



OPEN

Identification and characterisation of enteroaggregative *Escherichia coli* subtypes associated with human disease

Samuel J. Ellis^{1,2}, Lisa C. Crossman^{3,4}, Conor J. McGrath^{1,2}, Marie A. Chattaway⁵, Johanna M. Hölken^{1,2}, Bernard Brett⁶, Leah Bundy³, Gemma L. Kay^{1,2}, John Wain^{1,2} & Stephanie Schüller^{1,2}✉

Enteroaggregative *E. coli* (EAEC) are a major cause of diarrhoea worldwide. Due to their heterogeneity and carriage in healthy individuals, identification of diagnostic virulence markers for pathogenic strains has been difficult. In this study, we have determined phenotypic and genotypic differences between EAEC strains of sequence types (STs) epidemiologically associated with asymptomatic carriage (ST31) and diarrhoeal disease (ST40). ST40 strains demonstrated significantly enhanced intestinal adherence, biofilm formation, and pro-inflammatory interleukin-8 secretion compared with ST31 isolates. This was independent of whether strains were derived from diarrhoea patients or healthy controls. Whole genome sequencing revealed differences in putative virulence genes encoding aggregative adherence fimbriae, *E. coli* common pilus, flagellin and EAEC heat-stable enterotoxin 1. Our results indicate that ST40 strains have a higher intrinsic potential of human pathogenesis due to a specific combination of virulence-related factors which promote host cell colonization and inflammation. These findings may contribute to the development of genotypic and/or phenotypic markers for EAEC strains of high virulence.

Enteroaggregative *E. coli* (EAEC) are emerging pathogens most commonly associated with acute and persistent paediatric diarrhoea and growth retardation in developing nations¹. In addition, EAEC are a major cause of acute diarrhoea in travellers to developing countries² and persistent enteric infection in HIV/AIDS patients³. Notably, a large outbreak in Germany in 2011 was caused by a hybrid O104:H4 EAEC strain which carried the Shiga toxin gene from enterohaemorrhagic *E. coli* and resulted in over 4,300 cases of diarrhoea, 900 hospitalisations and 50 deaths⁴. EAEC are a heterogeneous group and have been traditionally defined by their characteristic “stacked brick”-like aggregative adherence (AA) to HEp-2 cells⁵. Pathogenesis involves adherence to the intestinal epithelium, stimulation of mucus secretion, biofilm formation, cytotoxic damage and mucosal inflammation. Several putative virulence factors involved in these processes have been identified^{6,7}. These include the aggregative adherence fimbriae (AAF) of which five major variants have been described so far (AAF/I to AAF/V)^{8–12}. AAF are encoded on the pAA virulence plasmid which also contains the major transcriptional regulator AggR and the surface protein dispersin which contributes to AAF dispersal^{13–15}. Notably, alternative adhesins such as the *E. coli* common pilus (ECP) can mediate aggregative adherence, particularly in the absence of AAF¹⁶. Further potential virulence factors of EAEC include the cytotoxins Pet, enteroaggregative *Escherichia coli* heat-stable enterotoxin 1 (EAST-1), and haemolysin E (HlyE) which promote epithelial damage and intestinal colonization⁶. In addition, the serine protease Pic exhibits mucinolytic activity and also stimulates intestinal mucus secretion^{17,18}. Despite the identification of multiple putative EAEC virulence candidates, not all EAEC strains harbouring these factors cause human disease, and genotypic studies have failed to consistently associate a single gene or combination of genes with EAEC virulence^{6,19}. To determine if certain EAEC lineages were more strongly associated with disease than others, Chattaway and colleagues analysed 564 EAEC isolates from three case-control studies in Bangladesh,

¹Norwich Medical School, University of East Anglia, Norwich, UK. ²Quadram Institute Bioscience, Norwich, UK. ³School of Biological Sciences, University of East Anglia, Norwich, UK. ⁴SequenceAnalysis.co.uk, Norwich Research Park, Norwich, UK. ⁵Gastrointestinal Bacteria Reference Unit, Public Health England, London, UK. ⁶Department of Gastroenterology, Norfolk and Norwich University Hospital, Norwich, UK. ✉e-mail: s.schuller@uea.ac.uk

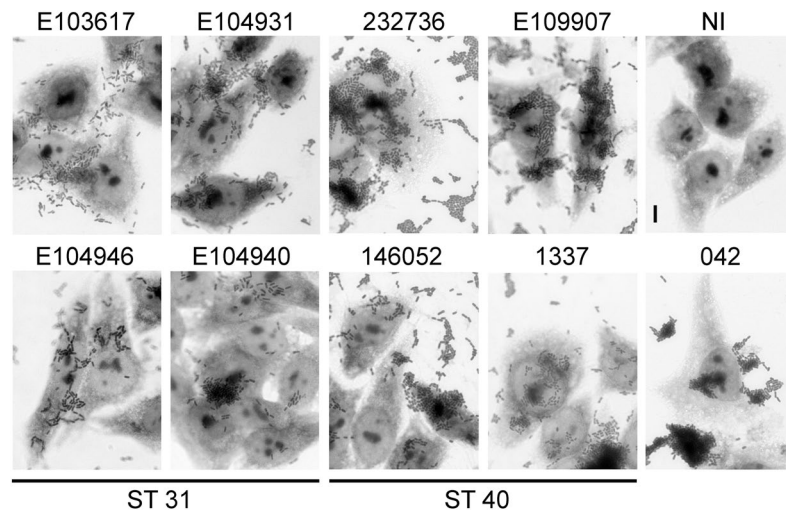


Figure 1. Aggregative adherence of EAEC to Hep-2 cells. Cells were infected with EAEC or left non-infected (NI) for 3 h, and adherent bacteria were visualised by Giemsa staining. Shown are representative images of two experiments performed in duplicate. Bar = 5 μ m.

Nigeria and the UK by Multi Locus Sequence Typing (MLST)²⁰. Of the 17 sequence types (STs) identified, ST40 and ST31 were significantly associated with disease and asymptomatic carriage, respectively. Here, we investigated if the association of EAEC ST40 with disease was linked to specific virulence phenotypes *in vitro*. To this aim, we selected eight strains each of ST40 (diarrhoea-associated) and ST31 (carriage-associated) and determined aggregative adherence (AA) to intestinal epithelium, interaction with mucus, biofilm formation, cytotoxicity and stimulation of an inflammatory response. In addition, all strains were genome-sequenced, and bioinformatic analysis was used to identify putative virulence genes associated with disease potential.

Results

ST40 strains adhere better to human colonic epithelium than ST31 isolates. For this study, eight strains from each ST31 and ST40 were selected including isolates from diarrhoea patients and healthy controls. The EAEC prototype strain 042 was included as positive control (Table S1). As many of the ST31 and ST40 strains were classified as EAEC by PCR probes only, all isolates were first tested for AA on Hep-2 cells which is the current gold-standard for the definition of the EAEC pathotype⁵. As shown in Fig. 1 (images shown for four strains of each ST only), all strains adhered in stacked-brick like aggregates to the cells and underlying coverslips. We further investigated adherence of the EAEC strains to human colonic epithelium *in vitro* and *ex vivo*. On confluent T84 colon carcinoma cells, adhesion of ST40 strains was significantly higher compared to that of ST31 isolates (Fig. 2A). In contrast, no differences in the number of cell-bound bacteria between isolates from cases and controls were detected. Quantification of EAEC growth in cell culture medium alone demonstrated replication of all strains during the 2 h incubation period which did not differ significantly between strains of ST31 and ST40 (Fig. S1). Similar results to those on T84 cells were obtained using IVOC of human colonic biopsies, although colonization levels of ST31 strains were more variable compared to those in cell culture. While all EAEC strains demonstrated AA to colonic tissue, colonization was significantly higher in EAEC strains of ST40 than ST31 (Fig. 2B,C). In addition, ST31 strains from controls adhered significantly better than ST31 isolates from patients with diarrhoea (Fig. 2B, $P < 0.01$).

Growth in mucus-containing medium is similar in ST31 and ST40 strains. Previous studies have suggested that utilisation of mucin as a nutrient source by EAEC might provide an advantage for intestinal colonization²¹. We therefore determined the growth of EAEC ST31 and ST40 strains in minimal M9 medium with or without porcine gastric mucin. Growth in M9 alone was low for all strains tested with significantly higher optical densities of ST31 versus ST40 strains ($p = 0.017$, Fig. 3A). However, addition of mucin enhanced growth of ST31 and ST40 strains to similar levels (Fig. 3A). No significant differences were detected in strains from cases and controls.

