(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
21 July 2016 (21.07.2016)

WIPO | PCT

(10) International Publication Number
WO 2016/115328 A1

(51) International Patent Classification:
A61K 39/108 (2006.01)

(21) International Application Number:
PCT/US2016/013377

(22) International Filing Date:
14 January 2016 (14.01.2016)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/103,549 14 January 2015 (14.01.2015) US

(71) Applicant: **WASHINGTON UNIVERSITY** [US/US]; One Brookings Drive, St. Louis, Missouri 63130 (US).

(72) Inventor: **FLECKENSTEIN, James M.**; c/o WASHINGTON UNIVERSITY, One Brookings Drive, St. Louis, Missouri 63130 (US).

(74) Agents: **RILEY-VARGAS, Rebecca C.** et al.; Polsinelli PC, 100 South Fourth Street, Suite 1000, St. Louis, Missouri 63102 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK,

[Continued on next page]

(54) Title: VACCINE COMPOSITIONS FOR USE AGAINST ENTEROTOXIGENIC ESCHERICHIA COLI

(57) Abstract: The present disclosure provides vaccine compositions for use against enterotoxigenic Escherichia coli (ETEC) comprising EtpA and EatA. EtpA and EatA are prevalent amount diverse ETEC isolates and display significant sequence conservation. In an aspect, the disclosure encompasses a vaccine composition comprising EtpA and EatA. In another aspect, the disclosure encompasses a method of protecting against intestinal colonization of enterotoxigenic Escherichia coli (ETEC) in a subject. The method comprises administering to the subject an effective amount of a vaccine composition comprising EatA and EtpA.

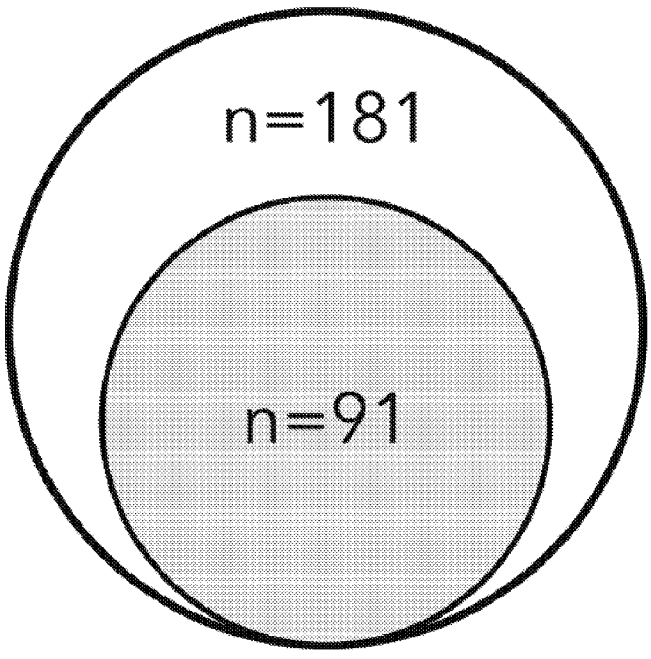


FIG. 1

WO 2016/115328 A1

SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

— *with sequence listing part of description (Rule 5.2(a))*

Published:

— *with international search report (Art. 21(3))*

VACCINE COMPOSITIONS FOR USE AGAINST ENTEROTOXIGENIC *ESCHERICHIA COLI*

GOVERNMENTAL RIGHTS

[0001] This invention was made with government support under R01A1089894 awarded by the NIH. The government has certain rights in the invention.

CROSS REFERENCE TO RELATED APPLICATIONS

[0002] This application claims the benefit of U.S. Provisional Application number 62/103,549, filed January 14, 2015, the disclosure of which is hereby incorporated by reference in its entirety.

FIELD OF THE INVENTION

[0003] The present disclosure provides vaccine compositions for use against enterotoxigenic *Escherichia coli* (ETEC) comprising EtpA and EatA. EtpA and EatA are prevalent amount diverse ETEC isolates and display significant sequence conservation.

BACKGROUND OF THE INVENTION

[0004] The enterotoxigenic *Escherichia coli* (ETEC) are among the most common causes of infectious diarrhea worldwide. Importantly, ETEC are disproportionately represented in cases of severe diarrheal illness as well as in deaths due to diarrhea among young children in developing countries [1].

[0005] These pathogens cause diarrhea by the elaboration and effective delivery of heat-labile and/or heat-stable enterotoxins to intestinal epithelial cells where they stimulate production of cyclic nucleotides ultimately activating the cystic fibrosis transmembrane regulator (CFTR) with resulting net efflux of fluid into the intestinal lumen [2]. Plasmid-encoded colonization factors (CFs), discovered [3] shortly after these organisms were identified as a causative agent of cholera-like diarrheal illness [4–6], are thought to be essential for effective colonization of the small intestine and required for ETEC pathogenesis.

[0006] Following early studies suggesting a pivotal role for these structures [7,8], CF antigens have defined the basis for most subsequent ETEC vaccine efforts [9,10]. However, one factor complicating development of a broadly protective vaccine for ETEC has been the general plasticity of *E. coli* genomes [11], and the significant antigenic heterogeneity of the CFs. To date, at least 26 antigenically distinct CF antigens have been described [12]. The lack of appreciable cross-protection afforded by these antigens combined with the complex landscape of CFs portrayed in ETEC molecular epidemiology studies continue to complicate rational CF antigen selection [13].

[0007] Antigenic heterogeneity, recent failure of LT-toxoid-based vaccine strategies [14,15], as well as the need to optimize the performance of live-attenuated vaccines currently in clinical trials [16–18] have highlighted the need to identify additional virulence molecules that might be targeted in ETEC vaccines.

SUMMARY OF THE INVENTION

[0008] In an aspect, the disclosure encompasses a vaccine composition comprising EtpA and EatA.

[0009] In another aspect, the disclosure encompasses a method of protecting against intestinal colonization of enterotoxigenic *Escherichia coli* (ETEC) in a subject. The method comprises administering to the subject an effective amount of a vaccine composition comprising EatA and EtpA.

[0010] In still another aspect, the disclosure encompasses a method of preventing or treating ETEC-associated diarrhea. The method comprises administering to the subject an effective amount of a vaccine composition comprising EatA and EtpA.

[0011] In still another aspect, the disclosure encompasses a method of preventing or treating an ETEC-associated infection in a subject. The method comprises administering to the subject an effective amount of a vaccine composition comprising EatA and EtpA.

BRIEF DESCRIPTION OF THE FIGURES

[0012] The application file contains at least one drawing executed in color. Copies of this patent application publication with color drawing(s) will be provided by the Office upon request and payment of the necessary fee.

[0013] **FIG. 1** depicts the relationship of strain subsets used in antigen expression studies, and strains with available whole genome sequences. All of the strains in the collection (n = 181) were examined for production of three secreted ETEC virulence proteins EtpA, EatA, and YghJ by immunoblotting of culture supernatants with the respective antibodies. A subset of these strains (n = 91) were recently sequenced at the Genome Sequencing Center for Infectious Diseases (GSCID).

[0014] **FIG. 2A, FIG. 2B** and **FIG. 2C** depict the conservation of novel pathotype-specific antigens EtpA and EatA among phylogenically distinct strains expressing different colonization factors. (**FIG. 2A**) Heatmap of EtpA and EatA showing the proportion of strains positive for expression of these antigens among different CF groups. CF antigen designation is shown at left of the heatmap. nd = no CF antigen detected. Below is the heatmap key depicting colors associated with each degree of antigen positivity. Density line in yellow depicts the relative number of map features assigned at each proportion. Primary data used to construct the heatmap can be found in **Table 5**. (**FIG. 2B**) Immunoblot detection of EtpA and EatA expression among strains from different phylogenies. The upper immunoblot demonstrates EtpA production in the prototype H10407 strain, ThroopD isolated in Dallas, TX in 1975, the Juruá_18/11 (Amazon, 1998), and phylogentically dispersed strains from icddr,b. The *etpA* mutant is included as a negative control. The lower blot demonstrates EatA production by H10407, phylogenically distributed strains from icddr,b and Envira_10/1, an additional isolate from cholera-like outbreaks in the Amazon. The *eatA* mutant is included as a negative control. (**FIG. 2C**) Phlyogram showing the phylogenetic distribution of selected ETEC strains (designations in blue) and reference *E. coli* strains (designations in black). Red circles and gold stars represent *eatA*+, and *etpA*+ strains, respectively.

[0015] **FIG. 3A, FIG. 3B, FIG. 3C** and **FIG. 3D** depict recognition of novel antigens during naturally occurring ETEC infections in Bangladesh. Shown are kinetic

ELISA data for four different recombinant antigens (**FIG. 3A**, rEtpA; **FIG. 3B**, the rEatA passenger domain; **FIG. 3C**, rYghJ; and **FIG. 3D**, rEaeH) obtained with 1:4096 dilutions of convalescent plasma from ETEC-infected patients hospitalized at ICDDR,B in Dhaka, Bangladesh (closed circles), or control patients not infected with enterotoxigenic *E. coli* (open circles). Horizontal bars represent geometric mean V_{max} kinetic ELISA values for each group. P values obtained by two-tailed Mann Whitney testing of groups are summarized (*<0.05; **<0.01; ***<0.001;****<0.0001). x-axis of each graph depicts the specificity of the secondary antibody used in the ELISA (IgG, IgA, IgM, and total IgG, IgA, and IgM).

[0016] **FIG. 4A, FIG. 4B, FIG. 4C** and **FIG. 4D** depict mice immunized with rEtpA and rEatA$_p$H134R are protected against ETEC infection. Serologic responses to (**FIG. 4A**) heat-labile toxin (LT), (**FIG. 4B**) the passenger domain of EatA (Eat$_p$), (**FIG. 4C**) EtpA. Shown are serum IgG responses following intranasal vaccination of mice with the LT adjuvant alone, or LT with 15 μg of either the proteolytically inactive passenger domain (EatA$_p$H134R), EtpA, or both antigens on days 0, 14, 28. Colonization of mice following immunization with the adjuvant alone compared with single and dual antigen vaccination. Comparisons between groups were by Mann Whitney two tailed nonparametric testing. (One mouse died during the vaccination period in the LT/EtpA group and was therefore excluded from the analysis). For mice with no detectable colonies following challenge, the number of cfu is arbitrarily reported as 1 (10^0) cfu, the theoretical limit of detection.

[0017] **FIG. 5A, FIG. 5B, FIG. 5C, FIG. 5D, FIG. 5E, FIG. 5F** and **FIG. 5G** depict EatA passenger domain alignments of predicted EatA sequences corresponding to multiple phylogenies from strains belonging to disparate geographic origins. The conserved catalytic triad at amino acids H78, D106, and S211 is highlighted by gray background shading. Geographic origin of strains is depicted in the color key at left of the alignment. Alignments were performed using CLUSTAL Omega (release 1.2.0 AndreaGiacomo) [40] algorithm plugin for CLC Main Workbench. H10407 (SEQ ID NO:2); ThroopD (SEQ ID NO:16); Envira (10_1) (SEQ ID NO:17); Jurua (18_11) (SEQ ID NO:18); 2850750 (SEQ ID NO:19); 2871950 (SEQ ID NO:20); P03050293.1 (SEQ ID

NO:17); P0304777.1 (SEQ ID NO:21); 2720900 (SEQ ID NO:22); 178900 (SEQ ID NO:23); 180200 (SEQ ID NO:19); 272950 (SEQ ID NO:16); TW10598 (SEQ ID NO:20); TW11681 (SEQ ID NO:16); B2C (SEQ ID NO:20); E24377A (SEQ ID NO:24); Consensus (SEQ ID NO:25).

[0018] **FIG. 6A, FIG. 6B, FIG. 6C, FIG. 6D** and **FIG. 6E** depict EtpA alignments of predicted sequences corresponding to multiple phylogenies from strains belonging to disparate geographic origins. Geographic origin of strains is depicted in the color key at left of the alignment. Alignments were performed using sequence alignment algorithm of CLC Main Workbench v6.9.1 with the following parameters: gap open cost = 10.0; gap extension cost = 1.0; end gap cost = as any other; alignment mode = very accurate (slow); redo alignments = no; use fixpoints = no. H10407 (SEQ ID NO:4); ThroopD (SEQ ID NO:26); Jurua(18/11) (SEQ ID NO:27); 2720900 (SEQ ID NO:4); 2850750 (SEQ ID NO:28); 2871950 (SEQ ID NO:29); P0302308.1 (SEQ ID NO:30); P0304777.1 (SEQ ID NO:31); P0299438.4 (SEQ ID NO:32); 178900 (SEQ ID NO:26); 180200 (SEQ ID NO:28); 1392/75 (SEQ ID NO:29); tw11681 (SEQ ID NO:26); tw14425 (SEQ ID NO:30); TW10828 (SEQ ID NO:28); tw10598 (SEQ ID NO:29); E24377A (SEQ ID NO:29).

[0019] **FIG. 7A, FIG. 7B, FIG. 7C** and **FIG. 7D** depict the immune response to selected ETEC proteins in infected patients and Bangledeshi and US controls. Shown are (IgG) kinetic ELISA responses (in V_{max}, milliunits/min) to recombinant proteins comparing convalescent plasma from patients hospitalized with acute ETEC infections at the International Centre for Diarrhoeal Disease Research in Dhaka, Bangladesh with controls (c) from Bangledeshi adults, and children, as well as plasma from age-matched children from Saint Louis Children's Hospital (slch). Antigens included two plasmid-encoded ETEC specific antigens (**FIG. 7A**) EtpA, and (**FIG. 7B**) the EatA passenger domain; and two chromosomally-encoded conserved antigens (**FIG. 7C**) YghJ, and (**FIG. 7D**) EaeH. All plasma samples were diluted 1:4096.

[0020] **FIG. 8** depicts immune responses to EtpA following volunteer challenge with ETEC H10407. All sera were diluted 1:4096 prior to testing against rEtpA-myc-$_6$His followed by detection of total antibody (IgM,IgG,IgA) in kinetic ELISA.

Pre and post values (open and closed circles, respectively) represent collective data from 2 independent ETEC H10407 challenge studies CIR218 and CIR193a. Data from CIR218 are shown as pre-challenge (d-2, open blue circles) and (d28, closed blue circles), while data from CIR193a appear as open grey circles (pre-challenge, d0) and closed grey circles (post challenge, d9). Dashed horizontal lines represent geometric means. P value represents comparison of pre and post-challenge samples by Mann Whitney 2-tailed analysis.

DETAILED DESCRIPTION OF THE INVENTION

[0021] The present disclosure encompasses vaccine compositions and methods for preventing and treating ETEC-associated infections. The vaccine compositions comprise EatA and EtpA. The inventors demonstrated that EatA and EtpA are broadly represented in a diverse collection of ETEC isolates comprising multiple phylogenetic backgrounds. The prevalence and sequence conservation of EatA and EtpA among ETEC isolates makes them exceptional vaccine candidates. Importantly, vaccine targets should be specific to the pathovar under study or restricted to pathogenic isolates, but not subject to significant antigenic variation. These features are demonstrated by EatA and EtpA.

I. COMPOSITIONS

[0022] Compositions of the invention are directed to vaccine compositions comprising EtpA and EatA. In certain embodiments, the EatA protein comprises a mutation that disrupts its enzymatic activity. Various aspects of the invention will be described in further detail below.

(a) EatA

[0023] In an aspect, the present disclosure encompasses a vaccine composition comprising EatA. The term “vaccine composition” as used herein means a composition that when administered to a subject, typically elicits a protective immune response, where a protective immune response is one that ameliorates one or more symptoms of the target disorder. As used herein, “EatA” refers to the enterotoxigenic *E. coli* autotransporter A. EatA is encoded by the *eatA* gene. EatA is a member of a family

of molecules referred to as serine protease autotransporters of the Enterobacteriaceae (SPATE). EatA modulates both adherence to epithelial cells and intestinal colonization in part by digesting EtpA. Additionally, EatA degrades MUC2, the major mucin secreted by intestinal epithelium. EatA comprises a passenger domain. As used herein, the "passenger domain" comprises a HDS (histidine-aspartate-serine) catalytic triad and a C-terminal β-domain required for extracellular secretion of the passenger domain. A skilled artisan would be able to identify the passenger domain in EatA from various species or strains based on the recited features.

[0024] The nucleotide sequence of *eatA* may be found at GenBank accession number AY163491.2. The amino acid sequence of EatA may be found at GenBank accession number AAO17297.1. Homologs can be found in other species or strains by methods known in the art. For example, sequence similarity may be determined by conventional algorithms, which typically allow introduction of a small number of gaps in order to achieve the best fit. In particular, "percent identity" of two polypeptides or two nucleic acid sequences is determined using the algorithm of Karlin and Altschul (Proc. Natl. Acad. Sci. USA 87:2264-2268, 1993). Such an algorithm is incorporated into the BLASTN and BLASTX programs of Altschul et al. (J. Mol. Biol. 215:403-410, 1990). BLAST nucleotide searches may be performed with the BLASTN program to obtain nucleotide sequences homologous to a nucleic acid molecule of the invention. Equally, BLAST protein searches may be performed with the BLASTX program to obtain amino acid sequences that are homologous to a polypeptide of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST is utilized as described in Altschul et al. (Nucleic Acids Res. 25:3389-3402, 1997). When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g., BLASTX and BLASTN) are employed. See www.ncbi.nlm.nih.gov for more details. Generally a homolog will have a least 80, 81, 82, 83, 84, 85, 86, 87, 88, or 89% homology. In another embodiment, the sequence may be at least 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100% homologous to EatA.

[0025] In a specific embodiment, an EatA of the vaccine composition comprises the full length sequence of ETEC H10407 EatA such as the sequence set

forth in SEQ ID NO:1 (MNKVFSLKYS FLAKGFIAVS ELARRVSVKG KLKSASSIII SPITIAIVSY APPSLAATVN ADISYQTFRD FAENKGAFIV GASNINIYDK NGVLVGVLDK APMPDFSSAT MNTGTLPPGD HTLYSPQYVV TAKHVNGSDI MSFGHIQNNY TVVGENNHNS LDIKIRRLNK IVTEVAPAEI SSVGAVNGAY QEGGRFKAFY RLGGGLQYIK DKNGNLTPVY TNGGFLTGGT ISALSSYNNG QMITAPTGDI FNPANGPLAN YLNKGDSGSP LFAYDSLDKK WVLVGVLSSG SEHGNNWVVT TQDFLHQQPK HDFDKTISYD SEKGSLQWRY NKNSGVGTLS QESVVWDMHG KKGGDLNAGK NLQFTGNNGE IILHDSIDQG AGYLQFFDNY TVTSLTDQTW TGGGIITEKG VNVLWQVNGV NDDNLHKVGE GTLTVNGKGV NNGGLKVGDG TVILNQRPDD NGHKQAFSSI NISSGRATVI LSDANQVNPD KISWGYRGGT LDLNGNNVNF TRLQAADYGA IVSNNNKNKS ELTLKLQTLN ENDISVDVKT YEVFGGHGSP GDLYYVPASN TYFILKSKAY GPFFSDLDNT NVWQNVGHDR DKAIQIVKQQ KIGESSQPYM FHGQLNGYMD VNIHPLSGKD VLTLDGSVNL PEGVITKKSG TLIFQGHPVI HAGMTTSAGQ SDWENRQFTM DKLRLDAATF HLSRNAHMQG DISAANGSTV ILGSSRVFTD KNDGTGNAVS SVEGSSIATT AGDQSYYSGN VLLENHSSLE VRENFTGGIE AYDSSVSVTS QNAIFDHVGS FVNSSLLLEK GAKLTAQSGI FTNNTMKIKE NASLTLTGIP SVGKPGYYSP VTSTTEGIHL GERASLSVKN MGYLSSNITA ENSAAIINLG DSNATIGKTD SPLFSTLMRG YNAVLQGNIM GPQSSVNMNN ALWHSDRNSE LKELKANDSQ IELGVRGHFA KLRVKELIAS NSVFLVHANN SQADQLNVTD KLQGSNNTIL VDFFNKAANG TNVTLITAPK GSDENTFKAG TQQIGFSNIT PEIRTENTDT ATQWVLTGYQ SVADARASKI ATDFMDSGYK SFLTEVNNLN KRMGDLRDSQ GDAGGWARIM NGTGSGESGY RDNYTHVQIG ADRKHELNGI DLFTGALLTY TDNNASSQAF SGKTKSLGGG VYASGLFESG AYFDLIGKYL HHDNRYTLNF ASLGERSYTS HSLYAGAEIG YRYHMSENTW VEPQMELVYG SVSGKSFNWK DQGMQLSMKD KDYHPLIGRT GVDVGRAFSG DTWKVTVRAG LGYQFDLLAN GETVLQDASG KKHFKGEKDS RMLMNVGTNV EVKDNMRFGL ELEKSAFGRY NIDNSINANF RYYF).

[0026] In another specific embodiment, an EatA of the vaccine composition comprises the passenger domain of ETEC H10407 EatA such as the

sequence set forth in SEQ ID NO:2 (ATVN ADISYQTFRD FAENKGAFIV GASNINIYDK NGVLVGVLDK APMPDFSSAT MNTGTLPPGD HTLYSPQYVV TAKHVNGSDI MSFGHIQNNY TVVGENNHNS LDIKIRRLNK IVTEVAPAEI SSVGAVNGAY QEGGRFKAFY RLGGGLQYIK DKNGNLTPVY TNGGFLTGGT ISALSSYNNG QMITAPTGDI FNPANGPLAN YLNKGDSGSP LFAYDSLDKK WVLVGVLSSG SEHGNNWVVT TQDFLHQQPK HDFDKTISYD SEKGSLQWRY NKNSGVGTLS QESVVWDMHG KKGGDLNAGK NLQFTGNNGE IILHDSIDQG AGYLQFFDNY TVTSLTDQTW TGGGIITEKG VNVLWQVNGV NDDNLHKVGE GTLTVNGKGV NNGGLKVGDG TVILNQRPDD NGHKQAFSSI NISSGRATVI LSDANQVNPD KISWGYRGGT LDLNGNNVNF TRLQAADYGA IVSNNNKNKS ELTLKLQTLN ENDISVDVKT YEVFGGHGSP GDLYYVPASN TYFILKSKAY GPFFSDLDNT NVWQNVGHDR DKAIQIVKQQ KIGESSQPYM FHGQLNGYMD VNIHPLSGKD VLTLDGSVNL PEGVITKKSG TLIFQGHPVI HAGMTTSAGQ SDWENRQFTM DKLRLDAATF HLSRNAHMQG DISAANGSTV ILGSSRVFTD KNDGTGNAVS SVEGSSIATT AGDQSYYSGN VLLENHSSLE VRENFTGGIE AYDSSVSVTS QNAIFDHVGS FVNSSLLLEK GAKLTAQSGI FTNNTMKIKE NASLTLTGIP SVGKPGYYSP VTSTTEGIHL GERASLSVKN MGYLSSNITA ENSAAIINLG DSNATIGKTD SPLFSTLMRG YNAVLQGNIM GPQSSVNMNN ALWHSDRNSE LKELKANDSQ IELGVRGHFA KLRVKELIAS NSVFLVHANN SQADQLNVTD KLQGSNNTIL VDFFNKAANG TNVTLITAPK GSDENTFKAG TQQIGFSNIT PEIRTENTDT ATQWVLTGYQ SVADARASKI ATDFMDSGYK SFLTEVNNLN KRMGDLRD).

[0027] In some embodiments, an EatA of the vaccine composition is a sequence of EatA comprising at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, or 89% identity to SEQ ID NO:1 or SEQ ID NO:2. In another embodiment, an EatA of the vaccine composition is a sequence of EatA comprising at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 100% identity to SEQ ID NO:1 or SEQ ID NO:2. In a specific

embodiment, an EatA of the vaccine composition is a sequence of EatA comprising about 95% to about 100% identity to SEQ ID NO:1 or SEQ ID NO:2.

[0028] In a specific embodiment, an EatA of the vaccine composition comprises a mutation that disrupts the serine protease activity. Methods known in the art may be used to determine if a mutation in EatA results in disruption in serine protease activity. For example, the ability of mutated EatA to cleave substrate may be determined. EatA possesses serine protease activity that is abolished by mutations within a serine protease catalytic triad formed by residues H134, D162, and S267. The catalytic triad is universally conserved within the passenger domain of EatA. Accordingly, an EatA of the vaccine composition may comprise one or more mutations in H134, D162, and/or S267, wherein the mutation disrupts serine protease activity. More specifically, an EatA of the vaccine composition may comprise one or mutations selected from the group consisting of H134A, D162A, and S267G, relative to SEQ ID NO:1.

[0029] In some embodiments, an EatA of the vaccine composition is a sequence of EatA comprising a mutation that disrupts the serine protease activity and has at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, or 89% identity to SEQ ID NO:1 or SEQ ID NO:2. In another embodiment, an EatA of the vaccine composition is a sequence of EatA comprising a mutation that disrupts the serine protease activity and has at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 100% identity to SEQ ID NO:1 or SEQ ID NO:2. In a specific embodiment, an EatA of the vaccine composition is a sequence of EatA comprising a mutation that disrupts serine protease activity and has about 95% to about 100% identity to SEQ ID NO:1 or SEQ ID NO:2.

[0030] In other embodiments, an EatA of the vaccine composition is a sequence of EatA comprising one or more mutations in H134, D162, and/or S267 that disrupts the serine protease activity and has at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, or 89% identity to SEQ ID NO:1 or SEQ ID NO:2. In another embodiment, an EatA of the

vaccine composition is a sequence of EatA comprising one or more mutations in H134, D162, and/or S267 that disrupts the serine protease activity and has at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 100% identity to SEQ ID NO:1 or SEQ ID NO:2. In a specific embodiment, an EatA of the vaccine composition is a sequence of EatA comprising one or more mutations in H134, D162, and/or S267 that disrupts serine protease activity and has about 95% to about 100% identity to SEQ ID NO:1 or SEQ ID NO:2.

[0031] In still other embodiments, an EatA of the vaccine composition is a sequence of EatA comprising one or more mutations selected from the group consisting of H134A, D162A, and S267G and has at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, or 89% identity to SEQ ID NO:1 or SEQ ID NO:2. In another embodiment, an EatA of the vaccine composition is a sequence of EatA comprising one or more mutations selected from the group consisting of H134A, D162A, and S267G that disrupts the serine protease activity and has at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 100% identity to SEQ ID NO:1 or SEQ ID NO:2. In a specific embodiment, an EatA of the vaccine composition is a sequence of EatA comprising one or more mutations selected from the group consisting of H134A, D162A, and S267G that disrupts serine protease activity and has about 95% to about 100% identity to SEQ ID NO:1 or SEQ ID NO:2.

[0032] In any of the foregoing embodiments, an EatA of the vaccine composition may be a truncated version of EatA provided it as the same activity as the full length or passenger domain of EatA (e.g. elicits a protective immune response or, stated another way, is antigenic). For example, the truncated version of EatA may be about 10, about 20, about 30, about 40, about 50, about 60, about 70, about 80, about 90, about 100, about 110, about 120, about 130, about 140, about 150, about 160, about 170, about 180, about 190, about 200, about 210, about 220, about 230, about 240, about 250, about 260, about 270, about 280, about 290, about 300, about 310,

about 320, about 330, about 340, about 350, about 360, about 370, about 380, about 390, about 400, about 410, about 420, about 430, about 440, about 450, about 460, about 470, about 480, about 490, about 500, about 510, about 520, about 530, about 540, about 550, about 560, about 570, about 580, about 590, about 600, about 610, about 620, about 630, about 640, about 650, about 660, about 670, about 680, about 690, about 700, about 710, about 720, about 730, about 740, about 750, about 760, about 770, about 780, about 790, about 800, about 810, about 820, about 830, about 840, about 850, about 860, about 870, about 880, about 890, about 900, about 910, about 920, about 930, about 940, about 950, about 960, about 970, about 980, about 990, about 1000, about 1010, about 1020, about 1030, about 1040, about 1050, about 1060, about 1070, about 1080, about 1090, about 1100, about 1110, about 1120, about 1130, about 1140, about 1150, about 1160, about 1170, about 1180, about 1190, about 1200, about 1210, about 1220, about 1230, about 1240, about 1250, about 1260, about 1270, about 1280, about 1290, about 1300, about 1310, about 1320, about 1330, about 1340, about 1350, about 1360, or about 1370 amino acids, provided it as the same activity as the full length or passenger domain of EatA (elicits a protective immune response or, stated another way, is antigenic).

(b) EtpA

[0033] In an aspect, the present disclosure encompasses a vaccine composition comprising EtpA. As used herein, "EtpA" refers to the ETEC two-partner secretion locus (*etpBAC*) protein A and is encoded by the *etpA* gene. EtpA is a member of a family of virulence proteins (generically referred to as TpsA proteins) that are secreted by two-partner secretion (TPS). EtpA is an exoprotein adhesin molecule and plays a critical role in bacterial adhesion *in vitro* and in the colonization of mucosal surfaces *in vivo*. EtpA has a conserved secretion domain in its amino terminus comprising Asn-Pro-Asn-Gly-Val (SEQ ID NO:5) at amino acids 150 to 154 and several repeat regions in the carboxy-terminus comprising four major repeat units (~226 amino acids) preceded by a 173-amino-acid partial repeat beginning at amino acid S648.

[0034] The nucleotide sequence of *etpA* may be found at GenBank accession number AY920525.2. The amino acid sequence of EtpA may be found at

GenBank accession number AAX13509.2. Homologs can be found in other species or strains by methods known in the art. For example, sequence similarity may be determined by conventional algorithms, which typically allow introduction of a small number of gaps in order to achieve the best fit. In particular, "percent identity" of two polypeptides or two nucleic acid sequences is determined using the algorithm of Karlin and Altschul (Proc. Natl. Acad. Sci. USA 87:2264-2268, 1993). Such an algorithm is incorporated into the BLASTN and BLASTX programs of Altschul et al. (J. Mol. Biol. 215:403-410, 1990). BLAST nucleotide searches may be performed with the BLASTN program to obtain nucleotide sequences homologous to a nucleic acid molecule of the invention. Equally, BLAST protein searches may be performed with the BLASTX program to obtain amino acid sequences that are homologous to a polypeptide of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST is utilized as described in Altschul et al. (Nucleic Acids Res. 25:3389-3402, 1997). When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g., BLASTX and BLASTN) are employed. See www.ncbi.nlm.nih.gov for more details. Generally a homolog will have a least 80, 81, 82, 83, 84, 85, 86, 87, 88, or 89% homology. In another embodiment, the sequence may be at least 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100% homologous to EtpA.

