



J.P. Morgan Healthcare Conference

Daphne Koller, CEO and Founder
January 2024

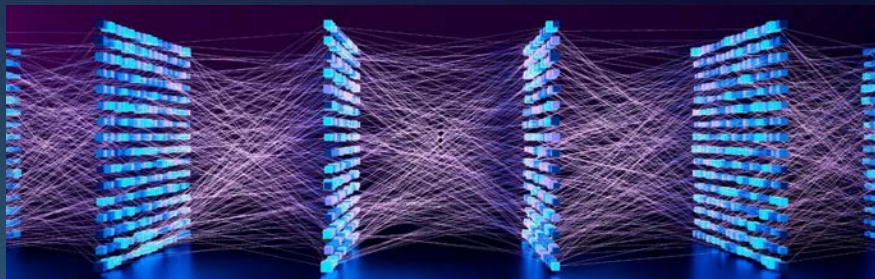


DISCLAIMER

The contents of this document and statements made during this presentation are not intended to market or promote the business of insitro or any particular products, services or capabilities. Representations concerning insitro's business or capabilities may be forward-looking, and, in the context of this presentation and discussion, are not meant as guidance concerning future business, undertakings, offerings or capabilities of insitro. Rather, the perspectives presented include speculation concerning current and future economic, scientific and business trends which are subject to change, and, in any event, may have little bearing on the future course of insitro's business.



The Convergence of Four Forces



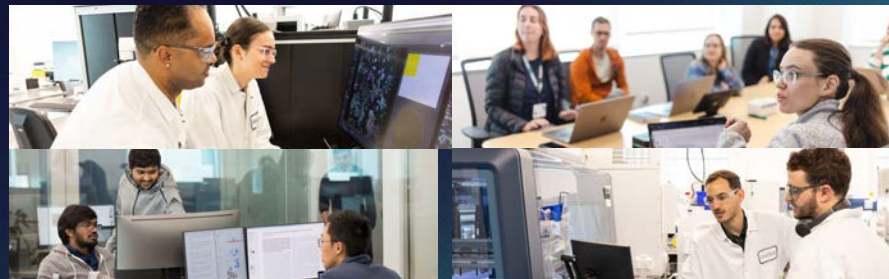
Advancements in ML algorithms



Unbelievably vast amounts of data



Growth in computational power



Growing base of ML talent

Can AI help us **make better medicines?**

We need the **right data**
in the hands of the **right people**
to build the **right ML models**

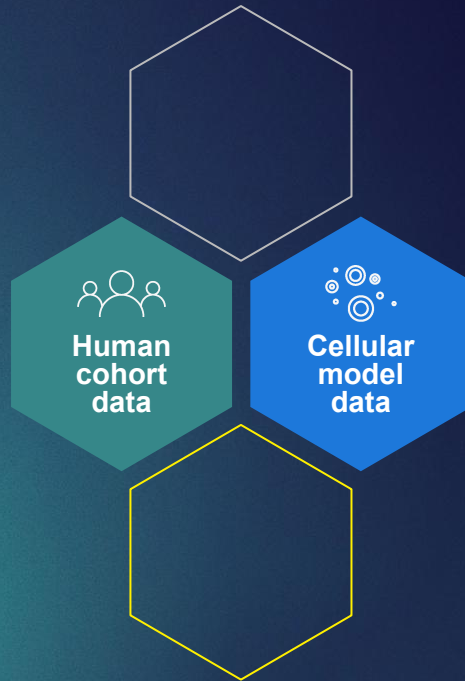
Data is Key



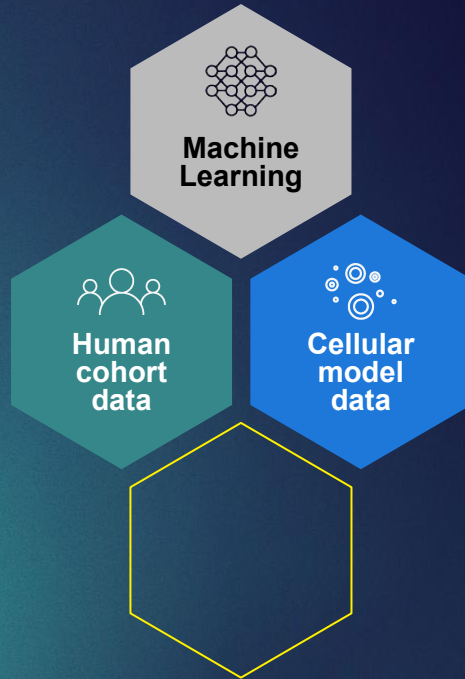
We have built a large and growing corpus of **human clinical data**...



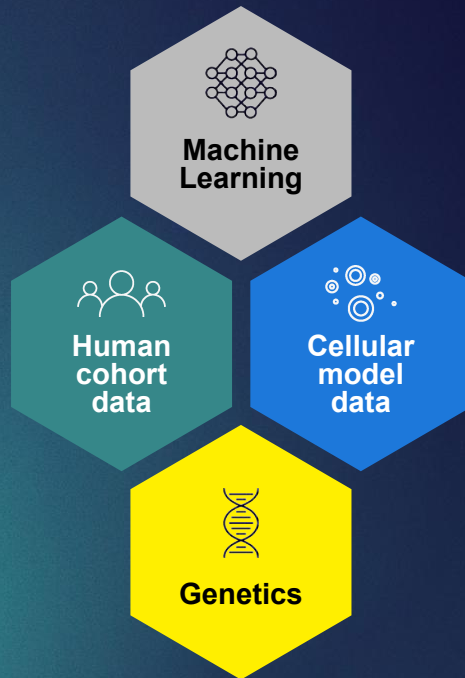
... And robust infrastructure
to print **cellular data at scale**



Machine learning enhances and interprets these data



And **genetics** provides a
ground truth for causality



We Collect Data in Humans and in Cells

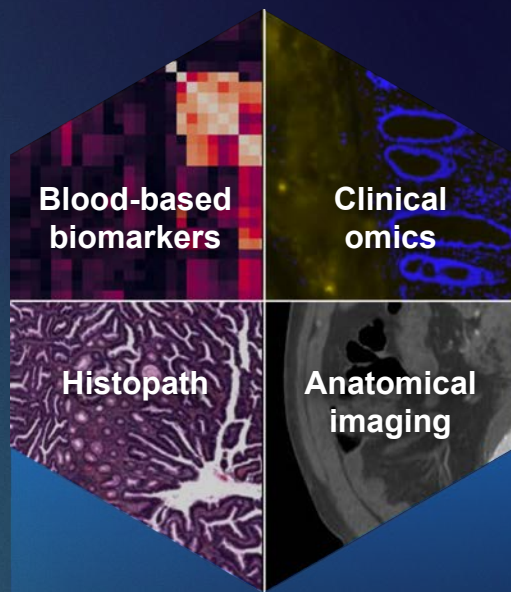


**Human
cohort
data**

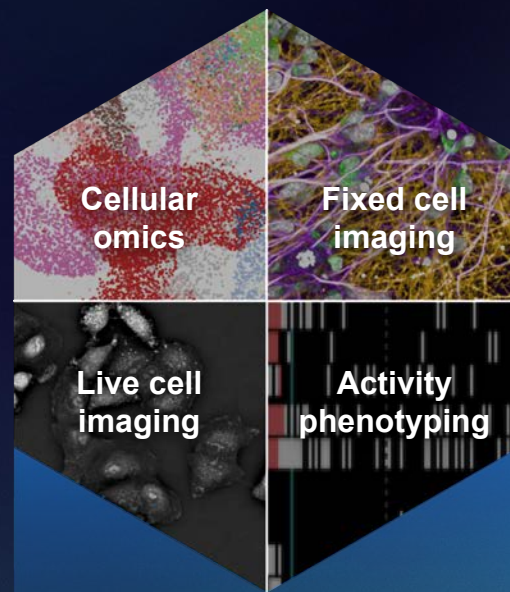


**Cellular
model
data**

... Across Multiple Phenotypic Modalities

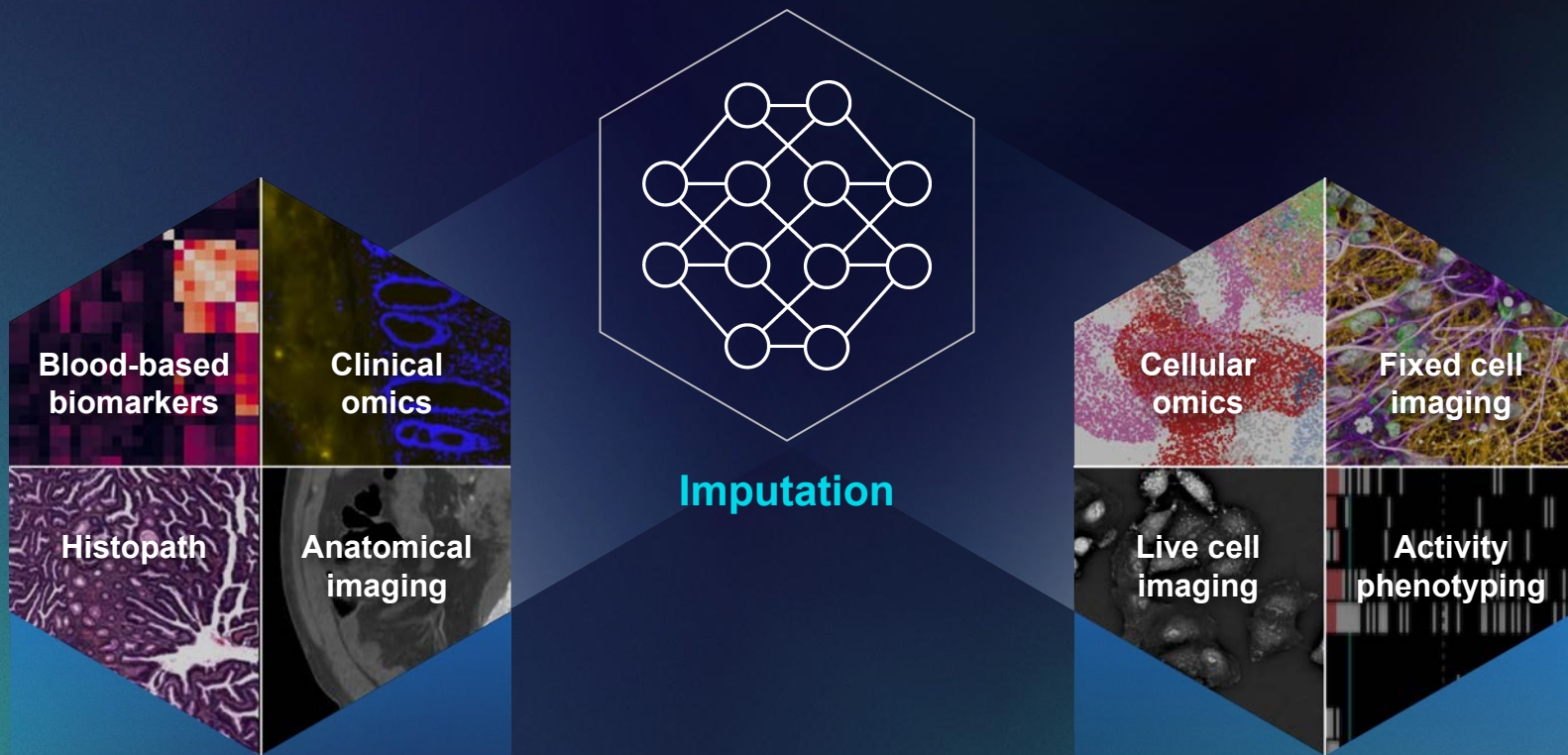


Human Cohort Data

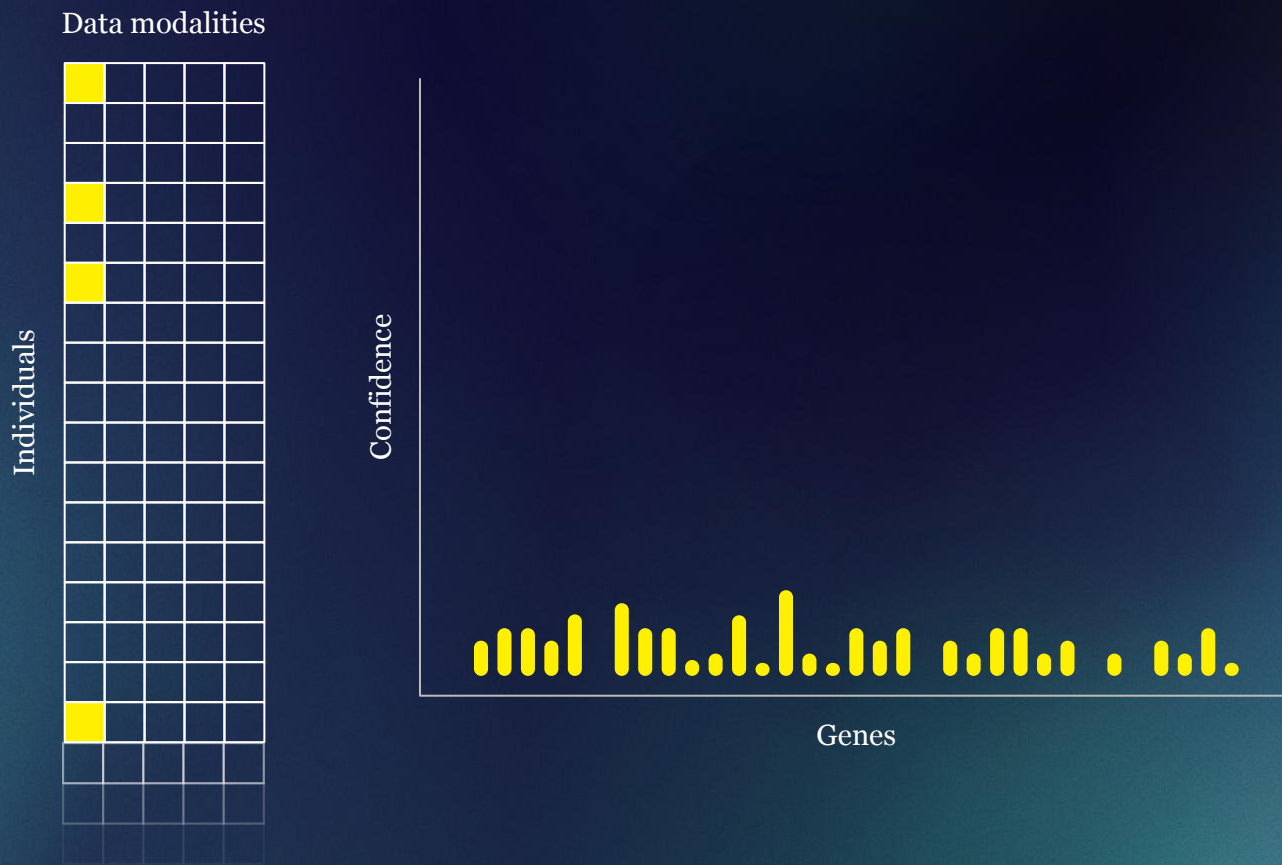


Cellular Model Data

And Hugely Expand Our Data via ML-enabled Imputation

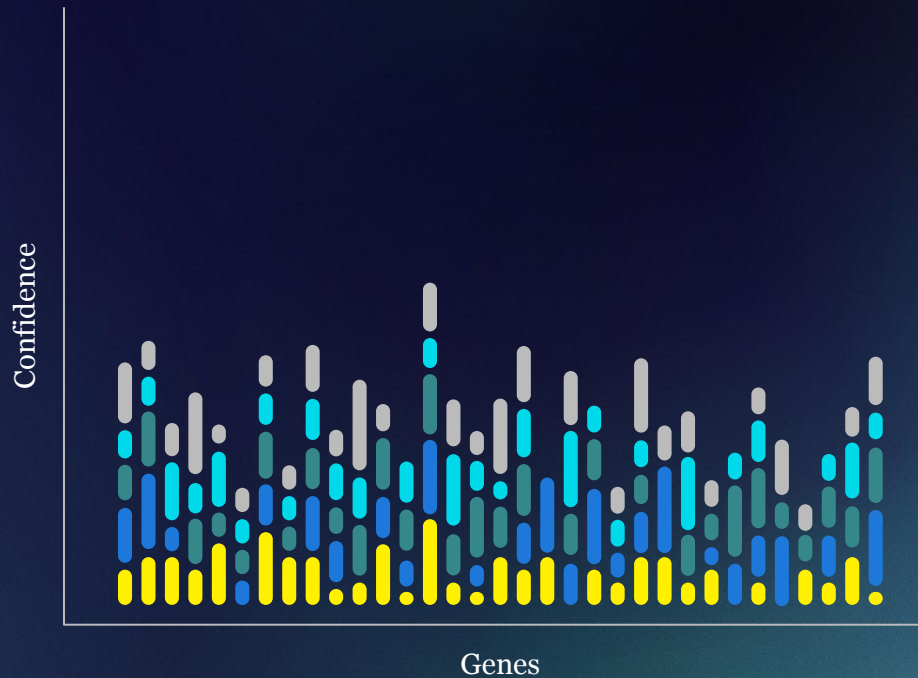
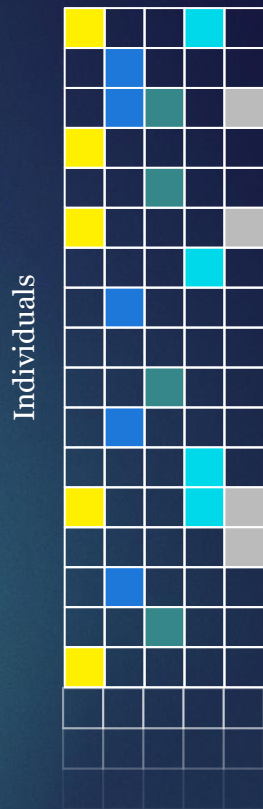


Many Genetic Analyses Are Underpowered for Discovery

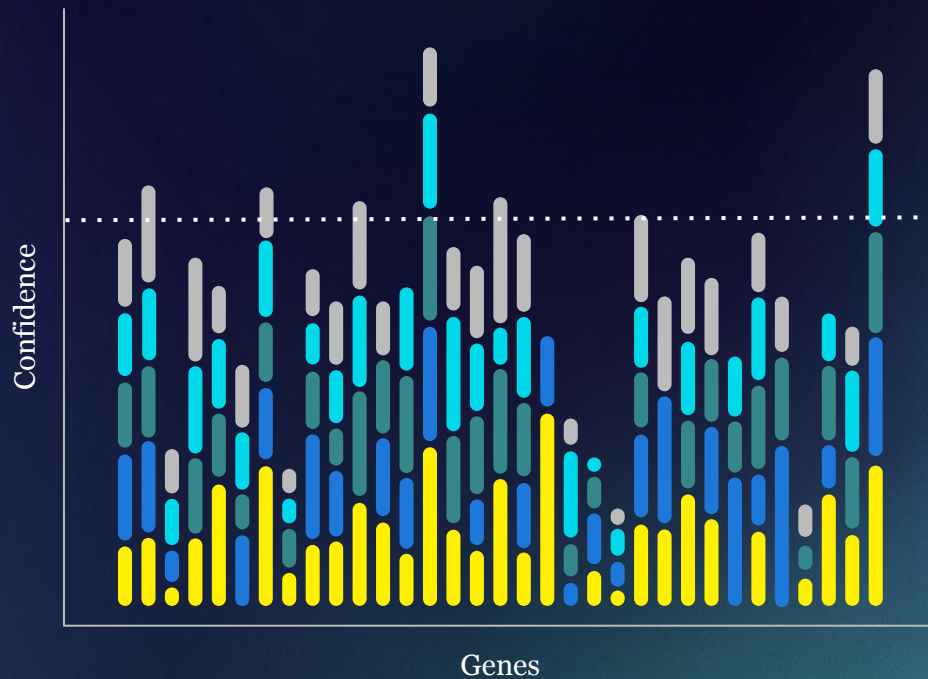
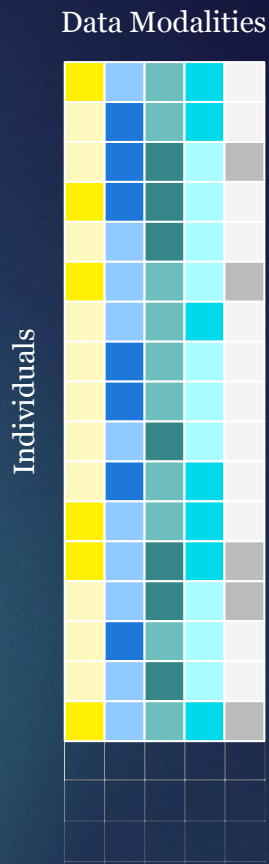


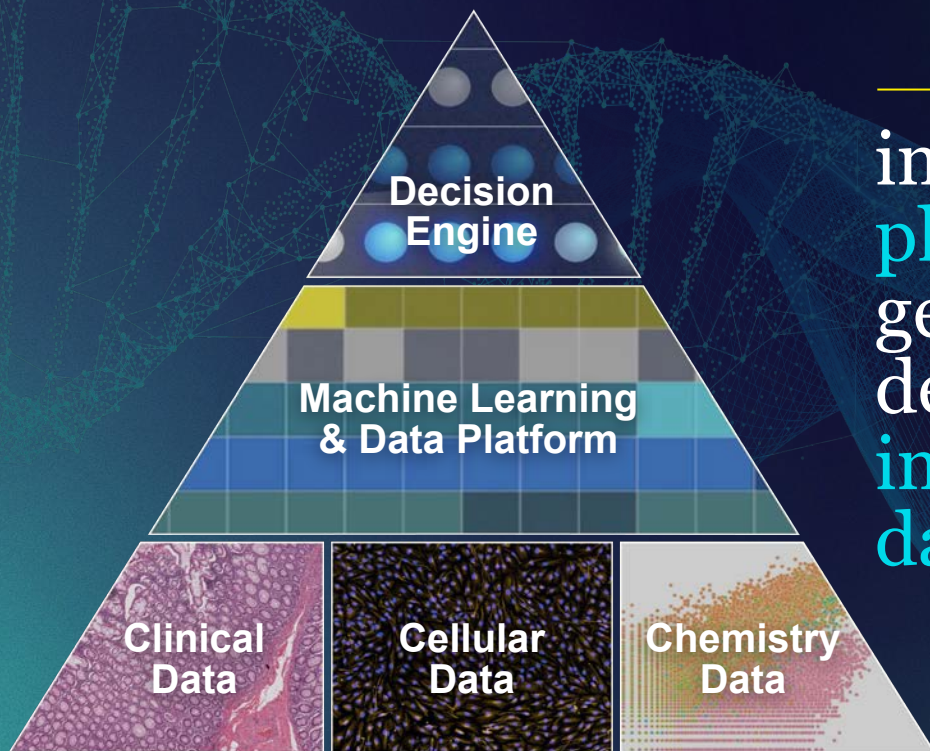
Multiple Modalities Can Increase Confidence

Data Modalities



ML-enabled Imputation Provides Huge Increase in Power





insitro has built a **unique platform** to robustly generate data at scale, derive **ML-enabled insights**, and make **data-driven decisions**

Our Platforms Improve Every Step along the Value Chain

Biology Discovery

The right targets in
the right patients

High confidence therapeutic
insights identified via our
Cell-ML and Clin-ML platforms

Drug Discovery

Better drugs, faster –
across modalities

Rapidly generating and
assessing molecules via our
Chem-ML and Cell-ML platforms

Clinical Strategy

Effective and efficient
clinical trials

Clin-ML platform generates
biomarkers for patient
selection, efficacy, and toxicity

So as to Deliver on Our Mission

PIPELINE

PLATFORM

To bring **better drugs faster** to the
patients who can benefit most
_____ through _____
machine learning and **data at scale**

We Leverage Our Platforms across Our Core TAs

Oncology

Neuroscience

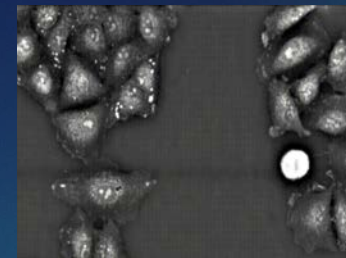
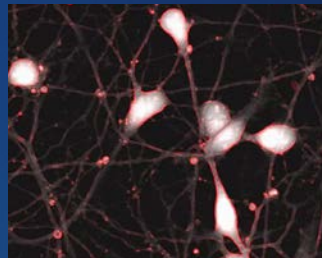
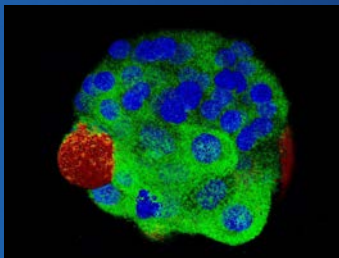
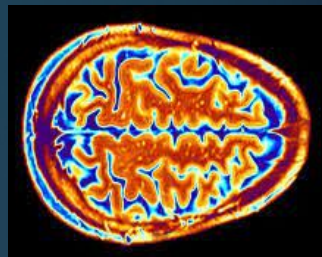
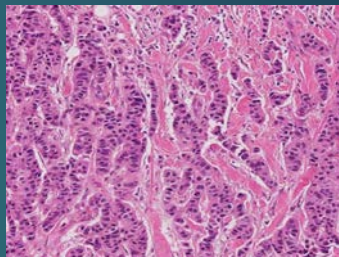
Metabolism



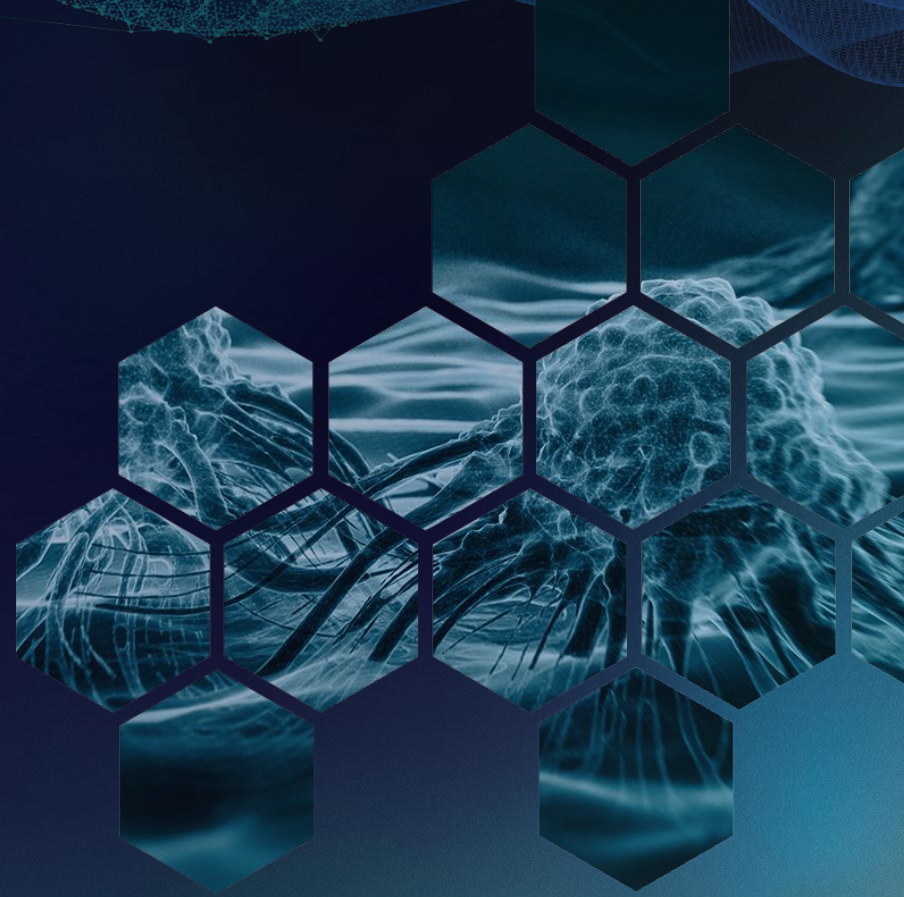
Human
cohort
data



Cellular
model
data

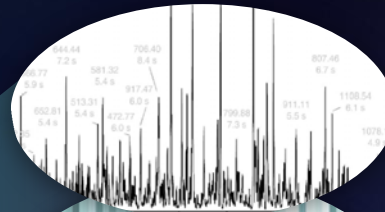
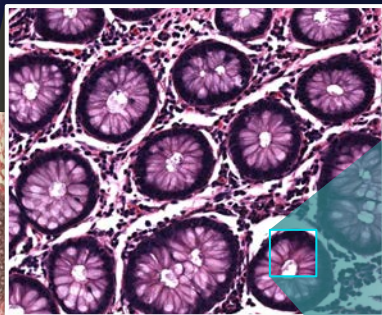
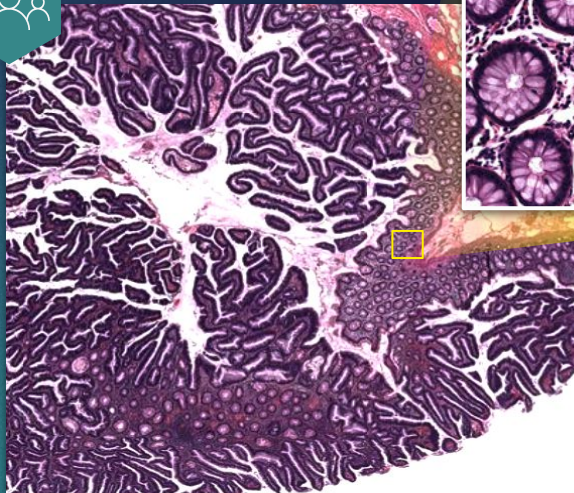


Oncology Discovery Using Clinical ML

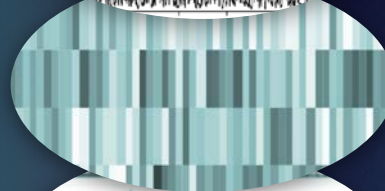


Histology as a Foundational Modality for Clinical-ML

- Routinely collected, increasingly digital
- High-content, multi-scale
- Beyond human interpretability



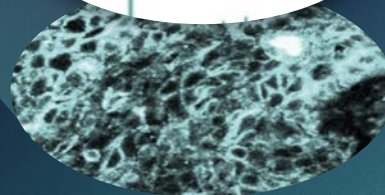
Proteomics,
Metabolomics,
Methylation



RNAseq



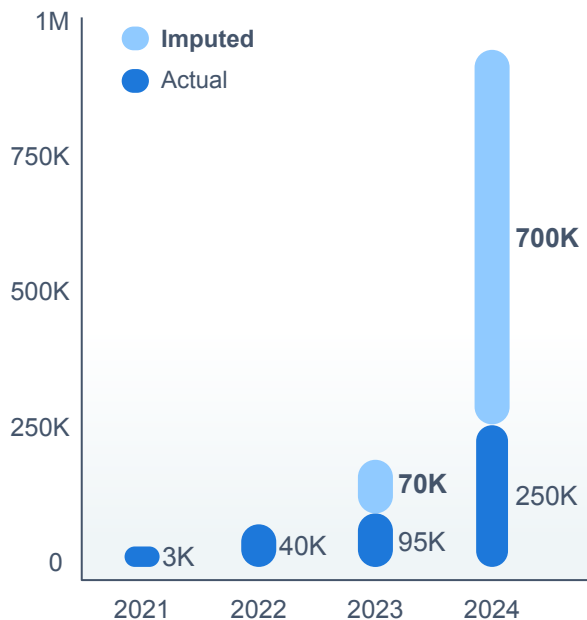
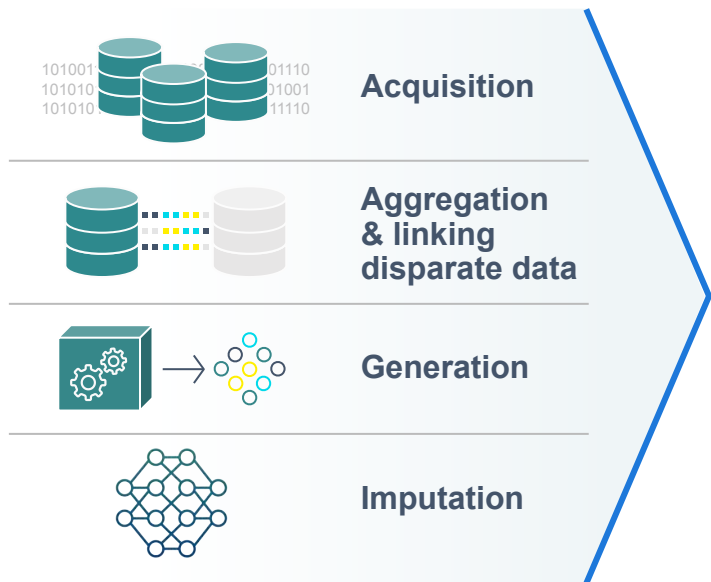
ATAC



Spatial
proteomics

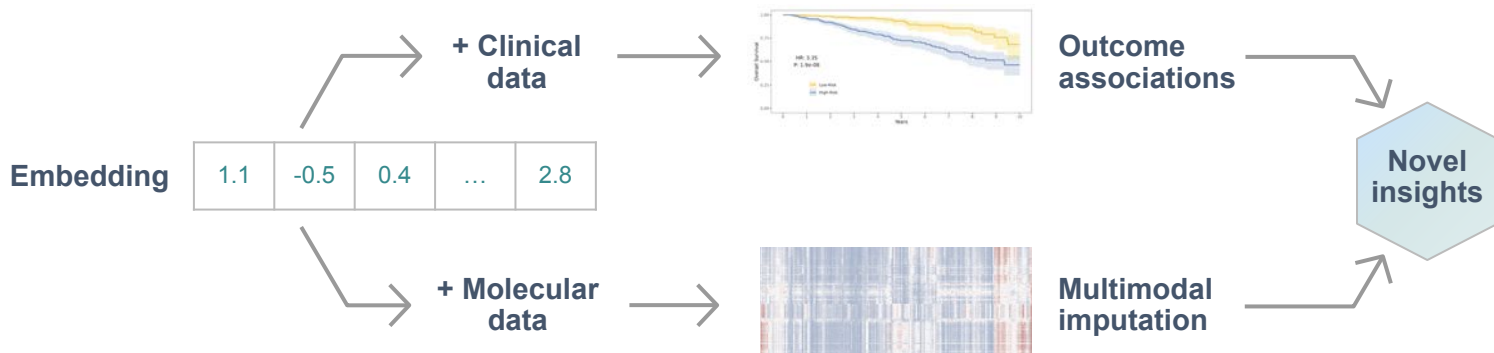
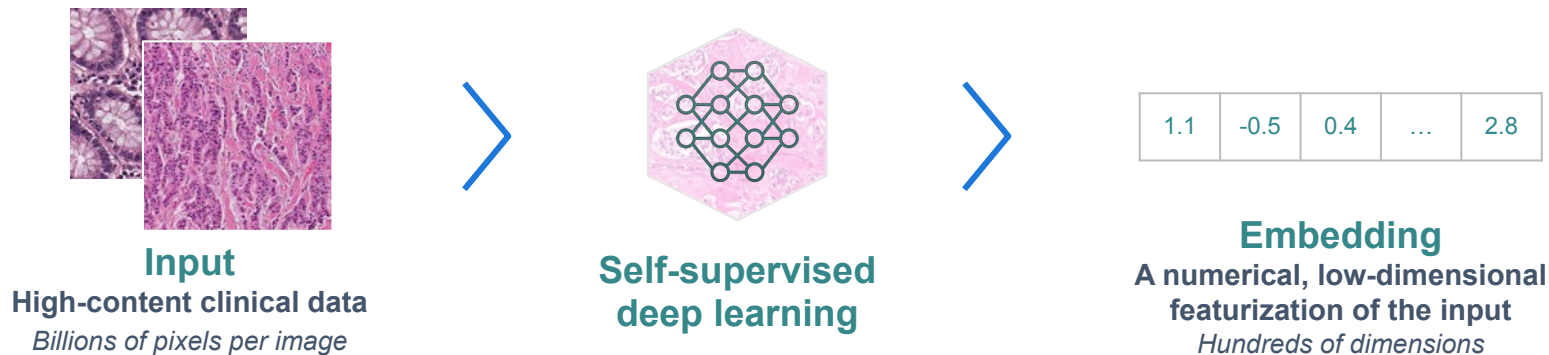
Multi-pronged Approach to Scale High-content Clinical Data Drives Platform Flywheel

We are building a rich data corpus spanning multiple modalities & disease contexts



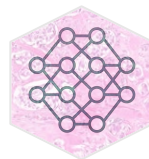
- ~100,000 curated cases with extensive multimodal data today
- Scaling to ~1M multimodal cases with imputation
- More data drives better and broader imputation
- Platform flywheel to massively increase prediction quality *without* massive costs

Self-Supervised Learning for High-Content Clinical Data

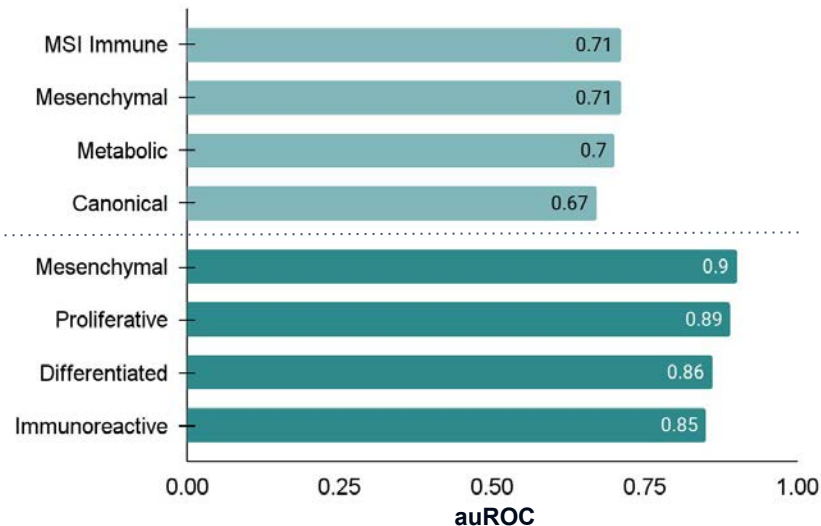
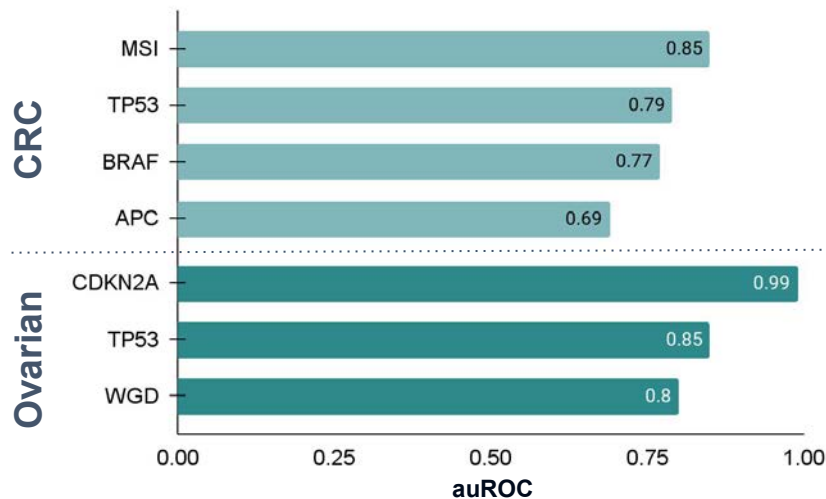


Accurate Predictions of Clinical Biomarkers from Histology

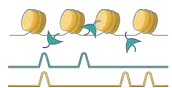
Predicting DNA alterations



Predicting RNA-defined subtypes

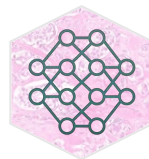


Predictions of Molecular Measurements Enable Broad Exploration of Therapeutic Insights



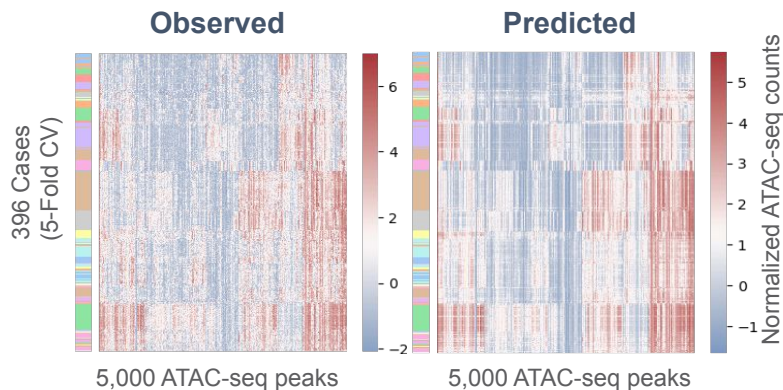
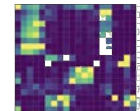
Imputing ATAC-seq

Model trained on **396 bulk samples**
spanning 23 tumor types

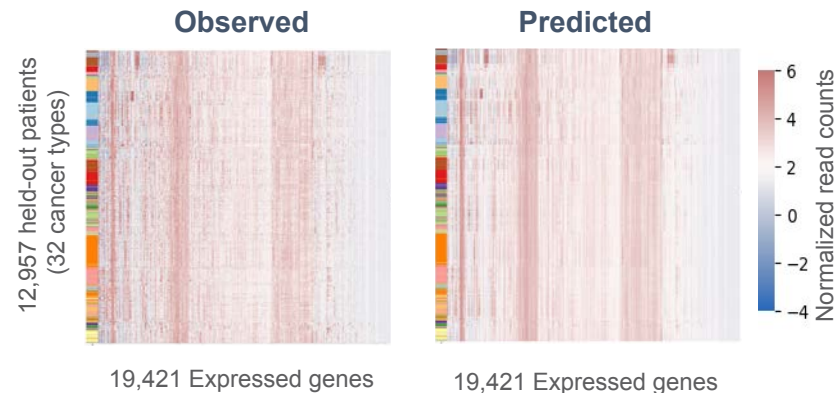


Imputing Gene Expression

Model trained on **~13K bulk samples**
spanning 32 cancer types



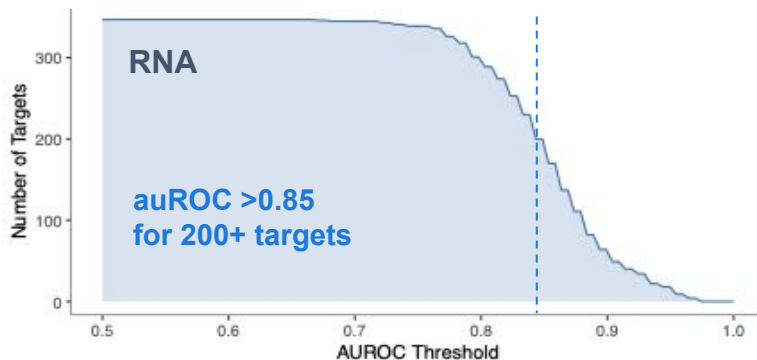
$r^2 = 0.61$ $p < 1e-30$



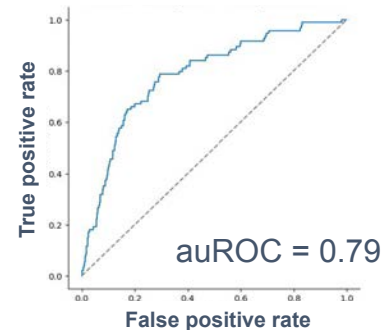
$r^2 = 0.64$ $p < 1e-30$

Therapeutic Insights from H&E Images Alone

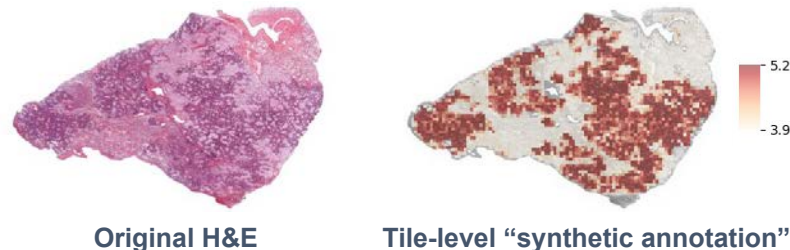
Prediction of gene overexpression
across multiple solid tumors for
targets of ~350 antibody-based drugs



Predicting individual
MET expression for
lung cancer patients



Predicting spatially resolved
MET expression



Impacting Oncology Drug Development

Target-derived biomarkers

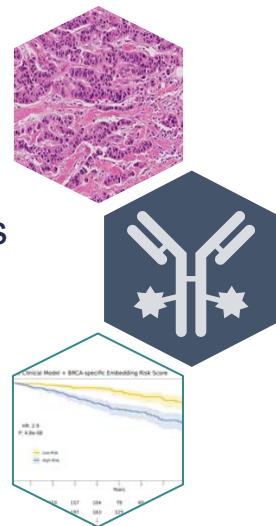
Improving outcomes with hyper-stratification of potential responders

Biomarker-inspired biologics design

Informed design of precisely targeted therapeutics

Target discovery & credentialing

Expanding precision treatments with novel target insights

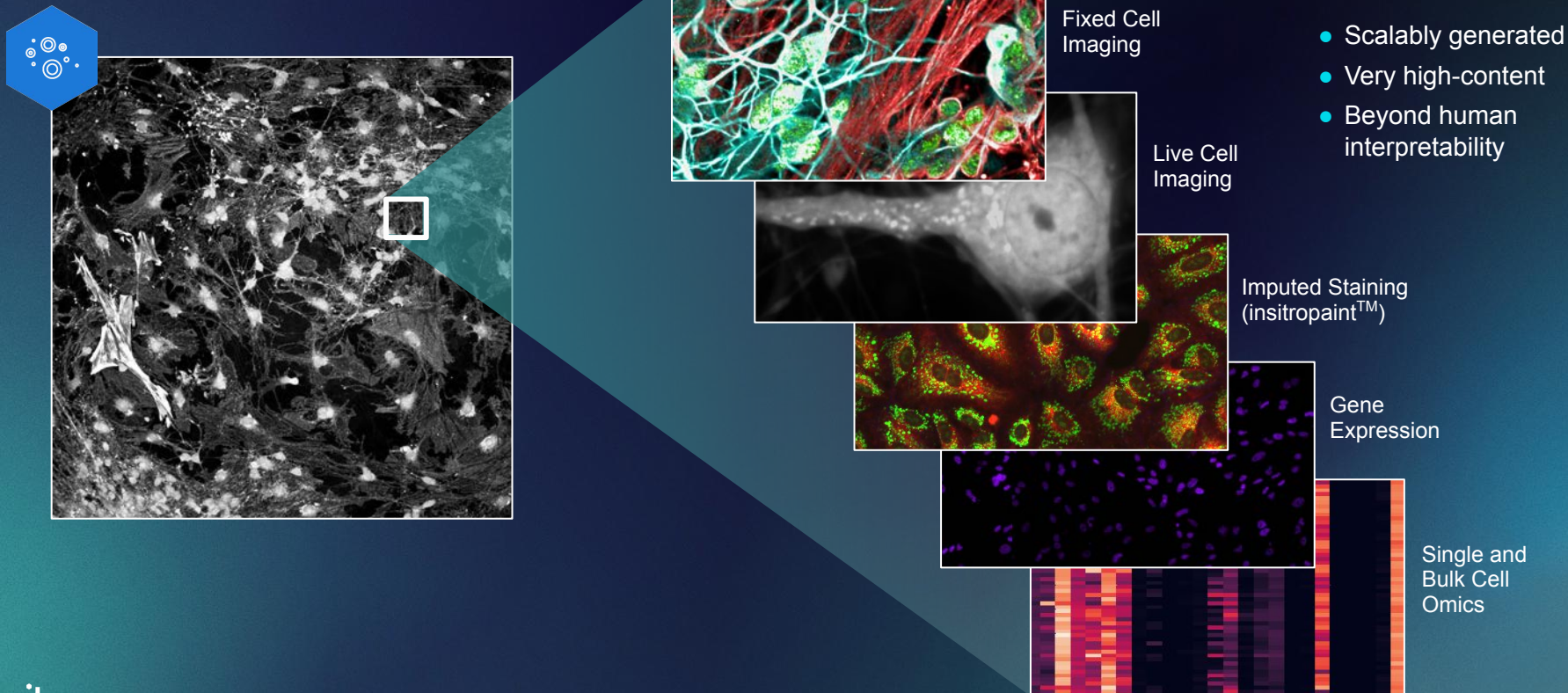


We are leveraging our ML capabilities to better develop oncology therapeutics, both by ourselves and with partners who can accelerate impact on patients

Neuroscience Discovery Using Cellular ML

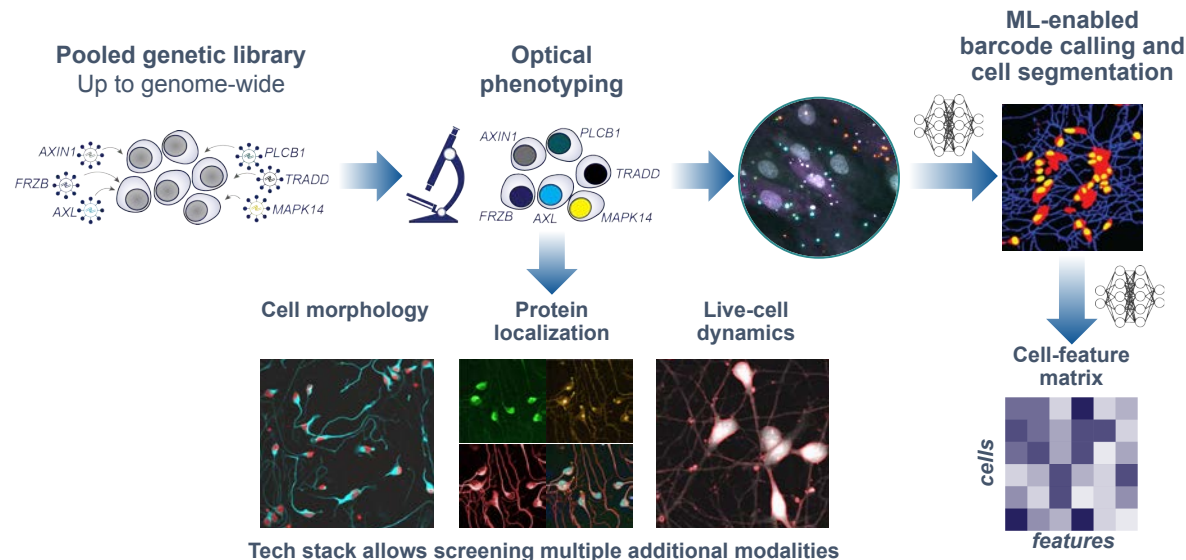


Imaging as a Foundational Modality for Cell-ML



High-Content Cellular Data Drives Biology Insights

Pooled Optical Screening in Human Cells (POSH) Platform



- High throughput gene editing, multimodal phenotypes, enabled by **advanced ML**
- Robust automation infrastructure in **~100k sq ft** of high-end labs
- **3 weeks** from disease model to modifier hits

Discovery in Tuberous Sclerosis Complex (TSC)

Causal Genetics	TSC1/TSC2 LoF
Patients	50K US; >1M WW
Symptoms	Tumors, epilepsy, neurological impairment
Unmet Need	Everolimus (Rapalog): mTOR inhibitor that modulates seizures in ~50% of patients



Platform-driven Biology and Drug Discovery in TSC

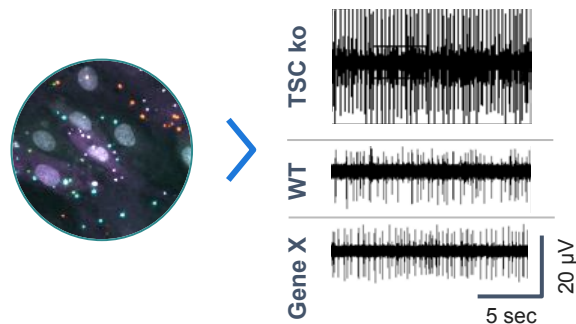
Build ML-enabled *in vitro* disease models

- Engineered iPSC neurons
- Multimodal phenotyping with imaging and transcriptomics
- ML to derive TSC disease axis



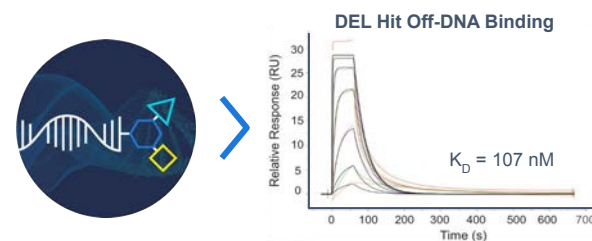
Identify and credential potential targets

- POSH¹ screens on ML disease axis
- ML selection of reverting targets
- Highly translatable functional electrophysiology assay
- >40% of targets reverted disease



Advance drug discovery

- Chem-ML drug discovery platform
- Hit finding: DELs + docking + ML
- Hit-to-lead: ML-enabled MPO
- Two novel targets in drug discovery



TSC Pioneered insitro's Cell-ML Platform

**Built differentiated
Cell-ML capabilities**

Automated, ML-enabled, high-content phenotyping and pooled genetic screening

**High-confidence, novel
targets emerged**

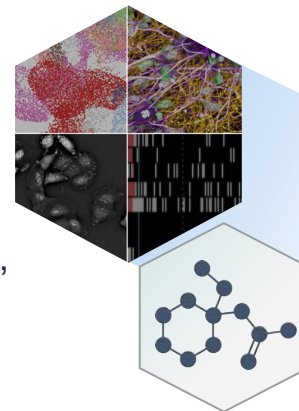
Targets credentialed in translatable assays, moving to drug discovery

**Unlocked rapid discovery
in 3 new genetic epilepsies**

From start to POSH-based modifier screens in ~3 months

**Strong foundation for
novel target discovery in
polygenic disease**

Familial disease variants as anchor points to build models that inform on sporadic disease

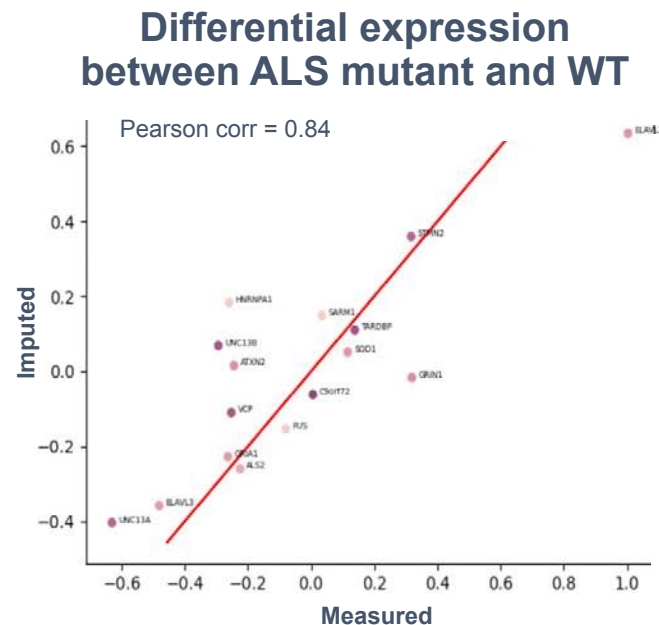
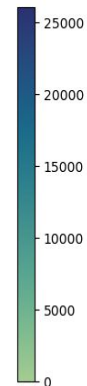


Discovery in Amyotrophic Lateral Sclerosis (ALS)

Causal Genetics	~10% of cases caused by ~25 known genetic drivers; 90% sporadic cases
Patients	~6000 new diagnoses WW per year
Symptoms	Progressive motor neuron death that leads to paralysis of nearly all skeletal muscles and death in 3-5 years
Unmet Need	Approved drugs provide only modest slowing of decline and ~3-5 months increase in lifespan



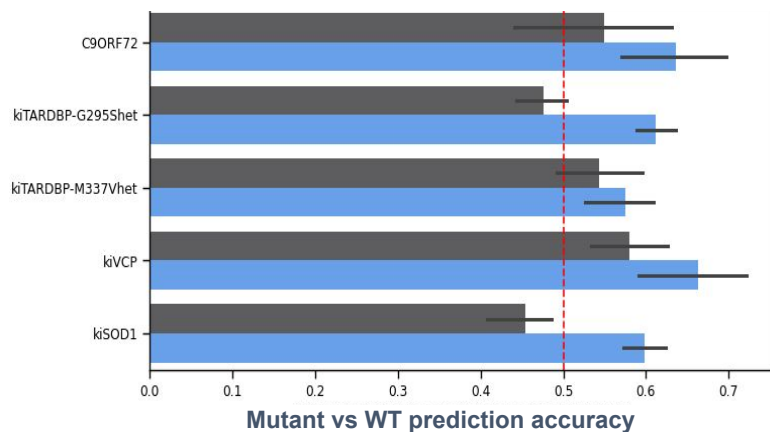
Work in collaboration between insitro and our partners at  Bristol Myers Squibb



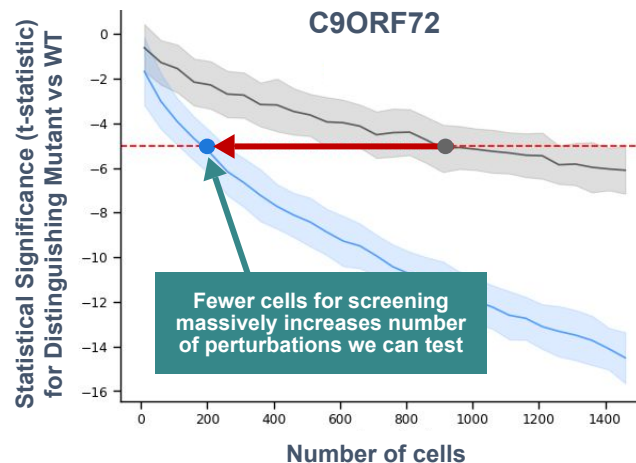
Imaging + ML embeddings allow imputation of disease-relevant molecular covariates

Building an ML-enabled ALS Disease Axis

ML disease axis more accurately distinguishes mutant from WT across familial ALS variants ...



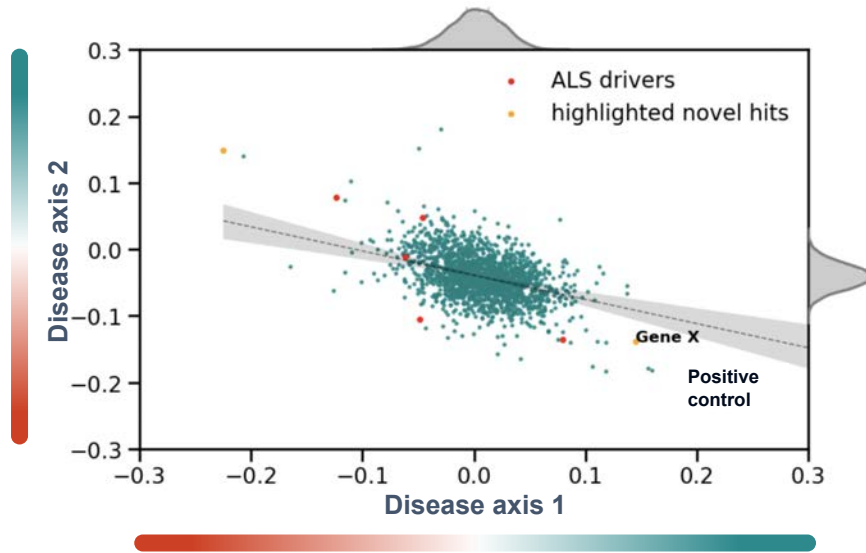
... which increases discrimination power and hence screen efficiency



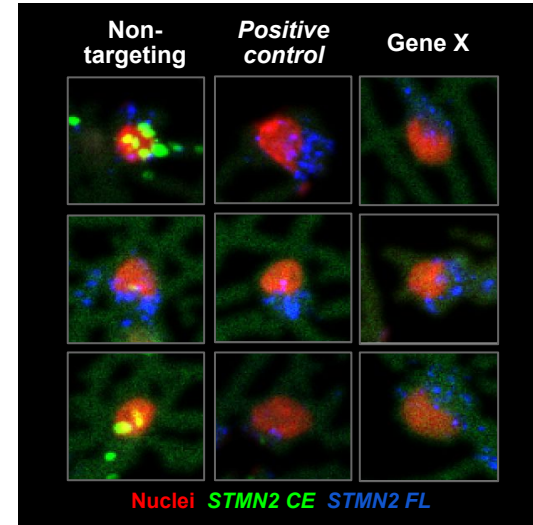
■ TDP-43 mislocalization anchor phenotype

■ ML Image-based disease axis

Pooled Screening Identified Modifiers of ALS *in vitro*



Representative tiles



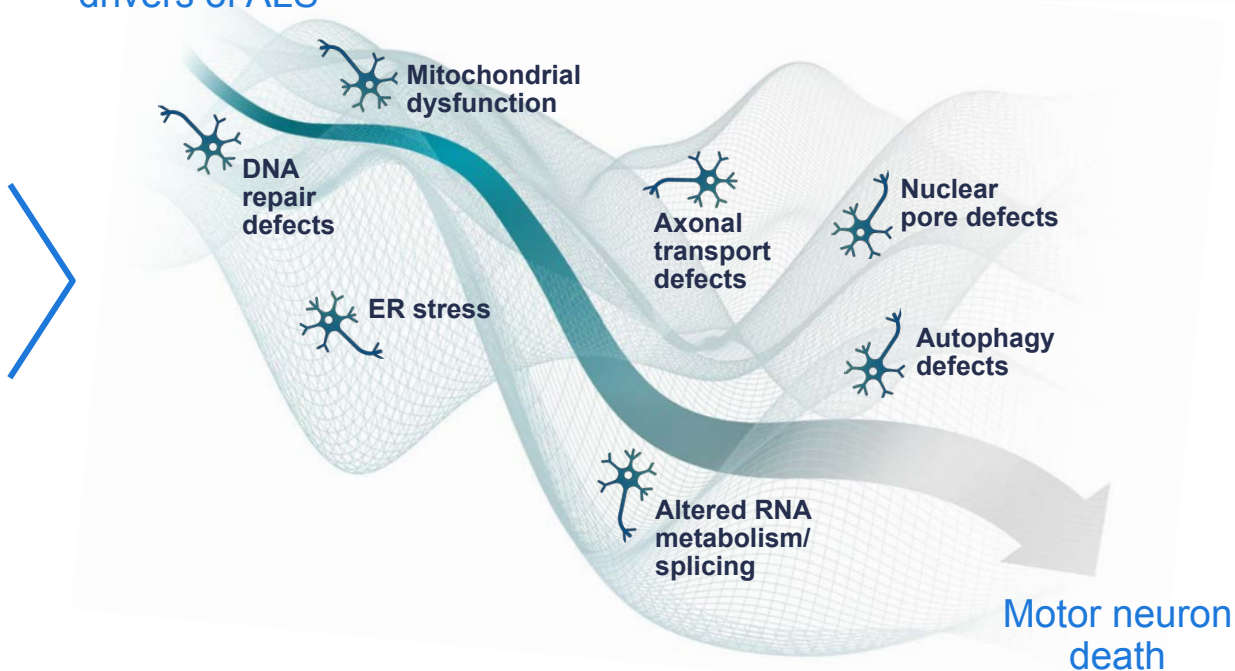
Pooled optical screening on iPSC-derived motor neurons identified novel gene knockouts that modulated ALS pathophysiology *in vitro*

Towards Novel Targets in ALS

We have established:

- >200 isogenic iPSC lines spanning ALS genetics
- Robust, automated generation of motor neurons in 10 days
- Infrastructure to capture and analyze multimodal phenotypic data at scale
- POSH screening to find genes that revert disease phenotypes
- Scaled processes for target credentialing

Known genetic drivers of ALS

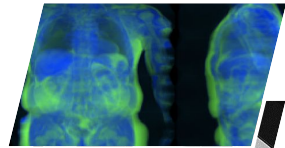


Metabolic Disease Discovery Using Clinical & Cellular ML



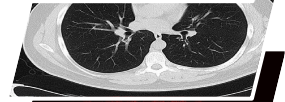
Integrated Approach for Liver Discovery and Validation

Integrated, multimodal
discovery & validation stack

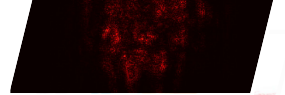


Body MRI (fat)

Abdominal MRI



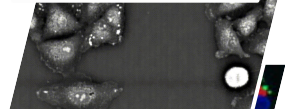
DXA



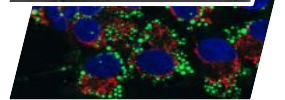
Biomarkers &
metabolites



Live cell



Fixed cell



Experiments
of Nature

Human
cohort
data

Clinical insights
(genes/pathways)

Prioritized
targets

Machine
Learning

Build the right models
for validation &
discovery

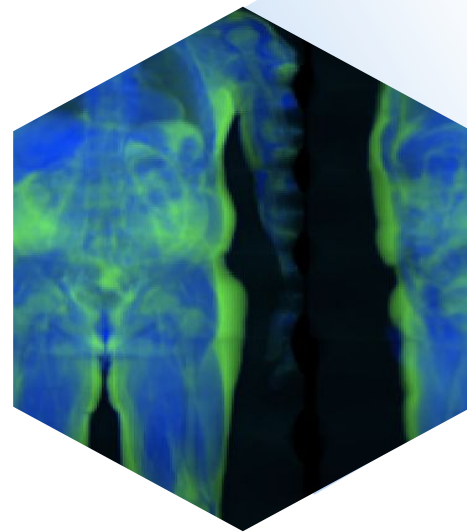
Cellular
model
data

Knockout
screens

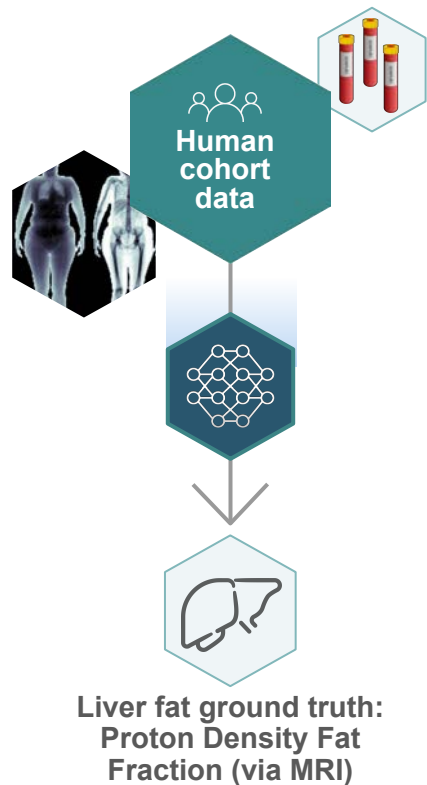
Experiments
in Lab

Discovery in Metabolic Dysfunction-Associated Steatotic Liver Disease (MASLD)

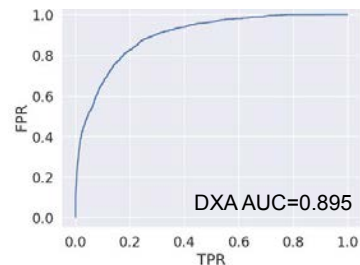
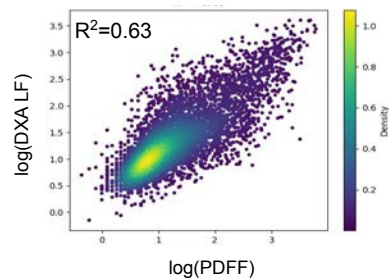
Disease Biology	Aberrant accumulation of liver fat Limited understanding of disease genetics, with most therapeutics targeting same small set of genes
Unmet Need	~30M patients in G7 countries alone who do not qualify for weight loss drugs (e.g., GLP-1s)
Differentiated Strategy	ML-imputed liver fat gives rise to novel, genetically supported insights and a mechanism orthogonal to current therapies
Therapeutic Approach	Rapid development of liver targeted therapeutics with potential for ML-derived efficacy biomarkers



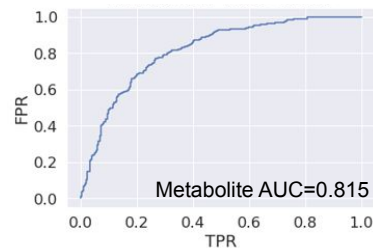
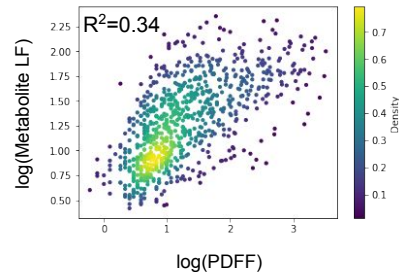
ML Enables Liver Fat Predictions from Multiple Data Modalities



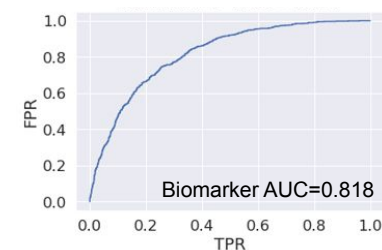
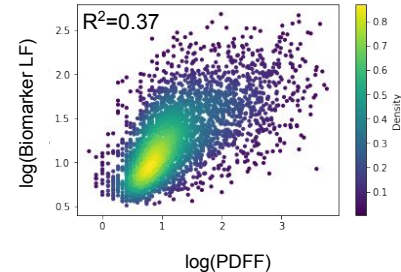
**DXA →
liver fat**



**Metabolites →
liver fat**

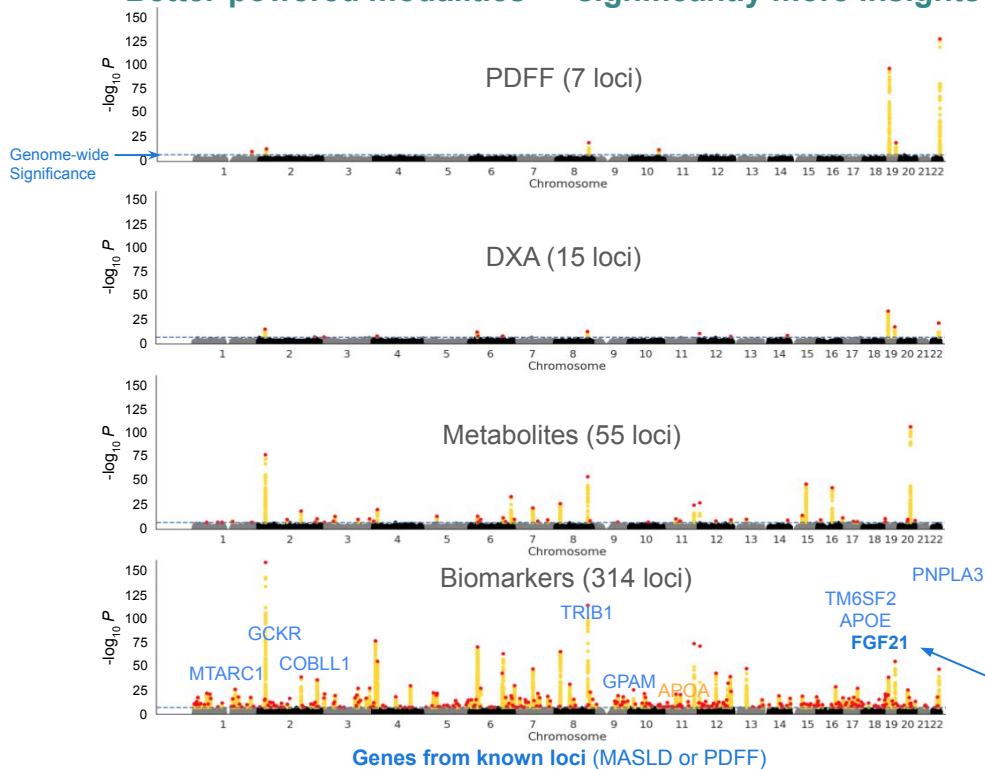


**Biomarkers →
liver fat**

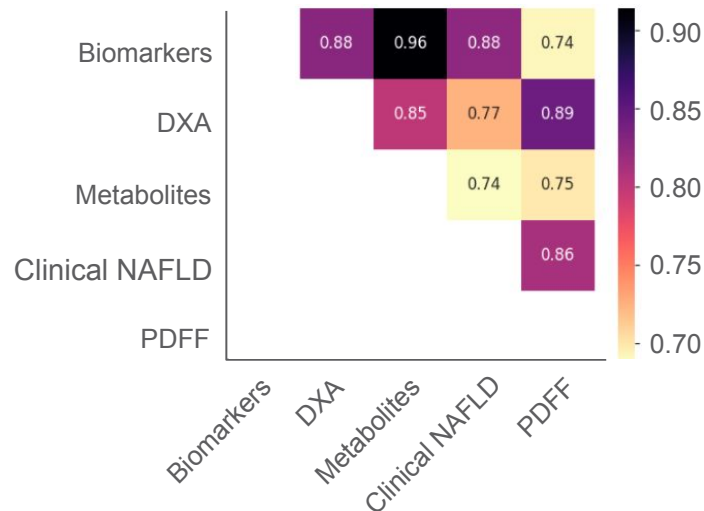


Liver Fat Imputations Empower Genetic Discovery

Better-powered modalities → significantly more insights



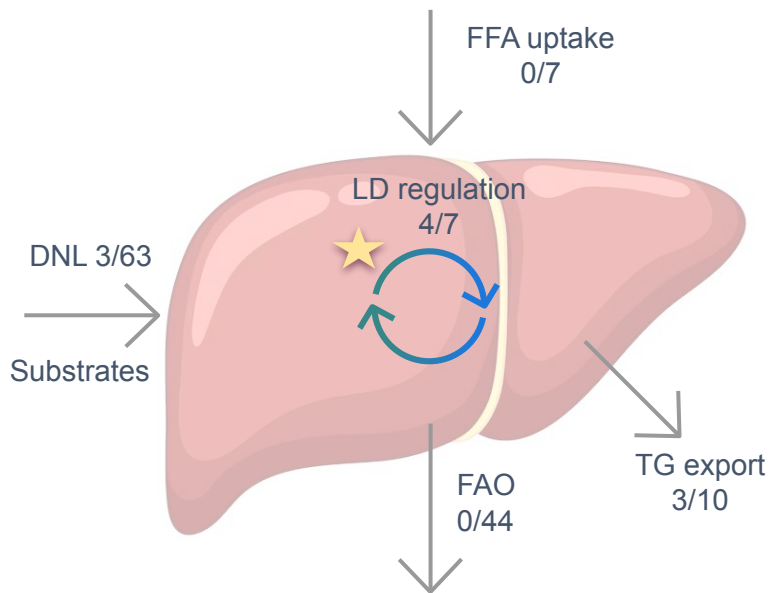
Correlated genetics across predicted phenotypic modalities support shared genetic basis



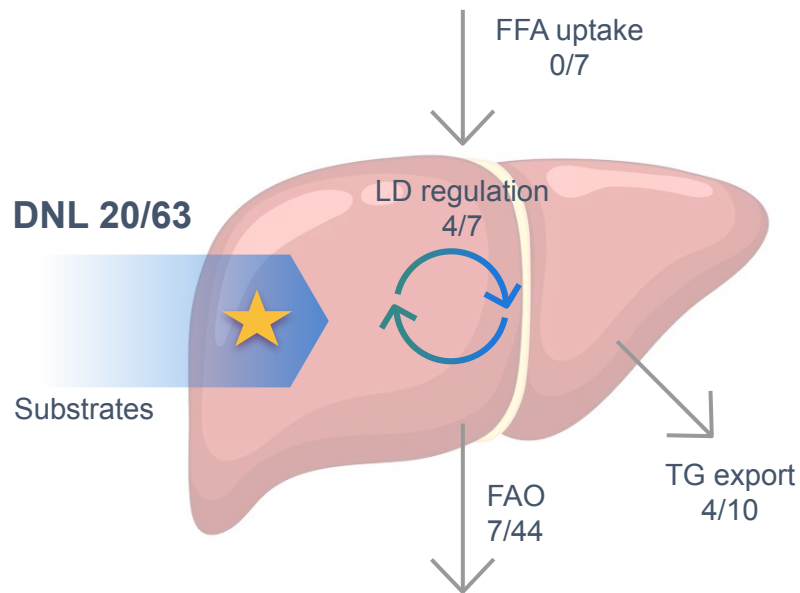
First human genetic support for FGF21

Novel Genetic Findings Transform Our Understanding of MASLD Pathophysiology

Prior literature genetic findings



insitro's genetic findings

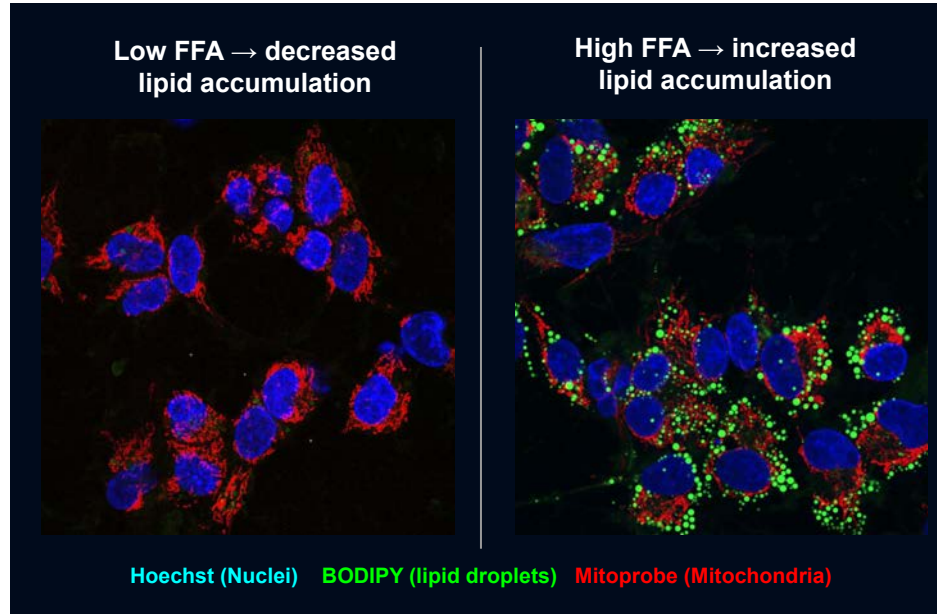


LD Regulation – lipid droplet regulation; **TG Export** – triglyceride export;
DNL – *de novo* lipogenesis; **FFA Uptake** – free fatty acid uptake;
FAO – fatty acid oxidation

Used Cell-ML Platform to Interrogate *de novo* Lipogenesis

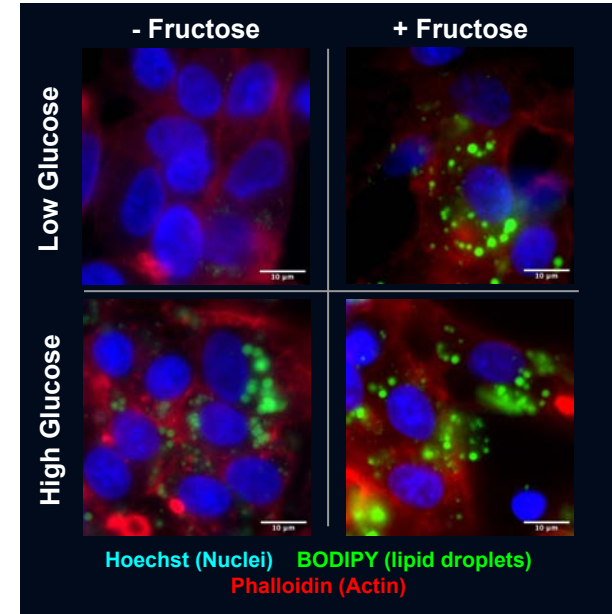
Creating pathway-relevant conditions and imaging phenotypes

FFA stressor



DNL stressor

High exposure of lipogenic substrates

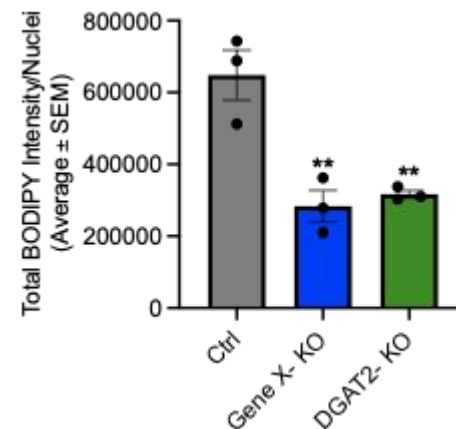
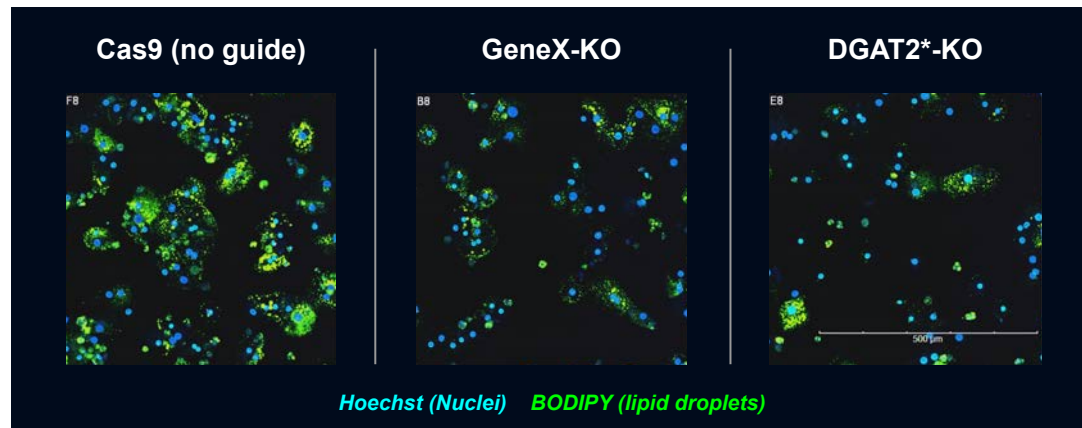


HepG2 cell system

DNL – *de novo* lipogenesis; FFA – free fatty acid

Novel Targets Modulate Lipid Accumulation *in vitro*

Lipid accumulation in pHep cells under DNL conditions



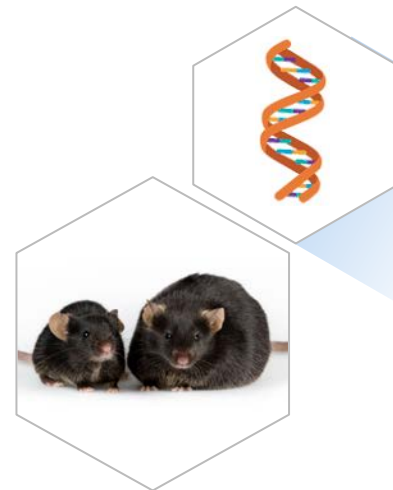
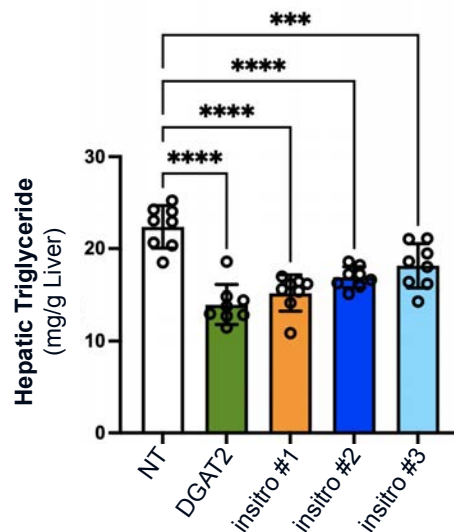
Fat reduction in primary hepatocytes with novel insitro targets comparable to DGAT2*

Targets Further Validated with siRNA in Obese Mice

insitro's novel targets

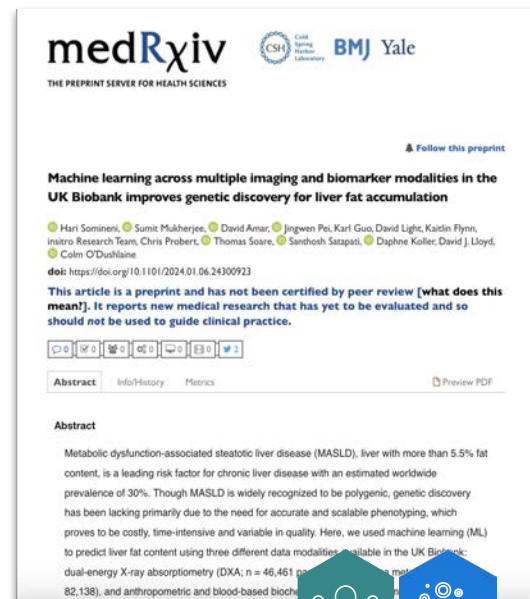
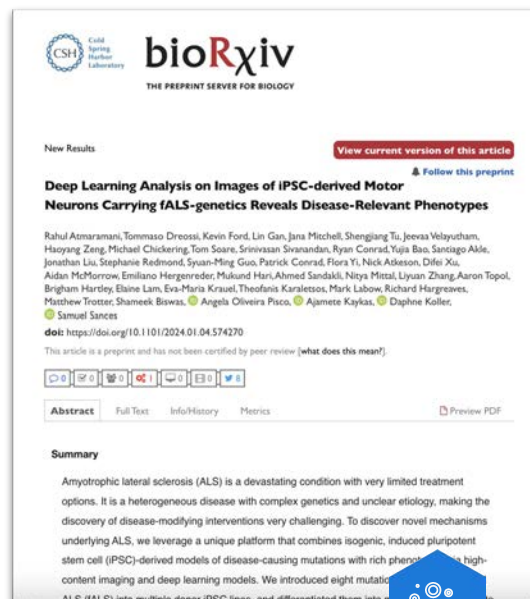
- Have strong genetic support
- Capture novel mechanisms targeting master regulators of *de novo* lipogenesis
- Show comparable triglyceride reduction to DGAT2, prior to compound optimization
- Form the basis for rapidly advancing, differentiated siRNA therapeutic programs

Liver fat reduction with siRNA



Targeting IND-enabling studies in 2024

Recent Publications Highlight Our Progress



Pre-prints in each of our three TAs posted January 2024

Looking Ahead



Experienced Senior Team Leading 260+ insitrocytes

Cellular Platform



AJAMETE "AJ" KAYKAS

Chief Exploration Officer and
Head of Neuroscience

ALLEN INSTITUTE NOVARTIS



MARTHA ROOK

Chief Technical Operations Officer

MILLIPORE SIGMA sigilon
therapeutics



KRISTINA GREEN

VP, Bioassays and Cell Based
Arrayed Screening

NOVARTIS MYOKARDIA RAPT



DAPHNE KOLLER
Founder & CEO

Calico coursera



Computation & Machine Learning



MATT RASMUSSEN

SVP, Software Engineering

Counsyl Myriad
genetics



COLM O'DUSHLAINE

Sr Director, Statistical Genetics

BROAD INSTITUTE REGENERON



JUSTIN LEE

Director, Imaging ML

Genentech PathAI



ANGELA OLIVEIRA PISCO

Director, Computational Biology

Cz KING'S
College
LONDON

Company Building



JEVAN SOO LENOX

Chief People Officer

STITCH FIX Square McKinsey
& Company



MARY ROZENMAN

CFO/CBO

aimmune Longitude
CAPITAL McKinsey
& Company



TAMARA TOMPKINS

Chief Legal Officer, General Counsel

amyris khosla ventures Neumora



ALLISON LAI

SVP, Accounting & Financial Operations

Genomic Health Adicet Bio pwc

Biology & Drug Discovery



PHILIP TAGARI

Chief Scientific Officer

MERCK AMGEN



ROBERT HILGRAF

VP, Molecular Design

BROAD INSTITUTE Celgene



SACHA HOLLAND

VP, Oncology

rigel CALITHERA



DAVID LLOYD

VP, Metabolic Disease

AMGEN CARMOT
THERAPEUTICS

S. Michael Rothenberg Joining as Chief Medical Officer



MASSACHUSETTS
GENERAL HOSPITAL
CANCER CENTER

LOXO
ONCOLOGY at Lilly

Pfizer



- Currently VP, Early Onc. Development & Clinical Research, Pfizer
- 20-year career across medical oncology, cancer research, and clinical trials in academia, biotech, and large pharma
- Significant contributions to many first-/best-in-class drugs
- Led development of selpercatinib (Retevmo, the first RET-targeted, pan-cancer drug) through NDA; elranatamab (Elrexio, CD3-targeting bispecific antibody) through PoC
- Innovative, accelerated clinical trial designs published in multiple high impact journals

“Mike's strong scientific acumen and expertise in designing and executing efficient trials will be invaluable in advancing the development of impactful medicines as Chief Medical Officer at insitro.”

— DR. SANDEEP MENON

Chief Development Officer at Alnylam
(former CSO, AI Digital & SVP, Head of Early Development Pfizer)

Support on Our 5+ Year Journey of Building insitro

Board Members



Daphne Koller
Founder & CEO



Paul McCracken
CPP Investments



Robert Nelson
ARCH



Vijay Pande
a16z



Roger Perlmutter
Independent Director
CEO Eikon, ex-President Merck R&D

Board Observers

Ziad Bakri
T. Rowe Price

Eli Casdin
Casdin Capital

Reid Huber
Third Rock

Krishna Yeshwant
GV

Top Tech and Biotech Backers

andreessen.
horowitz

ARCH
VENTURE
PARTNERS

BlackRock

CASDIN
CAPITAL
LIFE SCIENCE INVESTMENTS

CPP Investments



FORESITE CAPITAL



= SoftBank

TEMASEK

THIRD ROCK
VENTURES

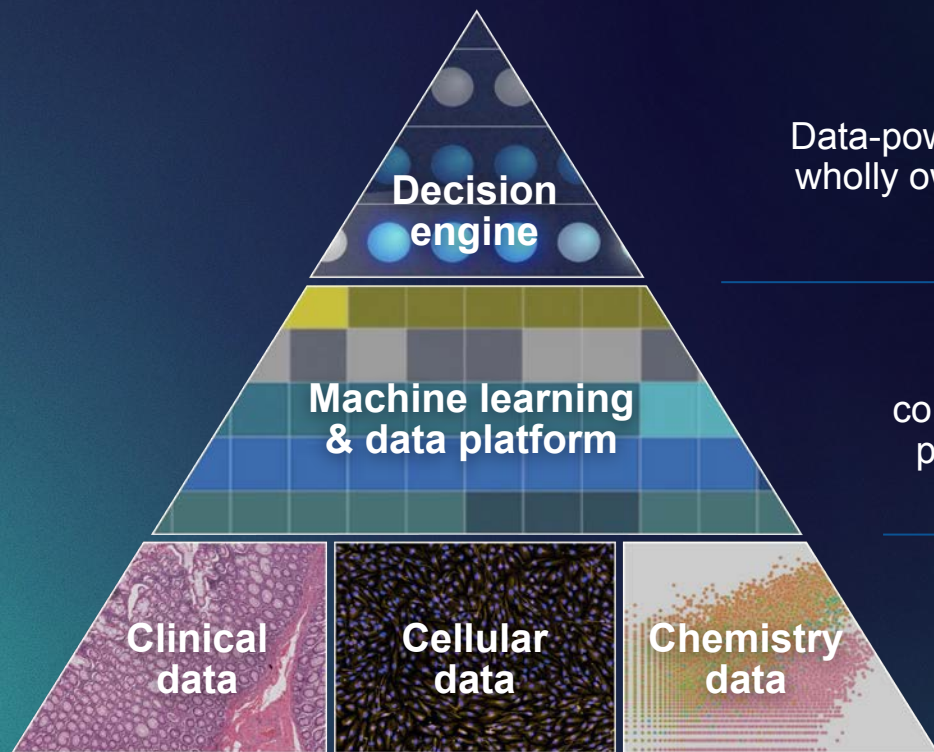
T.RowePrice



TWO SIGMA
VENTURES

- Raised >\$700M since founding in 2018
- YE2023 cash ~\$425M
- Strong operating runway into 2026+

Our Platform Is Powerful and Getting Better and Faster



Data-powered decisions create a strong pipeline of wholly owned and partnered therapeutic programs

Imputation-based data expansion and constantly improving ML predictive models produce high-confidence, novel insights

Multimodal data at scale across humans, cells, and molecules

Our Mission

PIPELINE

PLATFORM

To bring **better drugs faster** to the
patients who can benefit most
—— through ——
machine learning and **data at scale**



Thank you.

Learn more: www.insitro.com

