the naturally occurring peptide

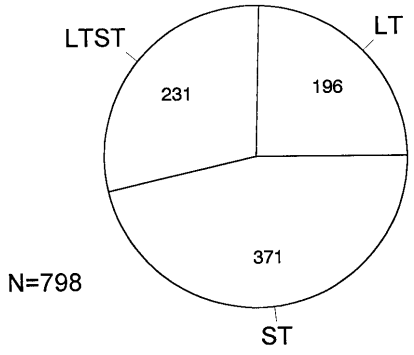

FIG. 5. Toxin distribution of ETEC isolates in the database.

guanylin and the receptor is widespread on epithelial cells (24, 57), it will be necessary to demonstrate that an immune response to ST does not interfere with necessary functions in humans by cross-reaction with guanylin.

Genes encoding LT and ST are carried on plasmids. Indeed, ent (for enterotoxin) plasmids were among the first to be associated with virulence (50, 53, 62, 65), and they played an important role in our present appreciation of transmissibility of virulence factors on plasmids and transposons. A single plasmid often carries a toxin and CFA, for example, CFA/I and ST (41, 44, 53), CFA/II and LT and ST (49, 50, 61), and CFA/IV and ST (71). This survey of a larger number of ETEC isolates than have been screened for plasmids supports the associations of CFA and toxins; LTST was strongly associated with CFA/II, and ST was associated with CFA/I and CFA/IV (Table 4). I find these strong associations surprising considering the bits and pieces of insertion sequences and transposases that flank the CFA operons that have been reported in the database of DNA sequences. This suggests that genes encoding toxins and CFA both arrived on transposons, elements that provide independent transfer. One explanation for the associations between toxin and CFA is that transposition occurred long ago and "winning combinations" of toxin and CFA genes persisted while sequences necessary for transposition became corrupted over time.

TABLE 4. Occurrence and association of toxins and CFAs[a]

Toxin	Total		CFA		
	N	N_L	*N*	Identity	N_L
LT	196	14	0	CFA/I	0
			16	CFA/II	4
			10	CFA/IV	4
			3	CS9	1
			33	CS17	4
			7	CS7	2
			1	CFA/III	1
			126	None	13
LTST	231	15	55	CFA/I	6
			132	CFA/II	14
			11	CFA/IV	5
			2	CS7	?
			1	PCFO159	1
			1	PCFO166	1
			29	None	8
ST	371	13	149	CFA/I	12
			8	CFA/II	3
			168	CFA/IV	9
			4	PCFO166	1
			42	None	8

[a] See Table 1, footnote *a*.

TABLE 5. Widespread ETEC Phenotypes[a]

No. of locations	No. of occurrences	O serogroup	H serogroup	CFA	Toxin
12	106	O6	H16	CFA/II	LTST
5	58	O153	H45	CFA/I	ST
	32	O78	H12	CFA/I	ST
	16	O148	H28	CFA/IV	ST
	7	O8	H9	CFA/II	LTST
4	14	O27	H7	CFA/IV	ST
	8	O169	H−	CFA/IV	ST
3	38	O25	H?	ND	LT
	21	O8	H9	CS17	LT
	18	O78	H12	CFA/I	LTST
	16	O128	H12	CFA/I	LTST
	10	O6	H−	CFA/II	LTST
	6	O126	H12	CFA/I	ST
	5	O8	H−	ND	LT
	4	O6	H16	ND	LTST

[a] See Table 1, footnote *a*.

ETEC isolates that express LT alone are the most likely to lack any of the CFAs that have been screened for in the collections, i.e., CFA/I, CFA/II, and CFA/IV. It seems likely that unknown CFAs will be found on these ETEC. In particular, CS17 seems likely to be present in many of the LT isolates. Only a few investigations have included a screen for CS17 (39), but even these few studies have found CS17 to be widespread and common.

