

2015

# Conservation and immunogenicity of novel antigens in diverse isolates of enterotoxigenic *Escherichia coli*

Qingwei Luo

*Washington University School of Medicine in St. Louis*

Firdausi Qadri

*International Centre for Diarrhoeal Disease Research, Bangladesh*

Rita Kansal

*University of Tennessee Health Science Center*

David A. Rasko

*University of Maryland at Baltimore*

Alaullah Sheikh

*Washington University School of Medicine in St. Louis*

*See next page for additional authors*

Follow this and additional works at: [http://digitalcommons.wustl.edu/open\\_access\\_pubs](http://digitalcommons.wustl.edu/open_access_pubs)

---

## Recommended Citation

Luo, Qingwei; Qadri, Firdausi; Kansal, Rita; Rasko, David A.; Sheikh, Alaullah; and Fleckenstein, James M., "Conservation and immunogenicity of novel antigens in diverse isolates of enterotoxigenic *Escherichia coli*." *PLoS Neglected Tropical Diseases*.9,1. e0003446. (2015).

[http://digitalcommons.wustl.edu/open\\_access\\_pubs/4622](http://digitalcommons.wustl.edu/open_access_pubs/4622)

---

**Authors**

Qingwei Luo, Firdausi Qadri, Rita Kansal, David A. Rasko, Alaullah Sheikh, and James M. Fleckenstein

RESEARCH ARTICLE

# Conservation and Immunogenicity of Novel Antigens in Diverse Isolates of Enterotoxigenic *Escherichia coli*

Qingwei Luo<sup>1</sup>, Firdausi Qadri<sup>2</sup>, Rita Kansal<sup>3</sup>, David A. Rasko<sup>4</sup>, Alaullah Sheikh<sup>5</sup>, James M. Fleckenstein<sup>1,5,6\*</sup>

**1** Department of Medicine, Division of Infectious Diseases, Washington University School of Medicine, St. Louis, Missouri, United States of America, **2** International Centre for Diarrhoeal Disease Research, Bangladesh, Mohakhali, Dhaka, Bangladesh, **3** Department of Microbiology, Immunology and Biochemistry, University of Tennessee Health Science Center, Memphis, Tennessee, United States of America, **4** Institute for Genome Sciences, Department of Microbiology and Immunology, University of Maryland School of Medicine, Baltimore, Maryland, United States of America, **5** Molecular Microbiology and Microbial Pathogenesis Program, Division of Biology and Biomedical Sciences, Washington University School of Medicine, St. Louis, Missouri, United States of America, **6** Medicine Service, Veterans Affairs Medical Center, St. Louis, Missouri, United States of America

\* [jflecken@wustl.edu](mailto:jflecken@wustl.edu)



**OPEN ACCESS**

**Citation:** Luo Q, Qadri F, Kansal R, Rasko DA, Sheikh A, Fleckenstein JM (2015) Conservation and Immunogenicity of Novel Antigens in Diverse Isolates of Enterotoxigenic *Escherichia coli*. PLoS Negl Trop Dis 9(1): e0003446. doi:10.1371/journal.pntd.0003446

**Editor:** Joseph M. Vinetz, University of California San Diego School of Medicine, UNITED STATES

**Received:** August 8, 2014

**Accepted:** November 26, 2014

**Published:** January 28, 2015

**Copyright:** This is an open access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the [Creative Commons CC0](https://creativecommons.org/licenses/by/4.0/) public domain dedication.

**Data Availability Statement:** Data for this manuscript can be found in the manuscript itself, in Supporting information or in online public repository at NCBI. Individual URLs for data are provided in the body of the manuscript.

**Funding:** These studies were funded in part by grant R01AI89894, and contract HHSN272200900009C from the National Institutes of Health/National Institute of Allergy and Infectious Diseases (<http://www.niaid.nih.gov/>); the Department of Veterans Affairs, Office of Research and Development (<http://www.research.va.gov/>); and the Bill & Melinda Gates Foundation

## Abstract

### Background

Enterotoxigenic *Escherichia coli* (ETEC) are common causes of diarrheal morbidity and mortality in developing countries for which there is currently no vaccine. Heterogeneity in classical ETEC antigens known as colonization factors (CFs) and poor efficacy of toxoid-based approaches to date have impeded development of a broadly protective ETEC vaccine, prompting searches for novel molecular targets.

### Methodology

Using a variety of molecular methods, we examined a large collection of ETEC isolates for production of two secreted plasmid-encoded pathotype-specific antigens, the EtpA extracellular adhesin, and EatA, a mucin-degrading serine protease; and two chromosomally-encoded molecules, the YghJ metalloprotease and the EaeH adhesin, that are not specific to the ETEC pathovar, but which have been implicated in ETEC pathogenesis. ELISA assays were also performed on control and convalescent sera to characterize the immune response to these antigens. Finally, mice were immunized with recombinant EtpA (rEtpA), and a protease deficient version of the secreted EatA passenger domain (rEatA<sub>PH134R</sub>) to examine the feasibility of combining these molecules in a subunit vaccine approach.

### Principal Findings

EtpA and EatA were secreted by more than half of all ETEC, distributed over diverse phylogenetic lineages belonging to multiple CF groups, and exhibited surprisingly little sequence variation. Both chromosomally-encoded molecules were also identified in a wide

(OPP1099494). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of any of the sponsoring agencies.

**Competing Interests:** The authors have declared that no competing interests exist.

variety of ETEC strains and YghJ was secreted by 89% of isolates. Antibodies against both the ETEC pathovar-specific and conserved *E. coli* antigens were present in significantly higher titers in convalescent samples from subjects with ETEC infection than controls suggesting that each of these antigens is produced and recognized during infection. Finally, co-immunization of mice with rEtpA and rEatAp<sub>H134R</sub> offered significant protection against ETEC infection.

## Conclusions

Collectively, these data suggest that novel antigens could significantly complement current approaches and foster improved strategies for development of broadly protective ETEC vaccines.

## Author Summary

Infectious diarrhea is one of the leading causes of death among young children in developing countries, and a major cause of morbidity in all age groups. The enterotoxigenic *Escherichia coli* contribute substantially to this burden of diarrheal illness, and have been a focus of vaccine development efforts for more than forty years following their discovery as a cause of severe diarrheal illness. The heat-labile, and/or heat stable enterotoxins that define ETEC are produced by a diverse population of *Escherichia coli*. This inherent genetic plasticity of *E. coli* has made it difficult to identify antigens specific to ETEC that are highly conserved. Therefore, identification of protective antigens shared by many ETEC strains will likely play an essential role in development of the next iteration of vaccines.

## Introduction

The enterotoxigenic *Escherichia coli* (ETEC) are among the most common causes of infectious diarrhea worldwide. Importantly, ETEC are disproportionately represented in cases of severe diarrheal illness as well as in deaths due to diarrhea among young children in developing countries [1].

These pathogens cause diarrhea by the elaboration and effective delivery of heat-labile and/or heat-stable enterotoxins to intestinal epithelial cells where they stimulate production of cyclic nucleotides ultimately activating the cystic fibrosis transmembrane regulator (CFTR) with resulting net efflux of fluid into the intestinal lumen [2]. Plasmid-encoded colonization factors (CFs), discovered [3] shortly after these organisms were identified as a causative agent of cholera-like diarrheal illness [4–6], are thought to be essential for effective colonization of the small intestine and required for ETEC pathogenesis.

Following early studies suggesting a pivotal role for these structures [7,8], CF antigens have defined the basis for most subsequent ETEC vaccine efforts [9,10]. However, one factor complicating development of a broadly protective vaccine for ETEC has been the general plasticity of *E. coli* genomes [11], and the significant antigenic heterogeneity of the CFs. To date, at least 26 antigenically distinct CF antigens have been described [12]. The lack of appreciable cross-protection afforded by these antigens combined with the complex landscape of CFs portrayed in ETEC molecular epidemiology studies continue to complicate rational CF antigen selection [13].

Antigenic heterogeneity, recent failure of LT-toxoid-based vaccine strategies [14,15], as well as the need to optimize the performance of live-attenuated vaccines currently in clinical trials [16–18] have highlighted the need to identify additional virulence molecules that might be targeted in ETEC vaccines. Recent efforts led to the identification of two loci discovered on the same virulence plasmid of ETEC strain H10407 that encodes the CFA/I colonization factor. These include the *etpBAC* two partner secretion locus responsible for production and export of EtpA [19], a novel adhesin molecule which bridges highly conserved regions of flagellin and the eukaryotic cell surface [20]. Also located on this plasmid is the *eatA* gene that encodes the EatA serine protease autotransporter molecule [21] capable of degrading EtpA [22] as well as MUC2 [23], the major gel-forming soluble mucin in the small intestine [24].

Recent immunoproteomic [25] and transcriptomic [26] analyses of H10407 have also highlighted two chromosomally encoded antigens that are not specific to the ETEC pathovar, but which nonetheless appear to be involved in the pathogenesis of these organisms. Conceivably, these molecules, Ygh [27], a secreted mucin-degrading metalloprotease, and EaeH [28], an adhesin, act in concert with colonization factors and other pathovar-specific virulence proteins like EatA and EtpA to promote toxin delivery.

While emerging data suggests that these novel proteins are highly immunogenic [25] and that EtpA and EatA are protective antigens [29–31] in a murine model of ETEC infection, additional data regarding their conservation among ETEC strains are needed to determine their suitability as vaccine targets. Here we demonstrate that these antigens are broadly represented in a diverse collection of ETEC isolates suggesting that they could be employed to augment existing approaches to ETEC vaccine development.

## Materials and Methods

### Bacterial strains and growth conditions

ETEC strains used in this study are detailed in [S1 Dataset](#). All strains were grown at 37° in Casamino acids yeast extract media [32] (CAYE: 2.0% Casamino Acids, 0.15% yeast extract, 0.25% NaCl, 0.871% K<sub>2</sub>HPO<sub>4</sub>, 0.25% glucose, and 0.1% (v/v) trace salts solution consisting of 5% MgSO<sub>4</sub>, 0.5% MnCl<sub>2</sub>, 0.5% FeCl<sub>3</sub>) from frozen glycerol stocks maintained at –80°C.

### Strain characterization by disease severity and colonization factor type

Strains from the International Centre for Diarrhoeal Disease Research (icddr,b) in Dhaka were selected based on their associated disease severity using modified WHO guidelines as previously outlined [33]. Expression of individual CFs was determined by dot immunoblotting with monoclonal antibodies specific to each respective CFs (CF-MAb) as previously described [34]. Briefly, 2 µl of a PBS suspension containing ~10<sup>6</sup> colony forming units of each ETEC strain was dotted onto nitrocellulose, air-dried, blocked with BSA in PBS, followed by detection with CF-MABs and goat anti-mouse IgG\_HRP conjugate. Bound MABs were then detected with 4-chloro-1-naphthol chromogen and H<sub>2</sub>O<sub>2</sub>.

