A ETFL based Machine Learning Approach for Discovering Mechanisms of Antibiotic Resistance

Proposal

Chunhao Gu

April 9, 2021

Antibiotics, a cornerstone of modern medicine, are threatened by the increasing burden of drug resistance and diminishing antimicrobial discovery pipelines [BW16]. In recent years, there is growing appreciation for studying secondary processes, for example, how bacterial metabolism interfaces with antibiotic lethality and resistance, which has the potential to open new drug discovery paradigms [MMP14].

Recently, the paper [Yan+19] proposed a new computational approach, called white-box machine learning, which shows the power in discovering bacterial metabolic mechanism on antibiotic lethality. The approach uses elastic net regression to correlate simulated fluxes of iJO1366 genome-scale metabolic model of *E. coli* and experimental IC_{50} values of antibiotic lethality. Since the metabolic model is mechanistically causal, the approach has the potential to discover the causality between the machine learning input and output. This is why the it is called white-box machine learning.

The metabolic simulation was done through parsimonious flux balance analysis (pFBA) [Lew+10] optimized for the biomass objective function and constrained for different chemical environments. Though the method is effective, the relatively low accuracy of pFBA might affect the validity of downstream machine learning and analysis.

On the other hand, the paper [SH20] put forward a unified formulation for **E**xpression and **T**hermodynamics-enabled **FL**ux models, called ETFL. It combines the previous work on ME-models [Ler+12; OBr+13; Llo+17] and TFA [HBH07], and changes a plain metabolic model to a hierarchical one, including thermodynamics, metabolism, RNA synthesis, etc., and is expected to provide higher accuracy of metabolic simulation.

In our research, we will aim to combine ETFL model and white-box machine learning and uses the new framework to explore mechanisms of antibiotic resistance. The research can be divided into three steps:

1) We use ETFL instead of pFBA model to simulate iJO1366 genome-scale metabolic model of *E. coli* under the same chemical environments as the paper [Yan+19], and under (all) the combinations of two chemical environments in the paper. In addition, we might also combine ETFL and pFBA.

2) To the new ETFL simulated metabolic fluxes and IC_{50} values provided by the paper [Yan+19], we apply the same machine learning approach as the paper [Yan+19]. In addition, we might alter the original machine learning approach to improve the performance.

3) With the trained machine learning model, we will do downstream analysis, such as enrichment analysis, to explore clues that might reveal mechanisms of antibiotic resistance and lethality.

In future, we could incorporate the new discovery of mechanisms of antibiotic resistance back into the ETFL model. The iteratively reconstructive ETFL model will be more accurate in simulating

metabolism and facilitate to discover new drugs that fuel antibiotic lethality or weaken antibiotic resistance.

References

[BW16] Eric D. Brown and Gerard D. Wright. "Antibacterial drug discovery in the resistance era." In: *Nature* 529.7586 (2016), pp. 336–343. ISSN: 0028-0836. DOI: 10.1038/nature17042.

[HBH07] Christopher S. Henry, Linda J. Broadbelt, and Vassily Hatzimanikatis. "Thermodynamics-Based Metabolic Flux Analysis." In: *Biophysical Journal* 92.5 (2007), pp. 1792–1805. ISSN: 0006-3495. DOI: 10.1529/biophysj.106.093138.

[Ler+12] Joshua A. Lerman et al. "In silico method for modelling metabolism and gene product expression at genome scale." In: *Nature Communications* 3.1 (2012), p. 929. DOI: 10.1038/ncomms1928.

[Lew+10] Nathan E Lewis et al. "Omic data from evolved E. coli are consistent with computed optimal growth from genome-scale models." In: *Molecular Systems Biology* 6.1 (2010), p. 390. ISSN: 1744-4292. DOI: 10.1038/msb.2010.47.

[Llo+17] Colton J. Lloyd et al. "COBRAme: A Computational Framework for Genome-Scale Models of Metabolism and Gene Expression." In: *bioRxiv* (2017), p. 106559. DOI: 10.1101/106559.

[MMP14] Paul Murima, John D. McKinney, and Kevin Pethe. "Targeting Bacterial Central Metabolism for Drug Development." In: *Chemistry & Biology* 21.11 (2014), pp. 1423–1432. ISSN: 1074-5521. DOI: https://doi.org/10.1016/j.chembiol.2014.08.020.

[OBr+13] Edward J O'Brien et al. "Genome-scale models of metabolism and gene expression extend and refine growth phenotype prediction." In: *Molecular Systems Biology* 9.1 (2013), p. 693. ISSN: 1744-4292. DOI: 10.1038/msb.2013.52.

[SH20] Pierre Salvy and Vassily Hatzimanikatis. "The ETFL formulation allows multi-omics integration in thermodynamics-compliant metabolism and expression models." In: *Nature Communications* 11.1 (2020), p. 30. DOI: 10.1038/s41467-019-13818-7.

[Yan+19] Jason H. Yang et al. "A White-Box Machine Learning Approach for Revealing Antibiotic Mechanisms of Action." In: *Cell* 177.6 (2019). ISSN: 0092-8674. DOI: 10.1016/j.cell.2019.04.016.