

Uncertainty-aware drug response prediction with chemical foundation models

Team novAlce





Chemical foundation models create informative biological representations

The Problem

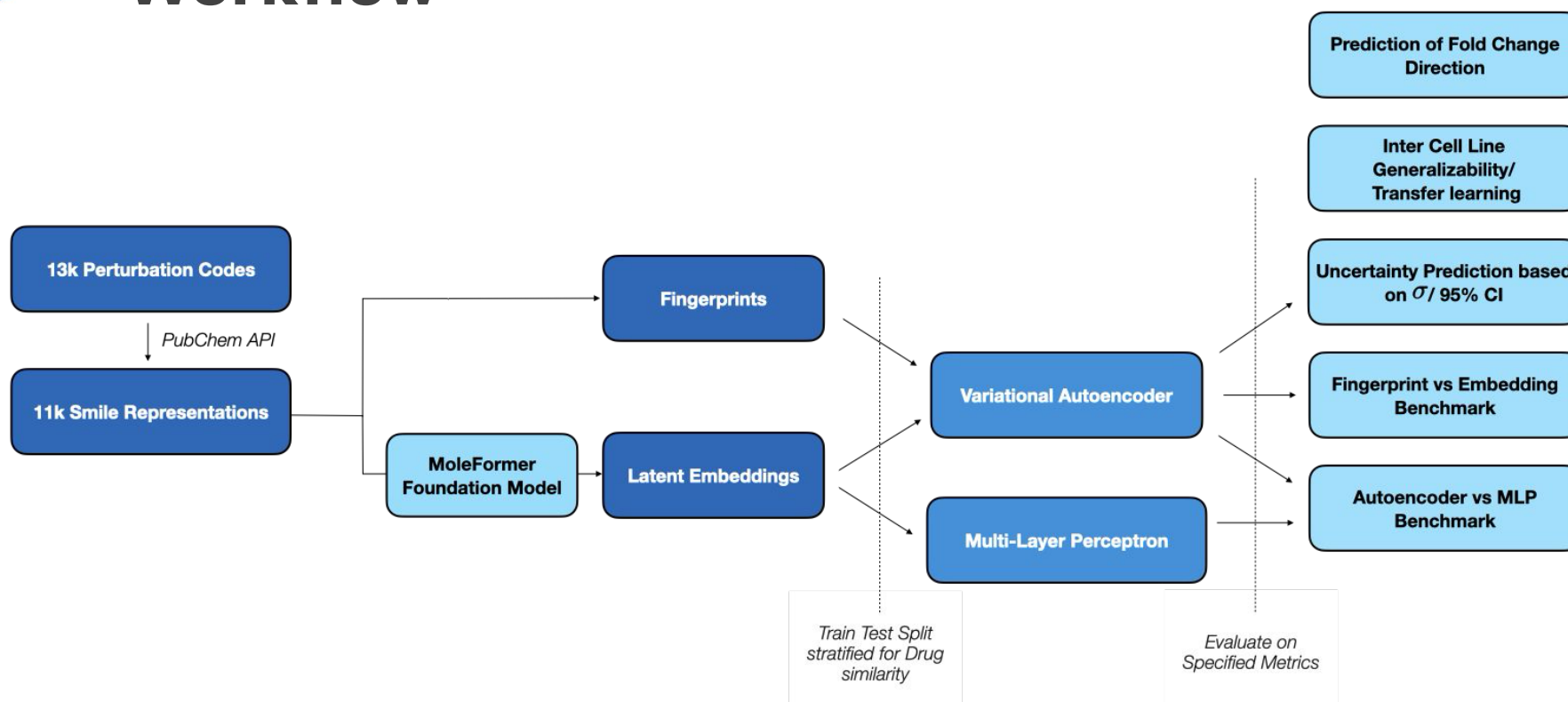
- Drug Development is **slow and expensive**

Our approach

- **Uncertainty aware models** – we predict an effect and quantify how confident we are
- **Foundation models for chemistry** – rich embeddings of compounds to generalize to unseen drugs
- **Fast, robust pipeline** – production style code so Novartis can iterate on new screens quickly

