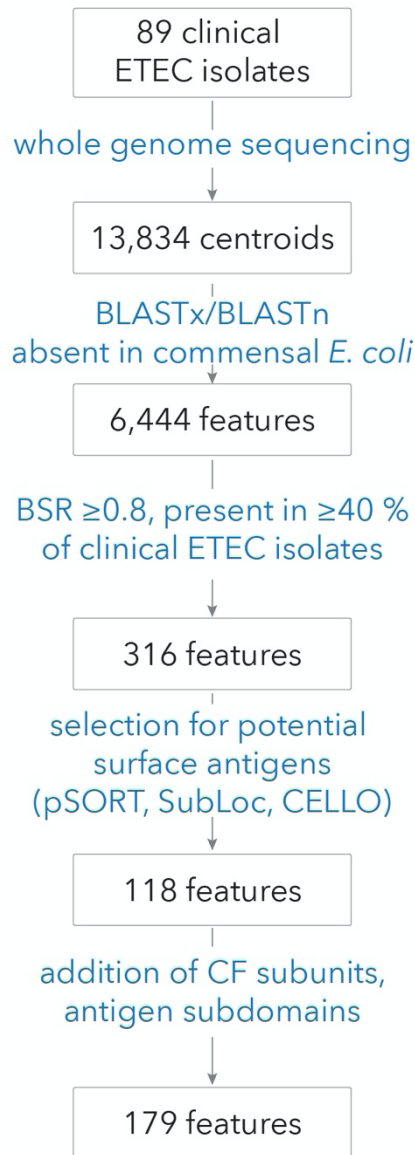


- 1 **supplementary material**
- 2 **supplementary figures**



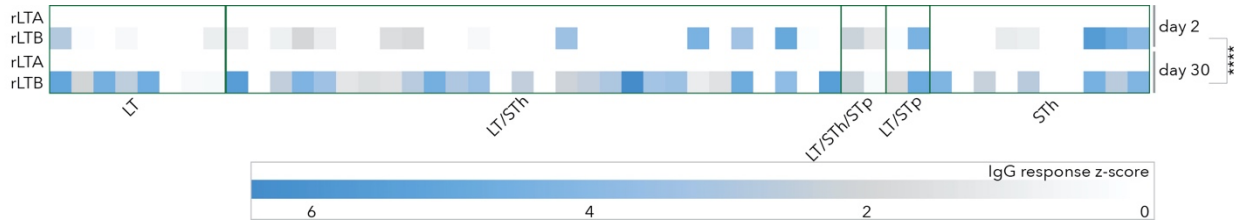
3

4 **supplementary figure 1.** reverse vaccinology strategy used to identify potential surface
5 antigens conserved within the ETEC pathovar. Whole genome sequencing was used to
6 identify candidate genes (centroids), subtracting elements common to commensal *E. coli*
7 strains. BLAST Score Ratio analysis (1) was then used to define unique features found in at
8 least 40% of ETEC. The resulting 316 features were then analyzed for potential surface
9 expression using pSORT (2), SubLoc (3), and CELLO (4).

