

Alejandra Migene Guzmán Del Carpio

**Análise genômica de uma cepa de *Escherichia coli*
enteroagregativa isolada de infecção de corrente
sanguínea**

Dissertação apresentada ao Programa de Pós-Graduação em Microbiologia do Instituto de Ciências Biomédicas da Universidade de São Paulo, para a obtenção do Título de Mestre em Ciências.

Área de concentração: Microbiologia

Orientador: Dr. Waldir Pereira Elias Junior

Versão corrigida

São Paulo

2022

RESUMO

DEL CARPIO A. M. G. **Análise genômica de uma cepa de *Escherichia coli* enteroagregativa isolada de infecção de corrente sanguínea.** 2022. 165 f. Dissertação (Mestrado em Microbiologia) – Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, 2022.

Escherichia coli tem sido relacionada entre as bactérias Gram negativas mais importantes envolvidas em infecções da corrente sanguínea. Vários são os fatores de virulência produzidos por *E. coli* patogênicas, dentre os quais destacam-se as proteínas autotransportadoras (AT), associadas à adesão, evasão ao sistema imune, formação de biofilme e toxicidade. Uma família de proteínas AT com atividades proteolíticas é conhecida por SPATE, ou *Serine Proteases Autotransporters of Enterobacteriaceae*. Em estudos prévios, nosso grupo analisou uma coleção de *E. coli* isolada de bacteremia, identificando a cepa EC092 que alberga os genes que codificam as SPATEs Pet, Pic, Sat e SepA, característica pouco comum em cepas intestinais ou extraintestinais de *E. coli*, além dos genes *aggR*, *aatA*, *aap*, *aaiA* e *aaiG*, característicos de *E. coli* enteroagregativa (EAEC). Como esses dados apontavam para o potencial híbrido dessa cepa, o objetivo do presente estudo foi avaliar características genotípicas e fenotípicas da cepa EC092, a fim de determinar o seu potencial híbrido. Para isso, a EC092 foi avaliada quanto à capacidade de secretar essas quatro SPATEs e de resistir à atividade bactericida do soro humano normal (SHN). Além disso, seu perfil plasmidial foi determinado e seu genoma completo foi analisado *in silico*. A secreção de Pet, Pic, Sat e SepA em sobrenadantes de cultura foi imunodetectada e a cepa EC092 foi capaz de resistir à atividade bactericida do SHN. O perfil plasmidial mostrou a presença de quatro plasmídeos de alto peso molecular. Após, o genoma completo da EC092 foi sequenciado e analisado, mostrando a sua classificação no filogruppo B1, ST278, sorotipo O165:H4, e detectando a presença de vários genes de virulência marcadores de EAEC (*aggR*, *aatA*, *aap*, *aaiA* e *aaiG*) e que codificam as quatro SPATEs (*pet*, *pic*, *sat* e *sepA*). O genoma da EC092 também contém genes que codificam fímbrias (ECP, LPF, HCP, Curli, F1C e Hra1), toxinas (EAST-1, HlyE e MchL), vários sistemas de captação de ferro (sideróforos e receptores), proteínas autotransportadoras da família AIDA-I (EhaG, Cah, YefA, YfaL e EhaC) e proteínas dos sistemas de secreção do tipo I, II, V e VI. Embora carregando o gene *aggR*, não foram detectados os genes que codificam as fímbrias de adesão agregativa (AAF), o fator de colonização CS22 e a fímbria AFP. Também não foram detectados os genes que definem o potencial de virulência extraintestinal ou uropatogênico. Análises filogenéticas com cepas de todos os patótipos de *E. coli* revelaram que o genoma da EC092 é relacionado a cepas de EAEC isoladas de fezes, pertencentes ao filogruppo B1. Também foi avaliada a sua relação com 98 genomas de EAEC isoladas de diferentes regiões geográficas, o que mostrou que quatro cepas de EAEC isoladas de fezes no Quênia (K44V1, K45V1 e K18V1) e no Egito (E13V1D) estão relacionadas a EC092 em um clado. *Heatmaps* construídos com fatores de virulência de vários genomas de EAEC mostraram que as cepas EC092, K45V1 e K44V1 são altamente semelhantes em termos de virulência. Finalmente, foram realizadas análises da vizinhança genômica entre os genomas dessas três cepas, para confirmar a ortologia dos genes *aggR*, *aap*, *pet*, *pic*, *sat* e *sepA*, além dos operons *aatA* e *aai*. A identidade na vizinhança genômica desses genes nos três genomas foi de 100%. Os resultados obtidos neste estudo mostraram que a cepa EC092 isolada de bacteremia é uma linhagem híbrida-patogênica entre EAEC e ExPEC, com potencial de virulência que permitiria a sua translocação do intestino até a corrente circulatória.

