Rakitina et al. BMC Genomics (2017) 18:544
DOI 10.1186/s12864-017-3917-x

BMC Genomics

RESEARCH ARTICLE Open Access

Genome analysis of *E. coli* isolated from Crohn's disease patients

Daria V. Rakitina[1*†], Alexander I. Manolov[1†], Alexandra V. Kanygina[2†], Sofya K. Garushyants[3,4†], Julia P. Baikova[1], Dmitry G. Alexeev[1,2], Valentina G. Ladygina[1], Elena S. Kostryukova[1], Andrei K. Larin[1], Tatiana A. Semashko[1], Irina Y. Karpova[1], Vladislav V. Babenko[1], Ruzilya K. Ismagilova[3,5], Sergei Y. Malanin[3,5], Mikhail S. Gelfand[3,4,6,7], Elena N. Ilina[1], Roman B. Gorodnichev[1], Eugenia S. Lisitsyna[1], Gennady I. Aleshkin[8], Petr L. Scherbakov[9], Igor L. Khalif[10], Marina V. Shapina[10], Igor V. Maev[11], Dmitry N. Andreev[11] and Vadim M. Govorun[1,2]

Abstract

Background: *Escherichia coli* (*E. coli*) has been increasingly implicated in the pathogenesis of Crohn's disease (CD). The phylogeny of *E. coli* isolated from Crohn's disease patients (CDEC) was controversial, and while genotyping results suggested heterogeneity, the sequenced strains of *E. coli* from CD patients were closely related.

Results: We performed the shotgun genome sequencing of 28 *E. coli* isolates from ten CD patients and compared genomes from these isolates with already published genomes of CD strains and other pathogenic and non-pathogenic strains. CDEC was shown to belong to A, B1, B2 and D phylogenetic groups. The plasmid and several operons from the reference CD-associated *E. coli* strain LF82 were demonstrated to be more often present in CDEC genomes belonging to different phylogenetic groups than in genomes of commensal strains. The operons include carbon-source induced invasion GimA island, prophage I, iron uptake operons I and II, capsular assembly pathogenetic island IV and propanediol and galactitol utilization operons.

Conclusions: Our findings suggest that CDEC are phylogenetically diverse. However, some strains isolated from independent sources possess highly similar chromosome or plasmids. Though no CD-specific genes or functional domains were present in all CD-associated strains, some genes and operons are more often found in the genomes of CDEC than in commensal *E. coli*. They are principally linked to gut colonization and utilization of propanediol and other sugar alcohols.

Keywords: Crohn's disease, *E. coli*, Genome, Propanediol

Background

Crohn's disease (CD), one of the major forms of inflammatory bowel disease (IBD), is a chronic generalized inflammation of the gastrointestinal tract. The histological picture of Crohn's disease includes thickened submucosa, transmural inflammation, fissuring ulceration, and non-caseating granulomas. The common complications in the intestine are presented by strictures, abscesses, fistulas and, in the long run, colon cancer. Extraintestinal complications include arthritis, erythema nodosum, uveitis, and primary sclerosing cholangitis.

Many factors, both genetic and environmental, are regarded to contribute to the CD pathogenesis. It is a general notion that CD is a result of abnormal immune response of genetically susceptible individuals to the imbalance in the intestinal microbiota (reviewed in [1, 2]). Host susceptibility factors include intestinal barrier dysfunctions (decreased levels of antimicrobial peptides defensins, discontinuous tight junctions and aberrant mucin assembly) and defects in innate immunity, autophagy, and phagocytosis. Polymorphisms in certain genes, e. g. *NOD2*, *ATG16L1*, and *IRGM*) involved in these processes have been reported to be associated with CD

* Correspondence: drakitina@list.ru
†Equal contributors
[1]Federal Research and Clinical Center of Physical-Chemical Medicine of Federal Medical Biological Agency, Moscow, Russia
Full list of author information is available at the end of the article