### Screening for ETEC virulence genes by PCR

We screened a total of 181 ETEC available isolates currently maintained as frozen glycerol stocks in our laboratories. The majority of these strains were collected between 1998 and 2011 in Bangladesh, and were obtained from the icddr,b in Dhaka. Complementing this collection were geographically disparate strains associated with severe diarrheal illness including strains from the Amazon region in Brazil [35], and ThroopD, an isolate from a patient with severe ETEC diarrheal illness who presented in Dallas in the 1970s [36]. Strains encoding *eatA* and

*etpA* were identified by PCR using primers directed against conserved regions of these genes as previously described [37]. Briefly, a small amount of frozen glycerol stock from each strain was introduced with a sterile pipette tip into a PCR mixture containing the respective primers and a master mix. Toxin genotypes were confirmed in these isolates using multiplex PCR screening for genes encoding heat-labile (LT), and heat-stable toxins (STp, and STh) as previously described [34]. Primer sequences are listed in [S1 Table](#).

### Immunoblotting for secreted ETEC virulence antigens

To determine production of secreted virulence antigens by different ETEC strains, supernatants from overnight cultures were first precipitated with trichloroacetic acid (TCA) [19] and resuspended in sample buffer before polyacrylamide gel electrophoresis. Western blotting was then performed using polyclonal rabbit antisera against recombinant versions of either EatA [21], EtpA [19], or YghJ [27] that were pre-absorbed against an *E. coli* lysate column (Pierce) and affinity-purified using the antigen immobilized on nitrocellulose membranes as previously described [31,38], followed by detection with affinity-purified secondary goat anti-rabbit-(IgG)-HRP conjugate (Santa Cruz Biotechnology, SC2004).

### Protein sequence comparisons of ETEC pathovar specific antigens

To examine antigenic conservation of EatA among ETEC isolates for which genomic DNA sequences are currently available, BLASTP [39] was used to search GenBank <https://www.ncbi.nlm.nih.gov/genbank/> using the full length sequence of the EatA protein from strain H10407 (<https://www.ncbi.nlm.nih.gov/protein/AAO17297.1>) as the query sequence. To construct alignments of EatA from positive strains, the 1042 residue passenger domain (corresponding to amino acids 57–1098 of EatA from H10407) was compared with EatA of ETEC isolates derived from different phylogenetic lineages using a CLUSTAL Omega (release 1.2.0 AndreaGiacomo) [40] algorithm plugin for CLC Main Workbench v6.9.1. A similar approach was used to compare the amino-terminal sequence of EtpA (amino acids 1–600, GenBank accession number AAX13509.2).

### Conservation heat mapping

Virulence protein expression data from the collection of 181 strains under study were included in the analysis. Heat maps were configured using R [41] version 3.1.0 (2014, <http://www.R-project.org/>) using `gplots` [42] and `RColorBrewer` [43] packages installed from <http://CRAN.R-project.org> using the `heatmap2` function within `gplots` (see [S2 Dataset](#)).

### Recombinant protein production

The antigens used in these studies were produced as polyhistidine-tagged recombinant proteins and purified by immobilized metal ion affinity chromatography (IMAC) as previously described [27,29,44,45]. Additional polishing steps including size exclusion or ion exchange chromatography were performed as needed to produce highly purified antigens. Purity of each antigen was assessed by SDS-PAGE followed by sensitive Coomassie Blue staining. Purified recombinant antigens were stored at  $-80^{\circ}\text{C}$ .

### Assessment of immune responses to novel ETEC virulence proteins

To quantify antibody concentrations directed at novel recombinant antigens, kinetic ELISA was performed on dilutions of plasma samples previously obtained from patients hospitalized at the International Centre for Diarrhoeal Disease Research in Dhaka, Bangladesh (icddr,b)

with acute symptomatic ETEC infections. Plasma samples from non-infected adults and children obtained at icddr,b, or specimens obtained from children at Saint Louis Children's Hospital were used as negative controls. Samples from human volunteer ETEC H10407 challenge studies were kindly provided by Dr. Robert Gormely and Dr. Stephen Savarino of National Naval Medical Center, Bethesda Maryland.

Use of these clinical materials was approved by the Institutional Review Boards of both icddr,b and Washington University School of Medicine. All plasma samples were maintained at 4°C in a humidified chamber prior to use in ELISA. Immune responses to purified recombinant proteins (rYghJ, rEaeH, rEtpA, rEatA<sub>p</sub>) were assessed by kinetic ELISA [46] as previously described [30,47]. Antigen binding to ELISA wells (Corning, Costar 2580) was first optimized to determine the optimal coating concentration and buffer system, using highly antigen-specific polyclonal rabbit antisera to detect binding by ELISA. Purified antigens were then diluted either in 50 mM carbonate buffer (pH 9.6) (rEtpA-myc-His<sub>6</sub>, 1 µg/ml; rEatA<sub>p</sub>, 10 µg/ml; rYghJ-myc-His<sub>6</sub>, 1 µg/ml); or in phosphate buffered saline (PBS, pH 7.4) (rEaeH-myc-His<sub>6</sub>, 1 µg/ml). ELISA plate wells were coated with 100 µl/well overnight at 4°C, washed with PBS containing 0.05% Tween-20 (PBS-T), and blocked for 1 h at 37°C with 1% BSA in PBS-T. All plasma samples were diluted at 1:4096 in blocking buffer. After incubation for 1 hour at 37°C, plates were washed with PBS-T, and secondary goat anti-human IgG(H+L)-HRP conjugated antibody (Pierce, 31410) was added at a final concentration of 1:10,000. After incubation for 30 minutes at 37°C, plates were washed and developed with TMB microwell peroxidase substrate [3,3',5,5'-Tetramethylbenzidine] (KPL, 50-76-00). Kinetic absorbance measurements were determined at a wavelength of 650 nm, and acquired at 40 s intervals for 20 minutes using a microplate spectrophotometer (Eon, BioTek). All data were recorded and analyzed using Gen5 software (BioTek) and reported as the Vmax expressed as milliunits/min. Statistical calculations were performed using Prism v4.0c (GraphPad Software), using nonparametric Mann-Whitney (two-tailed) comparisons of data.

## Mouse immunization and challenge studies

These studies were performed in strict accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health, using an established protocol approved by the Washington University School of Medicine Animal Studies Committee.

Four groups of twelve CD-1 mice were immunized intranasally with either 1 µg of LT (adjuvant only controls), or 1 µg of LT + 15 µg of rEatA<sub>p</sub>(H134R), or 1 µg of LT + 15 µg of rEtpA, or 1 µg of LT + 15 µg of rEatA(H134R)+15 µg of rEtpA on days 0, 14, 28. On day 40, mice were treated with streptomycin [5 g per liter] in drinking water for 24 hours, followed by drinking water alone for 18 hours. After administration of famotidine to reduce gastric acidity, mice were challenged with 10<sup>6</sup> cfu of the kanamycin-resistant (*lacZYA::KmR*) strain jf876 [48] by oral gavage as previously described [47]. Fecal samples (6 pellets/mouse) were collected on day 42 before oral gavage, re-suspended in buffer (10mM Tris, 100mM NaCl, 0.05% Tween 20, 5mM Sodium Azide, pH 7.4) overnight at 4°C, centrifuged to pellet insoluble material, and recover supernatant for fecal antibody testing (below). Twenty-four hours after infection, mice were sacrificed, sera were collected, and dilutions of saponin small-intestinal lysates were plated onto Luria agar plates containing kanamycin (50 µg/ml).

Murine immune responses to LT, EatA and EtpA were determined using previously described kinetic ELISA. Briefly, ELISA wells were coated with 1 µg/ml GM1, or 10 µg/ml of rEatA<sub>p</sub>(H134R), or 1 µg/ml rEtpA in carbonate buffer (15 mM Na<sub>2</sub>CO<sub>3</sub>, 35 mM NaHCO<sub>3</sub>, 0.2 g/L NaN<sub>3</sub>, pH8.6) overnight at 4°C. Wells were washed three times with



phosphate-buffered saline containing 0.05% Tween 20 (PBS-T), blocked with 1% bovine serum albumin (BSA) in PBS-T for 1 h at 37°C, and 100  $\mu$ l of fecal suspensions (undiluted) or sera (diluted 1:100 in PBS-T with 1% BSA) was added per ELISA well and incubated at 37°C for 1 h. Horseradish peroxidase-conjugated secondary antibodies were used and signal detected with TMB (3,3',5,5'-tetramethylbenzidine)-peroxidase substrate (KPL) substrate.

## Ethics statement

All animal studies were performed under protocols approved by the Animal Studies Committee of Washington University School of Medicine (protocol number 20110246A1). All procedures complied with Public Health Service guidelines, and The Guide for the [Care and Use of Laboratory Animals](#).

All human studies included were performed under a protocol approved by the Institutional Review Board of Washington University School of Medicine (IRB ID# 201110126). All of the human studies here report anonymous analysis of de-identified pre-existing sera previously stored from earlier studies for which no additional consent was obtained.

## Results

### Conservation of ETEC pathogen-specific secreted antigens

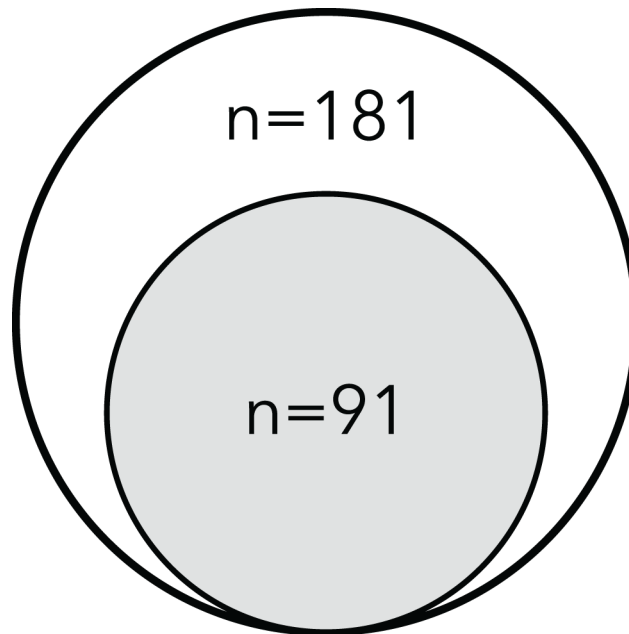
Two novel antigens, the EtpA adhesin, and the passenger domain of the EatA serine protease are encoded on the large 92 kilobase virulence plasmid of the prototypical ETEC strain H10407. Both of these secreted proteins[22,30] are required for H10407 to efficiently deliver heat-labile toxin to target epithelial cells. Furthermore, both of these antigens are immunogenic [25], and induce protective immune responses in a murine model of ETEC intestinal colonization[29,31]. To further assess their utility as potential vaccine antigens, we examined a large collection of ETEC strains that were well characterized with respect to associated clinical meta-data pertaining to disease severity and which had not undergone repeated serial passage in the laboratory.

