PNAS
www.pnas.org

Supplementary Information for

CEACAMs serve as toxin-stimulated receptors for enterotoxigenic *E. coli*

Alaullah Sheikh[1], Brunda Tumala[1], Tim J. Vickers[1], David Alvarado[2], Matthew A. Ciorba[2], Taufiqur Rhaman Bhuiyan[4], Firdausi Qadri[4], Bernhard B. Singer[5], and James M. Fleckenstein[1,3,6*]

corresponding
James M. Fleckenstein
Email: jflecken@wustl.edu

This PDF file includes:

Figures S1 to S7
Tables S1 to S7
SI References

www.pnas.org/cgi/doi/10.1073/pnas.2012480117

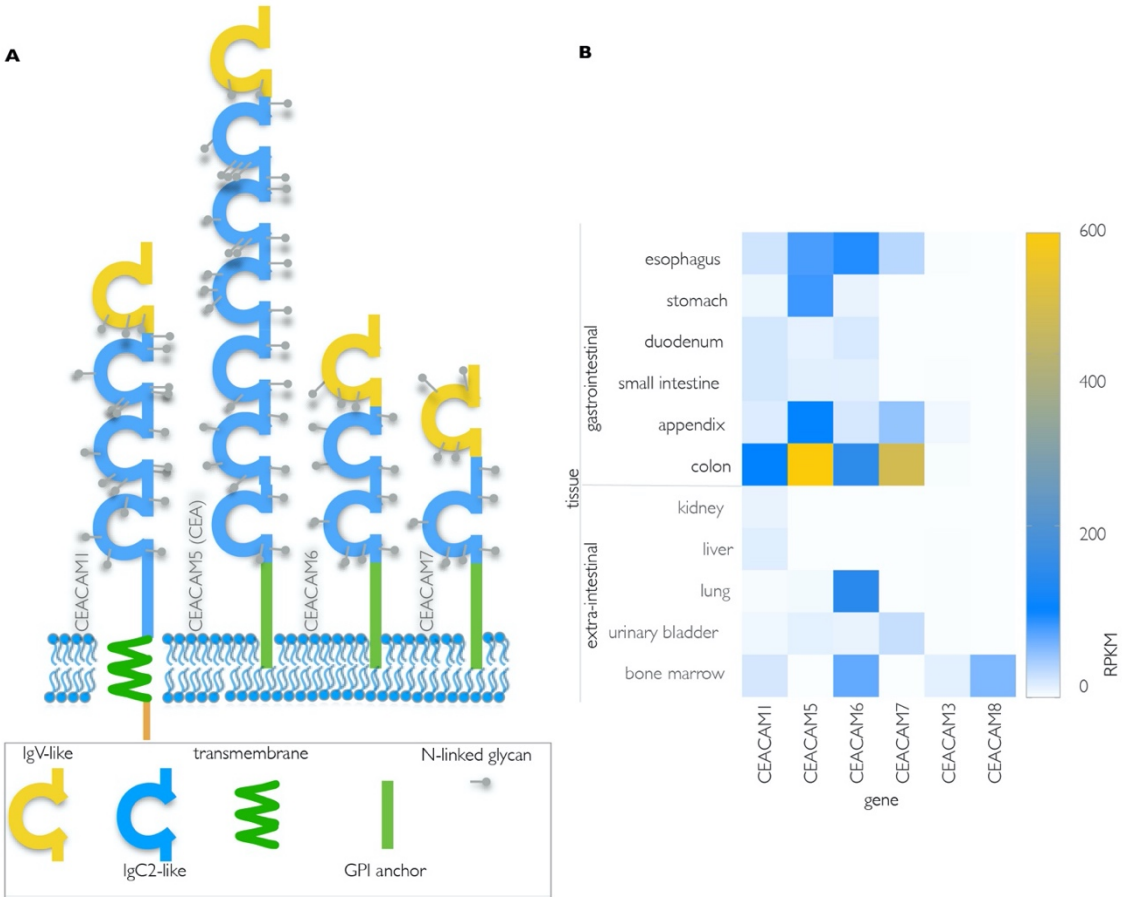

Fig. S1. structure and expression of individual CEACAMs. a. Structures of CEACAM molecules with significant expression in the gastrointestinal tract (adapted from(1)). IgV-like (yellow, amino terminus) = immunoglobulin variable-like domain (cd05741); IgC2-like (blue) immunoglobulin constant region like domain (pfam13895). **b**. tissue expression selected CEACAMs based on data from NCBI gene(2) for CEACAM1 (https://www.ncbi.nlm.nih.gov/gene/634), CEACAM3 (https://www.ncbi.nlm.nih.gov/gene/1084), CEACAM5 (https://www.ncbi.nlm.nih.gov/gene/1048), CEACAM6 (https://www.ncbi.nlm.nih.gov/gene/4680), CEACAM7 (https://www.ncbi.nlm.nih.gov/gene/1087), CEACAM8 (https://www.ncbi.nlm.nih.gov/gene/1088).

