bioRxiv preprint doi: https://doi.org/10.1101/2020.10.20.348110; this version posted October 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Contribution of noncanonical antigens to virulence and adaptive immunity in human infection with enterotoxigenic *E. coli*

Kuhlmann FM[1], Laine RO[1], Afrin S[1], Nakajima R[2], Akhtar M[3], Vickers T[1], Parker K[1†], Nizam NN[3], Grigura V[1], Goss CW[4], Felgner PL[2], Rasko DA[5], Qadri F[3], and Fleckenstein, JM[1,6].

[1]Department of Internal Medicine, Division of Infectious Diseases, Washington University, St. Louis, Missouri, United States of America

[2]Vaccine R & D Center, University of California, Irvine, Irvine, CA, USA

[3]Infectious Diseases Division, International Centre for Diarrhoeal Disease Research, Bangladesh (icddr,b), Dhaka, Bangladesh

4Division of Biostatistics, Washington University, St. Louis, Missouri, United States of America

[5]Department of Microbiology and Immunology and Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore, Maryland, United States of America

[6]Medicine Service, Veterans Affairs Medical Center, St. Louis, Missouri, United States of America

[†] current author address:

Corresponding Author:
James M. Fleckenstein
Washington University School of Medicine
Department of Medicine, Division of Infectious Diseases
660 S. Euclid Drive, Box 8051
Saint Louis, Missouri, USA 63110
jfleckenstein@wustl.edu

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.20.348110; this version posted October 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abstract

Enterotoxigenic *E. coli* (ETEC) contribute significantly to the substantial burden of infectious diarrhea among children living in low and middle income countries. In the absence of a vaccine for ETEC, children succumb to acute dehydration as well as non-diarrheal sequelae related to these infections including malnutrition. The considerable diversity of ETEC genomes has complicated canonical vaccine development approaches focused on a subset of antigens known as colonization factors (CFs). To identify additional conserved immunogens, we mined genomic sequences of 89 ETEC isolates, bioinformatically selected potential surface-exposed pathovar-specific antigens conserved in more than 40% of the genomes (n=118), and assembled the representative proteins onto microarrays, complemented with known or putative colonization factor subunit molecules (n=52), and toxin subunits to interrogate samples from individuals with acute symptomatic ETEC infections. Surprisingly, in this open-aperture approach, we found that immune responses were largely constrained to a small number of antigens including individual colonization factor antigens and EtpA, an extracellular adhesin. In a Bangladeshi cohort of naturally infected children < 2 years of age, both EtpA and a second noncanonical antigen, EatA, elicited significant serologic responses that were associated with protection from symptomatic illness. In addition, children infected with ETEC isolates bearing either *etpA or eatA* genes were significantly more likely to develop symptomatic disease. These studies support a role for more recently discovered noncanonical antigens in virulence and the development of adaptive immune responses during ETEC infections, findings that may inform vaccine design efforts to complement existing approaches.

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.20.348110; this version posted October 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Introduction

Enterotoxigenic *Escherichia coli* (ETEC) are one of the commonest causes of childhood diarrhea, accounting for 100s of millions of cases annually (1). This high burden of disease contributes a substantial risk of increased childhood morbidity and mortality (2), (3, 4). Repeated diarrheal infections, including those caused by ETEC, lead to the development of growth stunting and environmental enteropathy, which are life-long consequences of these enteric infections (5). Therefore, preventative efforts, including vaccination, could have a tremendous impact on global health (6). Despite the lack of a licensed ETEC vaccine, two important lines of evidence suggest ETEC vaccine development is feasible. First, controlled human infection models (CHIM) demonstrate that protective immunity develops following ETEC challenge (7, 8). In addition, the frequency of symptomatic infections in young children living in endemic regions wanes substantially with age (9, 10), suggesting that natural infections afford subsequent protection.

ETEC biology, and the extraordinary genetic plasticity of *E. coli*, has complicated the development of a broadly protective vaccine. Canonical approaches have focused primarily on surface features known as colonization factors (CFs) or CS antigens. However, the structural and antigenic diversity of these targets has proved challenging (11). Although toxoids that can elicit neutralizing antibodies against the heat-labile (LT) (12) and heat-stabile toxins (ST) (13) that define the ETEC pathovar are currently under development (14, 15), it is not yet clear whether these alone will afford sufficient, long-lasting protection.

While the ETEC pathovar exhibits high genetic diversity, the recent availability of multiple, genomic sequences from globally diverse ETEC affords the ability to apply reverse vaccinology approaches to the identification of conserved, surface-expressed antigens (16, 17). In addition, microarray-based profiling of immune responses in human volunteers to ETEC challenge has recently highlighted non-canonical antigens recognized during controlled experimental infection.(7, 18).

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.20.348110; this version posted October 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Application of these approaches to antigen discovery has reinforced the importance of several surface-expressed molecules common to the ETEC pathovar that are not currently targeted in classical vaccine approaches(19). These include two novel secreted molecules, the EtpA adhesin(20) and the EatA(21) autotransporter both originally identified in H10407, an ETEC strain isolated from a case of severe cholera-like diarrhea in Bangladesh. Recent work demonstrates that both antigens are globally distributed in the ETEC pathovar and are more highly conserved than the most common CFs(19, 22). Moreover, they are protective in murine models of infection(23-26) and immunogenic in human challenge trials(7, 18), suggesting that these molecules could provide additional antigenic targets for vaccine development. While much is known about EatA and EtpA under experimental conditions, less is known about their respective roles in natural infections. The present studies were designed to explore the role of these and other potential noncanonical antigens in shaping the adaptive immune response to ETEC infection and to examine their contribution to virulence.

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.20.348110; this version posted October 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Methods

clinical samples used in this study

Specimens used in these studies were obtained from archived studies on ETEC birth cohort carried out in Mirpur in Dhaka city (10) as well as other studies (27) The International Centre for Diarrhoeal Disease Research, Bangladesh (icddr,b). Frozen ETEC isolates were retrieved from storage (-80°C) and duplicate vials were shipped to Washington University and subsequent antigen detection.

microbial genome analysis and bioinformatic antigen selection

Genomes from 89 clinical ETEC isolates previously collected at icddrb were used to identify conserved surface proteins. Sequence data for all eighty-nine clinical isolates examined in this study are available in GenBank (28). Paired-end Illumina sequence data from each isolate were generated *de novo* and contigs were binned using a previously described protocol (28). The ETEC genomes were compared using LS-BSR as previously described (29-31). The predicted protein-encoding genes of each genome that had ≥90% nucleotide identity to each other were assigned to gene clusters using uclust (32). Representative sequences of each gene cluster were then compared to each genome using TBLASTN (33) with composition-based adjustment turned off, and the TBLASTN scores were used to generate a BSR value indicating the detection of each gene cluster in each of the genomes analyzed. The BSR value was determined by dividing the score of a gene compared to a genome by the score of the gene compared to its own sequence. The predicted protein function of each gene cluster was determined using an ergatis-based (34) in-house annotation pipeline (35). A total of 13,835 non-redundant putative genes (referred to here as 'centroids') were extracted from the eighty-nine genomes.

All 13,835 centroids in this study were subjected to a reverse vaccinology pipeline (Institute for Genome Sciences, Maryland, USA) to identify those features that contained features that suggested they were surface exposed. An additional subtractive analysis was conducted by filtering centroids (BLASTx and BLASTn) against the genome contents of six *E. coli* commensal and laboratory strains, yielding 6,444

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.20.348110; this version posted October 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

ETEC pathovar-specific centroids. These data were further refined by selecting centroids with a blast score ratio (BSR)(36) ≥ 0.8 (i.e. highly conserved) and present in at least 40% of the clinical isolates, yielding 316 conserved, virulence-linked genetic features for further analysis. BLASTx was next used to assign a putative function to these virulence-linked centroids. This analysis was coupled with results from pSORTv3.0 (37), SubLoc (38), and CELLO (39) to predict subcellular localization, altogether resulting in the down-selection to 118 potential surface-expressed molecules These features were complemented with all known and putative colonization factor subunits (n=52), toxin subunits, and subdomains of novel antigens for inclusion on the microarrays (supplemental dataset 1).

microarray production

Antigen-encoding regions selected for the microarrays were amplified by PCR using primers listed in supplemental dataset 2, and constructed as previously described(7, 18, 40) Recombinant versions of select antigens including EtpA, EatA, LT-A, LT-B, YghJ, ST-H, and EaeH were also included on the arrays.