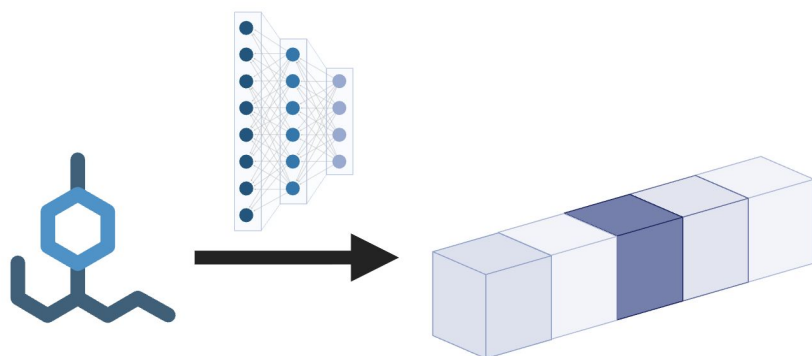


NovAI Genetic Expression Prediction Workflow

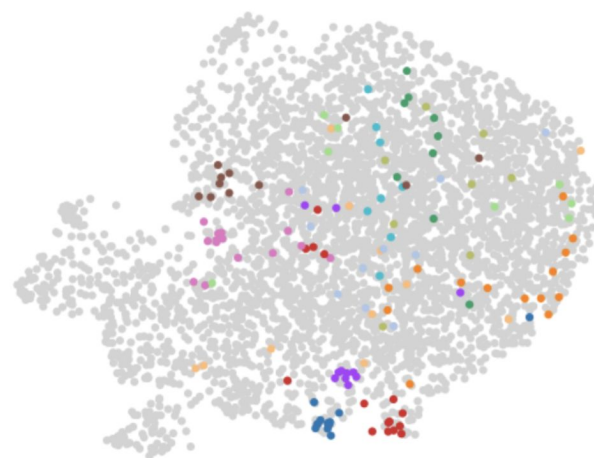




Chemical foundation models create informative biological representations



Latent representation

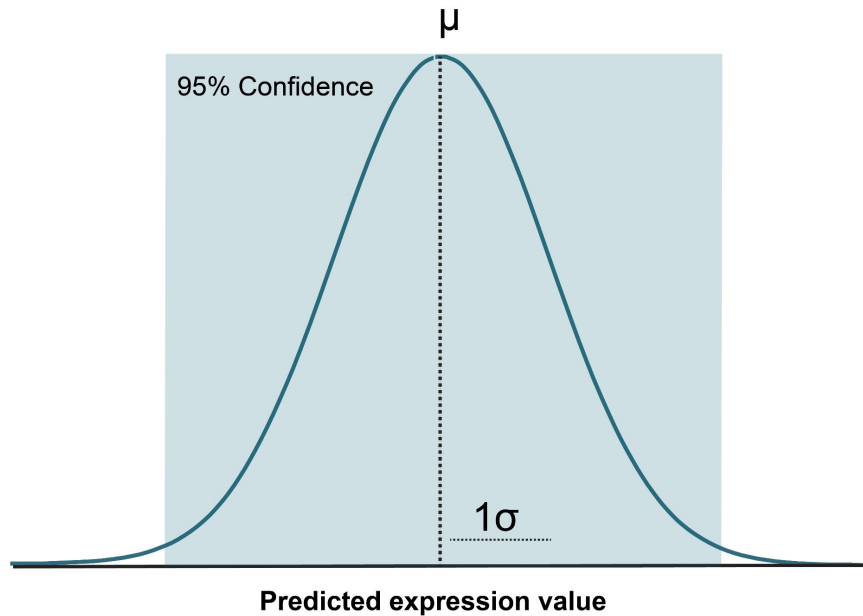


- 5-HT3 receptor agonist
- CDK Inhibitors
- Calcium-Sensing Receptor (CaSR) Antagonists
- Carbonic anhydrase inhibitor
- Carboxylesterase Inhibitors
- Dipeptidyl Peptidase IV (CD26;DPP-IV;DP-IV)
- Histone Deacetylase (HDAC) Inhibitors
- Phosphodiesterase IV Inhibitors
- Protein Kinase C (PKC) Inhibitors
- alpha1-Adrenoceptor Antagonists
- alpha-Adrenoceptor Antagonists
- mgluR5 Antagonists
- NA

