Interrogation of a live-attenuated enterotoxigenic *Escherichia coli* vaccine highlights features unique to wild-type infection

Subhra Chakraborty1, Arlo Randall2, Tim J. Vickers3, Doug Molina2, Clayton D. Harro1, Barbara DeNearing1, Jessica Brubaker1, David A. Sack1, A. Louis Bourgeois1, Philip L. Felgner2,4, Xiaowu Liang2, Sachin Mani5, Petra E. Gilmore6, Michael J. Darsley7, David A. Rasko8 and James M. Fleckenstein9

Enterotoxigenic *Escherichia coli* (ETEC) infections are a common cause of severe diarrheal illness in low- and middle-income countries. The live-attenuated ACE527 ETEC vaccine, adjuvanted with double mutant heat-labile toxin (dmLT), affords clear but partial protection against ETEC challenge in human volunteers. Comparatively, initial wild-type ETEC challenge completely protects against severe diarrhea on homologous re-challenge. To investigate determinants of protection, vaccine antigen content was compared to wild-type ETEC, and proteome microarrays were used to assess immune responses following vaccination and ETEC challenge. Although molecular interrogation of the vaccine confirmed expression of targeted canonical antigens, relative to wild-type ETEC, vaccine strains were deficient in production of flagellar antigens, immotile, and lacked production of the EtpA adhesin. Similarly, vaccination ± dmLT elicited responses to targeted canonical antigens, but relative to wild-type challenge, vaccine responses to some potentially protective non-canonical antigens including EtpA and the YghJ metalloprotease were diminished or absent. These studies highlight important differences in vaccine and wild-type ETEC antigen content and call attention to distinct immunologic signatures that could inform investigation of correlates of protection, and guide vaccine antigen selection for these pathogens of global importance.

**INTRODUCTION**

Enterotoxigenic *Escherichia coli* (ETEC) cause substantial morbidity due to diarrheal illness in resource-poor areas of the world where young children are disproportionately affected. In children under five years of age, these pathogens are among the leading causes of moderate-to-severe diarrhea and deaths due to acute diarrheal illness.1,2 ETEC also causes severe illness, clinically indistinguishable from cholera,3–5 and death in older individuals6 and remains the most common cause of travelers’ diarrhea. While oral rehydration therapy and other measures have contributed to a decline in deaths due to diarrheal illness, ETEC have been linked to post-diarrheal sequelae including malnutrition, growth stunting, and impaired cognitive development greatly compounding the post-diarrheal sequelae including malnutrition, growth stunting, and impaired cognitive development.

The ETEC pathovar is defined by the production and effective delivery of heat-stable (ST) and/or heat-labile (LT) enterotoxins to epithelial receptors in the small intestine. In the classical ETEC pathogenesis paradigm, plasmid-encoded colonization factor (CF) or coli surface (CS) antigens facilitate small intestinal colonization.8 Interaction with small intestinal enterocytes leads to toxin-induced alterations in salt and water transport that result in net fluid losses into the intestinal lumen and ensuing watery diarrheal illness ranging from mild to severe and cholera-like.5,9 ETEC infections among young children in endemic regions are thought to result in acquired immunity and a decreasing incidence of infection with age.9 Indeed, controlled human infection studies demonstrate that homologous re-challenge with the ETEC H10407 strain, which encodes CFA/I, results in robust protection against symptomatic ETEC infection.10 However, precise correlates of protection11 have not been established, and the majority of immunologic studies have focused on canonical virulence factors, namely the CF/CS antigens and heat-labile toxin. Nevertheless, recent studies indicate that the repertoire of immune responses following infection extends beyond these classical antigens.12

Because of inherent genetic plasticity of *E. coli*, no canonical virulence factor is universally conserved in ETEC. Therefore, to achieve broad coverage, most ETEC vaccines under development adopt a polyvalent approach targeting multiple CF/CS antigens and LT. ACE527, was developed as a live-attenuated vaccine combining three strains that collectively express CS1, CS2, CS3, CS5, CS6, CFA/I, and the B subunit of LT.13

In recent studies vaccination with ACE527 alone failed to protect against severe diarrhea upon challenge with H10407, while ACE527 adjuvanted with double mutant heat-labile toxin (dmLT)14 afforded significant protection (PE ~ 66%) (Clinical Trials Identifier NCT01739231).15 Comparatively, H10407 challenge elicits nearly complete protection against homologous re-challenge.10 To comprehensively assess the adaptive immune response to vaccination we examined the antigen continen of
ACE527 and used ETEC protein microarrays to examine antibody responses to the vaccine ± dmLT. These responses were then compared to vaccine placebo controls and to challenge with H10407 to profile potential benchmarks of protection.