microarray processing

Microarrays were shipped to icddrb where they were rehydrated for 10 minutes with 100 ul Array Blocking Buffer. *E. coli* lysate was reconstituted in a final volume of 20% in blocking buffer. Antibody in Lymphocyte Supernatant (ALS) prepared from blood of ETEC patients were diluted 1:2.5 in the resuspended lysate followed by loading onto the microarrays and incubated in the dark for 2 hours at 25°C on a rotating platform. Microarrays were then washed 3x with TBS-T (0.05% Tween in TBS, pH 7.5) followed by incubation for 5 minuites in TBS-T at 25°C. This process was repeated once with TBS followed by a final wash in distilled water. Slides were dried by centrifugation (10 minutes at 500 x *g*) then stored in desiccated boxes prior to shipping to the Felgner Laboratory, University of California, Irvine.

non-canonical antigen ELISA

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.20.348110; this version posted October 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

384 well plates (Corning, product number 3540) were coated with recombinant EatA passenger domain (rEatp, 10 micrograms/milliliter in carbonate buffer [15 mM Na_2CO_3, 35 mM $NaHCO_3$, 0.2g/l NaN_3, pH 9.6]) or recombinant EtpA (rEtpA, 1 microgram/milliliter in carbonate buffer) and shipped to icddr,b , being maintained at 4°C prior to use. The ELISA plates were manually washed three times with PBS-T (PBS with 0.05% tween) including brief centrifugation for 30 seconds at 200 x *g* on a tabletop centrifuge between washes. Plates were rehydrated with 1% BSA in PBS-T overnight at 4°C. The following day, serum or plasma samples and plates were warmed to ambient temperature (~25°C), serum was diluted 1: 200 in PBS-T with 1%BSA and briefly vortexed. 10 µl of diluted serum was added to the plates, centrifuged as above, sealed, and incubated at 37°C for 1 hour. After incubation, plates were washed 3 times with PBS-T as described above. 10 µl of HRP-conjugated anti-human IgG (Jackson ImmunoResearch Laboratories, Cat 309-035-006, West Grove, PA) was diluted 1:2000 in 1% BSA in PBS-T followed by incubation and washing as above. ELISA plates were read using 10 ul of 3,3′,5,5′-tetramethylbenzidine (TMB) substrate (Seracare, Cat# 50-76-00, Milford, MA) and the Vmax of the reaction was determined using a BioTek Plate reader with Gen5 Take3 software (v.2.00.18). Due to variations between ELISA plates, we analyzed data independently for each plate and in combination, adjusting for age to control for repeated measures.

strain characterization by PCR and immunoblotting

Frozen glycerol stocks of ETEC strains maintained at -80°C were used to inoculate lysogeny broth (LB) for overnight growth at 37°C, 250 rpm. 1 µl of the overnight culture was diluted in 100 µl of PBS of which 1 µl was used as the DNA template in initial PCR screening with primers in supplementary table 1. The thermocycler conditions for *eatA* and *etpA* were denaturation for 5 minutes at 95°C with 30 amplification cycles utilizing 95°C for 30 seconds, 52°C for 30 seconds, and 72°C for 2 minutes. The toxin multiplex assay (genes *eltB*, *estH*, and *estP*) were conducted as follows; 5 minutes at 95°C with 32 cycles of amplification using 94°C for 15 seconds, 55°C for 15 seconds, and finally 72°C for 30 seconds.

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.20.348110; this version posted October 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Amplicons were visualized as before using a 0.8% agarose gel with ethidium bromide. The H10407 strain (*eatA*, *etpA*, *estH*, *estP*, and *eltB* positive) was used as a positive control in for assays. To adjudicate discordant results PCR was performed using gDNA extraction with the Invitrogen PureLink Quick Plasmid Miniprep Kit (Cat# K210010, Thermo Fisher, Waltham, MA) Miniprep kit. If toxin multiplex PCRs were negative, isolates were deemed to have lost their original plasmid during storage, transportation, or culture passage and subsequently excluded from analysis. Immunoblotting for EatA and EtpA were performed on TCA-precipitated culture supernatants as previously described (19) using affinity-purified polyclonal rabbit antibodies against the passenger domain of EatA (21) or EtpA (20) (dilutions 1:1,000 and 1:5,000, respectively) Primary antibodies were detected using Horseradish Peroxidase (HRP)-conjugated anti-rabbit IgG secondary antibody (1:5,000 dilution, Invitrogen #A16110) for 1 hour at room temperature. HRP was detected with ECL Western blotting substrate (Bio-Rad, #ABIN412579).

statistical analysis

Categorical outcomes were analyzed using chi-square tests, Fisher's exact tests, or age-adjusted logistic regression analyses as appropriate. Serum data were analyzed using a linear repeated measures model with a compound symmetry covariance structure. p-values < 0.05 were considered significant. Analyses were conducted using SAS version 9.4 (SAS Institute Inc., Cary, NC, USA) or SPSS v.24 (IBM, Armonk, NY, USA), or GraphPad Prism v9.0.0.

Ethics Statement

These studies were approved by the Research Review and Ethical Review Committee of (icddr,b) and the Institutional Review Board of Washington University School of Medicine in Saint Louis.

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.20.348110; this version posted October 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Results

Antibodies following natural infection recognize a finite repertoire of ETEC proteins.

Both human experimental models(7) as well as natural infections(10) demonstrate that prior infection with ETEC affords substantial protection against symptomatic disease. Elucidation of the nature of protective adaptive immune responses to these mucosal pathogens can therefore inform vaccine development. While the majority of earlier ETEC vaccinology efforts have centered on colonization factor antigens, the present studies were designed to broadly profile antigenic responses and to assess the role of recently characterized non-canonical antigens. To assess the breadth of immune responses to ETEC during acute natural infection, we designed protein microarrays containing all known and putative colonization factor antigen subunits, and additional predicted surface-expressed proteins conserved in more than 40% of the ETEC pathovar including EtpA and EatA, secreted antigens expressed by a majority of ETEC strains in a global collection of isolates(19).

Despite the inclusion of multiple candidate surface molecules on the array predicted to be conserved among strains in Bangladesh from our *in silico* analysis, immune responses following infection were largely constrained to a small group of antigens including EtpA and EatA (figure 1A), LT (supplementary figure 1), select colonization factor subunits (supplementary figure 2) including CssB, one of two components of the CS6 polymer(41), a predominant immunogenic antigen among strains circulating in Bangladesh(27). Compared to control specimens obtained outside of the ETEC endemic area, both EatA and EtpA exhibited high levels of reactivity. Notably, for patients infected with EtpA-expressing strains, EtpA responses were significantly higher at day 30 following infection than those observed immediately following admission, whereas the converse was true in patients admitted with EtpA-negative strains. (figure 1B)

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.20.348110; this version posted October 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

In an open-aperture assessment of ALS specimens (42, 43) obtained from adults hospitalized at icddr,b Hospital in Dhaka, Bangladesh or from patients recruited at the Mirpur field site with acute symptomatic diarrheal illness, we again noted that immune responses following infection were largely constrained to a relatively small group of antigens including CS6, EtpA and EatA (supplementary dataset 1). When parsing antigen profiles of the infecting strain, we found that those infected with EtpA-expressing ETEC exhibited significant increases in both ALS IgA (p=0.005), and IgG (p=0.02) responses in the week following infection relative to those infected with EtpA-negative strains (figure 2). As anticipated, we also observed significant increases in ALS immunoreactivity to the CssB subunit of CS6 that correlated with the production of CS6 by the infecting strain (supplementary figure 3).

EatA and EtpA are immunogenic in young children.

Data from recent CHIM studies(7, 18) as well as earlier data from patients with natural ETEC infections(22), indicate that adults develop robust immune responses to non-canonical antigens including EtpA and EatA. However, in endemic areas young children are the population most severely impacted by ETEC with incidence declining after 24 months of age, presumably as protection develops subsequent to infection. Therefore, we examined sera from a cohort of Bangladeshi children followed from birth through 2 years of age (10) to profile development of antibody responses to EatA and EtpA over time (figure 3). During the first month of life, the majority of children were observed to have elevated IgG responses to both EatA and EtpA, presumably reflecting passive transfer of maternal antibodies(44). As anticipated, responses to both antigens decreased by three months of age, while mean responses to each antigen increased significantly through 24 months of age, likely reflecting early childhood infections with strains expressing EtpA and EatA.

Anti-EtpA or EatA responses relative to symptomatic diarrhea.