COMBINATIONS OF ANTIGENS

The most widely distributed and most frequently occurring ETEC phenotype was O6:H16 CFA/II LTST; it was found in more than twice as many locations as any other phenotype but still accounted for only 11% of the isolates (Table 5). O153:H45 CFA/I ST, O78:H12 CFA/I ST, O148:H28 CFA/IV ST, O8:H9 CFA/II LTST, O27:H7 CFA/IV ST, and O169:H− CFA/IV ST account for 24% of the isolates in the database. Their actual contribution is probably somewhat greater, since some isolates lacked one phenotypic characteristic when tested (for example O6:H− CFA/II LTST and O6:H16 CFA-negative LTST), and these may have had the full phenotype during infection; ETEC isolates readily lose CFAs and toxins during storage. Even so, the fact that the most common phenotype comprises such a low proportion of the database is one of the striking features of ETEC; shigellae and salmonellae have much less variety of phenotypes. The phenotypes are not clustered by location but are widespread. Even phenotypes that occurred in small numbers tended to be dispersed over distant locations.

CONCLUSIONS

ETEC isolates express many serologically distinct antigens, especially O and H. The most common O serogroup accounted for only 16% of the isolates in the database, and so there were a large number of phenotypes based on O serogroup, H serogroup, CFA, and toxin. There was no evidence for clustering of antigens by location; a given location yielded a variety of phe-

notypes. Even though there was a large variety of antigens, some common combinations were evident. These combinations can be exploited for detection and vaccine development against ETEC.

The diversity of O serogroups is particularly impressive, as is the relatively small proportion of isolates lacking a defined O antigen. This is undoubtedly related to the thorough work of clinical microbiologists who have developed O serotyping as a reliable method to differentiate pathogenic *E. coli* strains from commensal isolates.

The characterization of CFAs is not as complete as the characterization of O serogroups and deserves special attention. One goal of this review was to document the occurrence of CFAs to determine whether the total number of CFAs is small enough to make them attractive vaccine components and to identify phenotypes of ETEC lacking defined CFAs so that they can be further investigated. An important finding was that the method of detection was responsible for a high proportion of ETEC strains lacking CFAs. These data clearly suggest that assays for CFAs should be based on assays involving enzyme-linked antibodies, since the sensitivity of the method is critical for detection. Probably the best assays will include polyclonal antisera to detect antigenic variants, but this has not been documented. Even the more sensitive assays have yielded an unsatisfactorily high proportion of ETEC isolates lacking CFAs in Peru (39) and Thailand (42). Further studies are required to see if these sites have CFAs that have not yet been defined. However, the relatively high proportion of ETEC isolates with defined CFAs at the five other locations from three continents suggests that the current compilation of CFAs represents most of the ETEC.

Except for toxins, the data included for analysis are based on recognition of antigens by antisera. Since epitopes may differ by relatively minor alterations in primary sequence or slight changes in conformation, some associations may exist that are not apparent from the data. This is unlikely for CFAs because the protein sequences have been directly determined or deduced and their relationships have been established in that way (10). However, O and H serogroups may have relationships that are not apparent from these data. The structures of some of the more common O antigens have been determined, and they appear distinct, at least to me. Efforts to develop common epitopes for O antigens could be fruitful.

How some combinations of phenotypes became so widespread is not clear. Perhaps there was a common strain disseminated by human travelers, but many of these isolates were recovered from infections in young, native inhabitants living in remote locations, who were unlikely to have direct contact with travelers. Unfortunately, data from multilocus enzyme electrophoresis (58), plasmid profiles, or outer membrane protein patterns (36), which have been used to determine how closely *E. coli* strains are related, are not available. A technique based on PCR appears promising (47) but has not been widely applied yet.