RESULTS
Genomic and proteomic characterization of ACE527 vaccine strains

Whole genome DNA sequence data was used to verify the genotypes of wild-type parental isolates and the engineering of the ACE527 strains (summarized in supplementary table 1). As anticipated, the three engineered vaccine strains, ACAM2022, ACAM2025, and ACAM2027, collectively encoded the full complement of six CF/CS antigens (CFA/I, CS1, CS2, CS3, CS5, and CS6), and each encoded both the B subunit of heat-labile toxin, as well as the type II secretion system (T2SS) responsible for export of both LT and YghJ (supplementary table 3). Two of the three parental strains, WS1858B and WS3504D were noted to contain the plasmid etp84C locus which encodes the two-partner secretion system responsible for production and export of the EtpA adhesin, however analysis of the corresponding attenuated vaccine derivatives, ACAM2025 and ACAM2027 revealed that this locus had been lost in the vaccine strain construction (Fig. 1a, b). Similarly, the eatA gene which encodes a serine protease autotransporter protein that degrades MUC2 mucin, was present in each of the parents but absent from ACAM2025 (Fig. 1a, b).

The growth of each parental isolate and corresponding ACAM live-attenuated bacteria was compared to H10407. In general, the growth of the parental wild-type strains paralleled that of H10407.
while the growth of ACAM vaccine strains lagged slightly, potentially reflecting the combined effect of OmpC and OmpF mutations on growth. In contrast to recent studies of the H10407 proteome where the major protein subunit of flagella, flagellin (Flc, serotype H11), and the flagellar hook protein, FlgE were present in abundance, FlIC was not universally detected in the ace527 proteomes, and we were unable to detect the flagellar hook protein, FlgE (Fig. 1d). The wild-type strain of H10407 challenge strain included motility, production of key flagellar antigens, and other immunogenic proteins including the secreted EtpA blood group A lectin-adhesin.

Differential immune responses to ETEC antigens after vaccination and challenge
Recent controlled human infection studies with the H10407 strain of ETEC demonstrated that previously naive volunteers mounted strong mucosal antibody responses to classical vaccine targets including CFA/I and heat-labile toxin as well as a select number of molecules that have not been traditionally been targeted in vaccines. Here we set out to examine immune response to the antigenic proteins targeted in the present studies. As this serotype was not included in the vaccine, itself, and responses to H-serotype specific regions of the flagellins represented in the vaccine were noticeably absent (supplementary fig. 2), the responses to Flc H11 and other flagellins may reflect recognition of the non-serotype specific highly conserved alpha helical regions comprising the amino and carboxy terminal regions of flagellin molecules (supplementary fig. 3).

We likewise observed differential responses to two secreted non-canonical antigens, the passenger domain of the EatA autotransporter protein, and YghJ a metalloprotease secreted by the type 2 secretion system. Both proteins were recognized following vaccination with ace527 with 9/12 subjects responding to the EatAp and 6/12 responding to YghJ. These responses were enhanced slightly in the dmLT cohort (Table 1).

Immune responses to non-canonical proteins predominated after challenge
In naive subjects and in vaccine recipients challenged with H10407, we noted enhanced responses to a number of proteins that were not recognized following vaccination. Among the most striking differential responses following challenge compared to vaccination were those to flagellar proteins. 9/10 placebo vaccinated subjects responded to both full length and the serotype specific region (AA174-399) of Flc H11 (Table 3). Likewise, 6/10 volunteers responded to the FlgE flagellar hook protein, a response that was absent in vaccinees prior to challenge, consistent with the lack of motility in the ace527 strains and the absence of FlgE in the ace527 proteomes. In addition to flagellin, three secreted antigens YghJ (SslE), EatA, and EtpA, not currently targeted in classical ETEC vaccine approaches have been shown to be immunogenic in humans and associated with protection in animal models. Although we observed responses to YghJ following vaccination with either ace527 or ace527 + dmLT, the response to this protein was
significantly increased following challenge with H10407 (Fig. 3b). Similarly, we observed the most robust responses to the EatA passenger domain following challenge with H10407 (Fig. 3c). In contrast, in keeping with the absence of EtpA in the three vaccine strains, vaccinees did not recognize EtpA following immunization, while both the placebo group and the vaccinated subjects mounted robust responses to EtpA upon challenge with H10407 (Fig. 3d, Table 4).