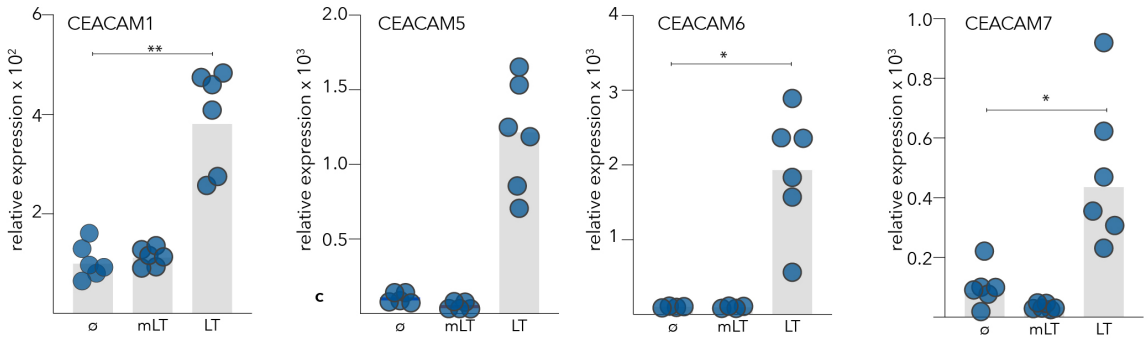

Fig. S2. heat-labile toxin induces transcription of genes encoding gastrointestinal CEACAMs. Shown are RT-PCR data demonstrating induction of CEACAM gene transcription following exposure of Caco-2 cells to heat-labile toxin (LT) relative to cells treated with inactive E112K mutant LT (mLT), or untreated cells (ø). Data include total of 5-6 technical replicates combined from two experiments. Bars represent geographic means. (*=<0.05, and **<0.01 by nonparametric ANOVA comparisons (Kruskal-Wallis).

