

Meeting abstract

Open Access

Bioinformatics determination of ETEC signature genes as potential targets for molecular diagnosis and reverse vaccinology

Heba M Amin¹, Abdel-Gawad M Hashem² and Ramy K Aziz^{*2}

Address: ¹Department of Microbiology and Immunology, Faculty of Pharmacy, MSA University, Cairo, Egypt and ²Department of Microbiology and Immunology, Faculty of Pharmacy, Cairo University, Cairo, Egypt

Email: Ramy K Aziz* - ramy.aziz@salmonella.org

* Corresponding author

from UT-ORNL-KBRIN Bioinformatics Summit 2009
Pikeville, TN, USA. 20–22 March 2009

Published: 25 June 2009

BMC Bioinformatics 2009, **10**(Suppl 7):A8 doi:10.1186/1471-2105-10-S7-A8

This abstract is available from: <http://www.biomedcentral.com/1471-2105/10/S7/A8>

© 2009 Amin et al; licensee BioMed Central Ltd.

Background

Genomes of the model bacterium, *Escherichia coli*, exhibit high plasticity caused by gene gain/loss via pathoadaptive mutations, genetic rearrangement, and horizontal gene transfer [1,2]. This genetic variability is also translated into a remarkable phenotypic and pathotypic diversity: while some *E. coli* strains normally inhabit the mammalian colon, other pathotypes cause a wide range of intestinal and extraintestinal diseases that include mild intestinal disturbance but also severe urinary tract infections and outbreaks of shigellosis-like dysentery or cholera-like watery diarrhea [1]. In this study, we focus on enterotoxigenic *E. coli* (ETEC), one of the world's deadliest infectious agents, which also represents a serious public health in Egypt's rural areas. Our aim is to integrate multiple bioinformatics tools to determine horizontally transferred, pathotype-specific signature genes as targets for specific, high-throughput molecular diagnostic tools and reverse vaccinology screens.

Methods and results

To estimate the extent of horizontal gene transfer in ETEC, we used a combination of bioinformatics tools, including GC%, comparative genomics analysis [3], and web-based prediction of pathogenicity islands via IslandPath <http://www.pathogenomics.sfu.ca/islandpath> [4]. Because *E. coli* strains are typically polylysogenic [5], we used the ACLAME Prophinder tool <http://aclame.ulb.ac.be/Tools/Prophinder> [6] to predict complete or rudimentary prophages scattered within the ETEC genome. To deter-

mine ETEC pathotype-specific genes or signature genes, we used comparative genomic tools available in the National Microbial Pathogen Data Resource (NMPDR) platform <http://www.nmpdr.org>, including the Signature Genes Tool and the Homolog Spreadsheet Tool [7]. We identified 128 genes that differentiate this pathotype from other *E. coli* strains, based on bidirectional-best-hit signature analysis. We also identified 94 genes that are characteristic to two closely related strains (24377A and 2348/69). Many of the ETEC-specific genes were mapped to prophages, prophage-like elements, and other pathogenicity islands; however, some of these signature genes, e.g., ORFs 21–39 in strain 24377A, seem to be rather lost in other *E. coli* strains (as they are conserved among other enterobacteria, e.g., *Shigella* and *Salmonella*). Our ongoing studies are testing some of these ETEC-specific genes as targets for multiplex PCR amplification to develop a rapid diagnostic typing method. Future studies will analyze the surface-association and antigenicity of these signature gene products as a first step in a reverse vaccinology strategy to develop novel ETEC vaccines.

References

1. Dobrindt U: **(Patho-)Genomics of Escherichia coli**. *Int J Med Microbiol* 2005, **295**(6–7):357–371.
2. Morschhauser J, Kohler G, Ziebuhr W, Blum-Oehler G, Dobrindt U, Hacker J: **Evolution of microbial pathogens**. *Philos Trans R Soc Lond B Biol Sci* 2000, **355**(1397):695–704.
3. Roten CA, Gamba P, Barblan JL, Karamata D: **Comparative Genomics (CG): a database dedicated to biometric comparisons of whole genomes**. *Nucleic Acids Res* 2002, **30**(1):142–144.

4. Hsiao W, Wan I, Jones SJ, Brinkman FS: **IslandPath: aiding detection of genomic islands in prokaryotes.** *Bioinformatics* 2003, **19(3)**:418-420.
5. Canchaya C, Proux C, Fournous G, Bruttin A, Brussow H: **Prophage genomics.** *Microbiol Mol Biol Rev* 2003, **67(2)**:238-276. table of contents.
6. Lima-Mendez G, Van Helden J, Toussaint A, Leplae R: **Prophinder: a computational tool for prophage prediction in prokaryotic genomes.** *Bioinformatics* 2008, **24(6)**:863-865.
7. McNeil LK, Reich C, Aziz RK, Bartels D, Cohoon M, Disz T, Edwards RA, Gerdes S, Hwang K, Kubal M, et al.: **The National Microbial Pathogen Database Resource (NMPDR): a genomics platform based on subsystem annotation.** *Nucleic Acids Res* 2007:D347-353.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