The common antigens are potential targets of exploitation for polyvalent vaccines against ETEC. Toxin- and CFA-based vaccines would have the broadest coverage with the fewest components, assuming that they can be developed as protective antigens. O and H antigens are too diverse to be practical unless common epitopes can be identified and exploited. Vaccines based on LT and ST alone would cover all ETEC isolates. LT is a promising antigen, but unfortunately ST is a poor antigen, and the constructs that are being developed have yet to be tested in humans. As for the CFAs, CFA/I, CS3, CS6, and CS17 would cover more than 75% of the ETEC isolates in the database. Addition of other pili may be indicated, if they are documented to be common. A pilus named longus has been suggested as a potential vaccine component based on the hybridization of an oligonucleotide DNA probe with 215 of 731 ETEC isolates (25). Longus is distinct from CFA and CS pili and was detected in ETEC strains expressing CFA/I, CFA/II, and CFA/IV as well as ETEC strains expressing none of these. If it is a protective antigen, it would be an important vaccine component.

Since most of the isolates that do not express CFA/I, CS3, or CS6 express LT as the sole toxin, a combination of the CFAs and an LT toxoid would cover 95% of the 988 ETEC isolates in the database and at least 85% of 1,699 ETEC isolates recovered from the 18 sites. The combination of CFA/I, CS3, CS6, and LT is therefore the most promising approach based on this analysis of ETEC isolates from 18 locations. Since there was generally not evidence for a predominance of one phenotype at one location, a polyvalent vaccine would be practical for all locations.

ACKNOWLEDGMENTS

I thank many coworkers at Walter Reed Army Institute of Research and Navy Medical Research Institute for encouragement and interest in this work. I especially thank Charles McQueen, Jeffrey Gambel, Andrew Trofa, Fred Cassels, and Tonia Agin for helpful comments on the manuscript.

This work was supported by the U.S. Army Medical Research and Material Command.

REFERENCES

1. **Agüero, M. E., L. Reyes, V. Prado, I. Orskov, F. Orskov, and F. C. Cabello.** 1985. Enterotoxigenic *Escherichia coli* in a population of infants with diarrhea in Chile. J. Clin. Microbiol. **22:**576–581.
2. **Armstrong, G. L., J. Hollingsworth, and J. G. Morris.** 1996. Emerging foodborne pathogens—*Escherichia coli* O157:H7 as a model of entry of a new pathogen into the food supply of the developed world. Epidemiol. Rev. **18:**29–51.
3. **Begaud, E., D. Mondet, and Y. Germani.** 1993. Molecular characterization of enterotoxigenic *Escherichia coli* (ETEC) isolated in New Caledonia (value of potential protective antigens in oral vaccine candidates). Res. Microbiol. **144:**721–728.
4. **Benoit, V., P. Raiche, M. G. Smith, J. Guthrie, E. F. Donnelly, E. M. Julian, R. Lee, S. DiMaio, M. Rittmann, and B. T. Matyas.** 1994. Foodborne outbreaks of enterotoxigenic *Escherichia coli*—Rhode Island and New Hampshire, 1993. Morbid. Mortal. Weekly Rep. **43:**81–89.
5. **Binsztein, N., M. J. Jouve, G. I. Viboud, L. L. Moral, M. Rivas, I. Orskov, C. Åhrén, and A. Svennerholm.** 1991. Colonization factors of enterotoxigenic *Escherichia coli* isolated from children with diarrhea in Argentina. J. Clin. Microbiol. **29:**1893–1898.
6. **Black, R. E.** 1993. Epidemiology of diarrhoeal disease: implications for control by vaccines. Vaccine **11:**100–106.
7. **Blanco, J., M. Blanco, E. A. Gonzalez, J. E. Blanco, M. P. Alonso, J. I. Garabal, and W. H. Jansen.** 1993. Serotypes and colonization factors of enterotoxigenic *Escherichia coli* isolated in various countries. Eur. J. Epidemiol. **9:**489–496.
8. **Blanco, J., E. A. Gonzalez, M. Blanco, J. I. Garabal, M. P. Alonso, S. Fernández, R. Villanueva, A. Aguilera, M. A. Garcia, J. Torres, A. Rey, W. H. Jansen, and P. A. M. Guinée.** 1991. Enterotoxigenic *Escherichia coli* associated with infant diarrhoea in Galicia, north-western Spain. J. Med. Microbiol. **35:**162–167.
9. **Cardenas, L., and J. D. Clements.** 1993. Development of mucosal protection against the heat-stable enterotoxin (ST) of *Escherichia coli* by oral immunization with a genetic fusion delivered by a bacterial vector. Infect. Immun. **61:**4629–4636.
10. **Cassels, F. J., and M. K. Wolf.** 1995. Colonization factors of diarrheagenic *E. coli* and their intestinal receptors. J. Ind. Microbiol. **15:**214–226.
11. **Changchawalit, S., P. Echeverria, D. N. Taylor, U. Leksomboon, C. Tirapat, B. Eampokalap, and B. Rowe.** 1984. Colonization factors associated with enterotoxigenic *Escherichia coli* isolated in Thailand. Infect. Immun. **45:**525–527.
12. **Clemens, J. D., D. A. Sack, J. R. Harris, J. Chakraborty, P. K. Neogy, B. Stanton, N. Huda, M. U. Khan, B. A. Kay, M. R. Khan, M. Ansarruzzaman, M. Yunus, M. R. Rao, A. Svennerholm, and J. Holmgren.** 1988. Cross-protection by B Subunit-whole cell cholera vaccine against diarrhea associated with heat-labile toxin-producing enterotoxigenic *Escherichia coli*: results of a large-scale field trial. J. Infect. Dis. **158:**372–377.