Table 1. Most differentially reactive purified protein antigens ALS IgA, responses day 0 to day 7 following ACE527 or ACE527 + dmLT

| Antigen          | Δa | Frequencyb | p    | Antigen          | Δa | Frequencyb | p    |
|------------------|----|------------|------|------------------|----|------------|------|
| CFA/I            | 4.34 | 1.00      | 5.4×10^-6 | CFA/I            | 4.81 | 1.00      | 8.4×10^-7 |
| PCF071           | 3.66 | 1.00      | 4.1×10^-6 | PCF071           | 3.08 | 1.00      | 5.2×10^-7 |
| LT-B             | 2.04 | 0.75      | 3.2×10^-3 | EatAp            | 2.74 | 0.85      | 1.6×10^-4 |
| EatAp            | 1.62 | 0.75      | 7.8×10^-3 | LT-B             | 2.90 | 0.77      | 7.7×10^-4 |
| CS2              | 1.41 | 0.58      | 1.5×10^-2 | CS2              | 1.09 | 0.62      | 2.0×10^-2 |
| YghJ             | 0.93 | 0.50      | 1.2×10^-2 | YghJ             | 0.80 | 0.62      | 3.0×10^-2 |
| CS17             | 0.82 | 0.50      | 3.4×10^-2 | CS17             | 1.14 | 0.54      | 1.7×10^-2 |
| CstH (CS3 pilin) | 1.04 | 0.33      | 7.0×10^-2 | CstH (CS3 pilin) | 1.34 | 0.31      | 3.0×10^-2 |

*aΔmean refers to difference in mean values at day 0 and day 7 with respect to the first vaccination
bFrequency refers to the proportion of subjects with at least a 50% increase in normalized signal intensity between day 0 and day 7 after the first vaccination
cPurified fimbriae
DISCUSSION

A systematic appraisal of immune response to candidate vaccines can aid in vaccine optimization and identification of mechanistic correlates of protection. Here, we combined genomic and proteomic interrogation of a live-attenuated vaccine with immunoproteomic analysis following vaccination and experimental human challenge with ETEC H10407, an extensively characterized strain originally obtained from a patient with severe cholera-like diarrheal illness. Previous studies demonstrating that challenge with wild-type ETEC H10407 bacteria affords virtually complete protection against severe diarrhea on subsequent re-challenge10 provide important benchmarks for comparison of candidate vaccines. Importantly, the studies reported here highlight a number of features that distinguish the vaccine strains from the challenge strain and the corresponding immune responses.

First, although the wild-type isolates used to construct the vaccine had previously been serotyped for flagellar (H) antigens, we found that production of flagellin, the major subunit of flagella, by both the parent and the vaccine strains was deficient, and none were motile upon testing. Similarly, contrasting with our earlier analysis of the H10407 challenge strain,12 we were unable to detect FlgE, the flagellar hook protein, in either the parental or vaccine strains, suggesting that despite the presence of the genes required for assembly, early steps involved in biosynthesis of flagella are deficient33,34 in the isolates selected for engineering of ACE527. These deficiencies may have negatively impacted the protection afforded by the vaccine in a number of ways including loss of TLR5 mediated stimulation of innate immunity, and the potent adjuvant activity of flagellin. While H serotypes of ETEC vary considerably,35 motility is a highly conserved virulence characteristic, and the lack of motility of the vaccine strains likely resulted in suboptimal antigen delivery to sampling sites within the intestinal mucosa.36 In addition, highly conserved regions of flagellin that flank serotype specific regions of the molecule may not only serve as potent stimuli of innate immunity, but contain cross-protective epitopes.29 Intriguingly, addition of dmLT as a mucosal adjuvant significantly enhanced the response to multiple flagellins independent of whether the specific serotype was present in the vaccine.

The genomes of the three vaccine strains and parents also revealed that the etpBAC locus, which encodes the two-partner secretion system responsible for production and export of the

![Fig. 3](image-url)
EtpA adhesin\textsuperscript{18} was present in two of the three parental strains but missing from the vaccine altogether, likely the result of engineering the vaccine strains to remove the plasmid-encoded toxins. EtpA is a high molecular weight glycoprotein secreted by ETEC that appears to facilitate bacterial adhesion by serving as a molecular bridge between flagella and the enterocyte surface where it binds to N-acetylgalactosamine (GalNAc) residues particularly when they are presented as the terminal glycan on human A blood group antigen. Interestingly, human challenge studies with H10407 demonstrate that individuals with A blood group are significantly more likely to experience severe diarrhea when challenged with this EtpA-producing strain.\textsuperscript{22} These and other recent studies\textsuperscript{12} demonstrate that EtpA is highly immunogenic and recognized by the majority of volunteers upon challenge with wild-type H10407, in distinct contrast to those immunized with ACE527 + dmLT. Although EtpA is required for optimal delivery of both LT and ST enterotoxins, and is protective against ETEC infection in a murine model, further studies are needed to assess its role as a potential protective antigen in a murine model, optimal delivery of both LT and ST enterotoxins, and is protective against ETEC infection in a murine model, further studies are needed to assess its role as a potential protective antigen in a murine model.

