Low-Concentration Cell Painting Images enable the Identification of Highly Potent Compounds

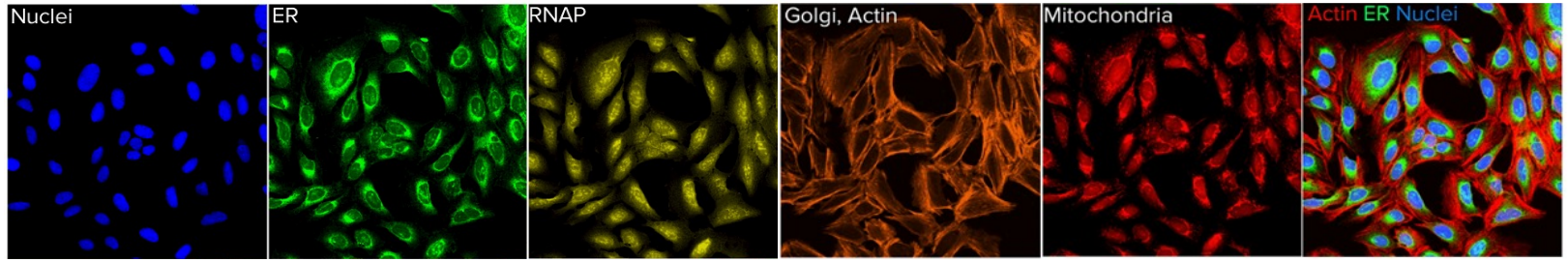
Son V. Ha[a,b], Steffen Jaensch[a], Lorena G. A. Freitas[a], Dorota Herman[a], Paul Czodrowski[b], and Hugo Ceulemans[a]

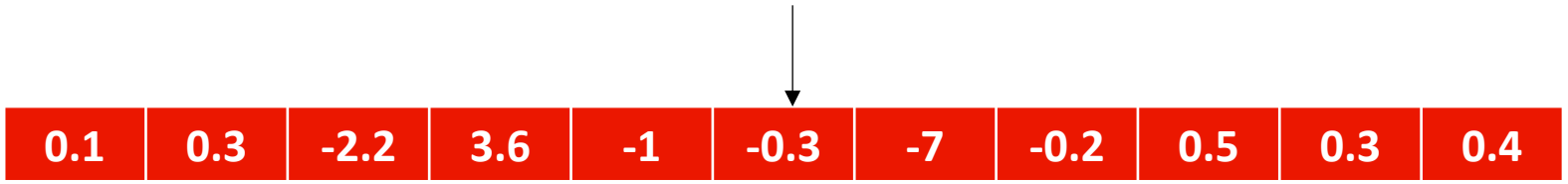
[a]Johnson & Johnson Innovative Medicine
[b]Department of Physical Chemistry, Johannes Gutenberg University Mainz, Germany