13. **Cravioto, A., R. E. Reyes, R. Ortega, G. Fernandez, R. Hernandez, and D. Lopez.** 1988. Prospective study of diarrhoeal disease in a cohort of rural Mexican children: incidence and isolated pathogens during the first two years of life. Epidemiol. Infect. **101:**123–134.
14. **Cravioto, A., R. E. Reyes, F. Trujillo, F. Uribe, A. Navarro, J. M. De La Roca, J. M. Hernandez, G. Perez, and V. Vazquez.** 1990. Risk of diarrhea during the first year of life associated with initial and subsequent colonization by specific enteropathogens. Am. J. Epidemiol. **131:**886–904.
15. **Dean, A. G., Y.-C. Ching, R. G. Williams, and L. B. Harden.** 1972. Test for *Escherichia coli* enterotoxin using infant mice: application in a study of diarrhea in children in Honolulu. J. Infect. Dis. **125:**407–411.
16. **Echeverria, P., D. N. Taylor, U. Lexsomboon, M. Bhaibulaya, N. R. Blacklow, K. Tamura, and R. Sakazaki.** 1989. Case-control study of endemic diarrheal disease in Thai children. J. Infect. Dis. **159:**543–548.
17. **Evans, D. G., and D. J. Evans, Jr.** 1978. New surface-associated heat-labile colonization factor antigen (CFA/II) produced by enterotoxigenic *Escherichia coli* of serogroups O6 and O8. Infect. Immun. **21:**638–647.
18. **Evans, D. G., D. J. Evans, Jr., A. R. Opekun, and D. Y. Graham.** 1988. Non-replicating oral whole cell vaccine protective against enterotoxigenic *Escherichia coli* (ETEC) diarrhea: stimulation of anti-CFA (CFA/I) and anti-enterotoxin (anti-LT) intestinal IgA and protection against challenge with ETEC belonging to heterologous serotypes. FEMS Microbiol. Immunol. **1:**117–125.
19. **Evans, D. G., D. Y. Graham, and D. J. Evans, Jr.** 1984. Administration of purified colonization factor antigens (CFA/I, CFA/II) of enterotoxigenic *Escherichia coli* to volunteers. Response to challenge with virulent enterotoxigenic *Escherichia coli*. Gastroenterology **87:**934–940.
20. **Evans, D. G., R. P. Silver, D. J. Evans, Jr., D. G. Chase, and S. L. Gorbach.** 1975. Plasmid-controlled colonization factor associated with virulence in *Escherichia coli* enterotoxigenic for humans. Infect. Immun. **12:**656–667.
21. **Fernández-Beros, M. E., V. Kissel, and M. E. Agüero.** 1988. Further characterization of *Escherichia coli* O153:H45, an ETEC serotype disseminated in Chile. Can. J. Microbiol. **34:**85–88.
22. **Freedman, D. J., C. O. Tacket, A. Delehanty, D. R. Maneval, and J. H. Crabb.** 1996. Milk antibodies against purified colonization factor antigens are protective in a human, enterotoxigenic *Escherichia coli* challenge study, abstr. P15, p. 371. *In* Abstracts of the 96th General Meeting of the American Society for Microbiology 1996. American Society for Microbiology, Washington, D.C.