Protective immunity to ETEC is likely complex and may represent the cumulative response to \textit{E. coli} core proteins, classical vaccine antigens, and more recently discovered proteins. Although the present studies, based on small numbers of human volunteers, do not permit us to establish clear mechanistic correlates of protection, they highlight the utility of combined genomic, proteomic and immunoproteomic platforms in interpreting the response to live-attenuated vaccines.

Comparison of the vaccine antigen content and immunologic responses to those observed with wild-type infection could inform the design, optimization, and engineering of next-generation ETEC vaccines to enhance protective efficacy. Moreover, the platforms used in the present studies could be generalized to interrogation of live-attenuated vaccines for other important pathogens.

**METHODS**

Bioinformatics and comparative genomics of ETEC isolates

To select candidate genes for protein expression, we analyzed the previously sequenced genomes of three parental ETEC isolates WS\_1858B, WS\_2773E, and WS3504D, used in the construction of ACE527,\textsuperscript{13} the genome of \textit{E. coli} H10407\textsuperscript{99} the challenge strain used in these studies, and the genomes of a diverse group of clinical isolates. Data describing these strains is presented in supplementary table 1. The genome content of these isolates was compared using Large-Scale BLAST Score Ratio Analysis\textsuperscript{40} and encoded products having a signal for potential secretion to the surface were identified using PSORT,\textsuperscript{41} TMHMM,\textsuperscript{42} and SignalP.\textsuperscript{43} The resulting dataset includes 800 antigens identified in H10407, 157 antigens present in one or more of the three isolates of the ACE527 vaccine lacking in the H10407 genome, and an additional 4168 features identified in comparative analysis of 207 clinical ETEC isolates (supplementary data s1).\textsuperscript{44,45} The selected 4168 gene features were present in more than 40\% of the ETEC isolates and were not present in the genomes of three common \textit{E. coli} commensal isolates \textit{E. coli} HS (GenBank Accession number NC\_009800), \textit{E. coli} K-12 (GenBank Accession number NC\_007791), \textit{E. coli} ATCC8739 (GenBank Accession number NC\_010468) or \textit{E. coli} IA1 (GenBank Accession number NC\_011741). Gene identifiers, DNA and predicted peptide sequences, and the isolates used as the template for isolation are included in supplementary_data_s1. Informatically selected features encompassed known ETEC antigens including the A and B subunits of heat-labile toxin (LT-A, and LT-B), CFs, the EatA\textsuperscript{19} serine protease, the EtpA adhesin,\textsuperscript{17} and the metalloprotease YghJ\textsuperscript{46} in addition to conserved and serotype specific regions of flagellin molecules represented in the challenge and vaccine strains.

Rapid Annotation using Subsystem Technology (RASTtk v 1.3.0, http://rast.nmpdr.org)\textsuperscript{47} was used to query completed ACE527 genomes for specific virulence factors. Multiplex PCR was used to verify the toxin profiles of parent and ACE527 vaccine strains using primers for estH, estP, and eatB encoding the STh, STp, and the B subunit of LT, respectively.\textsuperscript{48} PCR was also used to verify the presence or absence of the eatA and etpA genes\textsuperscript{49} (supplementary table 2). SerotypeFinder v 2.0 (https://cge.cbs.dtu.dk/services/SerotypeFinder/) was used to assign H serotypes from whole genome sequence data.\textsuperscript{50}

Microarray construction

Genes encoding candidate ETEC surface-expressed antigens were amplified by PCR, cloned into pXlT7,\textsuperscript{12,51} and expressed in a cell-free in vitro transcription—translation (IVTT) system as previously described.\textsuperscript{12} Each IVTT protein included 5\' polyhistidine (HIS) and 3\' hemaggulutinin (HA) epitopes. After robotic microarray printing onto nitrocellulose-coated glass slides, random slides were validated by probing with anti-His (mouse monoclonal clone HIS-1, Sigma-Aldrich, H1029-100UL) and anti-HA (rat monoclonal to HA peptide PVYDPDPDYA, clone 3F10, Sigma Aldrich, 11867423001) followed by fluorescent secondary antibodies.

Recombinant antigens including the EtpA adhesin,\textsuperscript{52} the passenger domain of EatA,\textsuperscript{19} YghJ,\textsuperscript{46} antigen 43,\textsuperscript{53} EaeH,\textsuperscript{54} LT-A and LT-B, and flagellin (FlIC)\textsuperscript{55} subunits were produced at Washington University in Saint Louis as
IgA (spots for IgA secondary antibody (without the dmLT adjuvant (25 μg per dose). At 6–7 months following the primary immunization, placebo and vaccine ± dmLT immunization subjects were rescreened for eligibility. Those enrolled were challenged with ~2 x 10^9 colony forming units (cfu) of the H10407 ETEC challenge strain after an overnight fast, and assessed in an inpatient setting for development of ETEC-associated diarrhea and other signs and symptoms of enteric illness. The challenge phase of this study included 13 volunteers who received three oral doses of ACE527, 13 volunteers who received three oral doses ACE527 adjuvanted with 25 μg of dmLT, and 10 placebo recipients (supplementary dataset s1).

