Genome and functional characterization of colonization factor antigen I- and CS6-encoding heat-stable enterotoxin-only enterotoxigenic Escherichia coli reveals lineage and geographic variation

Hazen, Tracy H.

Nagaraj, Sushma

Sen, Sunil

Permala-Booth, Jasnehta

Canto, Felipe Del

Vidal, Roberto

Barry, Eileen M.

Bitoun, Jacob P.

Chen, Wilbur H.

Tennant, Sharon M.

Rasko, David A.

 Enterotoxigenic Escherichia coli (ETEC) is a significant cause of childhood diarrhea and is a leading cause of traveler's diarrhea. ETEC strains encoding the heat-stable enterotoxin (ST) are more often associated with childhood diarrhea than ETEC strains that encode only the heat-labile enterotoxin (LT). Colonization factors (CFs) also have a demonstrated role in ETEC virulence, and two of the most prevalent CFs among ETEC that have caused diarrhea are colonization factor antigen I (CFA/I) and CS6. In the current report, we describe the genomes of 269 CS6- or CFA/ I-encoding ST-only ETEC isolates that were associated with human diarrhea. While the CS6 and CFA/I ETEC were identified in at least 13 different ETEC genomic lineages, a majority (85%; 229/269) were identified in only six lineages. Complete genome sequencing of selected isolates demonstrated that a conserved plasmid contributed to the dissemination of CFA/I whereas at least five distinct plasmids were involved in the dissemination of ST and/or CS6. Additionally, there were differences in gene content between CFA/I and CS6 ETEC

at the phylogroup and lineage levels and in association with their geographic location of isolation as well as lineage-related differences in ST production. Thus, we demonstrate that genomically diverse E. coli strains have acquired ST, as well as CFA/I or CS6, via one or more plasmids and that, in some cases, isolates of a particular lineage or geographic location have undergone additional modifications to their genome content. These findings will aid investigations of virulence and the development of improved diagnostics and vaccines against this important human diarrheal pathogen. IMPORTANCE Comparative genomics and functional characterization were used to analyze a global collection of CFA/I and CS6 ST-only ETEC isolates associated with human diarrhea, demonstrating differences in the genomic content of CFA/I and CS6 isolates related to CF type, lineage, and geographic location of isolation and also lineage-related differences in ST production. Complete genome sequencing of selected CFA/I and CS6 isolates enabled descriptions of a highly conserved ST-positive (ST) CFA/I plasmid and of at least five diverse ST and/or CS6 plasmids among the CS6 ETEC isolates. There is currently no approved vaccine for ST-only ETEC, or for any ETEC for that matter, and as such, the current report provides functional verification of ST and CF production and antimicrobial susceptibility testing and an in-depth genomic characterization of a collection of isolates that could serve as representatives of CFA/I- or CS6-encoding ST-only ETEC strains for future studies of ETEC pathogenesis, vaccine studies, and/or clinical trials.