[0035] In a specific embodiment, an EtpA of the vaccine composition comprises the full length sequence of ETEC H10407 EtpA such as the sequence set forth in SEQ ID NO:3 (MNRIYKLKFD KRRNELVVVS EITTGVGNAK ATGSVEGEKS PRRGVRAMAL SLLSGMMIMA HPAMSANLPT GGQIVAGSGS IQTPSGNQMN IHQNSQNMVA NWNSFDIGKG NTVQFDQPSS SAVALNRVVG GGESQIMGNL KANGQVFLVN PNGVLFGEGA SVSTSGFVAS TRDIKNDDFM NRRYTFSGGQ KAGAAIVNQG ELTTNAGGYI VLAADRVSNS GTIRTPGGKT VLAASERITL QLDNGGLMSV QVTGDVVNAL VENRGLVSAR DGQVYLTALG RGMLMNTVLN VSGVVEASGM HRQDGNIVLD GGDSGVVHLS GTLQADNASG QGGKVVVQGK NILLDKGSNI TATGGQGGGE VYVGGGWQGK DSNIRNADKV VMQGGARIDV SATQQGNGGT AVLWSDSYTN FHGQISAKGG ETGGNGGRVE TSSHGNLQAF GTVSASAKKG KAGNWLLDSA DITIVNGSNV SKTETTQSPP HTQFAPTAAG

SAVSNTSINN RLNNGTSVTI LTHRTRTGTA QGGNITVNAA INKSNGSDVN LTLQAGGNIT VNNSITSTEG KLNVNLSGAR TSNGSITISN NANITTNGGD ITVGTTNTSN RVNISINNTT LNASNGNIQL TGTGTDSGIL FAGNNRLTAS NIALTGNSTS GNAINLTGTA TLNATNNITL TGSSTSGNAI NLKGNNTLTA SNITLTGEST SGNAINLTDT TGTTTLNATN NITMQGTRVQ IKHSNITAGN FALNATVAGS EISNTTLTAT NNINLAAKTN SASSGVYLKD ARITSTNGSI TANGTATANG KATHLDGNVT LNASNGRIKL TGNGHGSASG ILFAGNNRLT ASNIALTGNS TSGNAINLTG TATLNATNDI TLTGSSTSGN AINLTGTATL NATNNITLTG SSTSGNAINL KGNNTLTASN ITLTGESTSG NAINLTDTTG TTTLNATNNI TMQGTRVQIK HSNITAGNFA LNATVAGSEI SNTTLTATNN INLAAKTNSA SSGVYLKDAR ITSTNGSITA NGTATANGKA THLDGNVTLN ASNGRIKLTG NGHGSASGIL FAGNNRLTAS NIALTGNSTS GNAINLTGTA TLNATNDITL TGSSTSGNAI NLTGTATLNA TNNITLTGSS TSGNAINLKG NNTLTASNIT LTGESTSGNA INLTDTTGTT TLNATNNITM QGTRVQIKHS NITAGNFALN ATVAGSEISN TTLTATNNIN LAAKTNSASS GVYLKDARIT STNGSITANG TATANGKATH LDGNVTLNAS NGRIKLTGNG HGSASGILFA GNNRLTASNI ALTGNSTSGN AINLTGTATL NATNDITLTG SSTSGNAINL TGTATLNATN NITLTGSSTS GNAINLKGNN TLTASNITLT GESTSGNAIN LTDTTGTTTL NATNNITMQG TRVQIKHSNI TAGNFALNAT VAGSEISNTT LTATNNINLA AKTNSASSGV YLKDARITST NGSITTNGTA TANGKATHLD GNVTLNASNG RIKLTGNGHG SASGILFAGN NRLTASNIAL TGNSTSGNAI NLTGTATLNA TNDITLTGSS TSGNAINLTG TATLNATNNI TLTGSSTSGN AINLKGNNTL TASNITLTGE STSGNAINLT DTTGTTTLNA TNNITMQGTR VQIKHSNITA GNFALNATVA GSEISNTTLT ATNNINLAAK TNSASSGVYL KDARITSTNG SITANGTAPA NDNATYLDGN VTLNASNGSI KLTGNGNGST SGILFAGNNT LTASNITLTG NSEVYWQ).

[0036] In another specific embodiment, an EtpA of the vaccine composition comprises the secreted portion of ETEC H10407 EtpA such as the sequence set forth in SEQ ID NO:4 (MNRIYKLKFD KRRNELVVVS EITTGVGNAK ATGSVEGEKS PRRGVRAMAL SLLSGMMIMA HPAMSANLPT GGQIVAGSGS

IQTPSGNQMN IHQNSQNMVA NWNSFDIGKG NTVQFDQPSS SAVALNRVVG GGESQIMGNL KANGQVFLVN PNGVLFGEGA SVSTSGFVAS TRDIKNDDFM NRRYTFSGGQ KAGAAIVNQG ELTTNAGGYI VLAADRVSNS GTIRTPGGKT VLAASERITL QLDNGGLMSV QVTGDVVNAL VENRGLVSAR DGQVYLTALG RGMLMNTVLN VSGVVEASGM HRQDGNIVLD GGDSGVVHLS GTLQADNASG QGGKVVVQGK NILLDKGSNI TATGGQGGGE VYVGGGWQGK DSNIRNADKV VMQGGARIDV SATQQGNGGT AVLWSDSYTN FHGQISAKGG ETGGNGGRVE TSSHGNLQAF GTVSASAKKG KAGNWLLDSA DITIVNGSNV SKTETTQSPP HTQFAPTAAG SAVSNTSINN RLNNGTSVTI LTHRTRTGTA QGGNITVNAA INKSNGSDVN LTLQAGGNIT VNNSITSTEG KLNVNLSGAR TSNGSITISN NANITTNGGD ITVGTTNTSN).

[0037] In some embodiments, an EtpA of the vaccine composition is a sequence of EtpA comprising at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, or 89% identity to SEQ ID NO:3 or SEQ ID NO:4. In another embodiment, an EtpA of the vaccine composition is a sequence of EtpA comprising at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 100% identity to SEQ ID NO:3 or SEQ ID NO:4. In a specific embodiment, an EtpA of the vaccine composition is a sequence of EtpA comprising about 94% to about 100% identity to SEQ ID NO:3 or SEQ ID NO:4.

[0038] In any of the foregoing embodiments, an EtpA of the vaccine composition may be a truncated version of EtpA provided it as the same activity as the full length or secreted portion of EtpA (elicits a protective immune response or, stated another way, is antigenic). For example, the truncated version of EtpA may be about 10, about 20, about 30, about 40, about 50, about 60, about 70, about 80, about 90, about 100, about 110, about 120, about 130, about 140, about 150, about 160, about 170, about 180, about 190, about 200, about 210, about 220, about 230, about 240, about 250, about 260, about 270, about 280, about 290, about 300, about 310, about 320, about 330, about 340, about 350, about 360, about 370, about 380, about 390, about 400, about 410, about 420, about 430, about 440, about 450, about 460, about 470,

about 480, about 490, about 500, about 510, about 520, about 530, about 540, about 550, about 560, about 570, about 580, about 590, about 600, about 610, about 620, about 630, about 640, about 650, about 660, about 670, about 680, about 690, about 700, about 710, about 720, about 730, about 740, about 750, about 760, about 770, about 780, about 790, about 800, about 810, about 820, about 830, about 840, about 850, about 860, about 870, about 880, about 890, about 900, about 910, about 920, about 930, about 940, about 950, about 960, about 970, about 980, about 990, about 1000, about 1010, about 1020, about 1030, about 1040, about 1050, about 1060, about 1070, about 1080, about 1090, about 1100, about 1110, about 1120, about 1130, about 1140, about 1150, about 1160, about 1170, about 1180, about 1190, about 1200, about 1210, about 1220, about 1230, about 1240, about 1250, about 1260, about 1270, about 1280, about 1290, about 1300, about 1310, about 1320, about 1330, about 1340, about 1350, about 1360, about 1370, about 1380, about 1390, about 1400, about 1410, about 1420, about 1430, about 1440, about 1450, about 1460, about 1470, about 1480, about 1490, about 1500, about 1510, about 1520, about 1530, about 1540, about 1550, about 1560, about 1570, about 1580, about 1590, about 1600, about 1610, about 1620, about 1630, about 1640, about 1650, about 1660, about 1670, about 1680, about 1690, about 1700, about 1710, about 1720, about 1730, about 1740, about 1750, about 1760, or about 1770 amino acids, provided it as the same activity as the full length or secreted portion of EtpA (elicits a protective immune response or, stated another way, is antigenic).

(c) combinations

[0039] In an embodiment, the vaccine composition may further comprise other immunogenic ETEC proteins. For example, the vaccine composition may further comprise colonization factor (CF) antigens. Colonization factors are proteinaceous surface appendages that facilitate adherence of bacteria to the intestinal epithelium of the host. Generally, CFs are fimbrial, fibrillar, or afimbrial structures. There are 22 different CF variants currently described (including colonization factor antigen I [CFA/I]), and the majority are usually designated “CS” followed by a number that indicates their placement in an order arranged according to the date of discovery. Non-limiting

examples of CFs include CFA/I, CS1, CS2, CS3, CS4, CS5, CS6, CS7, CS8, CS9, CS10, CS11, CS12, CS13, CS14, CS15, CS16, CS17, CS18, CS19, CS20, CS21, CS22, and CS23. The vaccine composition may also further comprise nonclassical adhesins such as Tia and TibA.

(d) composition

[0040] In an aspect, a vaccine composition of the invention comprises EatA and EtpA. In an embodiment, a vaccine composition of the invention comprises EatA and EtpA linked together. The EatA and EtpA may be linked together by various methods known in the art. Suitable linkers include amino acid chains and alkyl chains functionalized with reactive groups for coupling to EatA and EtpA. In an embodiment, the linker may include amino acid side chains, referred to as a peptide linker. Amino acid residue linkers are usually at least one residue and can be 40 or more residues, more often 1 to 10 residues, but do not comprise EatA or EtpA. Typical amino acid residues used for linking are tyrosine, cysteine, lysine, glutamic and aspartic acid, or the like.

[0041] In another embodiment, an alkyl chain linking group may be coupled to EatA and EtpA by reacting the amino group of the N-terminal residue of EatA with a first functional group on the alkyl chain, such as a carboxyl group or an activated ester. Subsequently, EtpA is attached to the alkyl chain to complete the formation of the complex by reacting a second functional group on the alkyl chain with an appropriate group on EtpA. The second functional group on the alkyl chain is selected from substituents that are reactive with a functional group on EtpA while not being reactive with the N-terminal residue of EatA. The process may also be reversed.

[0042] An alternative chemical linking group to an alkyl chain is polyethylene glycol (PEG), which is functionalized in the same manner as the alkyl chain described above. The EatA and/or EtpA of the invention may be PEGylated for improved systemic half-life and reduced dosage frequency. In an embodiment, PEG may be added to a linker.

[0043] The vaccine compositions of the invention may include a pharmaceutically acceptable excipient such as a suitable adjuvant. Suitable adjuvants

include an aluminum salt such as aluminum hydroxide or aluminum phosphate, but may also be a salt of calcium, iron or zinc, or may be an insoluble suspension of acylated tyrosine, or acylated sugars, or may be cationically or anionically derivatised saccharides, polyphosphazenes, biodegradable microspheres, monophosphoryl lipid A (MPL), lipid A derivatives (e.g. of reduced toxicity), 3-O-deacylated MPL [3D-MPL], quil A, Saponin, QS21, Freund's Incomplete Adjuvant (Difco Laboratories, Detroit, Mich.), Merck Adjuvant 65 (Merck and Company, Inc., Rahway, N.J.), AS-2 (Smith-Kline Beecham, Philadelphia, Pa.), CpG oligonucleotides, bioadhesives and mucoadhesives, microparticles, liposomes, polyoxyethylene ether formulations, polyoxyethylene ester formulations, muramyl peptides or imidazoquinolone compounds (e.g. imiquamod and its homologues). Human immunomodulators suitable for use as adjuvants in the invention include cytokines such as interleukins (e.g. IL-1, IL-2, IL-4, IL-5, IL-6, IL-7, IL-12, etc), macrophage colony stimulating factor (M-CSF), tumour necrosis factor (TNF), granulocyte, macrophage colony stimulating factor (GM-CSF) may also be used as adjuvants. In a specific embodiment, the adjuvant is heat-labile toxin (LT) or double mutant heat-labile toxin (LT).

[0044] Vaccines of the invention will typically, in addition to the antigenic and adjuvant components mentioned above, comprise one or more “pharmaceutically acceptable carriers or excipients”, which include any excipient that does not itself induce the production of antibodies harmful to the individual receiving the composition. Suitable excipients are typically large, slowly metabolised macromolecules such as proteins, saccharides, polylactic acids, polyglycolic acids, polymeric amino acids, amino acid copolymers, sucrose (Paoletti et al., 2001, Vaccine, 19:2118), trehalose (WO 00/56365), lactose and lipid aggregates (such as oil droplets or liposomes). Such carriers are well known to those of ordinary skill in the art. The vaccines may also contain diluents, such as water, saline, glycerol, etc. Additionally, auxiliary substances, such as wetting or emulsifying agents, pH buffering substances, and the like, may be present. Sterile pyrogen-free, phosphate buffered physiologic saline is a typical carrier. A thorough discussion of pharmaceutically acceptable excipients is available in

reference Gennaro, 2000, Remington: The Science and Practice of Pharmacy, 20th edition, ISBN:0683306472.

[0045] Compositions of the invention may be lyophilised or in aqueous form, i.e. solutions or suspensions. Liquid formulations of this type allow the compositions to be administered direct from their packaged form, without the need for reconstitution in an aqueous medium, and are thus ideal for injection. Compositions may be presented in vials, or they may be presented in ready filled syringes. The syringes may be supplied with or without needles. A syringe will include a single dose of the composition, whereas a vial may include a single dose or multiple doses (e.g. 2 doses).

[0046] Liquid vaccines of the invention are also suitable for reconstituting other vaccines from a lyophilized form. Where a vaccine is to be used for such extemporaneous reconstitution, the invention provides a kit, which may comprise two vials, or may comprise one ready-filled syringe and one vial, with the contents of the syringe being used to reconstitute the contents of the vial prior to injection.

[0047] Vaccines of the invention may be packaged in unit dose form or in multiple dose form (e.g. 2 doses). For multiple dose forms, vials are preferred to pre-filled syringes. Effective dosage volumes can be routinely established, but a typical human dose of the composition for injection has a volume of 0.5 mL.

[0048] In one embodiment, vaccines of the invention have a pH of between 6.0 and 8.0, in another embodiment, vaccines of the invention have a pH of between 6.3 and 6.9, e.g. 6.6±0.2. Vaccines may be buffered at this pH. Stable pH may be maintained by the use of a buffer. If a composition comprises an aluminum hydroxide salt, a histidine buffer may be used (WO03/009869). The composition should be sterile and/or pyrogen free.

[0049] Compositions of the invention may be isotonic with respect to humans.

[0050] Vaccines of the invention may include an antimicrobial, particularly when packaged in a multiple dose format. Antimicrobials may be used, such as 2-phenoxyethanol or parabens (methyl, ethyl, propyl parabens). Any preservative is

preferably present at low levels. Preservative may be added exogenously and/or may be a component of the bulk antigens which are mixed to form the composition (e.g. present as a preservative in pertussis antigens).

[0051] Vaccines of the invention may comprise a detergent e.g. a Tween (polysorbate), such as Tween 80. Detergents are generally present at low levels e.g. <0.01%.

[0052] Vaccines of the invention may include sodium salts (e.g. sodium chloride) to give tonicity. The composition may comprise sodium chloride. In one embodiment, the concentration of sodium chloride in the composition of the invention is in the range of 0.1 to 100 mg/mL (e.g. 1-50 mg/mL, 2-20 mg/mL, 5-15 mg/mL) and in a further embodiment the concentration of sodium chloride is 10±2 mg/mL NaCl e.g. about 9 mg/mL.

[0053] Vaccines of the invention will generally include a buffer. A phosphate or histidine buffer is typical.

[0054] Vaccines of the invention may include free phosphate ions in solution (e.g. by the use of a phosphate buffer) in order to favor non-adsorption of antigens. The concentration of free phosphate ions in the composition of the invention is in one embodiment between 0.1 and 10.0 mM, or in another embodiment between 1 and 5 mM, or in a further embodiment about 2.5 mM.

[0055] In other embodiments, the vaccine composition may be a live attenuated vaccine vector strain. A live attenuated vaccine vector strain is a strain that has reduced virulence but is still viable. In such an embodiment, EatA and EtpA are expressed in the attenuated strain. Importantly, in such an embodiment, the EatA comprises a mutation in a least two of the three catalytic triad amino acids. Additional inactivating mutations in EatA will limit the likelihood that a mutation *in vivo* will allow reversion. In certain embodiments, the live attenuated vaccine vector is *Salmonella typhi* Ty21a. In other embodiments, the live attenuated vaccine vector is any of the three attenuated strains of ACE527.

II. METHODS

[0056] In an aspect, the present disclosure encompasses a method of protecting against intestinal colonization of enterotoxigenic *Escherichia coli* (ETEC) in a subject. The method comprises administering to the subject an effective amount of a vaccine composition comprising EatA and EtpA. As used herein, “enterotoxigenic *Escherichia coli*” or “ETEC” is an *E. coli* strain able to produce at least one of two types of enterotoxins, the heat-labile toxin (LT) and/or the heat-stable toxin (ST). LT and/or ST are responsible for the movement of electrolytes and water from the intestinal cells to the intestinal lumen, resulting in watery diarrhea. Accordingly, the present disclosure also encompasses a method of preventing or treating ETEC-associated diarrhea.

[0057] In another aspect, the present disclosure encompasses a method of preventing or treating an ETEC-associated infection in a subject. The method comprises administering to the subject an effective amount of a vaccine composition comprising EatA and EtpA. The term "infection" as used herein includes presence of microbes, including bacteria, in or on a subject, which, if its growth were inhibited, would result in a benefit to the subject. As such, the term "infection" in addition to referring to the presence of bacteria also refers to normal flora which, are not desirable. The term "infection" includes infection caused by bacteria. An “ETEC-associated infection” is an infection caused by enterotoxigenic *Escherichia coli* (ETEC). Non-limiting examples of infections that may be prevented or treated using the compositions and/or methods of the invention include: infectious diarrhea, gastroenteritis or travelers’ diarrhea.

[0058] The term "treat", "treating" or "treatment" as used herein refers to administering a pharmaceutical composition of the invention for prophylactic and/or therapeutic purposes. The term "prophylactic treatment" refers to treating a subject who is not yet infected, but who is susceptible to, or otherwise at a risk of infection. The term "therapeutic treatment" refers to administering treatment to a subject already suffering from infection. The term "treat", "treating" or "treatment" as used herein also refers to administering a pharmaceutical composition of the invention in order to: (i) reduce or eliminate either an ETEC-associated infection or one or more symptoms of the ETEC-associated infection, or (ii) retard the progression of an ETEC-associated infection or of

one or more symptoms of the ETEC-associated infection, or (iii) reduce the severity of an ETEC-associated infection or of one or more symptoms of the ETEC-associated infection, or (iv) suppress the clinical manifestation of an ETEC-associated infection, or (v) suppress the manifestation of adverse symptoms of the ETEC-associated infection. Non-limiting examples of symptoms of an ETEC-associate infection include profuse watery diarrhea, abdominal cramping, fever, nausea with or without vomiting, chills, loss of appetite, headache, muscle aches and bloating.

[0059] The term "control" or "controlling" as used herein generally refers to preventing, reducing, or eradicating an ETEC-associated infection or inhibiting the rate and extent of such an infection, or reducing the microbial population, such as a microbial population present in or on a body or structure, surface, liquid, subject, etc, wherein such prevention or reduction in the ETEC-associated infection or microbial population is statistically significant with respect to untreated infection or population. In general, such control may be achieved by increased mortality amongst the microbial population.

[0060] The compositions of the present invention may be used to protect or treat a subject susceptible to infection by ETEC by means of administering said composition directly to a subject. The term "administration" or "administering" includes delivery of a composition or one or more pharmaceutically active ingredients to a subject, including for example, by any appropriate methods, which serves to deliver the composition or its active ingredients or other pharmaceutically active ingredients to the site of the infection. The method of administration can vary depending on various factors, such as for example, the components of the pharmaceutical composition or the type/nature of the pharmaceutically active or inert ingredients, the site of the potential or actual infection, the microorganism involved, severity of the infection, age and physical condition of the subject. Direct delivery may be accomplished by parenteral injection (intramuscularly, intraperitoneally, intradermally, subcutaneously, intravenously, or to the interstitial space of a tissue); or by rectal, oral, vaginal, topical, transdermal, intranasal, ocular, aural, pulmonary or other mucosal administration. In one embodiment, administration is by intramuscular injection to the thigh or the upper arm.

Injection may be via a needle (e.g. a hypodermic needle, electroporation device), but needle free injection may alternatively be used. A typical intramuscular dose is 0.5 mL. In another embodiment, administration is intranasal administration. The composition can be administered prophylactically (i.e. to prevent infection) or therapeutically (i.e. to treat infection). An immune response is preferably protective. The method may raise a booster response.

[0061] The invention provides a method for preventing of treating ETEC-associated infection in a subject, comprising the step of administering an effective amount of a composition of the invention. The term "effective amount" as used herein refers to an amount, which has a therapeutic effect or is the amount required to produce a therapeutic effect in a subject. For example, a therapeutically or pharmaceutically effective amount of a composition is the amount of the antigen required to produce a desired therapeutic effect as may be judged by clinical trial results and/or model animal infection studies. The effective or pharmaceutically effective amount depends on several factors, including but not limited to, the route of administration, the microorganism (e.g. bacteria) involved, characteristics of the subject (for example height, weight, sex, age and medical history), severity of infection, location of infection, and/or the particular type of antigen used. For prophylactic treatments, a therapeutically or prophylactically effective amount is that amount which would be effective to prevent a microbial (e.g. bacterial) infection.

[0062] The effective amount of antigen in each vaccine dose is selected as an amount which induces an immunoprotective response without significant adverse side effects in typical vaccines. Accordingly, the exact amount of the antigen that is required to elicit such a response will vary from subject to subject, depending on the species, age, and general condition of the subject, the severity of the condition being treated, the particular carrier or adjuvant being used and its mode of administration, and the like. Generally it is expected that each dose will comprise 1-1000 μg of total antigen, or 1-100 μg, or 1-40 μg, or 1-5 μg, or less than 1 μg. An optimal amount for a particular vaccine can be ascertained by studies involving observation of antibody titres and other responses in subjects. In certain embodiments, the vaccine composition is administered

at a dose ranging from about 50 to 150 µg. In another embodiment, the vaccine composition is administered at a dose of about 100 µg. In an exemplary embodiment, the vaccine composition is administered at a dose ranging from about 15 to about 30 µg.

[0063] Following initial administration of a vaccine composition of the disclosure, subjects may receive one or several additional administrations of the composition adequately spaced. Dosing treatment can be a single dose schedule or a multiple dose schedule. Suitable timing between doses (e.g. between 2-16 weeks) can be routinely determined.

[0064] In the prevention of an infection, a composition of the invention may be administered as multiple doses prior to infection. In the treatment of an infection, a composition of the invention may be administered as multiple doses following infection. Administration may be daily, twice daily, weekly, twice weekly, monthly, twice monthly, every 6 weeks, every 3 months, every 6 months or yearly. For example, administration may be every 2 weeks, every 3 weeks every 4 weeks, every 5 weeks, every 6 weeks, every 7 weeks, every 8 weeks, every 9 weeks, every 10 weeks, every 11 weeks or every 12 weeks. Alternatively, administration may be every 1 month, every 2 months, every 3 months, every 4 months, every 5 months, every 6 months, every 7 months, every 8 months, every 9 months, every 10 months, every 11 months or every 12 months. Still further, administration may be every 1 year, every 2 years, every 3 years, every 4 years, every 5 years, every 6 years, every 7 years, every 8 years, every 9 years, every 10 years, every 15 years or every 20 years. The duration of treatment can and will vary depending on the subject and the infection to be prevented or treated. For example, the duration of treatment may be for 1 week, 2 weeks, 3 weeks, 4 weeks, 5 weeks, 6 weeks. Alternatively, the duration of treatment may be for 1 month, 2 months, 3 months, 4 months, 5 months, 6 months, 7 months, 8 months, 9 months, 10 months, 11 months, or 12 months. In still another embodiment, the duration of treatment may be for 1 year, 2 years, 3 years, 4 years, 5 years, or greater than 5 years. It is also contemplated that administration may be frequent for a period of time and then administration may be spaced out for a period of time. For example, administration may

be every 4 weeks for 6 months to a year and then administration may be every year thereafter. The duration of treatment may also depend on the length of time the subject may be exposed to ETEC. For example, when the subject resides in or visits a region endemic for ETEC, the duration of treatment may be based on the length of time the subject spends in that region. In a specific embodiment, the duration of treatment may be once a day for the duration of time the subject is in the endemic region. A skilled artisan would be able to determine the effective dosing regimen based on the medical history and subject characteristics.

[0065] A vaccine composition of the disclosure may be administered in combination with standard treatments for ETEC-associated infection. Non-limiting examples of standard treatments for ETEC-associated infection include administration of clear liquids, packaged or premixed oral rehydration salts, chicken soup, bismuth subsalicylate, and antibiotics. Further, a vaccine composition of the disclosure may be administered in combination with other vaccine compositions for ETEC-associated infection. Non-limiting examples of other ETEC vaccine compositions include ACE527, ETVAX, LTR192G with dscCfaE, Vivotif + Dukoral, and ETEC/rCTB.

[0066] As used herein, "subject" or "patient" is used interchangeably. Suitable subjects include, but are not limited to, a human, a livestock animal, a companion animal, a lab animal, and a zoological animal. In one embodiment, the subject may be a rodent, e.g. a mouse, a rat, a guinea pig, etc. In another embodiment, the subject may be a livestock animal. Non-limiting examples of suitable livestock animals may include pigs, cows, horses, goats, sheep, llamas and alpacas. In yet another embodiment, the subject may be a companion animal. Non-limiting examples of companion animals may include pets such as dogs, cats, rabbits, and birds. In yet another embodiment, the subject may be a zoological animal. As used herein, a “zoological animal” refers to an animal that may be found in a zoo. Such animals may include non-human primates, large cats, wolves, and bears. In specific embodiments, the animal is a laboratory animal. Non-limiting examples of a laboratory animal may include rodents, canines, felines, and non-human primates. In certain embodiments, the

animal is a rodent. Non-limiting examples of rodents may include mice, rats, guinea pigs, etc. In a preferred embodiment, the subject is human.

[0067] A subject may be a subject at risk of ETEC infection. Non-limiting examples of subjects at risk of ETEC infection include subjects residing in an ETEC endemic region, subjects traveling to an ETEC endemic region, subjects residing in a location with a recent outbreak of ETEC, subjects suspected of ingesting ETEC-contaminated food or water. In certain embodiments, a subject may be an infant, toddler or young child.

EXAMPLES

[0068] The following examples are included to demonstrate preferred embodiments of the invention. It should be appreciated by those of skill in the art that the techniques disclosed in the examples that follow represent techniques discovered by the inventors to function well in the practice of the invention, and thus can be considered to constitute preferred modes for its practice. However, those of skill in the art should, in light of the present disclosure, appreciate that many changes can be made in the specific embodiments which are disclosed and still obtain a like or similar result without departing from the spirit and scope of the invention.

Introduction to the Examples.

[0069] The enterotoxigenic *Escherichia coli* (ETEC) are among the most common causes of infectious diarrhea worldwide. Importantly, ETEC are disproportionately represented in cases of severe diarrheal illness as well as in deaths due to diarrhea among young children in developing countries [1].

[0070] These pathogens cause diarrhea by the elaboration and effective delivery of heat-labile and/or heat-stable enterotoxins to intestinal epithelial cells where they stimulate production of cyclic nucleotides ultimately activating the cystic fibrosis transmembrane regulator (CFTR) with resulting net efflux of fluid into the intestinal lumen [2]. Plasmid-encoded colonization factors (CFs), discovered [3] shortly after these organisms were identified as a causative agent of cholera-like diarrheal illness [4–

6], are thought to be essential for effective colonization of the small intestine and required for ETEC pathogenesis.

[0071] Following early studies suggesting a pivotal role for these structures [7,8], CF antigens have defined the basis for most subsequent ETEC vaccine efforts [9,10]. However, one factor complicating development of a broadly protective vaccine for ETEC has been the general plasticity of *E. coli* genomes [11], and the significant antigenic heterogeneity of the CFs. To date, at least 26 antigenically distinct CF antigens have been described [12]. The lack of appreciable cross-protection afforded by these antigens combined with the complex landscape of CFs portrayed in ETEC molecular epidemiology studies continue to complicate rational CF antigen selection [13].

[0072] Antigenic heterogeneity, recent failure of LT-toxoid-based vaccine strategies [14,15], as well as the need to optimize the performance of live-attenuated vaccines currently in clinical trials [16–18] have highlighted the need to identify additional virulence molecules that might be targeted in ETEC vaccines.

Example 1. Conservation of ETEC pathogen-specific secreted antigens.

[0073] Two antigens, the EtpA adhesin, and the passenger domain of the EatA serine protease are encoded on the large 92 kilobase virulence plasmid of the prototypical ETEC strain H10407. Both of these secreted proteins [22,30] are required for H10407 to efficiently deliver heat-labile toxin to target epithelial cells. To assess their utility as potential vaccine antigens, we examined a large collection of ETEC strains that were well characterized with respect to associated clinical meta-data pertaining to disease severity and which had not undergone repeated serial passage in the laboratory.

[0074] Altogether, we found that these antigens are relatively conserved in the ETEC pathovar. Of the 181 strains examined in the present study (**FIG. 1**), we found that more than half of all strains produced EtpA (102/181, 56%) and/or EatA (106/181, 59%) (**Table 6, Table 7**), and that more than three quarters of all strains produced at least one of these antigens. Both EtpA and EatA were identified more than twice as

frequently as the most commonly identified CF (CS6), which was identified in 22% of strains in this collection (**Table 1**).

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

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

[1]expression as determined with corresponding monoclonal antibody
[2]refers to strains expressing proteins as determined by immunoblotting nd: not detected with any of the above monoclonal antibodies
[3](CFA/II pili are typically comprised of CS1 or CS2 in combination with CS3; CFA/IV pili may be formed by CS4 or CS5 in combination with CS6 or CS6 alone.)

Example 2. Relationship of plasmid-encoded virulence loci to colonization factor antigens.

[0075] Importantly, although the genes encoding the *etpBAC* secretion system [19] and the EatA auto-transporter [21] were initially discovered on the same large virulence plasmid of H10407, which also encodes the colonization factor (CF) CFA/I, we found that these loci were not restricted to strains expressing this particular CF, but were widely distributed among the different CFs, and were also present in strains for which no CF could be identified (**FIG. 2A**). Indeed, half of the strains for which no CF could be identified expressed either EtpA or EatA, suggesting that these antigens could complement existing vaccination strategies centered on CFs. As expected by the association with multiple CFs, we also found that EtpA and EatA were secreted by strains from multiple phylogenic lineages (**FIG. 2B, FIG. 2C**). Interestingly,

however we found a negative association between the etpBAC locus and strains expressing CFA/IV antigens [50,51] including CS5 in that none of the 23 strains possessing CS5 fimbriae secreted the EtpA adhesin. Similarly, among strains expressing CS6, which is frequently co-expressed with CS5, only a minority secreted EtpA.

[0076] As has been noted previously, ETEC strain H10407 causes more severe illness in human clinical challenge studies relative to other strains like B7A [53]. Because we had clinical metadata pertaining to disease severity for all of the strains in our collection, we questioned whether the production of either of EatA and/or EtpA was associated with strains isolated from more severe forms of infection. However, we did not find any clear association between either of these putative virulence loci and clinical outcome (**Table 6, Table 7**).

Example 3. Conservation of chromosomally-encoded antigens.

[0077] We also examined the conservation of two chromosomally-encoded antigens which are not specific to the ETEC pathovar. The *eaeH* gene was originally identified on the chromosome of ETEC strain H10407 by subtractive hybridization with *E. coli* MG1655 [54], is transcriptionally activated by cell contact [26], and under these conditions EaeH is produced by a diverse group of strains belonging to different phylogenies [28]. Using the EaeH peptide sequence from H10407 (GenBank accession AAZ57201), BLASTP searches of recently sequenced ETEC strains from Bangladesh and elsewhere (gscid.igs.umaryland.edu/wp.php?wp = comparative_genome_analysis_of_ enterotoxigenic_e._coli_isolates_from_infections_of_different_clinical_severity) also revealed that the *eaeH* gene was present in 63 out of 91 distinct isolates (69%) (**Table 9**). BLASTP searches of these data for another chromosomally encoded molecule, YghJ, a type II secretion system effector [55] recently shown to be involved in mucin degradation and toxin delivery [27] demonstrated that the *yghJ* gene was present on the chromosomes in 83 of 91 (91%) isolates. Similarly, we identified the YghJ protein in a majority (161/181, 89%) of ETEC culture supernatants (**Table 8**). This antigen was

produced across ETEC strains expressing multiple CF types including 31/36 strains that were CF-negative by monoclonal antibody screening.

Example 4. EatA and EtpA sequence conservation.

[0078] Ideally, putative vaccine targets should be specific to the pathovar under study or restricted to pathogenic isolates, but not subject to significant antigenic variation. Therefore to further examine the potential utility of two ETEC pathovar specific antigens, EtpA and EatA, as vaccine candidates, we used recently obtained DNA sequence information from multiple ETEC genomes belonging to different phylogenies and from temporally and geographically disparate sources to compare the predicted amino acid sequences of these proteins.

[0079] For the prototype EatA molecule, first described in ETEC H10407 [21], the 1042 residue region from amino acids 57–1098 is predicted for the secreted passenger domain that contains the serine protease catalytic triad [21] as well as protective epitopes [23]. We therefore compared this region of the molecule to those derived from the recently released genome sequences of multiple ETEC strains. Altogether, we found that the sequence of the EatA passenger domain ($EatA_p$) was very highly conserved across strains, and exhibited between 95–100% identity to the prototype H10407 $EatA_p$ (**Table 2**). Likewise, the predicted serine protease catalytic motif formed by the histidine, aspartic acid and serine residues at positions 134, 162, and 267, respectively were universally conserved within the passenger domains of these proteins (**FIG. 5**). Similarly, the predicted amino acid sequences of the secreted EtpA adhesin molecules from multiple strains exhibited between 94 and 100% identity to the H10407 prototype antigen (**Table 3**, **FIG. 6**).

[0080] Despite the fact that the comparator strains included here spanned isolates collected over nearly 40 years, belonging to different phylogenies and that strains originated in diverse locations in Asia, Africa and the Americas, both proteins appear to exhibit remarkably little antigenic variation. Likewise, in analysis of the genomes of strains isolated recently within Bangladesh both proteins demonstrated similar degrees of sequence conservation (**FIG. 5, FIG. 6**).

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

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

[1]based on BLAST-P searches against 1042 residues of predicted passenger domain of H10407.
[2]sequenced at GSCID (gscid.igs.umaryland.edu/wp.php?wp=comparative_genome_analysis_of_enterotoxigenic_e._coli_isolates_from_infections_of_different_clinical_severity)
[3]open reading frames corresponding to the *eatA* gene were first assembled from whole genome shotgun sequence contigs for these draft genomes; BLASTP for these homologues was conducted using CLC Main Workbench v6.9.1 and local database of predicted protein sequences derived from translation of assembled contigs from the respective sequencing projects.
[4](severity unknown)
[5]this study

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

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

tw10828	[3]AELC00000000.1	Guinea-Bissau	1996	diarrhea[4]	94	96	[49]
tw10598	[3]AELA00000000.1	Guinea-Bissau	1996	diarrhea[4]	95	96	[49]
E24377A	[3]NC_009786.1	Egypt	1980s	diarrhea[4]	95	96	[63]

[1]identity and similarity percentages reflect BLASTP comparisons of the first 600 residues of the EtpA sequence.
[2]sequenced at GSCID (gscid.igs.umaryland.edu/wp.php?wp=comparative_genome_analysis_of_enterotoxigenic_e._coli_isolates_from_infections_of_different_clinical_severity)
[3]open reading frames corresponding to the *etpA* gene were first assembled from whole genome shotgun sequence contigs for these draft genomes; BLASTP for these homologues was conducted using CLC Main Workbench v6.9.1 and local database of predicted protein sequences derived from translation of assembled contigs from the respective sequencing projects.
[4](severity unknown)
[5]this study

Example 5. Immunogenicity of novel virulence antigens.

[0081] Earlier immunoproteomic studies suggested that a variety of conserved *E. coli* proteins as well as ETEC pathovar specific proteins are recognized during the course of experimental infections in mice, and these responses parallel those observed using pooled convalescent sera from ETEC patients [25]. To further characterize the immune response to novel antigens, we focused on four, including two plasmid-encoded secreted ETEC pathovar-specific antigens: EatA protease, and the EtpA adhesin, as well as the highly conserved chromosomally-encoded YghJ metalloprotease and the EaeH adhesin protein.

[0082] In comparing convalescent plasma from patients hospitalized at icddr,b to uninfected controls from Bangladesh, we found that patients in general exhibited significantly greater total antibody (IgG, IgM, IgA) responses to each of these antigens following diarrheal illness (**FIG. 3**) suggesting that these proteins are expressed during the course of infection. Similar results were obtained in comparing plasma from un-infected children from a non-endemic area in the United States (**FIG. 7**).

[0083] We also examined the immune response to EtpA following infection by examining sera obtained before and after challenge of human volunteers with ETEC H10407. In sera obtained from two independent volunteer challenge studies, we also observed significant increases in immune responses to EtpA (**FIG. 8**), strongly suggesting that this secreted protein is specifically recognized following infection by ETEC strains including H10407 that secrete this antigen.

Example 6. Protective efficacy of combined EtpA-mutant EatA passenger vaccination.

[0084] The data above suggest that collectively these antigens might significantly extend coverage presently offered by classical approaches to ETEC vaccine development. We therefore questioned whether these two antigens could be successfully combined in a subunit approach. Because we have previously demonstrated that the native secreted EatA passenger domain will degrade intestinal mucin [29] as well as the EtpA adhesin molecule [22], we elected to vaccinate animals with a modified recombinant version of the EatA passenger that lacks protease activity (rEatA$_p$H134R).

[0085] Co-vaccination with rEtpA and the mutant rEatA$_p$H134R molecule elicited robust serologic responses to both molecules. As anticipated, each of the groups mounted strong serologic responses to the LT adjuvant (**FIG. 4A**), and both antigens retained their immunogenicity following co-immunization of EtpA with the rEatAH134R passenger domain (**FIG. 4B, FIG. 4C**) with responses that were at least comparable to those obtained following immunization with either antigen alone (**FIG. 4C**). Likewise, mice immunized with both antigens were significantly protected against colonization by ETEC (**FIG. 4D**). Collectively, these data suggest that co-immunization with these two antigens is feasible, and could be employed to expand present approaches to ETEC vaccine antigen selection.

Discussion for the Examples.

[0086] Enterotoxigenic *Escherichia coli* remain one of the most common causes of infectious diarrhea worldwide, and severe disease caused by these pathogens persists as leading cause of death among young children in developing countries [1]. Despite recognition of these toxin producing *E. coli* as a cause of severe cholera-like diarrheal illness more than forty years ago [57], there remains no effective broadly protective vaccine for ETEC.

[0087] Most vaccinology efforts to date have focused almost exclusively on a subset of plasmid-encoded antigens, namely the colonization factors (CFs) and heat-labile toxin [9]. Vaccines based on this strategy have faced several impediments.

First, the CFs are quite diverse with more than 26 distinct antigens described to date. In addition, a number of recent vaccine studies have suggested that simply engendering immune responses to CFs and/or heat-labile toxin may not be sufficient to provide sustained broad-based protection [14–16].

[0088] A major challenge to ETEC vaccine development in general is that the most highly conserved antigens of ETEC, typically encoded on core regions of the chromosome, are also shared with commensal *E. coli* [60]. Included among these chromosomally encoded conserved proteins are two antigens studied here, YghJ [27] and EaeH [28] that were recently shown to be important for ETEC virulence. While the present studies also demonstrate that these proteins are recognized during the course of ETEC infection, the degree to which these antigens can be safely targeted in vaccines without inadvertent disruption of the intestinal microflora remains to be studied.

[0089] The inherent plasticity of *E. coli* genomes contributes substantially to the difficulty in defining antigens unique to the ETEC pathovar that are widely conserved. No single antigen exclusive to these pathogens, but universally conserved in this pathovar, has been described to date. Some have suggested that this might be predicted based on the fact that the plasmid-encoded heat-labile and/or heat-stable toxins, which define the ETEC pathovar, could form a minimal complement of virulence genes in wide variety of *E. coli* host strains [61].

[0090] In this context, we examined the gene conservation and the actual production of these proteins in a large collection of well-characterized strains from Bangladesh, complemented by strains from other locations that were associated with severe disease and for which there were available clinical metadata. Notably, two plasmid-encoded ETEC pathotype-specific antigens, the EatA serine protease and the secreted EtpA adhesin molecule were shared broadly among strains belonging to different CF groups with the exception of strains that produced CFA/IV antigens CS4, CS5, CS6 which only infrequently produced EtpA.

[0091] In general, we found high degrees of concordance between the presence of these genes by PCR and production of the corresponding protein. The prevalence of EtpA and EatA was 56 and 59%, respectively, as determined by

examination of protein expression. Importantly, the strains that produced these antigens belonged to many different phylogenies suggesting that genes encoding these antigens have been widely dispersed.

[0092] The analyses of strains in this study largely focused on isolates from Bangladesh. However, these data are potentially relevant for vaccine development for a number of reasons. First, Bangladesh is highly endemic for enterotoxigenic *E. coli* infections, and consequently remains an important site for vaccine field trials. In addition, ETEC has been under study in this region since the discovery of this pathotype, permitting us to compare sequence variation in candidate antigens over four decades. Understanding both current prevalence and sequence conservation of potential novel vaccine antigens in this population over time will be particularly important for making rational decisions about their inclusion in future iterations of ETEC vaccines. Finally, the geographic and temporal dispersal of genes encoding EtpA and EatA in multiple phylogenetic backgrounds, further attests to importance of studying these molecules as vaccine targets.

[0093] The data presented here suggest that the novel pathovar-specific antigens could complement existing strategies for ETEC vaccine development by broadening the antigenic valency.

Methods for the Examples.

[0094] *Bacterial strains and growth conditions:* ETEC strains used in this study are detailed in **Table 6, Table 8, Table 9**. All strains were grown at 37° in Casamino acids yeast extract media [32] (CAYE: 2.0% Casamino Acids, 0.15% yeast extract, 0.25% NaCl, 0.871% K_2HPO_4, 0.25% glucose, and 0.1% (v/v) trace salts solution consisting of 5% $MgSO_4$, 0.5% $MnCl_2$, 0.5% $FeCl_3$) from frozen glycerol stocks maintained at −80°C.

[0095] *Strain characterization by disease severity and colonization factor type:* Strains from the International Centre for Diarrhoeal Disease Research (icddr,b) in Dhaka were selected based on their associated disease severity using modified WHO guidelines as previously outlined [33]. Expression of individual CFs was determined by dot immunoblotting with monoclonal antibodies specific to each respective CFs (CF-

MAb) as previously described [34]. Briefly, 2 µl of a PBS suspension containing 106 colony forming units of each ETEC strain was dotted onto nitrocellulose, air-dried, blocked with BSA in PBS, followed by detection with CF-MAbs and goat anti-mouse IgG_HRP conjugate. Bound MAbs were then detected with 4-chloro-1-naphthol chromogen and H_2O_2.

[0096] *Screening for ETEC virulence genes by PCR*: We screened a total of 181 ETEC available isolates currently maintained as frozen glycerol stocks in our laboratories. The majority of these strains were collected between 1998 and 2011 in Bangladesh, and were obtained from the icddr,b in Dhaka. Complementing this collection were geographically disparate strains associated with severe diarrheal illness including strains from the Amazon region in Brazil [35], and ThroopD, an isolate from a patient with severe ETEC diarrheal illness who presented in Dallas in the 1970s [36]. Strains encoding *eatA* and *etpA* were identified by PCR using primers directed against conserved regions of these genes as previous described [37]. Briefly, a small amount of frozen glycerol stock from each strain was introduced with a sterile pipette tip into a PCR mixture containing the respective primers and a master mix. Toxin genotypes were confirmed in these isolates using multiplex PCR screening for genes encoding heat-labile (LT), and heat-stable toxins (STp, and STh) as previously described [34]. Primer sequences are listed in **Table 4**.

Table 4. Oligonucleotide primers used in this study.

SEQ ID NO:	Gene	Primer designation	Sequence (5'>3')	Product size (bp)	Ref
6	*eatA*	jf082213.1-F	ATGTGCTTTGGCAGGTTAAT	1943	[37]
7		jf082213.2-R	ATATCCAGTCAGCACCCACT		
8	*etpA*	jf082213.3-F	GGTTCAGGCAGTATCCAGAC	999	[37]
9		jf082213.4-R	GGTGTAGCTGTCTGACCACA		
10	*eltB*	jf092313.3-F	ACGGCGTTACTATCCTCTC	273	[34]
11		jf092313.4-R	TGGTCTCGGTCAGATATGTG		
12	*est1A*	jf092313.5-F	TCTTTCCCCTCTTTTAGTCAG	166	[34]
13		jf092313.6-R	ACAGGCAGGATTACAACAAAG		

14 15	*est1B*	jf092313.8-F jf092313.7-R	AGTGGTCCTGAAAGCATG TACAAGCAGGATTACAACAC	64	[34]

[0097] *Immunoblotting for secreted ETEC virulence antigens*: To determine production of secreted virulence antigens by different ETEC strains, supernatants from overnight cultures were first precipitated with trichloroacetic acid (TCA) [19] and resuspended in sample buffer before polyacrylamide gel electrophoresis. Western blotting was then performed using polyclonal rabbit antisera against recombinant versions of either EatA [21], EtpA [19], or YghJ [27] that were pre-absorbed against an *E. coli* lysate column (Pierce) and affinity-purified using the antigen immobilized on nitrocellulose membranes as previously described [31,38], followed by detection with affinity-purified secondary goat anti-rabbit-(IgG)-HRP conjugate (Santa Cruz Biotechnology, SC2004).

[0098] *Protein sequence comparisons of ETEC pathovar specific antigens*: To examine antigenic conservation of EatA among ETEC isolates for which genomic DNA sequences are currently available, BLASTP [39] was used to search GenBank www.ncbi. nlm.nih.gov/genbank/ using the full length sequence of the EatA protein from strain H10407 (www.ncbi.nlm.nih.gov/protein/AAO17297.1) as the query sequence. To construct alignments of EatA from positive strains, the 1042 residue passenger domain (corresponding to amino acids 57–1098 of EatA from H10407) was compared with EatA of ETEC isolates derived from different phylogenic lineages using a CLUSTAL Omega (release 1.2.0 AndreaGia-como) [40] algorithm plugin for CLC Main Workbench v6.9.1. A similar approach was used to compare the amino-terminal sequence of EtpA (amino acids 1–600, GenBank accession number AAX13509.2).

[0099] *Conservation heat mapping*: Virulence protein expression data from the collection of 181 strains under study were included in the analysis. Heat maps were configured using R [41] version 3.1.0 (2014, www.R-project.org/) using gplots [42] and RColorBrewer [43] packages installed from CRAN.R-project.org using the heatmap2 function within gplots (see **Table 5**).

Table 5. Numerical data corresponding to the heatmap depicted in **FIG. 1A**.

#conservation	CFA/I	cs1	cs2	cs3	cs4	cs5	cs6	cs7	cs8	cs14	cs17	cs21	nd
EtpA	0.81	0.86	0.8	0.84	0.14	0	0.08	0.88	0	0.86	0.9	0.82	0.44
EatA	0.81	0.86	0.7	0.76	0.43	0.78	0.64	0.94	1	0.14	0.9	0.82	0.19
either	0.85	0.93	0.9	0.92	0.92	0.78	0.67	0.94	1	0.93	0.9	0.94	0.5

[0100] *Recombinant protein production*: The antigens used in these studies were produced as polyhistidine-tagged recombinant proteins and purified by immobilized metal ion affinity chromatography (IMAC) as previously described [27,29,44,45]. Additional polishing steps including size exclusion or ion exchange chromatography were performed as needed to produce highly purified antigens. Purity of each antigen was assessed by SDS-PAGE followed by sensitive Coomassie Blue staining. Purified recombinant antigens were stored at −80°C.

[0101] *Assessment of immune responses to novel ETEC virulence proteins*: To quantify antibody concentrations directed at novel recombinant antigens, kinetic ELISA was performed on dilutions of plasma samples previously obtained from patients hospitalized at the International Centre for Diarrhoeal Disease Research in Dhaka, Bangladesh (icddr,b) with acute symptomatic ETEC infections. Plasma samples from non-infected adults and children obtained at icddr,b, or specimens obtained from children at Saint Louis Children's Hospital were used as negative controls. Samples from human volunteer ETEC H10407 challenge studies were kindly provided by Dr. Robert Gormely and Dr. Stephen Savarino of National Naval Medical Center, Bethesda Maryland.

[0102] Use of these clinical materials was approved by the Institutional Review Boards of both icddr,b and Washington University School of Medicine. All plasma samples were maintained at 4°C in a humidified chamber prior to use in ELISA. Immune responses to purified recombinant proteins (rYghJ, rEaeH, rEtpA, rEatA$_p$) were assessed by kinetic ELISA [46] as previously described [30,47]. Antigen binding to ELISA wells (Corning, Costar 2580) was first optimized to determine the optimal coating concentration and buffer system, using highly antigen-specific polyclonal rabbit antisera to detect binding by ELISA. Purified antigens were then diluted either in 50 mM

carbonate buffer (pH 9.6) (rEtpA-myc-His_6, 1 µg/ml; $rEatA_p$, 10 µg/ml; rYghJ-myc-His_6, 1 µg/ml); or in phosphate buffered saline (PBS, pH 7.4) (rEaeH-myc-His_6, 1 µg/ml). ELISA plate wells were coated with 100 µl/well overnight at 4°C, washed with PBS containing 0.05% Tween-20 (PBS-T), and blocked for 1 h at 37°C with 1% BSA in PBS-T. All plasma samples were diluted at 1:4096 in blocking buffer. After incubation for 1 hour at 37°C, plates were washed with PBS-T, and secondary goat anti-human IgG(H+L)-HRP conjugated antibody (Pierce, 31410) was added at a final concentration of 1:10,000. After incubation for 30 minutes at 37°C, plates were washed and developed with TMB microwell peroxidase substrate [3,3',5,5'-Tetramethylbenzidine] (KPL, 50-76-00). Kinetic absorbance measurements were determined at a wavelength of 650 nm, and acquired at 40 s intervals for 20 minutes using a microplate spectrophotometer (Eon, BioTek). All data were recorded and analyzed using Gen5 software (BioTek) and reported as the Vmax expressed as milliunits/min. Statistical calculations were performed using Prism v4.0c (GraphPad Software), using nonparametric Mann-Whitney (two-tailed) comparisons of data.

[0103] *Mouse immunization and challenge studies*: These studies were performed in strict accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health, using an established protocol approved by the Washington University School of Medicine Animal Studies Committee.

[0104] Four groups of twelve CD-1 mice were immunized intranasally with either 1 µg of LT (adjuvant only controls), or 1 µg of LT + 15 µg of $rEatA_p$(H134R), or 1 µg of LT + 15 µg of rEtpA, or 1 µg of LT + 15 µg of rEatA(H134R)+15 µg of rEtpA on days 0, 14, 28. On day 40, mice were treated with streptomycin [5 g per liter] in drinking water for 24 hours, followed by drinking water alone for 18 hours. After administration of famotidine to reduce gastric acidity, mice were challenged with 10^6 cfu of the kanamycin-resistant (*lacZYA*::KmR) strain jf876 [48] by oral gavage as previously described [47]. Fecal samples (6 pellets/mouse) were collected on day 42 before oral gavage, re-suspended in buffer (10mM Tris, 100mM NaCl, 0.05% Tween 20, 5mM Sodium Azide, pH 7.4) overnight at 4°C, centrifuged to pellet insoluble material, and

recover supernatant for fecal antibody testing (below). Twenty-four hours after infection, mice were sacrificed, sera were collected, and dilutions of saponin small-intestinal lysates were plated onto Luria agar plates containing kanamycin (50 μg/ml).

[0105] Murine immune responses to LT, EatA and EtpA were determined using previously described kinetic ELISA. Briefly, ELISA wells were coated with 1 μg/ml GM1, or 10 μg/ml of rEatA$_p$(H134R), or 1 μg/ml rEtpA in carbonate buffer (15 mM Na_2CO_3, 35 mM $NaHCO_3$, 0.2 g/L NaN_3, pH8.6) overnight at 4°C. Wells were washed three times with phosphate-buffered saline containing 0.05% Tween 20 (PBS-T), blocked with 1% bovine serum albumin (BSA) in PBS-T for 1 h at 37°C, and 100 μl of fecal suspensions (undiluted) or sera (diluted 1:100 in PBS-T with 1% BSA) was added per ELISA well and incubated at 37°C for 1 h. Horseradish peroxidase-conjugated secondary antibodies were used and signal detected with TMB (3,3',5,5'-tetramethylbenzidine)-peroxidase substrate (KPL) substrate.

[0106] *Ethics statement*: All animal studies were performed under protocols approved by the Animal Studies Committee of Washington University School of Medicine (protocol number 20110246A1). All procedures complied with Public Health Service guidelines, and The Guide for the Care and Use of Laboratory Animals.

[0107] All human studies included were performed under a protocol approved by the Institutional Review Board of Washington University School of Medicine (IRB ID# 201110126). All of the human studies here report anonymous analysis of de-identified pre-existing sera previously stored from earlier studies for which no additional consent was obtained.

Table 6. EtpA, EatA, CF, LT, and ST prevalence.

Stock #(jf)	severity	Strain designation	origin	isolation year	EatA	EtpA	CFs	LT	ST	cfa/1	cs1	cs2	cs3	cs4	cs5	cs6	cs7	cs8	cs14	cs17	cs21	CFOve	EtpA/EatA
350	2	ETEC Throop D "1"	USA, Texas	1975	1	1		1	1														2
1388	2	ETEC Envira 8/11	Amazon, Brazil	1998	1	0	1	1	1						1	1							1
1389	2	ETEC Envira 10/1	Amazon, Brazil	1998	1	0	1	1	1						1	1							1
1390	2	ETEC Jurua 18/11	Amazon, Brazil	1998	1	1	1	1	0								1						2
1391	2	ETEC Jurua	Amazon,	1998	1	1	1	1	0								1						2

		20/10	Brazil																			
2007	2	P02997067.1	Bangladesh	2011	1	0		1	1													1
2021	2	MP021552.1	Bangladesh	2011	1	0		1	1													1
2036	2	MP021561.1	Bangladesh	2011	1	0		0	1													1
2053	2	MP021017.1	Bangladesh	2011	0	0		0	1													0
2063	2	MP020980.1	Bangladesh	2011	1	0		0	1													1
2097	1	2720900	Bangladesh	2007	1	1	1	0	1	1												2
2098	1	2719100	Bangladesh	2007	1	1	1	0	1	1										1		2
2099	1	2722950	Bangladesh	2007	1	1	1	1	1	1												2
2100	1	2726800	Bangladesh	2007	0	1	1	1	1									1				1
2101	1	2726950	Bangladesh	2007	1	1	1	0	1											1		2
2102	1	2729250	Bangladesh	2007	1	0	1	1	1	1												1
2103	2	2747800	Bangladesh	2007	1	0	1	1	1					1	1							1
2104	2	2741950	Bangladesh	2007	1	1	1	1	1		1		1							1		2
2105	2	2735000	Bangladesh	2007	1	1	1	1	0								1					2
2106	2	2730350	Bangladesh	2007	0	1	1	1	1	1										1		1
2107	2	2730450	Bangladesh	2007	1	1	1	1	1		1		1							1		2
2108	2	2770900	Bangladesh	2007	1	0	1	1	1					1	1							1
2109	2	2780750	Bangladesh	2007	0	0	0	0	1												1	0
2110	2	2731150	Bangladesh	2007	1	0	1	1	0							1		1				1
2111	2	2785200	Bangladesh	2007	0	0	0	1	0												1	0
2112	2	2733950	Bangladesh	2007	1	0	1	1	1					1		1						1
2113	2	2749250	Bangladesh	2007	1	0	1	1	1						1	1						1
2114	2	2756500	Bangladesh	2007	1	1	1	1	0								1					2
2115	2	2788150	Bangladesh	2007	0	0	0	1	1												1	0
2116	2	2762100	Bangladesh	2007	1	0	1	1	0							1		1				1
2117	1	2845350	Bangladesh	2008	0	0	1	1	1					1		1						0
2118	1	2845650	Bangladesh	2008	0	0	1	1	1					1		1						0
2119	1	2846750	Bangladesh	2008	1	1	1	1	1			1	1									2
2120	1	2848050	Bangladesh	2008	0	0	1	1	1						1	1						0
2121	2	2850400	Bangladesh	2008	1	1	1	1	0								1					2
2122	2	2850750	Bangladesh	2008	1	1	1	1	0										1			2
2123	1	2851500	Bangladesh	2008	1	1	1	1	0								1					2

2124	2	2853500	Bangladesh	2008	1	1	0	1	0													1	2
2125	2	2854350	Bangladesh	2008	0	1	1	1	0										1				1
2126	2	2860050	Bangladesh	2008	1	0	1	1	1						1	1							1
2127	2	2860650	Bangladesh	2008	1	0	1	1	1						1	1							1
2128	2	2861200	Bangladesh	2008	0	0	1	0	1							1							0
2129	2	2862600	Bangladesh	2008	0	0	1	0	1					1		1							0
2130	2	2863500	Bangladesh	2008	0	0	1	0	1	1													0
2131	2	2864350	Bangladesh	2008	0	0	0	1	1													1	0
2132	2	2865200	Bangladesh	2008	0	0	0	0	1													1	0
2133	2	2866350	Bangladesh	2008	1	0	1	1	1						1	1							1
2134	2	2866450	Bangladesh	2008	1	1	1	1	0								1						2
2135	2	2866550	Bangladesh	2008	1	1	1	1	0								1						2
2136	2	2866750	Bangladesh	2008	1	1	1	1	0								1						2
2137	2	2867750	Bangladesh	2008	1	1	1	1	0								1						2
2138	2	2871950	Bangladesh	2008	1	1	1	1	0			1	1										2
2139	2	2872000	Bangladesh	2008	1	1	1	1	0			1	1										2
2140	2	2872800	Bangladesh	2008	1	1	1	1	0								1						2
2141	2	2875000	Bangladesh	2008	1	1	1	1	0								1						2
2142	2	2875150	Bangladesh	2008	0	0	1	1	1					1		1							0
2143	2	199900.1	Bangladesh	2011	1	1	1	1	0											1			2
2144	2	199900.2	Bangladesh	2011	1	1	1	1	0											1			2
2145	2	21600.1	Bangladesh	2011	0	1	1	1	1										1				1
2146	2	21600.2	Bangladesh	2011	0	1	1	1	1										1				1
2147	2	21600.3	Bangladesh	2011	0	1	1	1	1										1				1
2148	2	21600.4	Bangladesh	2011	0	1	1	1	1										1				1
2149	0	BCE006, MS023	Bangladesh	2004	1	1	1	1	1		1		1										2
2150	0	BCE011, MS001	Bangladesh	2002	0	0	1	0	1	1													0
2151	0	BCE019, MS013	Bangladesh	2003	1	1	1	1	1		1		1										2
2152	0	BCE030, MS009	Bangladesh	2003	1	1	1	0	1	1													2
2153	0	BCE032, MS012	Bangladesh	2003	1	1	1	1	1		1		1										2
2154	0	BCE034, MS014	Bangladesh	2003	1	1	1	0	1	1													2
2155	2	P0304816.1	Bangladesh	2011	1	0		0	1														1
2170	2	P0305260.1	Bangladesh	2011	0	1		0	1														1