Altogether, we found that these antigens are relatively conserved in the ETEC pathovar, confirming the results of earlier studies that focused on strains from different phylogenies obtained in Guinea Bissau and Chile [37,49]. Of the 181 strains examined in the present study (Fig. 1), we found that more than half of all strains produced EtpA (102/181, 56%) and/or EatA (106/181, 59%) (S1 Dataset), and that more than three quarters of all strains produced at least one of these antigens. Both EtpA and EatA were identified more than twice as frequently as the most commonly identified CF (CS6), which was identified in 22% of strains in this collection (Table 1).

### Relationship of plasmid-encoded virulence loci to colonization factor antigens

Importantly, although the genes encoding the *etpBAC* secretion system[19] and the EatA auto-transporter[21] were initially discovered on the same large virulence plasmid of H10407, which also encodes the colonization factor (CF) CFA/I, we found that these loci were not restricted to strains expressing this particular CF, but were widely distributed among the different CFs, and were also present in strains for which no CF could be identified (Fig. 2A). Indeed, half of the strains for which no CF could be identified expressed either EtpA or EatA, suggesting that these antigens could complement existing vaccination strategies centered on CFs. As expected by the association with multiple CFs, we also found that EtpA and EatA were secreted by strains from multiple phylogenetic lineages (Fig. 2B, C). Interestingly, however we found a





**Figure 1. Relationship of strain subsets used in antigen expression studies, and strains with available whole genome sequences.** All of the strains in the collection (n = 181) were examined for production of three secreted ETEC virulence proteins EtpA, EatA, and YghJ by immunoblotting of culture supernatants with the respective antibodies. A subset of these strains (n = 91) were recently sequenced at the Genome Sequencing Center for Infectious Diseases (GSCID).

doi:10.1371/journal.pntd.0003446.g001

**Table 1. Distribution of EtpA and EatA among strains expressing different colonization factors.**

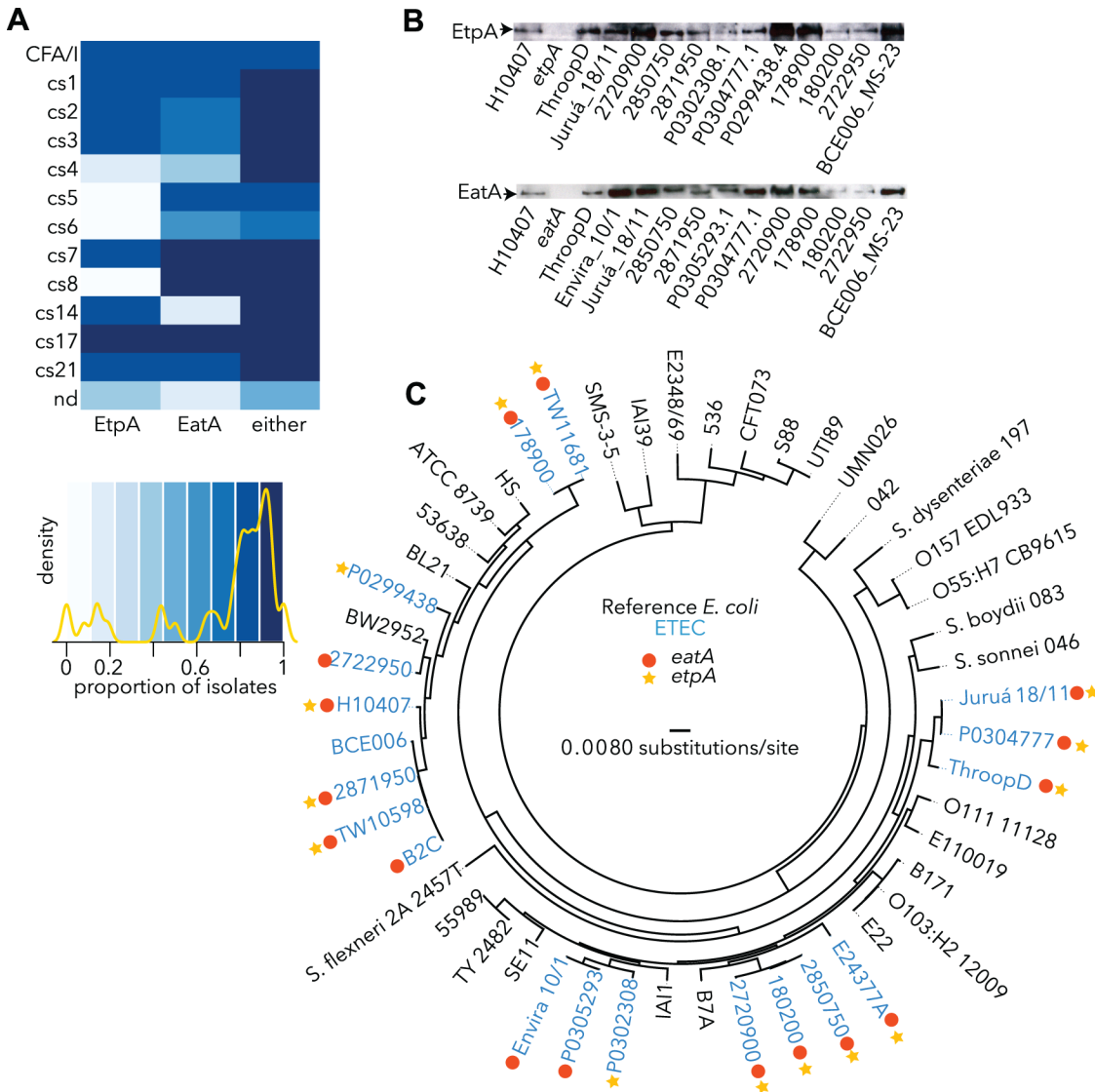
CF <sup>1</sup>	total (n)	EtpA n+ (%) <sup>2</sup>	EatA n+ (%) <sup>2</sup>	EtpA n+ or EatA n+ (%) <sup>2</sup>
CFA/I	21	17 (81)	17 (81)	18 (85)
CFA/II <sup>3</sup>				
cs1	14	12 (86)	12 (86)	13 (93)
cs2	10	8 (80)	7 (70)	9 (90)
cs3	25	21 (84)	19 (76)	23 (92)
CFA/IV <sup>3</sup>				
cs4	7	1 (14)	3 (43)	3 (43)
cs5	23	0 (0)	18 (78)	18 (78)
cs6	39	3 (8)	25 (64)	26 (67)
Other CFs				
cs7	16	14 (88)	15 (94)	15 (94)
cs8	2	0 (0)	2 (100)	2 (100)
cs14	14	12 (86)	2 (14)	13 (93)
cs17	10	9 (90)	9 (90)	9 (90)
cs21	17	14 (82)	14 (82)	16 (94)
nd	36	16 (44)	7 (19)	18 (50)

<sup>1</sup>expression as determined with corresponding monoclonal antibody

<sup>2</sup>refers to strains expressing proteins as determined by immunoblotting nd: not detected with any of the above monoclonal antibodies

<sup>3</sup>(CFA/II pili are typically comprised of CS1 or CS2 in combination with CS3; CFA/IV pili may be formed by CS4 or CS5 in combination with CS6 or CS6 alone.)

doi:10.1371/journal.pntd.0003446.t001



**Figure 2. Conservation of novel pathotype-specific antigens EtpA and EatA among phylogenetically distinct strains expressing different colonization factors.** a. Heatmap of EtpA and EatA showing the proportion of strains positive for expression of these antigens among different CF groups. CF antigen designation is shown at left of the heatmap. nd = no CF antigen detected. Below is the heatmap key depicting colors associated with each degree of antigen positivity. Density line in yellow depicts the relative number of map features assigned at each proportion. Primary data used to construct the heatmap can be found in [S2 Dataset](#). b. Immunoblot detection of EtpA and EatA expression among strains from different phylogenies. The upper immunoblot demonstrates EtpA production in the prototype H10407 strain, ThroopD isolated in Dallas, TX in 1975, the Jurua\_18/11 (Amazon, 1998), and phylogenetically dispersed strains from icddr,b. The *etpA* mutant is included as a negative control. The lower blot demonstrates EatA production by H10407, phylogenetically distributed strains from icddr,b and Envira\_10/1, an additional isolate from cholera-like outbreaks in the Amazon. The *eatA* mutant is included as a negative control. c. Phylogram showing the phylogenetic distribution of selected ETEC strains (designations in blue) and reference *E. coli* strains (designations in black). Red circles and gold stars represent *eatA*<sup>+</sup>, and *etpA*<sup>+</sup> strains, respectively.

doi:10.1371/journal.pntd.0003446.g002

negative association between the *etpBAC* locus and strains expressing CFA/IV antigens [50,51] including CS5 in that none of the 23 strains possessing CS5 fimbriae secreted the EtpA adhesin. Similarly, among strains expressing CS6, which is frequently co-expressed with CS5, only a minority secreted EtpA. These data are also consistent with our earlier observation that the prototype B7A strain, which also expresses CS6, lacks the *etpBAC* locus and does not secrete EtpA [19].

Interestingly, both the *eatA* and *etpBAC* loci were originally identified in ETEC strain H10407, originally isolated from an adult with severe, cholera-like illness in Bangladesh[52]. As has been noted previously, this strain also causes more severe illness in human clinical challenge studies relative to other strains like B7A that lack these loci[53]. Because we had clinical metadata pertaining to disease severity for all of the strains in our collection, we questioned whether the production of either of these antigens was associated with strains isolated from more severe forms of infection. However, we did not find any clear association between either of these putative virulence loci and clinical outcome (S1 Dataset).

## Conservation of chromosomally-encoded antigens

We also examined the conservation of two chromosomally-encoded antigens which are not specific to the ETEC pathovar, but have recently been shown to play a role in virulence. The *eaeH* gene was originally identified on the chromosome of ETEC strain H10407 by subtractive hybridization with *E. coli* MG1655[54], is transcriptionally activated by cell contact [26], and under these conditions EaeH is produced by a diverse group of strains belonging to different phylogenies[28]. Using the EaeH peptide sequence from H10407 (GenBank accession AAZ57201), BLASTP searches of recently sequenced ETEC strains from Bangladesh and elsewhere ([http://gscid.igs.umaryland.edu/wp.php?wp=comparative\\_genome\\_analysis\\_of\\_enterotoxigenic\\_e\\_coli\\_isolates\\_from\\_infections\\_of\\_different\\_clinical\\_severity](http://gscid.igs.umaryland.edu/wp.php?wp=comparative_genome_analysis_of_enterotoxigenic_e_coli_isolates_from_infections_of_different_clinical_severity)) also revealed that the *eaeH* gene was present in 63 out of 91 distinct isolates (69%) (S1 Dataset). BLASTP searches of these data for another chromosomally encoded molecule, YghJ, a type II secretion system effector[55] recently shown to be involved in mucin degradation and toxin delivery[27] demonstrated that the *yghJ* gene was present on the chromosomes in 83 of 91 (91%) isolates. Similarly, we identified the YghJ protein in a majority (161/181, 89%) of ETEC culture supernatants (S1 Dataset). This antigen was produced across ETEC strains expressing multiple CF types including 31/36 strains that were CF-negative by monoclonal antibody screening.