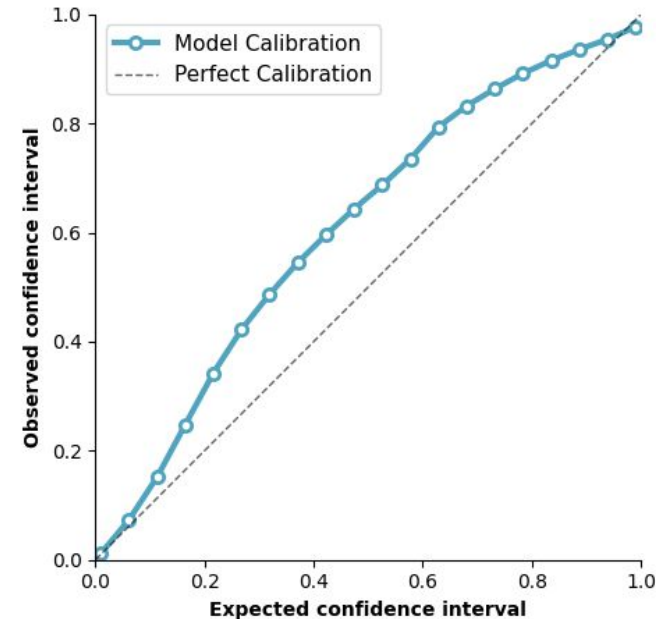
UMAP on Molformer embeddings



Uncertainty prediction with a probabilistic model



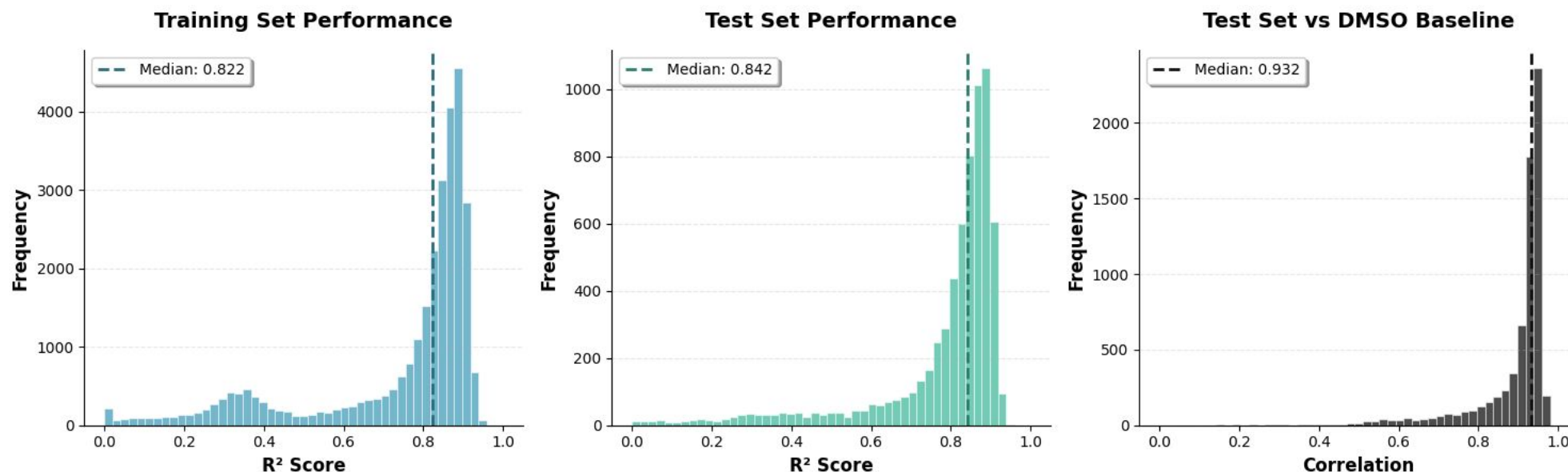
Probabilistic model (scVI generative model)



Agreement of expected confidence interval with observed confidence interval (test set)



