

Generating Drug-Induced Gene Expression Profiles through Continuous Dose-Response Curve Modeling

Martin Schmuahlek¹, Iljung Jin², and Hojung Nam^{2,3,*}

¹*Department of AI Convergence, Gwangju Institute of Science and Technology (GIST)*

²*School of Electrical Engineering and Computer Science, Gwangju Institute of Science and
Technology (GIST)*

³*AI Graduate School, Gwangju Institute of Science and Technology (GIST)*

*Corresponding author: hjnam@gist.ac.kr

Understanding the molecular effects of drugs on target cells is crucial for elucidating drug mechanisms and advancing pharmaceutical development. Drug-induced gene expression (DIGEX) profiles capture these molecular responses and serve as a valuable resource for uncovering mechanisms of action and identifying drug repurposing opportunities. However, experimentally measuring DIGEX across the vast landscape of drugs, cell lines, and concentrations is infeasible. To address this challenge, we present a novel deep learning approach that generates DIGEX profiles across arbitrary combinations of drugs, cell lines, and drug concentration. Our key innovation lies in generating DIGEX for arbitrary drug concentration within a Hill-slope equation framework, rather than treating it as a fixed input value like previous methods, enabling continuous interpolation of gene expression across dosages. In comprehensive evaluations using the LINCS L1000 dataset, the proposed model demonstrates superior performance compared to existing models, particularly when generalizing to previously unseen drugs and cell lines, demonstrating strong out-of-distribution (OOD) performance. The generated DIGEX profiles retain both drug-specific and cell-type-specific biological signals, as well as drug mechanism signals confirmed through network analysis. This framework offers a biologically grounded, scalable solution for modeling DIGEX changes to accelerate drug development pipelines.