

COMPUTATIONAL METHODS FOR BACTERIAL CHARACTERIZATION AND BACTERIA-HOST/ENVIRONMENT INTERACTION ANALYSES

Chao Zhang

Dr. Dong Xu, Dissertation Supervisor

ABSTRACT

With the rapid development of next-generation sequencing technologies, bacterial identification becomes a very important and essential step in processing genomic data, especially for metagenomic data. Many computational methods have been developed and some of them are widely used to address the problems in bacterial identification. First we review the algorithms of these methods, discuss their drawbacks, and propose future computational methods that use genomic data to characterize bacteria. Then, we tackle two specific computational problems in bacterial identification, namely, the detection of host-specific bacteria and the detection of disease-associated bacteria, by offering potential solutions as a starting point for those who are interested in the area. In addition, by utilizing our knowledge of *H. pylori*, we also predicted novel effectors for those known pathogens.