Immunologic correlates of protection against ETEC are currently unknown(45). The majority of clinical studies to date have examined the impact of prior infection with strains producing particular

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.20.348110; this version posted October 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

222 colonization factors and/or LT(9) as well as antibody acquisition on subsequent risk of infection with
223 similar strains(46, 47). We hypothesized that because EtpA and EatA are relatively common antigens in
224 the ETEC pathovar(19), higher antibody responses to these antigens may be associated with
225 subsequent protection against symptomatic infection. After excluding antibody responses at one
226 month of age, we examined the IgG antibody responses to EtpA and EatA preceding detection of ETEC
227 in either symptomatic or asymptomatic children between four and twenty-four months of age.
228 Interestingly, we observed elevated responses to both antigens prior to detection of ETEC in
229 asymptomatic children detection relative to symptomatic cases (figure 4), perhaps reflecting the
230 overall mitigating impact of prior exposure on development of diarrheal illness.
231
232 **etpA relative to blood group A upon first ETEC exposure.**
233 Recent studies have shown that the EtpA adhesin engages host cells via lectin interactions with N-
234 acetylgalactosamine residues, preferentially when presented as the terminal sugar on blood group A
235 glycans displayed on intestinal epithelia. Also, when challenged with ETEC H10407, an EtpA-producing
236 strain, human volunteers of A blood group are more likely to exhibit severe symptomatic illness(48),
237 recapitulating the observation that symptomatic ETEC infections were more common in children with
238 blood type A or AB in birth cohort studies of Bangladeshi children(10). We therefore examined this
239 cohort for potential associations between *etpA*, blood type, and disease status during natural ETEC
240 infections. In limiting analysis to the first ETEC isolation to avoid confounding effects of repeated
241 exposures, 41.2% of children that were blood type A or AB had symptomatic diarrhea during their first
242 ETEC infection (table 1) compared to 30.6% of those that were blood group B or O (30.6%, $p = 0.187$).
243 Blood group A or AB individuals were somewhat more likely to have an *etpA* positive strain recovered
244 from their first infection (74.5% vs. 66.7%, oR = 1.462, 95%CI 0.693-3.082), although these data were
245 not statistically significant.

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.20.348110; this version posted October 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

association of *eatA* and *etpA* with virulence.

Although both EatA, a mucin-degrading serine protease, and the EtpA blood group A lectin are secreted by a diverse population of ETEC strains(19), and contribute to virulence phenotypes *in vitro* as well as in small animal models of ETEC infection(23, 26, 49, 50), the role played by these antigens in human infections has yet to be explored in detail. To explore the association of *eatA* and *etpA* with symptomatic ETEC infection, we examined isolates collected in a birth cohort study in which stools were collected at monthly intervals in asymptomatic children (asymptomatic colonization) of during surveillance for diarrhea (symptomatic infection) (10). Notably, the presence of *etpA* or *eatA* significantly increased the odds of having symptomatic diarrhea (unadjusted odd ratios of 2.1 and 3.1, respectively (table 2). Similarly, after adjusting for age we observed significant associations between the presence of either EtpA (adjusted odds ratio 1.98, p=0.007) or EatA (adjusted odds ratio 2.91, p<0.001) and development of diarrheal disease.

The *eatA* gene (21) and *etpBAC* locus(20) encoding the two-partner secretion system responsible for EtpA secretion, were originally identified on the p948 plasmid of ETEC strain H10407, which also encodes the gene for STh (51), and our earlier studies suggested that both loci are more commonly associated with ST-producing strains (19). Importantly, large epidemiological studies have demonstrated an association between ST or ST/LT-producing ETEC and more severe disease relative to LT-only producing ETEC (52, 53). Similarly, we again found an association between ST-producing ETEC and symptomatic diarrhea, where 59.0% of colonizing ETEC isolates encode STh or STp (*estH* or *estP* positive) compared to 87.3% of diarrhea-associated isolates (adjusted odds ratio 4.66 [95% CI, 2.62, 8.85, p < 0.001]). We therefore asked whether the *eatA* or *etpA* associations with virulence were independent of ST. The presence of either gene was associated with higher risk of diarrheal illness independent of ST, although only the presence of *eatA* was significantly associated with illness adjusted for age. Collectively, however these data suggest that these more recently discovered non-canonical

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.20.348110; this version posted October 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

antigens, now frequently referred to as "accessory" virulence factors, could be important contributors to ETEC disease.

Discussion

ETEC were initially discovered in patients presenting with severe diarrheal illness that mimicked clinical cholera (54-56). Following seminal discoveries of the heat-labile (LT) and heat-stable (ST) toxins that define ETEC, and initial characterization of plasmid-encoded colonization factor antigens (CFs), a canonical approach to vaccine development focused on LT and CFs emerged. However, subsequent studies have revealed that the molecular pathogenesis of ETEC likely involves a number of other plasmid as well as chromosomally encoded features that may potentially expand the repertoire of target "non-canonical" antigens for use in ETEC vaccine development.

Among antigens that are unique to the ETEC pathovar are two high molecular weight secreted proteins, EtpA and EatA. The relative conservation of genes encoding their corresponding secretion systems within the ETEC pathovar(18, 19), their immunogenicity during natural(57) and experimental human infection(7, 18), and contribution to virulence *in vitro* and small animal studies have highlighted their potential utility as vaccine candidates. Nevertheless, our understanding of the importance of these antigens to ETEC virulence continues to evolve. Recent studies have revealed that the secreted 110 kD passenger domain of the EatA autotransporter protein functions as a mucin-degrading enzyme, capable of dissolving the MUC2 matrix that covers the surface of enterocytes, the target for ETEC binding and toxin delivery(23). EtpA, secreted by two-partner secretion mechanism that requires both the EtpB outer membrane pore and EtpC, a glycosyltransferase(20), functions as an adhesin by bridging the bacteria(50) and GalNAc-containing host cell glycans present on enterocytes(48). However, despite an emerging understanding of the function of these molecules, very little is known about their contribution to disease in human hosts.

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.20.348110; this version posted October 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

293 The present studies extend earlier observations to a cohort of naturally infected children in
294 Bangladesh(10) and suggest that these non-canonical antigens play critical roles in determining the
295 outcome of ETEC infections. The finding that genes encoding these antigens are significantly
296 associated with the development of symptomatic infection may have important implications for the
297 interpretation of large-scale epidemiologic studies that have employed population attributable fraction
298 methodology in which ETEC detected in cases of diarrheal illness are compared to asymptomatically
299 colonized controls(58). The present studies would seem to suggest that additional characterization of
300 ETEC beyond the pathovar-defining heat-labile or heat-stabile toxins could be required to accurately
301 assess the contribution of ETEC to the global burden of diarrheal disease.
302 In general, expanded open-aperture assessment of immune responses to natural ETEC
303 infections appears to reaffirm earlier observations in human volunteer studies(7). Namely, that there
304 are relatively few immunogenic targets in the potential repertoire of ETEC surface molecules, with
305 EtpA and EatA predominating among the pathovar-specific antigens.
306 Although we observed higher IgG serum antibody responses to both EtpA and EatA in children
307 who were simply colonized with ETEC compared to those with diarrhea, suggesting that these antigens
308 could afford some protection against symptomatic illness, these findings need to be interpreted
309 cautiously. Both EtpA and EatA are relatively common antigens among strains circulating in
310 Bangladesh, therefore the identification of antibodies could simply reflect prior infection that mitigates
311 infection through responses to other antigens. In addition, correlates of protection for ETEC, as well as
312 the protective role of serum IgG in enteric infections remain unclear with mucosal IgA responses
313 considered key to protection (45).
314 Altogether, however the findings reported here suggest that antigens which have not been part
315 of traditional approaches to vaccine development may play important roles in virulence, and in acquired

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.20.348110; this version posted October 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

316 immunity to ETEC. Further studies will clearly be needed to examine the efficacy of these more

317 recently discovered antigens as protective immunogens.