23. **Fujii, Y., Y. Okamuro, S. Hitotsubashi, A. Saito, N. Akashi, and K. Okamoto.** 1994. Effect of alterations of basic amino acid residues of *Escherichia coli* heat-stable enterotoxin II on enterotoxicity. Infect. Immun. **62:**2295–2301.
24. **Giannella, R. A.** 1995. *Escherichia coli* heat-stable enterotoxins, guanylins, and their receptors: what are they and what do they do? J. Lab. Clin. Med. **125:**173–181.
25. **Giron, J. A., G. I. Viboud, V. Sperandio, O. G. Gomez-Duarte, D. R. Maneval, M. J. Albert, M. M. Levine, and J. B. Kaper.** 1995. Prevalence and association of the longus pilus structural gene (lngA) with colonization factor antigens, enterotoxin types, and serotypes of enterotoxigenic *Escherichia coli*. Infect. Immun. **63:**4195–4198.
26. **Gomes, T. A., V. Rassi, K. L. MacDonald, S. R. Ramos, L. R. Trabulsi, M. A. Vieira, B. E. Guth, J. A. Candeias, C. Ivey, M. R. Toledo, et al.** 1991. Enteropathogens associated with acute diarrheal disease in urban infants in São Paulo, Brazil. J. Infect. Dis. **164:**331–337.
27. **Gonzalez, E. A., J. Blanco, I. Garabal, and M. Blanco.** 1991. Biotypes, antibiotic resistance and plasmids coding for CFA/I and STa in enterotoxigenic *Escherichia coli* strains of serotype O153:H45 isolated in Spain. J. Med. Microbiol. **34:**89–95.
28. **Gothefors, L., C. Åhrén, B. Stoll, D. K. Barua, F. Orskov, M. A. Salek, and A. Svennerholm.** 1985. Presence of colonization factor antigens on fresh isolates of fecal *Escherichia coli*: a prospective study. J. Infect. Dis. **152:**1128–1133.
29. **Guth, B. E. C., E. G. Aguiar, P. M. Griffin, S. R. T. D. S. Ramos, and T. A. T. Gomes.** 1994. Prevalence of colonization factor antigens (CFAs) and adherence to HeLa cells in enterotoxigenic *Escherichia coli* isolated from feces of children in São Paulo. Microbiol. Immunol. **38:**695–701.
30. **Harville, B. A., and L. A. Dreyfus.** 1995. Involvement of 5-hydroxytryptamine and prostaglandin E2 in the intestinal secretory action of *Escherichia coli* heat-stable enterotoxin B. Infect. Immun. **63:**745–750.
31. **Holmgren, J., C. Czerkinsky, N. Lycke, and A. Svennerholm.** 1994. Strategies for the induction of immune responses at mucosal surfaces making use of cholera toxin B subunit as immunogen, carrier, and adjuvant. Am. J. Trop. Med. Hyg. **50:**42–54.
32. **Jann, B., A. A. Shashkov, H. Kochanowski, and K. Jann.** 1994. Structural comparison or the O6 specific polysaccharides from *E. coli* O6:K2:H1, *E. coli* O6:K13:H1, and *E. coli* O6:K54:H10. Carbohydr. Res. **263:**217–225.
33. **Jansson, P.-E., B. Lindberg, C. Lönngren, C. Ortega, and S. B. Svenson.** 1984. Structural studies of the *Escherichia coli* O-antigen 6. Carbohydr. Res. **131:**277–283.
34. **Jansson, P.-E., B. Lindberg, G. Widmalm, and K. Leontein.** 1987. Structural studies of the *Escherichia coli* O78 O-antigen polysaccharide. Carbohydr. Res. **165:**87–92.