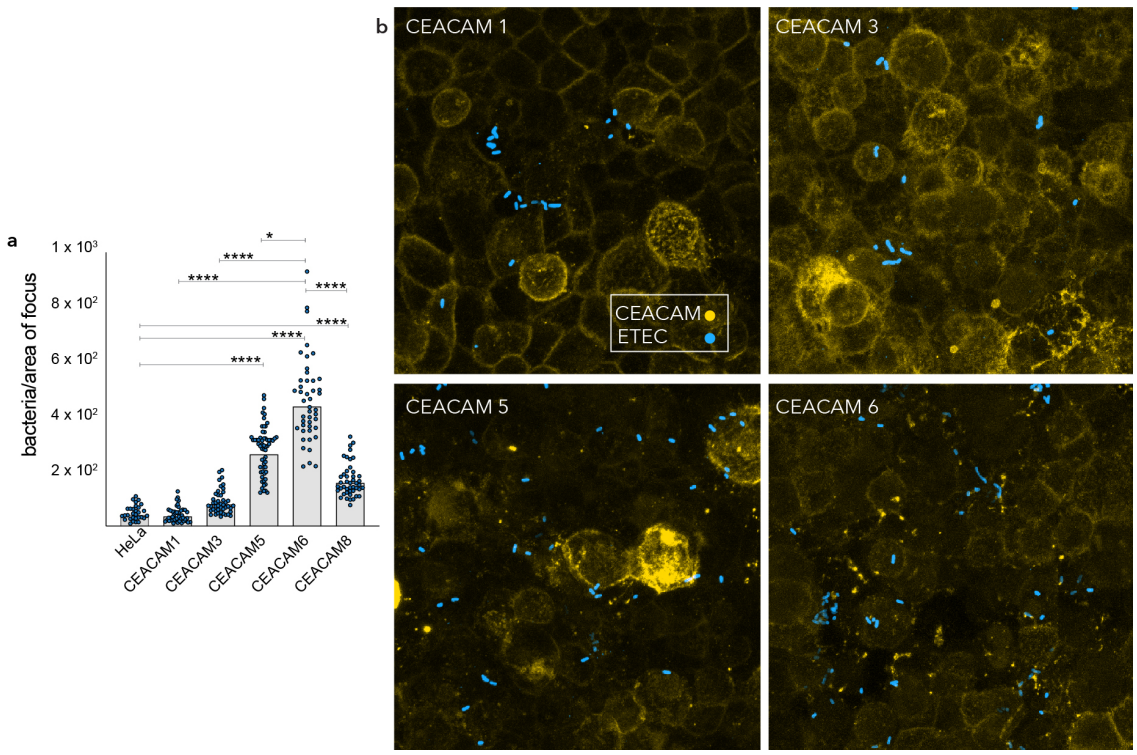

Fig S3. Preferential binding of ETEC to select CEACAMs.

a. ETEC H10407 adhesion to HeLa cells or HeLa cells expressing recombinant human CEACAMs. Data represent combined observations from n=3 experimental replicates. P values in a, b determined by ANOVA Kruskal-Wallis comparison of nonparametric data: **** <0.0001, *<0.05. Bars in a represent geographic means.

b. images depict representative CEACAM expression (pseudo-colored yellow) detected with specific monoclonal antibodies directed against each molecule followed by goat anti-mouse AlexaFluor 594 conjugate, binding of bacteria (detected with anti-O78 rabbit polyclonal antibody, and goat anti-rabbit AlexaFluor 488 conjugate, pseudo-colored blue).