318

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.20.348110; this version posted October 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Acknowledgements

These studies were supported in part by funding from National Institute of Allergy and Infectious Diseases (NIAID) of the National Institutes of Health (NIH) under Award Numbers R01AI089894, R01AI126887 (jmf), K23AI30389 (fmk) and The Department of Veterans Affairs I01BX004825 (jmf). The icddr,b is supported by the Governments of Bangladesh, Canada, Sweden, and the UK. James M. Fleckenstein is listed the inventor on patent 8323668

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.20.348110; this version posted October 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

figures

figure 1

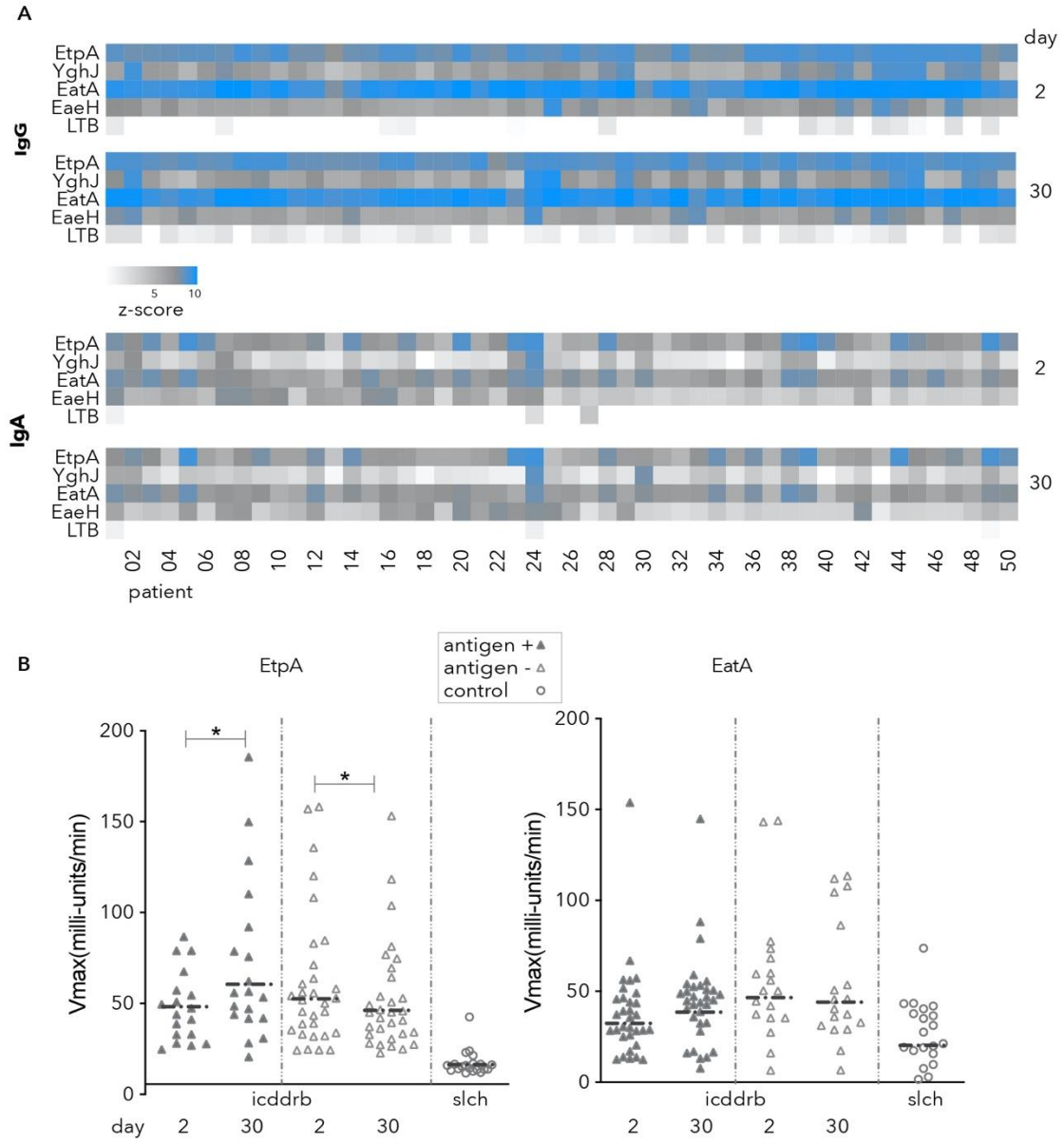


Figure 1. serologic response to non-canonical antigens following natural infection

A. heatmap indicates log 2 transformed z-score data indicating ETEC protein microarray responses from day 2, and 30 following presentation to icddrb to four non-canonical antigens EtpA, YghJ, the passenger domain of EatA, EaeH; and the B-subunit of ETEC heat-labile toxin (LT-B). **B.** kinetic ELISA responses to EtpA and EatA following infection. Data are segregated by the presence (closed symbols), or absence (open symbols) of each respective antigen in the strain recovered at presentation. Negative

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.20.348110; this version posted October 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

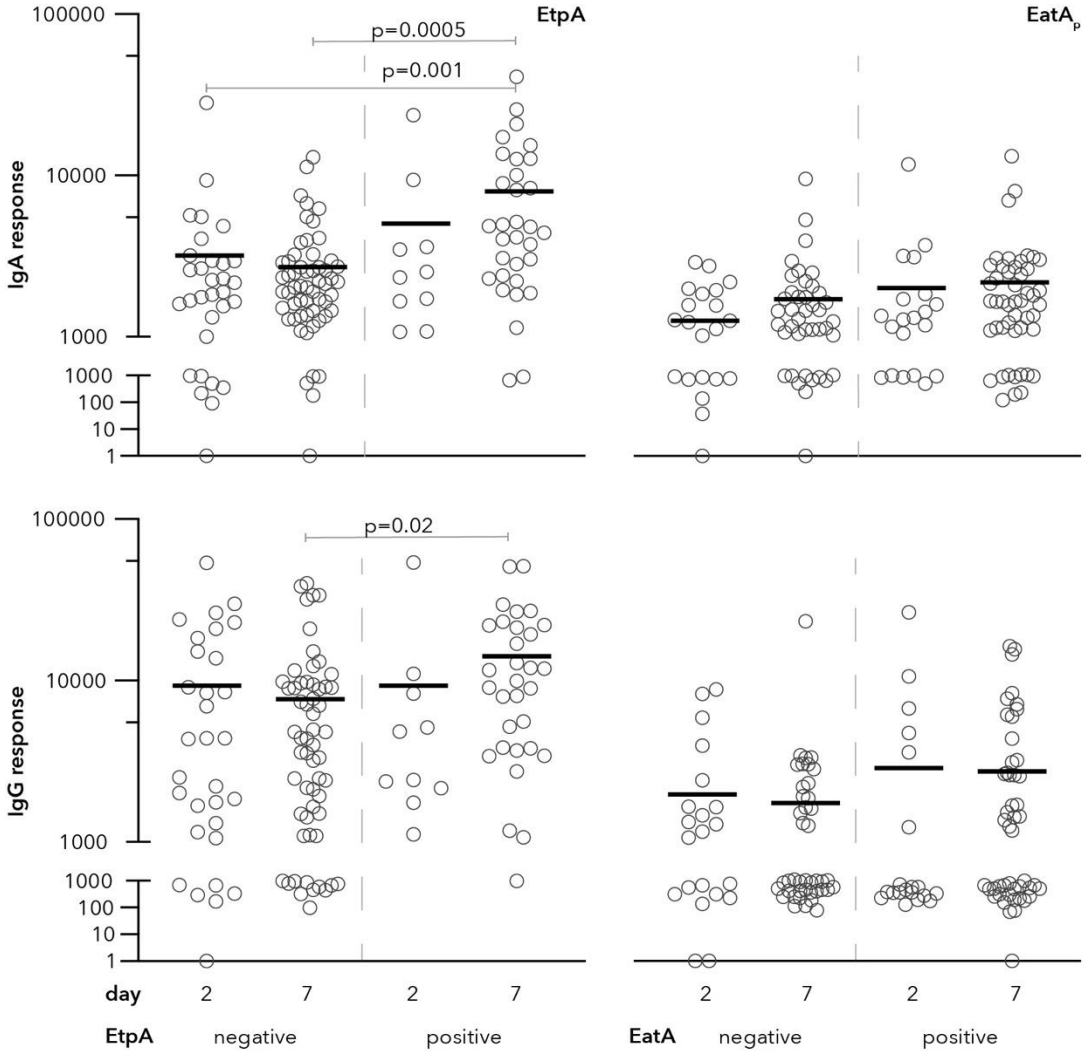
337 control samples from Saint Louis Children's Hospital (slch) are shown as open circles. *<0.05 by
338 Wilcoxon matched-pairs signed rank test.