## EatA and EtpA sequence conservation

Ideally, putative vaccine targets should be specific to the pathovar under study or restricted to pathogenic isolates, but not subject to significant antigenic variation. Therefore to further examine the potential utility of two ETEC pathovar specific antigens, EtpA and EatA, as vaccine candidates, we used recently obtained DNA sequence information from multiple ETEC genomes belonging to different phylogenies and from temporally and geographically disparate sources to compare the predicted amino acid sequences of these proteins.

For the prototype EatA molecule, first described in ETEC H10407[21], the 1042 residue region from amino acids 57–1098 is predicted for the secreted passenger domain that contains the serine protease catalytic triad[21] as well as protective epitopes[23]. We therefore compared this region of the molecule to those derived from the recently released genome sequences of multiple ETEC strains. Altogether, we found that the sequence of the EatA passenger domain (EatA<sub>p</sub>) was very highly conserved across strains, and exhibited between 95–100% identity to the prototype H10407 Eat<sub>p</sub> (Table 2). Likewise, the predicted serine protease catalytic motif formed by the histidine, aspartic acid and serine residues at positions 134, 162, and 267, respectively were universally conserved within the passenger domains of these proteins (S1 Fig.). Similarly, the predicted amino acid sequences of the secreted EtpA adhesin molecules from multiple strains exhibited between 94 and 100% identity to the H10407 prototype antigen (Table 3, S2 Fig.).

Despite the fact that the comparator strains included here spanned isolates collected over nearly 40 years, belonging to different phylogenies and that strains originated in diverse

**Table 2. EatA sequence conservation in geographically, temporally, and phylogenically disparate isolates.**

strain	accession	origin	isolation reported	phenotype	identity <sup>1</sup> (%)	Similarity (%)	ref.
H10407	<a href="#">Q84GK0.1</a>	Bangladesh	1973	cholera-like	-	-	[52]
Throop D	<sup>2</sup> <a href="#">EMW91712.1</a>	Dallas, TX	1976	cholera-like	98	99	[36]
Envira 10/1	<sup>2</sup> <a href="#">EMX71514.1</a>	Amazon, Brazil	1998	cholera-like	97	98	[35]
Juruá 18/11	<sup>2</sup> <a href="#">EMX58104.1</a>	Amazon, Brazil	1998	cholera-like	96	98	[35]
2850750	<sup>2</sup> <a href="#">EMW01116.1</a>	Bangladesh	2008	cholera-like	96	97	<sup>5</sup>
2871950	<sup>2</sup> <a href="#">EMV49671.1</a>	Bangladesh	2008	cholera-like	96	98	<sup>5</sup>
P0305293.1	<sup>2</sup> <a href="#">EMZ82436.1</a>	Bangladesh	2011	cholera-like	97	98	<sup>5</sup>
P0304777.1	<sup>2</sup> <a href="#">EMX02137.1</a>	Bangladesh	2011	cholera-like	96	97	<sup>5</sup>
2720900	<sup>2</sup> <a href="#">EMX91421.1</a>	Bangladesh	2007	cholera-like	100	100	<sup>5</sup>
178900	<sup>2</sup> <a href="#">ENA71292.1</a>	Bangladesh	2010	cholera-like	98	99	<sup>5</sup>
180200	<sup>2</sup> <a href="#">ENA61350.1</a>	Bangladesh	2010	cholera-like	96	98	<sup>5</sup>
2722950	<sup>2</sup> <a href="#">EMZ73955.1</a>	Bangladesh	2007	mild-self-limited	98	99	<sup>5</sup>
tw10598	<sup>3</sup> <a href="#">AELA00000000.1</a>	Guinea-Bissau	1996	diarrhea <sup>4</sup>	96	98	[49]
tw10722	<sup>3</sup> <a href="#">AELB00000000.1</a>	Guinea-Bissau	1996	diarrhea <sup>4</sup>	95	98	[49]
tw10828	<sup>3</sup> <a href="#">AELC00000000.1</a>	Guinea-Bissau	1996	diarrhea <sup>4</sup>	95	97	[49]
tw11681	<sup>3</sup> <a href="#">AELD00000000.1</a>	Guinea-Bissau	1996	diarrhea <sup>4</sup>	98	99	[49]
B2C	<a href="#">ETS27975.1</a>	Vietnam	1971	diarrhea <sup>4</sup>	96	98	[62]
E24377A	<a href="#">YP_001451588.1</a>	Egypt	1980s	diarrhea <sup>4</sup>	96	97	[63]

<sup>1</sup>based on BLAST-P searches against 1042 residues of predicted passenger domain of H10407.

<sup>2</sup>sequenced at GSCID ([http://gscid.igs.umaryland.edu/wp.php?wp= comparative\\_genome\\_analysis\\_of\\_enterotoxigenic\\_e\\_coli\\_isolates\\_from\\_infections\\_of\\_different\\_clinical\\_severity](http://gscid.igs.umaryland.edu/wp.php?wp= comparative_genome_analysis_of_enterotoxigenic_e_coli_isolates_from_infections_of_different_clinical_severity))

<sup>3</sup>open reading frames corresponding to the eatA gene were first assembled from whole genome shotgun sequence contigs for these draft genomes; BLASTP for these homologues was conducted using CLC Main Workbench v6.9.1 and local database of predicted protein sequences derived from translation of assembled contigs from the respective sequencing projects.

<sup>4</sup>(severity unknown)

<sup>5</sup>this study

doi:10.1371/journal.pntd.0003446.t002

locations in Asia, Africa and the Americas, both proteins appear to exhibit remarkably little antigenic variation. Likewise, in analysis of the genomes of strains isolated recently within Bangladesh both proteins demonstrated similar degrees of sequence conservation (S1–S2 Fig).

### Immunogenicity of novel virulence antigens

Earlier immunoproteomic studies suggested that a variety of conserved *E. coli* proteins as well as ETEC pathovar specific proteins are recognized during the course of experimental infections in mice, and these responses parallel those observed using pooled convalescent sera from ETEC patients [25]. To further characterize the immune response to novel antigens we focused on four proteins that have recently been shown to play a role in ETEC pathogenesis, including two plasmid-encoded secreted ETEC pathovar-specific antigens: EatA protease, and the EtpA adhesin, as well as the highly conserved chromosomally-encoded YghJ metalloprotease and the EaeH adhesin protein.

In comparing convalescent plasma from patients hospitalized at icddr,b to uninfected controls from Bangladesh, we found that patients in general exhibited significantly greater total antibody (IgG, IgM, IgA) responses to each of these antigens following diarrheal illness (Fig. 3) suggesting that these proteins are expressed during the course of infection. Similar results were

**Table 3. EtpA sequence conservation in geographically, temporally, and phylogenically disparate isolates.**

strain	accession	origin	isolation reported	phenotype	identity <sup>1</sup> (%)	Similarity (%)	ref.
H10407	<a href="#">AAX13509.2</a>	Bangladesh	1973	cholera-like	-	-	[52]
Throop D	<sup>2</sup> <a href="#">EMW91721.1</a>	Dallas, TX	1976	cholera-like	97	97	[36]
Juruá 18/11	<sup>2</sup> <a href="#">EMX66921.1</a>	Amazon, Brazil	1998	cholera-like	94	95	[35]
2720900	<sup>2</sup> <a href="#">EMX81000.1</a>	Bangladesh	2007	cholera-like	100	100	<sup>5</sup>
2850750	<sup>2</sup> <a href="#">EMW11804.1</a>	Bangladesh	2008	cholera-like	94	95	<sup>5</sup>
2871950	<sup>2</sup> <a href="#">EMV50451.1</a>	Bangladesh	2011	cholera-like	95	95	<sup>5</sup>
p0302308.1	<sup>2</sup> <a href="#">EMX03881.1</a>	Bangladesh	2011	cholera-like	98	98	<sup>5</sup>
p0304777.1	<sup>2</sup> <a href="#">EMX02126.1</a>	Bangladesh	2011	cholera-like	99	98	<sup>5</sup>
p0299438.4	<sup>2</sup> <a href="#">ENB94816.1</a>	Bangladesh	2011	cholera-like	99	99	<sup>5</sup>
178900	<sup>2</sup> <a href="#">ENA69446.1</a>	Bangladesh	2010	cholera-like	97	97	<sup>5</sup>
180200	<sup>2</sup> <a href="#">ENA71844.1</a>	Bangladesh	2010	cholera-like	94	95	<sup>5</sup>
1392/75	<a href="#">YP_003717617.1</a>	Hong Kong	-	diarrhea <sup>4</sup>	95	96	[61]
tw11681	<sup>3</sup> <a href="#">AELD00000000.1</a>	Guinea-Bissau	1996	diarrhea <sup>4</sup>	96	97	[49]
tw14425	<sup>3</sup> <a href="#">AELE00000000.1</a>	Guinea-Bissau	1996	diarrhea <sup>4</sup>	99	100	[49]
tw10828	<sup>3</sup> <a href="#">AELC00000000.1</a>	Guinea-Bissau	1996	diarrhea <sup>4</sup>	94	96	[49]
tw10598	<sup>3</sup> <a href="#">AELA00000000.1</a>	Guinea-Bissau	1996	diarrhea <sup>4</sup>	95	96	[49]
E24377A	<sup>3</sup> <a href="#">NC_009786.1</a>	Egypt	1980s	diarrhea <sup>4</sup>	95	96	[63]

<sup>1</sup>identity and similarity percentages reflect BLASTP comparisons of the first 600 residues of the EtpA sequence.

<sup>2</sup>sequenced at GSCID ([http://gscid.igs.umaryland.edu/wp.php?wp=comparative\\_genome\\_analysis\\_of\\_enterotoxigenic\\_e\\_coli\\_isolates\\_from\\_infections\\_of\\_different\\_clinical\\_severity](http://gscid.igs.umaryland.edu/wp.php?wp=comparative_genome_analysis_of_enterotoxigenic_e_coli_isolates_from_infections_of_different_clinical_severity))

<sup>3</sup>open reading frames corresponding to the *etpA* gene were first assembled from whole genome shotgun sequence contigs for these draft genomes; BLASTP for these homologues was conducted using CLC Main Workbench v6.9.1 and local database of predicted protein sequences derived from translation of assembled contigs from the respective sequencing projects.