Infection with EAEC ST31 and ST40 strains does not affect mucin levels in LS174T cells. Mucus secretion is a major feature in EAEC-mediated diarrhoea. To investigate the influence of EAEC infection on cellular mucin levels, mucin-producing LS174T cells were infected with EAEC, and production of the major secreted intestinal mucin MUC2 was determined by immunofluorescence staining (Fig. 3B). Subsequent densitometric analysis did not reveal any significant effect of EAEC infection on MUC2 production for any of the ST strains tested (Fig. 3C).

EAEC ST40 strains form enhanced biofilms compared to ST31 isolates. EAEC infection is associated with the formation of thick aggregating biofilms in the mucus layer and at the epithelial surface which impedes antibiotic penetration and efficacy of treatment^{22,23}. Interestingly, ST40 isolates demonstrated

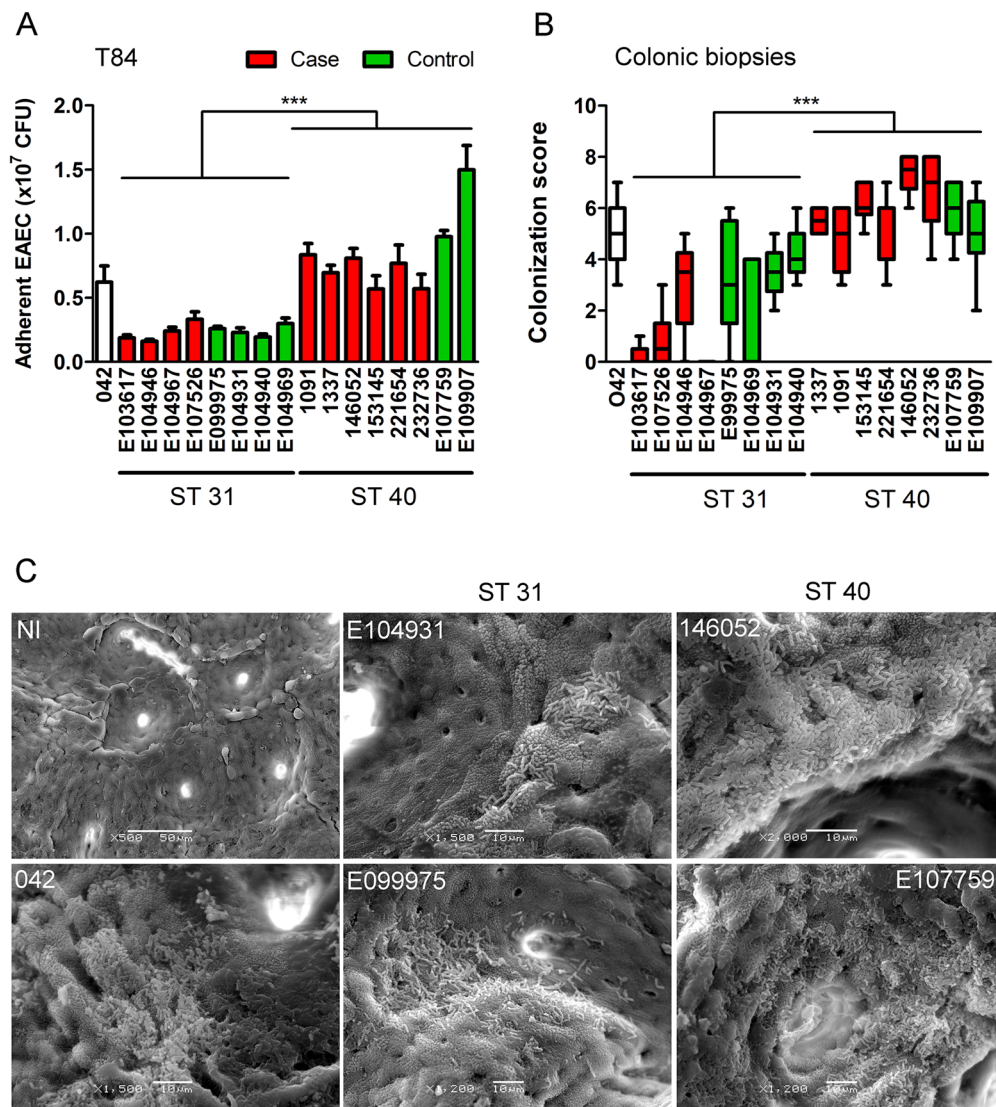


Figure 2. (A) Adhesion of EAEC ST strains to T84 cells after 2 h of infection. Adherent bacteria were quantified by determining colony-forming units (CFU) in cell lysates. Results represent the mean \pm standard error of the mean (SE) from three independent experiments in duplicate. Data from groups of ST31 and ST40 isolates were analysed using Student's unpaired t-test. $P^{***} < 0.001$. (B) Colonization of colonic biopsies after 7 h of incubation. Samples were evaluated by scanning electron microscopy, and colonization was ranked according to the size and frequency of bacterial aggregates as described in methods. Results from three independent experiments in duplicate are shown as box plots with medians. Significance was calculated using the non-parametric Mann-Whitney test. $P^{***} < 0.001$. (C) Scanning electron micrographs of colonic biopsies infected with EAEC or left non-infected (NI). Shown are representative images of three experiments performed in duplicate. Bar = 10 μ m (EAEC), 50 μ m (NI).

significantly enhanced biofilm formation on plastic surfaces compared to ST31 strains whilst reference strain 042 demonstrated the highest degree of biofilm formation (Figs. 4A and S2). No difference in biofilm intensity was detected between strains isolated from cases of diarrhoea and healthy controls within each ST.

EAEC ST31 and ST40 isolates exhibit similar cytotoxicity in T84 cells. A prominent feature of EAEC pathogenesis is the release of cytotoxins which promote epithelial cell death and mucosal damage^{6,24}. To determine the cytotoxic effect of ST31 and ST40 isolates on T84 cells, infections were performed for 8 h, and cell death was quantified by Trypan Blue stain. Infection with neither EAEC strain resulted in host cell detachment, and monolayer integrity after staining was confirmed by microscopy (Fig. S3, representative images shown for 042 and the highly cytotoxic strains E104946 and 221654). As shown in Fig. 4B, variable levels of cytotoxicity were observed, particularly among ST31 strains. However, there was no significant difference between the overall levels of T84 cell death caused by ST31 versus ST40 strains or between strains derived from cases or controls.

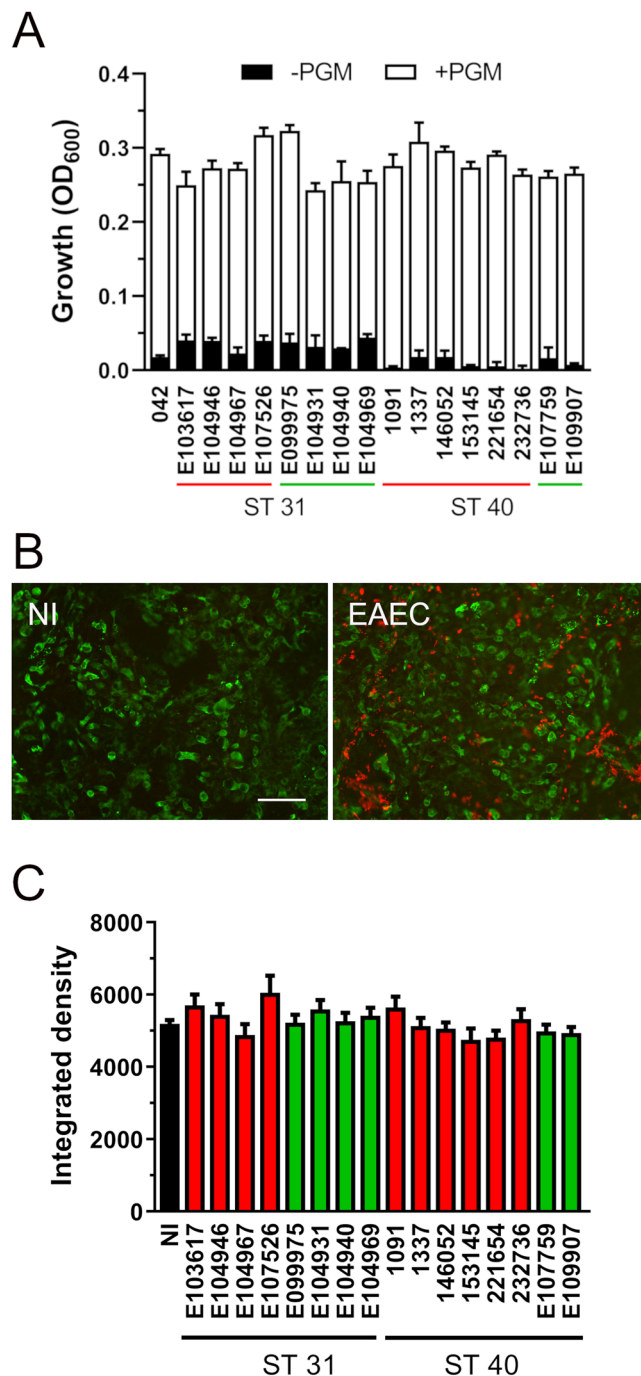


Figure 3. (A) Growth of EAEC ST strains in M9 medium with (+PGM) or without 0.5% porcine gastric mucin (-PGM) for 8 h was determined by optical density (OD_{600}). Strains from cases or controls are underlined in red and green, respectively. Results represent the mean \pm SE from three independent experiments. Data from groups of ST31 and ST40 isolates were analysed using Student's unpaired t-test. (B) Influence of EAEC infection on mucin production by LS174T cells. Cells were infected with EAEC strains for 4 h or left non-infected (NI). Mucus production was visualised by immunofluorescence staining for MUC2 (green), and EAEC were stained in red. Shown are representative images of NI samples and cells infected with E104931 (EAEC). Bar = 50 μ m. (C) MUC2 staining was quantified by integrated density. Results represent the mean \pm SE from at least four independent experiments in duplicate. Data from groups of ST31 and ST40 isolates were analysed using Student's unpaired t-test.

