REVIVE
by GARDP

Charting new frontiers in artificial intelligence for antibiotic design

Guest speakers: Jonathan Stokes & Kurt Thorn

Moderator: Akhila Kosaraju

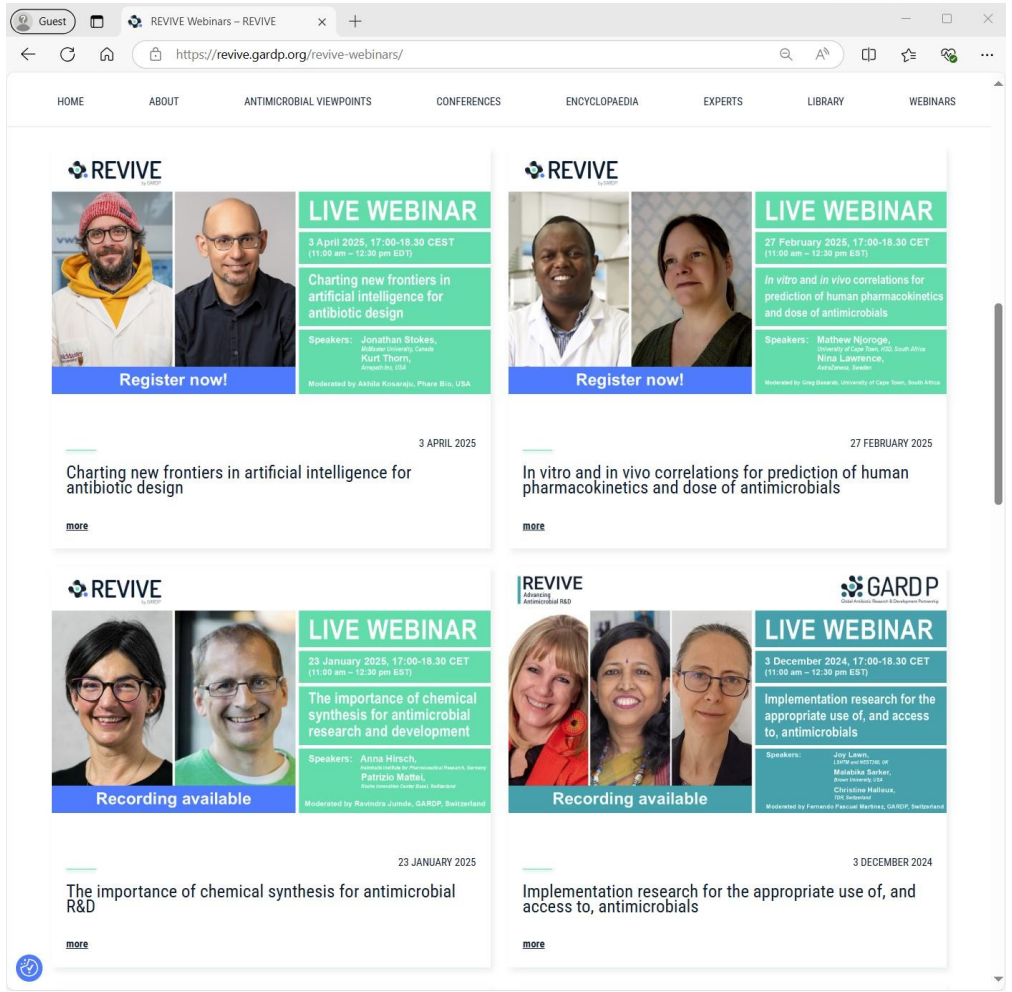
Host: Victor Kouassi

3 April 2025

REVIVE
by GARDP

Capture essential R&D technical knowledge and share expertise with the global community through the REVIVE website (revive.gardp.org).

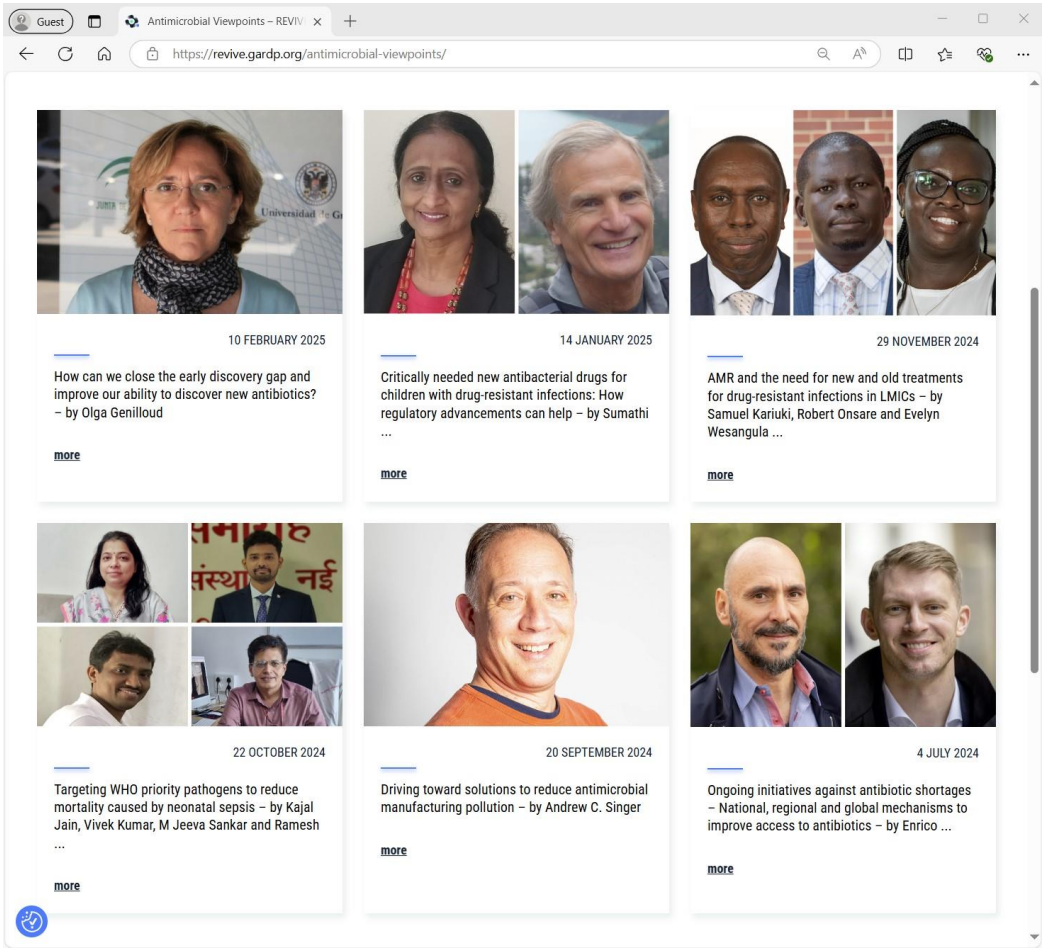
Webinar recordings



revive.gardp.org/webinars

Antimicrobial Viewpoints

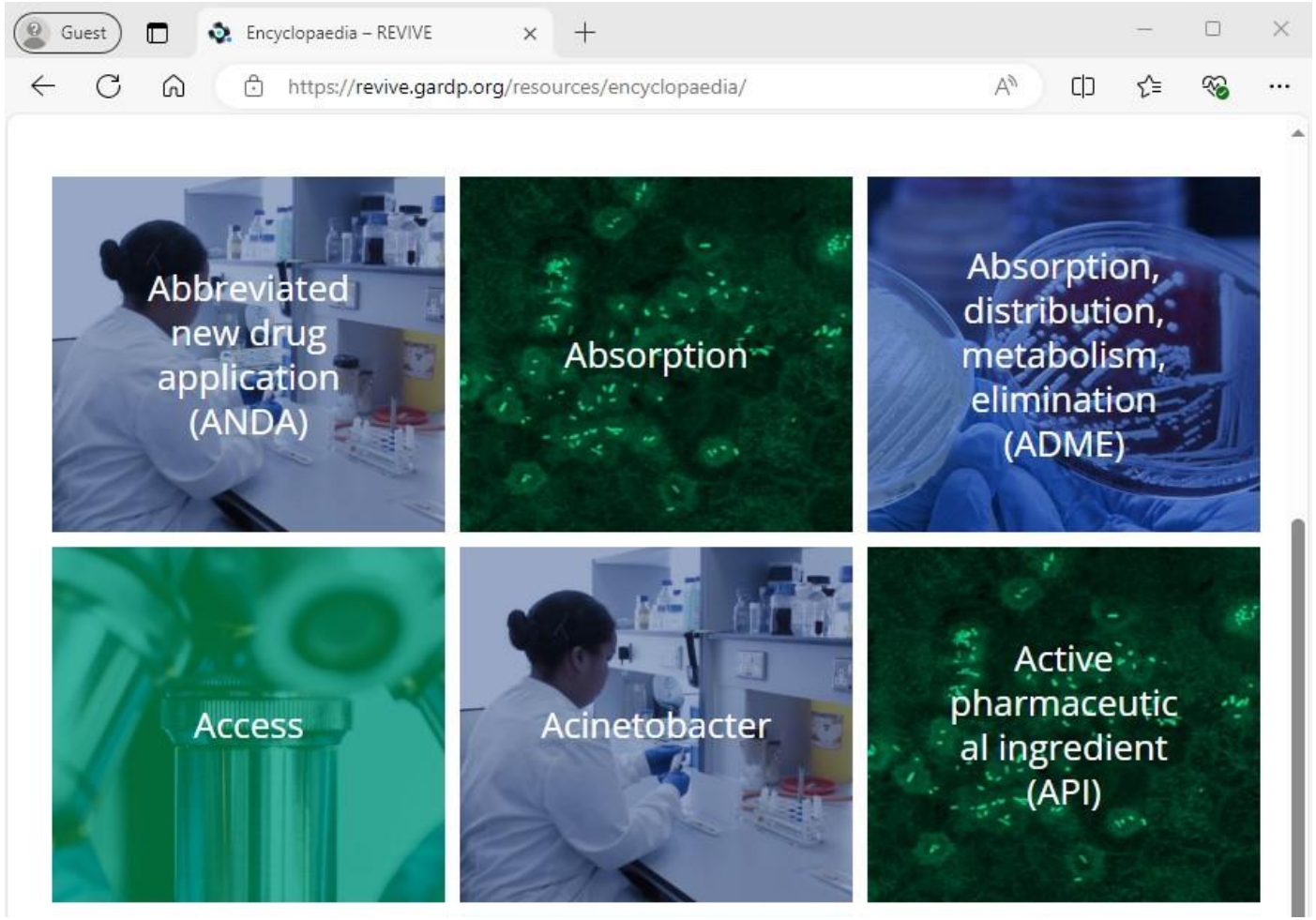
REVIVE by GARDP



revive.gardp.org/antimicrobial-viewpoints

Antimicrobial Encyclopaedia

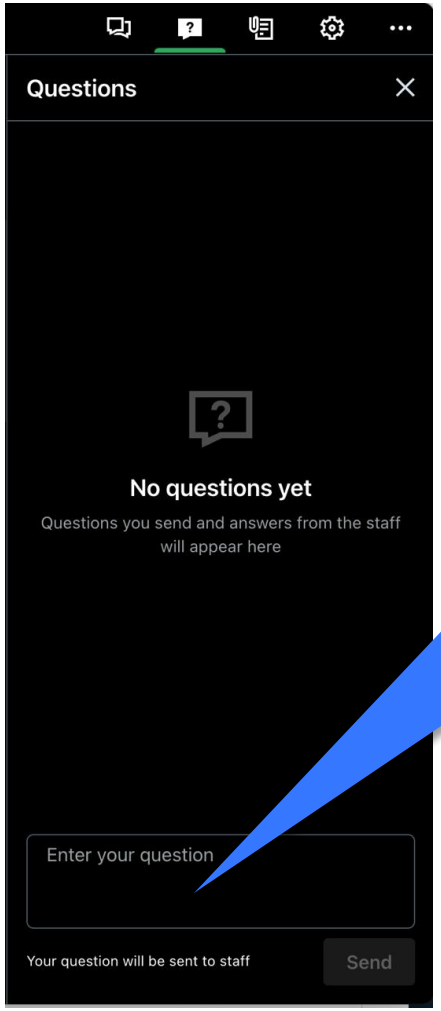
REVIVE by GARDP



revive.gardp.org/resources/encyclopaedia

How to submit your questions

If your question is addressed to a specific speaker, please include their name when submitting the question.



Please submit your questions through the box provided after clicking the ‘questions’ button. We will review all questions and respond to as many as possible after the presentation.

Today’s speakers

REVIVE by GARDP

Charting new frontiers in artificial intelligence for antibiotic design

Moderator:
Akhila Kosaraju
Phare Bio

Jonathan Stokes
McMaster University

Kurt Thorn
Arrepath, Inc.

REVIVE
by GARDP

Jonathan Stokes

Jonathan Stokes is an assistant professor in the Department of Biochemistry and Biomedical Sciences at McMaster University, Canada. He is also co-founder and Chief Scientific Officer of Stoked Bio.

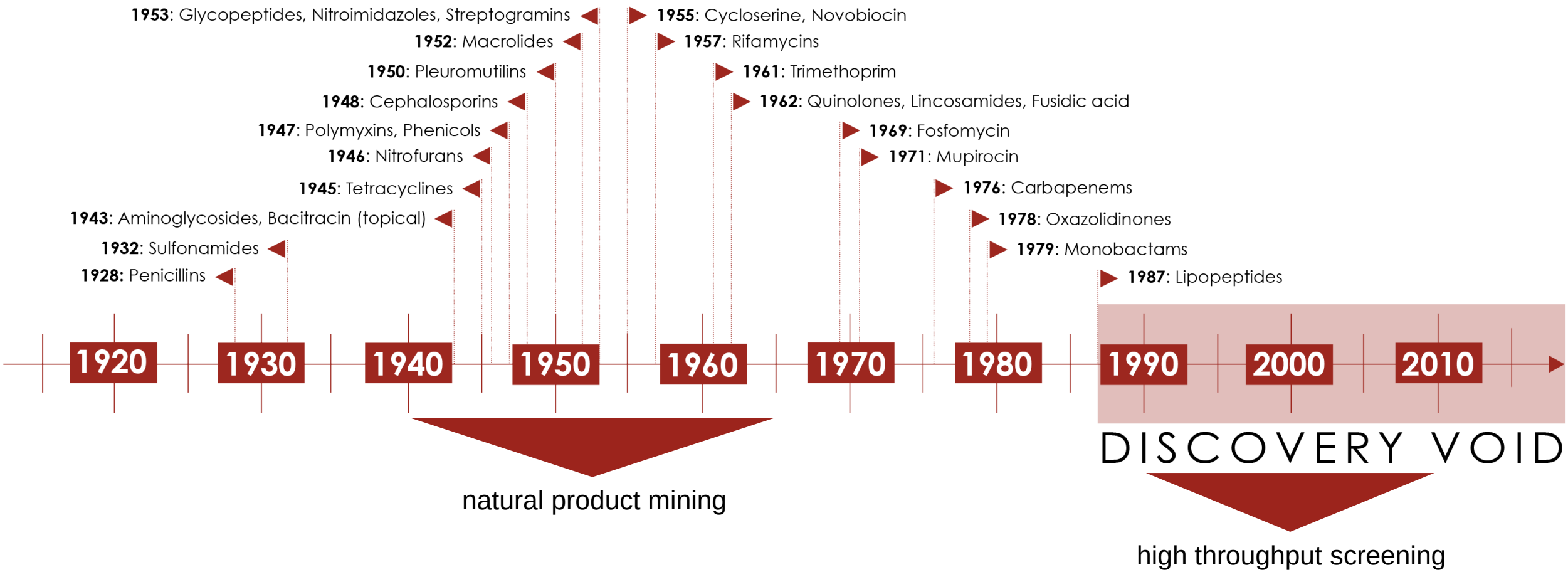
His research group focuses on the development and implementation of machine learning methods for drug discovery and design. They currently work on discovering novel antibiotics to overcome drug-resistant bacteria and new molecules to treat an aggressive form of brain cancer called glioblastoma.

Jonathan received his PhD in antimicrobial chemical biology in 2016 from McMaster University. From 2017 to 2021, he was a Banting postdoctoral fellow at the Broad Institute of MIT and Harvard in the USA.

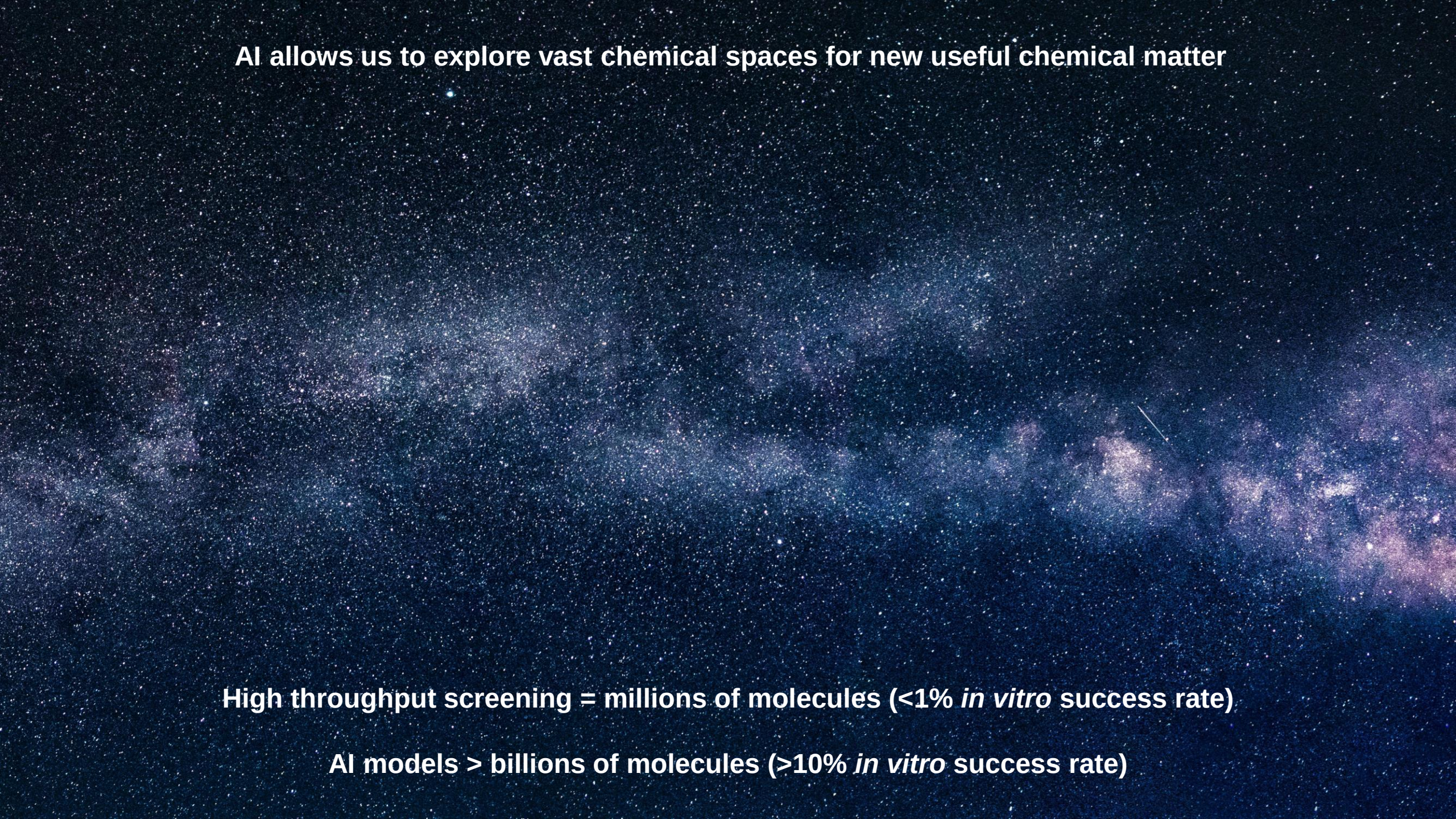
How should we be using AI to help us fight superbugs?

Jon Stokes
Assistant Professor
Department of Biochemistry and Biomedical Sciences
McMaster University

We are in an antibiotic discovery void

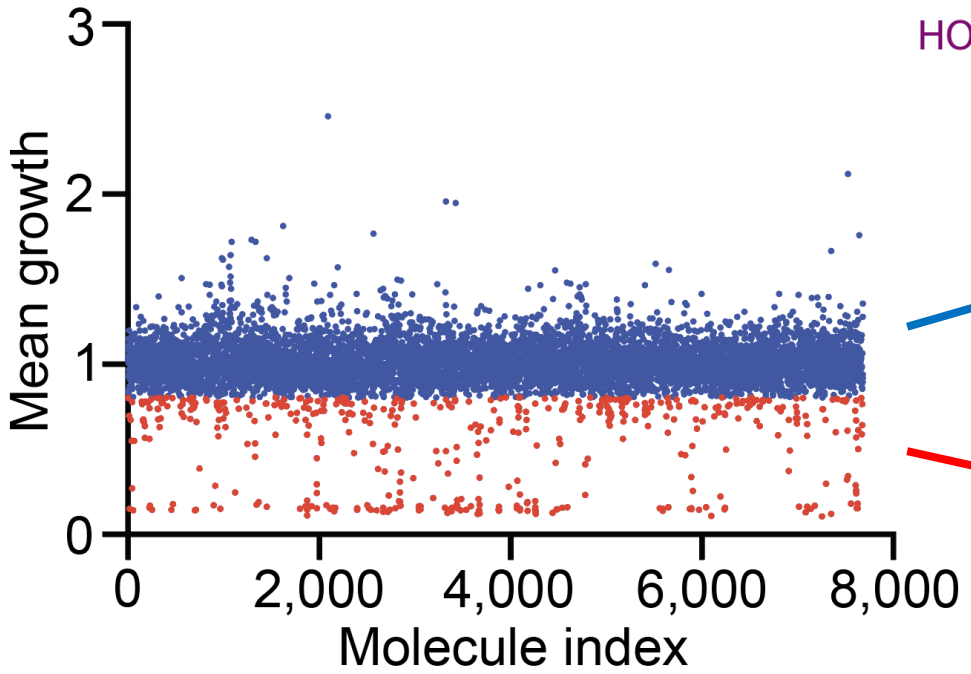


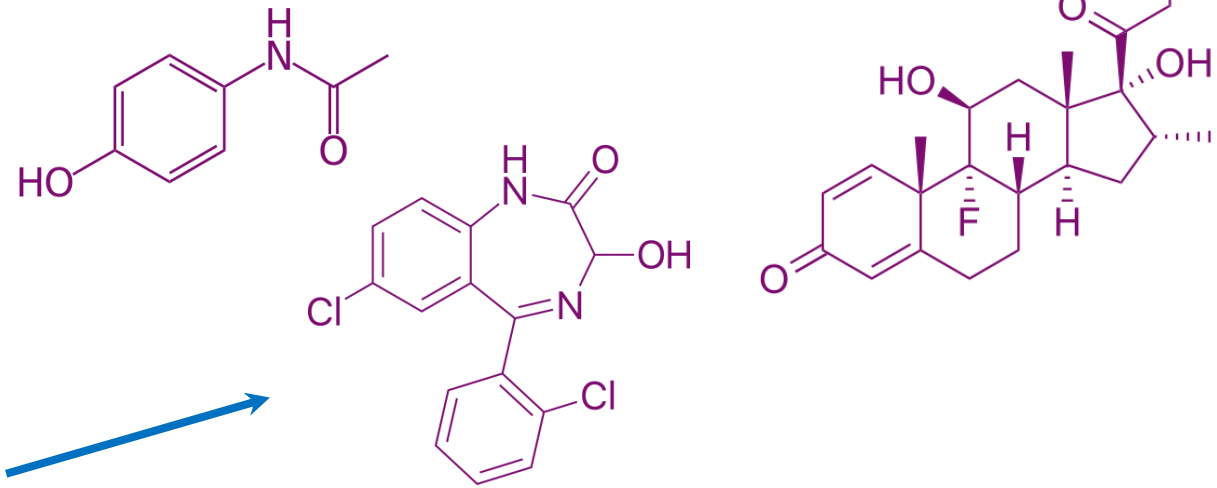
ReAct


AI allows us to explore vast chemical spaces for new useful chemical matter
High throughput screening = millions of molecules (<1% *in vitro* success rate)
AI models > billions of molecules (>10% *in vitro* success rate)

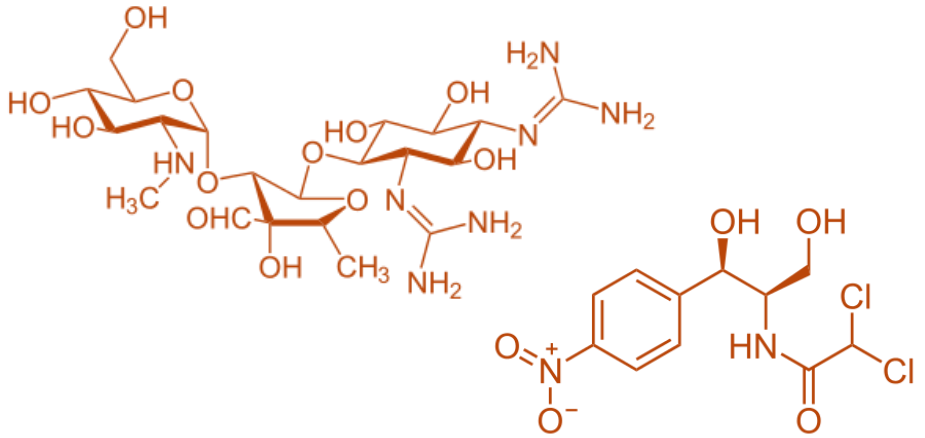
Acinetobacter baumannii is a challenging nosocomial gram-negative pathogen

We performed a screen of ~7,500 diverse molecules against *A. baumannii*





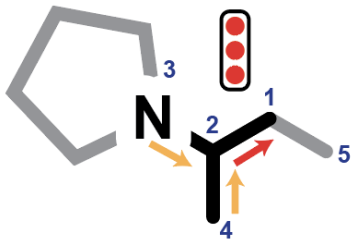
0

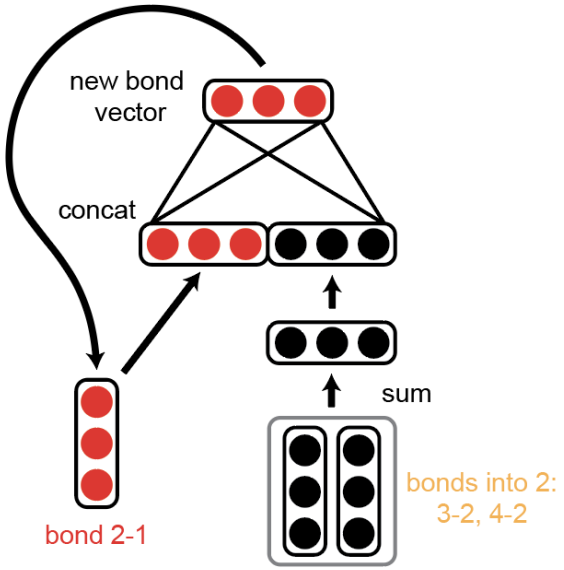


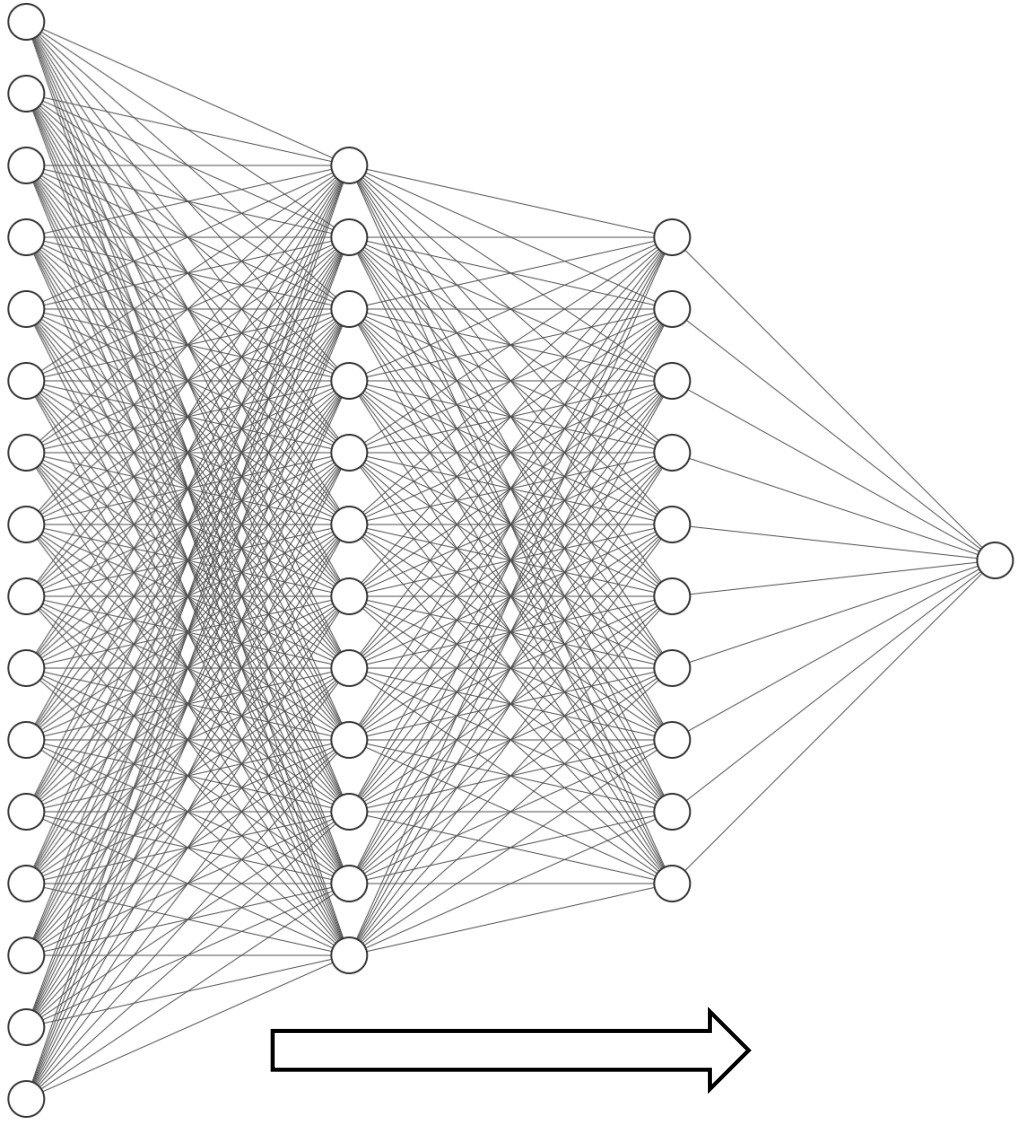
1

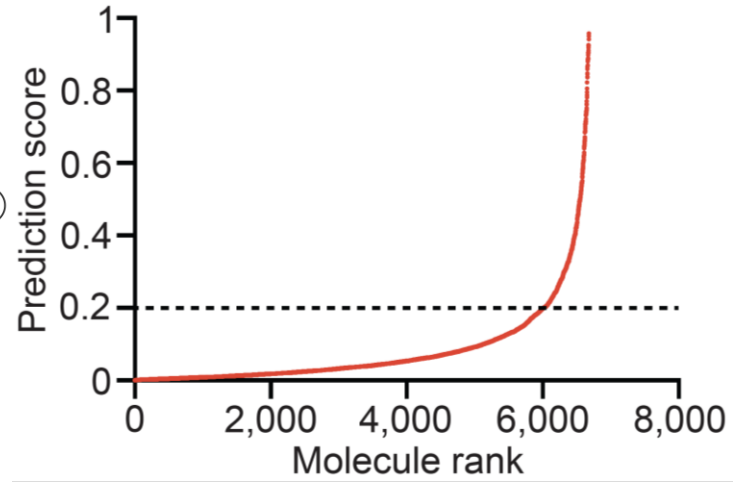
Liu, Catacutan, Rathod *et al.* Nature Chemical Biology (2023)

Using graph neural networks for antibiotic prediction against *A. baumannii*



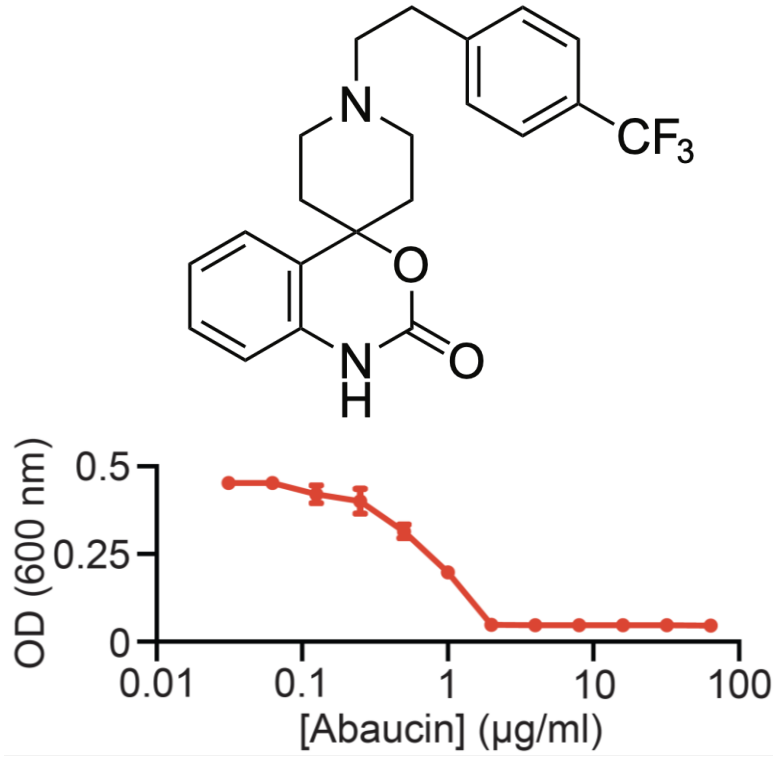


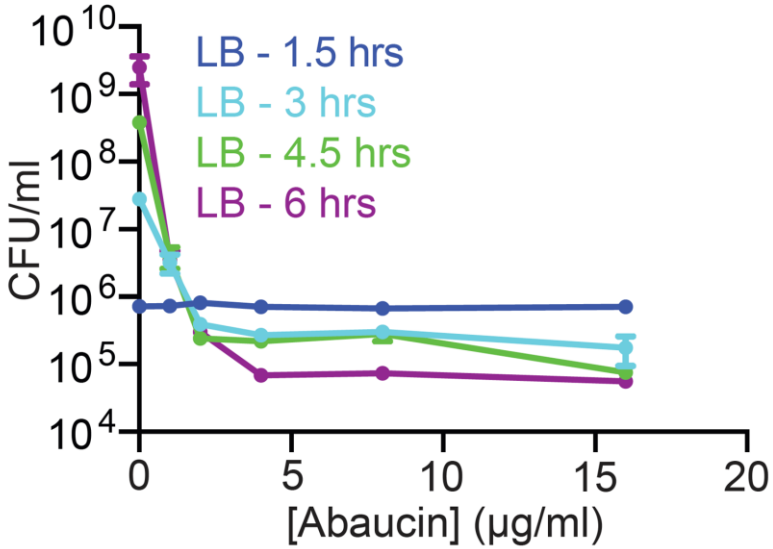




Input layer (vector representations of molecules)

Abaucin is active against *A. baumannii in vitro*





nature chemical biology

Article

https://doi.org/10.1038/s41589-023-01349-8

Deep learning-guided discovery of an antibiotic targeting *Acinetobacter baumannii*

Received: 25 March 2022

Accepted: 25 April 2023

Published online: 25 May 2023

Check for updates

Gary Liu[1,10], Denise B. Catacutan[1,10], Khushi Rathod[1,10], Kyle Swanson[2], Wengong Jin[2], Jody C. Mohammed[1], Anush Chiappino-Pepe[3,4], Saad A. Syed[5], Meghan Fragis[1,6], Kenneth Rachwalski[1], Jakob Magolan[1,6], Michael G. Surette[5], Brian K. Coombes[1], Tommi Jaakkola[2], Regina Barzilay[2,7], James J. Collins[3,7,8,9] ✉ & Jonathan M. Stokes[1] ✉

CBC

CNN

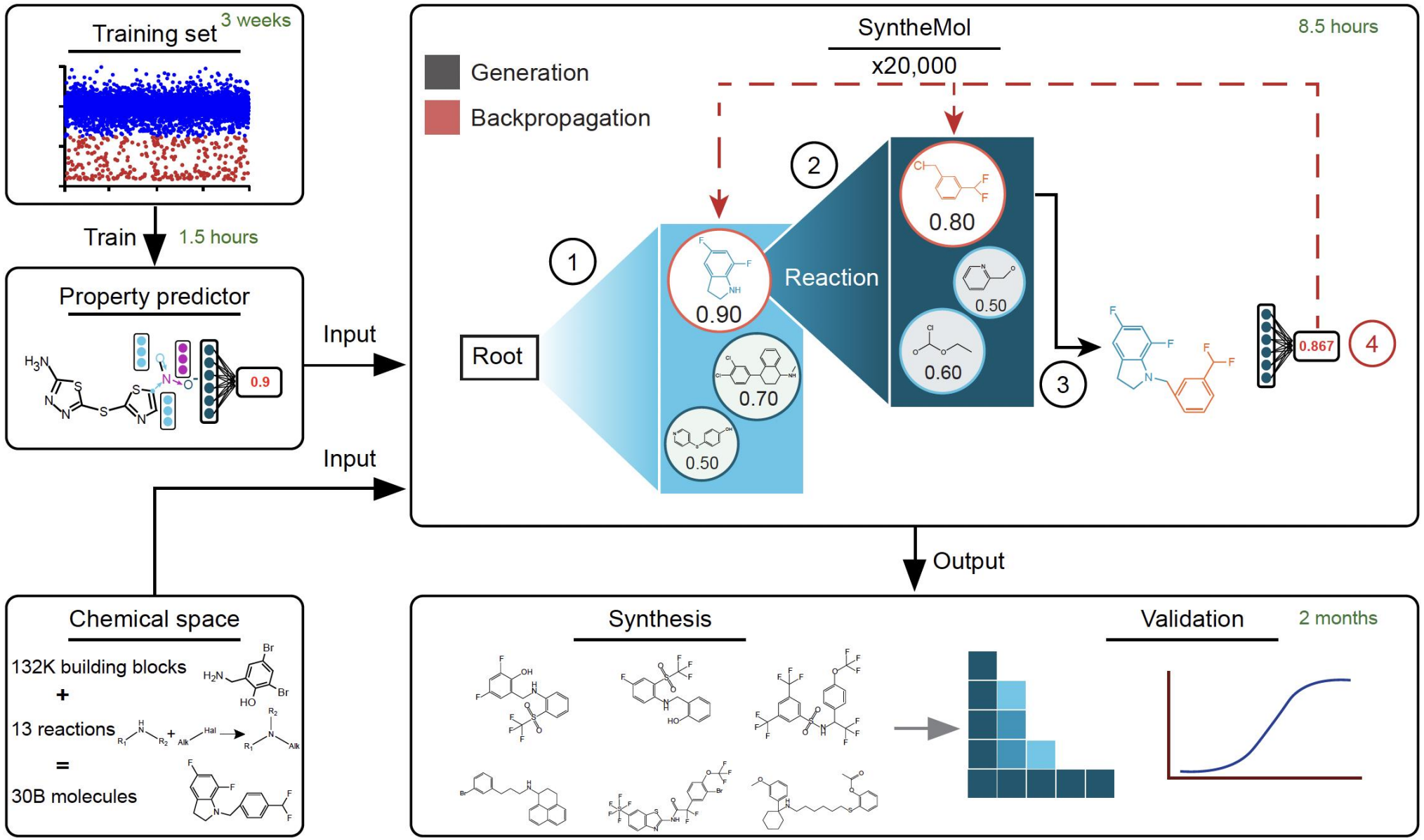
The Guardian

sky news

BBC

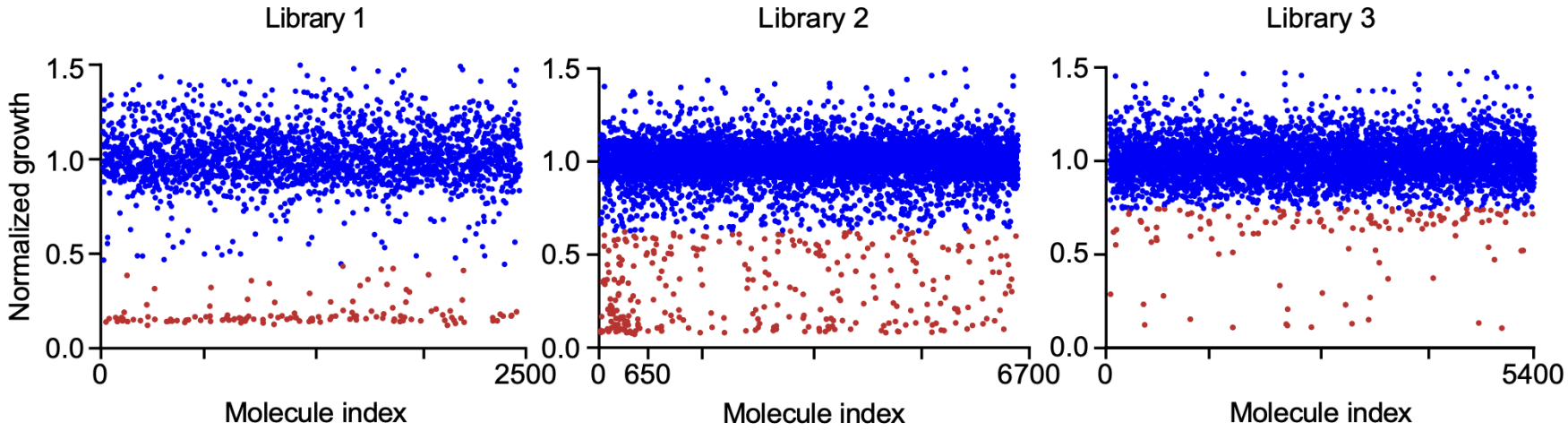
Discriminative models can explore ~10^{9} molecules, but drug-like chemical space is ~$10^{idontknow}$ molecules

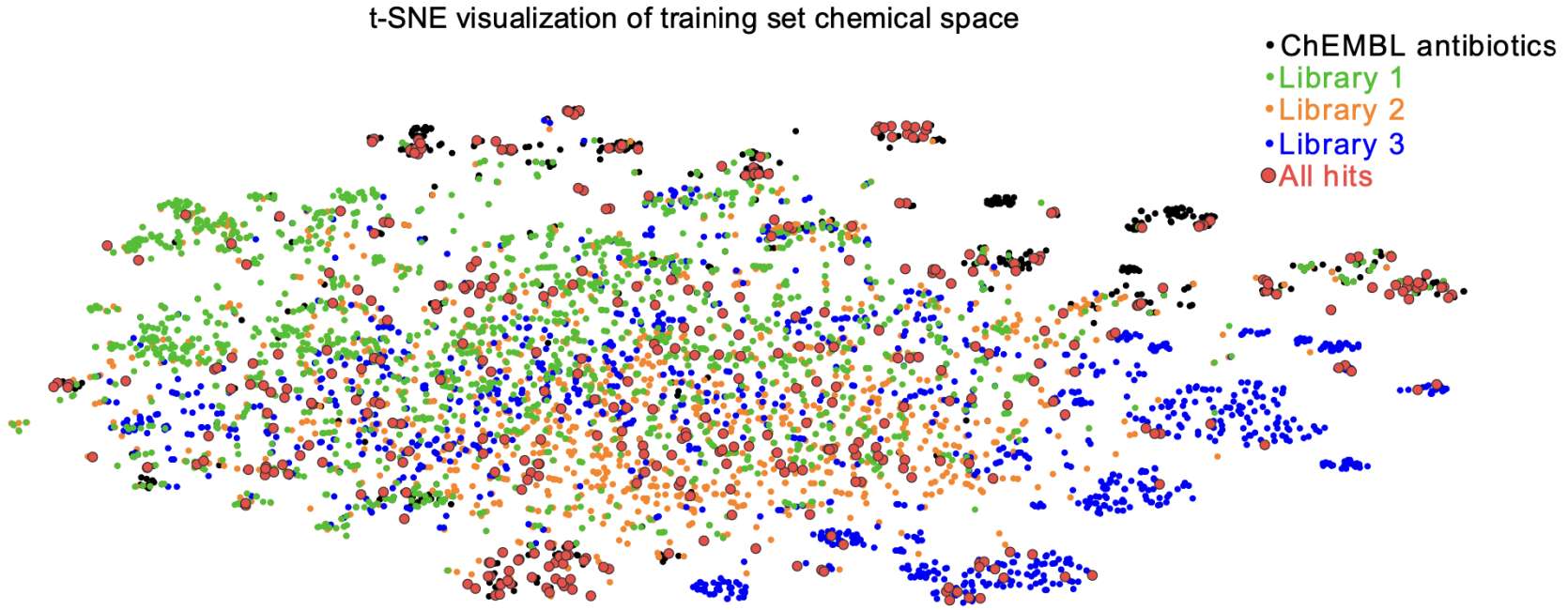
SyntheMol-MCTS is a generative AI algorithm for *de novo* antibiotic design tasks



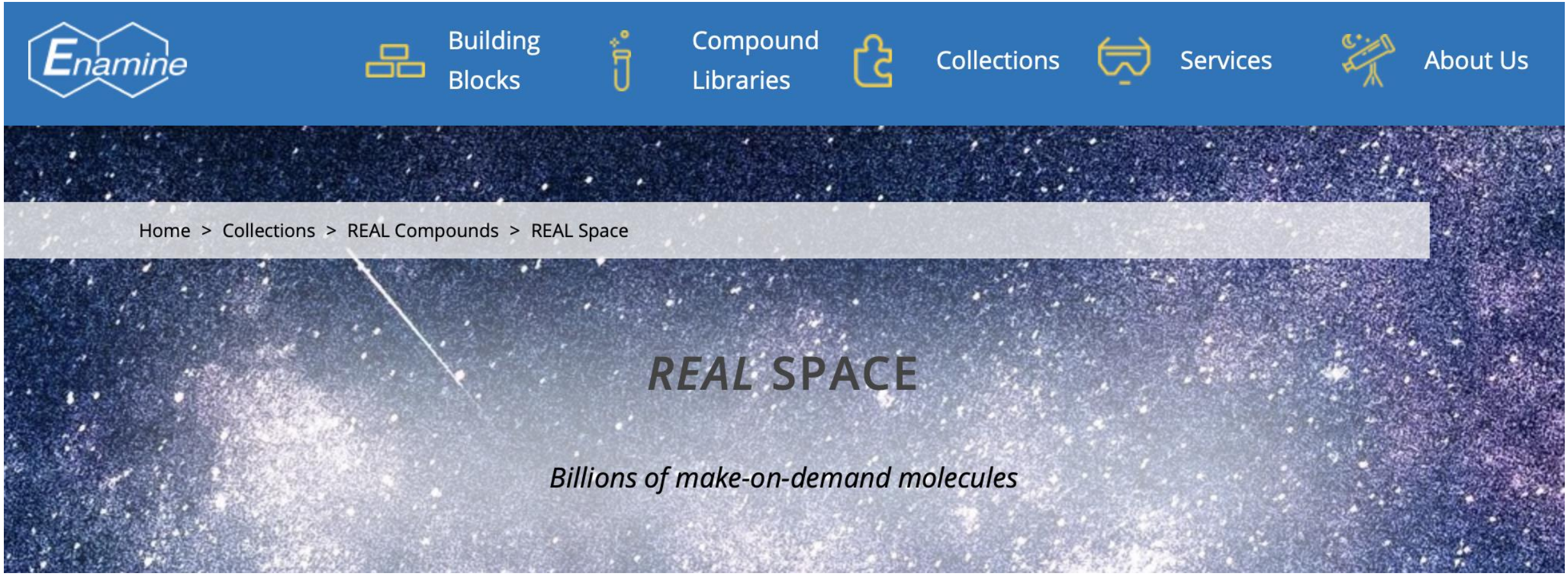
Swanson, Liu *et al.* Nature Machine Intelligence (2024)

We trained antibiotic property predictors on three chemical libraries totaling ~13,500 molecules

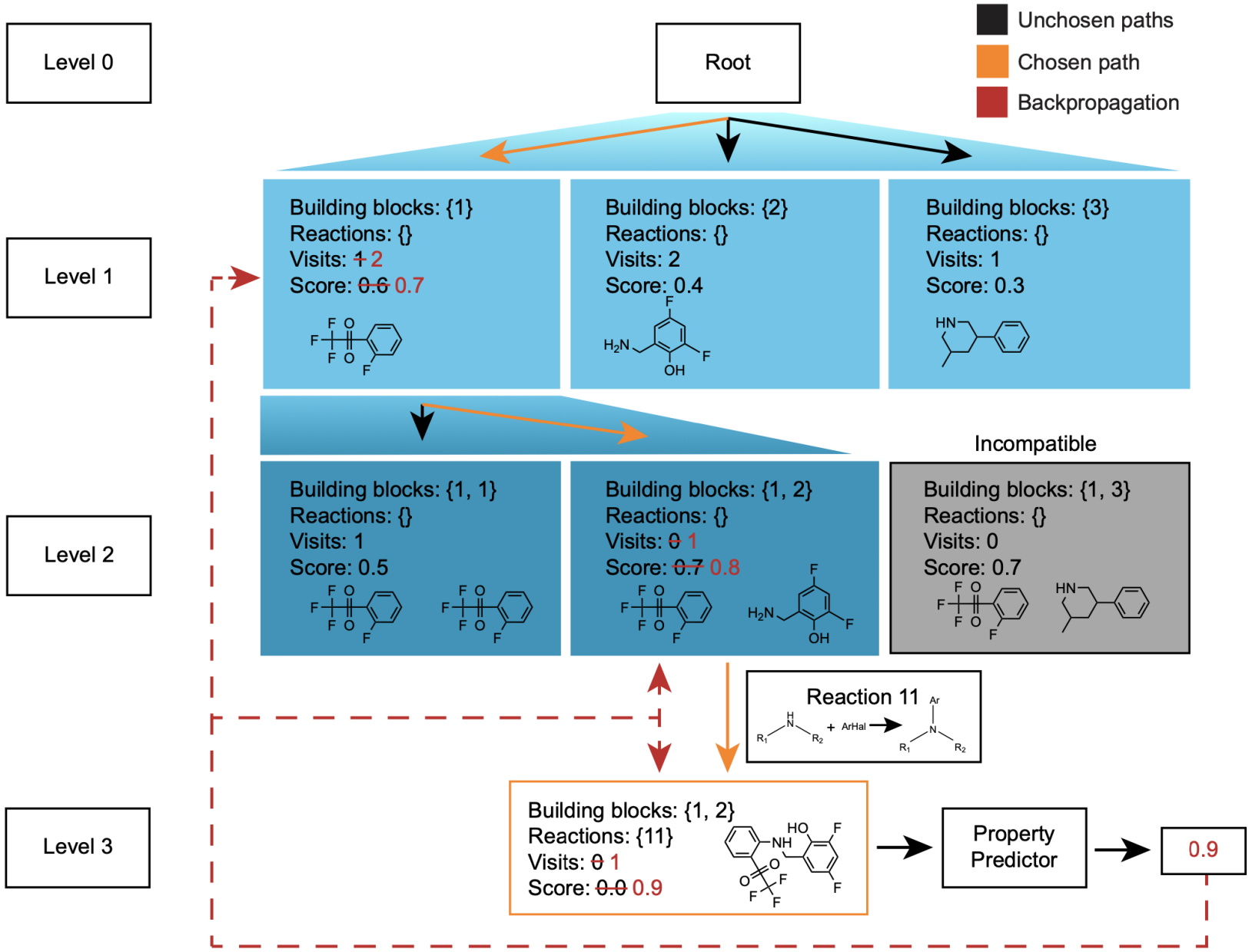




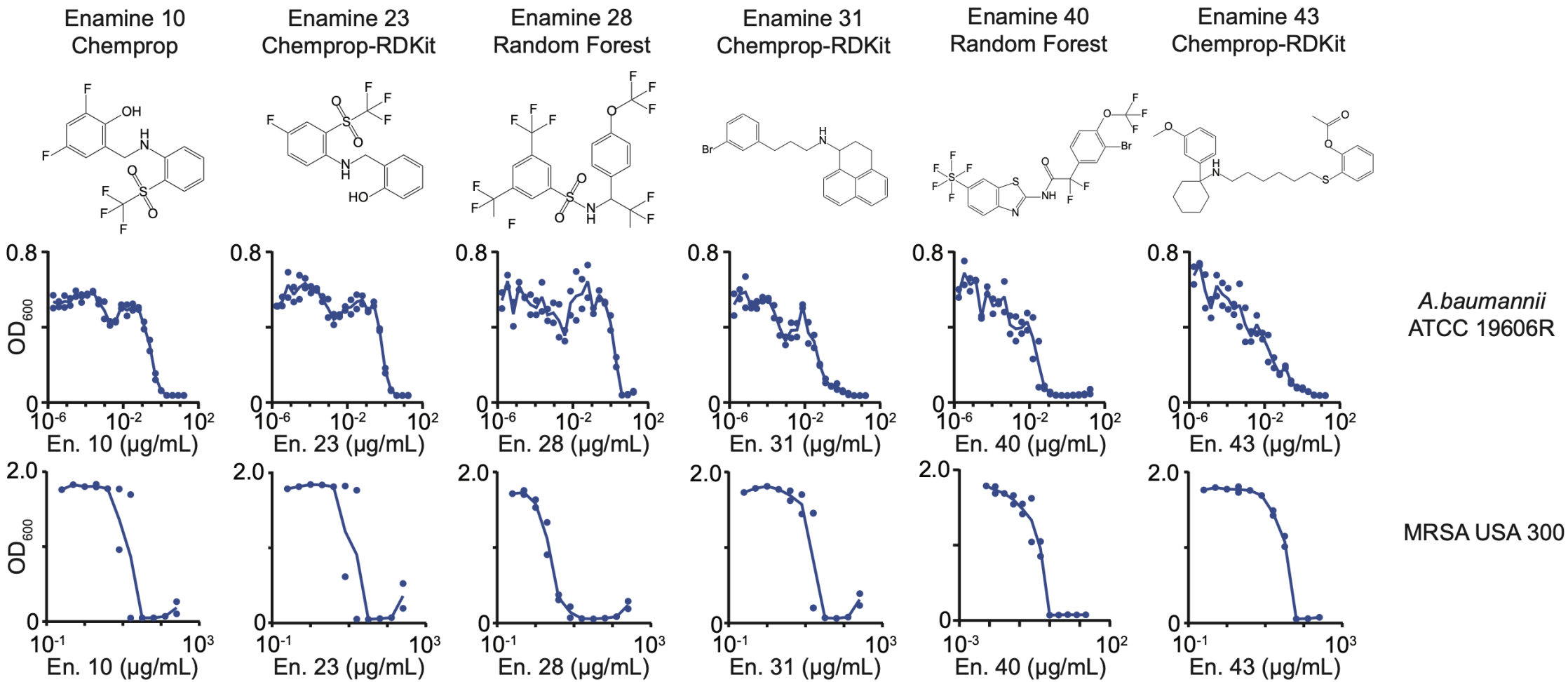
We leveraged ~132,000 molecular fragments and 13 reactions from the REAL space to generate novel molecules



Our molecule design process in a nutshell



Six of 58 synthesized molecules displayed activity *in vitro* against an array of ESKAPE pathogens



nature machine intelligence

Article

https://doi.org/10.1038/s42256-024-00809-7

Generative AI for designing and validating easily synthesizable and structurally novel antibiotics

Received: 10 March 2023

Accepted: 8 February 2024

Kyle Swanson [1,4], Gary Liu [2,4], Denise B. Catacutan [2], Autumn Arnold [2], James Zou [1,3] ✉ & Jonathan M. Stokes [2] ✉

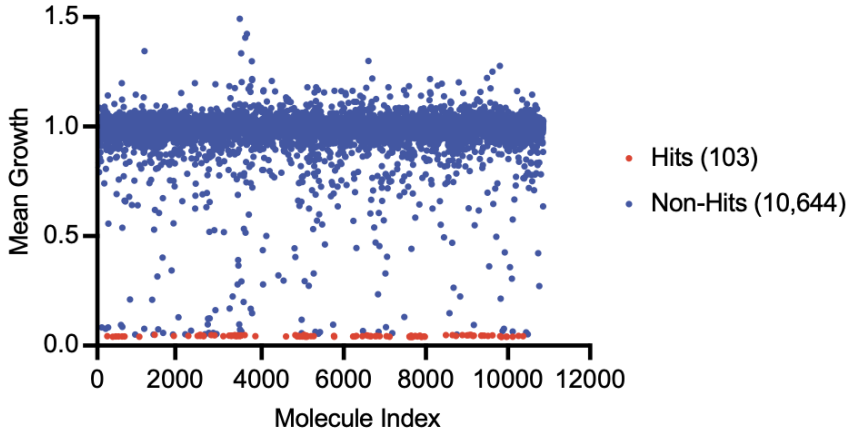
The New York Times

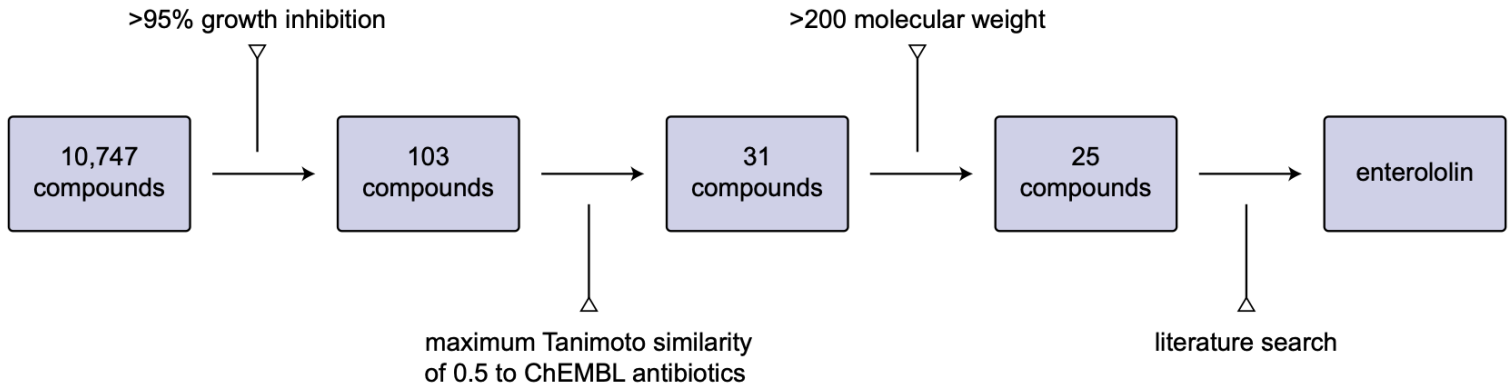
AI BUSINESS

TheScientist
EXPLORING LIFE, INSPIRING INNOVATION

We can use AI for stuff other than hit identification!

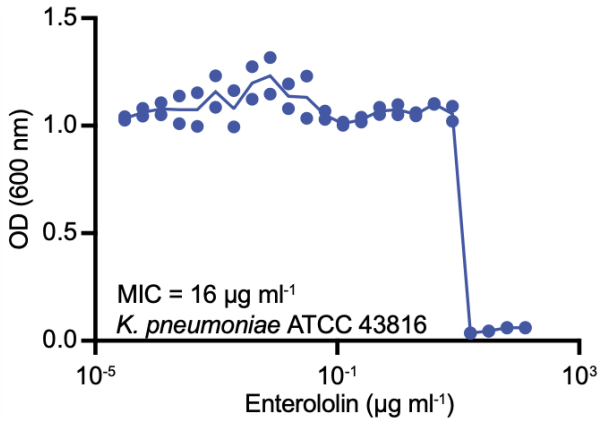
Discovery and AI-guided mechanistic elucidation of a novel Enterobacteriaceae-specific antibiotic

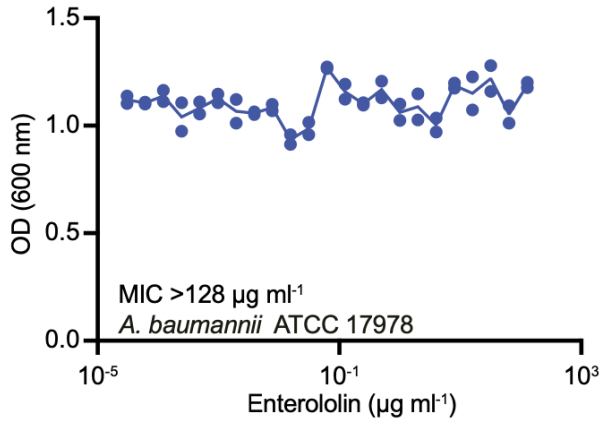


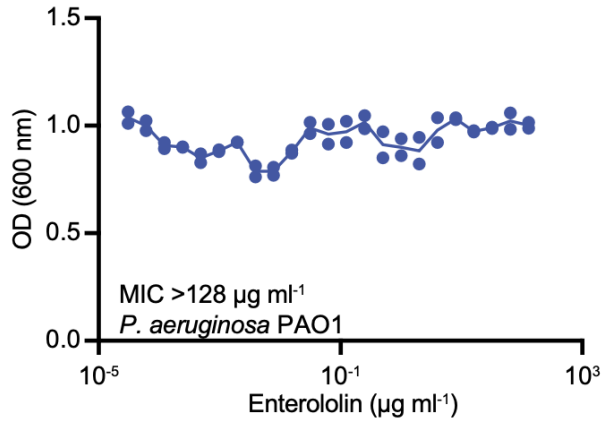


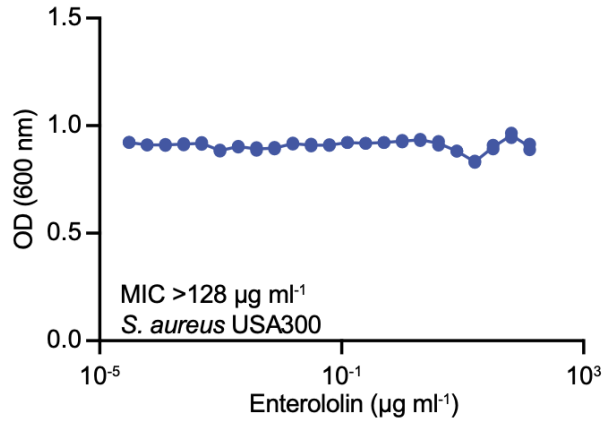
Catacutan *et al.* In revision

Enterololin has Enterobacteriaceae specific antibacterial activity – lab pathogens

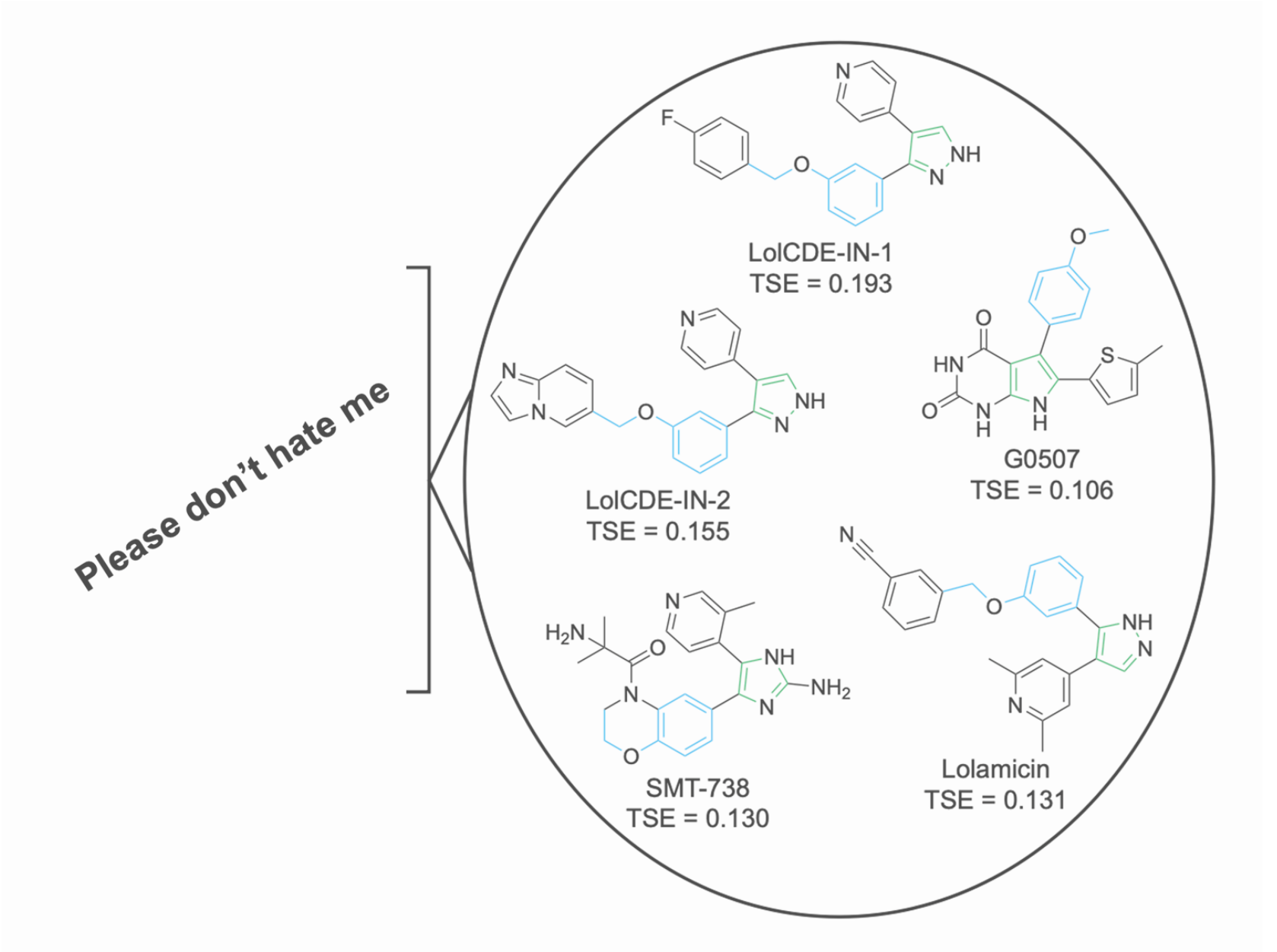




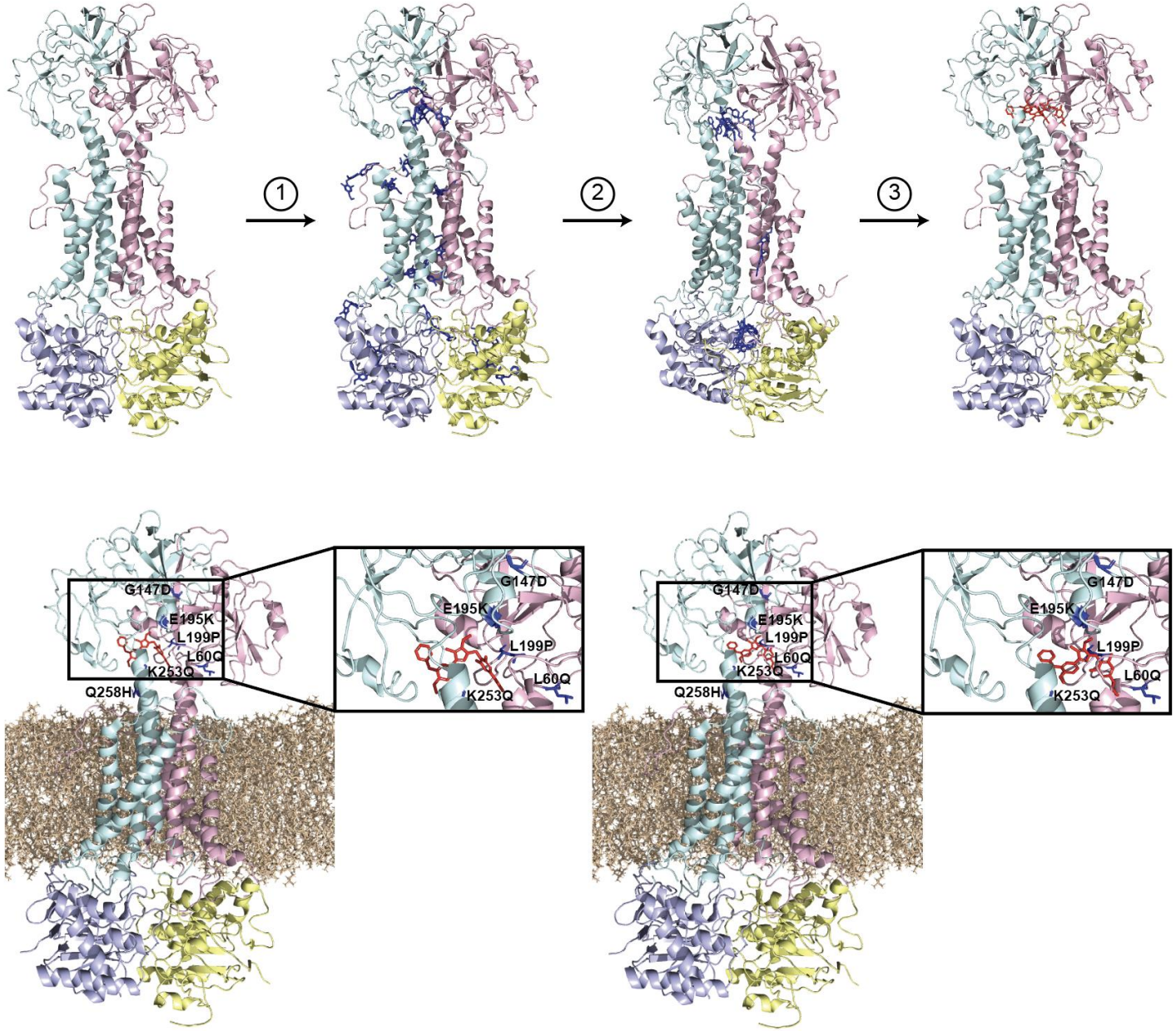




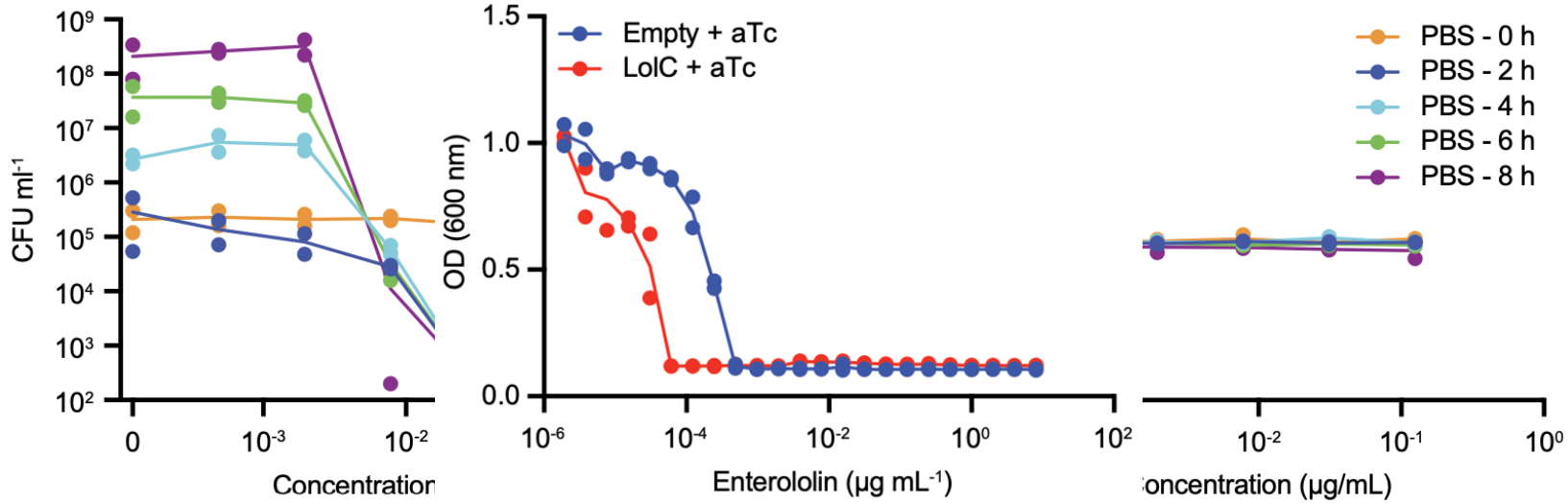
Some structural features of enterololin are (somewhat?) consistent with perturbation of lipoprotein trafficking

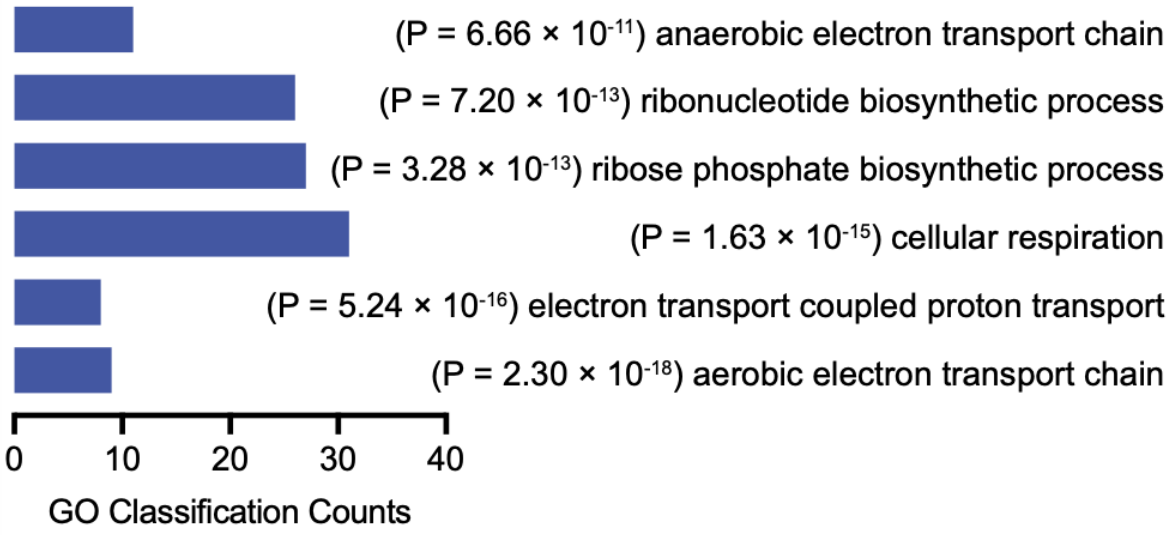


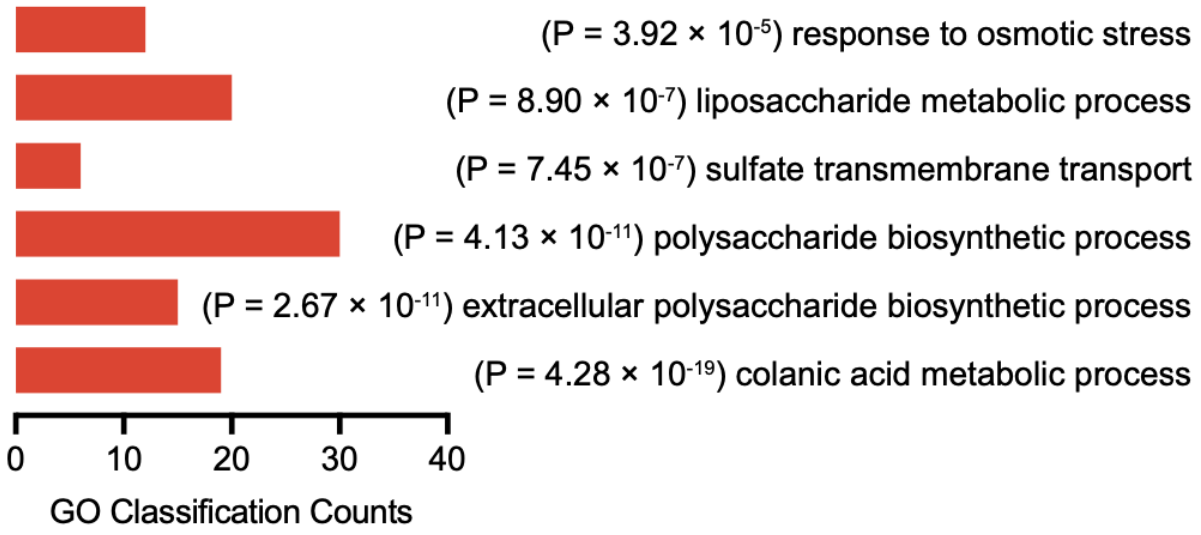
Machine learning model predictions of enterololin binding to the LolCDE complex – 100 seconds



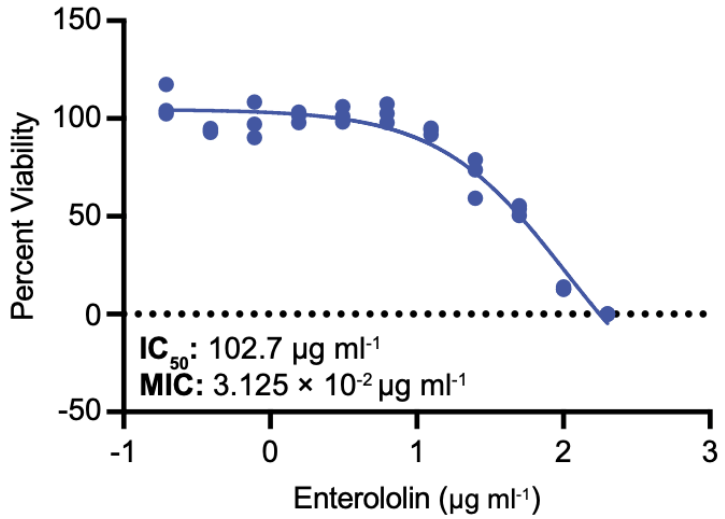
Wet lab validation of the mechanism of action of enterololin – like 6 months

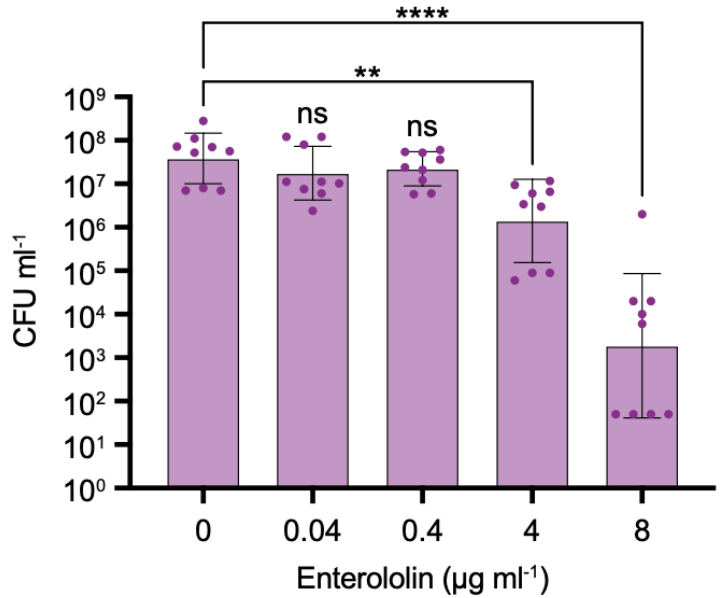


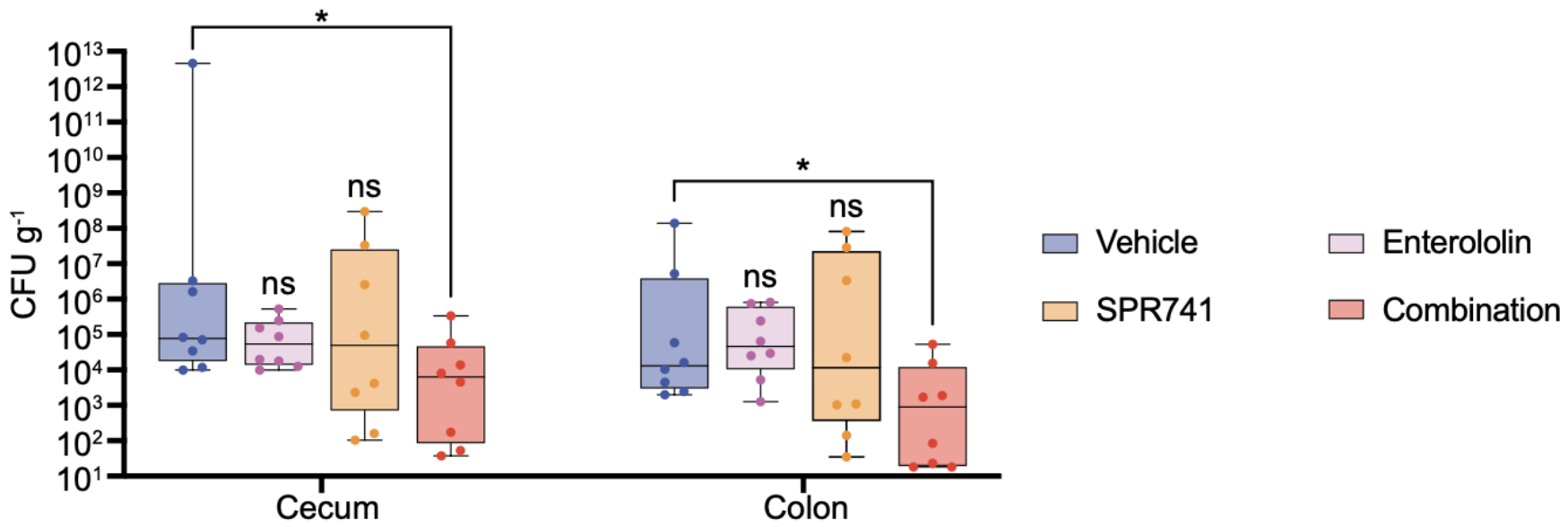




Enterololin shows activity in mouse models of AIEC infection







We’ve built a free online tool that you can all use if you’re interested

Arnold *et al.* In preparation

stokesjm@mcmaster.ca

INNOVATION.CA
CANADA FOUNDATION FOR INNOVATION | FONDATION CANADIENNE POUR L'INNOVATION

David Braley Centre for Antibiotic Discovery

Michael G. DeGroote Institute for Infectious Disease Research

Heersink Signature Scholarship

braintumour foundation of CANADA

Weston Family Foundation

CIHR IRSC

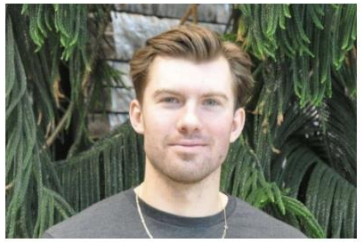
NSERC CRSNG

Jeremie Alexander
PhD Student
in silico | in vitro
alexaj16@mcmaster.ca

Autumn Arnold
MSc Student
in silico | in vitro
arnolp3@mcmaster.ca

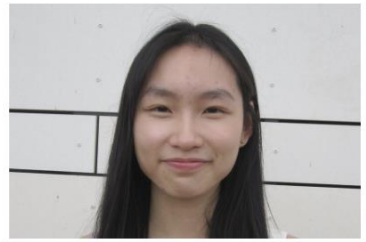
Denise Catacutan
PhD Student
in vitro | in vivo
catacud@mcmaster.ca

Gary Liu
PhD Student
in silico | in vitro
liug34@mcmaster.ca

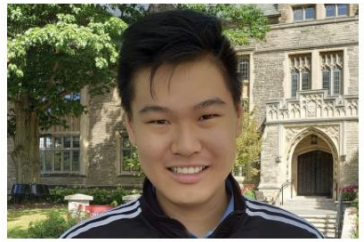


Stewart McLellan
MSc Student
in silico | in vitro
mclellas@mcmaster.ca

Richard Naiberg
Scientist
in silico | in vitro
rnaiber1@jhu.edu



Alice Pao
Undergraduate Student
in vitro | in vivo
paos1@mcmaster.ca



Wesley Ta
Undergraduate Student
in silico | in vitro
taw1@mcmaster.ca

Jennifer Tindall
Undergraduate Student
in silico | in vitro
tindallj@mcmaster.ca

Jon Stokes
Assistant Professor
PI
stokesjm@mcmaster.ca

Curly Fry
Dog
Being fun

Piglet
Tiny Dog
Being cute

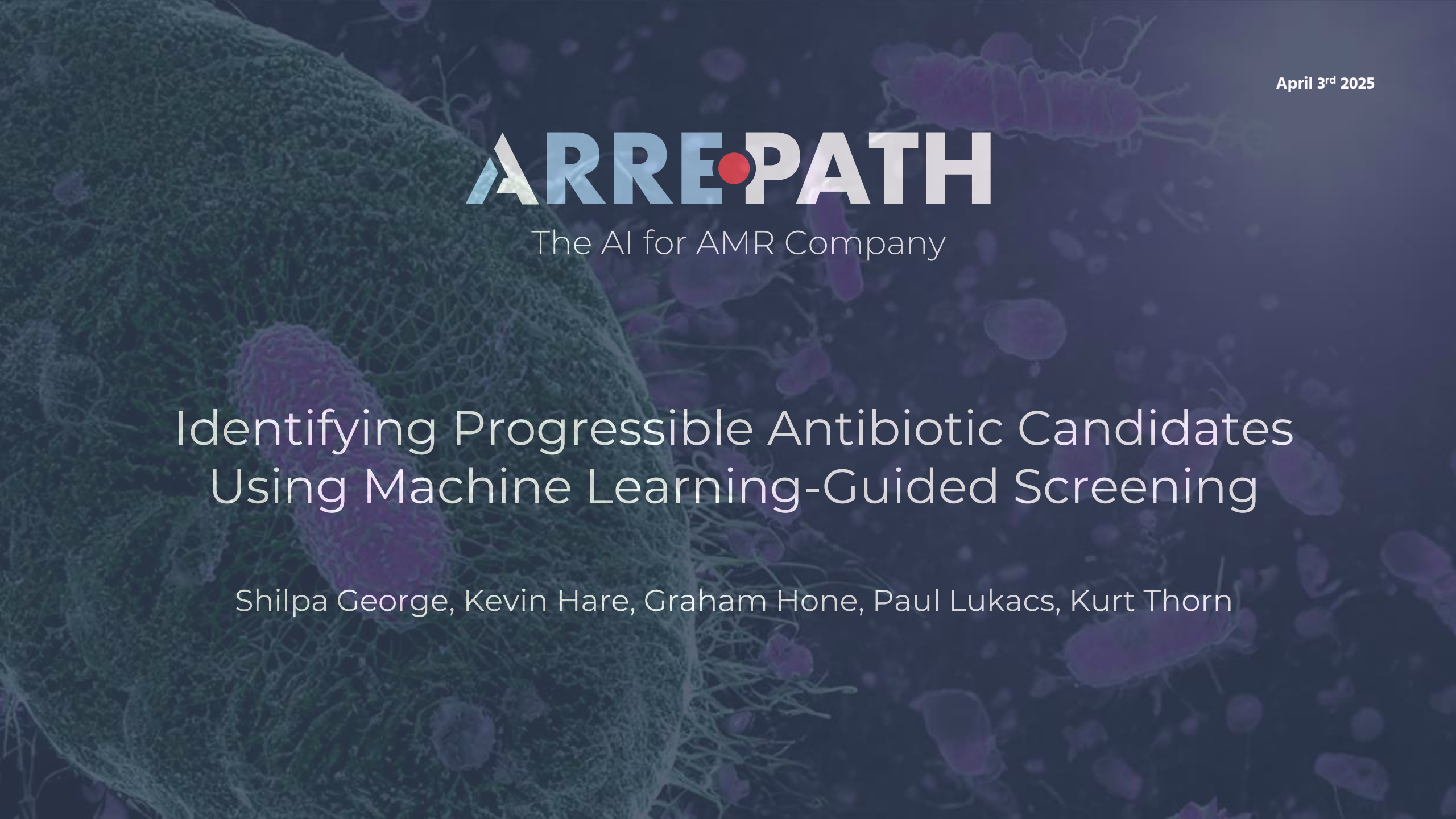
Kurt Thorn

Kurt Thorn is Chief Technology Officer at ArrePath, a biotech company that uses a combination of human intelligence and a cutting-edge AI/ML platform to discover and develop small molecule therapeutics to address significant unmet medical needs. Here, he leads efforts to develop tools to accelerate drug discovery for novel antibacterials.

Prior to this, Kurt was senior director of data science at Zymergen, a synthetic biology company. Before transitioning to industry, Kurt was a research fellow at Havard University and an associate professor at University of California, San Francisco (UCSF) and Director of the Nikon Imaging Center at UCSF.

Kurt has a PhD in biophysics from UCSF and received his bachelor's degree in chemistry from Princeton University.

arrepath.com/about/



April 3rd 2025

ARREPATH

The AI for AMR Company

Identifying Progressible Antibiotic Candidates Using Machine Learning-Guided Screening

Shilpa George, Kevin Hare, Graham Hone, Paul Lukacs, Kurt Thorn

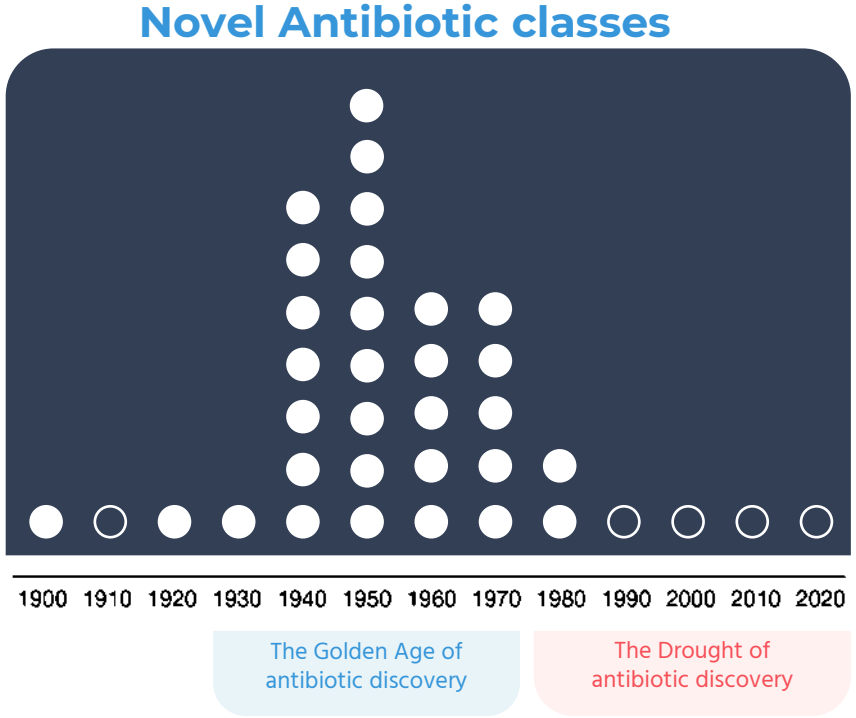
New approaches are needed to discover novel antibacterials

1.27 million deaths were caused by AMR worldwide in 2019

- Conventional (target-based) drug discovery approaches have largely failed to deliver new antibacterials, owing to the challenges of bacterial permeability and efflux (especially in Gram-negative bacteria).
- ArrePath is taking a three-pronged approach to address this problem:

1. ML models to predict novel whole-cell active compounds
2. Phenotypic screening of predicted whole cell active compounds
3. Image-based phenotyping to identify novel mechanisms of action

Antimicrobial resistance is a severe and growing problem



Pew Charitable Trust available at https://wellcome.org/news/its-time-fix-antibiotic-market
Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis, The Lancet, Jan 2022

ArrePath's goal: Efficiently discover whole-cell active antibiotics with novel mechanisms of action

Focus on Gram-negative bacteria

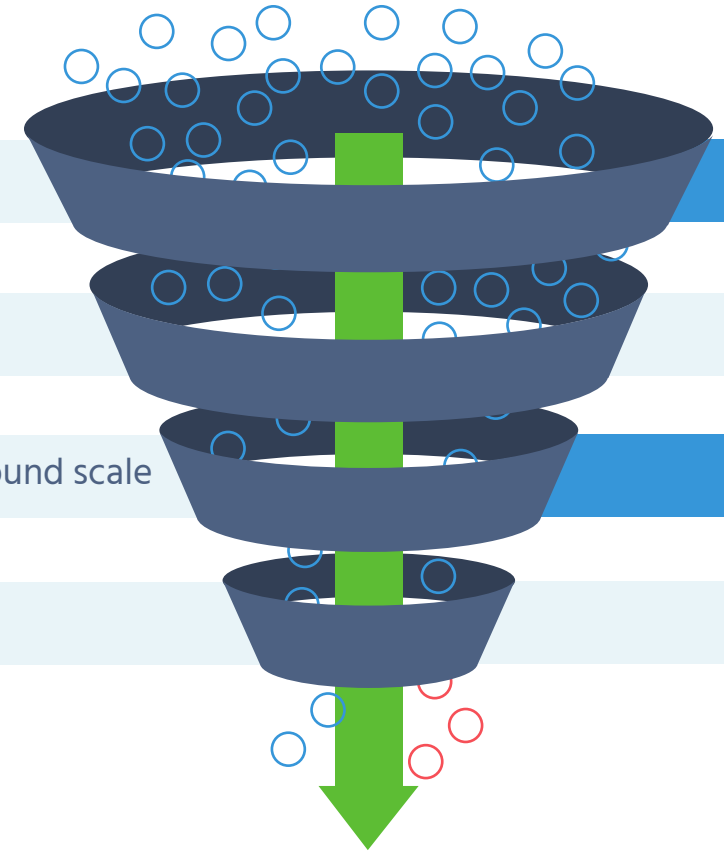
'Traditional' approach

Chemical feature selection

Phenotypic HTS: 100,000s compound scale

Traditional filters and profiling: 100-1000s compound scale

Chemical genetics: 10-100s compound scale



ArrePath's approach

ML-guided compound selection

Enrich for progressible compounds

Phenotypic HTS: 100,000s compound scale

Imaging: 100-1000s compound scale

Enrich for MoA novelty

Chemical genetics: 10-100s compound scale

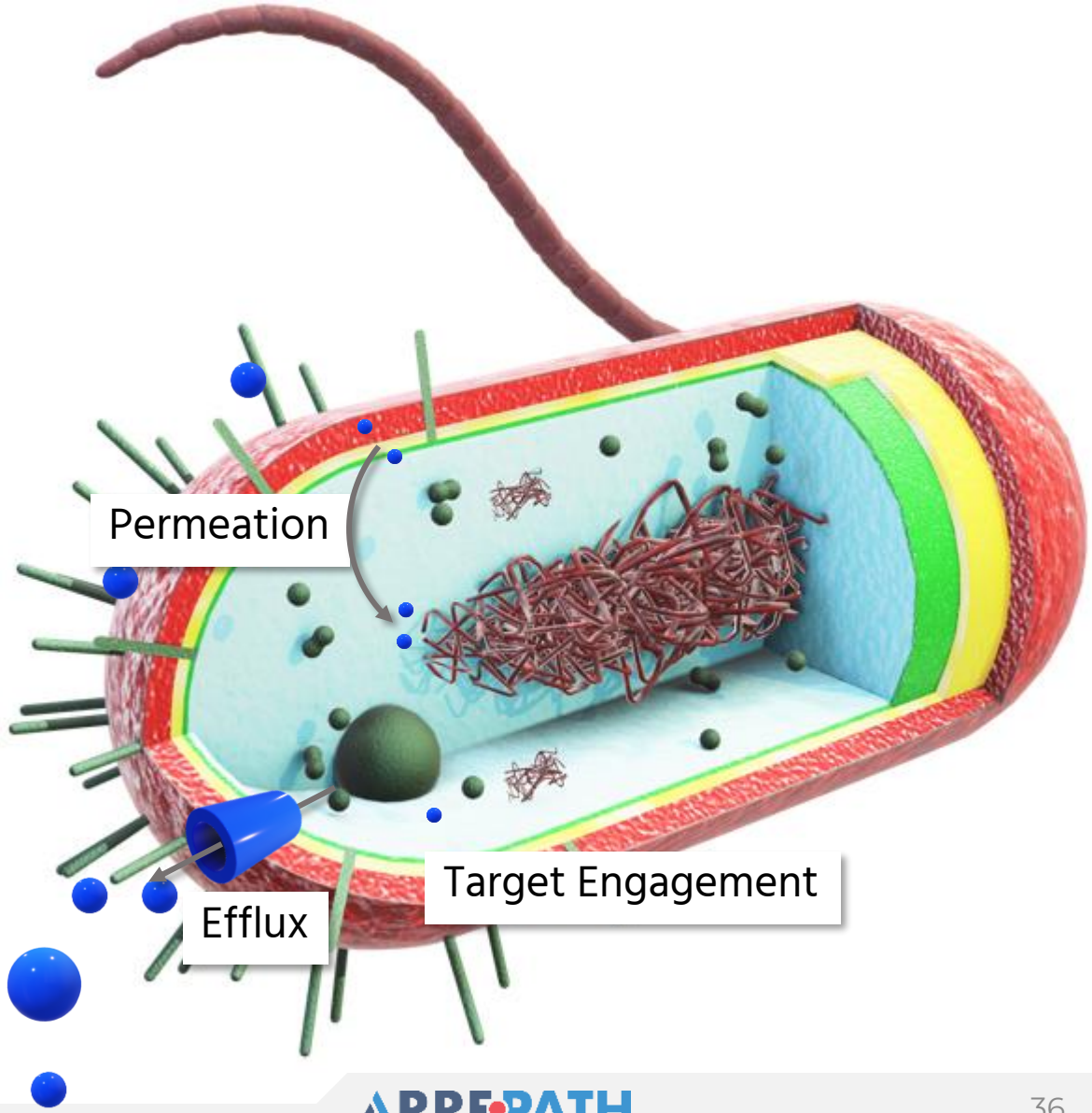
Hypothesis: ML models will learn target engagement, permeation, and efflux avoidance

Successful antibiotics must:

- Permeate the bacterial membrane(s) and cell wall
- Avoid efflux pumps
- Engage with the target

Our Hypothesis:

- A single ML model can learn chemical features associated with all three properties



Goals of ML-guided compound selection

Goal: maximize progressible hits by finding molecules that:

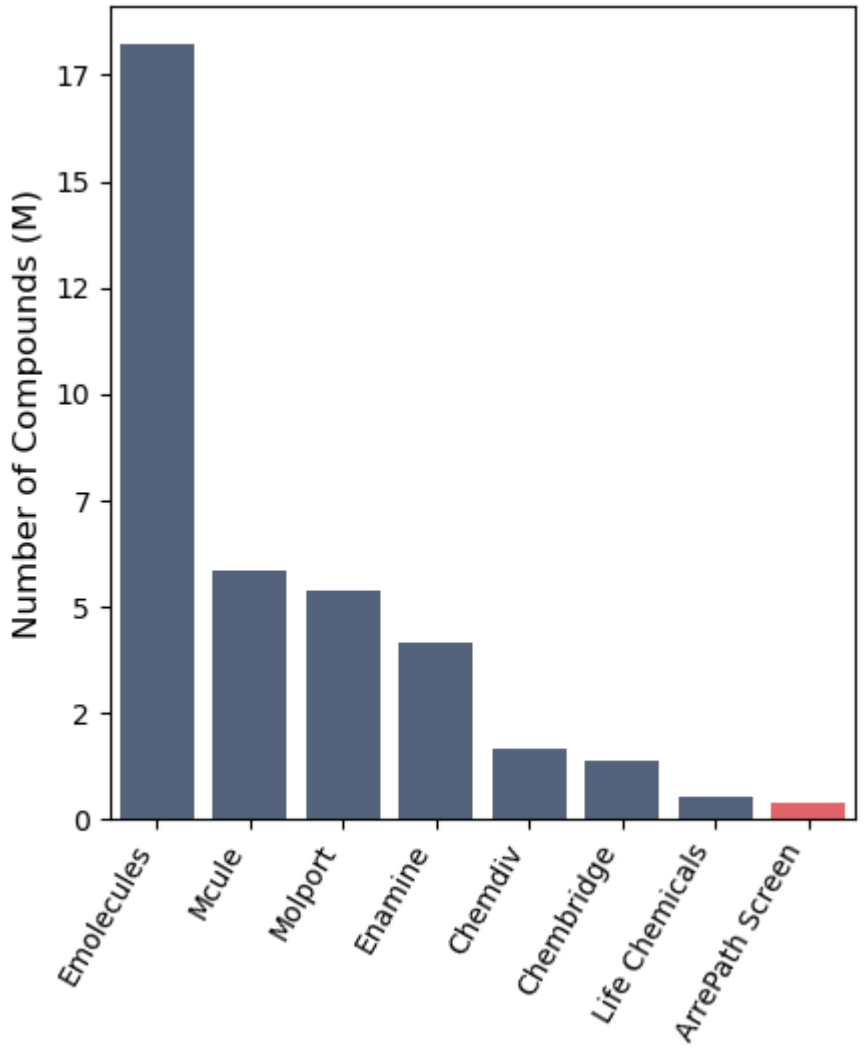
1. Kill bacteria
2. Don't kill mammalian cells
3. Have favorable medicinal chemistry features
4. Are chemically distinct from known antibacterial compounds

Modeling implications:

- Need models for *E. coli* activity and cytotoxicity that extrapolate well to new compounds
- Filter for unfavorable medicinal chemistry and PAINS
- Remove molecules structurally similar to commercial antibiotics

How do we build models with the best generalization?

- There are an estimated 10^{33} drug like molecules that can be made
- Commercial screening collections are ~5M compounds in size
- Goal of model is to identify the most active molecules in these collections – we don't need to predict an MIC, just an ordering.
- We want to maximize the amount and diversity of the compounds we can train our model on

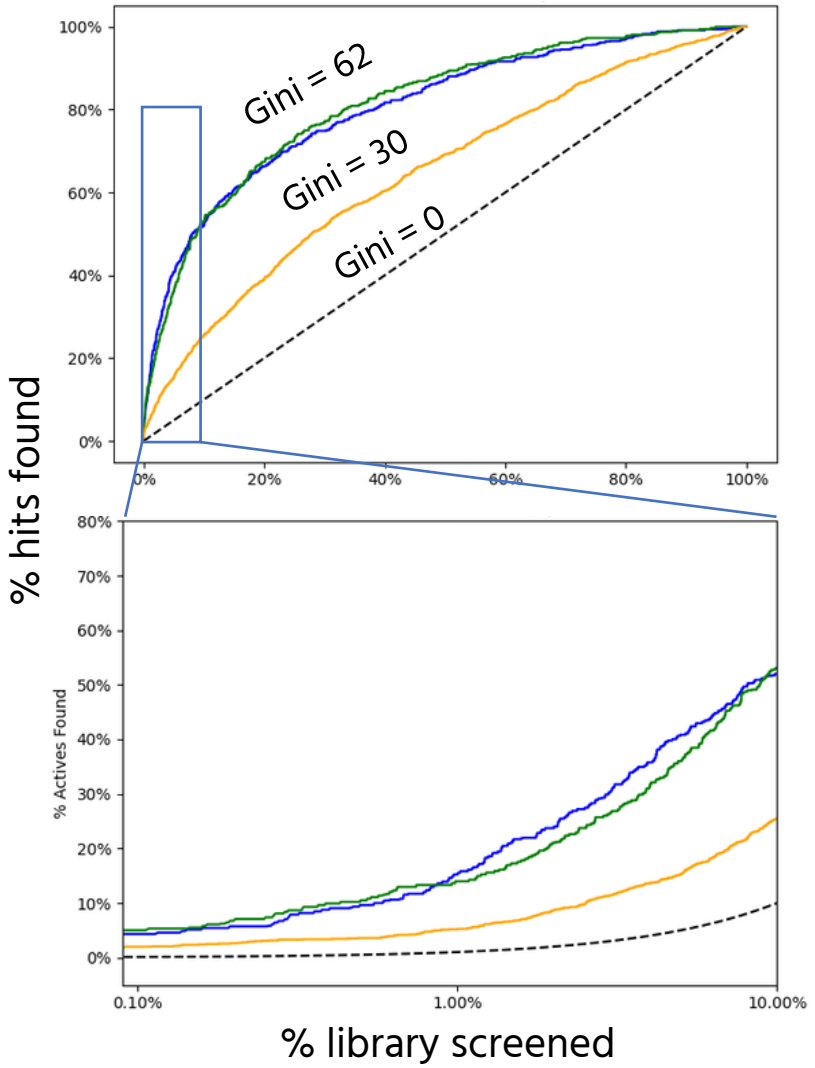


Models are designed to prioritize compounds for screening

- Datasets are scaffold-split – we evaluate model performance on Bemis-Murcko scaffolds different from those we train on
- Models are evaluated by how well they prioritize active molecules
- Virtual screening curves plot hits found as a function of library screened, if screened in the order ranked by the model

Scoring metrics:

- lift@X: 20% hits found at 1% screened compared to 1% from a random model → 20x lift
- Gini coefficient – a function of area under the curve. 0 = random; 100 = perfect



Bemis-Murcko scaffolds: The Properties of Known Drugs. 1. Molecular Frameworks. J. Med. Chem 1996

Antibiotic activity can be predicted across strains

More data from diverse strains outperforms less data from identical strains

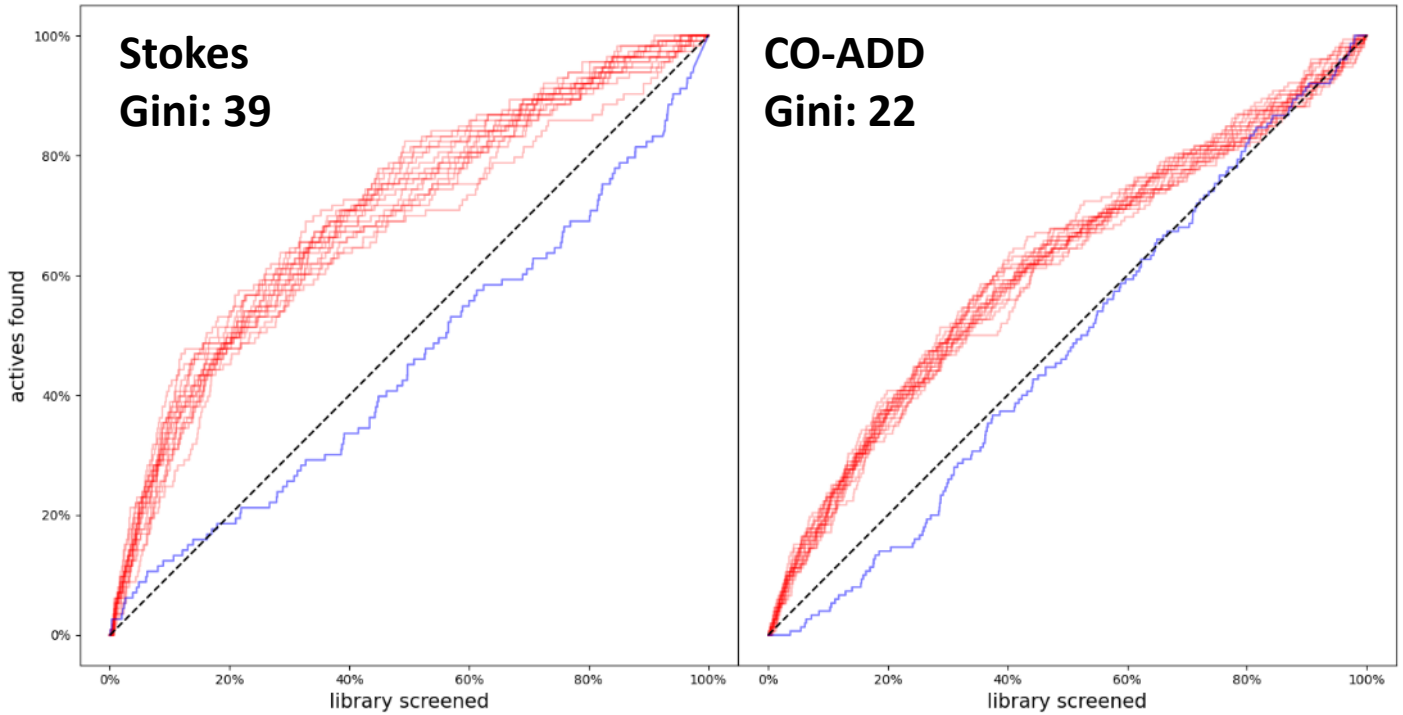
- Nearly all compounds active in WT *E. coli* are also active in permeabilized or efflux deficient strains
- Most (98%) of compounds are inactive in any strain.
- Mutant strain data improves WT activity prediction, even though most mutant-active compounds are inactive against WT.

% of actives also active in

	WT	*tolC*	*lpxC*
WT	100	**84**	**86.2**
tolC	9.3	100	30.6
lpxC	24.3	78.6	100

Data from CO-ADD

Red: WT *E. coli* predicted with 400k cpds measured in *lptD E. coli*
Blue: WT *E. coli predicted* with 500 cpds measured in WT *E. coli*

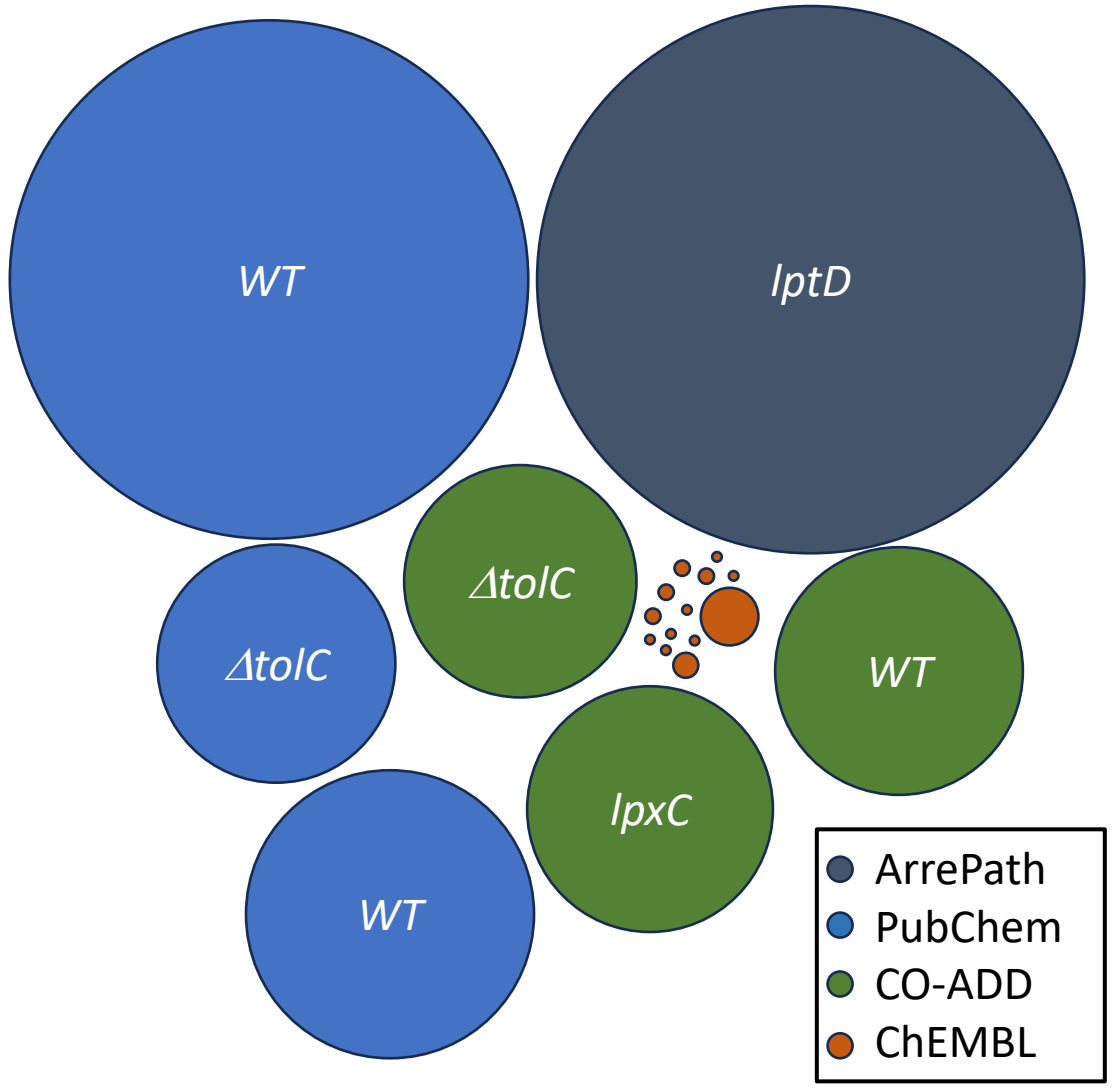


CO-ADD: *Blaskovich MAT, Zuegg J, Elliott AG, Cooper MA. Helping Chemists Discover New Antibiotics. ACS Infect Dis. American Chemical Society; 2015 Jul 10;1(7):285–287. PMID: 27622818*
Stokes: *Stokes JM, et al.. A Deep Learning Approach to Antibiotic Discovery. Cell. 2020 Feb 20;180(4):688-702.e13. PMCID: PMC8349178*

Leveraging lots of data

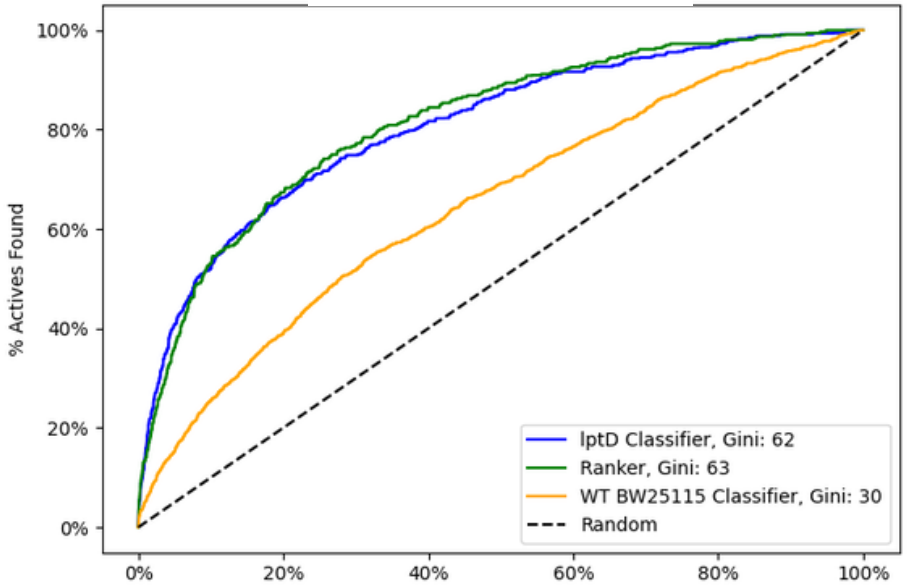
E. coli activity data is available on >1M compounds

- ArrePath has collected a large *E. coli* activity dataset, but there is even more public data
- These are measured in different strains and assay conditions → can't compute an MIC or a percentage inhibition
- Instead, train a model that can consistently rank each assay by activity

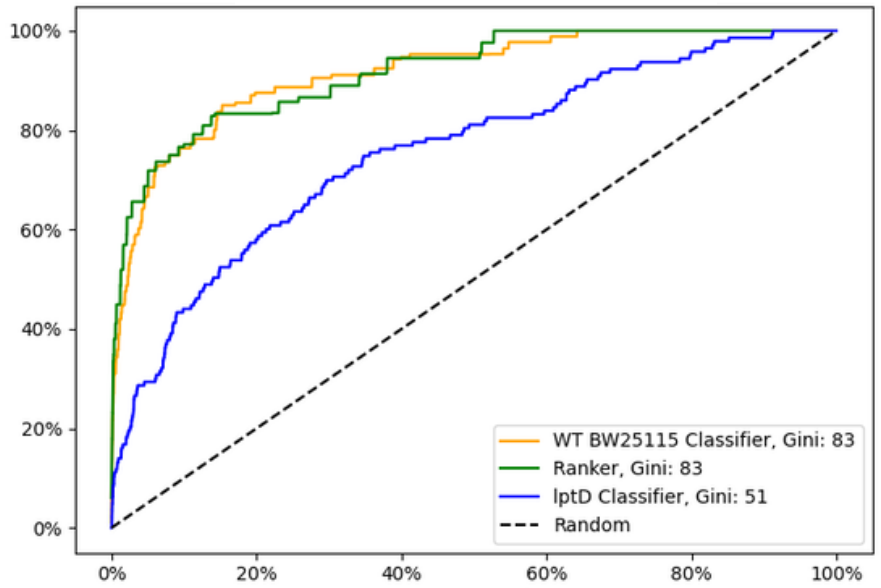


Antibacterial activity ranking models outperform single task classification models

Predicting lptD



Predicting BW25113



Model

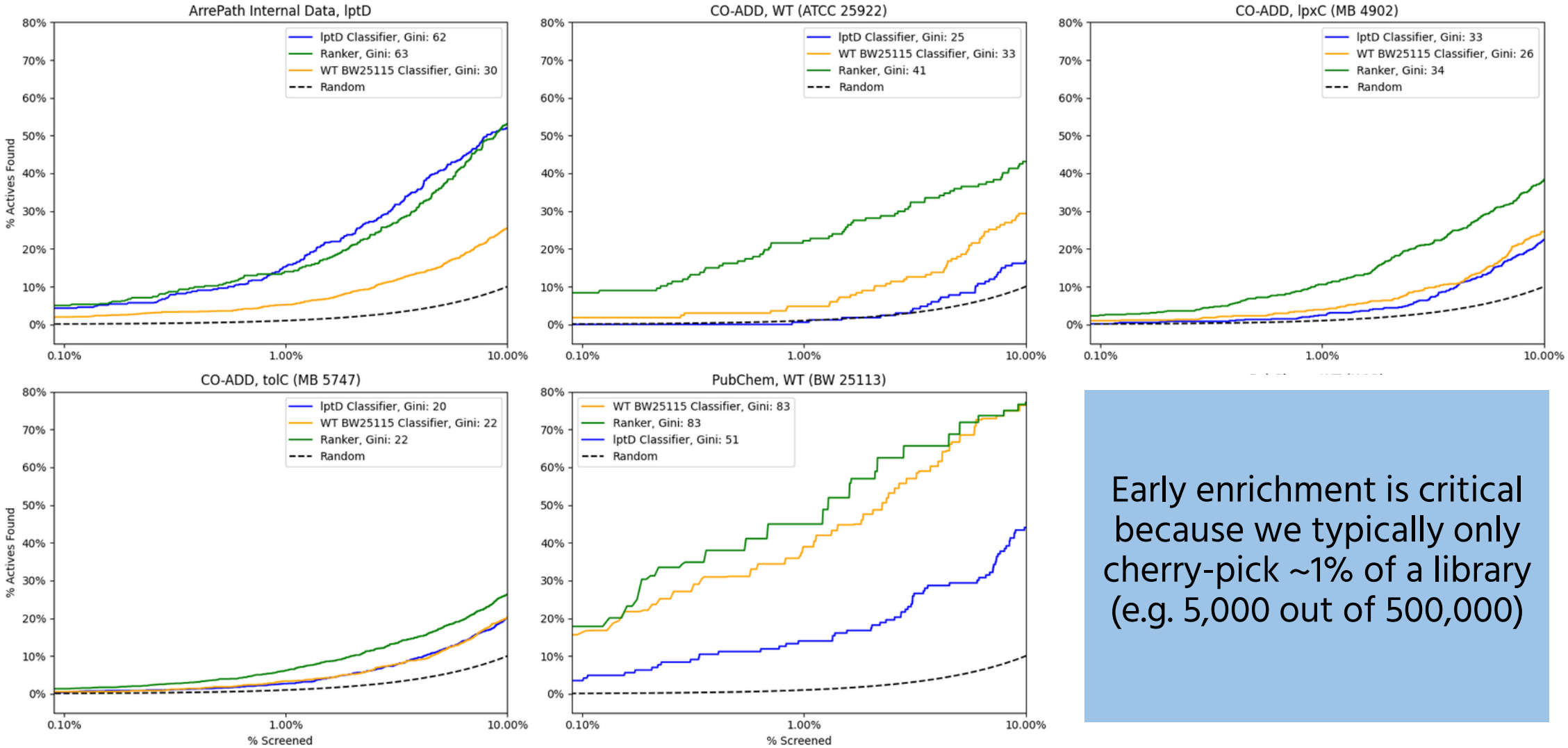
- Multitask ranking model
- Internal lptD classifier
- BW25113 WT classifier

- ArrePath has trained a multitask ranking model that predicts antibacterial activity. It is trained on 1.1M compounds across 9 datasets including WT, efflux deficient, and membrane compromised strains.
- Single task models trained on our internal permeabilized mutant data or a WT Pubchem screen perform well on the datasets they are trained on but generalize poorly to new data.
- The ranking model performs equivalently to the single task models on the data they are trained on and ***better*** on other datasets.

For classification training datasets: hold out 1/5th of data as test set and hold out all scaffolds in test set from training data.
For all other datasets, hold out entire dataset as test set.

Ranking model enriches for hits in early selection

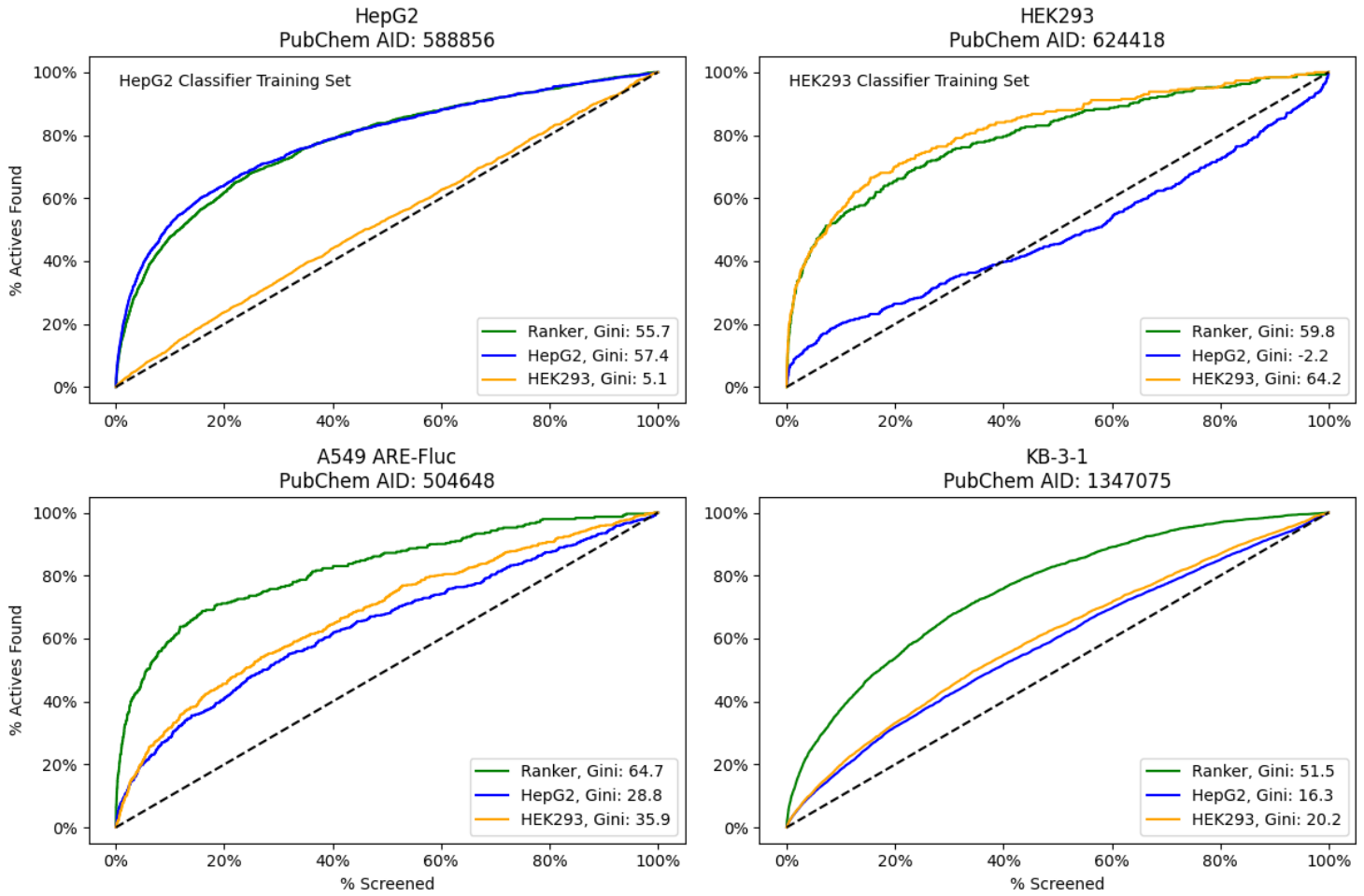
Enrichment factors up to 20x: 20% of actives identified after screening 1% of library



Early enrichment is critical because we typically only cherry-pick ~1% of a library (e.g. 5,000 out of 500,000)

Cytotoxicity ranking models outperform single task classification models

- ArrePath has trained a **multitask ranking model** that predicts cytotoxicity. It is trained on 2M compounds across 16 datasets, including measurements on HepG2, HEK293, and three other cell lines
- As in antibiotic activity, single task models perform well on the datasets they are trained on but generalize poorly to new data.
- The ranking model performs equivalently to the single task models on the data they are trained on and ***better*** on other datasets.

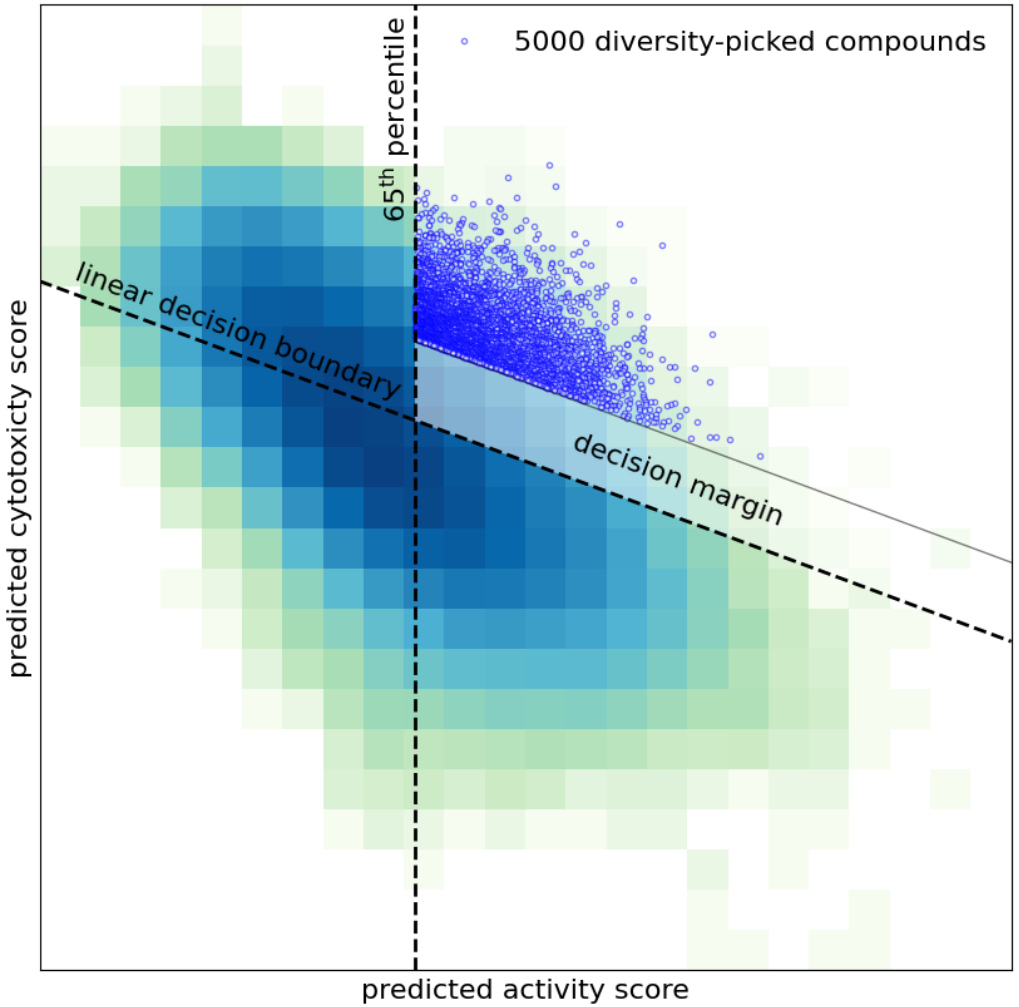


For training datasets: hold out 1/4th of data as test set and hold out all scaffolds in test set from training data.

Compound selection strategy

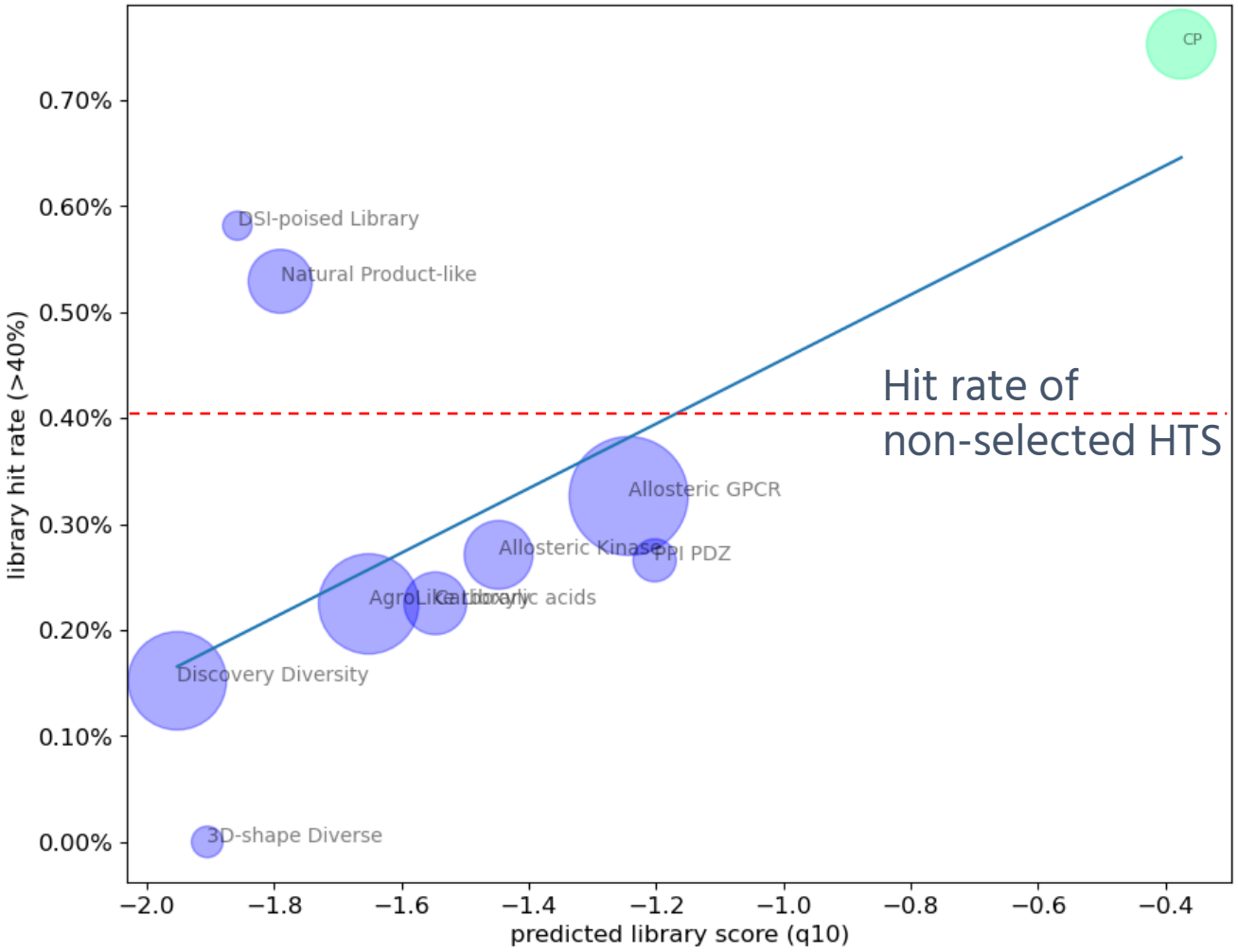
Source compounds from Enamine screening collections

- Excluding compounds structurally similar to any antibiotic approved or tested in the clinic.
- Pick compounds to maximize the distance from the cytotoxicity-antibacterial activity tradeoff boundary.
- We limit how many structurally similar compounds are selected to maximize diversity.
- Compounds are screened in *E. coli lptD* in M9 media in single replicates, followed by triplicate confirmation of hits.
- We also selected several pre-plated sets with high fractions of predicted actives.

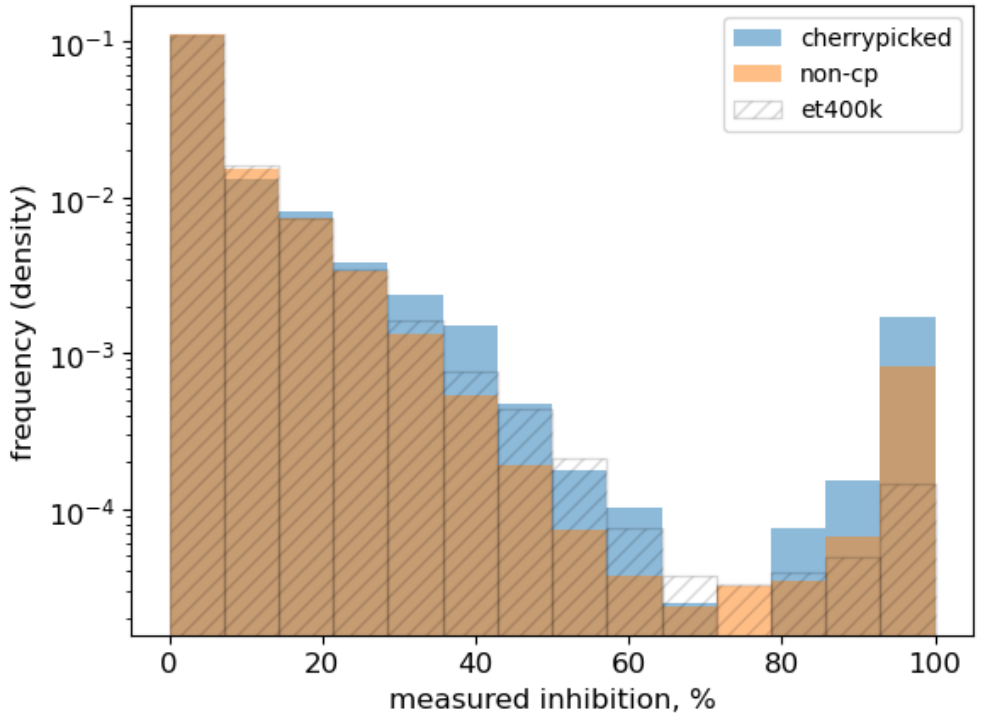


ArrePath's antibiotic activity model enriches for screening hits

2x increase in hit rate over unguided screening plus focus on tractability

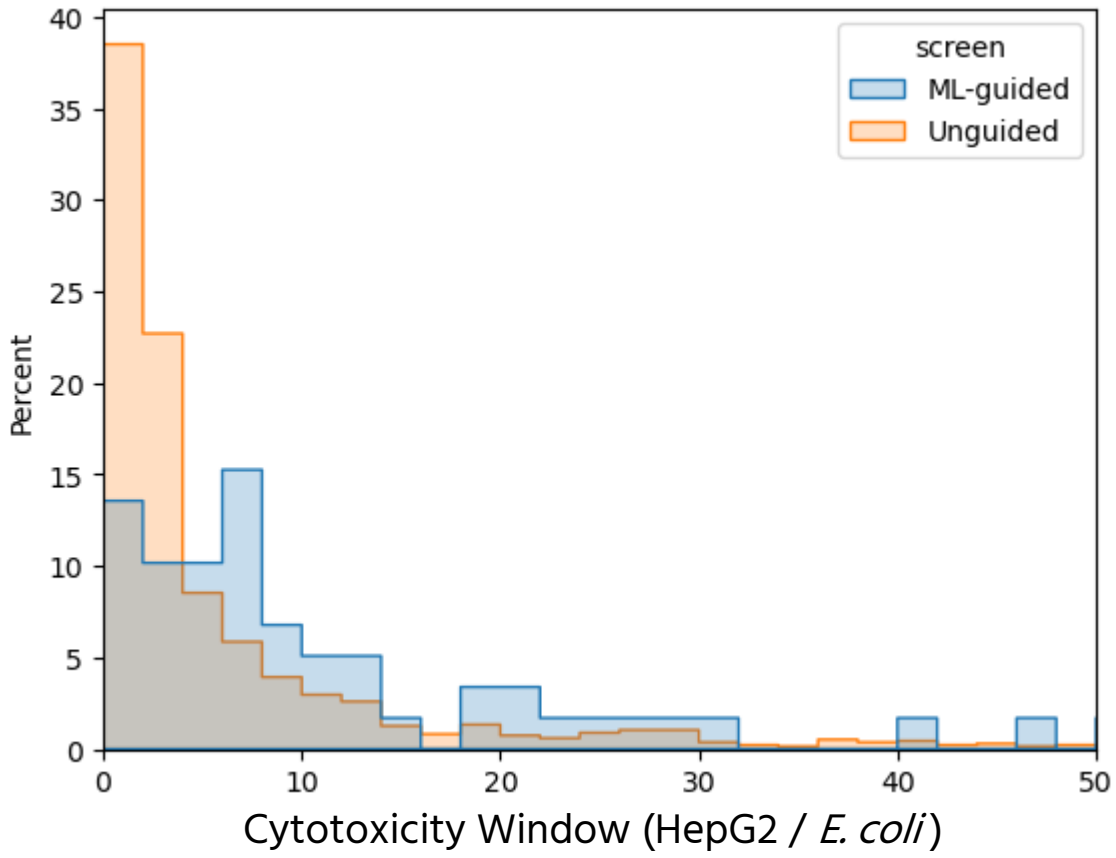


- ML-guided cherry-picking yielded twice as many hits as unguided selection, while selecting for low cytotoxicity and maximizing diversity
- Quality of hits improved significantly – fewer weak inhibitors



ArrePath's ML-guided screening enriches for progressible hits

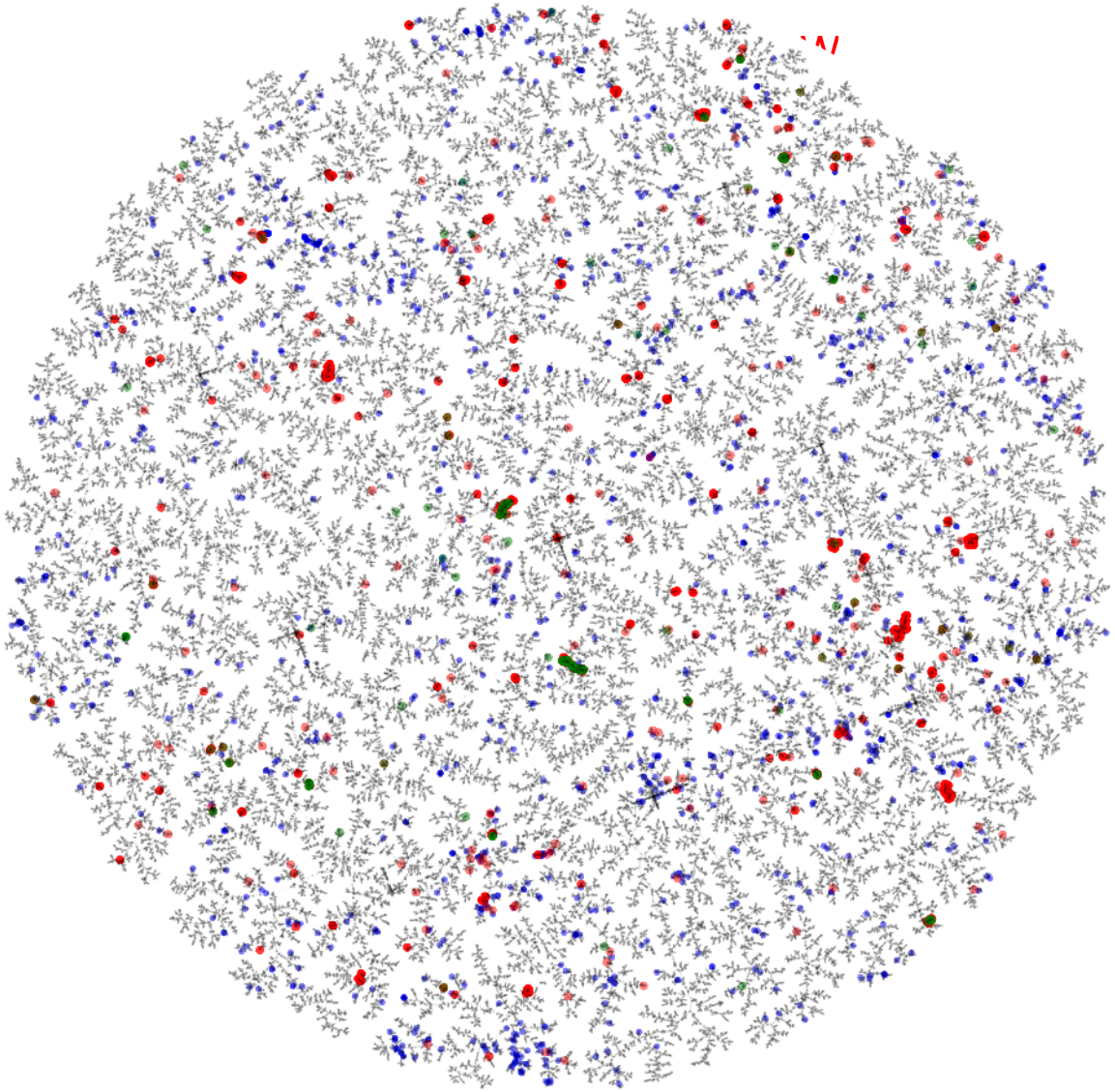
3x reduction in cytotoxic molecules over unguided screening



- **13% of ML-selected hits** "fail" due to cytotoxicity (cytotoxicity window <2)
- **37% of non-selected hits** "fail" due to cytotoxicity
- This represents an additional enrichment in progressible hits. Combined with the ~2x enrichment in hits, **this yields a ~3x efficiency gain in progressible hits from ML-guided screening**
- Compounds which do pass are on average less toxic (median cytotoxicity window 10 vs 6)
- We identify a larger fraction of WT-active compounds in ML-guided screening (67% vs 45%)

ArrePath's ML-guided screening identifies chemically-diverse hits

- The chemical space covered by our hits is distinct from both clinically used antibiotics and published non-clinical antibacterials
- The plot on the right shows a tmap projection of our screened chemical space along with our hits, clinical antibiotics, and antibacterials



- ArrePath ML-guided hits
- All screened compounds
- Approved antibiotics
- Non-clinical antibacterials of known mechanism

Selection for antibacterial activity alone yields a high hit rate

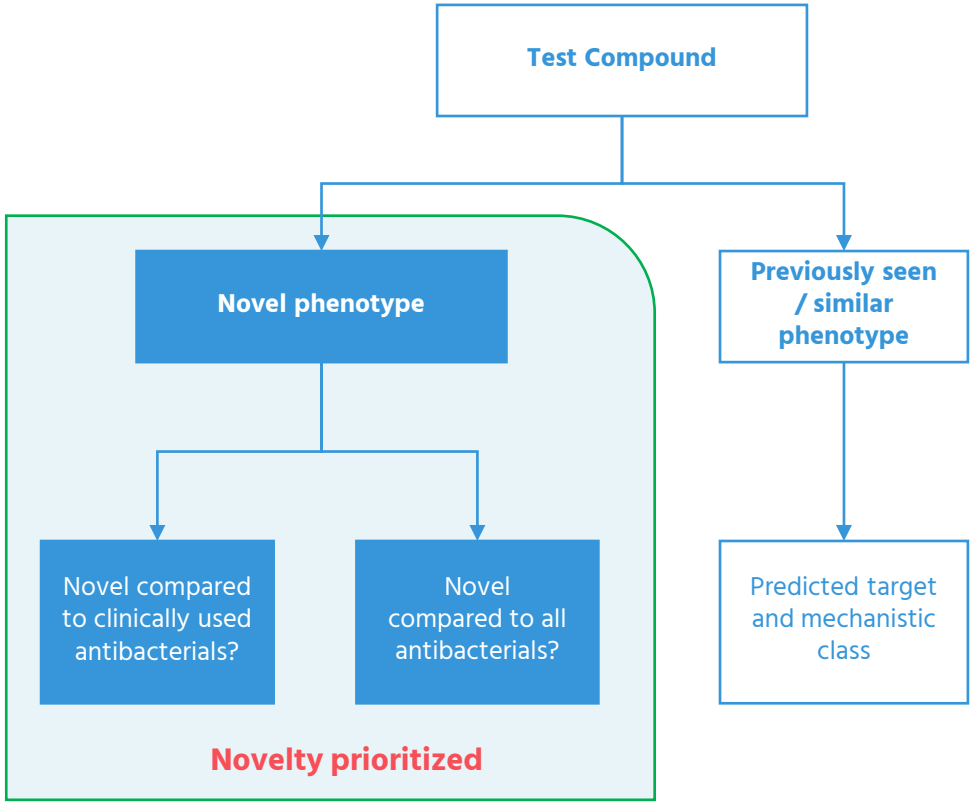
- Multiple ML models have been shown to yield high hit rates for predicting antibacterial activity
- These are reported on testing a small number of top-ranked compounds which increases hit rate
- Additionally, they frequently include compounds highly similar to known antibiotics (e.g. 47/51 for Stokes)
- **We estimate a hit rate in a similar range if we tested only the top ranked hits in our screen**
- Instead, we aim to prioritize selection of compounds with novel MoAs that can progress to hit-to-lead efforts

Source	Model	Compounds Tested	Hit Rate
Roche	GNEprop	345	24%
Stokes	Chemprop	99	51%
ArrePath*	Ranking	345	32%

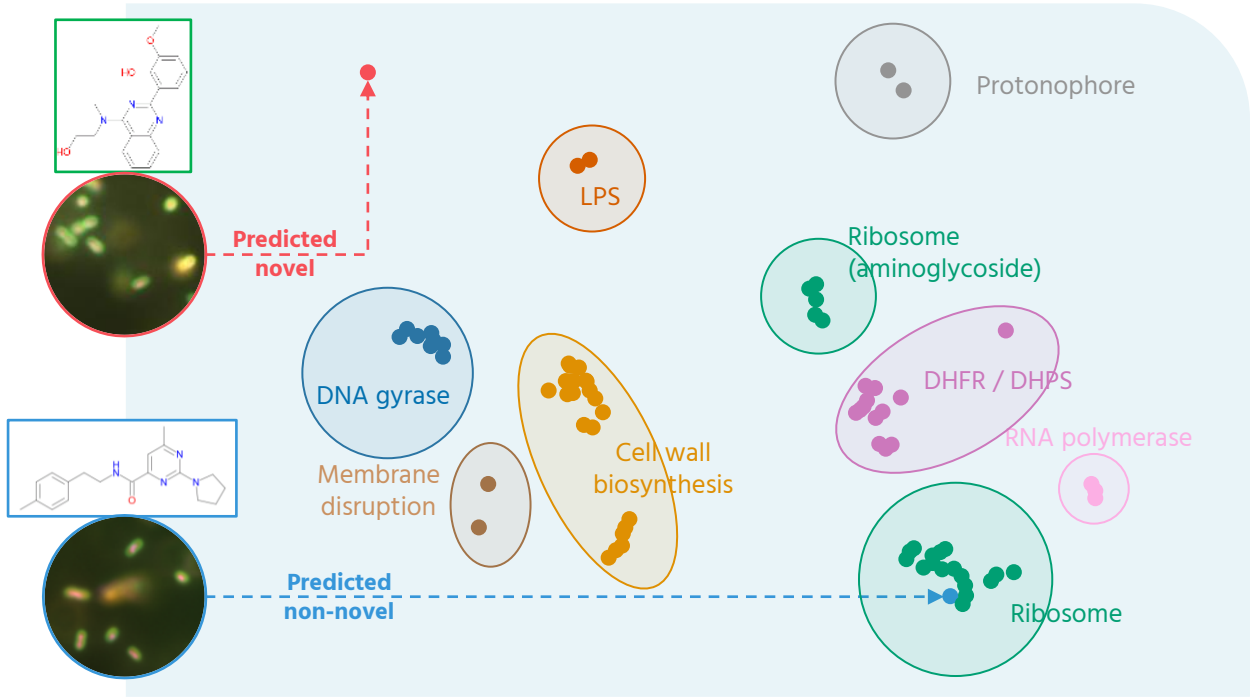
**Estimated*
Sources: PowerPoint Presentation (roche.com);
Stokes JM, et al. A Deep Learning Approach to Antibiotic Discovery. Cell. 2020 Feb 20;180(4):688-702.e13.

Mechanistically novel hits are identified with ArrePath's imaging platform

All screening hits are profiled on our imaging platform to predict mechanistic novelty and potential targets



Mapped antibacterial phenotypic space



Results of screen to date

Example hits shown below

- 104,779 compounds screened: ~30,000 cherry picked; ~70,000 in pre-plated libraries
- 3,376 primary hits; 1001 selected for follow up
- **287 active in in-house dose-response**
- **Example compounds:**

AP-00471778
MIC: 4 µg/mL
(*E. coli* lptD, M9 media)

Putative target:
Val/Ile Biosynthesis

AP-00469177
MIC: 1 µg/mL
(*E. coli* ATCC25922, M9 media)

Putative target:
Uracil Biosynthesis

AP-00477607
MIC: 8 µg/mL
(*E. coli* ATCC25922, M9 media)

Putative target:
Cysteine Biosynthesis

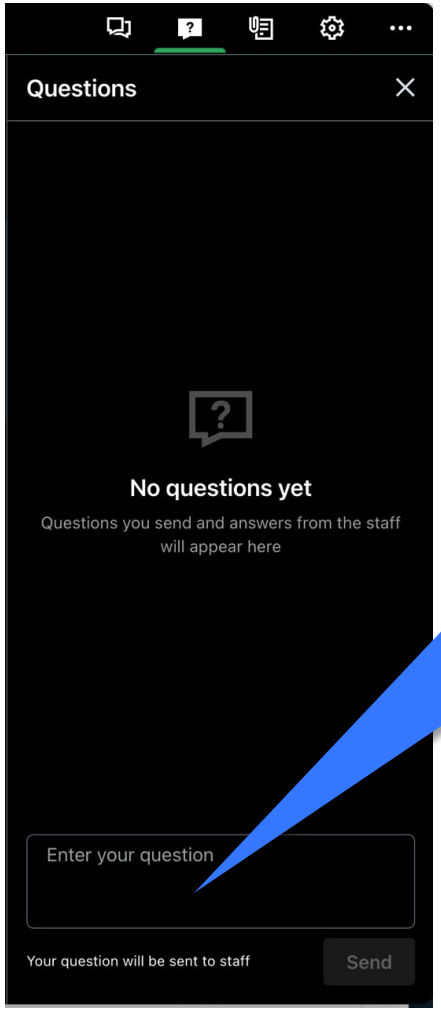
All three compounds show no HepG2 cytotoxicity at 100 µM

Conclusions

- Predictive antibacterial models can be trained from data across multiple bacterial assays, including across strains and media
- Screening compounds selected by antibacterial and cytotoxicity models increases antibacterial hit rate by 2x, and reduces cytotoxicity failures by 3x
- Screening compounds selected by antibacterial models increases wild type hit rate by 1.5x, indicating the models learn features associated with permeability and efflux avoidance
- Screening hits are prioritized by mechanistic novelty using imaging
- **ArrePath has used these approaches to identify a chemically diverse set of antibacterials active against multiple novel targets, resulting in our current LO project and another HTL project.**

How to submit your questions

If your question is addressed to a specific speaker, please include their name when submitting the question.



Please submit your questions through the box provided after clicking the ‘questions’ button. We will review all questions and respond to as many as possible after the presentation.

Today’s speakers

REVIVE by GARDP

Charting new frontiers in artificial intelligence for antibiotic design

Moderator:
Akhila Kosaraju
Phare Bio

Jonathan Stokes
McMaster University

Kurt Thorn
Arrepath, Inc.

The next REVIVE webinar will be announced soon!

Be the first to hear the latest REVIVE updates

- On the REVIVE website (revive.gardp.org/webinars)
- Subscribe to our newsletter
- On X (@gardp_amr) and LinkedIn

Thank you for joining us