<sup>4</sup>(severity unknown)

<sup>5</sup>this study

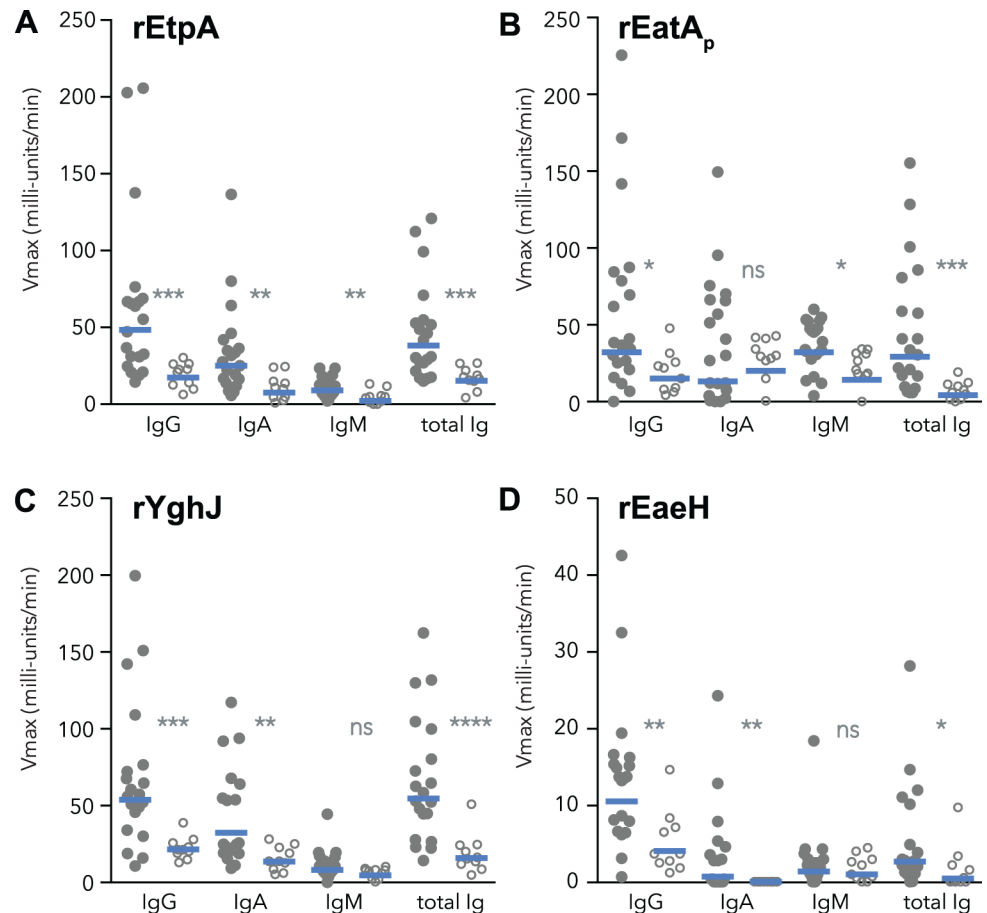
doi:10.1371/journal.pntd.0003446.t003

obtained in comparing plasma from un-infected children from a non-endemic area in the United States (S3 Fig.).

We also examined the immune response to EtpA following infection by examining sera obtained before and after challenge of human volunteers with ETEC H10407. In sera obtained from two independent volunteer challenge studies, we also observed significant increases in immune responses to EtpA (S4 Fig.), strongly suggesting that this secreted protein is specifically recognized following infection by ETEC strains including H10407 that secrete this antigen.

### Protective efficacy of combined EtpA-mutant EatA passenger vaccination

Previous studies have demonstrated that individually, vaccination with either EtpA [31,56] or the passenger domain of the EatA serine protease [29] affords protection against intestinal colonization in mice. The data above suggest that collectively these antigens might significantly extend coverage presently offered by classical approaches to ETEC vaccine development. We therefore questioned whether these two antigens could be successfully combined in a subunit approach. Because we have previously demonstrated that the native secreted EatA passenger domain will degrade intestinal mucin [29] as well as the EtpA adhesin molecule [22], we elected

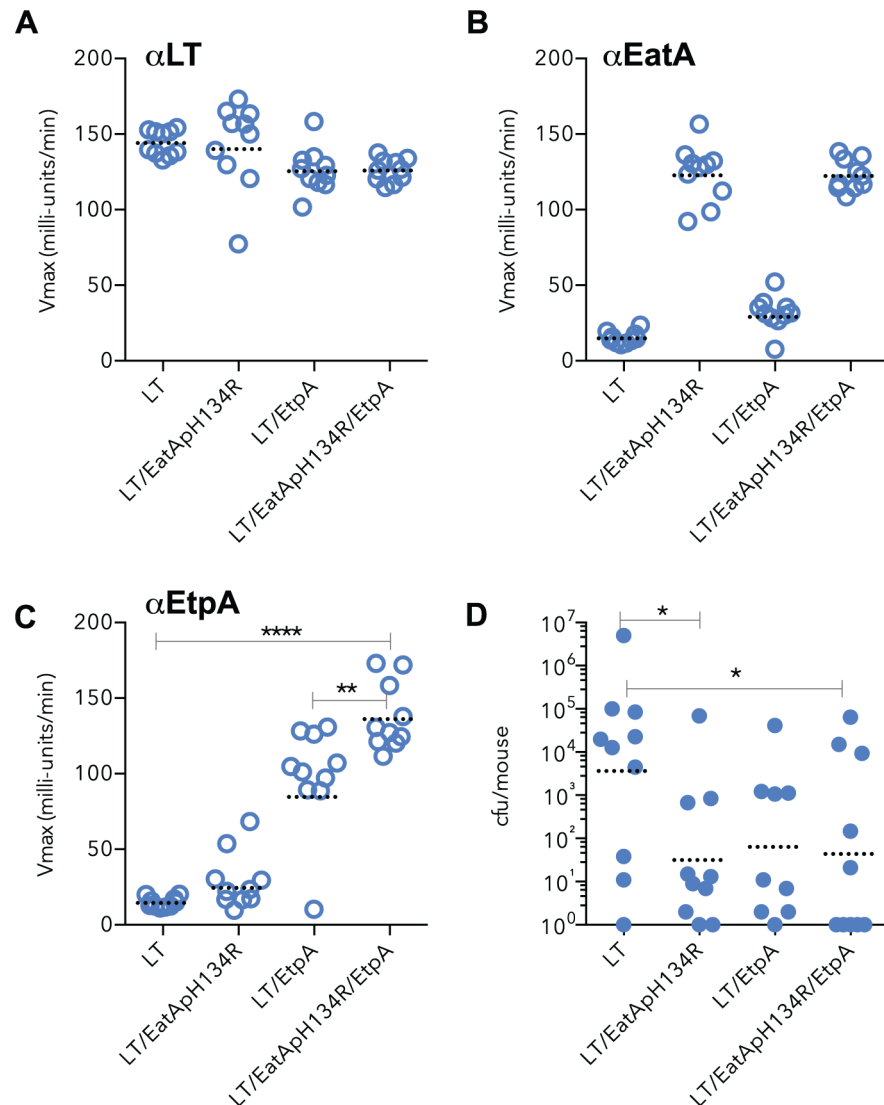


**Figure 3. Recognition of novel antigens during naturally occurring ETEC infections in Bangladesh.** Shown are kinetic ELISA data for four different recombinant antigens (a, rEtpA; b, the rEatA passenger domain; c, rYghJ; and d, rEaeH) obtained with 1:4096 dilutions of convalescent plasma from ETEC-infected patients hospitalized at ICDDR,B in Dhaka, Bangladesh (closed circles), or control patients not infected with enterotoxigenic *E. coli* (open circles). Horizontal bars represent geometric mean Vmax kinetic ELISA values for each group. P values obtained by two-tailed Mann Whitney testing of groups are summarized (\*<0.05; \*\*<0.01; \*\*\*<0.001; \*\*\*\*<0.0001). x-axis of each graph depicts the specificity of the secondary antibody used in the ELISA (IgG, IgA, IgM, and total Ig, IgA, and IgM).

doi:10.1371/journal.pntd.0003446.g003

to vaccinate animals with a modified recombinant version of the EatA passenger that lacks protease activity (rEatA<sub>PH134R</sub>).

Co-vaccination with rEtpA and the mutant rEatA<sub>PH134R</sub> molecule elicited robust serologic responses to both molecules that were comparable to vaccination with either antigen alone. As anticipated, each of the groups mounted strong serologic responses to the LT adjuvant (Fig. 4A), and both antigens retained their immunogenicity following co-immunization of EtpA with the rEatA<sub>PH134R</sub> passenger domain (Fig. 4B,C) with responses that were at least comparable to those obtained following immunization with either antigen alone (Fig. 4C). Likewise, mice immunized with both antigens were significantly protected against colonization by ETEC (Fig. 4D), although co-vaccination with both antigens did not appear to be more effective than vaccination with either antigen alone. Collectively, however, these data suggest that co-immunization with these two antigens is feasible, and could be employed to expand present approaches to ETEC vaccine antigen selection.



**Figure 4. Mice immunized with rEtpA and rEatA<sub>H134R</sub> are protected against ETEC infection.** Serologic responses to **a** heat-labile toxin (LT), **b** the passenger domain of EatA (Eatp), **c** EtpA. Shown are serum IgG responses following intranasal vaccination of mice with the LT adjuvant alone, or LT with 15 µg of either the proteolytically inactive passenger domain (EatA<sub>H134R</sub>), EtpA, or both antigens on days 0, 14, 28. Colonization of mice following immunization with the adjuvant alone compared with single and dual antigen vaccination. Comparisons between groups were by Mann Whitney two tailed nonparametric testing. (One mouse died during the vaccination period in the LT/EtpA group and was therefore excluded from the analysis). For mice with no detectable colonies following challenge, the number of cfu is arbitrarily reported as 1 (10<sup>0</sup>) cfu, the theoretical limit of detection.

doi:10.1371/journal.pntd.0003446.g004

## Discussion

Enterotoxigenic *Escherichia coli* remain one of the most common causes of infectious diarrhea worldwide, and severe disease caused by these pathogens persists as leading cause of death among young children in developing countries[1]. Despite recognition of these toxin producing *E. coli* as a cause of severe cholera-like diarrheal illness more than forty years ago[57], there remains no effective broadly protective vaccine for ETEC.



Most vaccinology efforts to date have focused almost exclusively on a subset of plasmid-encoded antigens, namely the colonization factors (CFs) and heat-labile toxin[9]. Vaccines based on this strategy have faced several impediments. First, the CFs are quite diverse with more than 26 distinct antigens described to date. In addition, a number of recent vaccine studies have suggested that simply engendering immune responses to CFs and/or heat-labile toxin may not be sufficient to provide sustained broad-based protection[14–16].