10

11

12

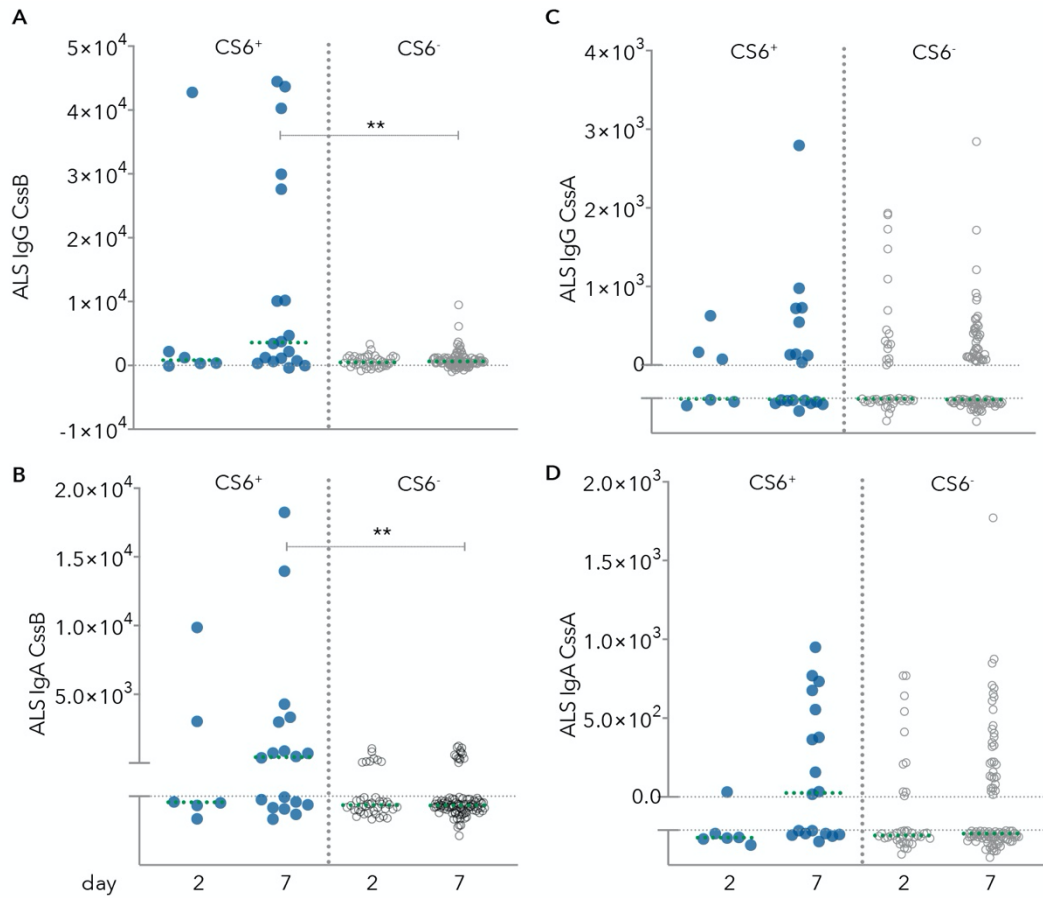


13

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

18

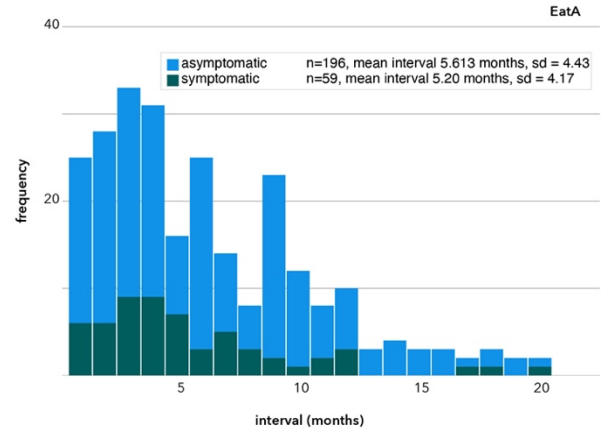
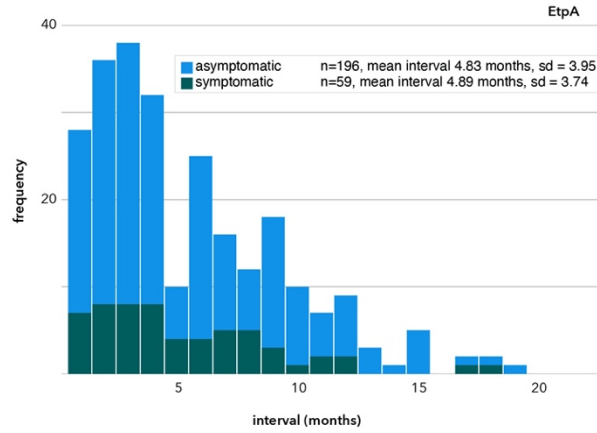
31



32

33 **supplementary figure 4.** ALS responses to CS6 subunits following ETEC infection. **(A)** IgG
34 response to CssB. **(B)** IgA response to CssB. ** $p < 0.01$, Kruskal-Wallis post-hoc analysis using
35 Dunn's test adjusted for multiple comparisons. **(C)** IgG response to CssA. **(D)** IgA response to
36 CssA.

37



38

39 **supplementary figure 5.** Interval between peak serum responses to EtpA (left), EatA (right)
 40 and identification of subsequent ETEC + samples in asymptomatic and symptomatic children.

41

42 **supplementary tables**43 **supplementary table 1**

supplementary table 1. primers used in strain interrogation				
gene	reference sequence GenBank	amplicon (bp)	Sequence (5'>3')	primer ID
<i>eatA</i>	AY163491.2	1943	ATGTGCTTTGGCAGGTAA	jf082213.1-F
			ATATCCAGTCAGCACCCACT	jf082213.2-R
<i>etpA</i>	AY920525.2	999	GGTTCAGGCAGTATCCAGAC	jf082213.3-F
			GGTGTAGCTGTCTGACCACA	jf082213.4-R
<i>eltB</i>	CBJ04425.1	273	ACGGCGTTACTATCCTCTC	jf092313.3-F
			TGGTCTCGGTCAGATATGTG	jf092313.4-R
<i>estP</i> (<i>sta1</i>)	CBJ04435.1	166	TCTTCCCCTCTTTAGTCAG	jf092313.5-F
			ACAGGCAGGATTACAACAAAG	jf092313.6-R
<i>estH</i> (<i>sta2</i>)	CBJ04483.1	64	TACAAGCAGGATTACAACAC	jf092313.7-F
			AGTGGTCCTGAAAGCATG	jf092313.8-R

44

45

46

47 **supplemental datasets**

48 **supplemental dataset 1**

49 Complete annotation of all array features, and primers used in amplification.

50 **supplemental dataset 2**

51 ALS IgG and ALS IgA array responses, and sample metadata.

52

53 **references**

- 54 1. Rasko DA, Myers GS, Ravel J. 2005. Visualization of comparative genomic analyses by
55 BLAST score ratio. *BMC Bioinformatics* 6:2.
- 56 2. Yu NY, Wagner JR, Laird MR, Melli G, Rey S, Lo R, Dao P, Sahinalp SC, Ester M, Foster
57 LJ, Brinkman FS. 2010. PSORTb 3.0: improved protein subcellular localization
58 prediction with refined localization subcategories and predictive capabilities for all
59 prokaryotes. *Bioinformatics* 26:1608-15.
- 60 3. Chen H, Huang N, Sun Z. 2006. SubLoc: a server/client suite for protein subcellular
61 location based on SOAP. *Bioinformatics* 22:376-7.
- 62 4. Yu CS, Lin CJ, Hwang JK. 2004. Predicting subcellular localization of proteins for
63 Gram-negative bacteria by support vector machines based on n-peptide
64 compositions. *Protein Sci* 13:1402-6.

65