2184	2	p03052 93.1	Bangla desh	2011	1	0		1	1														1
2199	2	p03018 67.1	Bangla desh	2011	1	1		0	1														2
2208	2	p03022 93.2	Bangla desh	2011	0	1		1	1														1
2216	2	p03023 08.1	Bangla desh	2011	0	1		1	1														1
2226	2	p03047 77.1	Bangla desh	2011	1	1		1	0														2
2267	2	MP0209 40.1	Bangla desh	2011	0	0		0	1														0
2282	2	P02989 42.1	Bangla desh	2011	1	0		1	1														1
2295	2	p02994 38.2	Bangla desh	2011	0	1		0	1														1
2305	2	p02999 17.1	Bangla desh	2011	0	0		1	1														0
2315	2	ETP050 002	Bangla desh	2005	1	1	1	0	1	1													2
2316	2	ETP050 003	Bangla desh	2005	0	0	0	0	1													1	0
2317	2	ETP050 007	Bangla desh	2005	0	0	1	1	1						1	1							0
2318	2	ETP050 008	Bangla desh	2005	0	1	0	0	1													1	1
2319	2	ETP050 009	Bangla desh	2005	0	0	1	1	1						1	1							0
2320	2	ETP050 010	Bangla desh	2005	0	1	1	1	1				1								1		1
2321	2	ETP050 011	Bangla desh	2005	0	1	1	1	1			1	1										1
2322	2	ETP050 012	Bangla desh	2005	0	0	1	0	1							1							0
2323	2	ETP050 015	Bangla desh	2005	0	0	1	1	1						1	1							0
2324	2	ETP050 016	Bangla desh	2005	0	1	0	0	1													1	1
2325	2	ETP050 017	Bangla desh	2005	1	0	1	1	1						1	1							1
2326	2	ETP050 019	Bangla desh	2005	0	1	0	1	1													1	1
2327	2	ETP050 20	Bangla desh	2005	0	0	1	1	0								1						0
2328	2	ETP050 26	Bangla desh	2005	1	1	1	1	1	1													2
2329	2	ETP050 035	Bangla desh	2005	0	0	0	1	1													1	0
2330	2	ETP050 038	Bangla desh	2005	0	0	1	1	1						1	1							0
2331	2	ETP050 039	Bangla desh	2005	0	1	1	1	1										1				1
2332	2	ETP050 044	Bangla desh	2005	0	0	1	0	1	1													0
2333	2	ETP050 046	Bangla desh	2005	1	1	1	0	1	1											1		2
2334	2	ETP050 047	Bangla desh	2005	0	0	0	1	1													1	0
2335	2	ETP050 050	Bangla desh	2005	0	0	1	1	0											1			0
2336	2	174750	Bangla desh	2010	1	0	1	1	1		1		1								1		1
2337	2	174900	Bangla desh	2010	0	0	0	0	1													1	0
2338	2	178200	Bangla desh	2010	0	0	0	0	1													1	0

2339	2	178850	Bangladesh	2010	1	0	0	1	1													1	1
2340	2	178900	Bangladesh	2010	1	1	1	0	1	1												1		2
2341	2	179100	Bangladesh	2010	0	0	0	1	1														1	0
2342	2	179550	Bangladesh	2010	1	1	1	1	0								1							2
2343	2	180200	Bangladesh	2010	1	1	1	1	1								1							2
2344	2	180600	Bangladesh	2010	1	1	1	1	1											1				2
2345	2	180050	Bangladesh	2010	1	1	1	1	0											1				2
2346	2	C034666	Bangladesh	2010	1	0	0	1	1														1	1
2347	2	C035134	Bangladesh	2010	1	1	1	1	1											1				2
2348	2	C035209	Bangladesh	2010	0	0	1	0	1			1	1											0
2349	2	C035213	Bangladesh	2010	1	1	1	1	1											1				2
2350	2	C035605	Bangladesh	2010	1	1	1	1	0											1				2
2351	2	C035662	Bangladesh	2010	0	0	0	1	1														1	0
2352	2	C035776	Bangladesh	2010	1	0	1	0	1					1		1								1
2353	2	C035959	Bangladesh	2010	0	1	1	1	1										1					1
2354	2	C036255	Bangladesh	2010	0	1	1	1	1										1					1
2355	2	173150	Bangladesh	2010	0	0	1	1	1		1		1									1		0
2356	0	BCE001, MS16	Bangladesh	2002	0	1	0	1	0														1	1
2357	0	BCE002, MS12	Bangladesh	2002	0	1	1	1	1		1		1											1
2358	0	BCE005, MS23	Bangladesh	2002	0	1	0	1	1														1	1
2359	0	BCE007, MS11	Bangladesh	2002	1	1	1	1	1		1		1									1		2
2360	0	BCE008, MS1	Bangladesh	2002	0	0	0	1	0														1	0
2361	0	BCE008, MS13	Bangladesh	2002	0	0	1	0	1							1								0
2362	0	BCE019, MS16	Bangladesh	2002	1	1	1	1	1			1	1											2
2363	0	BCE035, MS8	Bangladesh	2002	1	0	1	1	0								1							1
2364	0	BCE039, MS13	Bangladesh	2002	0	0	0	1	0														1	0
2365	0	BCE041, MS11	Bangladesh	2002	0	0	0	1	0														1	0
2366	0	BCE046, MS16	Bangladesh	2002	0	0	0	1	0														1	0
2367	0	BCE049, MS9	Bangladesh	2002	1	0	1	1	1			1	1									1		1
2368	0	BCE054, MS24	Bangladesh	2002	0	0	1	1	1										1					0
2369	0	BCE058, MS13	Bangladesh	2002	1	1	1	0	1	1														2
2370	0	BCE062, MS24	Bangladesh	2002	1	1	1	1	1			1	1											2
2371	0	BCE063, MS14	Bangladesh	2002	0	1	0	1	0														1	1

2372	0	BCE068, MS10	Bangladesh	2002	0	1	1	1	1			1	1										1
2373	0	BCE068, MS23	Bangladesh	2002	0	1	0	0	1													1	1
2374	0	BCE069, MS15	Bangladesh	2002	1	1	1	1	1	1													2
2375	0	BCE069, MS9	Bangladesh	2002	1	1	1	1	1		1		1								1		2
2381	2	mp021566.1	Bangladesh	2011	1	1		1	1														2
2409	2	ETP98109	Bangladesh	1998	1	1	1	0	1	1													2
2410	2	ETP98111	Bangladesh	1998	1	1	1	1	1					1		1							2
2411	2	ETP98112	Bangladesh	1998	1	1	1	1	1		1		1								1		2
2412	2	ETP98114	Bangladesh	1998	0	1	1	0	1										1				1
2413	2	ETP98115	Bangladesh	1998	0	1	1	1	1							1							1
2414	2	ETP98068	Bangladesh	1998	1	1	1	1	1	1													2
2415	2	ETP98073	Bangladesh	1998	1	0	1	1	1						1	1							1
2416	2	ETP98097	Bangladesh	1998	1	1	1	1	1										1				2
2417	2	ETP98103	Bangladesh	1998	1	1	1	1	1		1		1								1		2
2418	2	ETP98105	Bangladesh	1998	1	1	1	1	1		1		1								1		2
2419	2	ETP98053	Bangladesh	1998	1	0	1	1	1						1	1							1
2420	2	ETP98056	Bangladesh	1998	1	0	1	0	1							1							1
2421	2	ETP98061	Bangladesh	1998	1	1	1	0	1							1							2
2422	2	ETP98062	Bangladesh	1998	1	0	1	0	1										1				1
2423	2	ETP98066	Bangladesh	1998	0	0	1	1	1							1							0
2424	2	ETP98004	Bangladesh	1998	0	0	0	0	1													1	0
2425	2	ETP98015	Bangladesh	1998	1	0	1	0	1						1	1							1
2426	2	ETP98028	Bangladesh	1998	1	0	1	0	1						1	1							1
2427	2	ETP98038	Bangladesh	1998	1	1	1	0	1	1													2
2428	2	ETP98042	Bangladesh	1998	1	0	1	1	1						1	1							1
2430	1	BCE003, DS5	Bangladesh	2002	1	1	0	1	1													1	2
2431	1	BCE021, DS7	Bangladesh	2002	0	1	0	0	1													1	1
2432	1	BCE046, DS7	Bangladesh	2002	1	0	1	1	1						1	1							1
2433	1	BCE062, DS2	Bangladesh	2002	1	1	0	1	1													1	2
2434	1	BCE007, DS5	Bangladesh	2002	1	0	1	1	1						1	1							1
2435	1	BCE022,DS6	Bangladesh	2002	1	1	1	0	1	1													2
2436	1	BCE049, DS3	Bangladesh	2002	1	1	0	1	1													1	2
2437	1	BCE063, DS4	Bangladesh	2002	0	1	1	1	1										1				1

2438	1	BCE011, DS3	Bangladesh	2002	0	1	0	1	0													1	1
2439	1	BCE035, DS6	Bangladesh	2002	1	1	1	1	1			1	1										2
2440	1	BCE054, DS4	Bangladesh	2002	1	0	1	1	1						1	1							1
2441	1	BCE066, DS5	Bangladesh	2002	0	1	0	1	0													1	1
2442	1	BCE013, DS1	Bangladesh	2002	1	1	1	0	1	1													2
2443	1	BCE039, DS2	Bangladesh	2002	1	0	1	1	1						1	1							1
2444	1	BCE055, DS1	Bangladesh	2002	1	1	1	1	0											1			2
2445	1	BCE069, DS2	Bangladesh	2002	1	1	1	1	1		1		1								1		2
2446	1	BCE018, DS6	Bangladesh	2002	1	1	0	1	1													1	2
2447	1	BCE046, DS2	Bangladesh	2002	0	1	0	1	0													1	1
2448	1	BCE061, DS1	Bangladesh	2002	0	0	0	1	1													1	0
2449	1	BCE129, DS2	Bangladesh	2002	0	1		0	1														1
totals					106	102	126	134	142	21	14	10	25	7	23	39	16	2	14	10	17	36	

column A: frozen stock number in the Fleckenstein laboratory (used as internal reference and for ordering strains from the lab)
column B: disease severity 0=asymptomatic; 1=mild disease; 2=severe, cholera-like illness
column C: strain designation as originally defined
column D: country of origin
column E: date of isolation (year)
column F: EatA passenger domain detection in culture supernatant by immunoblotting 0=no; 1=yes
column G: EtpA detection in culture supernatant by immunoblotting 0=no; 1=yes
column H: CF detected 0=no; 1=yes
column I: LT detected (PCR) 0=no; 1=yes
column J: ST detected (PCR) 0=no; 1=yes
columns K-V: presence of given individual CFs; 0=no; 1=yes

Table 7. Distribution of EtpA and EatA in CF+ strains.

CF	cfa/i	cs1	cs2	cs3	cs4	cs5	cs6	cs7	cs8	cs14	cs17	cs21	cf nd
EatA+	17	12	7	19	3	18	25	15	2	2	9	14	7
EtpA+	17	12	8	21	1	0	3	14	0	12	9	14	16
EatA+(%)	0.80952381	0.86	0.7	0.76	0.43	0.78	0.64	0.938	1	0.14286	0.9	0.82353	0.19444444
EtpA+(%)	0.80952381	0.86	0.8	0.84	0.14	0	0.08	0.875	0	0.85714	0.9	0.82353	0.44444444
either +	18	13	9	23	3	18	26	15	2	13	9	16	18
either + (%)	86	93	90	92	43	78	67	94	100	93	90	94	50

Table 8. YghJ prevalence.

Stock #	**YghJ**	**Strain designation**	EatA	EtpA	CFs	LT	ST	Severity	Symptom	CFA/I	CS1	CS2	CS3	CS4	CS5	CS6	CS7	CS8	CS14	CS17	CS21	CF0ve
350	1	ETEC ThroopD "1"	1	1		1	1	2	D													
1388	1	ETEC Envira 8/11	1	0	1	1	1	2	D						CS5	CS6						
1389	1	ETEC Envira 10/1	1	0	1	1	1	2	D						CS5	CS6						
1390	1	ETEC Jurua 18/11	1	1	1	1	0	2	D								CS7					

1391	1	ETEC Jurua 20/10	1	1	1	1	0	2	D								CS7					
2007	1	P02997067.1	1	0		1	1	2	D													
2021	0	MP021552.1	1	0		1	1	2	D													
2036	1	MP021561.1	1	0		0	1	2	D													
2053	0	MP021017.1	0	0		0	1	2	D													
2063	1	MP020980.1	1	0		0	1	2	D													
2097	1	2720900	1	1	1	0	1	1	D	CFA/I												
2098	1	2719100	1	1	1	0	1	1	D	CFA/I											CS21	
2099	0	2722950	1	1	1	1	1	1	D	CFA/I												
2100	1	2726800	0	1	1	1	1	1	D										CS14			
2101	0	2726950	1	1	1	0	1	1	D												CS21	
2102	1	2729250	1	0	1	1	1	1	D	CFA/I												
2103	1	2747800	1	0	1	1	1	2	D						CS5	CS6						
2104	1	2741950	1	1	1	1	1	2	D		CS1		CS3								CS21	
2105	1	2735000	1	1	1	1	0	2	D								CS7					
2106	1	2730350	0	1	1	1	1	2	D	CFA/I											CS21	
2107	1	2730450	1	1	1	1	1	2	D		CS1		CS3								CS21	
2108	1	2770900	1	0	1	1	1	2	D						CS5	CS6						
2109	1	2780750	0	0	0	0	1	2	D													1
2110	1	2731150	1	0	1	1	0	2	D							CS6		CS8				
2111	1	2785200	0	0	0	1	0	2	D													1
2112	1	2733950	1	0	1	1	1	2	D					CS4		CS6						
2113	1	2749250	1	0	1	1	1	2	D						CS5	CS6						
2114	1	2756500	1	1	1	1	0	2	D								CS7					
2115	1	2788150	0	0	0	1	1	2	D													1
2116	1	2762100	1	0	1	1	0	2	D							CS6		CS8				
2117	1	2845350	0	0	1	1	1	1	D					CS4		CS6						
2118	1	2845650	0	0	1	1	1	1	D					CS4		CS6						
2119	1	2846750	1	1	1	1	1	1	D			CS2	CS3									
2120	0	2848050	0	0	1	1	1	1	D						CS5	CS6						
2121	1	2850400	1	1	1	1	0	2	D								CS7					
2122	1	2850750	1	1	1	1	0	2	D											CS17		
2123	1	2851500	1	1	1	1	0	1	D								CS7					
2124	1	2853500	1	1	0	1	0	2	D													1
2125	1	2854350	0	1	1	1	0	2	D										CS14			
2126	1	2860050	1	0	1	1	1	2	D						CS5	CS6						
2127	1	2860650	1	0	1	1	1	2	D						CS5	CS6						
2128	0	2861200	0	0	1	0	1	2	D							CS6						
2129	1	2862600	0	0	1	0	1	2	D					CS4		CS6						
2130	1	2863500	0	0	1	0	1	2	D	CFA												

										/I											
2131	1	2864350	0	0	0	1	1	2	D												1
2132	1	2865200	0	0	0	0	1	2	D												1
2133	1	2866350	1	0	1	1	1	2	D						CS5	CS6					
2134	1	2866450	1	1	1	1	0	2	D								CS7				
2135	1	2866550	1	1	1	1	0	2	D								CS7				
2136	1	2866750	1	1	1	1	0	2	D								CS7				
2137	1	2867750	1	1	1	1	0	2	D								CS7				
2138	1	2871950	1	1	1	1	0	2	D			CS2	CS3								
2139	1	2872000	1	1	1	1	0	2	D			CS2	CS3								
2140	1	2872800	1	1	1	1	0	2	D								CS7				
2141	1	2875000	1	1	1	1	0	2	D								CS7				
2142	1	2875150	0	0	1	1	1	2	D					CS4		CS6					
2143	1	199900.1	1	1	1	1	0	2	D											CS17	
2144	1	199900.2	1	1	1	1	0	2	D											CS17	
2145	1	21600.1	0	1	1	1	1	2	D										CS14		
2146	1	21600.2	0	1	1	1	1	2	D										CS14		
2147	1	21600.3	0	1	1	1	1	2	D										CS14		
2148	1	21600.4	0	1	1	1	1	2	D										CS14		
2149	1	BCE006, MS023	1	1	1	1	1	0	A		CS1		CS3								
2150	0	BCE011, MS001	0	0	1	0	1	0	A	CFA/1											
2151	1	BCE019, MS013	1	1	1	1	1	0	A		CS1		CS3								
2152	0	BCE030, MS009	1	1	1	0	1	0	A	CFA/I											
2153	1	BCE032, MS012	1	1	1	1	1	0	A		CS1		CS3								
2154	1	BCE034, MS014	1	1	1	0	1	0	A	CFA/1											
2155	1	P0304816.1	1	0		0	1	2	D												
2170	1	P0305260.1	0	1		0	1	2	D												
2184	1	p0305293.1	1	0		1	1	2	D												
2199	1	p0301867.1	1	1		0	1	2	D												
2208	1	p0302293.2	0	1		1	1	2	D												
2216	1	p0302308.1	0	1		1	1	2	D												
2226	1	p0304777.1	1	1		1	0	2	D												
2267	1	MP020940.1	0	0		0	1	2	D												
2282	1	P0298942.1	1	0		1	1	2	D												
2295	1	p0299438.2	0	1		0	1	2	D												
2305	1	p0299917.1	0	0		1	1	2	D												
2315	1	ETP050002	1	1	1	0	1	2	D	CFA/I											
2316	1	ETP050003	0	0	0	0	1	2	D												1
2317	1	ETP050007	0	0	1	1	1	2	D						CS5	CS6					
2318	1	ETP050008	0	1	0	0	1	2	D												1
2319	1	ETP050009	0	0	1	1	1	2	D						CS5	CS6					

2320	1	ETP050010	0	1	1	1	1	2	D				CS3								CS21	
2321	1	ETP050011	0	1	1	1	1	2	D			CS2	CS3									
2322	1	ETP050012	0	0	1	0	1	2	D							CS6						
2323	1	ETP050015	0	0	1	1	1	2	D						CS5	CS6						
2324	1	ETP050016	0	1	0	0	1	2	D													1
2325	1	ETP050017	1	0	1	1	1	2	D						CS5	CS6						
2326	1	ETP050019	0	1	0	1	1	2	D													1
2327	1	ETP05020	0	0	1	1	0	2	D								CS7					
2328	1	ETP05026	1	1	1	1	1	2	D	CFA/I												
2329	0	ETP050035	0	0	0	1	1	2	D													1
2330	1	ETP050038	0	0	1	1	1	2	D						CS5	CS6						
2331	1	ETP050039	0	1	1	1	1	2	D										CS14			
2332	0	ETP050044	0	0	1	0	1	2	D	CFA/I												
2333	1	ETP050046	1	1	1	0	1	2	D	CFA/I											CS21	
2334	1	ETP050047	0	0	0	1	1	2	D													1
2335	1	ETP050050	0	0	1	1	0	2	D											CS17		
2336	1	174750	1	0	1	1	1	2	D		CS1		CS3								CS21	
2337	1	174900	0	0	0	0	1	2	D													1
2338	1	178200	0	0	0	0	1	2	D													1
2339	0	178850	1	0	0	1	1	2	D													1
2340	1	178900	1	1	1	0	1	2	D	CFA/I											CS21	
2341	1	179100	0	0	0	1	1	2	D													1
2342	1	179550	1	1	1	1	0	2	D								CS7					
2343	1	180200	1	1	1	1	1	2	D								CS7					
2344	1	180600	1	1	1	1	1	2	D											CS17		
2345	1	180050	1	1	1	1	0	2	D											CS17		
2346	0	C034666	1	0	0	1	1	2	D													1
2347	1	C035134	1	1	1	1	1	2	D											CS17		
2348	1	C035209	0	0	1	0	1	2	D			CS2	CS3									
2349	1	C035213	1	1	1	1	1	2	D											CS17		
2350	1	C035605	1	1	1	1	0	2	D											CS17		
2351	1	C035662	0	0	0	1	1	2	D													1
2352	1	C035776	1	0	1	0	1	2	D					CS4		CS6						
2353	1	C035959	0	1	1	1	1	2	D										CS14			
2354	1	C036255	0	1	1	1	1	2	D										CS14			
2355	1	173150	0	0	1	1	1	2	D		CS1		CS3								CS21	
2356	1	BCE001, MS16	0	1	0	1	0	0	A													1
2357	1	BCE002, MS12	0	1	1	1	1	0	A		CS1		CS3									
2358	:	BCE005, MS23	0	1	0	1	1	0	A													1

2359	1	BCE007, MS11	1	1	1	1	1	0	A		CS1		CS3								CS21	
2360	1	BCE008, MS1	0	0	0	1	0	0	A													1
2361	0	BCE008, MS13	0	0	1	0	1	0	A							CS6						
2362	1	BCE019, MS16	1	1	1	1	1	0	A			CS2	CS3									
2363	1	BCE035, MS8	1	0	1	1	0	0	A								CS7					
2364	1	BCE039, MS13	0	0	0	1	0	0	A													1
2365	0	BCE041, MS11	0	0	0	1	0	0	A													1
2366	1	BCE046, MS16	0	0	0	1	0	0	A													1
2367	1	BCE049, MS9	1	0	1	1	1	0	A			CS2	CS3								CS21	
2368	1	BCE054, MS24	0	0	1	1	1	0	A										CS14			
2369	0	BCE058, MS13	1	1	1	0	1	0	A	CFA/I												
2370	1	BCE062, MS24	1	1	1	1	1	0	A			CS2	CS3									
2371	1	BCE063, MS14	0	1	0	1	0	0	A													1
2372	1	BCE068, MS10	0	1	1	1	1	0	A			CS2	CS3									
2373	1	BCE068, MS23	0	1	0	0	1	0	A													1
2374	1	BCE069, MS15	1	1	1	1	1	0	A	CFA/I												
2375	1	BCE069, MS9	1	1	1	1	1	0	A		CS1		CS3								CS21	
2381	1	mp021566.1	1	1		1	1	2	D													
2409	1	ETP98109	1	1	1	0	1	2	D	CFA/I												
2410	1	ETP98111	1	1	1	1	1	2	D					CS4		CS6						
2411	1	ETP98112	1	1	1	1	1	2	D		CS1		CS3								CS21	
2412	1	ETP98114	0	1	1	0	1	2	D										CS14			
2413	1	ETP98115	0	1	1	1	1	2	D							CS6						
2414	1	ETP98068	1	1	1	1	1	2	D	CFA/I												
2415	1	ETP98073	1	0	1	1	1	2	D						CS5	CS6						
2416	1	ETP98097	1	1	1	1	1	2	D										CS14			
2417	1	ETP98103	1	1	1	1	1	2	D		CS1		CS3								CS21	
2418	1	ETP98105	1	1	1	1	1	2	D		CS1		CS3								CS21	
2419	1	ETP98053	1	0	1	1	1	2	D						CS5	CS6						
2420	0	ETP98056	1	0	1	0	1	2	D							CS6						
2421	1	ETP98061	1	1	1	0	1	2	D							CS6						
2422	0	ETP98062	1	0	1	0	1	2	D										CS14			
2423	0	ETP98066	0	0	1	1	1	2	D							CS6						
2424	0	ETP98004	0	0	0	0	1	2	D													1
2425	1	ETP98015	1	0	1	0	1	2	D						CS5	CS6						
2426	1	ETP98028	1	0	1	0	1	2	D						CS5	CS6						

2427	1	ETP98038	1	1	1	0	1	2	D	CFA/I												
2428	1	ETP98042	1	0	1	1	1	2	D						CS5	CS6						
2430	1	BCE003, DS5	1	1	0	1	1	1	D													1
2431	1	BCE021, DS7	0	1	0	0	1	1	D													1
2432	1	BCE046, DS7	1	0	1	1	1	1	D						CS5	CS6						
2433	1	BCE062, DS2	1	1	0	1	1	1	D													1
2434	1	BCE007, DS5	1	0	1	1	1	1	D						CS5	CS6						
2435	1	BCE022,DS6	1	1	1	0	1	1	D	CFA/I												
2436	1	BCE049, DS3	1	1	0	1	1	1	D													1
2437	1	BCE063, DS4	0	1	1	1	1	1	D										CS14			
2438	1	BCE011, DS3	0	1	0	1	0	1	D													1
2439	1	BCE035, DS6	1	1	1	1	1	1	D			CS2	CS3									
2440	1	BCE054, DS4	1	0	1	1	1	1	D						CS5	CS6						
2441	1	BCE066, DS5	0	1	0	1	0	1	D													1
2442	1	BCE013, DS1	1	1	1	0	1	1	D	CFA/I												
2443	1	BCE039, DS2	1	0	1	1	1	1	D						CS5	CS6						
2444	1	BCE055, DS1	1	1	1	1	0	1	D											CS17		
2445	1	BCE069, DS2	1	1	1	1	1	1	D		CS1		CS3								CS21	
2446	1	BCE018, DS6	1	1	0	1	1	1	D													1
2447	1	BCE046, DS2	0	1	0	1	0	1	D													1
2448	1	BCE061, DS1	0	0	0	1	1	1	D													1
2449	1	BCE129, DS2	0	1		0	1	1	D													
totals	**161**		106	102	126	134	142															36

Similar data as **Table 6** showing presence (1=yes), or absence (0=no) of YghJ in culture supernatant as determined by immunoblotting (column B).

Table 9. Presence of eaeH gene in isolates sequenced at Genome Sequencing Center for Infectious Diseases (GSCID).