ST40 isolates induce higher levels of IL-8 secretion from T84 cells. To evaluate the inflammatory response associated with infection, confluent T84 cells were incubated with EAEC for 3 h. IL-8 concentrations in the supernatant were quantified by ELISA after an additional incubation of 20 h in gentamicin-containing

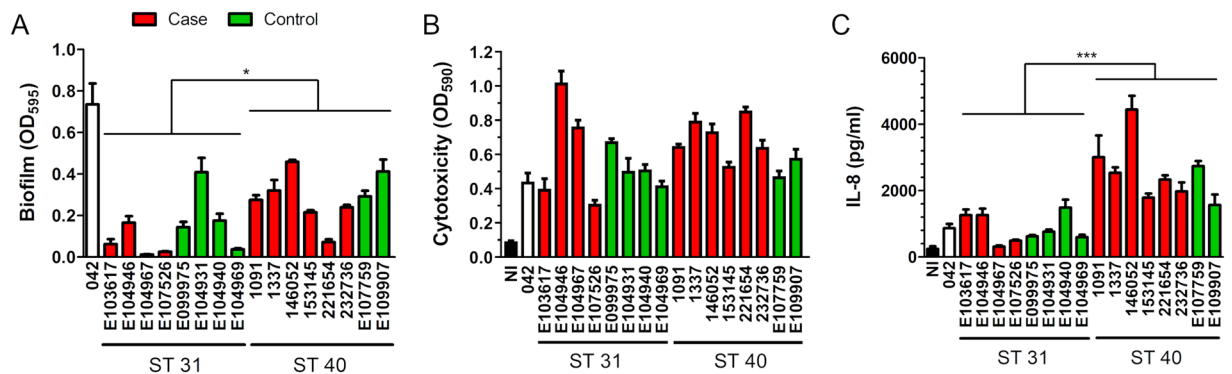


Figure 4. (A) Biofilm formation of EAEC ST strains. Bacteria were incubated in 96-well plates for 48 h, stained with crystal violet and absorbance was quantified (OD_{595}). (B) Cytotoxicity of EAEC infection. Death of T84 cells was quantified by Trypan Blue staining after 8 h of infection. Non-infected cells (NI) were included as controls. (C) Secretion of interleukin-8 (IL-8) by T84 cells infected with EAEC. IL-8 levels in supernatants were quantified by ELISA after 24 h of incubation. Results represent the mean \pm SE from three independent experiments in duplicate. Data from groups of ST31 and ST40 isolates were analysed using Student's unpaired t-test. $P^* < 0.05$, $P^{***} < 0.001$.

ST	Strain ^a	Serum agglutination	In silico typing
31	042	O44:H18	O44:H18
31	E099975	O15:H18	O15:H18
31	E104931	UD	O130:H27
31	E104940	O130:H27	O130:H27
31	E104969	O130:H25	O130:H27
31	E104946	UD	O15:H18
31	E104967	UD	O15:H18
31	E103617	O130:H27	O130:H27
31	E107526	UD	O130:H27
40	E107759	UD	O127:H21
40	E109907	O111ab:H-	O111:H21
40	1091	O111ac:H11	O111:H21
40	1337	O111ac:H11	O111:H21
40	146052	O154:H21	O25:H21
40	153145	O111:H21	O111:H21
40	221654	O111:H21	O111:H21
40	232736	O111:H21	O111:H21

Table 1. EAEC serotype determination. ^aIsolates from cases (bold) and controls. UD = undetermined.

medium, which allows host cell protein synthesis and secretion but prevents bacterial overgrowth. Trypan Blue staining of cell monolayers at the end of incubation confirmed that no significant cell death was caused by any of the strains compared to non-infected controls (data not shown). As demonstrated in Fig. 4C, infection with ST40 isolates resulted in significantly higher levels of IL-8 secretion compared to ST31 strains. In contrast, no association was observed between IL-8 concentrations and whether strains originated from cases or controls

ST31 and 40 strains exhibit differences in putative virulence genes. To determine genomic differences between the two STs which could explain their different phenotypic properties, whole genome sequencing was performed for all strains. Firstly, we determined the ST of all strains from the sequence reads and confirmed their previously established MLST²⁰. In addition, short read sequence typing was used to predict the serotype of each strain in silico which had only been partially established by serum agglutination testing before (Table 1). While ST31 strains belonged to serogroups O130:H27 or O15:H18, ST40 strains mainly displayed O111:H21. There was no considerable difference between serotypes of case or control strains, although more isolates need to be analysed to confirm this. Sequence analysis for EAEC-associated putative virulence genes obtained from GenBank (*aggR* FN554767.1, *pic* FN554766.1, *aap* FN554767.1, *astA* FN554767.1, *hlyE* FN554766.1, *pet* FN554767.1, *ecpA* FN554766.1, *aggA* U12894.1, *aafA* FN554767.1, *agg3A* AF411067.1, *hdaA* EU637023.1, and *aaf5A* AB571097.1) demonstrated that *astA* (EAST-1 toxin) was present in all ST40 but lacking in ST31 isolates (Table 2). Furthermore, *ecpA* (*E. coli* common pilus) was strongly associated with ST40 and encoded by only a quarter of ST31 strains. ST31 and 40 strains also differed in AAF type with ST40 strains harbouring *aaf5A*

	Strain ^a	Serotype	<i>aggR</i> ^b	<i>pic</i>	<i>aap</i>	<i>astA</i>	<i>hlyE</i>	<i>pet</i>	<i>ecpA</i>	<i>aggA</i> (AAF/I) ^c	<i>aafA</i> (AAF/II)	<i>agg3A</i> (AAF/III)	<i>hdaA</i> (AAF/IV)	<i>aaf5A</i> (AAF/V)
ST31	042	O44:H18	+	+	+	+	+	+	+		+			
	E103617	O130:H27	+	+	+		+		+	+				
	E104946	O15:H18	+	+	+		+					+		
	E104967	O15:H18	+	+	+		+					+		
	E107526	O?:H27		+	+		+			+				
	E099975	O15:H18	+	+	+		+					+		
	E104931	O130:H27	+	+	+		+		+	+				
	E104940	O130:H27	+	+	+		+			+				
	E104969	O130:H27	+	+	+		+			+				
ST40	1091	O111:H21	+	+	+	+	+		+					+
	1337	O111:H21	+	+	+	+	+		+					+
	146052	O?:H21		+	+	+	+		+					
	153145	O111:H21	+	+	+	+	+		+					+
	221654	O111:H21	+	+	+	+	+		+					+
	232736	O111:H21	+	+	+	+	+		+					+
	E107759	O127:H21	+	+	+	+	+		+					+
	E109907	O111:H21	+	+	+	+	+		+					+

Table 2. Detection of selected EAEC virulence genes in whole genome sequences of ST strains by using the Short Read Sequence Typing 2 tool (SRST2). ^aIsolates from cases (bold) and controls, ^bPresence of gene indicated by +, ^cCoverage threshold below 90% identity.