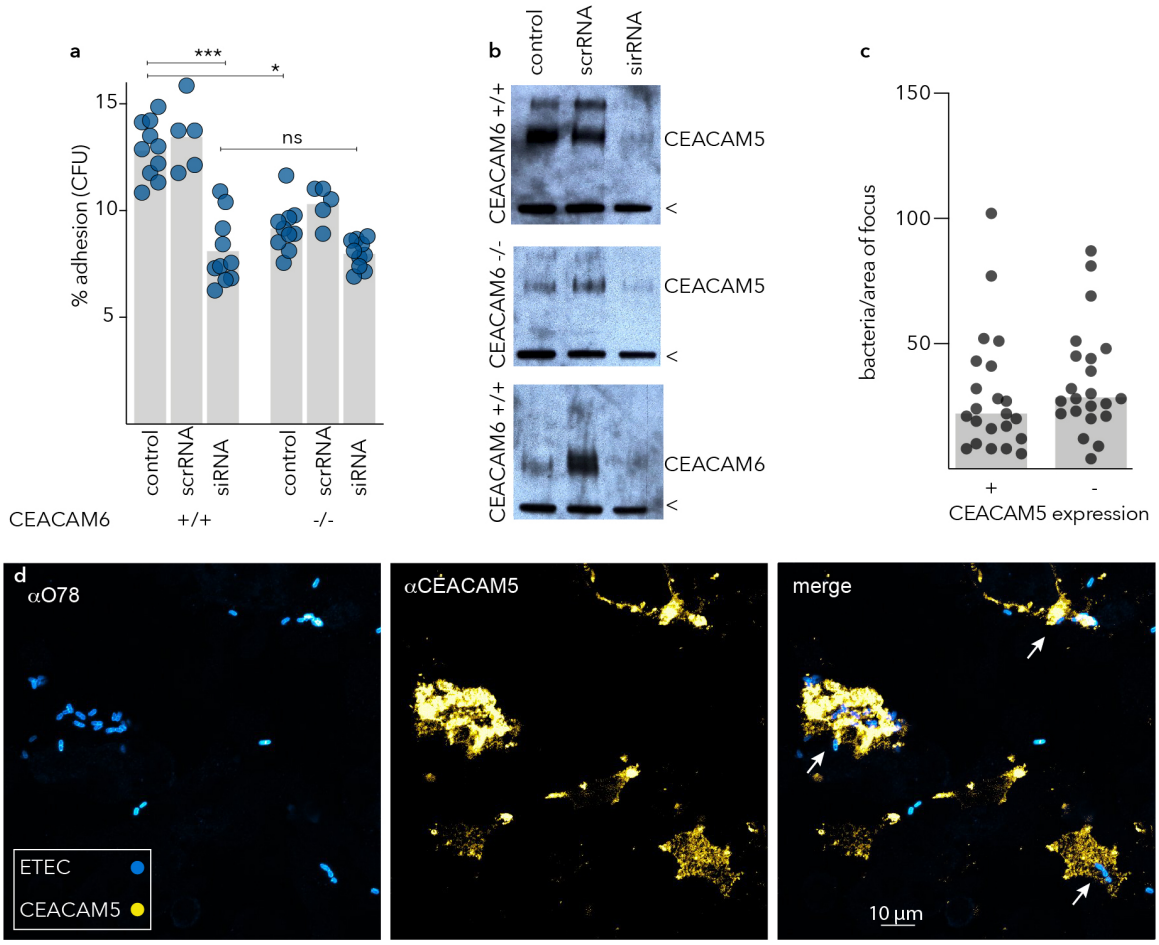

Fig S4. ETEC interactions with CEACAM5 expressed by Caco-2 cells. **a.,b** . impact of CEACAM5 siRNA knockdown on bacterial adhesion (**a**) and CEACAM5 and CEACAM6 expression in Caco-2 cells (**b**). (*** p=0.0001, *<0.05 by ANOVA, Kruskal-Wallis) Bars represent geographic means. Arrows (<) indicate tubulin as a loading control. Control, and scrRNA in a and b represent Lipofectamine transfection reagent alone and transfection with negative control siRNA, respectively. **c**. distribution of wild type ETEC (H10407) bacteria in areas exhibiting high or low CEACAM5 expression by confocal microscopy d. confocal images of ETEC, detected with anti-O78 antibody, in areas of robust CEACAM5 expression on Caco-2 cells.

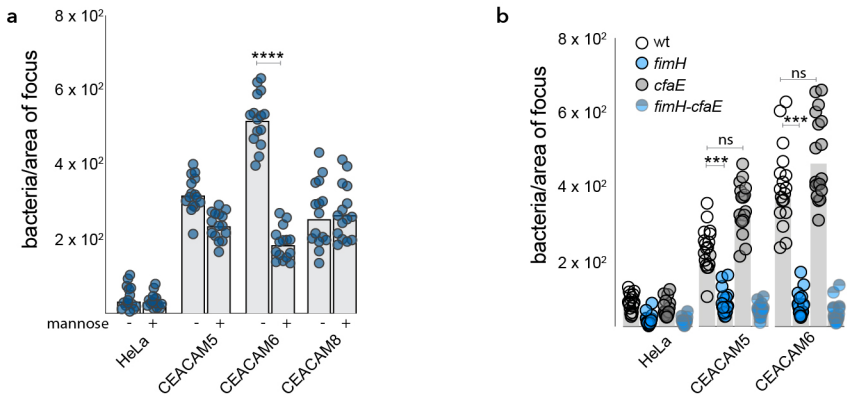

Fig S5. ETEC interaction with CEACAM6 depends on the mannose sensitive adhesin FimH. **a**. CEACAM6 interaction is inhibited by mannose**. b**. H10407-CEACAM interaction requires functional type 1 fimbriae, but does not rely on canonical CFA/I fimbriae. Graph depicts binding of wild type, *fimH* mutant, *cfaE* mutant, and *fimH/cfaE* mutant strains to HeLa cells expressing CEACAMs 5, or 6 compared to control HeLa cells. *** represents p≤0.001, and **** <0.0001 by ANOVA Kruskal-Wallis comparison of nonparametric data. Bars represent geographic means.