Stock #	**EaeH**	**Strain designation**	GENOTYPE	SOURCE	COMMENTS	severity
2150	1	BCE011, MS-01	asymptomatic colonization	icddr,b	CFA/1 ST+ 2002 asymptomatic colonization	0
2154	1	BCE034, MS-14	asymptomatic colonization	icddr,b	CFA/1 ST+ 2003 asymptomatic colonization	0
2170	1	P0305260.1	clinical E. coli isolate	icddr,b	tox-, severe cholera-like disease, hospitalized patient, isolate 1	2
1388	1	ETEC Envira 8/11	clinical ETEC isolate	from Brazil	CS5/CS6; LT; ST; severe cholera-like disease from 1998	2
1389	1	ETEC Envira 10/1	clinical ETEC isolate	from Brazil	CS5/CS6; LT; ST; severe cholera-like disease from 1998	2
2097	1	2720900	clinical ETEC isolate	icddr,b	CFA/I; ST; mild, self-limited diarrhea; isolate from 2007	1
2099	1	2722950	clinical ETEC isolate	icddr,b	CFA/I LT/ST 2007 mild, self-limited diarrhea	1

2100	1	2726800	clinical ETEC isolate	icddr,b	CS14 LT/ST 2007 mild, self-limited diarrhea	1
2101	1	2726950	clinical ETEC isolate	icddr,b	CS21 ST 2007 mild, self-limited diarrhea	1
2102	1	2729250	clinical ETEC isolate	icddr,b	CFA/I LT/ST 2007 mild, self-limited diarrhea	1
2103	1	2747800	clinical ETEC isolate	icddr,b	CS5+CS6 LT/ST 2007 severe cholera-like disease, hospitalized patient	2
2105	1	2735000	clinical ETEC isolate	icddr,b	CS7 LT 2007 severe cholera-like disease, hospitalized patient	2
2106	1	2730350	clinical ETEC isolate	icddr,b	CFA/I+CS21 LT/ST 2007 severe cholera-like disease, hospitalized patient	2
2108	1	2770900	clinical ETEC isolate	icddr,b	CS5+CS6 LT/ST 2007 severe cholera-like disease, hospitalized patient	2
2109	1	2780750	clinical ETEC isolate	icddr,b	CF-ve ST 2007 severe cholera-like disease, hospitalized patient	2
2110	1	2731150	clinical ETEC isolate	icddr,b	CS6+CS8 LT 2007 severe cholera-like disease, hospitalized patient	2
2111	1	2785200	clinical ETEC isolate	icddr,b	CF-ve LT 2007 severe cholera-like disease, hospitalized patient	2
2112	1	2733950	clinical ETEC isolate	icddr,b	CS4+CS6 LT/ST 2007 severe cholera-like disease, hospitalized patient	2
2113	1	2749250	clinical ETEC isolate	icddr,b	CS5+CS6 LT/ST 2007 severe cholera-like disease, hospitalized patient	2
2114	1	2756500	clinical ETEC isolate	icddr,b	CS7 LT 2007 severe cholera-like disease, hospitalized patient	2
2115	1	2788150	clinical ETEC isolate	icddr,b	CF-ve LT/ST 2007 severe cholera-like disease, hospitalized patient	2
2116	1	2762100	clinical ETEC isolate	icddr,b	CS6+CS8 LT 2007 severe cholera-like disease, hospitalized patient	2
2117	1	2845350	clinical ETEC isolate	icddr,b	CS4+CS6 LT/ST 2008 mild, self-limited diarrhea	1
2118	1	2845650	clinical ETEC isolate	icddr,b	CS4+CS6 LT/ST 2008 mild, self-limited diarrhea	1
2120	1	2848050	clinical ETEC isolate	icddr,b	CS5+CS6 LT/ST 2008 mild, self-limited diarrhea	1
2121	1	2850400	clinical ETEC isolate	icddr,b	CS7 LT 2008 severe cholera-like disease, hospitalized patient	2
2122	1	2850750	clinical ETEC isolate	icddr,b	CS17 LT 2008 severe cholera-like disease, hospitalized patient	2
2124	1	2853500	clinical ETEC isolate	icddr,b	CF-ve LT 2008 severe cholera-like disease, hospitalized patient	2
2125	1	2854350	clinical ETEC isolate	icddr,b	CS14 LT 2008 severe cholera-like disease, hospitalized patient	2
2126	1	2860050	clinical ETEC isolate	icddr,b	CS5+CS6 LT/ST 2008 severe cholera-like disease, hospitalized patient	2
2127	1	2860650	clinical ETEC isolate	icddr,b	CS5+CS6 LT/ST 2008 severe cholera-like disease, hospitalized patient	2
2129	1	2862600	clinical ETEC isolate	icddr,b	CS4+CS6 ST 2008 severe cholera-like disease, hospitalized patient	2
2131	1	2864350	clinical ETEC isolate	icddr,b	CF-ve LT/ST 2008 severe cholera-like disease, hospitalized patient	2
2132	1	2865200	clinical ETEC isolate	icddr,b	CF-ve ST 2008 severe cholera-like disease, hospitalized patient	2
2133	1	2866350	clinical ETEC isolate	icddr,b	CS5+CS6 LT/ST 2008 severe cholera-like disease, hospitalized patient	2
2134	1	2866450	clinical ETEC isolate	icddr,b	CS7 LT 2008 severe cholera-like disease, hospitalized patient	2
2135	1	2866550	clinical ETEC isolate	icddr,b	CS7 LT 2008 severe cholera-like disease, hospitalized patient	2
2136	1	2866750	clinical ETEC isolate	icddr,b	CS7 LT 2008 severe cholera-like disease, hospitalized patient	2
2137	1	2867750	clinical ETEC isolate	icddr,b	CS7 LT 2008 severe cholera-like disease, hospitalized patient	2
2140	1	2872800	clinical ETEC isolate	icddr,b	CS7 LT 2008 severe cholera-like disease, hospitalized patient	2
2142	1	2875150	clinical ETEC isolate	icddr,b	CS4+CS6 LT/ST 2008 severe cholera-like disease, hospitalized patient	2
2143	1	199900.1	clinical ETEC isolate	icddr,b	CS17 LT severe cholera-like disease, hospitalized patient; isolate 1	2

2155	1	P0304816.1	clinical ETEC isolate	icddr,b	STh, severe cholera-like disease, hospitalized patient, isolate 1	2
2184	1	p0305293.1	clinical ETEC isolate	icddr,b	LT/STh severe cholera-like disease, hospitalized patient, isolate 1	2
2216	1	p0302308.1	clinical ETEC isolate	icddr,b	LT/STh severe cholera-like disease, hospitalized patient, isolate 1	2
2226	1	p0304777.1	clinical ETEC isolate	icddr,b	LT severe cholera-like disease, hospitalized patient, isolate 1	2
2267	1	MP020940.1	clinical ETEC isolate	icddr,b	STh, severe cholera-like disease, hospitalized patient, isolate 1	2
2282	1	P0298942.1	clinical ETEC isolate	icddr,b	LT/STp, severe cholera-like disease, hospitalized patient, isolate 1	2
2295	1	p0299438.2	clinical ETEC isolate	icddr,b	STh, severe cholera-like disease, hospitalized patient, isolate 2	2
2305	1	p0299917.1	clinical ETEC isolate	icddr,b	LT/STh severe cholera-like disease, hospitalized patient, isolate 1	2
2337	1	174900	clinical ETEC isolate	icddr,b	CF-ve STh_2010 Diarr hospitalized patient	2
2338	1	178200	clinical ETEC isolate	icddr,b	CF-ve STh_2010 Diarr hospitalized patient	2
2339	1	178850	clinical ETEC isolate	icddr,b	CF-ve LT/STh_2010 Diarr hospitalized patient	2
2342	1	179550	clinical ETEC isolate	icddr,b	CS7 LT_2010 Diarr hospitalized patient	2
2343	1	180200	clinical ETEC isolate	icddr,b	CS7 LT/ST_2010 Diarr hospitalized patient	2
2344	1	180600	clinical ETEC isolate	icddr,b	CS17 LT/ST_2010 Diarr hospitalized patient	2
2345	1	180050	clinical ETEC isolate	icddr,b	CS17 LT_2010 Diarr hospitalized patient	2
2063	1	MP020980.1	ETEC isolate	icddr,b	ST; patient with severe ETEC diarrhea at Mirpur colony 1	2
2356	1	BCE001, MS16	ETEC; asymptomatic	icddr,b	CF-ve LT_2002-04 Asymptomatic community	0
2357	1	BCE002, MS12	ETEC; asymptomatic	icddr,b	CS1+CS3 LT/STh_2002-04 Asymptomatic community	0
2247	1	p0301904.1	non-ETEC. E. coli isolate	icddr,b	non-ETEC hospitalized patient, isolate 1	
2255	1	p0299483.1	non-ETEC. E. coli isolate	icddr,b	non-ETEC hospitalized patient, isolate 1	
2264	1	p0304799.1	non-ETEC. E. coli isolate	icddr,b	non-ETEC hospitalized patient, isolate 1	
2149	0	BCE006, MS-23	asymptomatic colonization	icddr,b	CS1+CS3 LT+ST+ 2004 asymptomatic colonization	0
2151	0	BCE019, MS-13	asymptomatic colonization	icddr,b	CS1+CS3 LT+ST+ 2003 asymptomatic colonization	0
2152	0	BCE030, MS-09	asymptomatic colonization	icddr,b	CFA/I ST+ 2003 asymptomatic colonization	0
2153	0	BCE032, MS-12	asymptomatic colonization	icddr,b	CS1+CS3 LT+ST+ 2003 asymptomatic colonization	0
350	0	ETEC ThroopD "1"	clinical ETEC isolate	R.A. Finkelstein	serotype 063:Nm	2
1390	0	ETEC Jurua 18/11	clinical ETEC isolate	from Brazil	CS7; LT; severe cholera-like disease from 1998	2
1391	0	ETEC Jurua 20/10	clinical ETEC isolate	from Brazil	CS7; LT; severe cholera-like disease from 1998	2
2098	0	2719100	clinical ETEC isolate	icddr,b	CFA/I+CS21; ST; mild, self limited diarrhea from 2007	1
2104	0	2741950	clinical ETEC isolate	icddr,b	CS1+CS3+CS21 LT/ST 2007 severe cholera-like disease, hospitalized patient	2
2107	0	2730450	clinical ETEC isolate	icddr,b	CS1+CS3+CS21 LT/ST 2007 severe cholera-like disease, hospitalized patient	2
2119	0	2846750	clinical ETEC isolate	icddr,b	CS2+CS3 LT/ST 2008 mild, self-limited diarrhea	1
2123	0	2851500	clinical ETEC isolate	icddr,b	CS7 LT 2008 mild, self-limited diarrhea	1
2128	0	2861200	clinical ETEC isolate	icddr,b	CS6 ST 2008 severe cholera-like disease, hospitalized patient	2
2130	0	2863500	clinical ETEC isolate	icddr,b	CFA/I ST 2008 severe cholera-like disease, hospitalized patient	2
2138	0	2871950	clinical ETEC isolate	icddr,b	CS2+CS3 LT 2008 severe cholera-like disease, hospitalized patient	2

2139	0	2872000	clinical ETEC isolate	icddr,b	CS2+CS3 LT 2008 severe cholera-like disease, hospitalized patient	2
2141	0	2875000	clinical ETEC isolate	icddr,b	CS7 LT 2008 severe cholera-like disease, hospitalized patient	2
2145	0	21600.1	clinical ETEC isolate	icddr,b	CS14 LT/ST severe cholera-like disease, hospitalized patient; isolate 1	2
2199	0	p0301867.1	clinical ETEC isolate	icddr,b	STh, severe cholera-like disease, hospitalized patient, isolate 1	2
2208	0	p0302293.2	clinical ETEC isolate	icddr,b	LT/STh severe cholera-like disease, hospitalized patient, isolate 2	2
2336	0	174750	clinical ETEC isolate	icddr,b	CS1+CS3+CS21 LT/STh_2010 Diarr hospitalized patient	2
2340	0	178900	clinical ETEC isolate	icddr,b	CFA/I, CS21 STh_2010 Diarr hospitalized patient	2
2341	0	179100	clinical ETEC isolate	icddr,b	CF-ve LT/STh_2010 Diarr hospitalized patient	2
2346	0	C-34666	clinical ETEC isolate	icddr,b	CF-ve LT/STh_2010 Diarr hospitalized patient	2
2381	0	mp021566.1	clinical ETEC isolate	icddr,b	LT/STh, Mirpur patient, isolate 1	2
2007	0	P02997067.1	ETEC isolate	icddr,b	LT/STh; patient with severe ETEC diarrhea at icddr,b colony 1	2
2021	0	MP021552.1	ETEC isolate	icddr,b	LT/STh; patient with severe ETEC diarrhea at Mirpur colony 1	2
2036	0	MP021561.1	ETEC isolate	icddr,b	STh; patient with severe ETEC diarrhea at Mirpur colony 1	2
2053	0	MP021017.1	ETEC isolate	icddr,b	ST; patient with severe ETEC diarrhea at Mirpur colony 1	2
2359	0	BCE007, MS11	ETEC; asymptomatic	icddr,b	CS1+CS3+CS21 LT/STh_2002-04 Asymptomatic community	0
2360	0	BCE008, MS1	ETEC; asymptomatic	icddr,b	CF-ve LT_2002-04 Asymptomatic community	0
2361	0	BCE008, MS13	ETEC; asymptomatic	icddr,b	CS6 STp_2002-04 Asymptomatic community	0
2358		BCE005, MS23	ETEC; asymptomatic	icddr,b	CF-ve LT/STh_2002-04 Asymptomatic community	0
	63					
60/91 GSCID ETEC isolates have EaeH gene in the genome.						

References for the Examples.

1. Kotloff KL, Nataro JP, Blackwelder WC, Nasrin D, Farag TH, et al. (2013) Burden and aetiology of diar-rhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): a prospective, case-control study. Lancet doi: 10.1016/S0140-6736(13)60844-2 PMID: 23680352
2. Fleckenstein JM, Hardwidge PR, Munson GP, Rasko DA, Sommerfelt H, et al. (2010) Molecular mech¬anisms of enterotoxigenic Escherichia coli infection. Microbes Infect 12: 89–98. doi: 10.1016/j.micinf. 2009.10.002 PMID: 19883790
3. Evans DG, Silver RP, Evans DJ Jr., Chase DG, Gorbach SL (1975) Plasmid-controlled colonization fac¬tor associated with virulence in Escherichia coli enterotoxigenic for humans. Infect Immun 12: 656– 667. PMID: 1100526
4. Carpenter CC, Barua D, Wallace CK, Sack RB, Mitra PP, et al. (1965) Clinical and physiological obser¬vations during an epidemic outbreak of non-vibrio

cholera-like disease in Calcutta. Bull World Health Organ 33: 665–671. PMID: 5295147

5. Gorbach SL, Banwell JG, Chatterjee BD, Jacobs B, Sack RB (1971) Acute undifferentiated human diar¬rhea in the tropics. I. Alterations in intestinal micrflora. J Clin Invest 50: 881–889. doi: 10.1172/ JCI106560 PMID: 4926260
6. Sack RB, Gorbach SL, Banwell JG, Jacobs B, Chatterjee BD, et al. (1971) Enterotoxigenic Escherichia coli isolated from patients with severe cholera-like disease. J Infect Dis 123: 378–385. doi: 10.1093/ infdis/123.4.378 PMID: 4938945
7. Evans DG, Satterwhite TK, Evans DJ Jr., DuPont HL (1978) Differences in serological responses and excretion patterns of volunteers challenged with enterotoxigenic Escherichia coli with and without the colonization factor antigen. Infect Immun 19: 883–888. PMID: 346488
8. Satterwhite TK, Evans DG, DuPont HL, Evans DJ Jr. (1978) Role of Escherichia coli colonisation factor antigen in acute diarrhoea. Lancet 2: 181–184. doi: 10.1016/S0140-6736(78)91921-9 PMID: 78384
9. Svennerholm AM, Lundgren A (2012) Recent progress toward an enterotoxigenic Escherichia coli vac¬cine. Expert review of vaccines 11: 495–507. doi: 10.1586/erv.12.12 PMID: 22551034
10. Zhang W, Sack DA (2012) Progress and hurdles in the development of vaccines against enterotoxigen-ic Escherichia coli in humans. Expert review of vaccines 11: 677–694. doi: 10.1586/erv.12.37 PMID: 22873126
11. Rasko DA, Rosovitz MJ, Myers GS, Mongodin EF, Fricke WF, et al. (2008) The pangenome structure of Escherichia coli: comparative genomic analysis of E. coli commensal and pathogenic isolates. J Bac-teriol 190: 6881–6893. doi: 10.1128/JB.00619-08 PMID: 18676672
12. Del Canto F, Botkin DJ, Valenzuela P, Popov V, Ruiz-Perez F, et al. (2012) Identification of the Coli Sur¬face Antigen 23 (CS23), a Novel Adhesin of Enterotoxigenic Escherichia coli. Infection and immunity doi: 10.1128/IAI.00263-12 PMID: 22645287

13. Isidean SD, Riddle MS, Savarino SJ, Porter CK (2011) A systematic review of ETEC epidemiology fo¬cusing on colonization factor and toxin expression. Vaccine 29: 6167–6178. doi: 10.1016/j.vaccine. 2011.06.084 PMID: 21723899
14. Riddle MS, Savarino SJ (2013) Moving beyond a heat-labile enterotoxin-based vaccine against entero-toxigenic Escherichia coli. Lancet Infect Dis. doi: 10.1016/S1473-3099(13)70355-4 PMID: 24291167
15. Behrens RH, Cramer JP, Jelinek T, Shaw H, von Sonnenburg F, et al. (2013) Efficacy and safety of a patch vaccine containing heat-labile toxin from Escherichia coli against travellers' diarrhoea: a phase 3, randomised, double-blind, placebo-controlled field trial in travellers from Europe to Mexico and Guate¬mala. Lancet Infect Dis. doi: 10.1016/S1473-3099(13)70297-4 PMID: 24291168
16. Darsley MJ, Chakraborty S, Denearing B, Sack DA, Feller A, et al. (2012) ACE527 Oral, Live Attenuat¬ed ETEC Vaccine Reduces the Incidence and Severity of Diarrhea in a Human Challenge Model of Di-arrheal Disease. Clin Vaccine Immunol doi: 10.1128/CVI.00364-12 PMID: 23035175
17. Harro C, Sack D, Bourgeois AL, Walker R, Denearing B, et al. (2011) A combination vaccine consisting of three live attenuated enterotoxigenic Escherichia coli strains expressing a range of colonization fac¬tors and LTB is well tolerated and immunogenic in a placebo-controlled double-blind Phase I trial in healthy adults. Clin Vaccine Immunol doi: 10.1128/CVI.05342-11 PMID: 21994354
18. Turner AK, Stephens JC, Beavis JC, Greenwood J, Gewert C, et al. (2011) Generation and characteri¬zation of a live attenuated enterotoxigenic Escherichia coli combination vaccine expressing six coloni¬zation factors and heat-labile toxin subunit B. Clin Vaccine Immunol 18: 2128–2135. doi: 10.1128/CVI. 05345-11 PMID: 21994355
19. Fleckenstein JM, Roy K, Fischer JF, Burkitt M (2006) Identification of a two-partner secretion locus of enterotoxigenic Escherichia coli. Infect Immun 74: 2245–2258. doi: 10.1128/IAI.74.4.2245-2258.2006 PMID: 16552055

20. Roy K, Hilliard GM, Hamilton DJ, Luo J, Ostmann MM, et al. (2009) Enterotoxigenic Escherichia coli EtpA mediates adhesion between flagella and host cells. Nature 457: 594–598. doi: 10.1038/ nature07568 PMID: 19060885
21. Patel SK, Dotson J, Allen KP, Fleckenstein JM (2004) Identification and molecular characterization of EatA, an autotransporter protein of enterotoxigenic Escherichia coli. Infect Immun 72: 1786–1794. doi: 10.1128/IAI.72.3.1786-1794.2004 PMID: 14977988
22. Roy K, Kansal R, Bartels SR, Hamilton DJ, Shaaban S, et al. (2011) Adhesin Degradation Accelerates Delivery of Heat-labile Toxin by Enterotoxigenic Escherichia coli. J Biol Chem 286: 29771–29779. doi: 10.1074/jbc.M111.251546 PMID: 21757737
23. Kumar P, Luo Q, Vickers TJ, Sheikh A, Lewis WG, et al. (2013) EatA, an Immununogenic Protective Antigen of Enterotoxigenic Escherichia coli Degrades Intestinal Mucin. Infect Immun. doi: 10.1128/IAI. 01078-13 PMID: 24478066
24. Johansson ME, Sjovall H, Hansson GC (2013) The gastrointestinal mucus system in health and dis¬ease. Nat Rev Gastroenterol Hepatol 10: 352–361. doi: 10.1038/nrgastro.2013.35 PMID: 23478383
25. Roy K, Bartels S, Qadri F, Fleckenstein JM (2010) Enterotoxigenic Escherichia coli elicits immune re-sponses to multiple surface proteins. Infect Immun 78: 3027–3035. doi: 10.1128/IAI.00264-10 PMID: 20457787
26. Kansal R, Rasko DA, Sahl JW, Munson GP, Roy K, et al. (2013) Transcriptional modulation of entero-toxigenic Escherichia coli virulence genes in response to epithelial cell interactions. Infect Immun 81: 259–270. doi: 10.1128/IAI.00919-12 PMID: 23115039
27. Luo Q, Kumar P, Vickers TJ, Sheikh A, Lewis WG, et al. (2014) Enterotoxigenic Escherichia coli Se¬cretes a Highly Conserved Mucin-Degrading Metalloprotease To Effectively Engage Intestinal Epitheli¬al Cells. Infect Immun 82: 509–521. doi: 10.1128/IAI.01106-13 PMID: 24478067

28. Sheikh A, Lou Q, Roy K, Shabaan S, Kumar P, et al. (2014) Contribution of the highly conserved EaeH surface protein to enterotoxigenic Escherichia coli pathogenesis. Infect Immun doi: 10.1128/IAI.01890-14 PMID: 24935979

29. Kumar P, Luo Q, Vickers TJ, Sheikh A, Lewis WG, et al. (2014) EatA, an Immunogenic Protective Anti¬gen of Enterotoxigenic Escherichia coli, Degrades Intestinal Mucin. Infect Immun 82: 500–508. doi: 10. 1128/IAI.01078-13 PMID: 24478066

30. Roy K, Hamilton DJ, Fleckenstein JM (2012) Cooperative role of antibodies against heat-labile toxin and the EtpA Adhesin in preventing toxin delivery and intestinal colonization by enterotoxigenic Escher-ichia coli. Clin Vaccine Immunol 19: 1603–1608. doi: 10.1128/CVI.00351-12 PMID: 22875600

31. Roy K, Hamilton D, Ostmann MM, Fleckenstein JM (2009) Vaccination with EtpA glycoprotein or flagel-lin protects against colonization with enterotoxigenic Escherichia coli in a murine model. Vaccine 27: 4601–4608. doi: 10.1016/j.vaccine.2009.05.076 PMID: 19523914

32. Mundell DH, Anselmo CR, Wishnow RM (1976) Factors influencing heat-labile Escherichia coli entero-toxin activity. Infect Immun 14: 383–388. PMID: 9363

33. Alam NH, Ashraf H (2003) Treatment of infectious diarrhea in children. Paediatric drugs 5: 151–165. PMID: 12608880

34. Sjoling A, Wiklund G, Savarino SJ, Cohen DI, Svennerholm AM (2007) Comparative analyses of phe¬notypic and genotypic methods for detection of enterotoxigenic Escherichia coli toxins and colonization factors. J Clin Microbiol 45: 3295–3301. doi: 10.1128/JCM.00471-07 PMID: 17687011

35. Vicente AC, Teixeira LF, Iniguez-Rojas L, Luna MG, Silva L, et al. (2005) Outbreaks of cholera-like diar¬rhoea caused by enterotoxigenic Escherichia coli in the Brazilian Amazon Rainforest. Trans R Soc Trop Med Hyg 99: 669–674. doi: 10.1016/j.trstmh.2005.03.007 PMID: 15975612

36. Finkelstein RA, Vasil ML, Jones JR, Anderson RA, Barnard T (1976) Clinical cholera caused by entero-toxigenic Escherichia coli. J Clin Microbiol 3: 382–384. PMID: 773963

37. Del Canto F, Valenzuela P, Cantero L, Bronstein J, Blanco JE, et al. (2011) Distribution of Classical and Nonclassical Virulence Genes in Enterotoxigenic Escherichia coli Isolates from Chilean Children and tRNA Gene Screening for Putative Insertion Sites for Genomic Islands. J Clin Microbiol 49: 3198– 3203. doi: 10.1128/JCM.02473-10 PMID: 21775541

38. Harlow E, Lane D, Harlow E (1999) Using antibodies: a laboratory manual. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press. xiv, 495 p. p.

39. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215: 403–410. doi: 10.1006/jmbi.1990.9999 PMID: 2231712

40. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, et al. (2011) Fast, scalable generation of high-qual¬ity protein multiple sequence alignments using Clustal Omega. Molecular systems biology 7: 539. doi: 10.1038/msb.2011.75 PMID: 21988835

41. Team RC (2014) R: A language and environment for statistical computing. Vienna, Austria: R Founda¬tion for Statistical Computing.

42. Warnes G, Bolker B, Bonebakker L, Gentleman R, Huber W, et al. (2014) gplots: Various R program¬ming tools for plotting data. R package version 2.13.0. ed.

43. Neuwirth E (2011) RColorBrewer: ColorBrewer palettes. R package version 1.0–5 ed.

44. Sheikh A, Luo Q, Roy K, Shaaban S, Kumar P, et al. Contribution of the highly conserved EaeH surface protein to enterotoxigenic Escherichia coli pathogenesis. Infect Immun in press.

45. Fleckenstein JM, Roy K (2009) Purification of recombinant high molecular weight two-partner secretion proteins from Escherichia coli. Nat Protoc 4: 1083–1092. doi: 10.1038/nprot.2009.87 PMID: 19707189

46. Tsang VC, Wilson BC, Maddison SE (1980) Kinetic studies of a quantitative single-tube enzyme-linked immunosorbent assay. Clinical chemistry 26: 1255–1260. PMID: 6772340

47. Harris JA, Roy K, Woo-Rasberry V, Hamilton DJ, Kansal R, et al. (2011) Directed evaluation of entero-toxigenic Escherichia coli autotransporter proteins as putative vaccine candidates. PLoS Negl Trop Dis 5: e1428. doi: 10.1371/journal.pntd.0001428 PMID: 22163060

48. Dorsey FC, Fischer JF, Fleckenstein JM (2006) Directed delivery of heat-labile enterotoxin by entero-toxigenic Escherichia coli. Cell Microbiol 8: 1516–1527. doi: 10.1111/j.1462-5822.2006.00736.x PMID: 16922869

49. Sahl JW, Steinsland H, Redman JC, Angiuoli SV, Nataro JP, et al. (2011) A comparative genomic anal¬ysis of diverse clonal types of enterotoxigenic Escherichia coli reveals pathovar-specific conservation. Infect Immun 79: 950–960. doi: 10.1128/IAI.00932-10 PMID: 21078854

50. Svennerholm AM, Vidal YL, Holmgren J, McConnell MM, Rowe B (1988) Role of PCF8775 antigen and its coli surface subcomponents for colonization, disease, and protective immunogenicity of enterotoxi-genic Escherichia coli in rabbits. Infection and immunity 56: 523–528. PMID: 3276628

51. McConnell MM, Thomas LV, Willshaw GA, Smith HR, Rowe B (1988) Genetic control and properties of coli surface antigens of colonization factor antigen IV (PCF8775) of enterotoxigenic Escherichia coli. In¬fect Immun 56: 1974–1980. PMID: 2456269

52. Evans DJ Jr., Evans DG (1973) Three characteristics associated with enterotoxigenic Escherichia coli isolated from man. Infect Immun 8: 322–328. PMID: 4581006

53. Coster TS, Wolf MK, Hall ER, Cassels FJ, Taylor DN, et al. (2007) Immune response, ciprofloxacin ac¬tivity, and gender differences after human experimental challenge by two strains of enterotoxigenic Escherichia coli. Infect Immun 75: 252–259. doi: 10.1128/IAI.01131-06 PMID: 17074855

54. Chen Q, Savarino SJ, Venkatesan MM (2006) Subtractive hybridization and optical mapping of the en-terotoxigenic Escherichia coli H10407 chromosome: isolation of unique sequences and demonstration of significant similarity to the

chromosome of E. coli K-12. Microbiology 152: 1041–1054. doi: 10.1099/mic.0.28648-0 PMID: 16549668

55. Moriel DG, Bertoldi I, Spagnuolo A, Marchi S, Rosini R, et al. (2010) Identification of protective and broadly conserved vaccine antigens from the genome of extraintestinal pathogenic Escherichia coli. Proc Natl Acad Sci U S A 107: 9072–9077. doi: 10.1073/pnas.0915077107 PMID: 20439758
56. Roy K, Hamilton D, Allen KP, Randolph MP, Fleckenstein JM (2008) The EtpA exoprotein of enterotoxi-genic Escherichia coli promotes intestinal colonization and is a protective antigen in an experimental model of murine infection. Infect Immun 76: 2106–2112. doi: 10.1128/IAI.01304-07 PMID: 18285493
57. Sack RB (2011) The discovery of cholera—like enterotoxins produced by Escherichia coli causing se¬cretory diarrhoea in humans. The Indian journal of medical research 133: 171–180. PMID: 21415491
58. Sheikh A, Luo Q, Roy K, Shabaan S, Kumar P, et al. (2014) Contribution of the highly conserved EaeH surface protein to enterotoxigenic Escherichia coli pathogenesis. Infect Immun 82: 3657–3666. doi: 10. 1128/IAI.01890-14 PMID: 24935979
59. Fleckenstein JM, Sheikh A (2014) Designing vaccines to neutralize effective toxin delivery by entero-toxigenic Escherichia coli. Toxins (Basel) 6: 1799–1812. doi: 10.3390/toxins6061799 PMID: 24918359
60. Fleckenstein J, Sheikh A, Qadri F (2014) Novel antigens for enterotoxigenic Escherichia coli vaccines. Expert review of vaccines 13: 631–639. doi: 10.1586/14760584.2014.905745 PMID: 24702311
61. Crossman LC, Chaudhuri RR, Beatson SA, Wells TJ, Desvaux M, et al. (2010) A commensal gone bad: complete genome sequence of the prototypical enterotoxigenic Escherichia coli strain H10407. Journal of bacteriology 192: 5822–5831. doi: 10.1128/JB.00710-10 PMID: 20802035
62. Madhavan TP, Steen JA, Hugenholtz P, Sakellaris H (2014) Genome Sequence of Enterotoxigenic Escherichia coli Strain B2C. Genome announcements 2. doi: 10.1128/genomeA.00247-14 PMID: 24723709

63. Scotland SM, McConnell MM, Willshaw GA, Rowe B, Field AM (1985) Properties of wild-type strains of enterotoxigenic Escherichia coli which produce colonization factor antigen II, and belong to serogroups other than O6. J Gen Microbiol 131: 2327–2333. PMID: 3906040

CLAIMS

What is claimed is:

1. A vaccine composition, the vaccine composition comprising EtpA and EatA.
2. The vaccine composition of claim 1, wherein the EatA comprises a mutation that disrupts serine protease activity.
3. The vaccine composition of claim 2, wherein the mutation is one or more mutations at histidine 134, aspartic acid 162, serine 267.
4. The vaccine composition of claim 3, wherein the mutation is one or more mutations selected from the group consisting of H134A, D162A, and S267G.
5. The vaccine composition of claim 1-4, wherein the EatA comprises the passenger domain.
6. The vaccine composition of claim 1, wherein the EatA comprises about 80% identity to SEQ ID NO:1 or SEQ ID NO:2.
7. The vaccine composition of claim 1, wherein the EatA comprises at least 95% identity to SEQ ID NO:1 or SEQ ID NO:2.
8. The vaccine composition of claim 1, wherein the EatA comprises SEQ ID NO:1 or SEQ ID NO:2.
9. The vaccine composition of claim 3, wherein the EatA comprises about 80% identity to SEQ ID NO:1 or SEQ ID NO:2.
10. The vaccine composition of claim 3, wherein the EatA comprises at least 95% identity to SEQ ID NO:1 or SEQ ID NO:2.
11. The vaccine composition of claim 1-10, wherein the EtpA comprises the secreted EtpA portion.

12. The vaccine composition of claim 1-11 wherein the EtpA comprises about 80% identity to SEQ ID NO:3 or SEQ ID NO:4.

13. The vaccine composition of claim 1-11, wherein the EtpA comprises at least 94% identity to SEQ ID NO:3 or SEQ ID NO:4.