Recent studies of ETEC pathogenesis suggest that a number of virulence factors in addition to the CFs are involved in efficient delivery of toxins to their cognate receptors on the epithelial surface[2,22,29,30,48,58]. Similarly, the immune response to ETEC infection appears to involve many proteins[25,47] in addition to the classical antigens that are the present focus of most vaccines. Collectively, these findings suggest that there may be additional molecules that could be targeted to interdict toxin delivery by these pathogens, expand the list of potential protective antigens, and complement existing approaches to vaccine development for ETEC[59,60].

A major challenge to ETEC vaccine development in general is that the most highly conserved antigens of ETEC, typically encoded on core regions of the chromosome, are also shared with commensal *E. coli*[60]. Included among these chromosomally encoded conserved proteins are two antigens studied here, YghJ[27] and EaeH[28] that were recently shown to be important for ETEC virulence. While the present studies also demonstrate that these proteins are recognized during the course of ETEC infection, the degree to which these antigens can be safely targeted in vaccines without inadvertent disruption of the intestinal microflora remains to be studied.

The inherent plasticity of *E. coli* genomes contributes substantially to the difficulty in defining antigens unique to the ETEC pathovar that are widely conserved. No single antigen exclusive to these pathogens, but universally conserved in this pathovar, has been described to date. Some have suggested that this might be predicted based on the fact that the plasmid-encoded heat-labile and/or heat-stable toxins, which define the ETEC pathovar, could form a minimal complement of virulence genes in wide variety of *E. coli* host strains[61]. Nevertheless, earlier studies conducted on phylogenically disparate strains from Guinea Bissau[49] and Chile[37] suggested that genes encoding two pathogen-specific antigens EatA and EtpA were present in a majority of strains.

In this context, we examined the gene conservation and the actual production of these proteins in a large collection of well-characterized strains from Bangladesh, complemented by strains from other locations that were associated with severe disease and for which there were available clinical metadata. Notably, two plasmid-encoded ETEC pathotype-specific antigens, the EatA serine protease and the secreted EtpA adhesin molecule were shared broadly among strains belonging to different CF groups with the exception of strains that produced CFA/IV antigens CS4, CS5, CS6 which only infrequently produced EtpA.

The studies reported here represent the largest screen for EtpA and EatA secretion in ETEC performed to date. Earlier studies reporting that genes encoding both proteins were highly conserved relied on either PCR[37] or screening of draft genomes[49] for the presence of the corresponding loci. In general, we found high degrees of concordance between the presence of these genes by PCR and production of the corresponding protein. We should point out however that draft genome assemblies typically fail to encompass the entire *etpA* gene as automated assembly algorithms cannot faithfully incorporate the multiple repeat regions comprising two thirds of *etpA*. This could impact interpretation of gene prevalence in ongoing large-scale ETEC genome sequencing projects. The prevalence of EtpA and EatA (56 and 59%, respectively) as determined by examination of protein expression in our study was slightly lower than previously reported in earlier studies that analyzed strains from Guinea Bissau, where both genes were present in 75% of strains [49]; or Chile, where *etpA* and *eatA*, were present in 71 and 75% of

strains, respectively[37]. Nevertheless similar to these earlier studies, the strains that produced these antigens belonged to many different phylogenies suggesting that genes encoding these antigens have been widely dispersed.

The analyses of strains in these studies largely focused on isolates from Bangladesh. However, these data are potentially relevant for vaccine development for a number of reasons. First, Bangladesh is highly endemic for enterotoxigenic *E. coli* infections, and consequently remains an important site for vaccine field trials. In addition, ETEC has been under study in this region since the discovery of this pathotype, permitting us to compare sequence variation in candidate antigens over four decades. Understanding both current prevalence and sequence conservation of potential novel vaccine antigens in this population over time will be particularly important for making rational decisions about their inclusion in future iterations of ETEC vaccines. Finally, the geographic and temporal dispersal of genes encoding EtpA and EatA in multiple phylogenetic backgrounds, further attests to importance of studying these molecules as potential vaccine targets as previously suggested by others[37,49].

The optimal formulation of an ETEC vaccine has yet to be defined, and many questions pertaining to the nature of protective immunity that develops following infections with these pathogens remain. Nevertheless, the data presented here do suggest that the novel pathovar-specific antigens could complement existing strategies for ETEC vaccine development by broadening the antigenic valency. Whether the expanded coverage afforded by inclusion of additional pathotype specific antigens would enhance vaccine efficacy beyond that presently achieved by targeting CFs and LT will need to be determined empirically.

## Supporting Information

**S1 Fig. EatA passenger domain alignments of predicted EatA sequences corresponding to multiple phylogenies from strains belonging to disparate geographic origins.** The conserved catalytic triad at amino acids H78, D106, and S211 is highlighted by gray background shading. Geographic origin of strains is depicted in the color key at left of the alignment. Alignments were performed using CLUSTAL Omega (release 1.2.0 AndreaGiacomo) [40] algorithm plugin for CLC Main Workbench.

(PDF)

**S2 Fig. EtpA alignments of predicted sequences corresponding to multiple phylogenies from strains belonging to disparate geographic origins.** Geographic origin of strains is depicted in the color key at left of the alignment. Alignments were performed using sequence alignment algorithm of CLC Main Workbench v6.9.1 with the following parameters: gap open cost = 10.0; gap extension cost = 1.0; end gap cost = as any other; alignment mode = very accurate (slow); redo alignments = no; use fixpoints = no.

(PDF)

**S3 Fig. Immune response to selected ETEC proteins in infected patients and Bangladeshi and US controls.** Shown are (IgG) kinetic ELISA responses (in Vmax, milli-units/min) to recombinant proteins comparing convalescent plasma from patients hospitalized with acute ETEC infections at the International Centre for Diarrhoeal Disease Research in Dhaka, Bangladesh with controls (c) from Bangladeshi adults, and children, as well as plasma from age-matched children from Saint Louis Children's Hospital (slch). Antigens included two plasmid-encoded ETEC specific antigens (a) EtpA, and (b) the EatA passenger domain; and two chromosomally-encoded conserved antigens (c) YghJ, and (d) EaeH. All plasma samples were diluted 1:4096.

(PDF)

**S4 Fig. Immune responses to EtpA following volunteer challenge with ETEC H10407.** All sera were diluted 1:4096 prior to testing against rEtpA-myc-6His followed by detection of total antibody (IgM,IgG,IgA) in kinetic ELISA. Pre and post values (open and closed circles, respectively) represent collective data from 2 independent ETEC H10407 challenge studies CIR218 and CIR193a. Data from CIR218 are shown as pre-challenge (d-2, open blue circles) and (d28, closed blue circles), while data from CIR193a appear as open grey circles (pre-challenge, d0) and closed grey circles (post challenge, d9). Dashed horizontal lines represent geometric means. P value represents comparison of pre and post-challenge samples by Mann Whitney 2-tailed analysis.

(PDF)

**S1 Table. Oligonucleotide primers used in this study.** Shown in the table are oligonucleotide pairs, and predicted product sizes for ETEC pathotype-specific virulence gene amplification in these studies.

(PDF)

**S1 Dataset. S1 Dataset.xlsx contains the complete listing of all 181 strains analyzed for production of the secreted EtpA, EatA antigens (tab 1), and YghJ (tab 2).** Tab 3 shows 91 unique isolates sequenced by the Genome Sequencing Center for Infectious Diseases (GSCID) which were screened for the presence of the *eaeH* gene by BLASTP homology searches. Note that in the dataset, disease severity is classified numerically (0 = isolate from asymptomatic colonization, 1 = mild disease, 2 = severe cholera-like illness). The presence or absence of a given protein or gene in each strain is defined in a binary fashion where 1 = yes; 0 = no.

(XLSX)

**S2 Dataset. S2 Dataset.csv contains numerical data corresponding to the heatmap depicted in Fig. 1A.**

(CSV)

## Author Contributions

Conceived and designed the experiments: JMF FQ DAR QL. Performed the experiments: QL RK AS. Analyzed the data: QL DAR JF. Contributed reagents/materials/analysis tools: FQ DAR. Wrote the paper: QL JF DAR.

## References

1. Kotloff KL, Nataro JP, Blackwelder WC, Nasrin D, Farag TH, et al. (2013) Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): a prospective, case-control study. *Lancet* doi: [10.1016/S0140-6736\(13\)60844-2](https://doi.org/10.1016/S0140-6736(13)60844-2) PMID: [23680352](https://pubmed.ncbi.nlm.nih.gov/23680352/)
2. Fleckenstein JM, Hardwidge PR, Munson GP, Rasko DA, Sommerfelt H, et al. (2010) Molecular mechanisms of enterotoxigenic *Escherichia coli* infection. *Microbes Infect* 12: 89–98. doi: [10.1016/j.micinf.2009.10.002](https://doi.org/10.1016/j.micinf.2009.10.002) PMID: [19883790](https://pubmed.ncbi.nlm.nih.gov/19883790/)
3. Evans DG, Silver RP, Evans DJ Jr., Chase DG, Gorbach SL (1975) Plasmid-controlled colonization factor associated with virulence in *Escherichia coli* enterotoxigenic for humans. *Infect Immun* 12: 656–667. PMID: [1100526](https://pubmed.ncbi.nlm.nih.gov/1100526/)
4. Carpenter CC, Barua D, Wallace CK, Sack RB, Mitra PP, et al. (1965) Clinical and physiological observations during an epidemic outbreak of non-vibrio cholera-like disease in Calcutta. *Bull World Health Organ* 33: 665–671. PMID: [5295147](https://pubmed.ncbi.nlm.nih.gov/5295147/)
5. Gorbach SL, Banwell JG, Chatterjee BD, Jacobs B, Sack RB (1971) Acute undifferentiated human diarrhea in the tropics. I. Alterations in intestinal microflora. *J Clin Invest* 50: 881–889. doi: [10.1172/JCI106560](https://doi.org/10.1172/JCI106560) PMID: [4926260](https://pubmed.ncbi.nlm.nih.gov/4926260/)