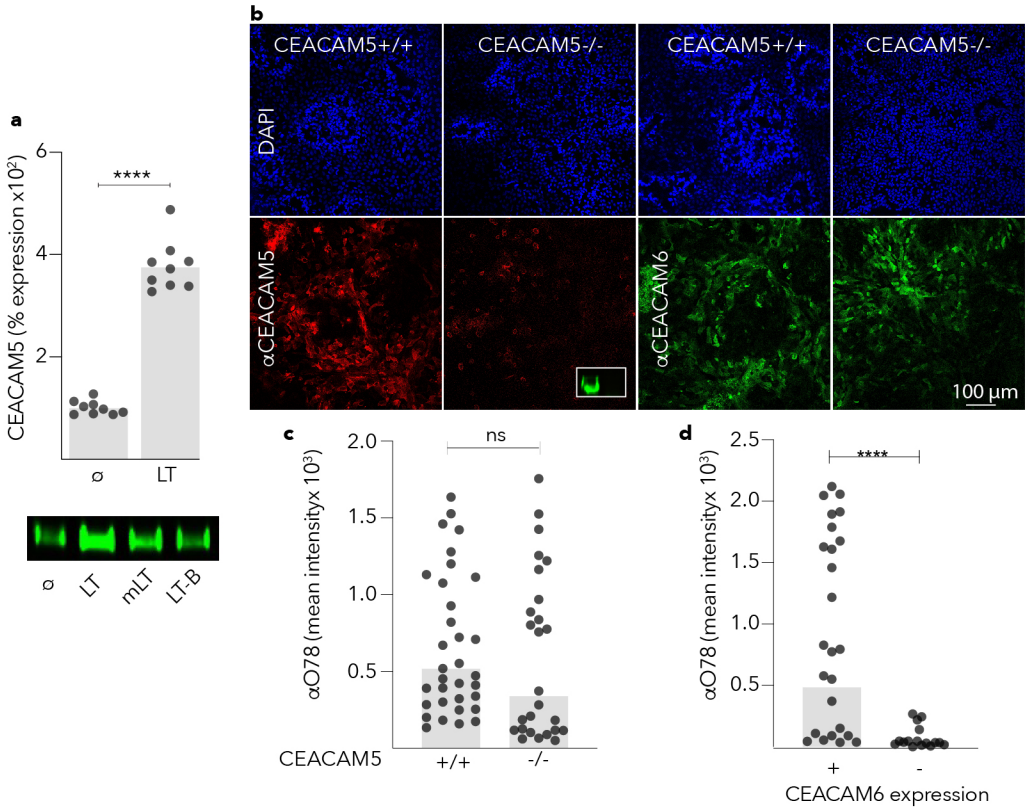

Fig S6. ETEC interaction with CEACAM5 expressed on small intestinal epithelia. **a**. RT-PCR data demonstrating increased expression of CEACAM5 in small intestinal enteroid cells following heat-labile toxin treatment (LT) relative to untreated cells (ø). a-CEACAM5 immunoblot below shows relative expression of CEACAM5 following treatment with LT, mLT (E112K), or LT-B subunit alone. **b**. CEACAM expression in CRISPR/Cas9-generated CEACAM5-/- ileal enteroid line. Confocal images showing CEACAM5 (red) and CEACAM6 (green) expression in wild type and CEACAM5 -/- knockout cells. Inset immunoblot also demonstrates lack of CEACAM5 production in CEACAM5 -/- knockout cells relative to parental cell line. **c**. ETEC adhesion, detected by anti-O78 signal, to wild type and CEACA5 -/- knockout cells. Bars represent geometric means, ns= not significant, **d**. CEACAM6 directs bacterial association on CEACAM5 -/- cells. Bars represent geometric means and **** represents $p<0.00001$. Data were generated from at least 3 individual experiments, and p values were calculated using nonparametric Mann-Whitney test. Bars throughout represent geographic means.

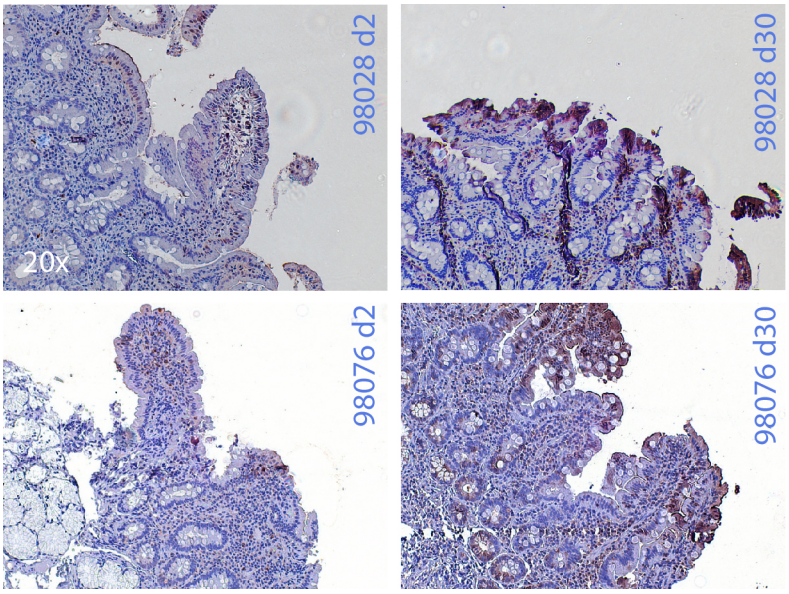

Fig S7. serial small intestinal biopsies from two ETEC-infected patients obtained on day 2 of hospitalization at icddrb and day 30 post presentation.