Predicting transcriptional responses to drug perturbations in cell lines

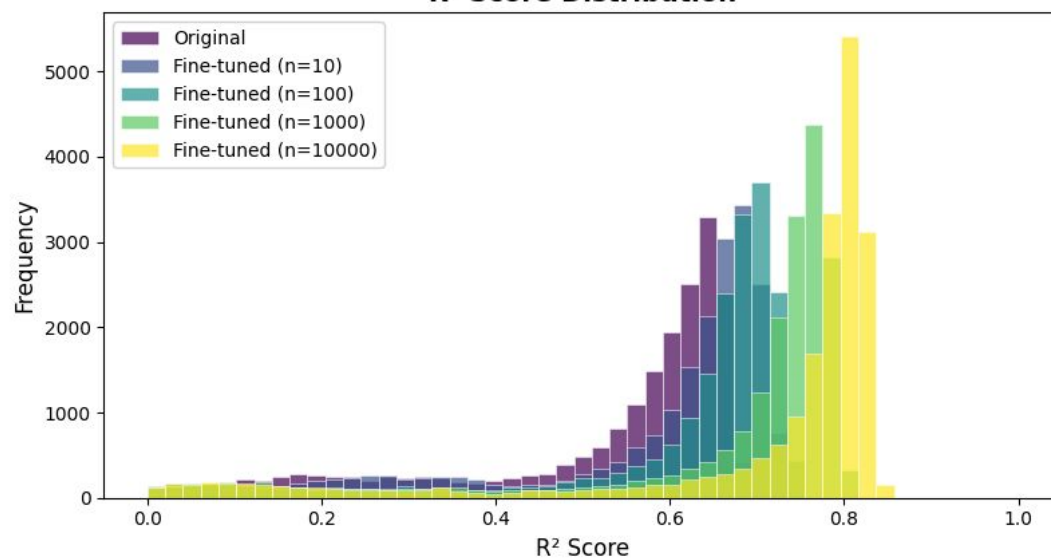


MCE Bioactive Compounds (HEK293T, 10 μ M) on 2000 highly variable genes. **0.84 median R²** across test set, **44% accuracy in directional change prediction**. Prediction on 20% test split stratified by compounds

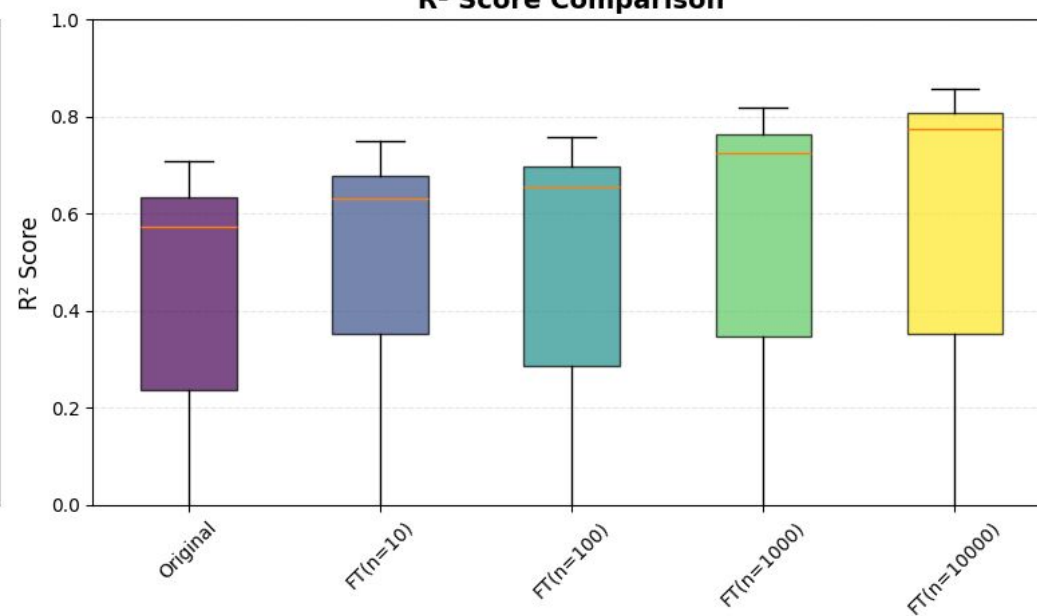


Transfer learning improves generalization between cell lines

R² Score Distribution



R² Score Comparison





24 hours of code. Years of screening time saved.

What we do

- Predict transcriptome responses in silico
- Prioritize compounds that hit disease pathways
- Cut down wet lab compounds, doses and cell lines