TITLE: A NEW WAY TO IMPROVE THE ACCESS TO BIOINFORMATIC TOOLS FOR BACTERIAL GENOME ANALYSES USING TELEGRAM APP

AUTHOURS: SANTOS, F. F.[1]; TROVÃO, L. O.[1]; GOMES, T. A. T.[1]; KITAMURA, F. C.[2]

INSTITUTION: [1]DEPARTAMENTO DE MICROBIOLOGIA, IMUNOLOGIA E PARASITOLOGIA, ESCOLA PAULISTA DE MEDICINA, UNIVERSIDADE FEDERAL DE SÃO PAULO, R. BOTUCATU, 862, 3º. ANDAR, VILA CLEMENTINO, CEP 04023-062, SÃO PAULO - SP, BR. [2]DEPARTAMENTO DE DIAGNÓSTICO POR IMAGEM, ESCOLA PAULISTA DE MEDICINA, UNIVERSIDADE FEDERAL DE SÃO PAULO, R. NAPOLEÃO DE BARROS, 800, VILA CLEMENTINO, CEP 04024-000, SÃO PAULO - SP, BR.

ABSTRACT:

The bioinformatics software developed in later years allowed to convert the large amount of bacterial genome data, provided by next generation sequencing, into relevant and useful information for biological and health research. Nevertheless, a great number of bacterial proteins are still characterized as hypothetical or have their function poorly described. To cover this gap, machine-learning based approaches have been created to classify proteins as virulence factors, such as a specific tool for identifying adhesins and adhesin-like proteins, SPAAN. Considering the significant contribution of this software to search for genes involved bacterial pathogenesis, the two objectives of this work were (1) to make the access to SPAAN easier to all researchers, including those who are not familiar with bioinformatic tools based on Linux command line, creating an even more accessible interface to SPAAN via a free smartphone App (Telegram) available for Android and iOS (2) to demonstrate that this mobile app could be used as an innovative way to improve the access to bioinformatic resources. A Telegram Bot was created using the python-telegram-bot wrapper, which receives an input FASTA file, sent by the user through the Telegram Bot, and processes it using the askquery executable pipeline, which runs the SPAAN binaries in a specific order. The output of SPAAN is then sent back to the user as a comma separated values (CSV) file inside the chat. The Bot provides simple usability instructions and reference to the SPAAN publication. The Telegram Bot was successfully created and deployed in Amazon Web Services virtual machine. The instructions to use the Bot are: (1) download Telegram in Apple Store or GooglePlay; (2) sign in with your phone number; (3) search for "Adhesin_Finder"; (4) once in the chat with the bot, click the "Start" button; (4) Within the chat with the Bot, send any FASTA file with a protein sequence. After a few seconds, the user should receive the "query.out" file containing the predictions. The user who needs a fast setup interface may use SPAAN via Telegram web or smartphone app. Considering this, we expect that this interface could widen the usage of bioinformatic tools for bacterial genome analyses.

Keywords: adhesin prediction, fast setup interface, protein function prediction, SPAAN.