(AAF/V) and ST31 O130 and O15 strains expressing *aggA* (AAF/I) and *agg3A* (AAF/III), respectively (Table 2). While a coverage threshold of 90% was applied for all read alignments, this was reduced to 80% for *aggA*. In contrast to differences between STs, virulence gene profiles were similar in isolates from cases or controls within each ST. Further differences between ST31 and ST40 were identified using the Virulence Factors of Bacterial Pathogens database (VFDB), and genes found in specific STs only are listed in Table 3. While VFDB analysis confirmed the association of AAF/I and AAF/III with ST31 and the prevalence of EAST1 in ST40 strains, other ST40-specific fimbriae and adhesins (*cfa*, *ehaA*, *stg*, *ycb*) were identified which may contribute to the enhanced host cell binding of this ST. In addition to the 16 strains phenotypically characterised in this study, a further 31 ST31 and 8 ST40 clinical EAEC isolates from PHE were sequenced and analysed. Similar to our previous results, O130:H27 was the predominant serotype in ST31 strains (16 out of 31 strains, Table S2). In addition, three strains were identified as O25:H2, while two isolates belonged to serotype O15:H18 identified in ST31 previously. For ST40, half of the strains encoded O111:H21 (Table S2). Analysis of all ST31 and ST40 whole genome sequences for EAEC-specific virulence gene alleles obtained from the VFDB confirmed the prevalence of *astA* in ST40 (88%) versus ST31 strains (26%, Fig. 5). In addition, all but one ST40 strain harboured *ecpA*, while this gene was present in only 49% of ST31 isolates. Interestingly, one particular allele (*ecpA_CVF625_VFG034419*) was detected in 81% of ST40 versus only 5% of ST31 strains. With regards to AAF, *aaf5A* was the prominent variant in ST40 (75%), whereas more than half of ST31 strains did not harbour any AAF type, and one third were positive for *aggA*. Another interesting finding emerging from the VFDB analysis was that although all ST31 and ST40 isolates possessed a gene for haemolysin E (*hlyE*), different alleles dominated in ST31 (VFG036058) and ST40 strains (VFG036061, Fig. 5).

Discussion

Clinical pathology and laboratory research suggest that EAEC pathogenesis involves adherence to the intestinal mucosa, mucus secretion, biofilm formation, cytotoxic damage and mucosal inflammation due to cytokine release²⁵. Importantly, the majority of EAEC *in vitro* and *in vivo* studies have been conducted using prototype strain 042 which we have included as positive control in all phenotypic assays. Although 042 belongs to ST31, it does not share any characteristics with the other ST31 strains investigated in this study. This may be due to differences in origin (Peru versus UK) or changes in genotype/phenotype during prolonged laboratory passage and underlines the heterogeneity of the EAEC pathotype and necessity to open up research to a wider range of EAEC isolates.

Adhesion to the intestinal mucosa is the first step in EAEC colonization of the human gut^{22,26}, and increased adherence of ST40 versus ST31 strains could explain the higher association of this ST with diarrhoeal disease. While AA to cervix-derived HEP-2/HeLa cells is supported by all five variants of AAFs^{8–12}, a role for binding to human intestinal epithelium has only been demonstrated for AAF/I and AAF/II^{8,27} while the relevance of AAF/III–V remains unknown. Genotypic analysis showed that the ST40 strains investigated in our study possessed the fimbrial variant AAF/V whereas ST31 strains harboured the gene for AAF/I or AAF/III. Notably, alignment of *aggA* (encoding the major pilin subunit of AAF/I) to the query sequence of EAEC strain 17-2 showed less than 90% identity suggesting the presence of a different fimbrial variant. However, high variability with up to 83% identity has been previously reported for the major pilin subunit Agg5A¹². Furthermore, the presence of the AAF/I accessory genes *aggB–D* in respective ST31 strains supports the identification of AAF/I. Interestingly, expression of AAF/V in *E. coli* HB101 resulted in higher adhesion to HEP-2 cells compared to expression of AAF/

Gene	Present in all strains of	Function
<i>EC042_4532</i>	ST31	Putative type VI secretion protein
<i>air/eaex</i>	ST31	Enteroggregative immunoglobulin-repeat protein with predicted homology to intimin from enteropathogenic <i>E. coli</i>
<i>aslA</i>	ST31	Putative Ser-type periplasmic non-aryl sulfatase
<i>chuA, S-Y</i>	ST31	Haem utilisation operon
<i>ehaB</i>	ST31	Autotransporter linked to adhesion and biofilm formation
<i>kpsD</i>	ST31	Polysialic acid transport protein
<i>pkgA</i>	ST31	Putative type III secretion effector
<i>sitA-D</i>	ST31	Iron/manganese transport operon
<i>aggB-D</i>	ST31 (O130)	AAF/I accessory genes
<i>gspC-M</i>	ST31 (O130) & ST40	Cryptic type 2 secretion system
<i>agg3A</i>	ST31 (O15:H18)	AAF/III structural subunit
<i>hlyA-D</i>	ST31 (O15:H18)	Alpha-haemolysin
<i>papA-K, X</i>	ST31 (O15:H18)	P pilin operon
<i>agg3 B-D</i>	ST31 (O15:H18) & ST40	AAF/III accessory genes
<i>astA</i>	ST40	EAST-1 toxin
<i>cfaA-D</i>	ST40	CS1 fimbrial operon
<i>ehaA</i>	ST40	Autotransporter linked to adhesion and biofilm formation
<i>stgC,D</i>	ST40	Part of Stg fimbrial operon
<i>ycbE,R-V</i>	ST40	Putative chaperone-usher fimbrial operon

Table 3. Prevalence of bacterial virulence genes in ST31 and ST40 strains as determined from whole genome sequences using Virulence Factor Database.

III¹² suggesting different binding affinities of AAF variants to epithelial cell receptors. Previous studies have identified the mucin glycoprotein MUC1, intermediate filament-associated cytokeratin 8, and the extracellular matrix proteins fibronectin, laminin, and type IV collagen as cellular binding partners for AAF/II^{28–30}. While AAF/I and AAF/II bind to fibronectin by electrostatic interactions, AAF/V does not recognize fibronectin or other extracellular matrix proteins^{31,32}. This might indicate a preference for more accessible host receptors on the apical cell surface (e.g. MUC1) which facilitate initial binding to the intestinal epithelium. Notably, more than half of the ST31 strains sequenced in our study did not harbour any AAF variant, and it remains to be investigated if this contributes to the lower disease potential of this clonal complex. Apart from AAF, other adhesins have been implicated in EAEC - host cell binding. A previous study characterising 130 EAEC strains of diverse origin reported that only 3.1% and 5.4% of isolates were positive for AAF/I and AAF/II genes, respectively, while 96% contained the ECP structural gene *ecpA*, and 63% of those produced ECP when adhering to HEp-2 cells¹⁶. In addition, an *ecpA* deletion mutant in enterohaemorrhagic *E. coli* was deficient in HEp-2 cell binding. Therefore, the prevalence of *ecpA* and particularly, its specific allele VFG034419, in ST40 versus ST31 strains might suggest a role of ECP in adherence to the gut epithelium. In addition to AAF/V and ECP, other ST40-specific adhesins were identified which include CFA/I fimbriae mediating adherence of enterotoxigenic *E. coli* to human intestinal mucosa¹, Stg fimbriae promoting epithelial attachment of *Salmonella* Typhi³³, and the type V secretion system autotransporter EhaA involved in enterohaemorrhagic *E. coli* aggregation and adhesion to bovine rectal epithelial cells³⁴.

In addition to epithelial adherence, EAEC interaction with mucus appears to play an important role in infection as evidenced by the presence of mucus in stools of diarrhoeal patients and the thick bacterial biofilms observed in the intestinal mucus layer⁶. Previous studies have shown that the serine protease Pic exhibits mucinolytic activity and enables EAEC growth in mouse caecal mucus and minimal medium containing bovine submaxillary mucin^{17,21}. This agrees with our results demonstrating that both EAEC ST31 and ST40 harbour *pic* and can use porcine gastric mucin as a nutrient source. In our previous studies, we have demonstrated that the metalloprotease StcE degrades MUC2 and thereby facilitates adherence of enterohaemorrhagic *E. coli* to human LS174T cells and colonic biopsy samples³⁵. However, we did not detect any influence of EAEC infection on MUC2 production by LS174T cells in this study. In addition to its mucinolytic activity, Pic also acts as a secretagogue and stimulates mucus secretion in rat ileal loops¹⁸. As this is mediated by an increase in goblet cell numbers rather than increased mucin expression by existing goblet cells, an *in vivo* system would be required to investigate this effect.