Johnson&Johnson
Innovative Medicine

Introduction
Cell Painting (CP) Protocol

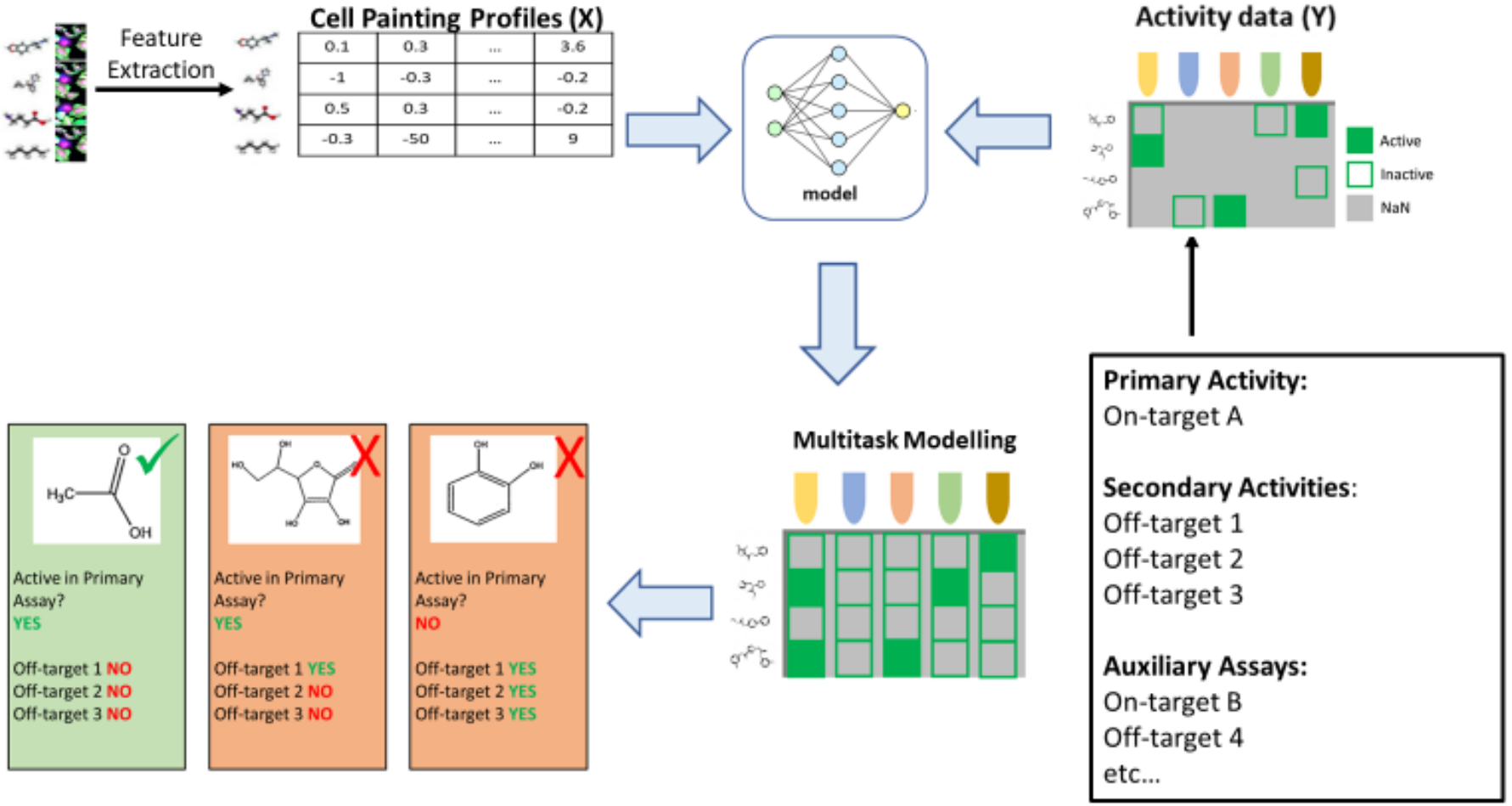
- **Morphological Profiling** of compounds. (How cells ‘look’)
- Microscopy images of cells perturbed with a compounds.
- CP can be performed at different concentrations
  - 0.16uM, 0.8uM, 4uM, 10uM, 20uM



