Göteborg, 2016

Whole genome sequencing of enterotoxigenic *Escherichia coli* (ETEC)

Identification of ETEC lineages and novel colonization factors

Akademisk avhandling

Som för avläggande av medicine doktorsexamen vid Sahlgrenska akademin vid Göteborgs universitet kommer att offentligen försvaras i hörsal
Ivan Östholm, Medicinaregatan 13, Göteborg

Fredagen den 27:e maj 2016, klockan 9:00

av **Astrid von Mentzer**

Fakultetsopponent:
Professor Ann-Beth Jonsson
Stockholms Universitet, Sverige

Avhandlingen baseras på följande delarbeten

I. **Identification of enterotoxigenic *Escherichia coli* (ETEC) clades with long-term global distribution**
von Mentzer A, Connor TR, Wieler LH, Semmler T, Iguchi A, Thomson NR, Rasko DA, Joffre E, Corander J, Pickard D, Wiklund G, Svennerholm A-M, Sjöling Å and Dougan G
Nature Genetics, *2014. 46:1321-26*.

II. **Identification and characterization of a novel colonization factor based on whole genome sequencing in enterotoxigenic *Escherichia coli* (ETEC)**
von Mentzer A, Tobias J, Wiklund G, Aslett M, Dougan G,
Sjöling Å and Svennerholm A-M
Submitted, 2016.

III. **Identification of candidate novel enterotoxigenic *Escherichia coli* (ETEC) colonization factors based on phenotypic and genotypic analyses**
von Mentzer A, Wiklund G, Dougan G, Sjöling Å and Svennerholm A-M
In manuscript.

SAHLGRENSKA AKADEMIN
INSTITUTIONEN FÖR BIOMEDICIN

GÖTEBORGS UNIVERSITET

Göteborg, 2016

Whole genome sequencing of enterotoxigenic *Escherichia coli* (ETEC)

Identification of ETEC lineages and novel colonization factors

Astrid von Mentzer

Department of Microbiology and Immunology, Institute of Biomedicine, Sahlgrenska akademin, Göteborgs universitet, Sverige, 2016.

Abstract

Enterotoxigenic *Escherichia coli* (ETEC) infections are a common cause of diarrhea but little is known about the evolution and genomic composition of ETEC. The main aim of this thesis was to generate a large collection of whole genome sequenced ETEC isolates to study the evolution of such bacteria on a global level and to search for novel colonization factors (CFs) using both classical and cutting-edge genomic techniques.

Using whole genome sequencing and epidemiological data of 362 human ETEC isolates collected during three decades from all over the world, we studied the population structure of ETEC. We could show that major ETEC lineages comprising isolates with specific virulence profiles, i.e. CFs and LT and/or ST toxins, are stable and spread worldwide. These findings suggest that the virulence genes have been acquired once and then spread through clonal expansion and that a vaccine based on the most prevalent CFs could be protective against a large proportion of ETEC diarrhea cases.

At least 30% of all clinical ETEC isolates lack a known CF. Therefore, we examined whole genome sequences of 94 “CF negative” isolates with the aim to identify novel CFs using two different approaches. **I)** By comparative genomics we have characterized a novel CF, CS30, which is related to the porcine CF 987P (F6). The major subunit of CS30 is 18.5 kD in size and the assembly of the fimbriae is dependent on its expression. CS30-positive bacteria are heavily fimbriated, as shown by electron microscopy, which promotes binding to human intestinal (Caco-2) cells. Furthermore, CS30 expression is thermo-regulated. **II)** By means of phenotypic analyses (SDS-PAGE, Caco-2 adhesion assays and electron microscopy) we have identified a number of isolates that may harbor additional putative fimbrial and non-fimbrial novel CFs. Nine candidate isolates with putative novel CFs were identified from 35 isolates: these were shown to express thermo-regulated proteins of 12-25 kD by SDS-PAGE analyses indicating the presence of major subunits and these isolates were found to bind well to Caco-2 cells. Based on further analyses of these isolates, using comparative genomics to identify CF related genes/operons, four AFA/Dr/AAF-like operons and an operon related to the porcine CF K88, were identified. The findings in this thesis have improved the knowledge of ETEC genomics and will provide a basis for future studies of ETEC transmission and pathogenicity.

Keywords: ETEC, whole genome sequencing, evolution, colonization factor, reverse genetics

ISBN: 978-91-628-9800-7 (PRINT) http://hdl.handle.net/2077/41845
ISBN: 978-91-628-9801-4 (PDF)