Table S1. gastrointestinal CEACAMS

Antigen	other designation(s)	ncbi gene number	domains	membrane anchor	tissue expression	references
CEACAM1	BGP, CD66	634	cd05774 Ig_CEACAM_D1 smart00410 Ig-like	TM	colon, small intestine, marrow/ WBCs	(1, 2)
CEACAM5	CEA, CD66e	1048	cd05774 Ig_CEACAM_D1 smart00410 Ig-like cd05740 Ig_CEACAM_D4	GPI	colon, small intestine	(2)
CEACAM6	NCA, CD66c	4680	cd05774 Ig_CEACAM_D1 smart00410 Ig-like	GPI	colon, lung, small intestine, marrow/WBCs, bladder	(2, 3)
CEACAM7		1087	cd05774 Ig_CEACAM_D1 smart00410 Ig-like	GPI	colon	(2)

TM=transmembrane; GPI=glycosylphosphotidylinositol; WBCs=white blood cells

Table S2. bacterial strains and plasmids

Bacterial strains and plasmids used in these studies		
designation	Description	Souce/Reference(s)
Strain		
H10407	Wild-type ETEC strain O78:H11;CFA/1;LT+/ST+;EtpA+	(4, 5)
jf570	*eltA*::FRT[a], antibiotic sensitive	(6)
jf2944	*fimH*::Km^R, *fimH* gene is replaced with kanamycin resistant cassette	27
jf1862	*cfaE*::Km^R *cfaE* gene replaced with kanamycin resistant cassette	27
jf2945	*fimH*:: Cm^R, *cfaE*::Km^R *fimH-cfaE* double mutants; *fimH* gene is replaced with chloramphenicol resistant cassette and *cfaE* gene is replaced with kanamycin resistant cassette	27
Stbl3	*F⁻mcrB mrrhsdS20(r_B^-, m_B^-) recA13 supE44 ara-14 galK2 lacY1 proA2 rpsL20(Str^R) xyl-5 λ⁻leumtl-1*	ThermoFisher
plasmid		
pRC202454	CEACAM6 expression plasmid; CMV promotor; Neo^R/Km^R	OriGene
pcDNA3.4_mIgG2-mMxr8a	expression plasmid containing IL-2 signal peptide and C-terminal 6His tag Amp^R/Neo^R/	Diamond lab (7)
pcDNA3.4_CEACAM6-His	858 bp CEACAM6 sequence amplified from pRC202454 with primers jf030218.1/.3 cloned by Gibson assembly into pcDNA3.4_mIg2-mMxr8a digested with *EcoRI/AgeI*	this study//addgene 149337
pFCIV	lentiviral packaging vector YFP; Amp^R, $Bleo^R$; ColE1ori;	hope center for neurological disorders https://hopecenter.wustl.edu/?page_id=7068 PMC5008082
pFCIV-CEACAM6	1035bp CEACAM6 insert amplified from pRC202454 using primers jf021518.1/.2 and cloned into BamHI and AgeI sites of pFCIV by Gibson assembly	this study/addgene 149303
psPAX2	2nd generation lentiviral packaging plasmid (Addgene#12260)	Addgene
pMD2.G	VSV-G envelope expressing plasmid for lentiviral packaging (Addgene #12259)	Addgene
plentiCRISPRv2[b]	Cas9 guide RNA plasmid addgene 52961, Amp^R	(8)
pCEACAM5gRNA	CEACAM5 guide RNA introduced into plentiCRISPRv2 at *BsmBI* sites by Gibson assembly	this study
pfimHLD-his	Expression of FimH lectin domain	(9)
pfimHLD:Q133K-his	Expression of Q133K mutated FimH lectin domain	(9)

[a]FRT=FLP recombinase recognition target.
[b]lentiCRISPR v2 was a gift from Feng Zhang (Addgene plasmid # 52961 ; http://n2t.net/addgene:52961 ; RRID:Addgene_52961).