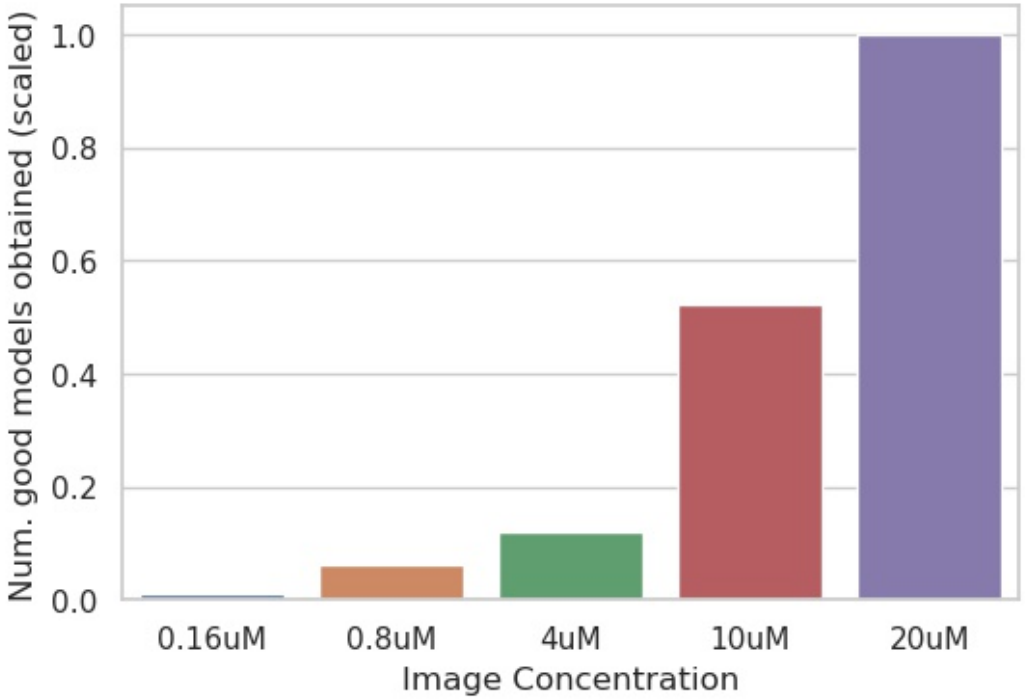


Introduction

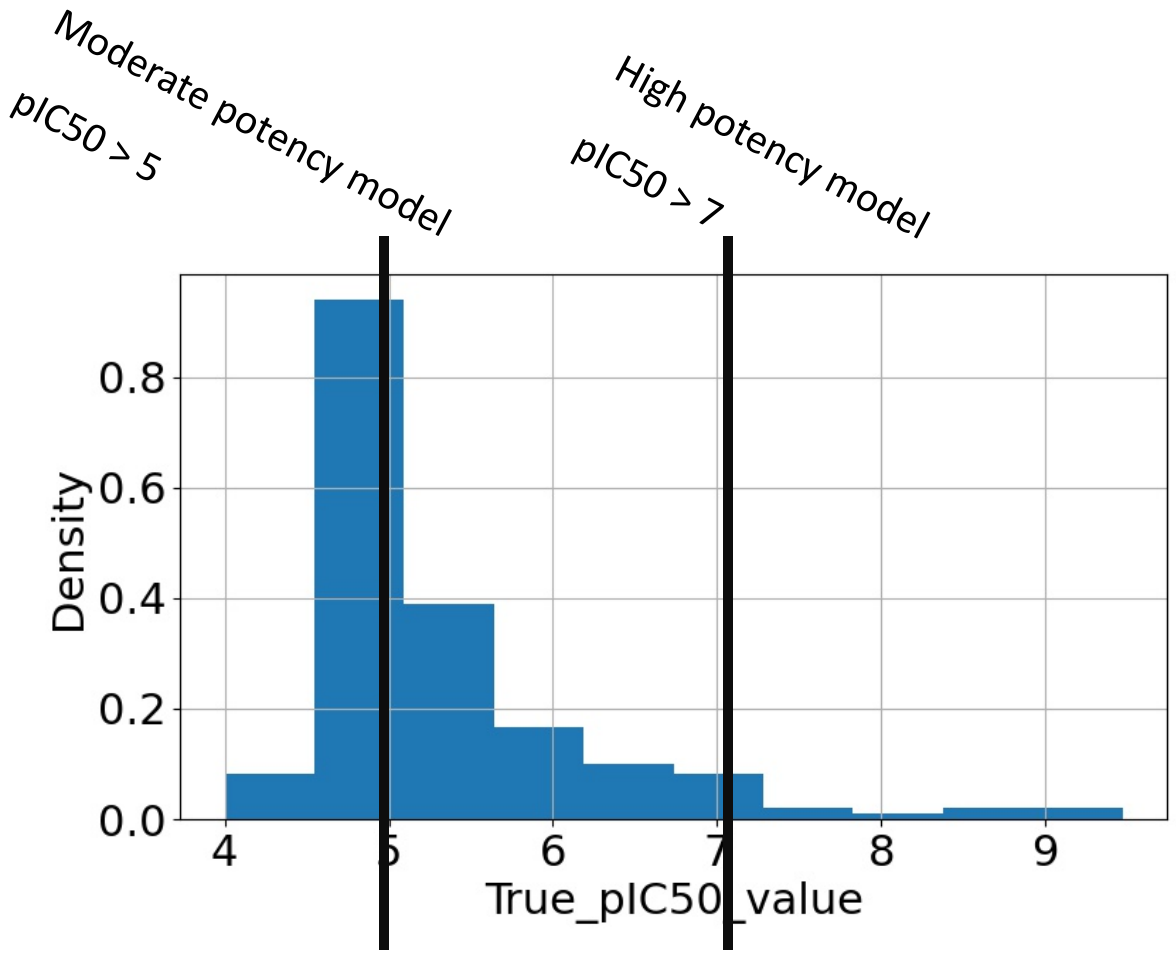
Image-informed Ligand-based Multitask Activity Modelling