14. The vaccine composition of claim 1-10, wherein the EtpA comprises SEQ ID NO:3 or SEQ ID NO:4.

15. The vaccine composition of claims 1-14, wherein the vaccine composition comprises EtpA and EatA linked together.

16. The vaccine composition of claim 1-15, further comprising colonization factor antigens.

17. The vaccine composition of claim 1-16, wherein the vaccine composition comprises a suitable adjuvant.

18. The vaccine composition of claim 23, wherein the adjuvant is selected from the group consisting of heat-labile toxin and double mutant heat-labile toxin (dmLT).

19. A method of protecting against intestinal colonization of enterotoxigenic *Escherichia coli* (ETEC) in a subject, the method comprising administering to the subject an effective amount of a vaccine composition comprising EatA and EtpA.

20. A method of preventing or treating ETEC-associated diarrhea, the method comprising administering to the subject an effective amount of a vaccine composition comprising EatA and EtpA.

21. A method of preventing or treating an ETEC-associated infection in a subject, the method comprising administering to the subject an effective amount of a vaccine composition comprising EatA and EtpA.

22. The method of claims 19-21, wherein the vaccine composition is administered at a dose ranging from about 15 to about 100 μg.

23. The method of claim 19-22, wherein the vaccine composition is administered every 2 to 4 weeks.

24. The method of claim 19-23, wherein the vaccine composition is administered intranasally.

25. The method of claim 19-24, further comprising administering of standard treatments for ETEC-associated infection.

26. The method of claim 19-24, wherein the subject is at risk of ETEC infection.

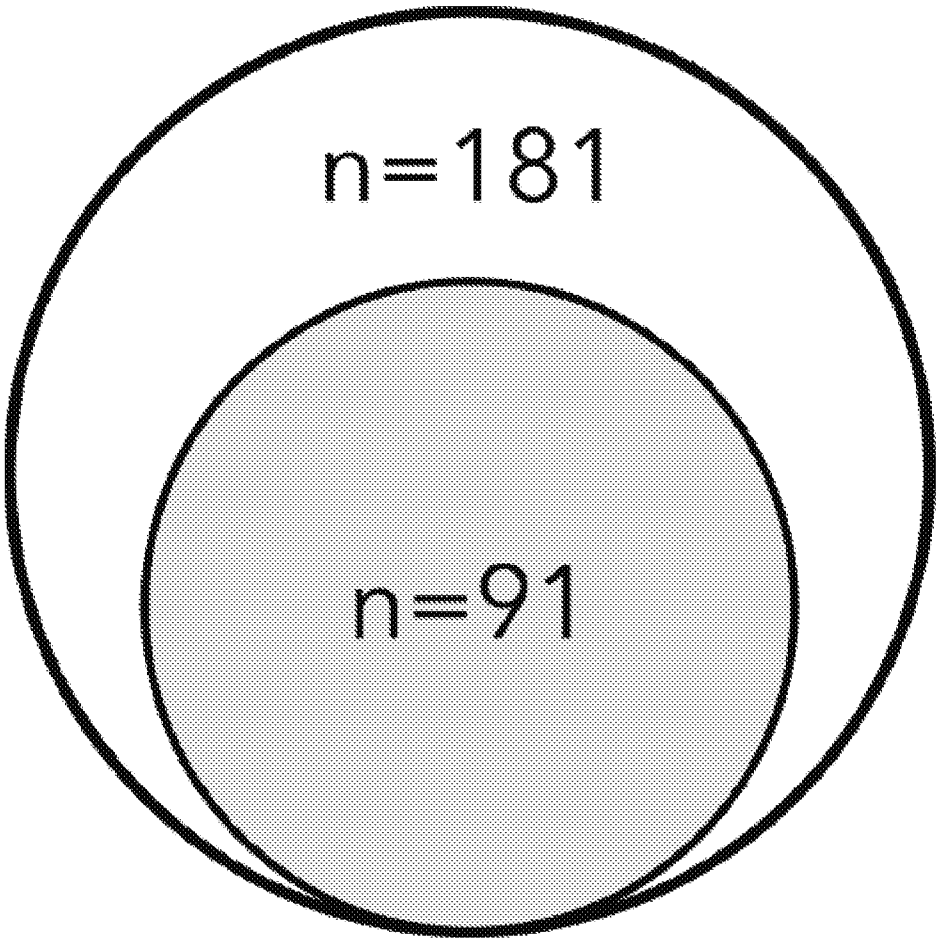


FIG. 1

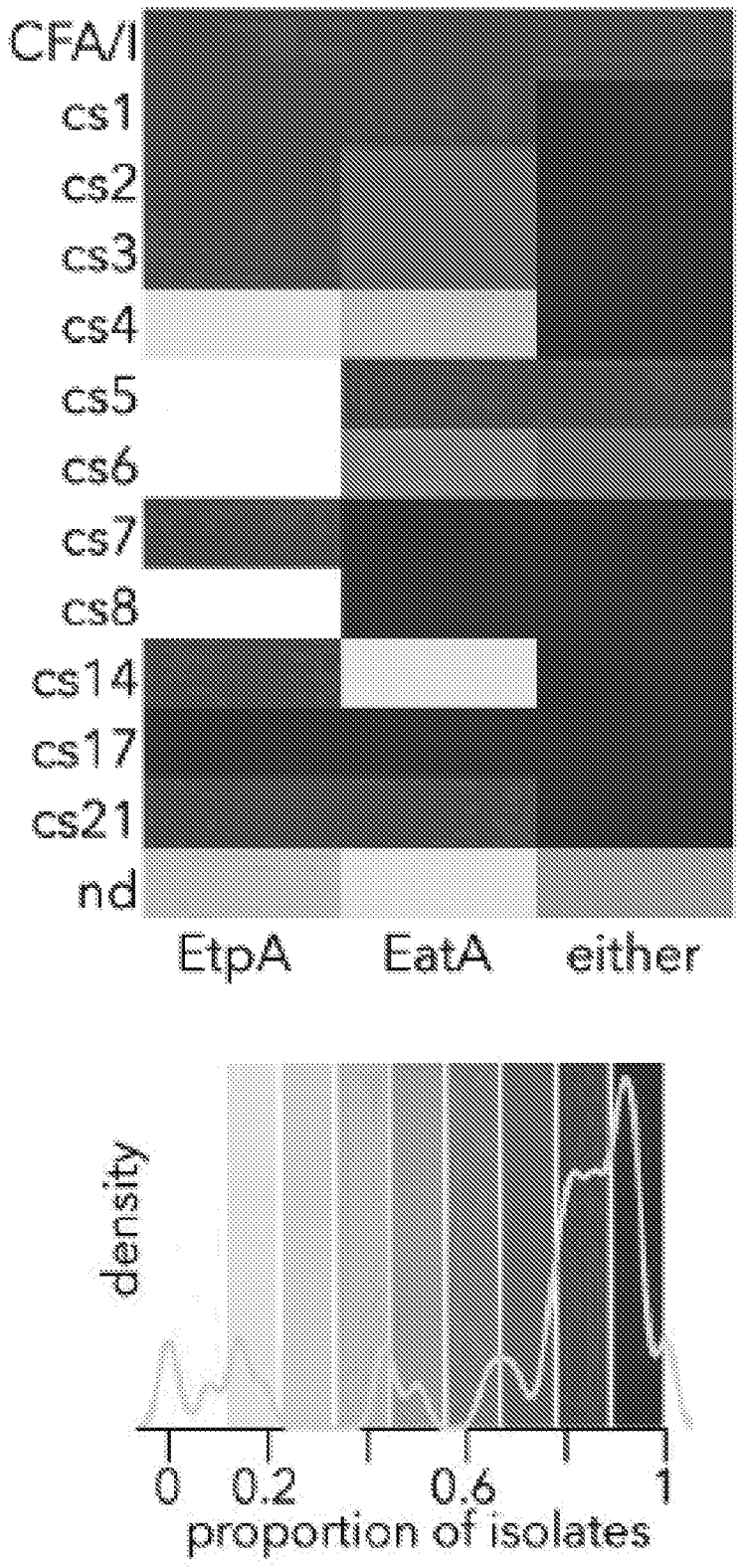

CFA/I
cs1
cs2
cs3
cs4
cs5
cs6
cs7
cs8
cs14
cs17
cs21
nd
EtpA
EatA
either
density
0
0.2
0.6
1
proportion of isolates

FIG. 2A

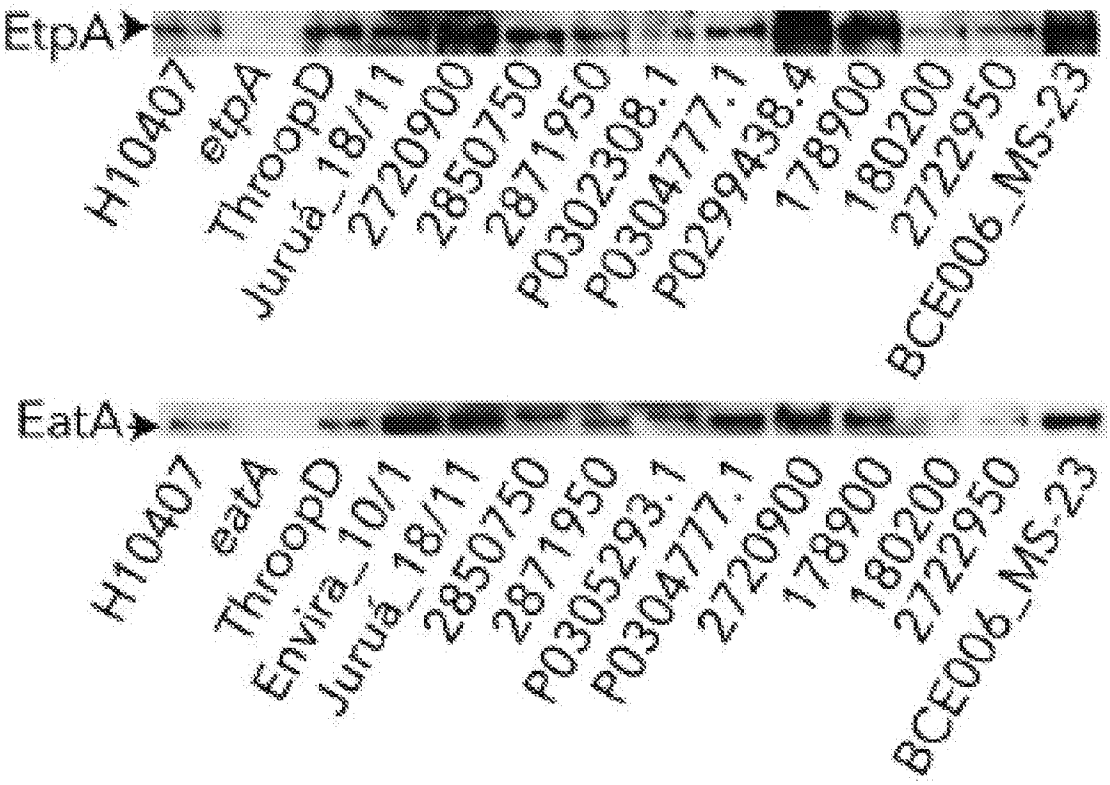

EtpA
H10407
etpA
ThroopD
Juruá_18/11
2720900
2850750
2871950
P0302308.1
P0304777.1
P0299438.4
178900
180200
2722950
BCE006_MS-23
EatA
H10407
eatA
ThroopD
Envira_10/1
Juruá_18/11
2850750
2871950
P0305293.1
P0304777.1
2720900
178900
180200
2722950
BCE006_MS-23

FIG. 2B

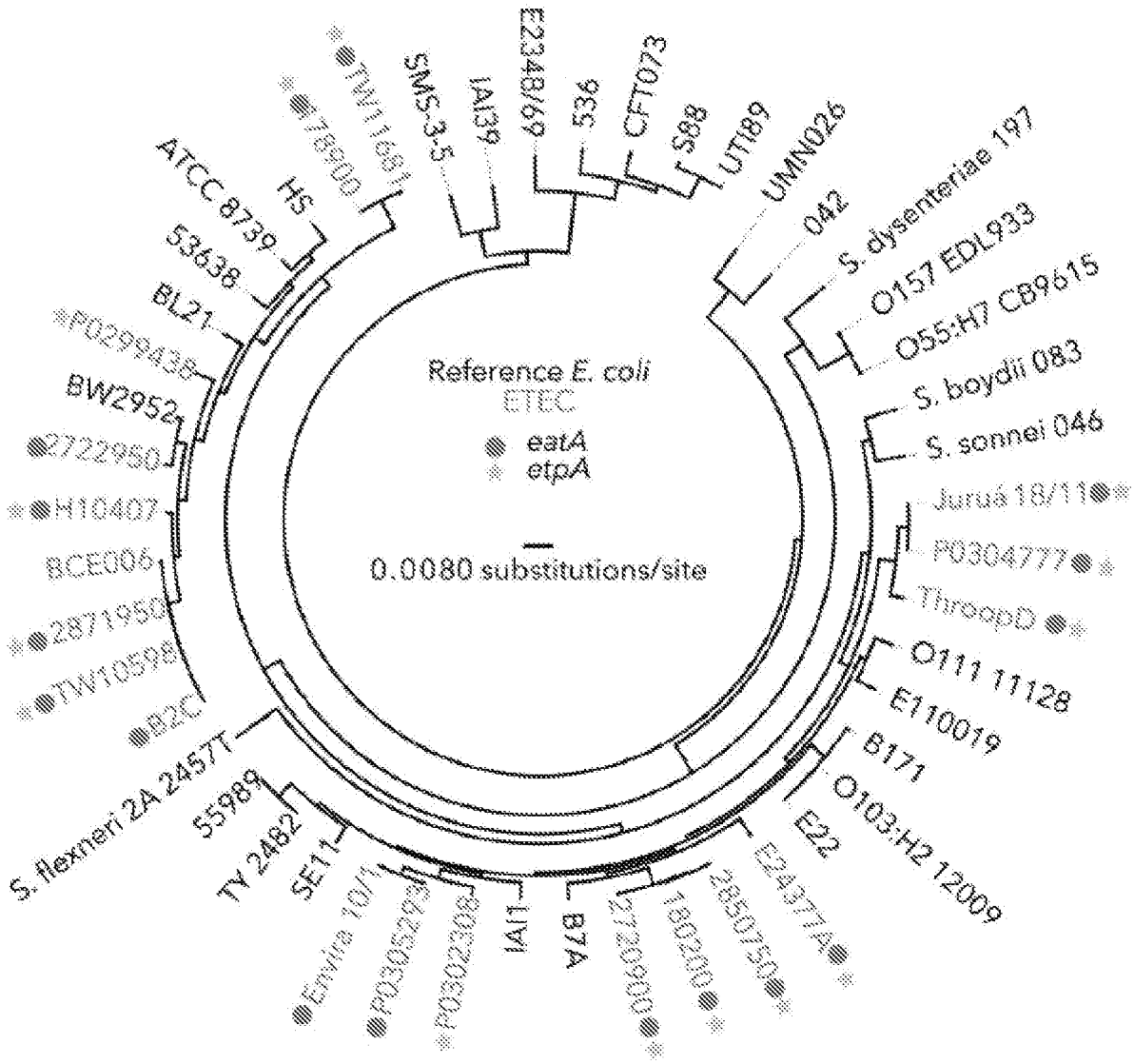

Reference E. coli
ETEC
eatA
etpA
0.0080 substitutions/site
E2348/69
536
CFT073
S88
UTI89
UMN026
042
S. dysenteriae 197
O157 EDL933
O55:H7 CB9615
S. boydii 083
S. sonnei 046
Jurua 18/11
F0304777
ThroopD
O111 11128
E110019
B171
O103:H2 12009
E22
E24377A
2850750
160200
2720900
B7A
UMN
F030232
F030529
Envira 10/1
5E11
TY 2482
55989
S. flexneri 2A 2457T
B2C
TW10598
2871950
BCE006
H10407
2722950
BW2952
F0299438
BL21
53638
ATCC 8739
HS
78900
TW11681
SMS-3-5
IAI39

FIG. 2C

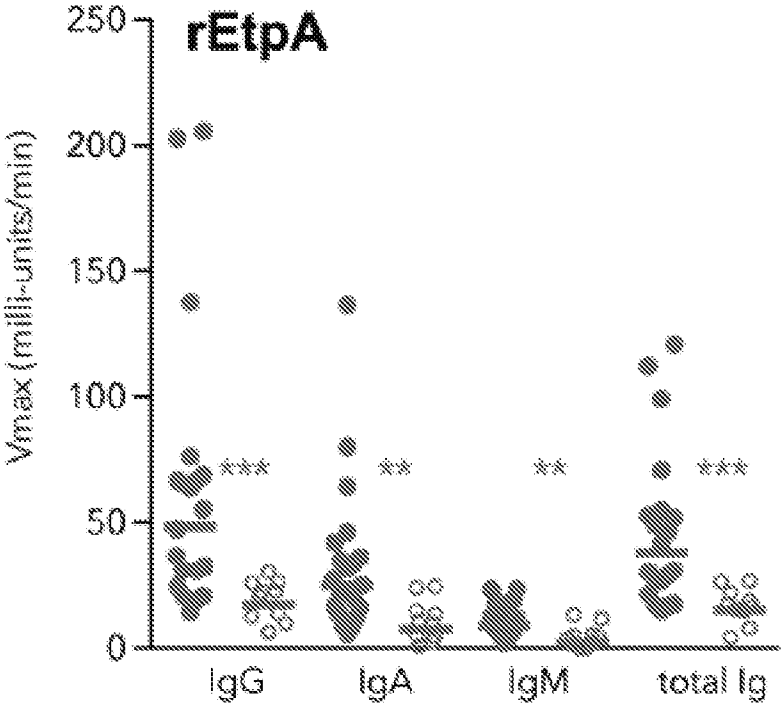

rEtpA
Vmax (milli-units/min)
250
200
150
100
50
0
***
**
**
***
IgG
IgA
IgM
total Ig

FIG. 3A

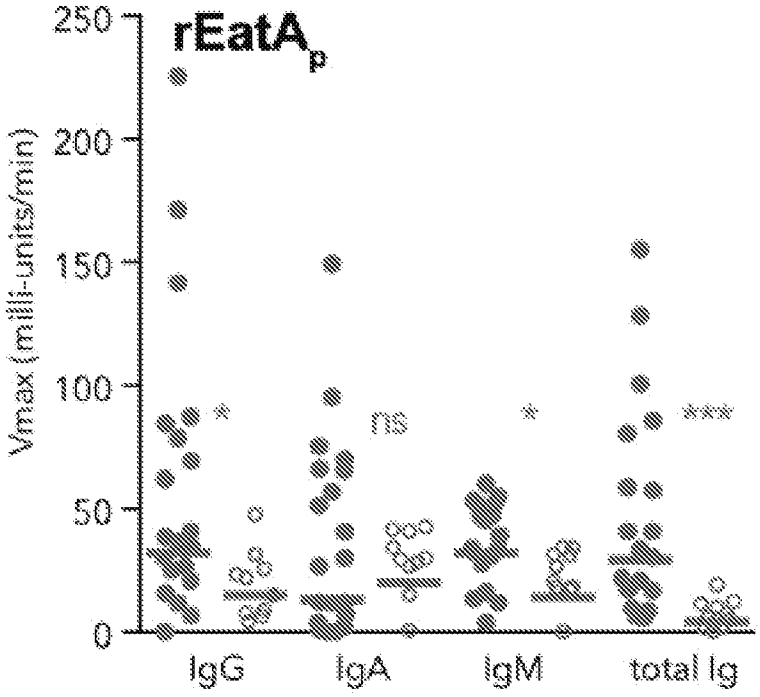

rEatA$_p$
Vmax (milli-units/min)
250
200
150
100
50
0
*
ns
*
***
IgG
IgA
IgM
total Ig

FIG. 3B

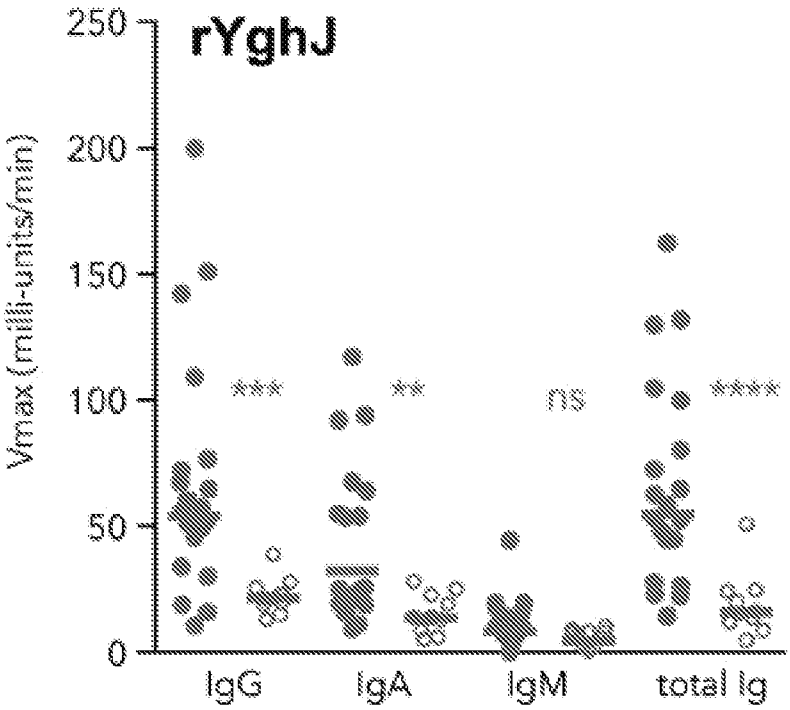

rYghJ
Vmax (milli-units/min)
250
200
150
100
50
0
***
**
ns
****
IgG
IgA
IgM
total Ig

FIG. 3C

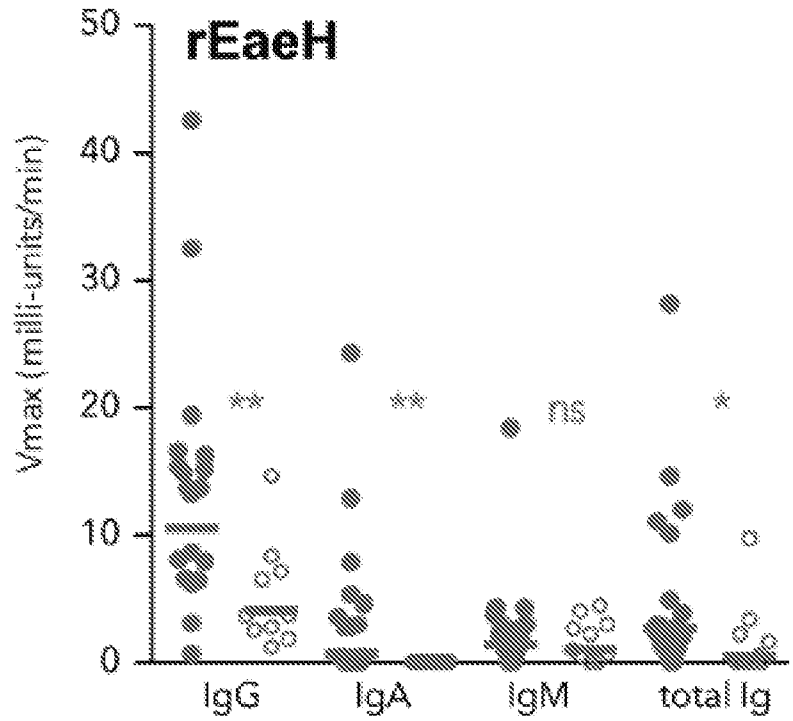

rEaeH
Vmax (milli-units/min)
50
40
30
20
10
0
**
**
ns
*
IgG
IgA
IgM
total Ig

FIG. 3D

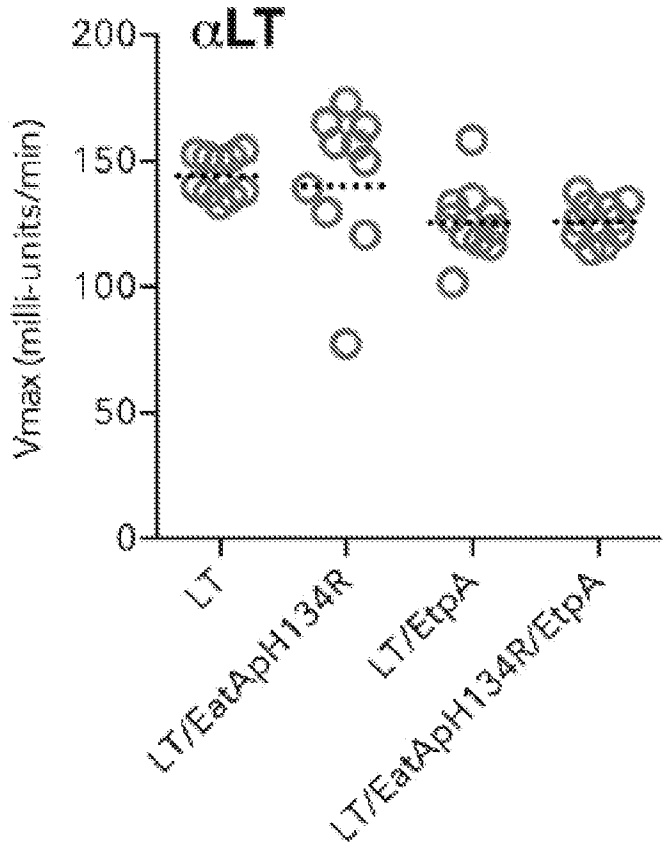

αLT
Vmax (milli-units/min)
200
150
100
50
0
LT
LT/EatApH134R
LT/EtpA
LT/EatApH134R/EtpA

FIG. 4A

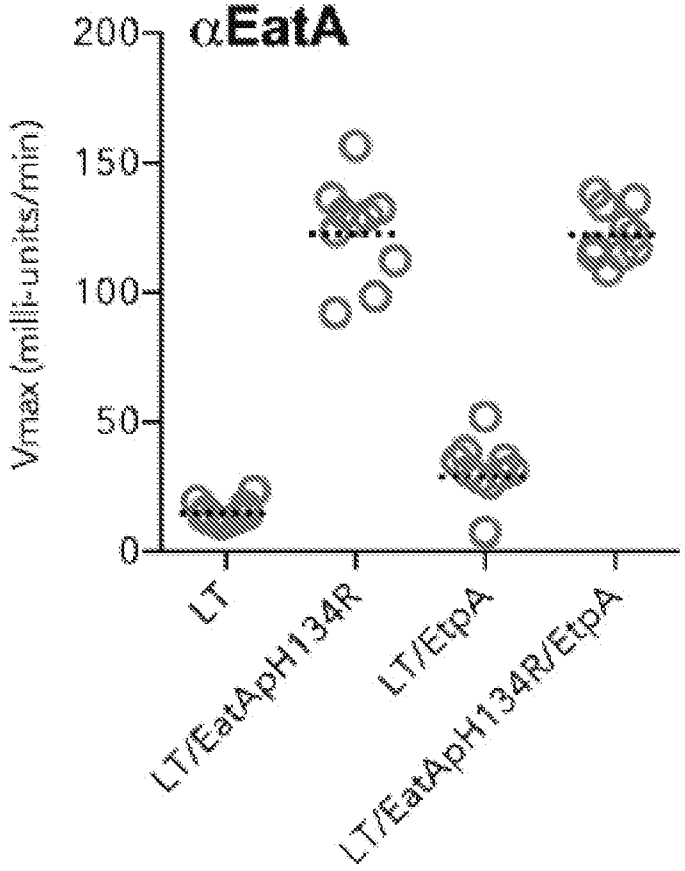

αEatA
Vmax (milli-units/min)
200
150
100
50
0
LT
LT/EatApH134R
LT/EtpA
LT/EatApH134R/EtpA

FIG. 4B

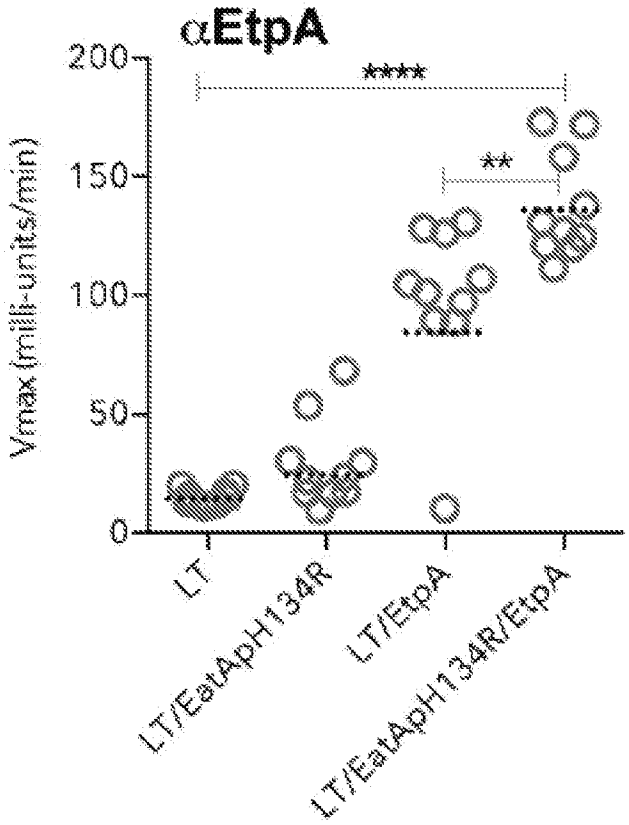

αEtpA
Vmax (milli-units/min)
200
150
100
50
0
****
**
LT
LT/EatApH134R
LT/EtpA
LT/EatApH134R/EtpA