Mucus secretion facilitates biofilm formation on intestinal mucosa which is a characteristic virulence feature of EAEC^{24,36}. Previous studies have shown that AAF/I-V are required for biofilm formation in several EAEC strains including prototype strain 042^{9,12,37}. In addition, there may be accessory factors modulating biofilm composition and density as the gene *shf* has been shown to be required for the establishment of firm biofilms by EAEC 042³⁸. Interestingly, type 1 fimbriae and flagella which have been implicated in biofilm formation by uropathogenic *E. coli* and *E. coli* K-12^{39,40}, did not affect aggregation of EAEC 042³⁷. Our results demonstrated increased

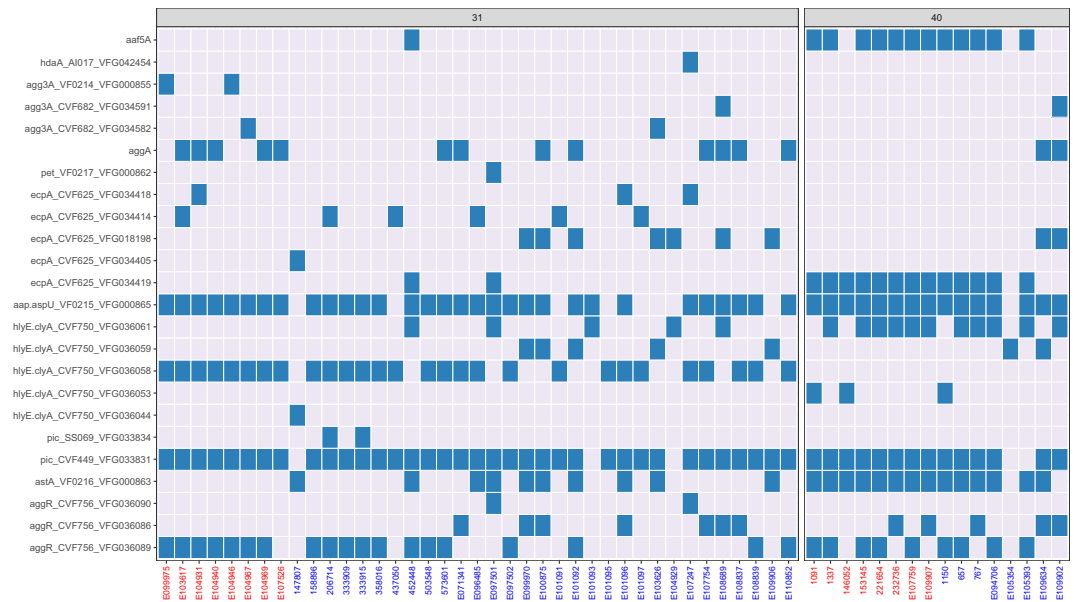


Figure 5. Presence of EAEC virulence-related gene alleles in whole genome sequences of 39 ST31 and 16 ST40 isolates. Strains characterised phenotypically in this study are highlighted in red. Several alleles present in the VFDB were detected for *aggR*, *pic*, *hlyE*, *ecpA*, and *agg3A*, and the presence of each allele is indicated by a blue box. Gene presence was determined by SRST2, except for *aggA* where only a partial read alignment was indicated. In this case, BLAST analysis was used which showed an average nucleotide identity of 86.5%.

biofilm formation by EAEC ST40 versus ST31 strains, although variation within ST groups was more pronounced compared to that observed in adherence assays. Similar to cell adhesion, it could be argued that biofilm formation might be dependent on AAF variant, with AAF/V present in ST40 promoting stronger aggregation than AAF/I or AAF/III in ST31. However, this does not apply to ST31 strain E104931 and ST40 strain 221654 which formed stronger and weaker aggregates than the other ST31 and ST40 isolates, respectively. Therefore, additional factors to AAF must play a role in biofilm formation by these EAEC strains.

Cytotoxicity is another prominent feature of EAEC pathogenesis and has been demonstrated in human intestinal biopsies and T84 cells²⁴. Previous studies in our laboratory have shown that EAEC binding to T84 cells induces gene expression of the toxins EAST-1, HlyE and Pet suggesting a role in EAEC virulence⁴¹. As Pet is absent in all ST31 and ST40 strains characterised in this study, and EAST-1 is expressed in ST40 isolates only, these two toxins are unlikely the cause for the cytotoxic effects observed in T84 monolayers infected with EAEC ST31 and ST40 isolates. In contrast, the pore-forming toxin HlyE is encoded by both ST31 and ST40 strains, although different alleles predominate in each ST, and the contribution of HlyE to epithelial damage warrants further investigation.

Mucosal inflammation also contributes to EAEC pathogenesis, and elevated levels of the pro-inflammatory cytokines IL-1 β and IL-8 have been detected in stool samples from patients with EAEC diarrhoea^{42,43}. Several EAEC proteins have been identified which promote secretion of the neutrophil chemoattractant IL-8 *in vitro*, and varying results have been obtained depending on the cell line used and its differentiation status. While the IL-8 response in non-polarized T84 cells infected with EAEC 042 was mainly dependent on flagellin, AafB (minor pilin protein of AAF/II) also contributed to IL-8 secretion by polarized T84 epithelia⁴⁴. Stimulation of IL-8 secretion by flagella and AafB was further confirmed in HEp-2 and Caco-2 cells^{45,46}. Another study showed that deletion of the gene for AafA (major pilin protein of AAF/II) severely reduced IL-8 driven neutrophil transmigration across 042-infected T84 epithelia, whilst deletion of the genes encoding AafB or flagellin did not have any effect⁴⁷. In addition, expression of AAF/I, AAF/III, or AAF/IV induced neutrophil transmigration by other EAEC prototype strains, and the dependency of EAEC-induced inflammation on AAF was confirmed in human intestinal xenograft mice⁴⁷. Our studies demonstrated higher IL-8 secretion by T84 cells infected with ST40 versus ST31 strains which might be caused by differences in AAF variant and/or flagellin type. While ST40 strains expressed AAF/V and H21, ST31 strains encoded either AAF/I and H27 or AAF/III and H18. Differences in pro-inflammatory responses to AAF variants have been demonstrated in human intestinal biopsies with higher IL-8 secretion in tissues infected with AAF/II- versus AAF/I- or AAF/III-expressing EAEC strains⁴⁸. Furthermore, purified AAF/II induced greater neutrophil transmigration across T84 cells than AAF/I when applied with *E. coli* K-12²⁸. Similarly, uropathogenic *E. coli* expressing H4 flagella induced higher levels of IL-10 secretion compared to isogenic mutants harbouring H1 and H7 alleles⁴⁹.

In addition to differences in ECP, AAF and flagellin alleles, all ST40 isolates in our phenotypic study contained *astA* which was absent in ST31 strains. This gene encodes the low-molecular-weight EAEC heat-stable enterotoxin 1 (EAST1) which shares significant homology with the heat-stable enterotoxin of enterotoxigenic *E. coli*⁵⁰ and stimulates anion secretion by T84 epithelia⁵¹. While *astA* was significantly associated with infant diarrhoea in Brazil versus asymptomatic controls⁵², the gene was commonly detected in both clinical and non-clinical EAEC

strains in China⁵³. Therefore, its relevance as a marker for pathogenic EAEC remains uncertain, and future studies are needed to further establish the role of EAST1 in intestinal disease.

While we established a correlation between certain virulence traits and genes with ST31 and ST40 strains, no difference was observed between isolates from diarrhoea cases and asymptomatic carriers within each group. Notably, only two of the investigated ST40 isolates were derived from healthy controls, so further confirmation is required for this ST. Nevertheless, the lack of correlation between virulence trait and strain origin might be due to differences in host susceptibility which can be influenced by genetic and environmental factors and represent a considerable confounding factor in case-control association studies. Single nucleotide polymorphisms in the promoter region of IL-8 and CD14 and the lactoferrin gene have been linked to increased susceptibility to EAEC infection^{54–56}. Another predisposing factor appears to be co-infection with other enteropathogens which is considerably higher in EAEC-positive cases than in asymptomatic controls⁵⁷.

Taken together, our studies suggest that EAEC ST40 strains have a higher intrinsic potential to cause enteric disease due to their enhanced ability to bind to human colonic epithelium, form biofilms and induce pro-inflammatory IL-8 secretion. Differences in AAF and flagellin alleles and/or the presence of EAST1 and ECP might represent the underlying cause for these phenotypes. Future studies will ascertain the suitability of these genotypic and phenotypic markers for identification of highly pathogenic EAEC which will improve diagnosis of diarrhoea. In addition, our findings in conjunction with previous data⁴⁵ implicate mucosal adherence as an efficient target for future antimicrobial and vaccine strategies against EAEC infection.

Methods

Bacterial culture and identification. All EAEC strains used in this study are listed in Table S1 and S2. EAEC were defined by *aat* gene/CVD432 probe reaction, presence of *aggR* or aggregative adherence phenotype. For all experiments, strains were grown standing overnight in Lysogeny broth (LB Lennox) at 37 °C.

Cell culture

Human epithelial HEP-2 cells (ECACC 86030501) and LS174T colon adenocarcinoma cells (ECACC 87060401) were cultured in Dulbecco's Modified Eagle's Medium (DMEM, Sigma) with 4 mM L-glutamine, 1% non-essential amino acids and 10% foetal bovine serum (FBS, Sigma). T84 colon carcinoma cells (ECACC 88021101) were maintained in DMEM/F-12 1:1 mixture (Sigma) containing 2.5 mM L-glutamine and 10% FBS and used between passage 46 and 63. All cells were grown at 37 °C in a 5% CO₂ atmosphere.