Introduction
Research Question



1. What are the uses of **low-concentration images**?



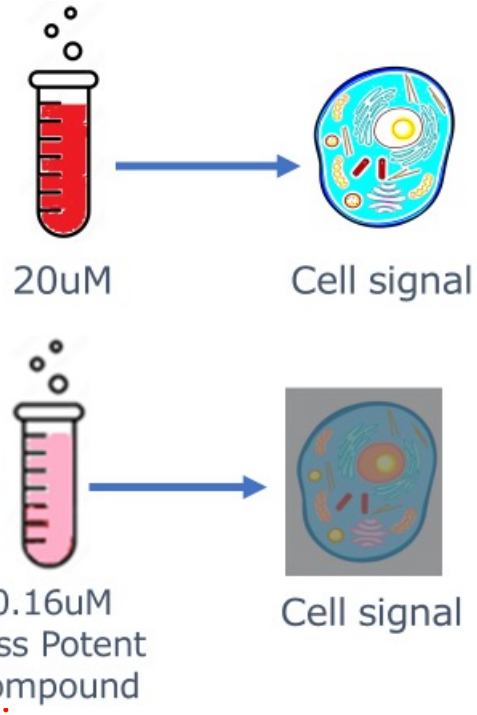
2. How to obtain more **good high potency model**?

Hypothesis
Why Low Concentration Images Lead to Bad Models?

A lot of cell signal is needed for modelling

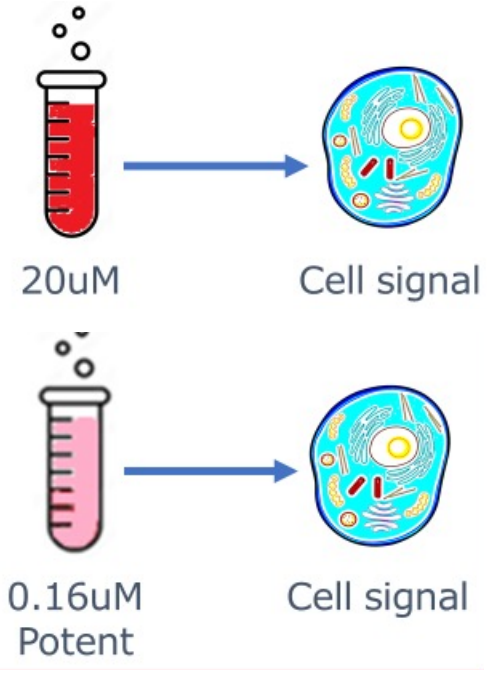
Less Potent Compound

(common)



Highly Potent Compound

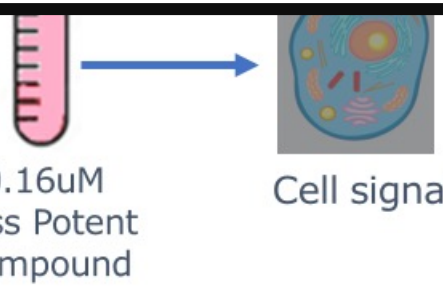
(rare)

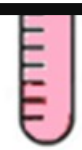


Hypothesis
Why Low Concentration Images Lead to Bad Models?

A lot of cell signal is needed for modelling

Can we leverage this behavior to specifically retrieve highly potent compounds?

