Literature Review

Yanru Liao, Breanna Richards, Jina Yang

2022-10-23

Introduction and Literature Review

In recent years, the excessive and uncontrolled use of antibiotics for the treatment of infectious diseases in humans and animals has become a major area of concern for human health. The primary mechanism of antibiotics ineffectiveness medicine is antibiotic residues in daily food including raw meat and milk. One of the main consequences of antibiotic residues in foods of animal origin is the proliferation of antibiotic-resistant bacteria. The presence of antibiotic-resistant pathogenic bacteria in foods may lead to foodborne infections in humans with an increased incidence of treatment failure and severity of disease. Additionally, they can transfer the resistance genes to other microorganisms through the food chain (Hassani et al., 2022; Rajaei et al., 2021). Food products of animal origin such as meat and meat products are the main media for the transmission of food-borne zoonotic bacterial pathogens. Escherichia coli (E.coli), and Salmonella have been known as the major zoonotic bacterial pathogens. These pathogens are associated with many cases of foodborne illness and death in humans following the consumption of contaminated food in the world (Hashempour-Baltork et al., 2019; Rajaei et al., 2021). Effective treatment of these infections requires knowledge of the organism's susceptibility to various antibiotics. However, the traditional method to obtain this information is culturing bacteria in the clinical laboratory and subsequent testing for commonly used antibiotics, which is heavily time-consuming (Ren et al., 2022; Van Camp et al., 2020). High-throughput sequencing technology based on genomic data of bacteria has the potential to fast guide the clinical decision-making process. Our target pathogenic bacteria isolates in this study include Salmonella enterica, E.coli Shigella, and Campylobacter jejuni. In this study, we aim to predict drug resistance and susceptibility of bacteria isolates based on genomics (i.e., resistant genes and genetic variants) and other characteristics (i.e., location, source type, host disease, etc) and capture their changes over time. This enables us to aid the prevention of bacteria-related foodborne disease outbreaks. Clinicians and policymakers can impede the dissemination of these human diseases in advanced based on these predictions.

Our methodology and analyses conducted in this study are an application of Van Camp, et al.'s paper published in the Frontiers in Microbiology Journal in 2020. Specifically, we will predict whether a pathogen is susceptible or resistant to antibiotics including tetracycline, streptomycin, sulfisoxazole, ampicillin, ceftriaxone, etc. In similar previous studies, machine learning methods and penalized logistic regression have performed well in drug resistance prediction (Ren et al., 2022; Van Camp et al., 2020). We will convert all intermediate antibiotics resistance and demonstrated resistance levels to 'Resistance' and dose-dependent susceptibility levels to 'Susceptibility' to project data to a binary classification problem. The whole genome sequencing data in terms of coverage of known antibiotics resistance genes is our input data. The National Center for Biotechnology Information's (NCBI) Pathogen database has a high sequence similarity (including polymorphism in strains or sequences derived from closely related species) that cannot be classified by the alignment techniques. In the preprocessing stage, we will cluster similar antibiotic resistance genes using the cluster_fast module of the USERCH algorithm to group genes with greater than 90% sequence identity together. Those clustered genes will be named after the most representative gene as defined by USERCH (i.e., cluster centroid) (Edgar, 2010). We chose the threshold of 90% for grouping as a suboptimal compromise between balancing the number of genes per cluster, performance of the model, and the biological relevance of the genes grouped together (Van Camp et al., 2020). We will evaluate our models by comparing the predicted antibiotics resistance and susceptibility to laboratory-confirmed drug susceptibility through AUC.

Bibliography

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. Bioinformatics (Oxford, England), 26(19), 2460–2461. https://doi.org/10.1093/bioinformatics/btq461

Hashempour-Baltork, F., Hosseini, H., Shojaee-Aliabadi, S., Torbati, M., Alizadeh, A. M., & Alizadeh, M. (2019). Drug Resistance and the Prevention Strategies in Food Borne Bacteria: An Update Review. Advanced Pharmaceutical Bulletin, 9(3), 335–347. https://doi.org/10.15171/apb.2019.041

Hassani, S., Moosavy, M.-H., Gharajalar, S. N., Khatibi, S. A., Hajibemani, A., & Barabadi, Z. (2022). High prevalence of antibiotic resistance in pathogenic foodborne bacteria isolated from bovine milk. Scientific Reports, 12(1), Article 1. https://doi.org/10.1038/s41598-022-07845-6

Rajaei, M., Moosavy, M.-H., Gharajalar, S. N., & Khatibi, S. A. (2021). Antibiotic resistance in the pathogenic foodborne bacteria isolated from raw kebab and hamburger: Phenotypic and genotypic study. BMC Microbiology, 21(1), 272. https://doi.org/10.1186/s12866-021-02326-8

Ren, Y., Chakraborty, T., Doijad, S., Falgenhauer, L., Falgenhauer, J., Goesmann, A., Hauschild, A.-C., Schwengers, O., & Heider, D. (2022). Prediction of antimicrobial resistance based on whole-genome sequencing and machine learning. Bioinformatics, 38(2), 325–334. https://doi.org/10.1093/bioinformatics/btab681

Van Camp, P.-J., Haslam, D. B., & Porollo, A. (2020). Prediction of Antimicrobial Resistance in Gram-Negative Bacteria From Whole-Genome Sequencing Data. Frontiers in Microbiology, 11, 1013. https://doi.org/10.3389/fmicb.2020.01013