

## Project Proposal: Genomically Predicting Antibiotic Resistance in Gonorrhea

Gonorrhea is the second most commonly reported bacterial communicable disease in the United States with an estimated 1,568,000 new infections per year.<sup>1</sup> This sexually transmitted infection (STI) is caused by the bacteria *N. gonorrhoeae* and is treated with antibiotics. Untreated gonorrhea can cause health problems including ectopic pregnancy, infertility, and possible increased risk for contracting HIV,<sup>2</sup> and new infections cost the U.S. healthcare system \$271 M each year.<sup>3</sup> However, strains of gonorrhea have developed antibiotic resistance with almost half of all infections now resistant to at least one treatment.<sup>4</sup> While the current primary treatment, ceftriaxone, is effective against almost all strains, its common usage risks further increases in resistances and raises the threat of antimicrobial-resistant gonorrhea.

Identifying opportunities to efficiently utilize other antibiotics, such as azithromycin and cefixime, provides a course of action to mitigate further resistance to ceftriaxone. The mechanisms that result in resistance are coded into the DNA of strains, and as sequencing continues to become more efficient and affordable, genetic testing presents an increasingly more viable means to predict antibiotic resistance of a strain and recommend effective treatment. This has the potential to minimize the cost of new infections on healthcare systems in addition to slowing the rising antibacterial resistance.

Data from 12 studies that sequenced antibiotic-resistant strains of gonorrhea include the strains, common shared DNA (unitigs) present in each strain, and resistance to azithromycin, cefixime, and ceftriaxone measured by Minimal Inhibitory Concentration (MIC).<sup>5</sup> **I propose using this data to create a model that predicts the resistance of a strain for azithromycin, ciprofloxacin, and cefixime from the unitigs present in the strain with a recall of 85%.**

This project will focus solely on these three antibiotics as they are the three latest frontline treatments and are present in the data.<sup>6</sup> A constraint of the project is that strains are constantly mutating and more recent mechanisms of resistance will have different genetic codes not identified in these unitigs. Therefore, while this data can provide an effective tool, updated sequences will maximize the predictive power of the model.

---

<sup>1</sup> <https://www.cdc.gov/std/treatment-guidelines/gonorrhea-adults.htm>

<sup>2</sup> <https://www.cdc.gov/std/gonorrhea/stdfact-gonorrhea.htm>

<sup>3</sup> <https://www.cdc.gov/std/statistics/prevalence-2020-at-a-glance.htm>

<sup>4</sup> <https://www.cdc.gov/std/statistics/2020/overview.htm>

<sup>5</sup> <https://www.kaggle.com/datasets/nwheeler443/gono-unitigs>

<sup>6</sup> <https://www.mdpi.com/2079-6382/7/3/60>