Aggregative adherence to HEP-2 cells. HEP-2 cells were seeded at a density of 1×10^5 cells/well into 24-well plates containing circular coverslips and grown to 60% confluency. Cells were inoculated with 10^7 CFU of EAEC and incubated at 37 °C in air/5% CO₂ for 3 h. Coverslips were subsequently washed thrice with phosphate-buffered saline (PBS), fixed in 70% methanol for 15 min and stained with 10% Giemsa Modified Solution (Sigma) for 30 min. Samples were observed under brightfield with a Zeiss Axio Imager M2 Microscope.

Adherence to T84 cells. T84 cells were seeded at a density of 1.5×10^5 cells/well into 24-well plates and cultured for 6–8 days until fully confluent. Cells were inoculated with 10^7 CFU of EAEC and incubated at 37 °C in air/5% CO₂ for 2 h. After removal of non-adherent bacteria by three washes in PBS, cells were lysed in 1% Triton X-100 in PBS for 10 min. Serial dilutions of cell lysates and bacterial inocula were plated on LB agar plates, and CFU were counted the next day.

Bacterial growth in DMEM/F-12 medium. EAEC overnight cultures adjusted to OD₆₀₀ 1.0 were diluted 1:100 in DMEM/F-12 medium and dispensed into 96-well plates in 100 µl aliquots. After incubation at 37 °C for 2 h, optical density was determined at 600 nm.

In vitro organ culture of human colonic biopsies and scanning electron microscopy. This study was performed with approval from the University of East Anglia Faculty of Medicine and Health Ethics Committee (Ref 2010/11-030). All samples were registered with the Norwich Research Park Biorepository (REC reference 19/BE/0089) and all research was performed in accordance with relevant guidelines and regulations. After obtaining informed consent, biopsy samples from the sigmoid colon were taken from adult patients undergoing routine endoscopy at the Gastroenterology Department of the Norfolk and Norwich University Hospital. Biopsy specimens were orientated with the mucosal surface facing upwards on a foam support and submersed in IVOC medium (DMEM/NCTC-135 medium (1:1), 10% newborn calf serum, 1.1 g/l sodium bicarbonate, 0.5% D-mannose, Sigma). Samples were inoculated with 2.5×10^7 CFU of EAEC or LB medium as negative control. Biopsies were incubated at 37 °C in air/ 5% CO₂ for 7 h on a rocking platform with medium exchange after 4 and 6 h to prevent bacterial overgrowth. At the end of the incubation, biopsies were washed in PBS to remove non-adherent bacteria and fixed with 2.5% glutaraldehyde in PBS at 4 °C overnight. Specimens were subsequently dehydrated in a graded acetone series, dried using tetramethylsilane (Sigma), mounted on aluminium stubs, sputter-coated with gold (Polaron SC7640, Quorum Technologies) and observed using a JSM 4900 LV scanning electron microscope (JEOL). Bacterial colonization was scored on a relative scale based on size of bacterial aggregates (1 = less than 10; 2 = ca 10–100; 3 = ca 100–1,000; 4 = more than 1,000 bacteria) and frequency (1 = less than 10%, 2 = ca 25%, 3 = ca 50%, 4 = more than 50% bacterial coverage of biopsy surface), and both values were added.

Growth in mucus-containing medium. EAEC overnight cultures adjusted to OD₆₀₀ 1.0 were diluted 1:100 in 200 µl M9 medium (6.8 g/l Na₂HPO₄, 3.0 g/l KH₂PO₄, 1.0 g/l NH₄Cl, 0.5 g/l NaCl, 0.5 g/l glucose) with or

without 0.5% (w/v) porcine gastric mucin (Sigma) and added to a 96-well plate. Plates were incubated at 37 °C for 8 h, and bacterial growth was quantified by optical density (OD₆₀₀, Benchmark Plus, Bio-Rad).

Quantification of mucus production in LS174T cells by immunostaining. LS174T cells were seeded at a density of 1.5×10^5 cells/well on circular coverslips placed in 24-well plates and grown for 5 days. Cells were inoculated with 10^7 CFU of EAEC and incubated at 37 °C in air/5% CO₂ for 4 h. Non-adherent bacteria were removed by three washes in PBS, and cells were fixed in ice-cold methanol/ acetone (1:1) for 4 min on ice. For immunofluorescence staining, cells were blocked with 0.5% (w/v) bovine serum albumin in PBS for 20 min followed by incubation in mouse anti-MUC2 (1:250, Santa Cruz) for 60 min and detection by Alexa Flour 488-conjugated donkey anti-mouse IgG (1:400, Life Technologies). EAEC were stained with goat anti-*E. coli* (1:400, Abcam) and Alexa Flour 568-conjugated donkey anti-goat IgG (1:400, Life Technologies). Coverslips were mounted in Vectashield (Vector laboratories) and examined with a Zeiss Axio Imager M2 Microscope. MUC2 staining was quantified by integrated density using ImageJ software.

Biofilm formation. EAEC overnight cultures adjusted to OD₆₀₀ 1.0 were diluted 1:100 in DMEM, and 100 µl of suspension were added to a 96-well plate. After incubation at 37 °C in air/5% CO₂ for 48 h, media were removed, and wells were washed twice with sterile water. Bacterial biofilms at the bottom of the well were stained with 0.1% crystal violet (w/v) for 10 min. After removal of excess stain with two washes in water, plates were dried, and the crystal violet stain was solubilised in 30% acetic acid and quantified by optical density (OD₅₉₅).

Cytotoxicity assay. Confluent T84 cells in 24-well plates were inoculated with 2×10^7 CFU of EAEC and incubated at 37 °C in air/5% CO₂ for 8 h. After removal of the medium, dead cells were stained with 0.05% Trypan Blue for 15 min at 37 °C. Unbound dye was removed by washing with PBS, and monolayer integrity was confirmed by microscopy. Internalised dye was subsequently released by cell lysis in 1% SDS, and absorbance was determined at OD₅₉₀.

Interleukin-8 secretion. T84 cells were seeded out at a density of 4×10^5 cells/well into 12-well plates and grown for 6 days until fully confluent. Cells were infected with 1.5×10^7 CFU and incubated for 3 h at 37 °C in air/5% CO₂. After removal of non-adherent bacteria, fresh medium containing 50 µg/ml gentamicin (Sigma) was added, and cells were incubated for another 20 h. IL-8 in supernatants was determined by ELISA (Peprotech) according to the manufacturer's instructions. In addition, cytotoxicity was quantified by Trypan Blue staining as described above.

Genome sequencing. Extracted bacterial DNA was converted into Nextera XT libraries according to the library preparation guide (Illumina) for sequencing on an Illumina NextSeq platform. Libraries were quantified using HS dsDNA Qubit assay as per the manufacturer's instructions. Libraries were pooled in equimolar amounts, denatured, and sequenced on the Illumina NextSeq 500/550 platform using V2 2×150 bp chemistry with paired-end protocol. Alternatively, sequencing was undertaken using an Illumina HiSeq 2500 system to produce 100 bp paired-end sequence fragments (Illumina, Cambridge, United Kingdom). Genome sequencing was performed at a minimum average depth of $30 \times$. All nucleotide sequences were deposited at DDBJ/EMBL/GenBank under accession number ERA1598958.

Bioinformatics. Bioinformatic analyses were performed using tools accessed via the MRC Cloud Infrastructure for Microbial Bioinformatics (MRC CLIMB) platform, the University of East Anglia Norwich High performance cluster (UEA-HPC) or local platforms as appropriate. FASTQ paired-end reads were quality trimmed using Trimmomatic 0.36, removing adaptors and low quality or N bases from the leading and trailing 30 bp. The read was scanned with a 4-base sliding window and cut when the average quality per base dropped below 15. If the read length post trimming was less than 36 bp, the read and its pair were discarded. MLST determination was achieved using SRST2 and processed locally or on the Galaxy computational biology platform. SRST2 was also used for *E. coli* serotype prediction and virulence factor genotyping. While short read serotyping analysis employed the EcOH database, *Escherichia* virulence gene sequences were obtained from the Virulence Factors of Bacterial Pathogens database (VFDB; <http://www.mgc.ac.cn/cgi-bin/VFs/genus.cgi?Genus=Escherichia>) with sequences added from NCBI GenBank as required and pre-clustered with cd-hit at 90% identity.

Statistics. Statistical analysis was performed using GraphPad Prism software (version 6). A P-value of <0.05 was considered significant.

Data sharing. All data generated or analysed during this study are included in this published article (and its Supplementary Information files). All nucleotide sequences were deposited at DDBJ/EMBL/GenBank under accession number ERA1598958.