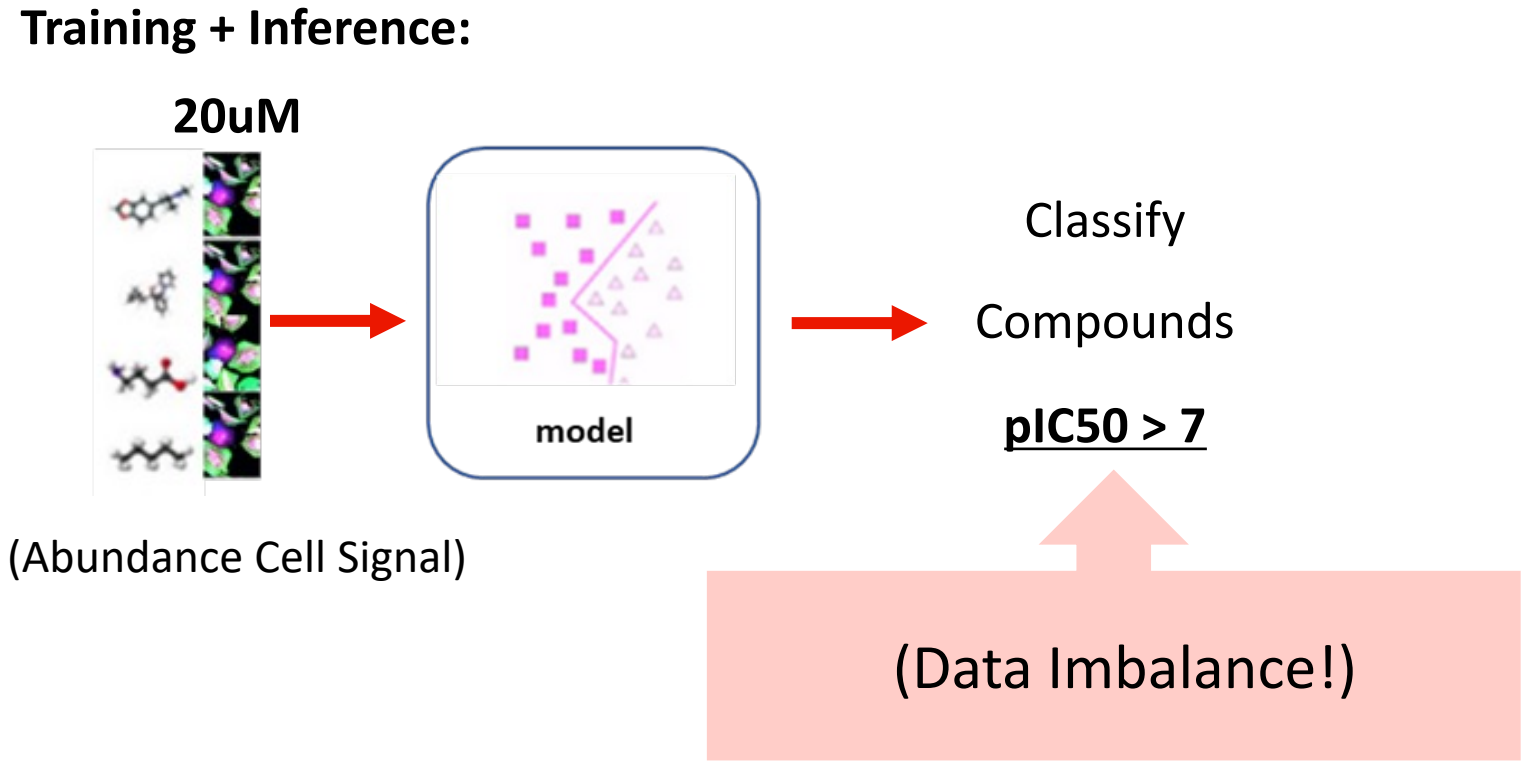




Method
Traditional Method

Aim: Classify compounds with pIC50 > 7

Method: Build a Level 7 Model



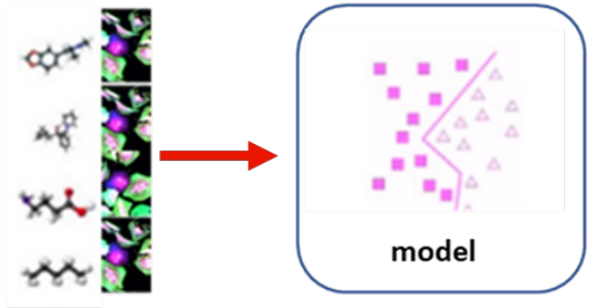
Method

Proposed Approach: Training/Inference with Different Concentration Images

Aim: Classify compounds with pIC50 > 7

Training

20uM



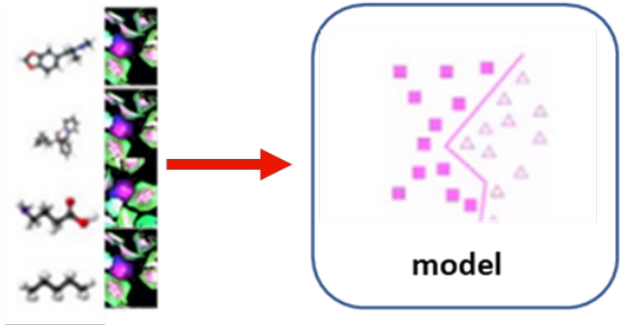