### Table 3. Differentially reactive IVTT antigens following ETEC H10407 challenge of naive subjects

| Antigen                  | Δ mean | Frequency | p-value  |
|--------------------------|--------|-----------|----------|
| Flic H11                 | 6.06   | 0.90      | 5.8 x 10^-5 |
| Antigen 43               | 3.77   | 0.70      | 2.4 x 10^-3 |
| Flic H11 AA 174,299 (H11 serotype specific) | 2.88 | 0.90 | 1.4 x 10^-4 |
| EatA passenger domain AA 533–1045 | 2.68 | 0.70 | 1.1 x 10^-2 |
| YghJ metalloprotease AA 695–1493 | 2.50 | 0.80 | 1.8 x 10^-3 |
| FlgE flagellar hook      | 1.51   | 0.60      | 2.8 x 10^-2 |
| OmpW                     | 1.45   | 0.70      | 5.4 x 10^-3 |
| YghJ metalloprotease AA 1–800 | 1.19 | 0.70 | 4.3 x 10^-3 |
| Aida-I family autotransporter YfaL | 1.09 | 0.60 | 1.4 x 10^-2 |
| EatA serine protease AA 1–1364 | 0.99 | 0.40 | 3.9 x 10^-2 |

*Δ mean refers to difference in mean values at day 0 prior to challenge and day 7 post challenge

Frequency refers to the proportion of subjects with at least a 50% increase in normalized signal intensity between day 0 and day 7 post challenge

### Table 4. Differentially reactive purified protein antigens following ETEC H10407 challenge of naive subjects

| Antigen                  | Δ mean | Frequency | p-value  |
|--------------------------|--------|-----------|----------|
| YghJ                     | 5.40   | 0.90      | 1.1 x 10^-4 |
| CS14                     | 4.85   | 0.90      | 1.8 x 10^-5 |
| EtpA amino-terminal domain | 4.08 | 0.90 | 1.6 x 10^-3 |
| CS2                      | 2.74   | 0.90      | 4.7 x 10^-3 |
| CS3                      | 2.37   | 0.90      | 2.5 x 10^-3 |
| CS1                      | 2.29   | 0.80      | 5.6 x 10^-3 |
| Pcf071                   | 2.29   | 0.80      | 1.9 x 10^-3 |
| EatAp                    | 2.64   | 0.70      | 5.7 x 10^-3 |
| CS17                     | 2.37   | 0.70      | 1.2 x 10^-2 |
| EtpA full length         | 1.34   | 0.70      | 8.1 x 10^-3 |

*Δ mean refers to difference in mean values at day 0 prior to challenge and day 7 post challenge

Frequency refers to the proportion of subjects with at least a 50% increase in normalized signal intensity between day 0 and day 7 post challenge

Microarray antigen content

The microarrays used in this study were comprised of IVTT expressed proteins selected above and purified proteins representing known ETEC antigens (n = 38) (Supplementary dataset s1). Also included on the array are IVTT control spots (n = 28), positive control spots for human IgG (n = 16), positive control spots for human IgA (n = 16), and positive controls for human lysozyme (n = 16).

Vaccination with ACE527 and controlled human infection studies

Samples analyzed in the present study were derived from Phase I/2b trial of ACE527 conducted in healthy human volunteers at the Centre for Immunization Research at Johns Hopkins University School of Public Health (clinical trial number NCT01739231). In this earlier trial, subjects were randomly assigned to three groups: the placebo group or groups that received three doses of ACE527 (~10^10 cfu/dose) given orally with or without the dmLT adjuvant (25 μg per dose). At 6–7 months following the primary immunization, placebo and vaccine ± dmLT immunization subjects were re-screened for eligibility. Those enrolled were challenged with ~2 x 10^9 colony forming units (cfu) of the H10407 ETEC challenge strain after an overnight fast, and assessed in an inpatient setting for development of ETEC-associated diarrhea and other signs and symptoms of enteric illness. The challenge phase of this study included 13 volunteers who received...
compare normalized signal data with a Kruskal–Wallis test for multiple comparisons.

Human studies
Use of the archived biospecimens, and data in the present study was performed with the approval of the Institution Review Boards of Johns Hopkins University School of Medicine and Washington University School of Medicine.

Reporting summary
Further information on research design is available in the Nature Research
Reporting Summary linked to this article.