339

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.20.348110; this version posted October 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

340

341 figure 2



342

343 **Figure 2. ALS responses to EtpA or EatA**. Shown are microarray data for IgA (top panels) and IgG

344 responses to EtpA (left) and the passenger domain of EatA (EatAp, right) on days 2, and 7 following

345 hospitalization. Data in each graph are segregated according to antigen expression in the infecting

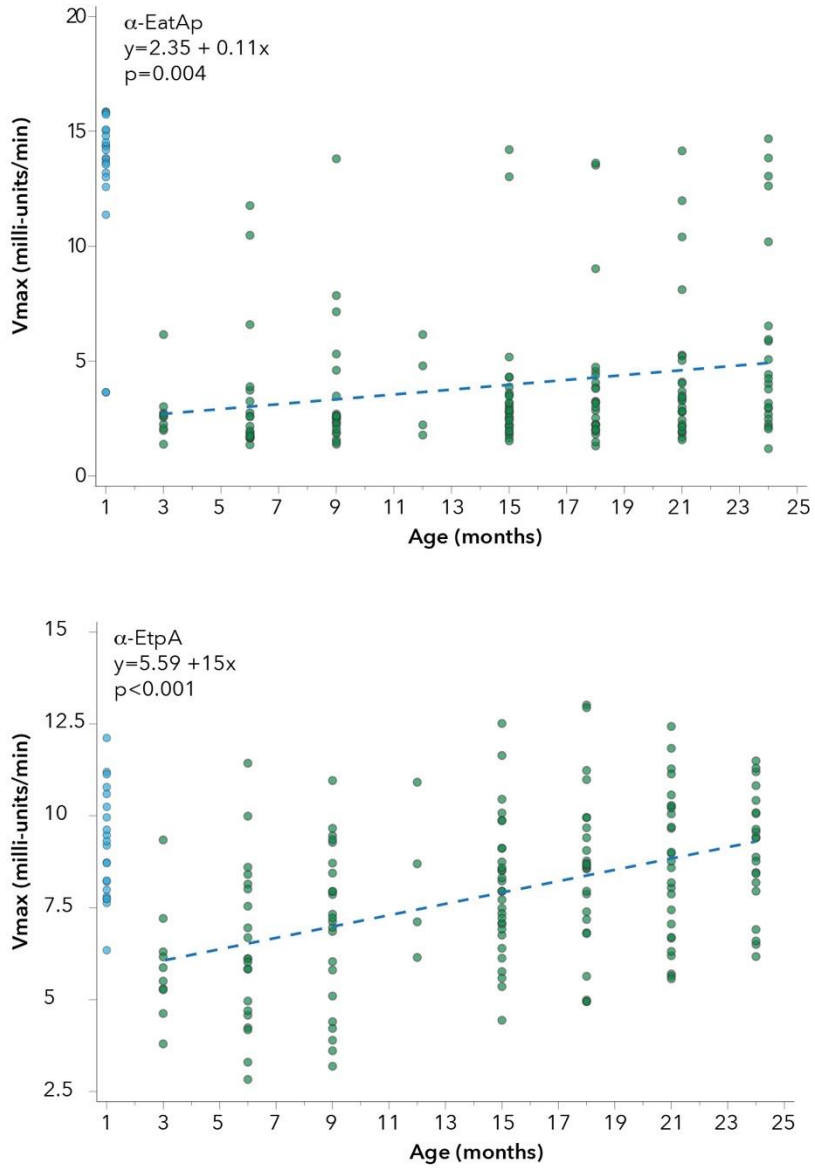
346 strain (negative or positive). p values reflect Kruskall-Wallis, with post-hoc analysis using Dunn's test

347 adjusted for multiple comparisons for between group analysis.

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.20.348110; this version posted October 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

348

349 **figure 3**



350

351 **Figure 3.** Anti-EtpA or Anti-EatA IgG responses increase with age. Shown are representative kinetic

352 ELISA data for serum IgG samples obtained from children ages 1-24 months enrolled in a birth cohort

353 study. Scatter plots of Anti-EtpA and Anti-EatA IgG plotted against data for ages 3-24 months with

354 regression lines from linear repeated measures models overlaid (dotted line) demonstrate significant

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.20.348110; this version posted October 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

355 increases over time in responses to the passenger domain of EatA ($EatA_p$, top) and EtpA (bottom). See

356 Figure S4 for additional plots.

357

358

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.20.348110; this version posted October 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

figure 4

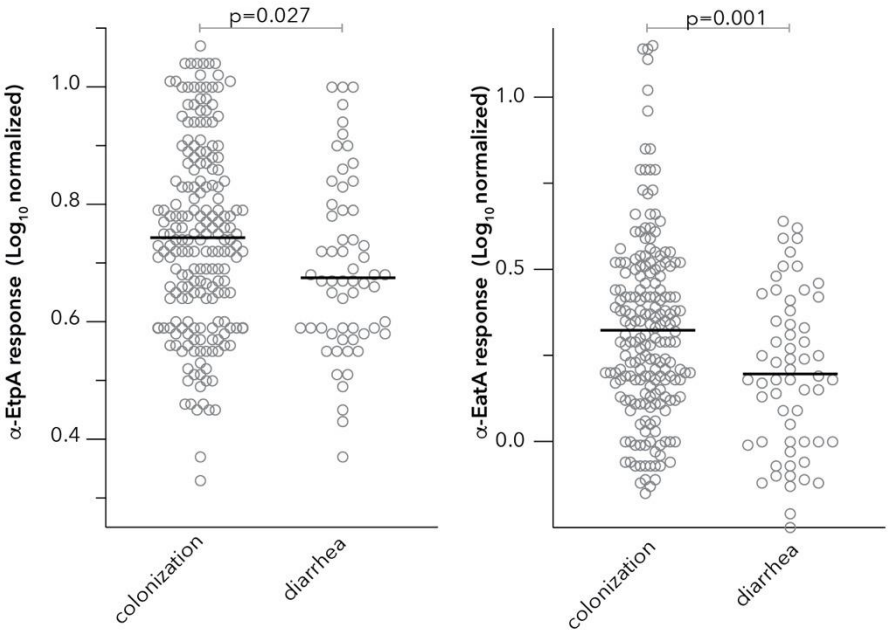


Figure 4. Serum IgG responses preceding asymptomatic ETEC colonization and diarrhea. Shown are peak serum IgG responses for EtpA (left) or the EatA passenger domain (right) preceding either asymptomatic colonization or diarrheal illness with ETEC. Data shown are Log_{10} transformed IgG antibody responses determined by kinetic ELISA. Bars represent mean values.

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.20.348110; this version posted October 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

tables

table 1. associations of blood group, etpA expression and diarrheal disease

table 2. association of blood group, *etpA*, and diarrheal illness

	clinical presentation		*etpA* status	
	asymptomatic	symptomatic	negative	positive
blood type	n (%)	n (%)	n (%)	n (%)
O or B	75 (69.4)	33 (30.6)	36 (33.3)	72 (66.7)
A or AB	30 (58.8)	21 (41.2)	13 (25.5)	38 (74.5)
		p=0.187*		p=0.317†

* comparison of symptomatic infections in A/AB vs B/O

† comparison of *etpA* positivity in A/AB vs B/O individuals with first episodes of diarrhea

p values determined by chi-square test

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.20.348110; this version posted October 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

non-canonical antigen analysis of Bangladeshi birth cohort samples

		EtpA *antigen* status		Unadjusted Odds Ratio (95%CI)		Age Adjusted Odds Ratio	
Toxin(s)	Diarrhea	Negative	Positive		p-value		p-value
All	-	110 (40.6%)	161 (59.4%)	2.10 (1.28, 3.45)	0.003	1.98 (1.21, 3.29)	0.007
	+	27 (24.5%)	83 (75.5%)				
ST or ST/LT	-	55 (34.4%)	105 (65.6%)	1.57 (0.89, 2.77)	0.116	1.51 (0.86, 2.68)	0.156
	+	24 (25%)	72 (75%)				
LT only	-	55 (49.5%)	56 (50.5%)	3.60 (0.95, 13.61)	0.047	2.59 (0.78, 10.79)	0.144
	+	3 (21.4%)	11 (78.6%)				

		EatA *antigen* status		Unadjusted Odds Ratio (95%CI)		Age Adjusted Odds Ratio	
Toxin(s)	Diarrhea	Negative	Positive		p-value		p-value
All	-	152(56.1%)	119 (43.9%)	3.11 (1.93, 5.01)	<.001	2.91 (1.81, 4.75)	<.001
	+	32 (29.1%)	78 (70.9%)				
ST or ST/LT	-	67 (41.9%)	93 (58.1%)	2.05 (1.18, 3.56)	0.011	1.91 (1.1, 3.38)	0.024
	+	25 (26%)	71 (74%)				
LT only	-	85 (76.6%)	26 (23.4%)	3.27 (1.05, 10.18)	0.051	2.36 (0.74, 7.46)	0.142
	+	7 (50%)	7 (50%)				

p-values for unadjusted odds ratios obtained from simple chi-square or Fisher's exact tests, and p-values for age-adjusted odds ratios obtained from logistic regressions that included age as a covariate.

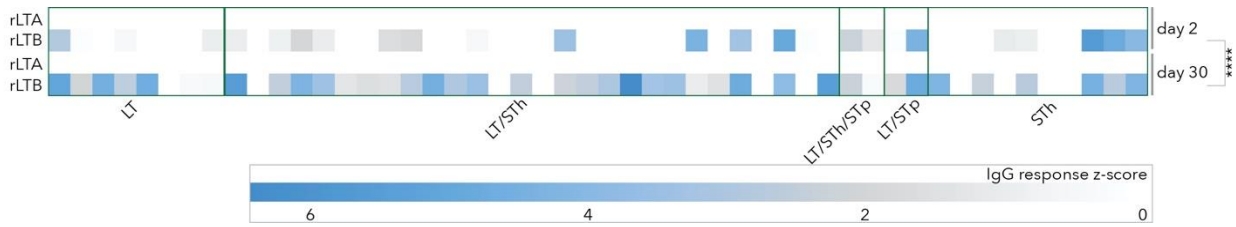
table 2. relationship of *eatA*, and *etpA*, to symptomatic ETEC.