Classify Compounds **pIC50 > 5**

(Abundance Cell Signal)

(No Data Imbalance)

Inference:

0.16 uM



Classify Compounds **pIC50 > 7**

(Only Potent compounds show signal)

Build a Level 5 Model…

… then repurpose it into a Level 7 Model

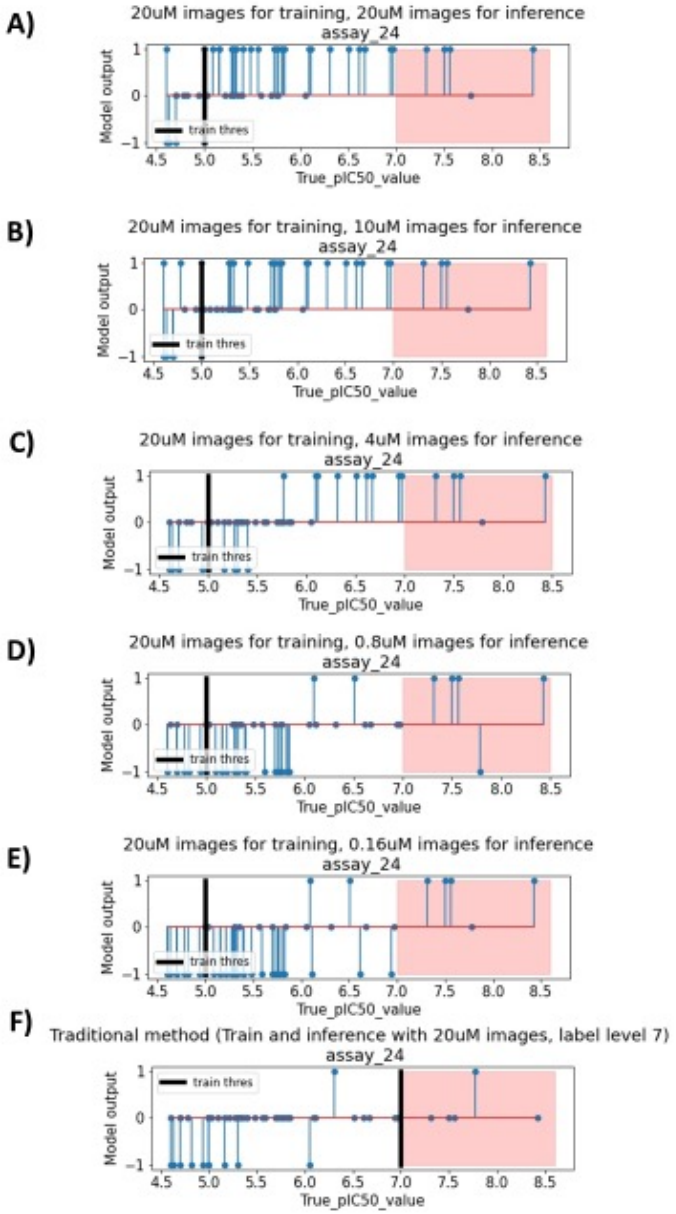
Results

1. Activity Models can be repurposed for higher potency classification

TL,DR:

The approach works as intended

→ Our hypothesis holds



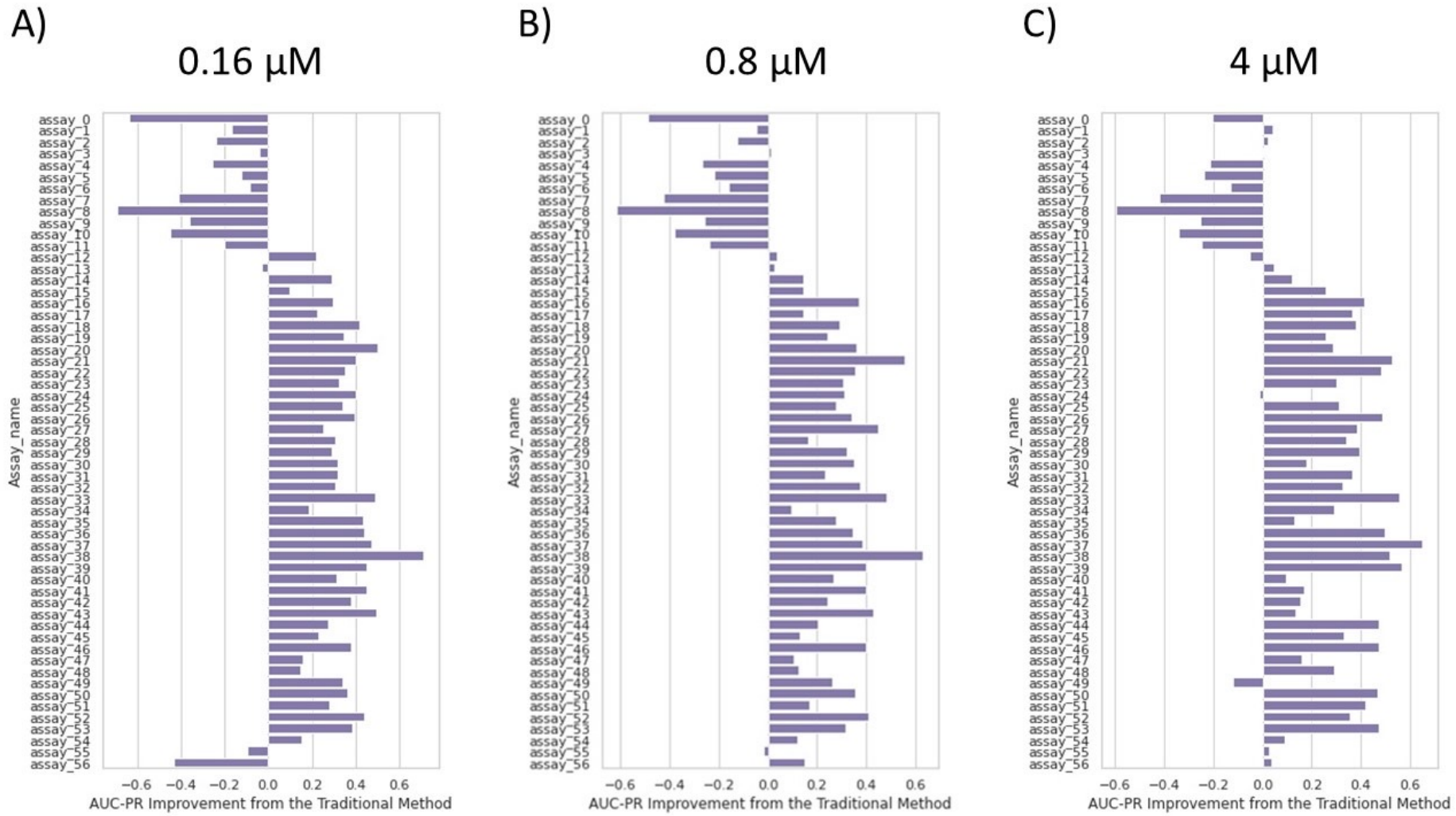
Y-axis: Model output

X-axis: True pIC50 Values

Observation: Low concentration images for inference → Model only ‘focus’ on highly potent compounds

Results

2. Our Approach Improves High Potency Classification in ~75% of assays over Traditional Method



Conclusion
Benefits to the Drug Discovery Pipeline

- Prioritizing hits for experimental follow-up (based on potency).
- Deprioritizing compounds with potent off-target activities in hit-triaging.