35. **Jansson, P. E., J. Lönngren, G. Widmalm, K. Leontein, K. Slettengren, S. B. Svenson, G. Wrangsell, A. Dell, and P. R. Tiller.** 1985. Structural studies of the O-antigen polysaccharides of *Klebsiella* O5 and *Escherichia coli* O8. Carbohydr. Res. **145:**59–66.
36. **Kapur, V., D. G. White, R. A. Wilson, and T. S. Whittam.** 1992. Outer membrane protein patterns mark clones of *Escherichia coli* O2 and O78 strains that cause avian septicemia. Infect. Immun. **60:**1687–1691.
37. **Knutton, S., M. M. McConnell, B. Rowe, and A. S. McNeish.** 1989. Adhesion and ultrastructural properties of human enterotoxigenic *Escherichia coli* producing colonization factor antigens III and IV. Infect. Immun. **57:**3364–3371.
38. **Levine, M. M.** 1990. Vaccines against enterotoxigenic *Escherichia coli* infections I. Vaccines based predominantly or entirely on antibacterial immunity, p. 649–660. *In* G. C. Woodrow and M. M. Levine (ed.), New generation vaccines. Marcel Dekker, Inc., New York, N.Y.
39. **McConnell, M. M., M. L. Hibberd, M. E. Penny, S. M. Scotland, T. Cheasty, and B. Rowe.** 1991. Surveys of human enterotoxigenic *Escherichia coli* from three different geographical areas for possible colonization factors. Epidemiol. Infect. **106:**477–484.
40. **McConnell, M. M. and B. Rowe.** 1989. Prevalence of the putative colonization factors CFA/III and PCFO159:H4 in enterotoxigenic *Escherichia coli*. J. Infect. Dis. **159:**582–586.
41. **McConnell, M. M., H. R. Smith, G. A. Willshaw, A. M. Field, and B. Rowe.** 1981. Plasmids coding for colonization factor antigen I and heat-stable enterotoxin production isolated from enterotoxigenic *Escherichia coli*: comparison of their properties. Infect. Immun. **32:**927–936.
42. **McConnell, M. M., L. V. Thomas, N. P. Day, and B. Rowe.** 1985. Enzyme-linked immunosorbent assays for the detection of adhesion factor antigens of enterotoxigenic *Escherichia coli*. J. Infect. Dis. **152:**1120–1127.
43. **Moens, S., and J. Vanderleyden.** 1996. Functions of bacterial flagella. Crit. Rev. Microbiol. **22:**67–100.
44. **Murray, B. E., D. J. Evans, Jr., M. E. Peñaranda, and D. G. Evans.** 1983. CFA/I-ST plasmids: comparison of enterotoxigenic *Escherichia coli* (ETEC) of serogroups O25, O63, O78, and O128 and mobilization from an R factor-containing epidemic ETEC isolate. J. Bacteriol. **153:**566–570.
45. **Nandi, A., R. Mathew, and S. S. Visweswariah.** 1996. Expression of the extracellular domain of the human heat-stable enterotoxin receptor in *Escherichia coli* and generation of neutralizing antibodies. Protein Expression Purif. **8:**151–159.