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.20.348110; this version posted October 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

supplementary data

supplementary figures

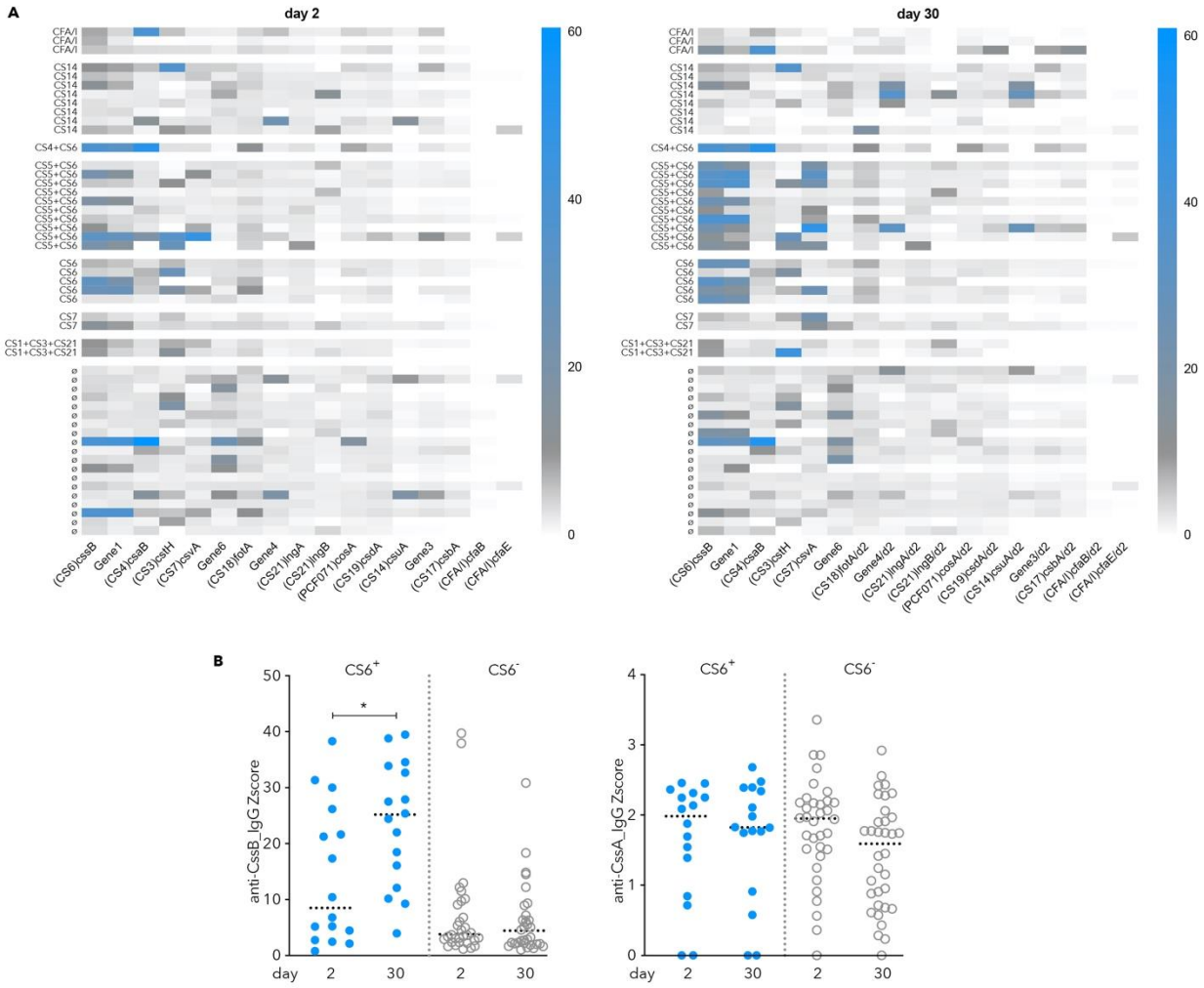
supplementary figure 1



supplementary figure 1. IgG responses to LT subunits LT-A and LT-B following infection. Shown are array z-score data from days 2 and 30 following presentation to icddrb. Data are segregated by the toxin profile of the ETEC strain isolated at presentation. ****=p<0.0001 by Wilcoxon matched pairs comparison of day 2 and day 30 LT-B responses.

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.20.348110; this version posted October 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

supplementary figure 2



supplementary figure 2. (**A**) Serum IgG responses to colonization factor subunits following ETEC infection, on day 2 and day 30 following presentation to icddrb. Shown are normalized zscore data for select CF antigens (bottom) printed on the microarrays for 50 individuals infected with ETEC. Antigens expressed by the infecting strain for each patient are shown at the left of the heatmap. Responses following infection with CS6-expressing strains are outlined in green. **B**. Serologic responses to CssB (left), and CssA (right) subunits of CS6 following infection. Data are parsed by antigen expression in the infecting strain with symbols in blue indicating CS6+ samples. *p=0.016 by Wilcoxon matched-pairs signed rank testing.

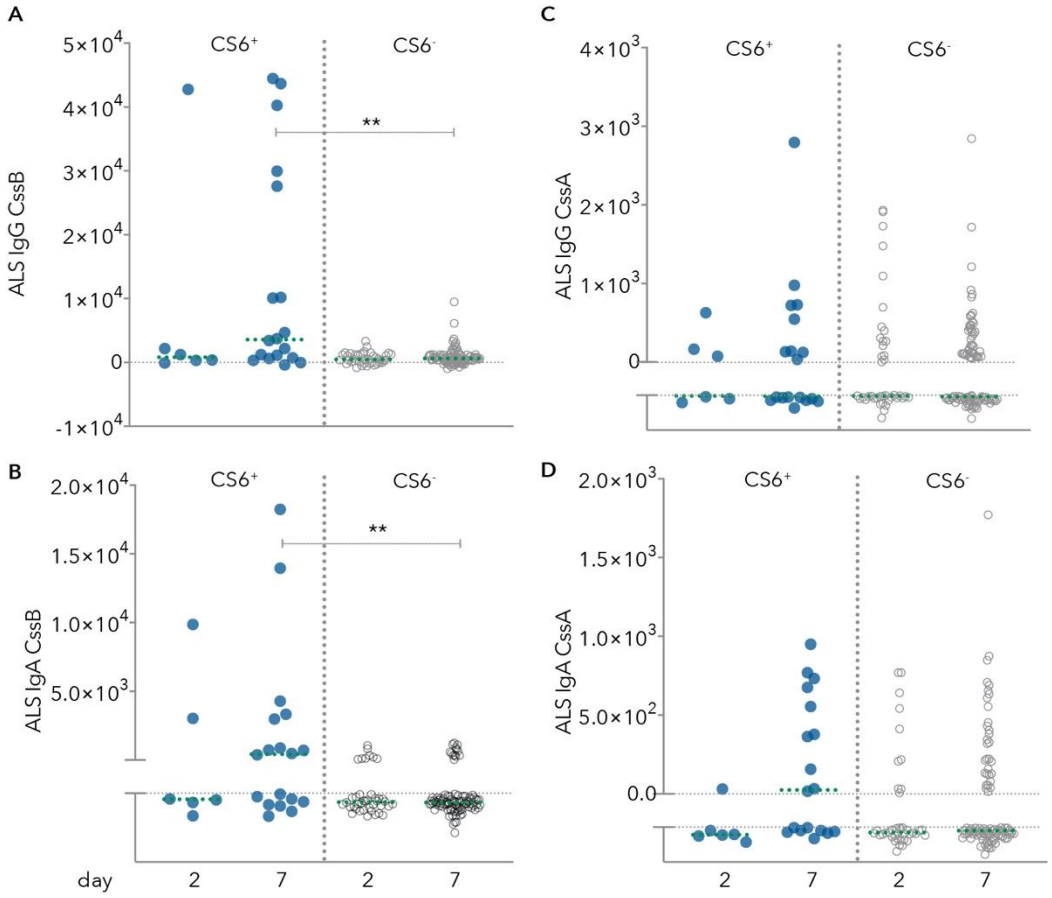
bioRxiv preprint doi: https://doi.org/10.1101/2020.10.20.348110; this version posted October 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

403

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.20.348110; this version posted October 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

404 supplementary figure 3

405



406

407 **supplementary figure 3**. ALS responses to CS6 subunits following ETEC infection. (**A**) IgG response to

408 CssB. (**B**) IgA response to CssB. **p<0.01**,** Kruskal-Wallis post-hoc analysis using Dunn's test adjusted

409 for multiple comparisons. (**C**) IgG response to CssA. (**D**) IgA response to CssA.