FIG. 4C

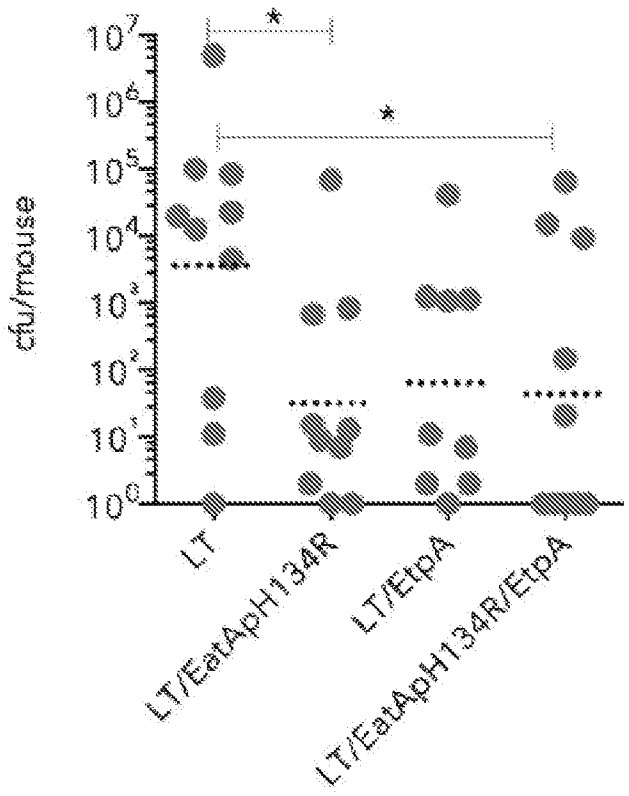

cfu/mouse
*
*
LT
LT/EatApH134R
LT/EtpA
LT/EatApH134R/EtpA

FIG. 4D

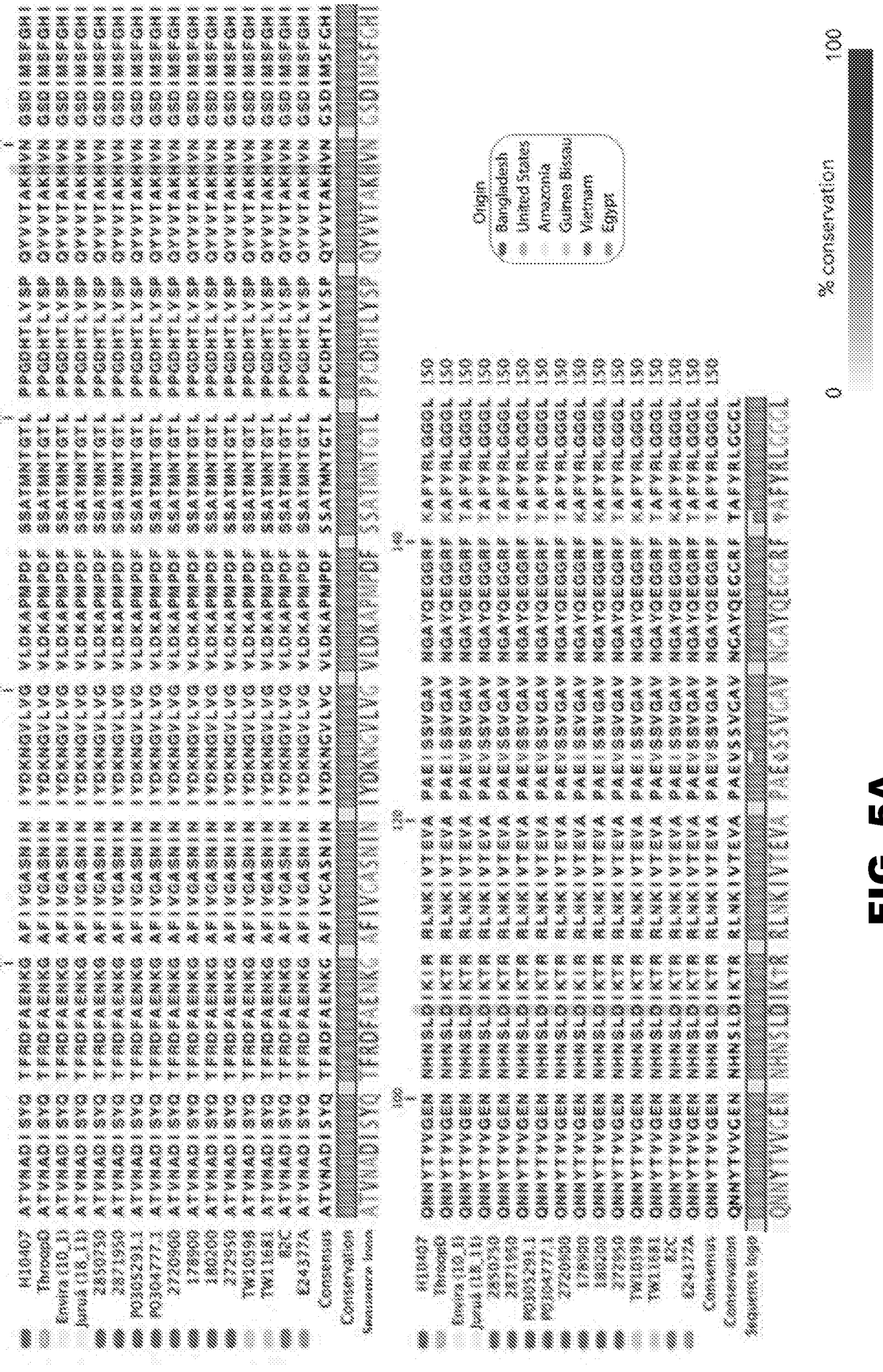

Origin
Bangladesh
United States
Amazonia
Guinea Bissau
Vietnam
Egypt
% conservation
0
100

FIG. 5A

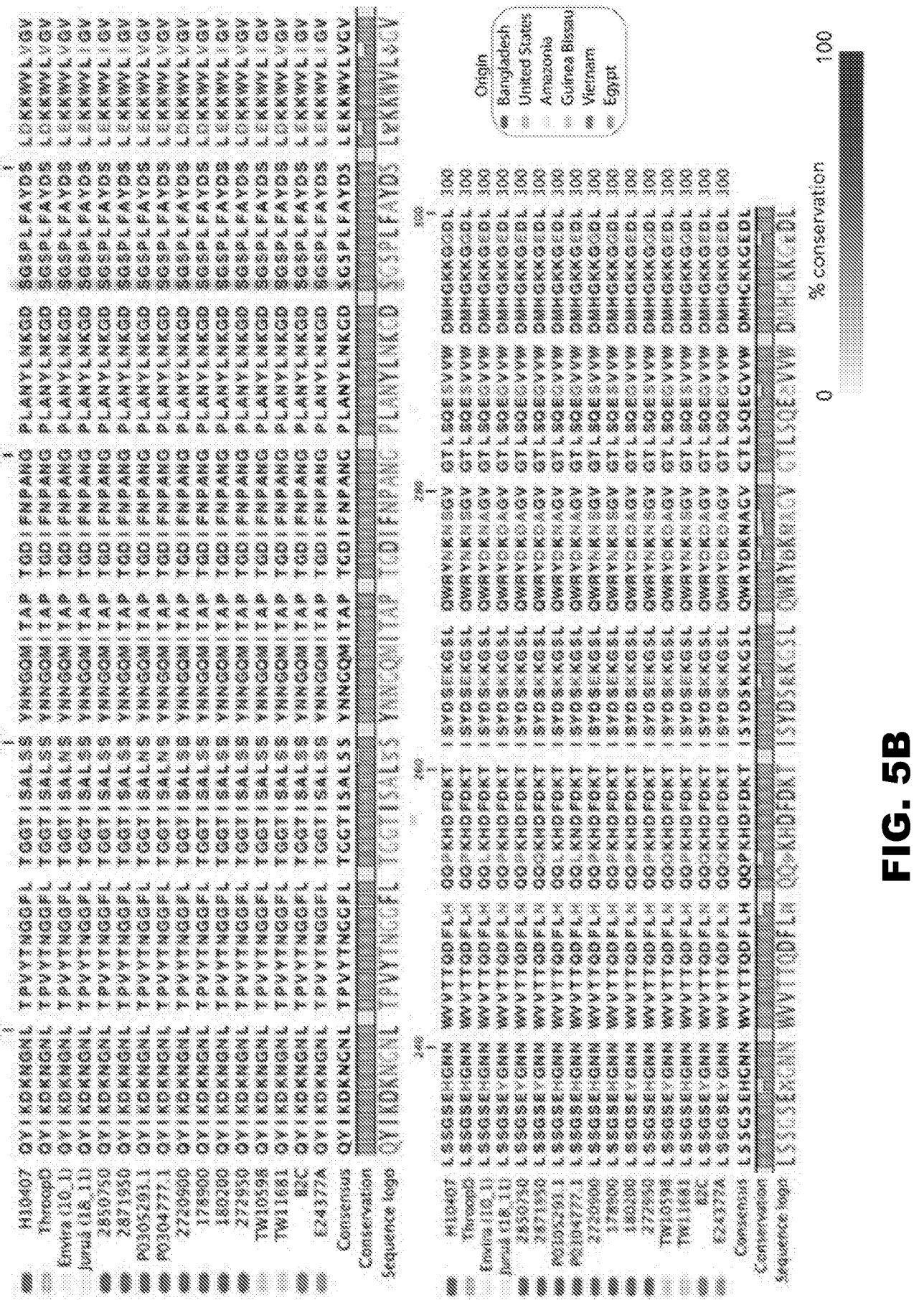

Origin
Bangladesh
United States
Amazonia
Guinea Bissau
Vietnam
Egypt
% conservation
0
100

FIG. 5B

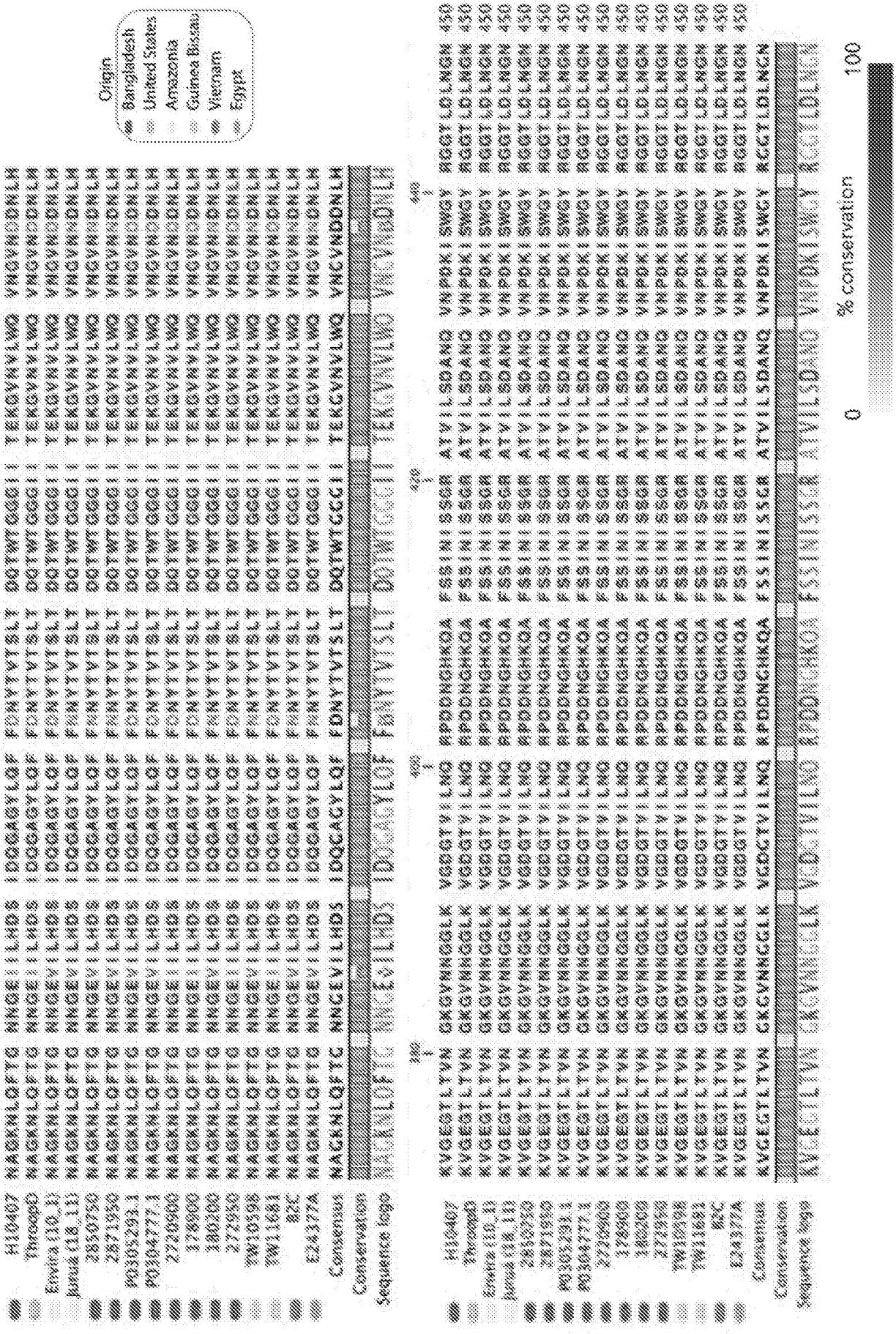

Origin
Bangladesh
United States
Amazonia
Guinea Bissau
Vietnam
Egypt
% conservation
0
100
Consensus
Conservation
Sequence logo

FIG. 5C

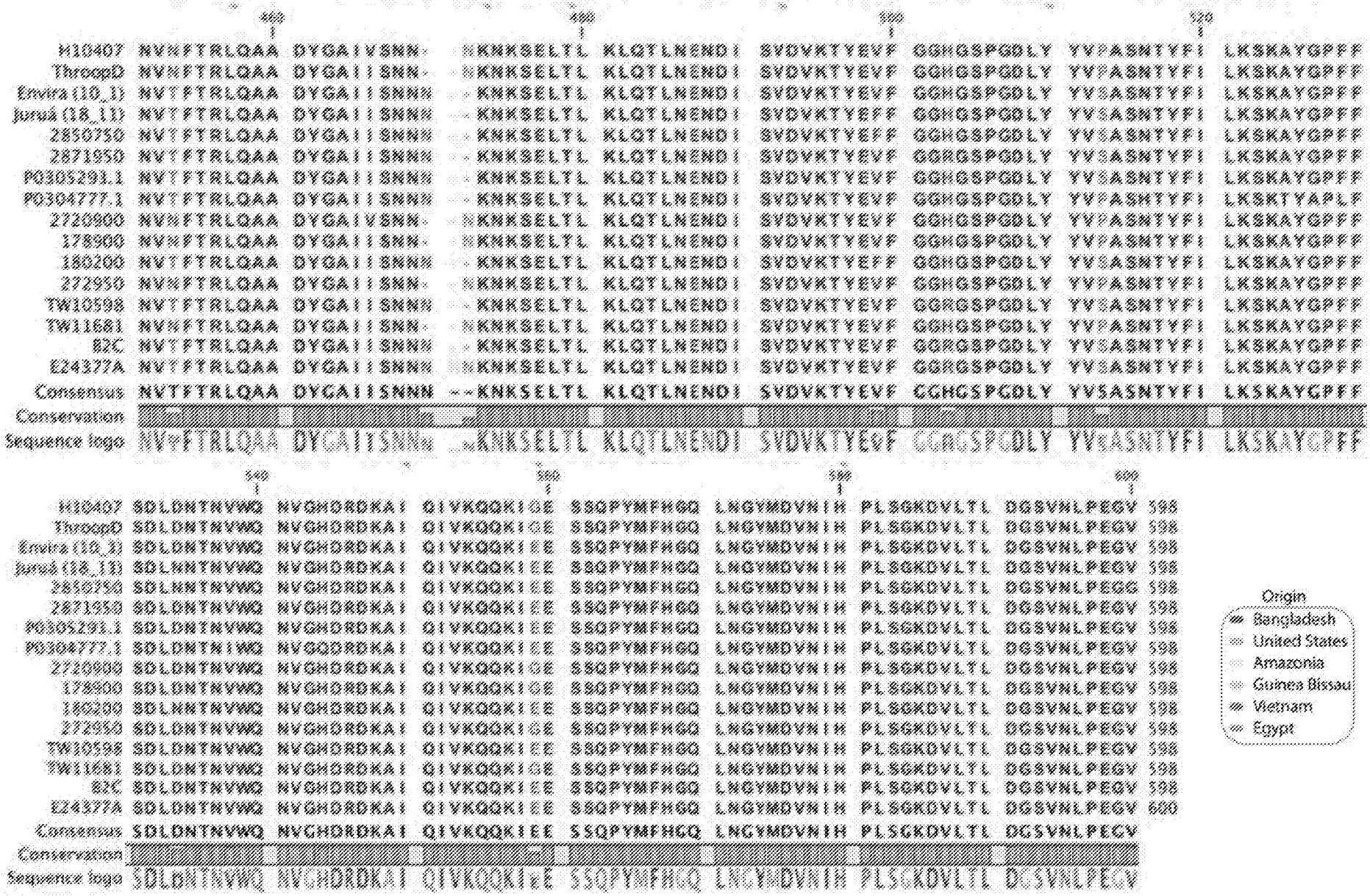

H10407
ThroopD
Envira (10_1)
Juruá (18_11)
2850750
2871950
PO305291.1
PO304777.1
2720900
178900
180200
272950
TW10598
TW11681
82C
E24377A
Consensus
Conservation
Sequence logo
Origin
Bangladesh
United States
Amazonia
Guinea Bissau
Vietnam
Egypt
0
% conservation
100

FIG. 5D

Origin
Bangladesh
United States
Amazonia
Guinea Bissau
Vietnam
Egypt
% conservation
0
100

FIG. 5E

H10407 SLLLEKGAKL TAQSGIFTNM TMKIKENASL TLTGIPSVGK PGYYSPVTST TEGIHLGERA SLSVKNMGYL
ThrosoD SLLLEKGAKL TAQSGIFTNM TMEIKENASL TLTGIPSVGK PGYYSPVIST TEGIHLGERA SLSVKNMGYL
Enviro (10_1) SLLLEKGAKL TAQSGIFTNM TMEIKENASL TLTGIPSVGK PGYYSPVIST TEGIHLGERA SLSVKNMGYL
Jurnl (18_11) SLLLEKGAKL TAQSGIFTNM TMEIKENASL TLTGIPSVGK PGYYSPVIST TEGIHLGERA SLSVKNMGYL
2850750 SLLLEKGAKL TAQSGIFTNM TMEIKENASL TLTGIPSVGK PGYYSPVIST TEGIHLGERA SLSVKNMGYL
2871950 SLLLEKGAKL TAQSGIFTNM TMEIKENASL TLTGIPSVGK PGYYSPVIST TEGIHLGERA SLSVKNMGYL
P03052931 SLLLEKGAKL TAQSGIFTNM TMEIKENASL TLTGIPSVGK PGYYSPVIST TEGIHLGERA SLSVKNMGYL
P03047771 SLLLEKGAKL TAQSGIFTNM TMEIKENASL TLTGIPSVGK PGYYSPVIST TEGIHLGERA SLSVKNMGYL
2720500 SLLLEKGAKL TAQSGIFTNM TMKIKENASL TLTGIPSVGK PGYYSPVTST TEGIHLGERA SLSVKNMGYL
178900 SLLLEKGAKL TAQSGIFTNM TMEIKENASL TLTGIPSVGK PGYYSPVIST TEGIHLGERA SLSVKNMGYL
1802000 SLLLEKGAKL TAQSGIFTNM TMEIKENASL TLTGIPSVGK PGYYSPVIST TEGIHLGERA SLSVKNMGYL
272950 SLLLEKGAKL TAQSGIFTNM TMEIKENASL TLTGIPSVGK PGYYSPVIST TEGIHLGERA SLSVKNMGYL
TW10598 SLLLEKGAKL TAQSGIFTNM TMEIKENASL TLTGIPSVGK PGYYSPVIST TEGIHLGERA SLSVKNMGYL
TW11681 SLLLEKGAKL TAQSGIFTNM TMEIKENASL TLTGIPSVGK PGYYSPVIST TEGIHLGERA SLSVKNMGYL
82C SLLLEKGAKL TAQSGIFTNM TMEIKENASL TLTGIPSVGK PGYYSPVIST TEGIHLGERA SLSVKNMGYL
E24377A SLLLEKGAKL TAQSGIFTNM TMEIKENASL TLTGIPSVGK PGYYSPVIST TEGIHLGERA SLSVKNMGYL
Consensus SLLLEKGAKL TAQSGIFTNM TMEIKENASL TLTGIPSVGK PGYYSPVIST TEGIHLGERA SLSVKNMGYL
Conservation
Sequence logo SLLLEKGAKL TAQSGIFTNM TMEIKENASL TLTGIPSVGK PGYYSPVIST TEGIHLGERA SLSVKNMGYL
H10407 SSNIIAEDSA AIINLGDSNA TIGKTDSPLF STLMKGYNAV LQGNIMGPQS SVNMNNALWH SDRNSEIKEL KANDSQIELG ESS
ThrosoD SSNIIAEDSA AIINLGDSNA TIGKTDSPLF HTLMKGYNAV LQGNIMGPQS SVNMNNALWH SDRNSEIKEL KANDSQIELG ESS
Enviro (10_1) SSNIIAEDSA AIINLGDSNA TIGKTDSPLF HTLMKGYNAV LQGNIMGPQS SVNMNNALWH SDRNSEIKEL KANDSQIELG ESS
Jurnl (18_11) SSNIIAEDSA AIINLGDSNA TIGKTDSPLF HTLMKGYNAV LQGNIMGPQS SVNMNNALWH SDRNSEIKEL KANDSQIELG ESS
2850750 SSNIIAEDSA AIINLGDSNA TIGKTDSPLF HTLMKGYNAV LQGNIMGPQS SVNMNNALWH SDRNSEIKEL KANDSQIELG ESS
2871950 SSNIIAEDSA AIINLGDSNA TIGKTDSPLF HTLMKGYNAV LQGNIMGPQS SVNMNNALWH SDRNSEIKEL KANDSQIELG ESS
P03052931 SSNIIAEDSA AIINLGDSNA TIGKTDSPLF HTLMKGYNAV LQGNIMGPQS SVNMNNALWH SDRNSEIKEL KANDSQIELG ESS
P03047771 SSNITAEDSA AIINLGDSNA TIGKTDSPLF STLMKGYNAV LQGNIMGPQS SVNMNNALWH SDRNSEIKEL KANDSQIELG ESS
2720500 SSNIIAEDSA AIINLGDSNA TIGKTDSPLF HTLMKGYNAV LQGNIMGPQS SVNMNNALWH SDRNSEIKEL KANDSQIELG ESS
178900 SSNIIAEDSA AIINLGDSNA TIGKTDSPLF HTLMKGYNAV LQGNIMGPQS SVNMNNALWH SDRNSEIKEL KANDSQIELG ESS
1802000 SSNIIAEDSA AIINLGDSNA TIGKTDSPLF HTLMKGYNAV LQGNIMGPQS SVNMNNALWH SDRNSEIKEL KANDSQIELG ESS
272950 SSNIIAEDSA AIINLGDSNA TIGKTDSPLF HTLMKGYNAV LQGNIMGPQS SVNMNNALWH SDRNSEIKEL KANDSQIELG ESS
TW10598 SSNIIAEDSA AIINLGDSNA TIGKTDSPLF HTLMKGYNAV LQGNIMGPQS SVNMNNALWH SDRNSEIKEL KANDSQIELG ESS
TW11681 SSNIIAEDSA AIINLGDSNA TIGKTDSPLF HTLMKGYNAV LQGNIMGPQS SVNMNNALWH SDRNSEIKEL KANDSQIELG ESS
82C SSNIIAEDSA AIINLGDSNA TIGKTDSPLF HTLMKGYNAV LQGNIMGPQS SVNMNNALWH SDRNSEIKEL KANDSQIELG ESS
E24377A SSNIIAEDSA AIINLGDSNA TIGKTDSPLF HTLMKGYNAV LQGNIMGPQS SVNMNNALWH SDRNSEIKEL KANDSQIELG ESS
Consensus SSNIIAEDSA AIINLGDSNA TIGKTDSPLF NTLMKGYNAV LQGNIMGPQS SVNMNNALWH SDRNSEIKEL KANDSQIELG ESS
Conservation
Sequence logo SSNIIAEDSA AIINLGDSNA TIGKTDSPLF NTLMKGYNAV LQGNIMGPQS SVNMNNALWH SDRNSEIKEL KANDSQIELG ESS

FIG. 5F

Origin
Bangladesh
United States
Amazonia
Guinea Bissau
Vietnam
Egypt
% conservation
0
100

FIG. 5G

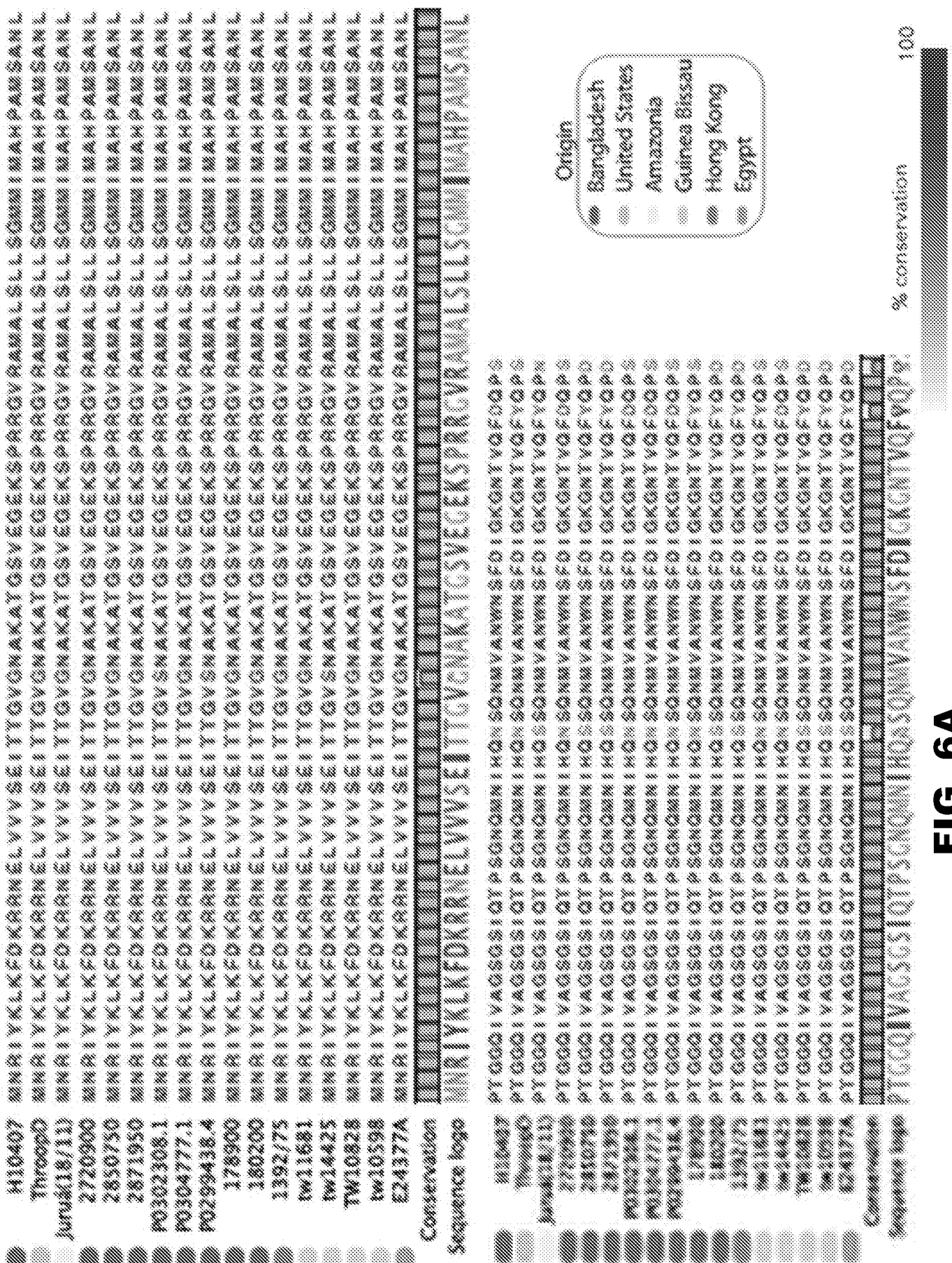

H10407
Throop/0
Jurua/18/111
2720900
2650750
28711950
P03023108.1
P03047777.1
P02994384
178900
180200
1392/75
tw11681
tw14425
TW10828
tw10598
E24377A
Conservation
Sequence logo
Origin
Bangladesh
United States
Amazonia
Guinea Bissau
Hong Kong
Egypt
% conservation
100