6. Sack RB, Gorbach SL, Banwell JG, Jacobs B, Chatterjee BD, et al. (1971) Enterotoxigenic *Escherichia coli* isolated from patients with severe cholera-like disease. *J Infect Dis* 123: 378–385. doi: [10.1093/infdis/123.4.378](https://doi.org/10.1093/infdis/123.4.378) PMID: [4938945](https://pubmed.ncbi.nlm.nih.gov/4938945/)
7. Evans DG, Satterwhite TK, Evans DJ Jr., DuPont HL (1978) Differences in serological responses and excretion patterns of volunteers challenged with enterotoxigenic *Escherichia coli* with and without the colonization factor antigen. *Infect Immun* 19: 883–888. PMID: [346488](https://pubmed.ncbi.nlm.nih.gov/346488/)
8. Satterwhite TK, Evans DG, DuPont HL, Evans DJ Jr. (1978) Role of *Escherichia coli* colonisation factor antigen in acute diarrhoea. *Lancet* 2: 181–184. doi: [10.1016/S0140-6736\(78\)91921-9](https://doi.org/10.1016/S0140-6736(78)91921-9) PMID: [78384](https://pubmed.ncbi.nlm.nih.gov/78384/)
9. Svennerholm AM, Lundgren A (2012) Recent progress toward an enterotoxigenic *Escherichia coli* vaccine. *Expert review of vaccines* 11: 495–507. doi: [10.1586/erv.12.12](https://doi.org/10.1586/erv.12.12) PMID: [22551034](https://pubmed.ncbi.nlm.nih.gov/22551034/)
10. Zhang W, Sack DA (2012) Progress and hurdles in the development of vaccines against enterotoxigenic *Escherichia coli* in humans. *Expert review of vaccines* 11: 677–694. doi: [10.1586/erv.12.37](https://doi.org/10.1586/erv.12.37) PMID: [22873126](https://pubmed.ncbi.nlm.nih.gov/22873126/)
11. Rasko DA, Rosovitz MJ, Myers GS, Mongodin EF, Fricke WF, et al. (2008) The pangenome structure of *Escherichia coli*: comparative genomic analysis of *E. coli* commensal and pathogenic isolates. *J Bacteriol* 190: 6881–6893. doi: [10.1128/JB.00619-08](https://doi.org/10.1128/JB.00619-08) PMID: [18676672](https://pubmed.ncbi.nlm.nih.gov/18676672/)
12. Del Canto F, Botkin DJ, Valenzuela P, Popov V, Ruiz-Perez F, et al. (2012) Identification of the Coli Surface Antigen 23 (CS23), a Novel Adhesin of Enterotoxigenic *Escherichia coli*. *Infection and immunity* doi: [10.1128/IAI.00263-12](https://doi.org/10.1128/IAI.00263-12) PMID: [22645287](https://pubmed.ncbi.nlm.nih.gov/22645287/)
13. Isidean SD, Riddle MS, Savarino SJ, Porter CK (2011) A systematic review of ETEC epidemiology focusing on colonization factor and toxin expression. *Vaccine* 29: 6167–6178. doi: [10.1016/j.vaccine.2011.06.084](https://doi.org/10.1016/j.vaccine.2011.06.084) PMID: [21723899](https://pubmed.ncbi.nlm.nih.gov/21723899/)
14. Riddle MS, Savarino SJ (2013) Moving beyond a heat-labile enterotoxin-based vaccine against enterotoxigenic *Escherichia coli*. *Lancet Infect Dis*. doi: [10.1016/S1473-3099\(13\)70355-4](https://doi.org/10.1016/S1473-3099(13)70355-4) PMID: [24291167](https://pubmed.ncbi.nlm.nih.gov/24291167/)
15. Behrens RH, Cramer JP, Jelinek T, Shaw H, von Sonnenburg F, et al. (2013) Efficacy and safety of a patch vaccine containing heat-labile toxin from *Escherichia coli* against travellers' diarrhoea: a phase 3, randomised, double-blind, placebo-controlled field trial in travellers from Europe to Mexico and Guatemala. *Lancet Infect Dis*. doi: [10.1016/S1473-3099\(13\)70297-4](https://doi.org/10.1016/S1473-3099(13)70297-4) PMID: [24291168](https://pubmed.ncbi.nlm.nih.gov/24291168/)
16. Darsley MJ, Chakraborty S, Denearing B, Sack DA, Feller A, et al. (2012) ACE527 Oral, Live Attenuated ETEC Vaccine Reduces the Incidence and Severity of Diarrhea in a Human Challenge Model of Diarrheal Disease. *Clin Vaccine Immunol* doi: [10.1128/CVI.00364-12](https://doi.org/10.1128/CVI.00364-12) PMID: [23035175](https://pubmed.ncbi.nlm.nih.gov/23035175/)
17. Harro C, Sack D, Bourgeois AL, Walker R, Denearing B, et al. (2011) A combination vaccine consisting of three live attenuated enterotoxigenic *Escherichia coli* strains expressing a range of colonization factors and LTb is well tolerated and immunogenic in a placebo-controlled double-blind Phase I trial in healthy adults. *Clin Vaccine Immunol* doi: [10.1128/CVI.05342-11](https://doi.org/10.1128/CVI.05342-11) PMID: [21994354](https://pubmed.ncbi.nlm.nih.gov/21994354/)
18. Turner AK, Stephens JC, Beavis JC, Greenwood J, Gewert C, et al. (2011) Generation and characterization of a live attenuated enterotoxigenic *Escherichia coli* combination vaccine expressing six colonization factors and heat-labile toxin subunit B. *Clin Vaccine Immunol* 18: 2128–2135. doi: [10.1128/CVI.05345-11](https://doi.org/10.1128/CVI.05345-11) PMID: [21994355](https://pubmed.ncbi.nlm.nih.gov/21994355/)
19. Fleckenstein JM, Roy K, Fischer JF, Burkitt M (2006) Identification of a two-partner secretion locus of enterotoxigenic *Escherichia coli*. *Infect Immun* 74: 2245–2258. doi: [10.1128/IAI.74.4.2245-2258.2006](https://doi.org/10.1128/IAI.74.4.2245-2258.2006) PMID: [16552055](https://pubmed.ncbi.nlm.nih.gov/16552055/)
20. Roy K, Hilliard GM, Hamilton DJ, Luo J, Ostmann MM, et al. (2009) Enterotoxigenic *Escherichia coli* EtpA mediates adhesion between flagella and host cells. *Nature* 457: 594–598. doi: [10.1038/nature07568](https://doi.org/10.1038/nature07568) PMID: [19060885](https://pubmed.ncbi.nlm.nih.gov/19060885/)
21. Patel SK, Dotson J, Allen KP, Fleckenstein JM (2004) Identification and molecular characterization of EatA, an autotransporter protein of enterotoxigenic *Escherichia coli*. *Infect Immun* 72: 1786–1794. doi: [10.1128/IAI.72.3.1786-1794.2004](https://doi.org/10.1128/IAI.72.3.1786-1794.2004) PMID: [14977988](https://pubmed.ncbi.nlm.nih.gov/14977988/)
22. Roy K, Kansal R, Bartels SR, Hamilton DJ, Shaaban S, et al. (2011) Adhesin Degradation Accelerates Delivery of Heat-labile Toxin by Enterotoxigenic *Escherichia coli*. *J Biol Chem* 286: 29771–29779. doi: [10.1074/jbc.M111.251546](https://doi.org/10.1074/jbc.M111.251546) PMID: [21757737](https://pubmed.ncbi.nlm.nih.gov/21757737/)
23. Kumar P, Luo Q, Vickers TJ, Sheikh A, Lewis WG, et al. (2013) EatA, an Immunogenic Protective Antigen of Enterotoxigenic *Escherichia coli* Degrades Intestinal Mucin. *Infect Immun*. doi: [10.1128/IAI.01078-13](https://doi.org/10.1128/IAI.01078-13) PMID: [24478066](https://pubmed.ncbi.nlm.nih.gov/24478066/)
24. Johansson ME, Sjovall H, Hansson GC (2013) The gastrointestinal mucus system in health and disease. *Nat Rev Gastroenterol Hepatol* 10: 352–361. doi: [10.1038/nrgastro.2013.35](https://doi.org/10.1038/nrgastro.2013.35) PMID: [23478383](https://pubmed.ncbi.nlm.nih.gov/23478383/)
25. Roy K, Bartels S, Qadri F, Fleckenstein JM (2010) Enterotoxigenic *Escherichia coli* elicits immune responses to multiple surface proteins. *Infect Immun* 78: 3027–3035. doi: [10.1128/IAI.00264-10](https://doi.org/10.1128/IAI.00264-10) PMID: [20457787](https://pubmed.ncbi.nlm.nih.gov/20457787/)