410

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.20.348110; this version posted October 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

411 # supplementary tables

412 ## supplementary table 1

supplementary table 1. primers used in strain interrogation

gene	reference sequence GenBank	amplicon (bp)	Sequence (5'>3')	primer ID
eatA	AY163491.2	1943	ATGTGCTTTGGCAGGTTAA	jf082213.1-F
			ATATCCAGTCAGCACCCACT	jf082213.2-R
etpA	AY920525.2	999	GGTTCAGGCAGTATCCAGAC	jf082213.3-F
			GGTGTAGCTGTCTGACCACA	jf082213.4-R
eltB	CBJ04425.1	273	ACGGCGTTACTATCCTCTC	jf092313.3-F
			TGGTCTCGGTCAGATATGTG	jf092313.4-R
estP (*sta1*)	CBJ04435.1	166	TCTTTCCCCTCTTTTAGTCAG	jf092313.5-F
			ACAGGCAGGATTACAACAAAG	jf092313.6-R
estH (*sta2*)	CBJ04483.1	64	TACAAGCAGGATTACAACAC	jf092313.7-F
			AGTGGTCCTGAAAGCATG	jf092313.8-R

413

414

415

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.20.348110; this version posted October 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References:

1. Collaborators GBDDD. 2018. Estimates of the global, regional, and national morbidity, mortality, and aetiologies of diarrhoea in 195 countries: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Infect Dis 18:1211-1228.
2. Anderson JDt, Bagamian KH, Muhib F, Amaya MP, Laytner LA, Wierzba T, Rheingans R. 2019. Burden of enterotoxigenic Escherichia coli and shigella non-fatal diarrhoeal infections in 79 low-income and lower middle-income countries: a modelling analysis. Lancet Glob Health 7:e321-e330.
3. Hosangadi D, Smith PG, Giersing BK. 2019. Considerations for using ETEC and Shigella disease burden estimates to guide vaccine development strategy. Vaccine 37:7372-7380.
4. Lanata CF, Black RE. 2018. Estimating the true burden of an enteric pathogen: enterotoxigenic Escherichia coli and Shigella spp. Lancet Infect Dis 18:1165-1166.
5. Watanabe K, Petri WA, Jr. 2016. Environmental Enteropathy: Elusive but Significant Subclinical Abnormalities in Developing Countries. EBioMedicine 10:25-32.
6. Hosangadi D, Smith PG, Kaslow DC, Giersing BK, Who E, Shigella Vaccine Consultation Expert G. 2018. WHO consultation on ETEC and Shigella burden of disease, Geneva, 6-7th April 2017: Meeting report. Vaccine doi:10.1016/j.vaccine.2017.10.011.
7. Chakraborty S, Randall A, Vickers TJ, Molina D, Harro CD, DeNearing B, Brubaker J, Sack DA, Bourgeois AL, Felgner PL, Liang X, Mani S, Wenzel H, Townsend RR, Gilmore PE, Darsley MJ, Rasko DA, Fleckenstein JM. 2018. 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.
8. Levine MM, Nalin DR, Hoover DL, Bergquist EJ, Hornick RB, Young CR. 1979. Immunity to enterotoxigenic Escherichia coli. Infect Immun 23:729-36.
9. Steinsland H, Valentiner-Branth P, Gjessing HK, Aaby P, Molbak K, Sommerfelt H. 2003. Protection from natural infections with enterotoxigenic Escherichia coli: longitudinal study. Lancet 362:286-91.
10. Qadri F, Saha A, Ahmed T, Al Tarique A, Begum YA, Svennerholm AM. 2007. Disease burden due to enterotoxigenic Escherichia coli in the first 2 years of life in an urban community in Bangladesh. Infect Immun 75:3961-8.
11. Gaastra W, Svennerholm AM. 1996. Colonization factors of human enterotoxigenic Escherichia coli (ETEC). Trends Microbiol 4:444-52.
12. Clements JD, Norton EB. 2018. The Mucosal Vaccine Adjuvant LT(R192G/L211A) or dmLT. mSphere 3.
13. Fleckenstein JM, Kuhlmann FM. 2019. Enterotoxigenic Escherichia coli Infections. Curr Infect Dis Rep 21:9.
14. Zegeye ED, Govasli ML, Sommerfelt H, Puntervoll P. 2019. Development of an enterotoxigenic Escherichia coli vaccine based on the heat-stable toxin. Hum Vaccin Immunother 15:1379-1388.
15. Taxt AM, Diaz Y, Aasland R, Clements JD, Nataro JP, Sommerfelt H, Puntervoll P. 2016. Towards Rational Design of a Toxoid Vaccine against the Heat-Stable Toxin of Escherichia coli. Infect Immun 84:1239-1249.
16. Sahl JW, Sistrunk JR, Baby NI, Begum Y, Luo Q, Sheikh A, Qadri F, Fleckenstein JM, Rasko DA. 2017. Insights into enterotoxigenic Escherichia coli diversity in Bangladesh utilizing genomic epidemiology. Sci Rep 7:3402.

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.20.348110; this version posted October 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

17. Mottram L, Chakraborty S, Cox E, Fleckenstein J. 2019. How genomics can be used to understand host susceptibility to enteric infection, aiding in the development of vaccines and immunotherapeutic interventions. Vaccine doi:10.1016/j.vaccine.2019.01.016.
18. Chakraborty S, Randall A, Vickers TJ, Molina D, Harro CD, DeNearing B, Brubaker J, Sack DA, Bourgeois AL, Felgner PL, Liang X, Mani S, Wenzel H, Townsend RR, Gilmore PE, Darsley MJ, Rasko DA, Fleckenstein JM. 2019. Interrogation of a live-attenuated enterotoxigenic Escherichia coli vaccine highlights features unique to wild-type infection. NPJ Vaccines 4:37.
19. Kuhlmann FM, Martin J, Hazen TH, Vickers TJ, Pashos M, Okhuysen PC, Gomez-Duarte OG, Cebelinski E, Boxrud D, Del Canto F, Vidal R, Qadri F, Mitreva M, Rasko DA, Fleckenstein JM. 2019. Conservation and global distribution of non-canonical antigens in Enterotoxigenic Escherichia coli. PLoS Negl Trop Dis 13:e0007825.
20. Fleckenstein JM, Roy K, Fischer JF, Burkitt M. 2006. Identification of a two-partner secretion locus of enterotoxigenic Escherichia coli. Infect Immun 74:2245-58.
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 72:1786-94.
22. 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:e0003446.
23. Kumar P, Luo Q, Vickers TJ, Sheikh A, Lewis WG, Fleckenstein JM. 2014. EatA, an Immunogenic Protective Antigen of Enterotoxigenic Escherichia coli, Degrades Intestinal Mucin. Infect Immun 82:500-8.
24. Luo Q, Vickers TJ, Fleckenstein JM. 2016. Immunogenicity and Protective Efficacy against Enterotoxigenic Escherichia coli Colonization following Intradermal, Sublingual, or Oral Vaccination with EtpA Adhesin. Clin Vaccine Immunol 23:628-37.
25. 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 27:4601-8.
26. 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 76:2106-12.
27. Qadri F, Ahmed T, Ahmed F, Bhuiyan MS, Mostofa MG, Cassels FJ, Helander A, Svennerholm AM. 2007. Mucosal and systemic immune responses in patients with diarrhea due to CS6-expressing enterotoxigenic Escherichia coli. Infect Immun 75:2269-74.
28. Sahl JW, Sistrunk JR, Fraser CM, Hine E, Baby N, Begum Y, Luo Q, Sheikh A, Qadri F, Fleckenstein JM, Rasko DA. 2015. Examination of the Enterotoxigenic Escherichia coli Population Structure during Human Infection. MBio 6.
29. Sahl JW, Caporaso JG, Rasko DA, Keim P. 2014. The large-scale blast score ratio (LS-BSR) pipeline: a method to rapidly compare genetic content between bacterial genomes. PeerJ 2:e332.
30. Hazen TH, Sahl JW, Fraser CM, Donnenberg MS, Scheutz F, Rasko DA. 2013. Refining the pathovar paradigm via phylogenomics of the attaching and effacing *Escherichia coli*. Proc Natl Acad Sci USA 110:12810-5.
31. Sahl JW, Gillece JD, Schupp JM, Waddell VG, Driebe EM, Engelthaler DM, Keim P. 2013. Evolution of a pathogen: a comparative genomics analysis identifies a genetic pathway to pathogenesis in Acinetobacter. PLoS One 8:e54287.
32. Edgar RC. 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26:2460-1.