FIG. 6A

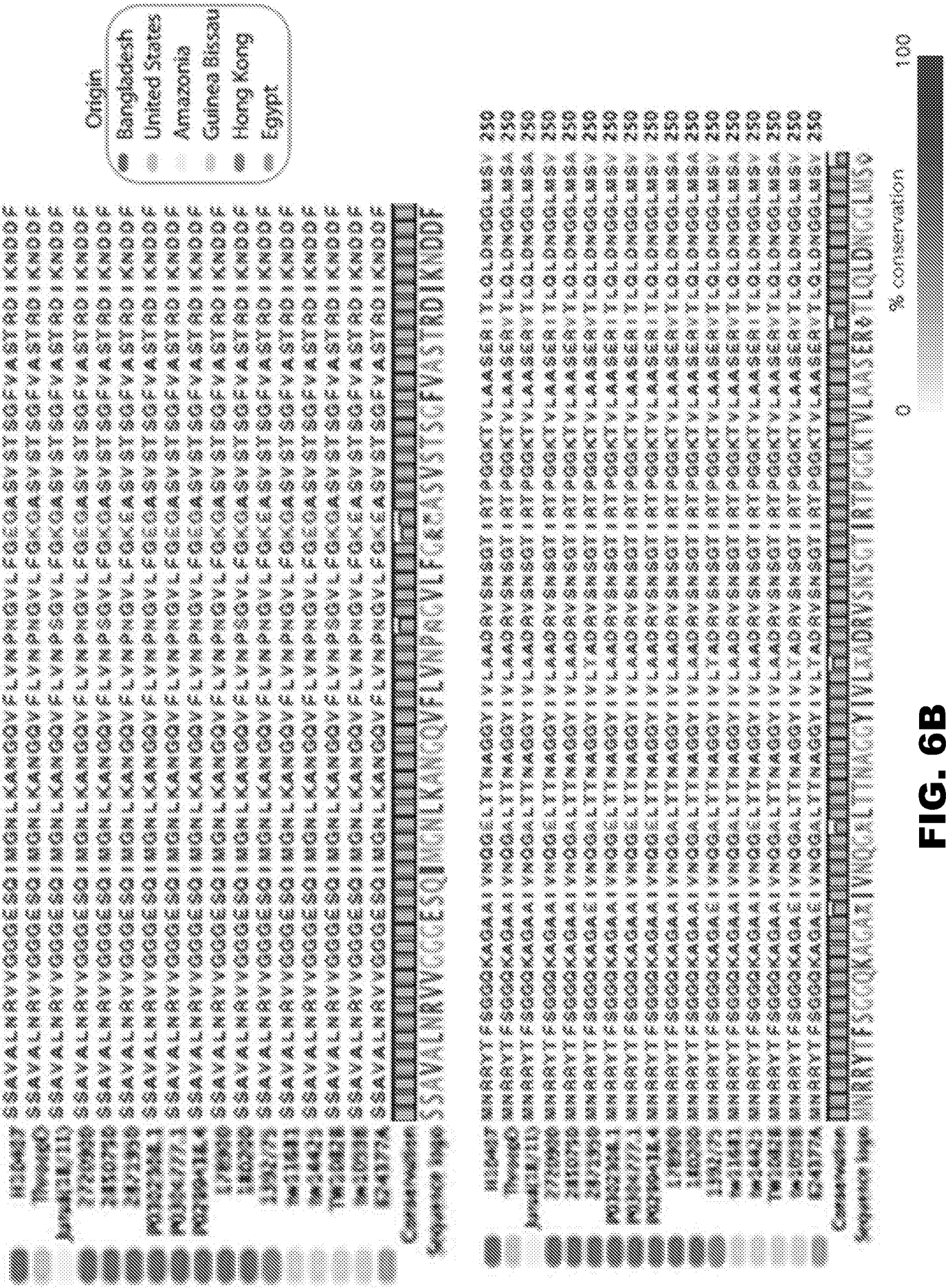

Origin
Bangladesh
United States
Amazonia
Guinea Bissau
Hong Kong
Egypt
% conservation
0
100
Consensus
Sequence logo

FIG. 6B

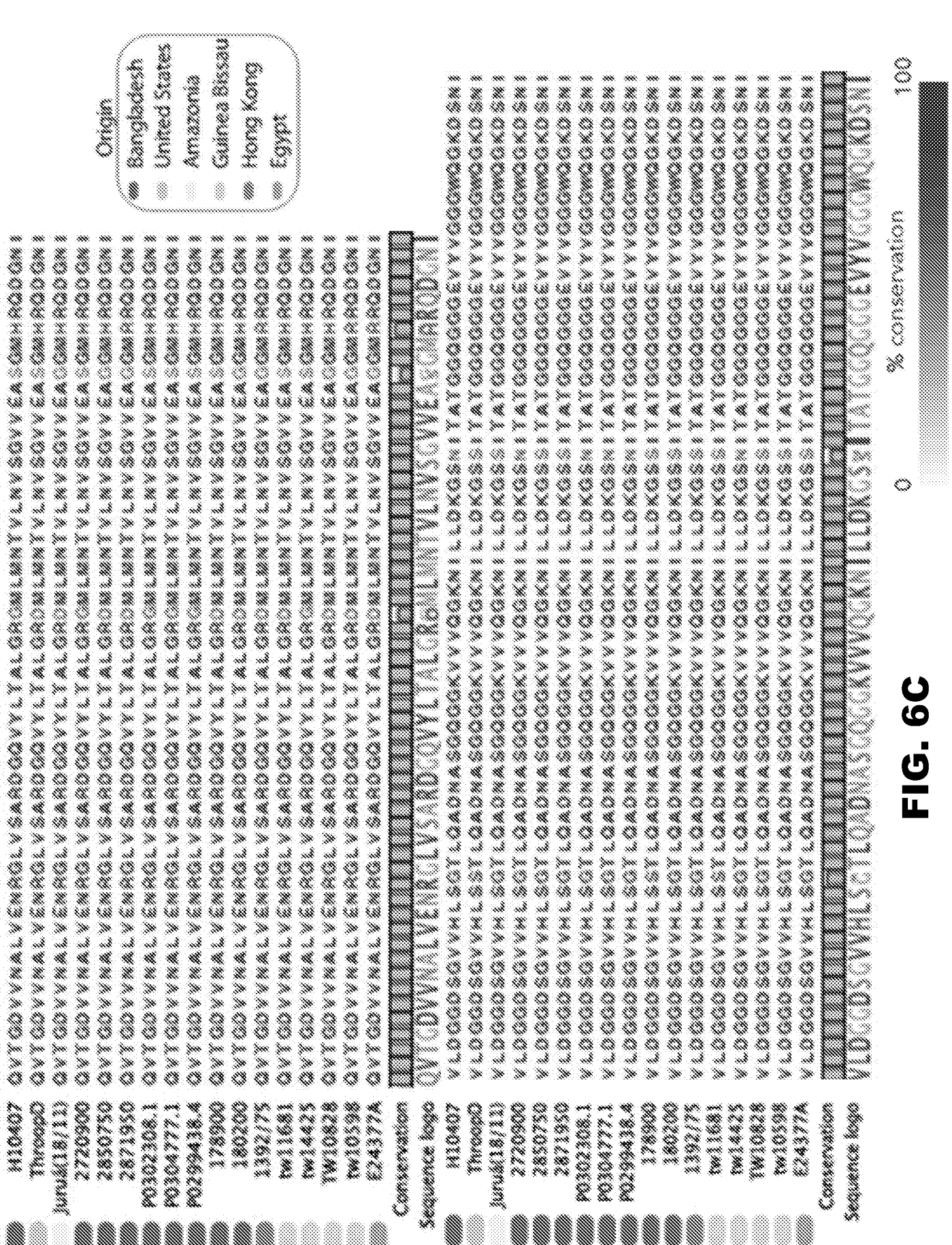

H10407
Thrasp0
Jurukl18/111
2720900
2850750
2871950
P03023108.1
P03047777.1
P02994138.4
178900
180200
1392/75
tw11681
tw14425
TW10828
tw10398
E24377A
Conservation
Sequence logo
Origin
Bangladesh
United States
Amazonia
Guinea Bissau
Hong Kong
Egypt
0
% conservation
100

FIG. 6C

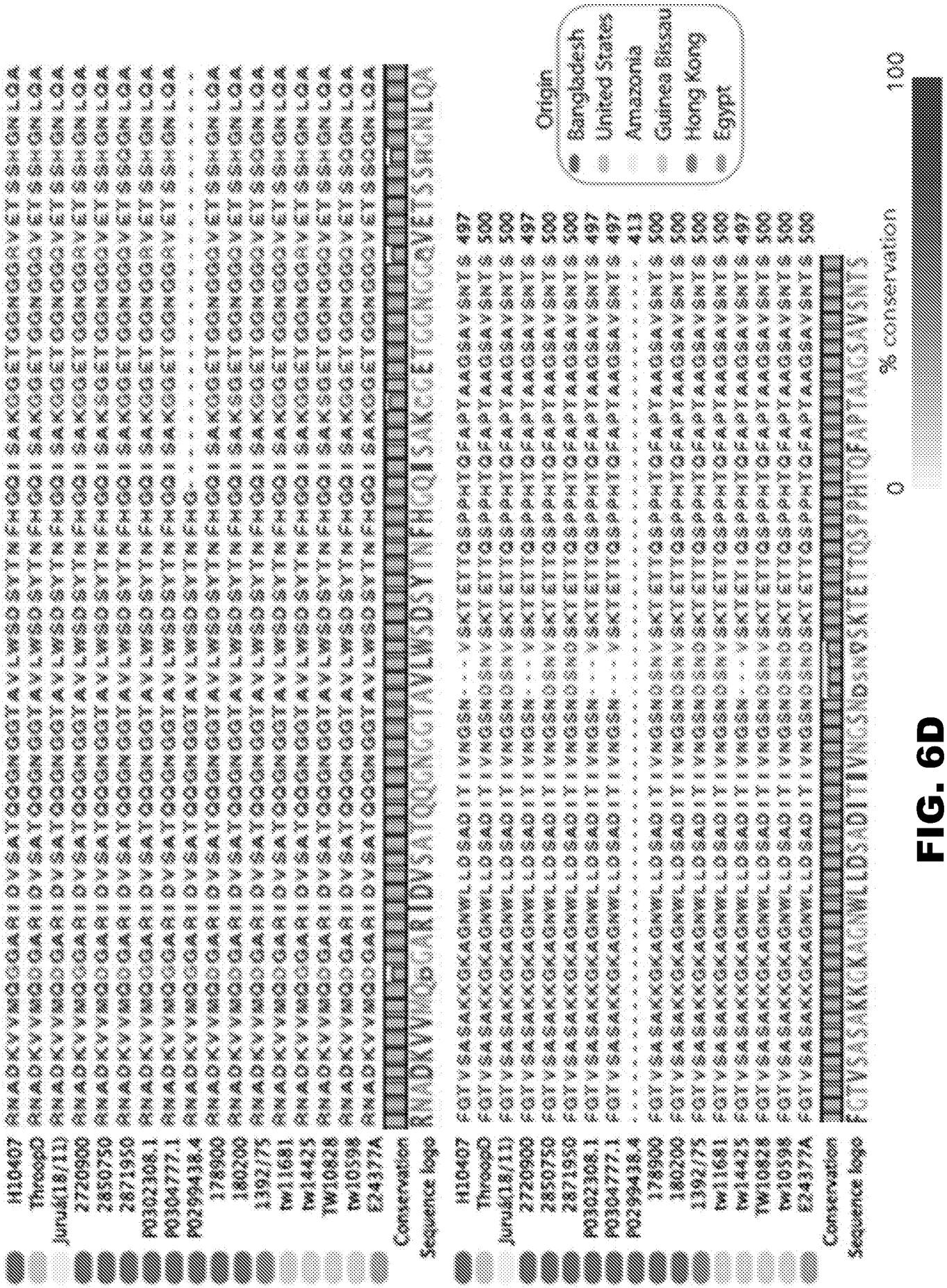

H10407
ThrocpD
Jun4(18/111)
2720900
2850750
2871950
P03023081
P03047771
P02994184
178900
180200
1392/75
tw11681
tw14425
TW10828
tw10598
E24377A
Conservation
Sequence logo
Origin
Bangladesh
United States
Amazonia
Guinea Bissau
Hong Kong
Egypt
0
% conservation
100

FIG. 6D

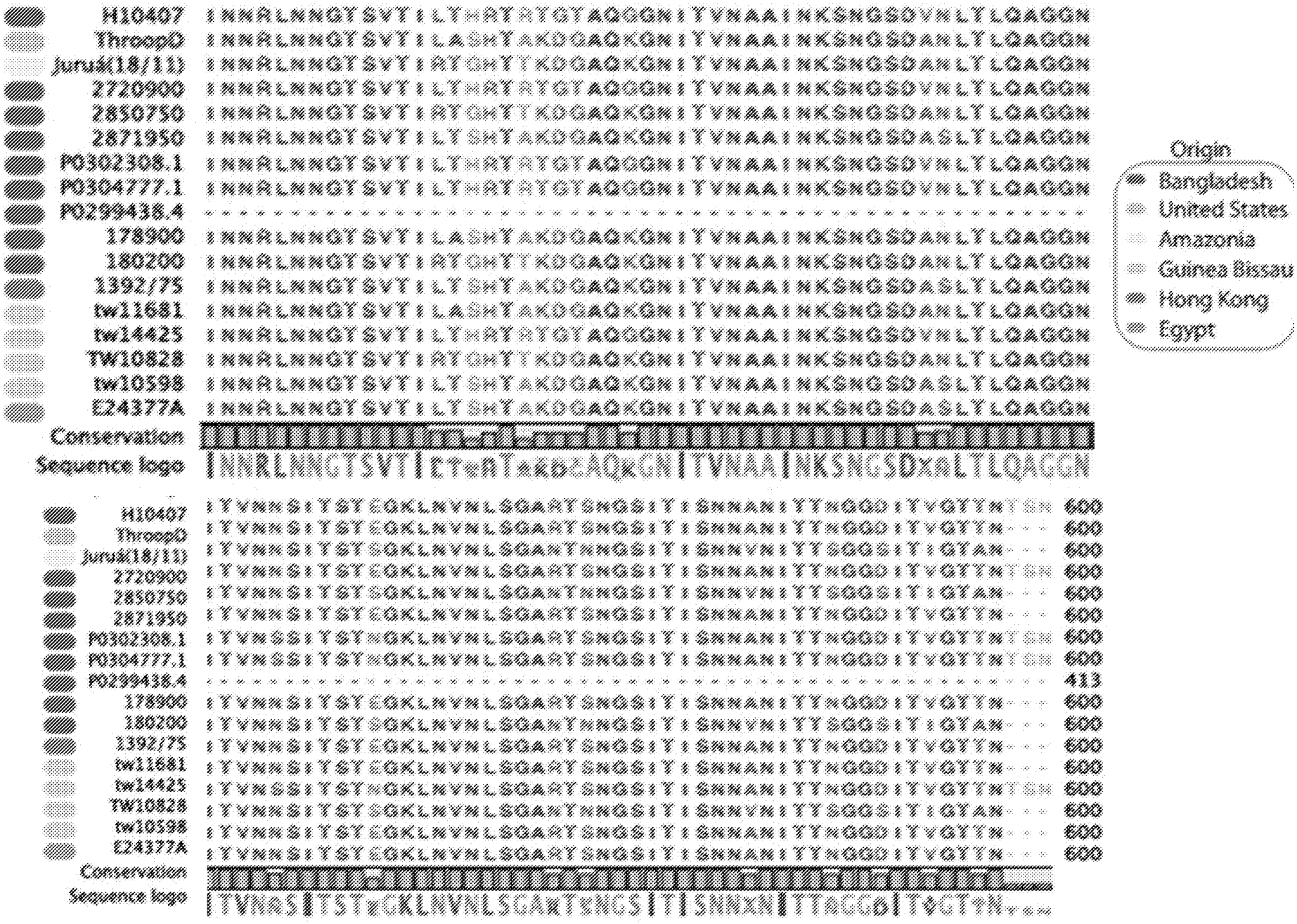

H10407
ThroopO
Juruá(18/11)
2720900
2850750
2871950
P0302308.1
P0304777.1
P0299438.4
178900
180200
1392/75
tw11681
tw14425
TW10828
tw10598
E24377A
Conservation
Sequence logo
Origin
Bangladesh
United States
Amazonia
Guinea Bissau
Hong Kong
Egypt
0
% conservation
100

FIG. 6E

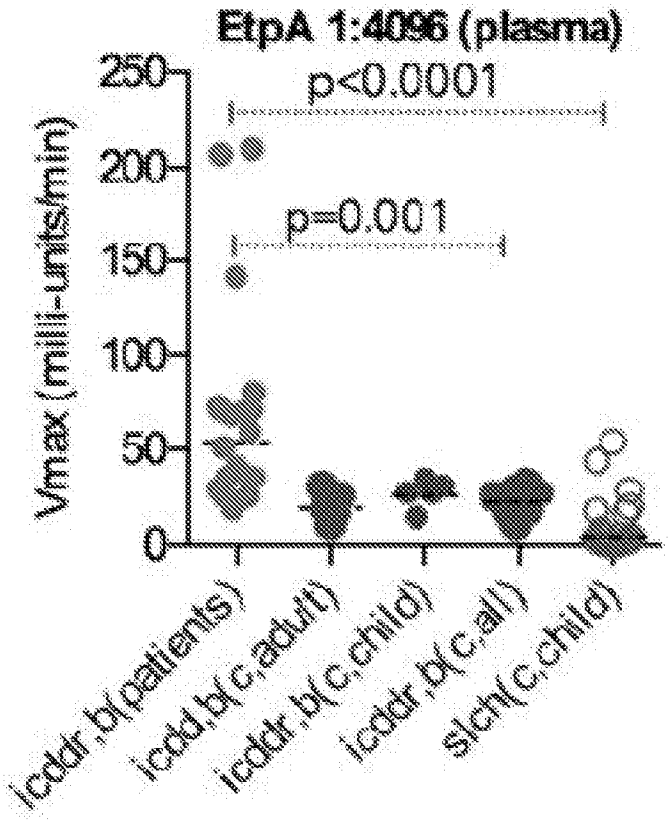

EtpA 1:4096 (plasma)
p<0.0001
p=0.001
Vmax (milli-units/min)
250
200
150
100
50
0
icddr,b(patients)
icdd,b(c,adult)
icddr,b(c,child)
icddr,b(c,all)
sickn(c,child)

FIG. 7A

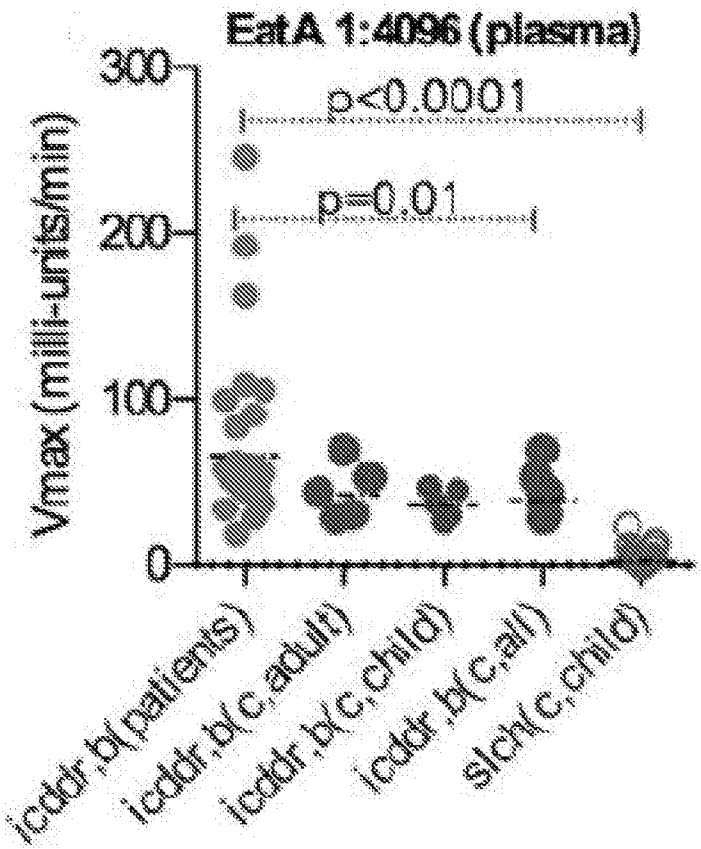

EatA 1:4096 (plasma)
p<0.0001
p=0.01
Vmax (milli-units/min)
300
200
100
0
icddr,b(patients)
icddr,b(c,adult)
icddr,b(c,child)
icddr,b(c,all)
sickn(c,child)

FIG. 7B

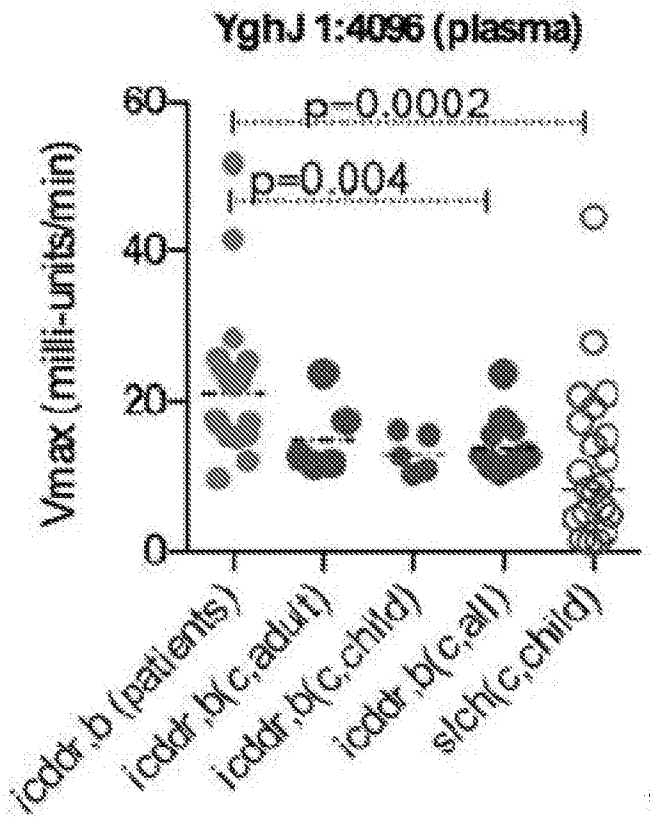

YghJ 1:4096 (plasma)
Vmax (milli-units/min)
60
40
20
0
p=0.0002
p=0.004
icddr,b (patients)
icddr,b(c,adult)
icddr,b(c,child)
icddr,b(c,all)
sjch(c,child)

FIG. 7C

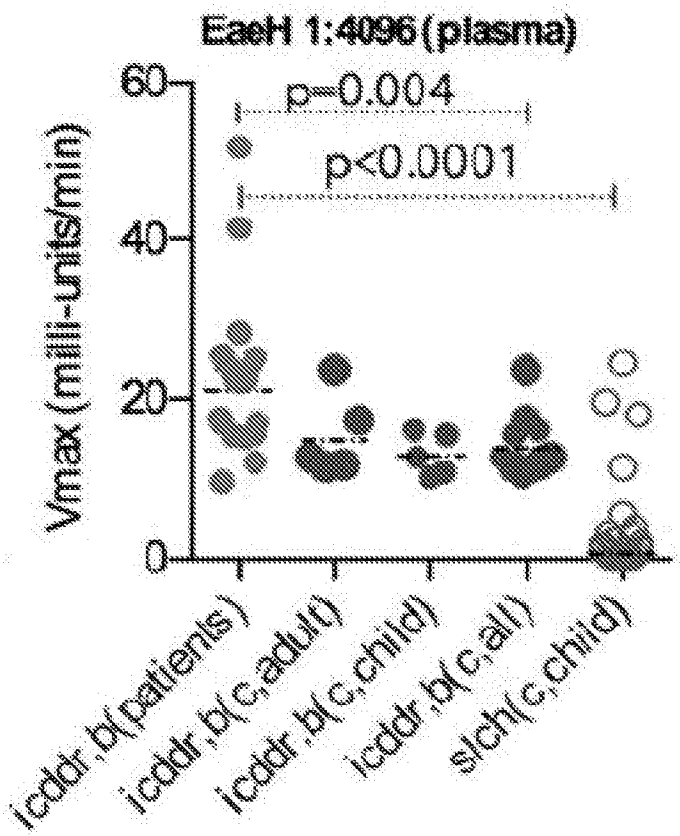

EaeH 1:4096 (plasma)
Vmax (milli-units/min)
60
40
20
0
p=0.004
p<0.0001
icddr,b(patients)
icddr,b(c,adult)
icddr,b(c,child)
icddr,b(c,all)
sjch(c,child)

FIG. 7D

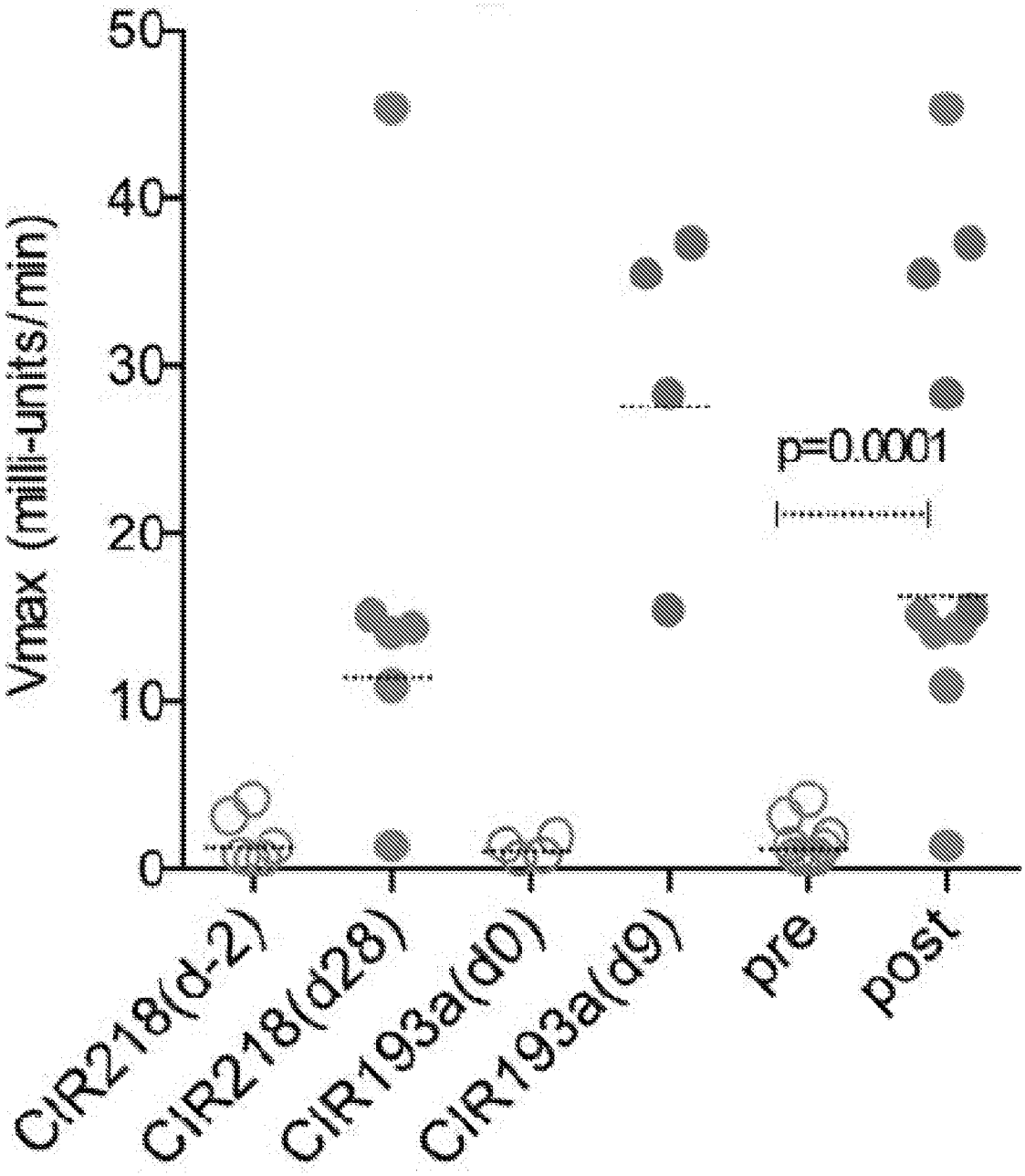

Vmax (milli-units/min)
50
40
30
20
10
0
p=0.0001
CIR218(d-2)
CIR218(d28)
CIR193a(d0)
CIR193a(d9)
pre
post

FIG. 8

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2016/013377

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - A61K 39/108 (2016.01)
CPC - A61K 39/0258 (2016.02)
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC(8) - A61K 39/02, 39/108; A61P 31/04; C07K 14/245 (2016.01)
CPC - A61K 2039/70, 39/02, 39/0258; C07K 14/245 (2016.02)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
USPC - 424/169.1, 184.1, 241.1, 257.1 (keyword delimited)

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
PatBase, Google Patents, PubMed.

Search terms used: ETEC vaccine EtpA EatA

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	KUMAR et al. "EatA, an immunogenic protective antigen of enterotoxigenic Escherichia coli, degrades intestinal mucin," Infect Immun. 2014 11 November 2013 (11.11.2013), Vol. 82, Pgs. 500-8. entire document	1-8
Y	ROY et al. "The EtpA exoprotein of enterotoxigenic Escherichia coli promotes intestinal colonization and is a protective antigen in an experimental model of murine infection," Infect Immun. 19 February 2008 (19.02.2008), Vol. 76, Pgs. 2106-12. entire document	1-8, 19-22
Y	PATEL et al. "Identification and molecular characterization of EatA, an autotransporter protein of enterotoxigenic Escherichia coli," Infect Immun. March 2004, Vol. 72, Pgs. 1786-94. entire document	4
Y	US 2007/0128183 A1 (MEINKE et al) 07 July 2007 (07.07.2007) entire document	6-8
Y	US 2011/0206694 A1 (FLECKENSTEIN) 25 August 2011 (25.08.2011) entire document	19-22
P, X	LUO et al. "Conservation and immunogenicity of novel antigens in diverse isolates of enterotoxigenic Escherichia coli," PLoS Negl Trop Dis. 28 January 2015 (28.01.2015), Vol. 9: e0003446, Pgs. 1-19. entire document	1-8

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
16 March 2016	04 APR 2016

Name and mailing address of the ISA/	Authorized officer
Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, VA 22313-1450 Facsimile No. 571-273-8300	Blaine R. Copenheaver PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2016/013377

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:

   a. ☒ forming part of the international application as filed:

      ☒ in the form of an Annex C/ST.25 text file.

      ☐ on paper or in the form of an image file.

   b. ☐ furnished together with the international application under PCT Rule 13*ter*.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.

   c. ☐ furnished subsequent to the international filing date for the purposes of international search only:

      ☐ in the form of an Annex C/ST.25 text file (Rule 13*ter*.1(a)).

      ☐ on paper or in the form of an image file (Rule 13*ter*.1(b) and Administrative Instructions, Section 713).

2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

SEQ ID NOs:1-4 were searched.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2016/013377

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☒ Claims Nos.: 11-18, 23-26
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.