46. **Ozaki, H., T. Sato, H. Kubota, Y. Hata, Y. Katsube, and Y. Shimonishi.** 1991. Molecular structure of the toxin domain of heat-stable enterotoxin produced by a pathogenic strain of *Escherichia coli*. A putative binding site for a binding protein on rat intestinal epithelial cell membranes. J. Biol. Chem. **266:**5934–5941.
47. **Pacheco, A. B. F., B. E. C. Guth, D. F. D. Almeida, and L. C. S. Ferreira.** 1996. Characterization of enterotoxigenic *Escherichia coli* by random amplification of polymorphic DNA. Res. Microbiol. **147:**175–182.
48. **Peltola, H., A. Siitonen, H. Kyronseppa, I. Simula, L. Mattila, P. Oksanen, M. J. Kataja, and M. Cadoz.** 1991. Prevention of travellers' diarrhoea by oral B- subunit/whole-cell cholera vaccine. Lancet **338:**1285–1289.
49. **Peñaranda, M. E., D. G. Evans, B. E. Murray, and D. J. Evans, Jr.** 1983. ST:LT:CFA/II plasmids in enterotoxigenic *Escherichia coli* belonging to serogroups O6, O8, O80, O85, and O139. J. Bacteriol. **154:**980–983.
50. **Peñaranda, M. E., M. Mann, D. G. Evans, and D. J. Evans, Jr.** 1980. Transfer of an ST:LT:CFA/II plasmid into *Escherichia coli* K-12 strain RR1 by co-transformation with pSC301 plasmid DNA. FEMS Microbiol. Lett. **8:**251–254.
51. **Polotsky, Y., E. Dragunsky, and T. Khavkin.** 1994. Morphologic evaluation of the pathogenesis of bacterial enteric infections. Crit. Rev. Microbiol. **20:**161–208.
52. **Ratnayake, S., A. Weintraub, and G. Widmalm.** 1994. Structural studies of the enterotoxigenic *Escherichia coli* (ETEC) O153 O-antigenic polysaccharide. Carbohydr. Res. **265:**113–120.
53. **Reis, M. H. L., H. T. Affonso, L. R. Trabulsi, A. J. Mazaitis, R. Maas, and W. K. Maas.** 1980. Transfer of a CFA/I-ST plasmid promoted by a conjugative plasmid in a strain of *Escherichia coli* of serotype O128ac:H12. Infect. Immun. **29:**140–143.
54. **Sanchez, J., R. Argotte, and A. Buelna.** 1997. Engineering of cholera toxin A-subunit for carriage of epitopes at its amino end. FEBS Lett. **401:**95–97.
55. **Sanchez, J., S. Johansson, B. Löwenadler, A. M. Svennerholm, and J. Holmgren.** 1990. Recombinant cholera toxin B subunit and gene fusion proteins for oral vaccination. Res. Microbiol. **141:**971–979.
56. **Sato, T., H. Ozaki, Y. Hata, Y. Kitagawa, Y. Katsube, and Y. Shimonishi.** 1994. Structural characteristics for biological activity of heat-stable enterotoxin produced by enterotoxigenic *Escherichia coli*: X-ray crystallography of weakly toxic and nontoxic analogs. Biochemistry **33:**8641–8650.
57. **Sears, C. L., and J. B. Kaper.** 1996. Enteric bacterial toxins: mechanisms of action and linkage to intestinal secretion. Microbiol. Rev. **60:**167–215.
58. **Selander, R. K., D. A. Caugant, H. Ochman, J. M. Musser, M. N. Gilmour, and T. S. Whittam.** 1986. Methods of multilocus enzyme electrophoresis for bacterial population genetics and systematics. Appl. Environ. Microbiol. **51:**873–884.