DATA AVAILABILITY
Proteomics data have been uploaded to the ProteomeXchange56,57 database accessible at http://www.proteomexchange.org/ via accession number PXD014724. Protein Microarray data have been uploaded to the GEO database8,9,9 https://www.ncbi.nlm.nih.gov/geo/ under accession number GSE134792. Source data for the figures and tables is also provided in the supplementary information.

ACKNOWLEDGEMENTS
Research reported in this publication was supported by PATH: funding from National Institute of Allergy and Infectious Diseases (NIAID) of the National Institutes of Health (NIH) under Award Numbers R01AI089894, R01AI126687 (jmf), the Washington University School of Medicine and the Department of Veterans Affairs (S01BX001469, jmf). This work was supported by the Intramural Research Program of the NIH, NIAID, and the National Institute of Allergy and Infectious Diseases. The authors thank Edward Oakas (Walter Reed Army Institute of Research), Stephen Savarino, and Stephen Poole (Naval Medical Research Center) for providing select purified antigens that were used in production of the arrays.

AUTHOR CONTRIBUTIONS
Conceived and designed experiments: S.C., J.M.F., S.M., A.L.B., D.A.S., P.L.F., X.L., H.W. and M.J.D.; performed experiments: T.V., D.M., C.D.H., B.D., J.B. and P.E.G. Analyzed the data: A.R., R.R.T., P.E.G., J.M.F., D.A.R., A.L.B., A.R. and S.C. Each of the listed co-authors made substantial contributions to the work through design and conception, and/or acquisition, analysis, and interpretation of the data.

ADDITIONAL INFORMATION
Supplementary information accompanies the paper on the npj Vaccines website (https://doi.org/10.1038/s41551-019-0131-7).

Competing interests: The corresponding author (J.M.F.) is listed as an inventor on patent 8,323,668 related to the EtpA protein. X.L. and A.R. received grants from the Bill and Melinda Gates Foundation during the conduct of the study.