26. Kansal R, Rasko DA, Sahl JW, Munson GP, Roy K, et al. (2013) Transcriptional modulation of enterotoxigenic *Escherichia coli* virulence genes in response to epithelial cell interactions. *Infect Immun* 81: 259–270. doi: [10.1128/IAI.00919-12](https://doi.org/10.1128/IAI.00919-12) PMID: [23115039](https://pubmed.ncbi.nlm.nih.gov/23115039/)
27. Luo Q, Kumar P, Vickers TJ, Sheikh A, Lewis WG, et al. (2014) Enterotoxigenic *Escherichia coli* Secretes a Highly Conserved Mucin-Degrading Metalloprotease To Effectively Engage Intestinal Epithelial Cells. *Infect Immun* 82: 509–521. doi: [10.1128/IAI.01106-13](https://doi.org/10.1128/IAI.01106-13) PMID: [24478067](https://pubmed.ncbi.nlm.nih.gov/24478067/)
28. Sheikh A, Lou Q, Roy K, Shabaan S, Kumar P, et al. (2014) Contribution of the highly conserved EaeH surface protein to enterotoxigenic *Escherichia coli* pathogenesis. *Infect Immun* doi: [10.1128/IAI.01890-14](https://doi.org/10.1128/IAI.01890-14) PMID: [24935979](https://pubmed.ncbi.nlm.nih.gov/24935979/)
29. Kumar P, Luo Q, Vickers TJ, Sheikh A, Lewis WG, et al. (2014) EatA, an Immunogenic Protective Antigen of Enterotoxigenic *Escherichia coli*, Degrades Intestinal Mucin. *Infect Immun* 82: 500–508. doi: [10.1128/IAI.01078-13](https://doi.org/10.1128/IAI.01078-13) PMID: [24478066](https://pubmed.ncbi.nlm.nih.gov/24478066/)
30. Roy K, Hamilton DJ, Fleckenstein JM (2012) Cooperative role of antibodies against heat-labile toxin and the EtpA Adhesin in preventing toxin delivery and intestinal colonization by enterotoxigenic *Escherichia coli*. *Clin Vaccine Immunol* 19: 1603–1608. doi: [10.1128/CAI.00351-12](https://doi.org/10.1128/CAI.00351-12) PMID: [22875600](https://pubmed.ncbi.nlm.nih.gov/22875600/)
31. Roy K, Hamilton D, Ostmann MM, Fleckenstein JM (2009) Vaccination with EtpA glycoprotein or flagellin protects against colonization with enterotoxigenic *Escherichia coli* in a murine model. *Vaccine* 27: 4601–4608. doi: [10.1016/j.vaccine.2009.05.076](https://doi.org/10.1016/j.vaccine.2009.05.076) PMID: [19523914](https://pubmed.ncbi.nlm.nih.gov/19523914/)
32. Mundell DH, Anselmo CR, Wishnow RM (1976) Factors influencing heat-labile *Escherichia coli* enterotoxin activity. *Infect Immun* 14: 383–388. PMID: [9363](https://pubmed.ncbi.nlm.nih.gov/9363/)
33. Alam NH, Ashraf H (2003) Treatment of infectious diarrhea in children. *Paediatric drugs* 5: 151–165. PMID: [12608880](https://pubmed.ncbi.nlm.nih.gov/12608880/)
34. Sjoling A, Wiklund G, Savarino SJ, Cohen DI, Svennerholm AM (2007) Comparative analyses of phenotypic and genotypic methods for detection of enterotoxigenic *Escherichia coli* toxins and colonization factors. *J Clin Microbiol* 45: 3295–3301. doi: [10.1128/JCM.00471-07](https://doi.org/10.1128/JCM.00471-07) PMID: [17687011](https://pubmed.ncbi.nlm.nih.gov/17687011/)
35. Vicente AC, Teixeira LF, Iniguez-Rojas L, Luna MG, Silva L, et al. (2005) Outbreaks of cholera-like diarrhoea caused by enterotoxigenic *Escherichia coli* in the Brazilian Amazon Rainforest. *Trans R Soc Trop Med Hyg* 99: 669–674. doi: [10.1016/j.trstmh.2005.03.007](https://doi.org/10.1016/j.trstmh.2005.03.007) PMID: [15975612](https://pubmed.ncbi.nlm.nih.gov/15975612/)
36. Finkelstein RA, Vasil ML, Jones JR, Anderson RA, Barnard T (1976) Clinical cholera caused by enterotoxigenic *Escherichia coli*. *J Clin Microbiol* 3: 382–384. PMID: [773963](https://pubmed.ncbi.nlm.nih.gov/773963/)
37. Del Canto F, Valenzuela P, Cantero L, Bronstein J, Blanco JE, et al. (2011) Distribution of Classical and Nonclassical Virulence Genes in Enterotoxigenic *Escherichia coli* Isolates from Chilean Children and tRNA Gene Screening for Putative Insertion Sites for Genomic Islands. *J Clin Microbiol* 49: 3198–3203. doi: [10.1128/JCM.02473-10](https://doi.org/10.1128/JCM.02473-10) PMID: [21775541](https://pubmed.ncbi.nlm.nih.gov/21775541/)
38. Harlow E, Lane D, Harlow E (1999) Using antibodies: a laboratory manual. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press. xiv, 495 p. p.
39. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215: 403–410. doi: [10.1006/jmbi.1990.9999](https://doi.org/10.1006/jmbi.1990.9999) PMID: [2231712](https://pubmed.ncbi.nlm.nih.gov/2231712/)
40. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, et al. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular systems biology* 7: 539. doi: [10.1038/msb.2011.75](https://doi.org/10.1038/msb.2011.75) PMID: [21988835](https://pubmed.ncbi.nlm.nih.gov/21988835/)
41. Team RC (2014) R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.
42. Warnes G, Bolker B, Bonebakker L, Gentleman R, Huber W, et al. (2014) gplots: Various R programming tools for plotting data. R package version 2.13.0. ed.
43. Neuwirth E (2011) RColorBrewer: ColorBrewer palettes. R package version 1.0–5 ed.
44. Sheikh A, Luo Q, Roy K, Shaaban S, Kumar P, et al. Contribution of the highly conserved EaeH surface protein to enterotoxigenic *Escherichia coli* pathogenesis. *Infect Immun* in press.
45. Fleckenstein JM, Roy K (2009) Purification of recombinant high molecular weight two-partner secretion proteins from *Escherichia coli*. *Nat Protoc* 4: 1083–1092. doi: [10.1038/nprot.2009.87](https://doi.org/10.1038/nprot.2009.87) PMID: [19707189](https://pubmed.ncbi.nlm.nih.gov/19707189/)
46. Tsang VC, Wilson BC, Maddison SE (1980) Kinetic studies of a quantitative single-tube enzyme-linked immunosorbent assay. *Clinical chemistry* 26: 1255–1260. PMID: [6772340](https://pubmed.ncbi.nlm.nih.gov/6772340/)
47. Harris JA, Roy K, Woo-Rasberry V, Hamilton DJ, Kansal R, et al. (2011) Directed evaluation of enterotoxigenic *Escherichia coli* autotransporter proteins as putative vaccine candidates. *PLoS Negl Trop Dis* 5: e1428. doi: [10.1371/journal.pntd.0001428](https://doi.org/10.1371/journal.pntd.0001428) PMID: [22163060](https://pubmed.ncbi.nlm.nih.gov/22163060/)
48. Dorsey FC, Fischer JF, Fleckenstein JM (2006) Directed delivery of heat-labile enterotoxin by enterotoxigenic *Escherichia coli*. *Cell Microbiol* 8: 1516–1527. doi: [10.1111/j.1462-5822.2006.00736.x](https://doi.org/10.1111/j.1462-5822.2006.00736.x) PMID: [16922869](https://pubmed.ncbi.nlm.nih.gov/16922869/)



49. Sahl JW, Steinsland H, Redman JC, Angiuoli SV, Nataro JP, et al. (2011) A comparative genomic analysis of diverse clonal types of enterotoxigenic *Escherichia coli* reveals pathovar-specific conservation. *Infect Immun* 79: 950–960. doi: [10.1128/IAI.00932-10](https://doi.org/10.1128/IAI.00932-10) PMID: [21078854](https://pubmed.ncbi.nlm.nih.gov/21078854/)
50. Svennerholm AM, Vidal YL, Holmgren J, McConnell MM, Rowe B (1988) Role of PCF8775 antigen and its coli surface subcomponents for colonization, disease, and protective immunogenicity of enterotoxigenic *Escherichia coli* in rabbits. *Infection and immunity* 56: 523–528. PMID: [3276628](https://pubmed.ncbi.nlm.nih.gov/3276628/)
51. McConnell MM, Thomas LV, Willshaw GA, Smith HR, Rowe B (1988) Genetic control and properties of coli surface antigens of colonization factor antigen IV (PCF8775) of enterotoxigenic *Escherichia coli*. *Infect Immun* 56: 1974–1980. PMID: [2456269](https://pubmed.ncbi.nlm.nih.gov/2456269/)
52. Evans DJ Jr., Evans DG (1973) Three characteristics associated with enterotoxigenic *Escherichia coli* isolated from man. *Infect Immun* 8: 322–328. PMID: [4581006](https://pubmed.ncbi.nlm.nih.gov/4581006/)
53. Coster TS, Wolf MK, Hall ER, Cassels FJ, Taylor DN, et al. (2007) Immune response, ciprofloxacin activity, and gender differences after human experimental challenge by two strains of enterotoxigenic *Escherichia coli*. *Infect Immun* 75: 252–259. doi: [10.1128/IAI.01131-06](https://doi.org/10.1128/IAI.01131-06) PMID: [17074855](https://pubmed.ncbi.nlm.nih.gov/17074855/)
54. Chen Q, Savarino SJ, Venkatesan MM (2006) Subtractive hybridization and optical mapping of the enterotoxigenic *Escherichia coli* H10407 chromosome: isolation of unique sequences and demonstration of significant similarity to the chromosome of *E. coli* K-12. *Microbiology* 152: 1041–1054. doi: [10.1099/mic.0.28648-0](https://doi.org/10.1099/mic.0.28648-0) PMID: [16549668](https://pubmed.ncbi.nlm.nih.gov/16549668/)
55. Moriel DG, Bertoldi I, Spagnuolo A, Marchi S, Rosini R, et al. (2010) Identification of protective and broadly conserved vaccine antigens from the genome of extraintestinal pathogenic *Escherichia coli*. *Proc Natl Acad Sci U S A* 107: 9072–9077. doi: [10.1073/pnas.0915077107](https://doi.org/10.1073/pnas.0915077107) PMID: [20439758](https://pubmed.ncbi.nlm.nih.gov/20439758/)
56. Roy K, Hamilton D, Allen KP, Randolph MP, Fleckenstein JM (2008) The EtpA exoprotein of enterotoxigenic *Escherichia coli* promotes intestinal colonization and is a protective antigen in an experimental model of murine infection. *Infect Immun* 76: 2106–2112. doi: [10.1128/IAI.01304-07](https://doi.org/10.1128/IAI.01304-07) PMID: [18285493](https://pubmed.ncbi.nlm.nih.gov/18285493/)
57. Sack RB (2011) The discovery of cholera—like enterotoxins produced by *Escherichia coli* causing secretory diarrhoea in humans. *The Indian journal of medical research* 133: 171–180. PMID: [21415491](https://pubmed.ncbi.nlm.nih.gov/21415491/)
58. Sheikh A, Luo Q, Roy K, Shabaan S, Kumar P, et al. (2014) Contribution of the highly conserved EaeH surface protein to enterotoxigenic *Escherichia coli* pathogenesis. *Infect Immun* 82: 3657–3666. doi: [10.1128/IAI.01890-14](https://doi.org/10.1128/IAI.01890-14) PMID: [24935979](https://pubmed.ncbi.nlm.nih.gov/24935979/)
59. Fleckenstein JM, Sheikh A (2014) Designing vaccines to neutralize effective toxin delivery by enterotoxigenic *Escherichia coli*. *Toxins (Basel)* 6: 1799–1812. doi: [10.3390/toxins6061799](https://doi.org/10.3390/toxins6061799) PMID: [24918359](https://pubmed.ncbi.nlm.nih.gov/24918359/)
60. Fleckenstein J, Sheikh A, Qadri F (2014) Novel antigens for enterotoxigenic *Escherichia coli* vaccines. *Expert review of vaccines* 13: 631–639. doi: [10.1586/14760584.2014.905745](https://doi.org/10.1586/14760584.2014.905745) PMID: [24702311](https://pubmed.ncbi.nlm.nih.gov/24702311/)
61. Crossman LC, Chaudhuri RR, Beatson SA, Wells TJ, Desvaux M, et al. (2010) A commensal gone bad: complete genome sequence of the prototypical enterotoxigenic *Escherichia coli* strain H10407. *Journal of bacteriology* 192: 5822–5831. doi: [10.1128/JB.00710-10](https://doi.org/10.1128/JB.00710-10) PMID: [20802035](https://pubmed.ncbi.nlm.nih.gov/20802035/)
62. Madhavan TP, Steen JA, Hugenholtz P, Sakellaris H (2014) Genome Sequence of Enterotoxigenic *Escherichia coli* Strain B2C. *Genome announcements* 2. doi: [10.1128/genomeA.00247-14](https://doi.org/10.1128/genomeA.00247-14) PMID: [24723709](https://pubmed.ncbi.nlm.nih.gov/24723709/)
63. Scotland SM, McConnell MM, Willshaw GA, Rowe B, Field AM (1985) Properties of wild-type strains of enterotoxigenic *Escherichia coli* which produce colonization factor antigen II, and belong to serogroups other than O6. *J Gen Microbiol* 131: 2327–2333. PMID: [3906040](https://pubmed.ncbi.nlm.nih.gov/3906040/)