Table S3. cell lines

Cell line	Description	Source
Caco-2	Colonic epithelial cells	ATCC HTB-37
G2 (CEACAM6 -/-)	CRISPR-Cas9-generated CEACAM6-negative Caco-2 cells	This study
G2-CEACAM6	CEACAM6 trans-complemented G2 cells	This study
HeLa	Cervical epithelial cells	ATCC CCL-2
HeLa-GFP-CEACAM6	CEACAM6-expressing GFP positive HeLa cells	This study
HeLa_CEACAM1	Stably transfected HeLa cells expressing CEACAM1	(10)
HeLa_CEACAM3	Stably transfected HeLa cells expressing CEACAM3	(10)
HeLa_CEACAM5	Stably transfected HeLa cells expressing CEACAM5	(10)
HeLa_CEACAM6	Stably transfected HeLa cells expressing CEACAM6	(10)
HeLa_CEACAM8	Stably transfected HeLa cells expressing CEACAM8	(10)
Expi293F	Human embryonic Kidney cells	Thermo Fisher A14527
HEK293T	Derivative of human embryonic kidney 293 cells suitable for high titer lentivirus production, ATCC CRL-3216	ATCC
CEACAM5-/-	CRISPR-Cas9-generated CEACAM5-negative human small intestinal enteroid cell line	This study
Hu135D	human small intestinal (ileal) enteroid, blood group A	(11)
Hu235D	human small intestinal (ileal) enteroid, blood group O	(11)
Hu248D	human small intestinal (ileal) enteroid, blood group B	(11)

Table S4. primers

primers used in these studies	
cloning primers	
primer name	sequence 5′-3′
f021518.1	tgggctgcaggtcgactctagaggatccATGGGACCCCCCTCAGCCCCTCCCTGCA
jf021518.2	ggggagggagaggggcggatcaccggtTTATATCAGAGCCACCCTGGCCA
jf030218.1	TAAGTCTTGCACTTGTCACGAATTCGATATCGAAGCTCACTATTGAATCCACG
jf030218.3	CTCAATGGTGATGGTGATGATGACCGGTTGCAGAGACTGTGATCATCGTG

RT-PCR primers					
gene name	Entrez gene ID[1]	Primer Bank ID[2]	forward primer	reverse primer	amplicon (bp)
CEACAM1	634	329112546c1	TGCTCTGATAGCAGTAGCCCT	TGCCGGTCTTCCCGAAATG	53
CEACAM5	1048	98986444c2	TCTTGGCTGATTGATGGGAAC	CACTGGCTGAGTTATTGGCCT	112
CEACAM6	4680	PMID: 18223215	CACCGTCGGCATCACGA	GAAGAATTCAGGGTCTGGTCCA	113
CEACAM7	1087	257153399c3	CTTCAATCCGGTGGAGAACAA	CGCTGAGTAGAACGAGGGTC	155
GAPDH	2597	PMID: 28885617	CTTCAATCCGGTGGAGAACAA	CGCTGAGTAGAACGAGGGTC	101

[1]https://www.ncbi.nlm.nih.gov/gene/

[2] https://pga.mgh.harvard.edu/primerbank/index.html

Table S5. siRNA

siRNAs			
gene symbol	siRNA designation	chromosome location	source
CEACAM6	s9283	Chr.19: 41755421 - 41772211 (Build GRCh38)	ThermoFisher (4392420)
CEACAM5	s2887	Chr.19: 41708585 - 41730433 (Build GRCh38)	ThermoFisher (4392420)
-	negative control #2	none	ThermoFisher (4390846)

Table S6. antibodies

antibodies			
CEACAM-specific antibodies			
Antibody	Specificity	description	source/reference
B3-17	CEACAM1	a-human CEACAM1, mouse monoclonal	(12)
308/3-3	CEACAM3,5	a-human CEACAM3,5, mouse monoclonal	Sigma, MABT326
A0115	CEACAM/CEA	a-human Carcinoembryonic Antigen (CEA), polyclonal Rabbit	DAKO, A0115
M1545	CD66e (CEA)	a-human CD66e mouse monocolonal	PeliCluster
5C8C4	CEACAM5	a-human CEACAM5, mouse monoclonal	(12)
1H7-H4B	CEACAM6	a-human CEACAM6, mouse monoclonal	(12)
6/40c	CEACAM8	a-human CEACAM8, mouse monoclonal	(13)
ETEC H10407-specific (serotype O78,H11)			
a-O78	O78 polysaccharide	a O78 polysaccharide rabbit polyclonal, cross-absorbed	Penn State *E. coli* Reference Center[b]
secondary antibodies			
Antibody	Specificity	description	source/reference
AB_2338586[a]	mouse IgG (H+L)	Goat- a mouse Fab, biotin-SP (long spacer), affinity purified	Jackson ImmunoResearch (115-067-003)
A-11070	rabbit IgG (H+L)	F(ab')2-Goat a-Rabbit IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 488	ThermoFisher
A-11032	mouse IgG (H+L)	Goat a-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 594	ThermoFisher
A-11072	rabbit IgG (H+L)	F(ab')2-Goat a -Rabbit IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 594	ThermoFisher
A-11029	mouse IgG (H+L)	Goat a -Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 488	ThermoFisher
[a]Resource Identification Portal at https://scicrunch.org/resources [b]https://foodscience.psu.edu/research/centers/ecoli			