59. **Sen, D., U. Ganguly, M. R. Saha, S. K. Bhattacharya, P. Datta, D. Datta, A. K. Mukherjee, R. Chakravarty, and S. C. Pal.** 1984. Studies on *Escherichia coli* as a cause of acute diarrhoea in Calcutta. J. Med. Microbiol. **17:**53–58.
60. **Skelton, N. J., K. C. Garcia, D. V. Goeddel, C. Quan, and J. P. Burnier.** 1994. Determination of the solution structure of the peptide hormone guanylin: observation of a novel form of topological stereoisomerism. Biochemistry **33:**13581–13592.
61. **Smith, H. R., S. M. Scotland, and B. Rowe.** 1983. Plasmids that code for production of colonization factor antigen II and enterotoxin production in strains of *Escherichia coli*. Infect. Immun. **40:**1236–1239.
62. **Smith, H. W., and M. A. Linggood.** 1971. Observations on the pathogenic properties of the K88 Hly and Ent plasmids of *Escherichia coli* with particular reference to porcine diarrhoea. J. Med. Microbiol. **4:**467–485.
63. **Smyth, C. J.** 1982. Two mannose-resistant haemagglutinins of enterotoxigenic *Escherichia coli* of serotype O6:H16 or H− isolated from travellers' and infantile diarrhoea. J. Gen. Microbiol. **128:**2081–2096.
64. **Snider, D. P.** 1995. The mucosal adjuvant activities of ADP-ribosylating bacterial enterotoxins. Crit. Rev. Immunol. **15:**317–348.
65. **So, M., F. Heffron, and B. J. McCarthy.** 1979. The *E. coli* gene encoding heat stable toxin is a bacterial transposon flanked by inverted repeats of *IS*1. Nature **277:**453–456.
66. **Spangler, B. D.** 1992. Structure and function of cholera toxin and the related *Escherichia coli* heat-labile enterotoxin. Microbiol. Rev. **56:**622–647.
67. **Svennerholm, A., J. Holmgren, and D. A. Sack.** 1989. Development of oral vaccines against enterotoxigenic *Escherichia coli* diarrhoea. Vaccine **7:**196–198.
68. **Thomas, L. V., A. Cravioto, S. M. Scotland, and B. Rowe.** 1982. New fimbrial antigenic type (E8775) that may represent a colonization factor in enterotoxigenic *Escherichia coli* in humans. Infect. Immun. **35:**1119–1124.
69. **Thomas, L. V., M. M. McConnell, B. Rowe, and A. M. Field.** 1985. The possession of three novel coli surface antigens by enterotoxigenic *Escherichia coli* strains positive for the putative colonization factor PCF8775. J. Gen. Microbiol. **131:**2319–2326.
70. **Thomas, L. V., and B. Rowe.** 1982. The occurrence of colonisation factors (CFA/I, CFA/II and E8775) in enterotoxigenic *Escherichia coli* from various countries in South East Asia. Med. Microbiol. Immunol. **171:**85–90.
71. **Thomas, L. V., B. Rowe, and M. M. McConnell.** 1987. In strains of *Escherichia coli* O167 a single plasmid encodes for the coli surface antigens CS5 and CS6 of putative colonization factor PCF8775, heat-stable enterotoxin, and colicin Ia. Infect. Immun. **55:**1929–1931.
72. **Willshaw, G. A., T. Cheasty, B. Rowe, H. R. Smith, D. N. Faithfull-Davies, and T. G. J. Brooks.** 1995. Isolation of enterotoxigenic *Escherichia coli* from British troops in Saudi Arabia. Epidemiol. Infect. **115:**455–463.
73. **Wolf, M. K., D. N. Taylor, E. C. Boedeker, K. C. Hyams, D. R. Maneval, M. M. Levine, K. Tamura, R. A. Wilson, and P. Echeverria.** 1993. Characterization of enterotoxigenic *Escherichia coli* isolated from U.S. troops deployed to the Middle East. J. Clin. Microbiol. **31:**851–856.
74. **Wood, L. V., W. H. Wolfe, G. Ruiz-Palacios, W. S. Foshee, L. I. Corman, F. McCleskey, J. A. Wright, and H. L. DuPont.** 1983. An outbreak of gastroenteritis due to a heat-labile enterotoxin-producing strain of *Escherichia coli*. Infect. Immun. **41:**931–934.
75. **Zaki, A. M., H. L. DuPont, M. A. el Alamy, R. R. Arafat, D. Amin, M. M. Awad, L. Bassiouni, I. Z. Imam, G. S. el Malih, A. el Marsafie, M. S. Mohieldin, T. Naguib, M. A. Rakha, M. Sidaros, N. Wasef, C. E. Wright, and R. G. Wyatt.** 1986. The detection of enteropathogens in acute diarrhea in a family cohort population in rural Egypt. Am. J. Trop. Med. Hyg. **35:**1013–1022.