Received: 10 June 2019; Accepted: 7 April 2020;

Published online: 04 May 2020

References

- Nataro, J. P. & Kaper, J. B. Diarrheagenic *Escherichia coli*. *Clin. Microbiol. Rev.* **11**, 142–201 (1998).
- Lääveri, T., Vilkinen, K., Pakkanen, S. H., Kirveskari, J. & Kantele, A. A prospective study of travellers' diarrhoea: analysis of pathogen findings by destination in various (sub)tropical regions. *Clin. Microbiol. Infect.* **24**, 908 e909–908 e916, <https://doi.org/10.1016/j.cmi.2017.10.034> (2018).

3. Durrer, P. *et al.* Intestinal infection due to enteroaggregative *Escherichia coli* among human immunodeficiency virus-infected persons. *J. Infect. Dis.* **182**, 1540–1544, <https://doi.org/10.1086/315885> (2000).
4. Bielaszewska, M. *et al.* Characterisation of the *Escherichia coli* strain associated with an outbreak of haemolytic uraemic syndrome in Germany, 2011: a microbiological study. *Lancet Infect. Dis.* **11**, 671–676 (2011).
5. Nataro, J. P. *et al.* Patterns of adherence of diarrheagenic *Escherichia coli* to HEp-2 cells. *Pediatr. Infect. Dis. J.* **6**, 829–831 (1987).
6. Estrada-Garcia, T. & Navarro-Garcia, F. Enteroaggregative *Escherichia coli* pathotype: a genetically heterogeneous emerging foodborne enteropathogen. *FEMS Immunol. Med. Microbiol.* **66**, 281–298, <https://doi.org/10.1111/j.1574-695X.2012.01008.x> (2012).
7. Hebbelstrup Jensen, B., Olsen, K. E., Struve, C., Krogfelt, K. A. & Petersen, A. M. Epidemiology and clinical manifestations of enteroaggregative *Escherichia coli*. *Clin. Microbiol. Rev.* **27**, 614–630, <https://doi.org/10.1128/cmr.00112-13> (2014).
8. Czeczulin, J. R. *et al.* Aggregative adherence fimbriae II, a second fimbrial antigen mediating aggregative adherence in enteroaggregative *Escherichia coli*. *Infect. Immun.* **65**, 4135–4145 (1997).
9. Boisen, N., Struve, C., Scheutz, F., Krogfelt, K. A. & Nataro, J. P. New adhesin of enteroaggregative *Escherichia coli* related to the Afa/Dr/AAF family. *Infect. Immun.* **76**, 3281–3292, <https://doi.org/10.1128/iai.01646-07> (2008).
10. Nataro, J. P. *et al.* Aggregative adherence fimbriae I of enteroaggregative *Escherichia coli* mediate adherence to HEp-2 cells and hemagglutination of human erythrocytes. *Infect. Immun.* **60**, 2297–2304 (1992).
11. Bernier, C., Gounon, P. & Le Bouguenec, C. Identification of an aggregative adhesion fimbria (AAF) type III-encoding operon in enteroaggregative *Escherichia coli* as a sensitive probe for detecting the AAF-encoding operon family. *Infect. Immun.* **70**, 4302–4311 (2002).
12. Jönsson, R. *et al.* A novel Aggregative Adherence Fimbriae (AAF/V) of enteroaggregative *Escherichia coli*. *Infect. Immun.* **83**, 1396–1405, <https://doi.org/10.1128/iai.02820-14> (2015).
13. Morin, N., Santiago, A. E., Ernst, R. K., Guillot, S. J. & Nataro, J. P. Characterization of the AggR regulon in enteroaggregative *Escherichia coli*. *Infect. Immun.* **81**, 122–132, <https://doi.org/10.1128/IAI.00676-12> (2013).
14. Nataro, J. P., Yikang, D., Yingkang, D. & Walker, K. AggR, a transcriptional activator of aggregative adherence fimbria I expression in enteroaggregative *Escherichia coli*. *J. Bacteriol.* **176**, 4691–4699 (1994).
15. Sheikh, J. *et al.* A novel dispersin protein in enteroaggregative *Escherichia coli*. *J. Clin. Invest.* **110**, 1329–1337 (2002).
16. Avelino, F. *et al.* The majority of enteroaggregative *Escherichia coli* strains produce the *E. coli* common pilus when adhering to cultured epithelial cells. *Int. J. Med. Microbiol.* **300**, 440–448, <https://doi.org/10.1016/j.ijmm.2010.02.002> (2010).
17. Henderson, I. R., Czeczulin, J., Eslava, C., Noriega, F. & Nataro, J. P. Characterization of Pic, a secreted protease of *Shigella flexneri* and enteroaggregative *Escherichia coli*. *Infect. Immun.* **67**, 5587–5596 (1999).
18. Navarro-Garcia, F. *et al.* Pic, an Autotransporter Protein Secreted by Different Pathogens in the Enterobacteriaceae Family, Is a Novel Mucus Secretagogue. *Infect. Immun.* **78**, 4101–4109, <https://doi.org/10.1128/IAI.00523-10> (2010).
19. Jenkins, C., Chart, H., Willshaw, G. A., Cheasty, T. & Tompkins, D. S. Association of putative pathogenicity genes with adherence characteristics and fimbrial genotypes in typical enteroaggregative *Escherichia coli* from patients with and without diarrhoea in the United Kingdom. *Eur. J. Clin. Microbiol. Infect. Dis.* **26**, 901–906, <https://doi.org/10.1007/s10096-007-0388-z> (2007).
20. Chattaway, M. A. *et al.* Enteroaggregative *Escherichia coli* have evolved independently as distinct complexes within the *E. coli* population with varying ability to cause disease. *Plos One* **9**, e112967, <https://doi.org/10.1371/journal.pone.0112967> (2014).
21. Harrington, S. M. *et al.* The Pic protease of enteroaggregative *Escherichia coli* promotes intestinal colonization and growth in the presence of mucin. *Infect. Immun.* **77**, 2465–2473, <https://doi.org/10.1128/iai.01494-08> (2009).
22. Hicks, S., Candy, D. C. & Phillips, A. D. Adhesion of enteroaggregative *Escherichia coli* to pediatric intestinal mucosa *in vitro*. *Infect. Immun.* **64**, 4751–4760 (1996).
23. Hall, C. W. & Mah, T. F. Molecular mechanisms of biofilm-based antibiotic resistance and tolerance in pathogenic bacteria. *FEMS Microbiol. Rev.* **41**, 276–301, <https://doi.org/10.1093/femsre/fux010> (2017).
24. Nataro, J. P., Hicks, S., Phillips, A. D., Vial, P. A. & Sears, C. L. T84 cells in culture as a model for enteroaggregative *Escherichia coli* pathogenesis. *Infect. Immun.* **64**, 4761–4768 (1996).
25. Huang, D. B. & Dupont, H. L. Enteroaggregative *Escherichia coli*: an emerging pathogen in children. *Semin. Pediatr. Infect. Dis.* **15**, 266–271 (2004).
26. Knutton, S. *et al.* Ability of enteroaggregative *Escherichia coli* strains to adhere *in vitro* to human intestinal mucosa. *Infect. Immun.* **60**, 2083–2091 (1992).
27. Boisen, N. *et al.* The presence of the pAA plasmid in the German O104:H4 Shiga toxin (Stx) type 2a-producing enteroaggregative *Escherichia coli* strain promotes the translocation of Stx2a across an epithelial cell monolayer. *J. Infect. Dis.*, <https://doi.org/10.1093/infdis/jiu399> (2014).
28. Boll, E. J. *et al.* Enteroaggregative *Escherichia coli* adherence fimbriae drive inflammatory cell recruitment via interactions with epithelial MUC1. *mBio* **8**, e00717–00717, <https://doi.org/10.1128/mBio.00717-17> (2017).
29. Izquierdo, M. *et al.* Identification of cell surface-exposed proteins involved in the fimbria-mediated adherence of enteroaggregative *Escherichia coli* to intestinal cells. *Infect. Immun.* **82**, 1719–1724, <https://doi.org/10.1128/iai.01651-13> (2014).
30. Farfan, M. J., Inman, K. G. & Nataro, J. P. The major pilin subunit of the AAF/II fimbriae from enteroaggregative *Escherichia coli* mediates binding to extracellular matrix proteins. *Infect. Immun.* **76**, 4378–4384, <https://doi.org/10.1128/iai.00439-08> (2008).
31. Berry, A. A. *et al.* Structural insight into host recognition by aggregative adherence fimbriae of enteroaggregative *Escherichia coli*. *Plos Pathog.* **10**, e1004404, <https://doi.org/10.1371/journal.ppat.1004404> (2014).
32. Jönsson, R. *et al.* Structural and functional studies of *Escherichia coli* aggregative adherence fimbriae (AAF/V) reveal a deficiency in extracellular matrix binding. *Biochim. Biophys. Acta* **1865**, 304–311, <https://doi.org/10.1016/j.bbapap.2016.11.017> (2017).
33. Forest, C. *et al.* Contribution of the *stg* fimbrial operon of *Salmonella enterica* serovar Typhi during interaction with human cells. *Infect. Immun.* **75**, 5264–5271, <https://doi.org/10.1128/iai.00674-07> (2007).
34. Wells, T. J. *et al.* EhaA is a novel autotransporter protein of enterohemorrhagic *Escherichia coli* O157:H7 that contributes to adhesion and biofilm formation. *Environ. Microbiol.* **10**, 589–604, <https://doi.org/10.1111/j.1462-2920.2007.01479.x> (2008).
35. Hews, C. L. *et al.* The StcE metalloprotease of enterohaemorrhagic *Escherichia coli* reduces the inner mucus layer and promotes adherence to human colonic epithelium *ex vivo*. *Cell. Microbiol.* **19**, <https://doi.org/10.1111/cmi.12717> (2017).
36. Tzipori, S. *et al.* Studies with enteroaggregative *Escherichia coli* in the gnotobiotic piglet gastroenteritis model. *Infect. Immun.* **60**, 5302–5306 (1992).
37. Sheikh, J., Hicks, S., Dall'Agnol, M., Phillips, A. D. & Nataro, J. P. Roles for Fis and YafK in biofilm formation by enteroaggregative *Escherichia coli*. *Mol. Microbiol.* **41**, 983–997 (2001).
38. Fujiyama, R. *et al.* The *shf* gene of a *Shigella flexneri* homologue on the virulent plasmid pAA2 of enteroaggregative *Escherichia coli* O42 is required for firm biofilm formation. *Curr. Microbiol.* **56**, 474–480, <https://doi.org/10.1007/s00284-008-9115-y> (2008).
39. Pratt, L. A. & Kolter, R. Genetic analysis of *Escherichia coli* biofilm formation: roles of flagella, motility, chemotaxis and type I pili. *Mol. Microbiol.* **30**, 285–293 (1998).
40. Hung, C. *et al.* *Escherichia coli* biofilms have an organized and complex extracellular matrix structure. *mBio* **4**, e00645–00613, <https://doi.org/10.1128/mBio.00645-13> (2013).
41. Ellis, S. J., Yasir, M., Browning, D. F., Busby, S. J. W. & Schuller, S. Oxygen and contact with human intestinal epithelium independently stimulate virulence gene expression in enteroaggregative *Escherichia coli*. *Cell. Microbiol.* **21**, e13012, <https://doi.org/10.1111/cmi.13012> (2019).