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.20.348110; this version posted October 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

33. Gertz EM, Yu YK, Agarwala R, Schaffer AA, Altschul SF. 2006. Composition-based statistics and translated nucleotide searches: improving the TBLASTN module of BLAST. BMC Biol 4:41.
34. Orvis J, Crabtree J, Galens K, Gussman A, Inman JM, Lee E, Nampally S, Riley D, Sundaram JP, Felix V, Whitty B, Mahurkar A, Wortman J, White O, Angiuoli SV. 2010. Ergatis: a web interface and scalable software system for bioinformatics workflows. Bioinformatics 26:1488-92.
35. Galens K, Orvis J, Daugherty S, Creasy HH, Angiuoli S, White O, Wortman J, Mahurkar A, Giglio MG. 2011. The IGS standard operating procedure for automated prokaryotic annotation. Stand Genomic Sci 4:244-51.
36. Rasko DA, Myers GS, Ravel J. 2005. Visualization of comparative genomic analyses by BLAST score ratio. BMC Bioinformatics 6:2.
37. Yu NY, Wagner JR, Laird MR, Melli G, Rey S, Lo R, Dao P, Sahinalp SC, Ester M, Foster LJ, Brinkman FS. 2010. PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. Bioinformatics 26:1608-15.
38. Chen H, Huang N, Sun Z. 2006. SubLoc: a server/client suite for protein subcellular location based on SOAP. Bioinformatics 22:376-7.
39. Yu CS, Lin CJ, Hwang JK. 2004. Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on n-peptide compositions. Protein Sci 13:1402-6.
40. Davies DH, Liang X, Hernandez JE, Randall A, Hirst S, Mu Y, Romero KM, Nguyen TT, Kalantari-Dehaghi M, Crotty S, Baldi P, Villarreal LP, Felgner PL. 2005. Profiling the humoral immune response to infection by using proteome microarrays: high-throughput vaccine and diagnostic antigen discovery. Proc Natl Acad Sci U S A 102:547-52.
41. Roy SP, Rahman MM, Yu XD, Tuittila M, Knight SD, Zavialov AV. 2012. Crystal structure of enterotoxigenic Escherichia coli colonization factor CS6 reveals a novel type of functional assembly. Mol Microbiol 86:1100-15.
42. Chang HS, Sack DA. 2001. Development of a novel in vitro assay (ALS assay) for evaluation of vaccine-induced antibody secretion from circulating mucosal lymphocytes. Clin Diagn Lab Immunol 8:482-8.
43. Forrest BD. 1988. Identification of an intestinal immune response using peripheral blood lymphocytes. Lancet 1:81-3.
44. Bhuiyan TR, Saha A, Lundgren A, Qadri F, Svennerholm AM. 2010. Immune responses to Helicobacter pylori infection in Bangladeshi children during their first two years of life and the association between maternal antibodies and onset of infection. J Infect Dis 202:1676-84.
45. Holmgren J, Parashar UD, Plotkin S, Louis J, Ng SP, Desauziers E, Picot V, Saadatian-Elahi M. 2017. Correlates of protection for enteric vaccines. Vaccine 35:3355-3363.
46. Tobias J, Andersson K, Bialik A, Cohen D. 2008. Preexisting antibodies to homologous colonization factors and heat-labile toxin in serum, and the risk to develop enterotoxigenic Escherichia coli-associated diarrhea. Diagn Microbiol Infect Dis 60:229-31.
47. Rao MR, Wierzba TF, Savarino SJ, Abu-Elyazeed R, El-Ghoreb N, Hall ER, Naficy A, Abdel-Messih I, Frenck RW, Jr., Svennerholm AM, Clemens JD. 2005. Serologic correlates of protection against enterotoxigenic Escherichia coli diarrhea. J Infect Dis 191:562-70.
48. Kumar P, Kuhlmann FM, Chakraborty S, Bourgeois AL, Foulke-Abel J, Tumala B, Vickers TJ, Sack DA, DeNearing B, Harro CD, Wright WS, Gildersleeve JC, Ciorba MA, Santhanam S, Porter CK, Gutierrez RL, Prouty MG, Riddle MS, Polino A, Sheikh A, Donowitz M, Fleckenstein JM. 2018. Enterotoxigenic Escherichia coli-blood group A interactions intensify diarrheal severity. J Clin Invest 128:3298-3311.

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.20.348110; this version posted October 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

555 49. Roy K, Kansal R, Bartels SR, Hamilton DJ, Shaaban S, Fleckenstein JM. 2011. Adhesin degradation
556 accelerates delivery of heat-labile toxin by enterotoxigenic Escherichia coli. J Biol Chem
557 286:29771-9.
558 50. Roy K, Hilliard GM, Hamilton DJ, Luo J, Ostmann MM, Fleckenstein JM. 2009. Enterotoxigenic
559 Escherichia coli EtpA mediates adhesion between flagella and host cells. Nature 457:594-8.
560 51. Crossman LC, Chaudhuri RR, Beatson SA, Wells TJ, Desvaux M, Cunningham AF, Petty NK, Mahon
561 V, Brinkley C, Hobman JL, Savarino SJ, Turner SM, Pallen MJ, Penn CW, Parkhill J, Turner AK,
562 Johnson TJ, Thomson NR, Smith SG, Henderson IR. 2010. A commensal gone bad: complete
563 genome sequence of the prototypical enterotoxigenic Escherichia coli strain H10407. J Bacteriol
564 192:5822-31.
565 52. Kotloff KL, Nataro JP, Blackwelder WC, Nasrin D, Farag TH, Panchalingam S, Wu Y, Sow SO, Sur D,
566 Breiman RF, Faruque AS, Zaidi AK, Saha D, Alonso PL, Tamboura B, Sanogo D, Onwuchekwa U,
567 Manna B, Ramamurthy T, Kanungo S, Ochieng JB, Omore R, Oundo JO, Hossain A, Das SK,
568 Ahmed S, Qureshi S, Quadri F, Adegbola RA, Antonio M, Hossain MJ, Akinsola A, Mandomando I,
569 Nhampossa T, Acacio S, Biswas K, O'Reilly CE, Mintz ED, Berkeley LY, Muhsen K, Sommerfelt H,
570 Robins-Browne RM, Levine MM. 2013. Burden and aetiology of diarrhoeal disease in infants and
571 young children in developing countries (the Global Enteric Multicenter Study, GEMS): a
572 prospective, case-control study. Lancet 382:209-22.
573 53. Platts-Mills JA, Babji S, Bodhidatta L, Gratz J, Haque R, Havt A, McCormick BJ, McGrath M,
574 Olortegui MP, Samie A, Shakoor S, Mondal D, Lima IF, Hariraju D, Rayamajhi BB, Qureshi S, Kabir
575 F, Yori PP, Mufamadi B, Amour C, Carreon JD, Richard SA, Lang D, Bessong P, Mduma E, Ahmed
576 T, Lima AA, Mason CJ, Zaidi AK, Bhutta ZA, Kosek M, Guerrant RL, Gottlieb M, Miller M, Kang G,
577 Houpt ER, Investigators M-EN. 2015. Pathogen-specific burdens of community diarrhoea in
578 developing countries: a multisite birth cohort study (MAL-ED). Lancet Glob Health 3:e564-75.
579 54. De SN, Bhattacharya K, Sarkar JK. 1956. A study of the pathogenicity of strains of Bacterium coli
580 from acute and chronic enteritis. J Pathol Bacteriol 71:201-9.
581 55. Nalin DR, Bhattacharjee AK, Richardson SH. 1974. Cholera-like toxic effect of culture filtrates of
582 Escherichia coli. J Infect Dis 130:595-607.
583 56. Sack RB, Gorbach SL, Banwell JG, Jacobs B, Chatterjee BD, Mitra RC. 1971. Enterotoxigenic
584 Escherichia coli isolated from patients with severe cholera-like disease. J Infect Dis 123:378-85.
585 57. Roy K, Bartels S, Qadri F, Fleckenstein JM. 2010. Enterotoxigenic Escherichia coli elicits immune
586 responses to multiple surface proteins. Infect Immun 78:3027-35.
587 58. Lanata CF, Black RE. 2018. Estimating the true burden of an enteric pathogen: enterotoxigenic
588 Escherichia coli and Shigella spp. Lancet Infect Dis doi:10.1016/S1473-3099(18)30546-2.

589