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

REFERENCES
1. Khalil, I. A. et al. Morbidity and mortality due to shigellosis and enterotoxigenic Escherichia coli diarrhea: the Global Burden of Disease Study 1990–2016. Lancet Infect. Dis. 18, 1229–1240 (2018).
2. Kotsloff, K. L. et al. Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEIMS): a prospective, case-control study. Lancet 382, 209–222 (2013).
3. Sack, R. B. et al. Enterotoxigenic Escherichia coli isolated from patients with severe cholera-like disease. J. Infect. Dis. 123, 378–385 (1971).
4. Finkelstein, R. A., Vasil, M. L., Jones, J. R., Anderson, R. A. & Barnard, T. Clinical cholera caused by enterotoxigenic Escherichia coli. J. Clin. Microbiol. 3, 382–384 (1976).
5. Vicente, A. C. et al. Outbreaks of cholera-like diarrhoea caused by enterotoxigenic Escherichia coli in the Brazilian Amazon Rainforest. Trans. R. Soc. Trop. Med Hyg. 99, 669–674 (2005).
6. Lamberti, L. M., Bourgeois, A. L., Fischer Walker, C. L., Black, R. E. & Sack, D. Estimating diarrheal illness and deaths attributable to Shigella and enterotoxigenic Escherichia coli among older children, adolescents, and adults in South Asia and Africa. PLoS Negl. Trop. Dis. 8, e2705 (2014).
7. Anderson, J. D. et al. Burden of enterotoxigenic Escherichia coli and shigellosis non-fatal diarrhoeal infections in 79 low-income and lower middle-income countries: a modelling analysis. Lancet Glob. Health 7, e321–e330 (2019).
8. Fleckenstein, J. M. et al. Molecular mechanisms of enterotoxigenic Escherichia coli infection. Microbes Infect. 12, 89–98 (2010).
9. Qadri, F. et al. Disease burden due to enterotoxigenic Escherichia coli in the first 2 years of life in an urban community in Bangladesh. Infect. Immun. 75, 3961–3968 (2007).
10. Harro, C. et al. Refinement of a human challenge model for evaluation of enterotoxigenic Escherichia coli vaccines. Clin. Vaccine Immunol. 18, 1719–1727 (2011).
11. Holmgren, J. et al. Correlates of protection for enteric vaccines. Vaccine 35, 3335–3363 (2017).
12. Chakraborty, S. et al. Human experimental challenge with enterotoxigenic Escherichia coli elicits immune responses to canonical and novel antigens relevant to vaccine development. J. Infect. Dis. 218, 1436–1446 (2018).
13. Turner, A. K. et al. 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 (2011).
14. Norton, E. B., Lawson, L. B., Freytag, L. C. & Clements, J. D. Characterization of a mutant Escherichia coli heat-labile toxin, LT(R192G/L211A), as a safe and effective oral adjuvant. Clin. Vaccine Immunol. 18, 546–551 (2011).
15. Harro, C. et al. Live attenuated enterotoxigenic Escherichia coli (ETEC) vaccine with dmLT adjuvant protects human volunteers against virulent experimental ETEC challenge. Vaccine 37, 1978–1986 (2019).
16. Tauschek, M., Gorrell, R. J., Strugnell, R. A. & Robins-Browne, R. M. Identification of a protein secretory pathway for the secretion of heat-labile enterotoxin by an enterotoxigenic strain of Escherichia coli. Proc. Natl Acad. Sci. USA 99, 7066–7071 (2002).
17. Baldi, D. L. et al. The type II secretion system and its ubiquitous lipoprotein substrate, SseE, are required for biofilm formation and virulence of enteropathogenic Escherichia coli. Infect. Immun. 80, 2042–2052 (2012).
18. Fleckenstein, J. M., Roy, K., Fischer, J. F. & Burkit, M. Identification of a two-partner secretion locus of enterotoxigenic Escherichia coli. Infect. Immun. 74, 2245–2258 (2006).
19. Kumar, P. et al. EtaA, an immunogenic protective antigen of enterotoxigenic Escherichia coli, degrades intestinal mucin. Infect. Immun. 82, 500–508 (2014).
20. Kamaraju, M. et al. OmpC and OmpF are required for growth under hyper-osmotic stress above pH 8 in Escherichia coli. Lett. Appl Microbiol. 42, 195–201 (2006).
21. Guillier, M. & Gottesman, S. Remodelling of the Escherichia coli outer membrane by two small regulatory RNAs. Mol. Microbiol. 59, 231–247 (2006).
22. Kumar, P. et al. Enterotoxigenic Escherichia coli blood group A interactions intensify diarrheal severity. J. Clin. Invest. 128, 3298–3311 (2018).
23. Qadri, F. et al. Antigen-specific immunoglobulin A antibodies secreted from circulating B cells are an effective marker for recent local immune responses in patients with cholera: comparison to antibody-secreting cell responses and other immunological markers. Infect. Immun. 71, 4808–4814 (2003).
24. Chang, H. S. & Sack, D. A. Development of a novel in vitro assay (ALS assay) for evaluation of vaccine-induced antibody secretion from circulating mucosal lymphocytes. Clin. Diagn. Lab Immunol. 8, 482–488 (2001).
25. Anantha, R. P. et al. Evolutionary and functional relationships of colonization factor antigen i and other class 5 adhesive fibriniae of enterotoxigenic Escherichia coli. Infect. Immun. 72, 7190–7201 (2004).
26. Sakellari, H., Penumalli, V. R. & Scott, J. R. The level of expression of the minor pilin subunit, CooD, determines the number of C51 pili assembled on the cell surface of Escherichia coli. J. Bacteriol. 181, 1694–1697 (1999).
27. Harro, C. et al. 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. 18, 2118–2127 (2011).
28. Nesta, B. et al. SfEı elicits functional antibodies that impair in vitro mucinase activity and in vivo colonization by both intestinal and extraintestinal Escherichia coli strains. PLoS Pathog. 10, e1004124 (2014).