42. Steiner, T. S., Lima, A. A., Nataro, J. P. & Guerrant, R. L. Enteroaggregative *Escherichia coli* produce intestinal inflammation and growth impairment and cause interleukin-8 release from intestinal epithelial cells. *J. Infect. Dis.* **177**, 88–96 (1998).
43. Greenberg, D. E., Jiang, Z. D., Steffen, R., Verenker, M. P. & DuPont, H. L. Markers of inflammation in bacterial diarrhea among travelers, with a focus on enteroaggregative *Escherichia coli* pathogenicity. *J. Infect. Dis.* **185**, 944–949, <https://doi.org/10.1086/339617> (2002).
44. Harrington, S. M., Strauman, M. C., Abe, C. M. & Nataro, J. P. Aggregative adherence fimbriae contribute to the inflammatory response of epithelial cells infected with enteroaggregative *Escherichia coli*. *Cell. Microbiol.* **7**, 1565–1578, <https://doi.org/10.1111/j.1462-5822.2005.00588.x> (2005).
45. Edwards, L. A., Bajaj-Elliott, M., Klein, N. J., Murch, S. H. & Phillips, A. D. Bacterial-epithelial contact is a key determinant of host innate immune responses to enteropathogenic and enteroaggregative *Escherichia coli*. *Plos One* **6**, e27030, <https://doi.org/10.1371/journal.pone.0027030> (2011).
46. Steiner, T. S., Nataro, J. P., Poteet-Smith, C. E., Smith, J. A. & Guerrant, R. L. Enteroaggregative *Escherichia coli* expresses a novel flagellin that causes IL-8 release from intestinal epithelial cells. *J. Clin. Invest.* **105**, 1769–1777 (2000).
47. Boll, E. J. *et al.* The fimbriae of enteroaggregative *Escherichia coli* induce epithelial inflammation *in vitro* and in a human intestinal xenograft model. *J. Infect. Dis.* **206**, 714–722, <https://doi.org/10.1093/infdis/jis417> (2012).
48. Gupta, D., Sharma, M., Sarkar, S., Thapa, B. R. & Chakraborti, A. Virulence determinants in enteroaggregative *Escherichia coli* from North India and their interaction in *in vitro* organ culture system. *FEMS Microbiol. Lett.* **363**, fnw189, <https://doi.org/10.1093/femsle/fnw189> (2016).
49. Kakkanat, A. *et al.* The role of H4 flagella in *Escherichia coli* ST131 virulence. *Sci. Rep.* **5**, 16149, <https://doi.org/10.1038/srep16149>, <https://www.nature.com/articles/srep16149#supplementary-information> (2015).
50. Savarino, S. J. *et al.* Enteroaggregative *Escherichia coli* heat-stable enterotoxin 1 represents another subfamily of *E. coli* heat-stable toxin. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 3093–3097 (1993).
51. Veilleux, S., Holt, N., Schultz, B. D. & Dubreuil, J. D. *Escherichia coli* EAST1 toxin toxicity of variants 17-2 and O42. *Comp. Immunol., Microbiol. Infect. Dis.* **31**, 567–578, <https://doi.org/10.1016/j.cimid.2007.10.003> (2008).
52. Zamboni, A., Fabbri, S. H., Fagundes-Neto, U. & Scaletsky, I. C. A. Enteroaggregative *Escherichia coli* virulence factors are found to be associated with infantile diarrhea in Brazil. *J. Clin. Microbiol.* **42**, 1058–1063, <https://doi.org/10.1128/JCM.42.3.1058-1063.2004> (2004).
53. Zhang, R. *et al.* Comparative genetic characterization of enteroaggregative *Escherichia coli* strains recovered from clinical and non-clinical settings. *Sci. Rep.* **6**, 24321, <https://doi.org/10.1038/srep24321> (2016).
54. Jiang, Z. D. *et al.* Genetic susceptibility to enteroaggregative *Escherichia coli* diarrhea: polymorphism in the interleukin-8 promoter region. *J. Infect. Dis.* **188**, 506–511, <https://doi.org/10.1086/377102> (2003).
55. Mohamed, J. A. *et al.* Single nucleotide polymorphisms in the promoter of the gene encoding the lipopolysaccharide receptor CD14 are associated with bacterial diarrhea in US and Canadian travelers to Mexico. *Clin. Infect. Dis.* **52**, 1332–1341, <https://doi.org/10.1093/cid/cir228> (2011).
56. Mohamed, J. A. *et al.* A novel single-nucleotide polymorphism in the lactoferrin gene is associated with susceptibility to diarrhea in North American travelers to Mexico. *Clin. Infect. Dis.* **44**, 945–952, <https://doi.org/10.1086/512199> (2007).
57. Chattaway, M. A. *et al.* Investigating the link between the presence of enteroaggregative *Escherichia coli* and infectious intestinal disease in the United Kingdom, 1993 to 1996 and 2008 to 2009. *Euro Surveill.* **18** (2013).

Acknowledgements

We thank Claire Jenkins for providing EAEC strains and Marlene Smith for collecting colonic biopsy samples. We are grateful to Oliver Charity and Rob Kingsley for help with bioinformatic analysis and Bertrand Lézé for support with electron microscopy. This work was supported by the BBSRC (Doctoral Training studentship BB/J014524/1 to S.E. and BB/M011216/1 to C.M., Quadram Institute Bioscience Strategic Programme grant BB/R012504/1 to J.W.), MRC (New Investigator Research Grant MR/J002062/1 to S.S.), the Gastrointestinal Bacteria Reference unit (PHE), and the European Commission (Erasmus+ studentship to J.H.).

Author contributions

S.S. and J.W. designed the project. S.E., C.M., J.H. and L.B. performed the laboratory work. M.C. selected and provided bacterial strains and serotyping data. B.B. provided endoscopic biopsy samples. G.K. carried out whole genome sequencing. L.C. and S.E. performed bioinformatic analyses. S.S. and S.E. drafted the manuscript, and all authors read and helped edit the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-64424-3>.

Correspondence and requests for materials should be addressed to S.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020