Palavras-chave: *Escherichia coli*. Bacteremia. SPATEs. EAEC.

ABSTRACT

DEL CARPIO A. M. G. **Genomic analyses of an enteroaggregative *Escherichia coli* strain isolated from bloodstream infection.** 2022. 165 p. Master thesis (Microbiology) – Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, 2022.

Escherichia coli has been considered as one of the most important Gram-negative bacteria involved in bloodstream infections. Pathogenic *E. coli* produces several virulence factors, including autotransporter (AT) proteins associated with adhesion, immune evasion, biofilm formation and toxicity. A family of AT proteins with proteolytic activities is known as SPATE, or Serine Proteases Autotransporters of Enterobacteriaceae. In previous studies, our group analyzed a collection of *E. coli* isolated from bacteremia and identified the strain EC092, which contains the genes encoding the SPATEs Pet, Pic, Sat and SepA, an uncommon characteristic in intestinal or extraintestinal *E. coli* strains, in addition to *aggR*, *aatA*, *aap*, *aaiA* and *aaiG*, characteristic genes of enteroaggregative *E. coli* (EAEC). As these data indicated the hybrid potential of this strain, the objective of this study was to evaluate genotypic and phenotypic characteristics of EC092, in order to determine its hybrid potential. For this purpose, EC092 was evaluated for its capacity to secrete the four SPATEs and to resist to the bactericidal activity of normal human serum (NHS). Moreover, its plasmideal profile was determined and its whole genome sequenced and analyzed *in silico*. Secretion of Pet, Pic, Sat and SepA was detected in culture supernatants by immunoassays, and EC092 was able to resist to the bactericidal activity of NHS. The plasmideal analysis showed the presence of four high-molecular-weight plasmids. Then, the complete genome of EC092 was sequenced and analyzed, showing its classification into phylogroup B1, ST278, serotype O165:H4, and detecting the presence of several EAEC virulence markers (*aggR*, *aatA*, *aap*, *aaiA* and *aaiG*) and those encoding SPATEs (*pet*, *pic*, *sat*, and *sepA*). The genome of EC092 also contains genes encoding fimbriae (ECP, LPF, Curli, F1C and Hra1), toxins (EAST-1, HlyE and MmcL), various iron uptake systems (siderophores and receptors), autotransporter proteins of the AIDA-I family (EhaG, Cah, YejA, YfaL and EhaC), and proteins of the type I, II, V, and VI secretion systems. While carrying *aggR*, genes encoding the aggregative adhesion fimbriae (AAF), colonization factor CS22 and AFP fimbriae were not detected. Also, genes defining extra-intestinal or uropathogenic virulence potential were not detected. Phylogenetic analyses with strains of all *E. coli* pathotypes revealed that EC092 is related to EAEC strains isolated from feces belonging to the phylogroup B1. The relationship of EC092 with 98 EAEC genomes isolated from different geographic regions was also evaluated, which showed that four EAEC strains isolated from feces in Kenya (K44V1, K45V1 and K18V1) and in Egypt (E13V1D) are related to EC092 in a clade. Heatmaps constructed with virulence factors from several EAEC genomes showed that strains EC092, K45V1 and K44V1 are highly similar in terms of virulence. Finally, analyzes of the genomic neighborhood between the genomes of these three strains were performed to confirm the orthology of the *aggR*, *aap*, *pet*, *pic*, *sat* and *sepA* genes, and the *aatA* and *aai* operons. The identity in the genomic neighborhood of these genes of the three genomes was 100%. The results obtained in this study showed that strain EC092, isolated from bacteremia, is a hybrid-pathogenic strain between EAEC and ExPEC, with virulence potential that would allow its translocation from the intestine to the bloodstream.

Keywords: *Escherichia coli*. Bacteremia. SPATEs. EAEC.