29. Roy, K., Hamilton, D., Ostmann, M. M. & Fleckenstein, J. M. Vaccination with EtpA glycoprotein or flagellin protects against colonization with enterotoxigenic Escherichia coli in a murine model. Vaccine 27, 4601–4608 (2009).
30. Roy, K., Hamilton, D., Allen, K. P., Randolph, M. P. & Fleckenstein, J. M. 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 (2008).
31. Roy, K., Hamilton, D. J. & Fleckenstein, J. M. Cooperative role of antibodies against heat-labile toxin and the EtpA Adhesin in preventing toxin delivery and intestinal colonization by enterotoxigenic Escherichia coli. Clin. Vaccin. Immunol. 19, 1603–1608 (2012).
32. Evans, D. J. Jr. & Evans, D. G. Three characteristics associated with enterotoxigenic Escherichia coli isolated from man. Infect. Immun. 8, 322–328 (1973).
33. Chilcott, G. S. & Hughes, K. T. Coupling of flagellar gene expression to flagellar assembly in Salmonella enterica serovar typhimurium and Escherichia coli. Microbiol. Mol. Biol. Rev. 64, 694–708 (2000).
34. Chevance, F. F. & Hughes, K. T. Coordinating assembly of a bacterial macromolecular machine. Nat. Rev. Microbiol. 6, 455–465 (2008).
35. Wolf, M. K. Occurrence, distribution, and associations of O and H serogroups, colonization factor antigens, and toxins of enterotoxigenic Escherichia coli. Clin. Microbiol. Rev. 10, 569–584 (1997).
36. Dorsey, F. C., Fischer, J. F. & Fleckenstein, J. M. Directed delivery of heat-labile enterotoxin by enterotoxigenic Escherichia coli. Cell Microbiol. 8, 1516–1527 (2006).
37. Roy, K. et al. Enterotoxigenic Escherichia coli EtpA mediates adhesion between flagella and host cells. Nature 457, 594–598 (2009).
38. Kumar, P. et al. Dynamic interactions of a conserved enterotoxigenic Escherichia coli adhesin with intestinal mucins govern epithelial engagement and toxin delivery. Infect. Immun. 84, 3608–3617 (2016).
39. Crossman, L. C. et al. A commensal gone bad: complete genome sequence of the prototypical enterotoxigenic Escherichia coli strain H10407. J. Bacteriol. 192, 5822–5831 (2010).
40. Sahl, J. W., Caporaso, J. G., Rasko, D. A. & Keim, P. The large-scale blast score ratio (LS-BSR) pipeline: a method to rapidly compare genetic content between bacterial genomes. PeerJ 2, e332 (2014).
41. Nakai, K. & Horton, P. PSORT: a program for detecting sorting signals in pro- and eukaryotic proteins and predicting their subcellular localization. Trends Biochem Sci. 24, 34–36 (1999).
42. Sonnhammer, E. L., von Heijne, G. & Krogh, A. A hidden Markov model for predicting transmembrane helices in protein sequences. Proc. Int. Conf. Intell. Syst. Mol. Biol. 6, 175–182 (1998).
43. Petersen, T. N., Brunak, S., von Heijne, G. & Nielsen, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat. Methods 8, 785–786 (2011).
44. Sahl, J. W. et al. Insights into enterotoxigenic Escherichia coli diversity in Bangladesh utilizing genomic epidemiology. Sci. Rep. 7, 3402 (2017).
45. Sahl, J. W. et al. Examination of the enterotoxigenic Escherichia coli population structure during human infection. mBio 6, e00501–15 (2015). https://doi.org/10.1128/mBio.00501-15.
46. Luo, Q. et al. Enterotoxigenic Escherichia coli secretes a highly conserved mucin-degrading metalloprotease to effectively engage intestinal epithelial cells. Infect. Immun. 82, 509–521 (2014).
47. Brettin, T. et al. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. Sci. Rep. 5, 8365 (2015).
48. Sjöling, A., Wiklund, G., Savarino, S. J., Cohen, D. I. & Svennerholm, A. M. Comparative analyses of phenotypic and genotypic methods for detection of enterotoxigenic Escherichia coli toxins and colonization factors. J. Clin. Microbiol. 45, 3295–3301 (2007).
49. Del Canto, F. et al. 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 (2011).
50. Joensen, K. G., Tetzschner, A. M., Iguchi, A., Aarestrup, F. M. & Schuetz, F. Rapid and easy in silico serotyping of Escherichia coli isolates by use of whole-genome sequencing data. J. Clin. Microbiol. 53, 2410–2426 (2015).
51. Davies, D. H. et al. Profiling the humoral immune response to infection by using proteome microarrays: high-throughput vaccine and diagnostic antigen discovery. Proc. Natl Acad. Sci. USA 102, 547–552 (2005).
52. Fleckenstein, J. M. & Roy, K. Purification of recombinant high molecular weight two-partner secretion proteins from Escherichia coli. Nat. Protoc. 4, 1083–1092 (2009).
53. Harris, J. A. et al. Directed evaluation of enterotoxigenic Escherichia coli auto-transporter proteins as putative vaccine candidates. PLoS Negl. Trop. Dis. 5, e1428 (2011).
54. Sheikh, A. et al. Contribution of the highly conserved EaeH surface protein to enterotoxigenic Escherichia coli pathogenesis. Infect. Immun. 82, 3657–3666 (2014).
55. Luo, Q. et al. Conservation and immunogenicity of novel antigens in diverse isolates of enterotoxigenic Escherichia coli. PLoS Negl. Trop. Dis. 9, e003446 (2015).
56. Vizcaino, J. A. et al. ProteomeXchange provides globally coordinated proteomics data submission and dissemination. Nat. Biotechnol. 32, 223–226 (2014).
57. Deutsch, E. W. et al. The ProteomeXchange consortium in 2017: supporting the cultural change in proteomics public data deposition. Nucleic Acids Res. 45, D1100–D1106 (2017).
58. Edgar, R., Domrachev, M. & Lash, A. E. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res. 30, 207–210 (2002).
59. Barrett, T. et al. NCBI GEO: archive for functional genomics data sets—update. Nucleic Acids Res. 41, D991–D995 (2013).

© The Author(s) 2019

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/.