Table S7. original image files

links to original image files*		
image name	description	DOI†
figure 1c	immunoblots	10.6084/m9.figshare.12137544
figure 2a	immunoblots of CEACAM6 and CEACAM5 in Caco-2 and G2 (CRISPR-Cas9 engineered CEACAM6-/- cells)	10.6084/m9.figshare.12141033
figure 3a-c	immunoblots of FimH-CEACAM6 molecular interactions	10.6084/m9.figshare.12141228
figure 4a	TEM, CEACAM6 enteroid immunogold	10.6084/m9.figshare.12137658
figure 4b	TEM, CEACAM6 enteroid immunogold	10.6084/m9.figshare.12137652
figure 4d	confocal images of ETEC on polarized small intestinal enteroids in context of CEACAM6 expression	10.6084/m9.figshare.12137655
figure 4f.	immunoblots of small intestinal enteroids treated with LT vs control untreated cells. three technical replicates.	10.6084/m9.figshare.12141138
figure 4h	immunoblot	10.6084/m9.figshare.12137661
figure 4i	immunohistochemistry for CEACAM6 in infected intestinal sections	10.6084/m9.figshare.12138297
supplemental figure 4b	immunoblots of CEACAM5,6 expression following CEACAM5 siRNA treatment	10.6084/m9.figshare.12142092
supplemental figure 5a	immunoblot of CEACAM5 expression following LT induction of small intestinal enteroids	10.6084/m9.figshare.12140994
*deposited at figshare.com † https://doi.org/		

SI references

1. S. D. Gray-Owen, R. S. Blumberg, CEACAM1: contact-dependent control of immunity. *Nat Rev Immunol* **6**, 433-446 (2006).
2. L. Fagerberg *et al.*, Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Mol Cell Proteomics* **13**, 397-406 (2014).
3. M. Sadarangani, A. J. Pollard, S. D. Gray-Owen, Opa proteins and CEACAMs: pathways of immune engagement for pathogenic Neisseria. *FEMS Microbiol Rev* **35**, 498-514 (2011).
4. D. G. Evans, R. P. Silver, D. J. Evans, Jr., D. G. Chase, S. L. Gorbach, Plasmid-controlled colonization factor associated with virulence in *Escherichia coli* enterotoxigenic for humans. *Infect Immun* **12**, 656-667 (1975).
5. D. J. Evans, Jr., D. G. Evans, Three characteristics associated with enterotoxigenic Escherichia coli isolated from man. *Infect Immun* **8**, 322-328 (1973).
6. F. C. Dorsey, J. F. Fischer, J. M. Fleckenstein, Directed delivery of heat-labile enterotoxin by enterotoxigenic Escherichia coli. *Cell Microbiol* **8**, 1516-1527 (2006).
7. R. Zhang *et al.*, Mxra8 is a receptor for multiple arthritogenic alphaviruses. *Nature* **557**, 570-574 (2018).
8. N. E. Sanjana, O. Shalem, F. Zhang, Improved vectors and genome-wide libraries for CRISPR screening. *Nat Methods* **11**, 783-784 (2014).
9. A. Sheikh *et al.*, Highly conserved type 1 pili promote enterotoxigenic E. coli pathogen-host interactions. *PLoS Negl Trop Dis* **11**, e0005586 (2017).
10. B. B. Singer *et al.*, Deregulation of the CEACAM expression pattern causes undifferentiated cell growth in human lung adenocarcinoma cells. *PLoS One* **5**, e8747 (2010).
11. P. Kumar *et al.*, Enterotoxigenic Escherichia coli-blood group A interactions intensify diarrheal severity. *J Clin Invest* **128**, 3298-3311 (2018).
12. H. T. Muturi *et al.*, Tumor and endothelial cell-derived microvesicles carry distinct CEACAMs and influence T-cell behavior. *PLoS One* **8**, e74654 (2013).
13. B. B. Singer *et al.*, Soluble CEACAM8 interacts with CEACAM1 inhibiting TLR2-triggered immune responses. *PLoS One* **9**, e94106 (2014).