A study of genetic diversity of Escherichia coli BH100 through structural and comparative genomics

Gustavo Santos de Oliveira, Andreia Maria Amaral Nascimento, Edmar Chartone de Souza

Abstract

Genetic variability can be seen as the driving force to evolution. In prokaryotes, many mechanisms emerged such as microorganisms could enhance variability, allowing their spread throughout different ecological niches. In a clinical context, major attention is given to the capacity of some microorganisms in colonize human tissues and developing diseases. From many different bacterial infections currently known, the urinary tract infection (UTI), which is caused mainly by Escherichia coli, can be highlighted as a major concern for public health. The comprehension of mechanisms associated with the emergency of variability and pathogenicity only was possible thanks to the advancements of Molecular Biology and more recently to Bioinformatics. In the present work, we aimed to use and develop bioinformatic tools in order to assembly, annotate and fully characterize the complete genome of E. coli BH100, which was isolated in 1973 in Belo Horizonte, Brazil, from urine of a patient suffering from UTI. This strain is resistant to ampicillin, tetracycline, kanamycin, chloramphenicol, inorganic mercury and in some cases to streptomycin. This multiresistance is due the presence of a mobilizable plasmid carrying a beta-lactamase gene (bla) and a self-transmissible R plasmid carrying the genetic resistance marks of all other elements. Thanks to the use of Next Generation Sequencing (NGS) we were able to assemble the genome of this strain as well as other variants built from curation of the original plasmids, such as we could interrogate the effects of these elements to the emergence of variability. The complete genome of E. coli BH100 resulted in a chromosome of $\tilde{5}$.2 Mb, the smaller mobilizable plasmid of $\tilde{1}$5 kb, and a self-transmissible plasmid of $\tilde{1}$07 kb. This strain has shown a considerable number of insertion sequences from the family IS3 spread differently along the chromosome of BH100 and its variants. The comparative analysis indicate that this strain might be an authentic uropathogenic E. coli (UPEC) causing pyelonephritis. The functional annotation confirmed the presence of all resistance marks in transposons (Tn). Special attention was given to the presence of Tn21, identical to the one found in plasmid NR1, and to a potential new transposon carrying a gene for kanamycin resistance flanked by IS5 elements. In the end, we propose a new mechanism capable to explain the emergence of unstable and diversified resistance to streptomycin within the population of E. coli BH100. In this model, we suggest streptomycin resistance shows up due an increase in the copy number of the gene aadA1 present in Tn21

Funding: