

AI for Biomedicine:

Lecture 2 *AI in biomedical research: opportunities and real-world applications*

Yasuko Iwakiri, PhD, Yale University School of Medicine

Croucher Foundation, November 27, 2025

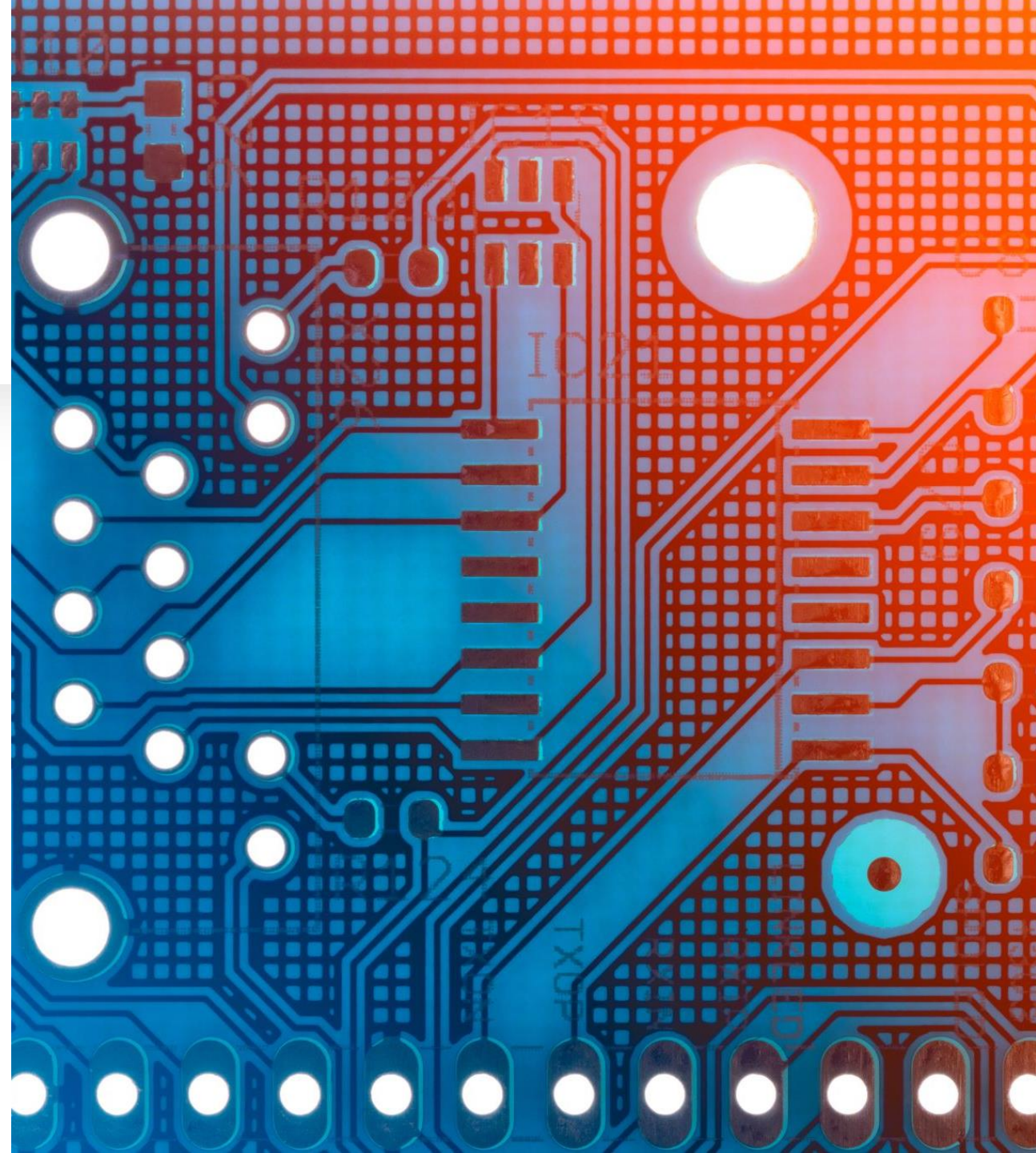


Croucher Foundation
裘槎基金會



Lecture 2 - Overview

- Introduction
- AI in next generation experimental models
 - Virtual cell model
 - Organoid
 - Organ-on-a-Chip
- AI-enhanced image analysis
- AI-driven integration of diverse data
- AI for drug discovery and tests
- Key takeaways



Section I

Recap of Lecture 1

Recap of Lecture 1



Research is moving toward more human-centered, AI-assisted experimental system



Animal models still help, but their limits are becoming clearer, esp. for human-specific biology



Biology is complex, and cross-species differences make translation hard



Data remain scattered across omics, imaging, and clinical platforms, which slows progress



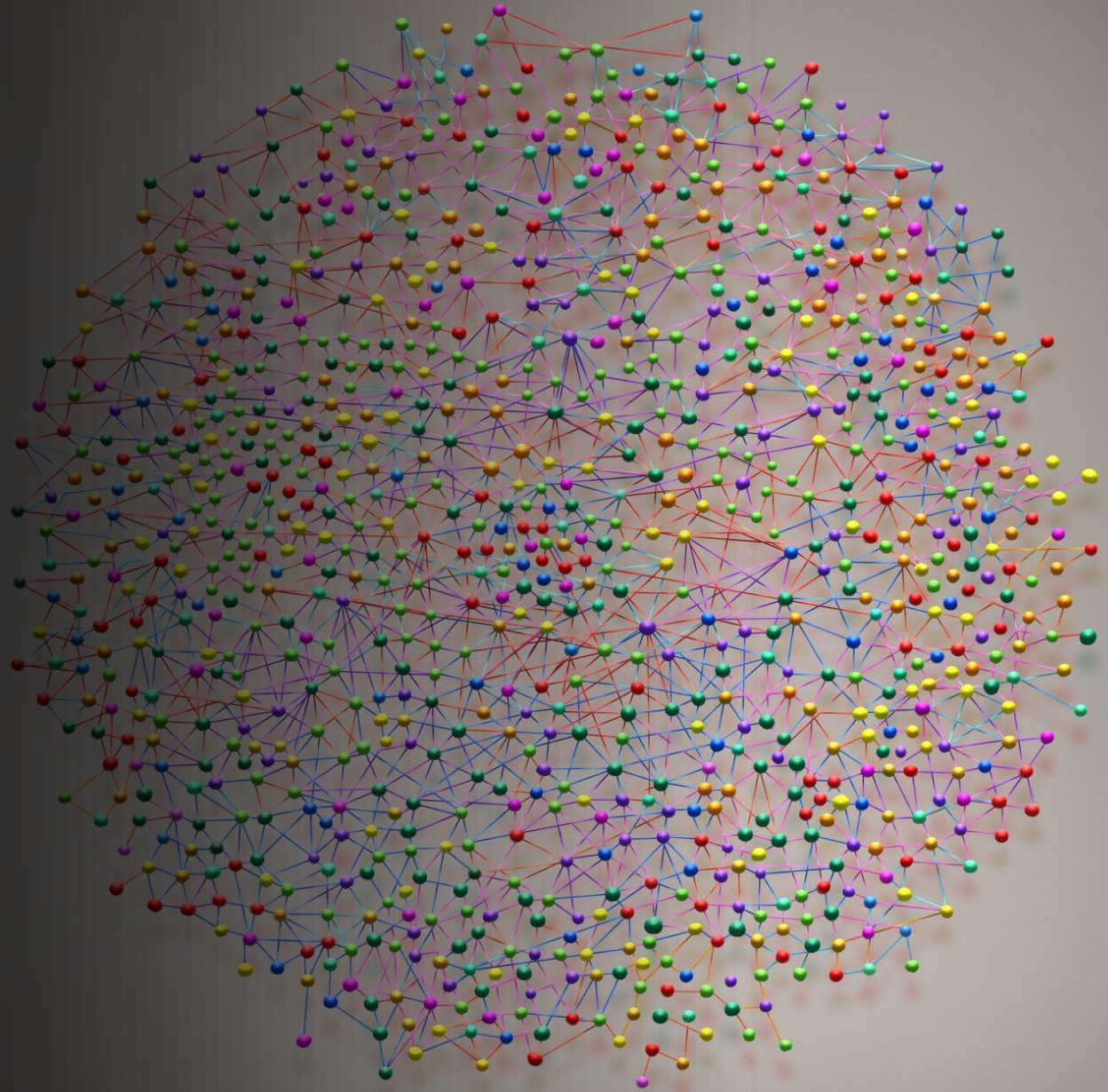
Lecture 2:

How AI supports experimental tools and show current examples.

Section II

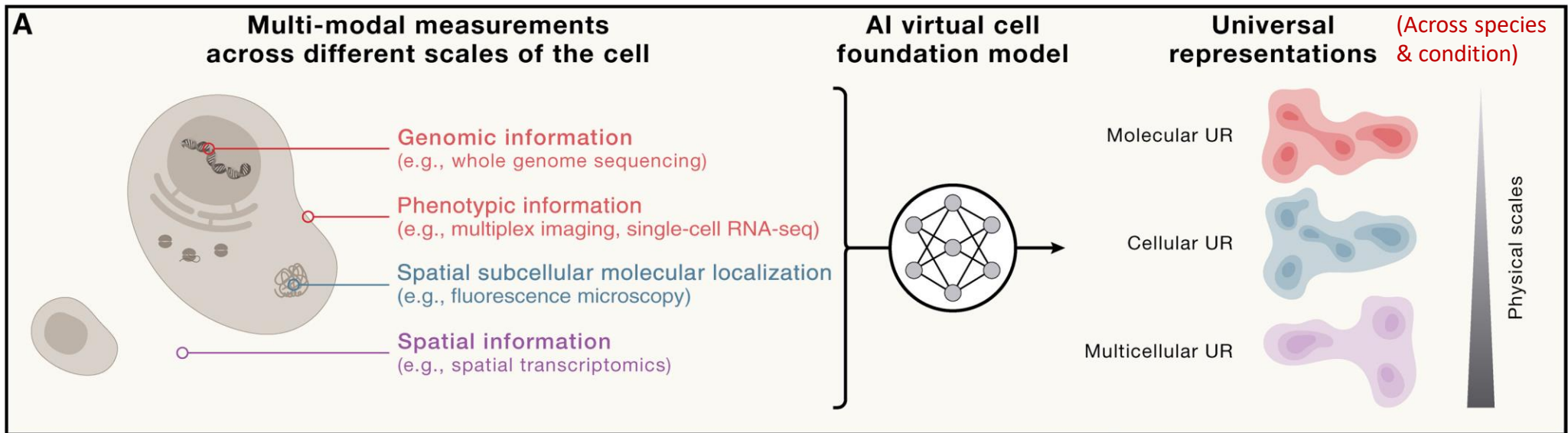
AI in next generation experimental models

AI virtual cell (AIVC)



AI virtual cell (AIVC) model:

- **A computational model** that simulates the biological functions and interactions of a cell
- *To understand a cell's function, scientists have attempted to construct virtual cell models to simulate, predict, and steer cell behavior*

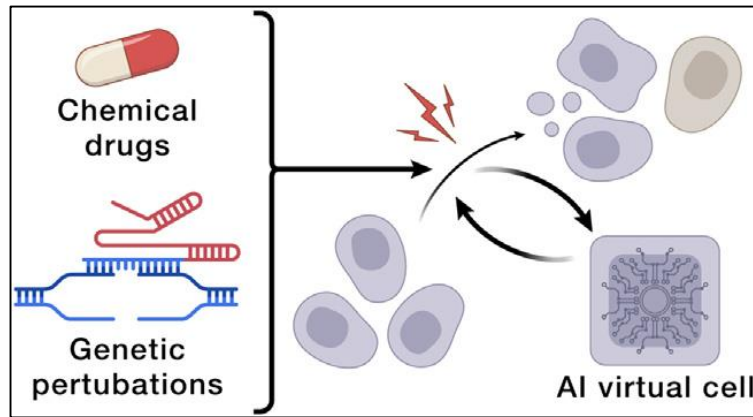


Reference: "How to build the virtual cell with AI: Priorities and opportunities"
Bunne C, ... Quake SR. *Cell*. 2025 (review). PMID: 39672099

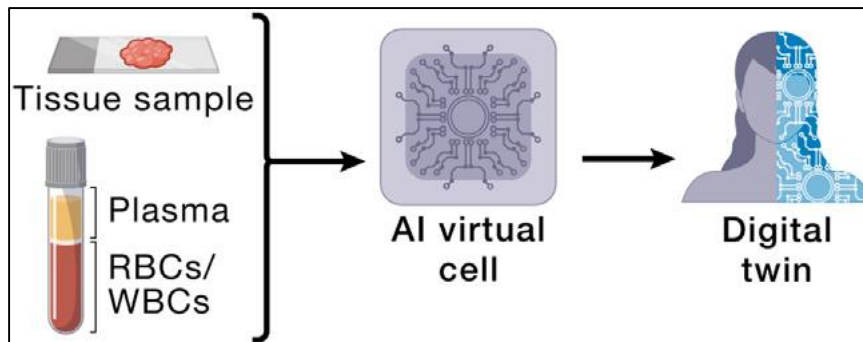
AI virtual cell model (AIVC):

Applications:

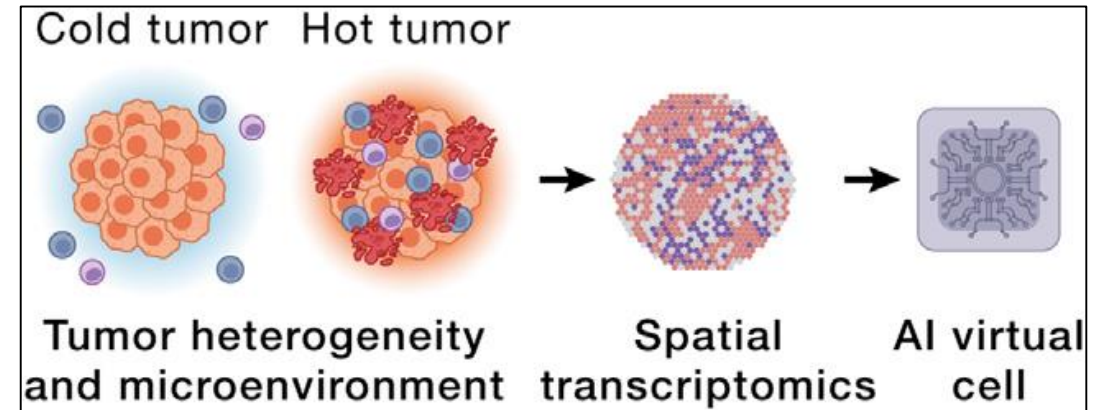
Drug discovery & cell-based therapeutics



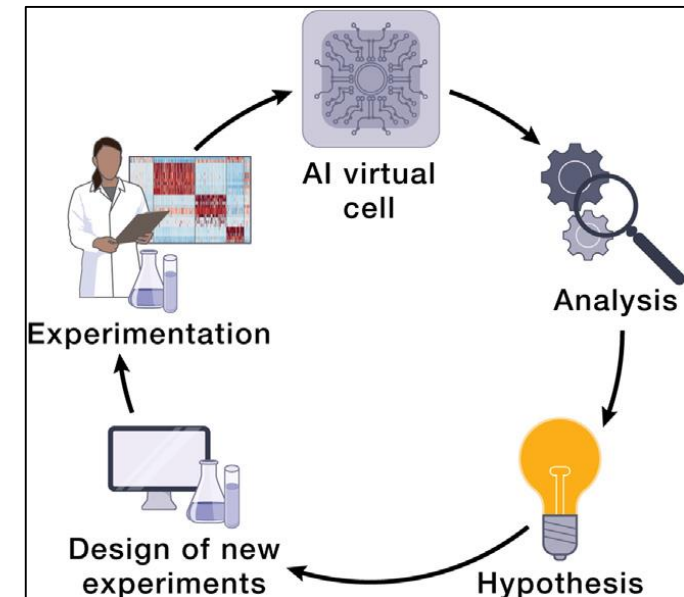
Personalized diagnosis for patients



Uncover cancer microenvironment

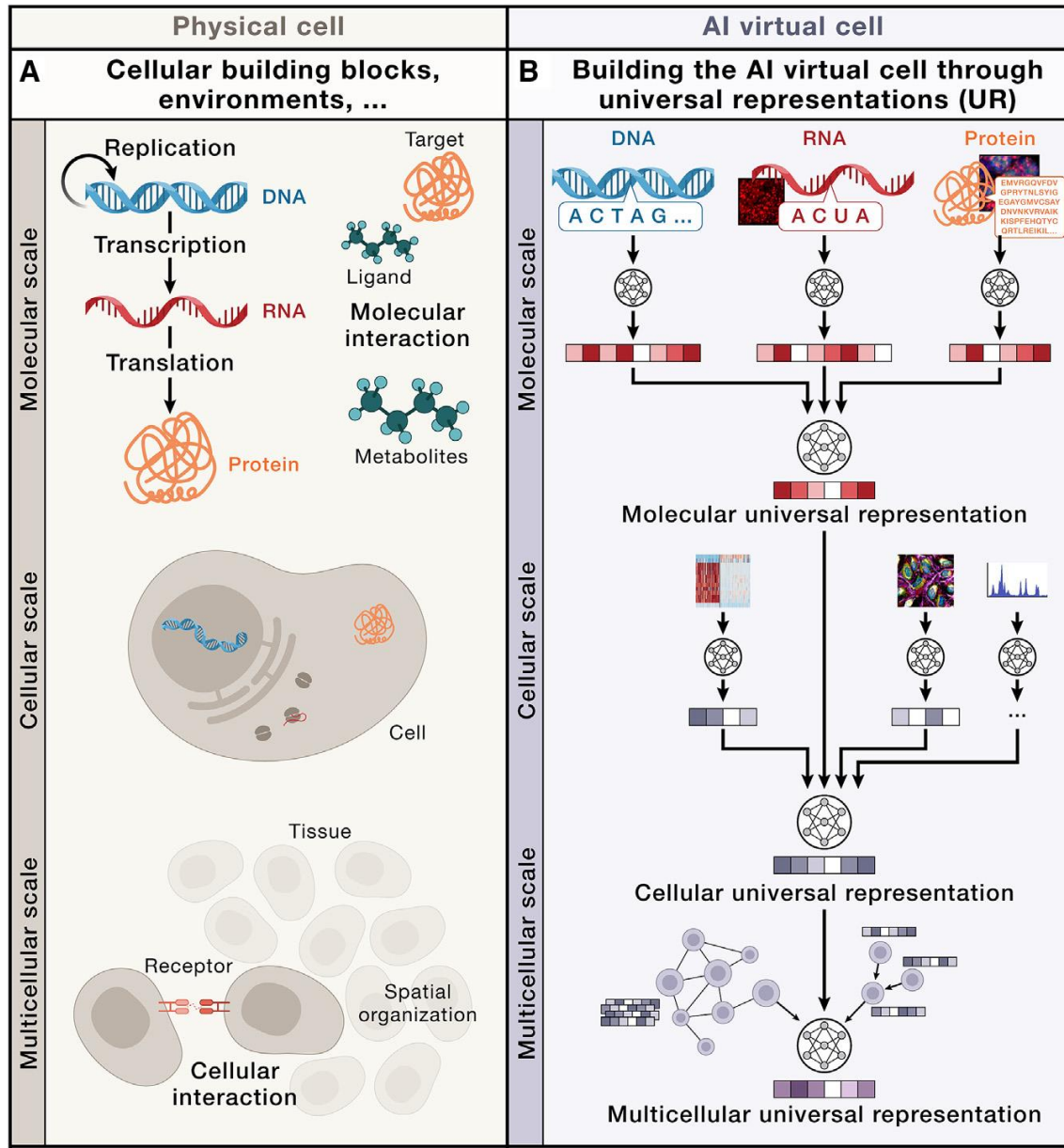


Hypothesis generation for scientific research



AI virtual cell (AIVC) model:

Building the AIVC:



1. Molecular level

2. Cellular level

3. Multi-cellular level

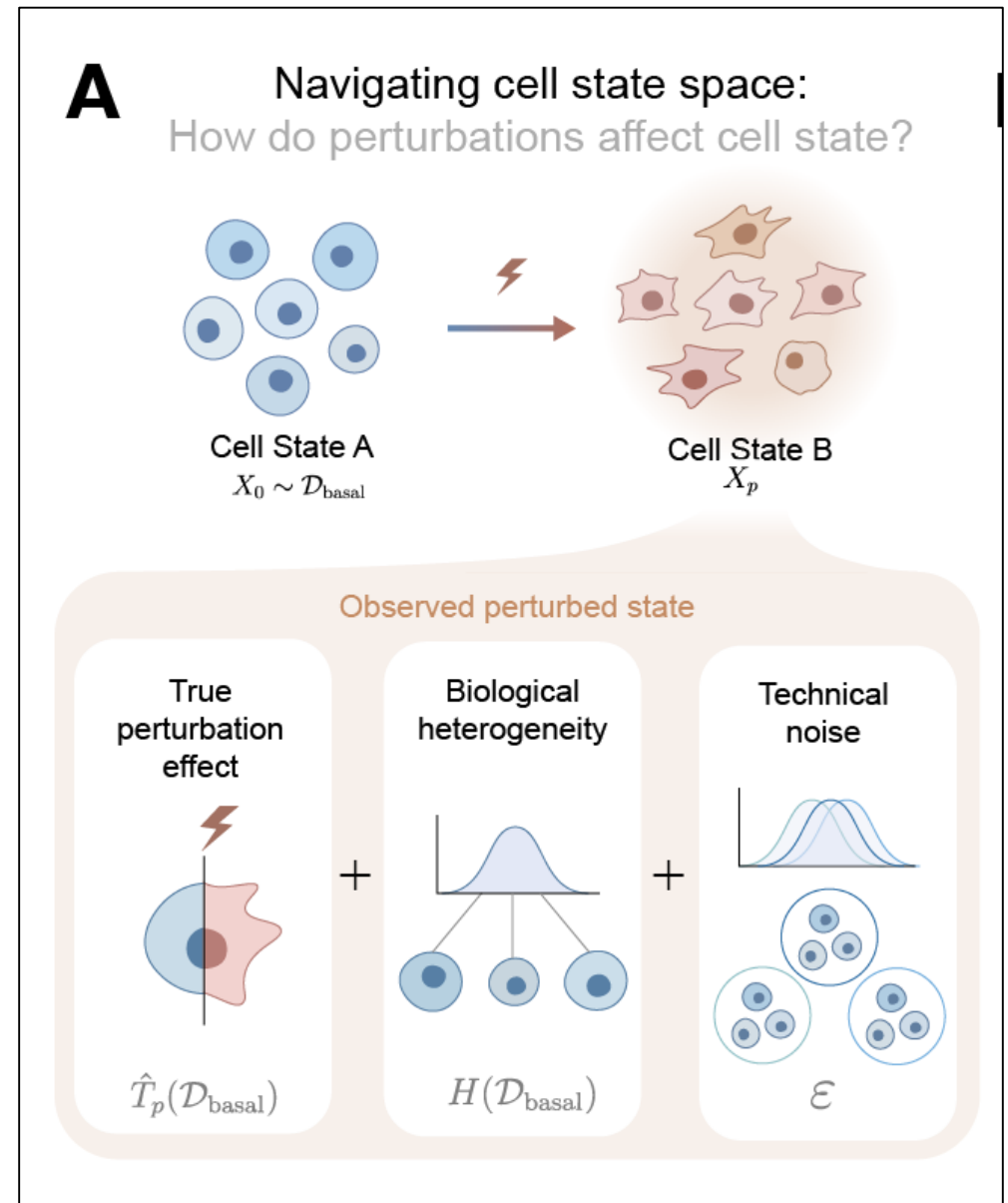
Predicting cellular responses to perturbation across diverse contexts with **State**

Adduri AK, ... Roohani YH. *bioRxiv*. 2025.

Unmet needs driving this study:

- Poor generalization to new cell types
- Heterogeneity masks true effects
- Technical variation adds noise
- Costly and hard to scale

→ *Large-scale screens are expensive and difficult to run in many contexts*

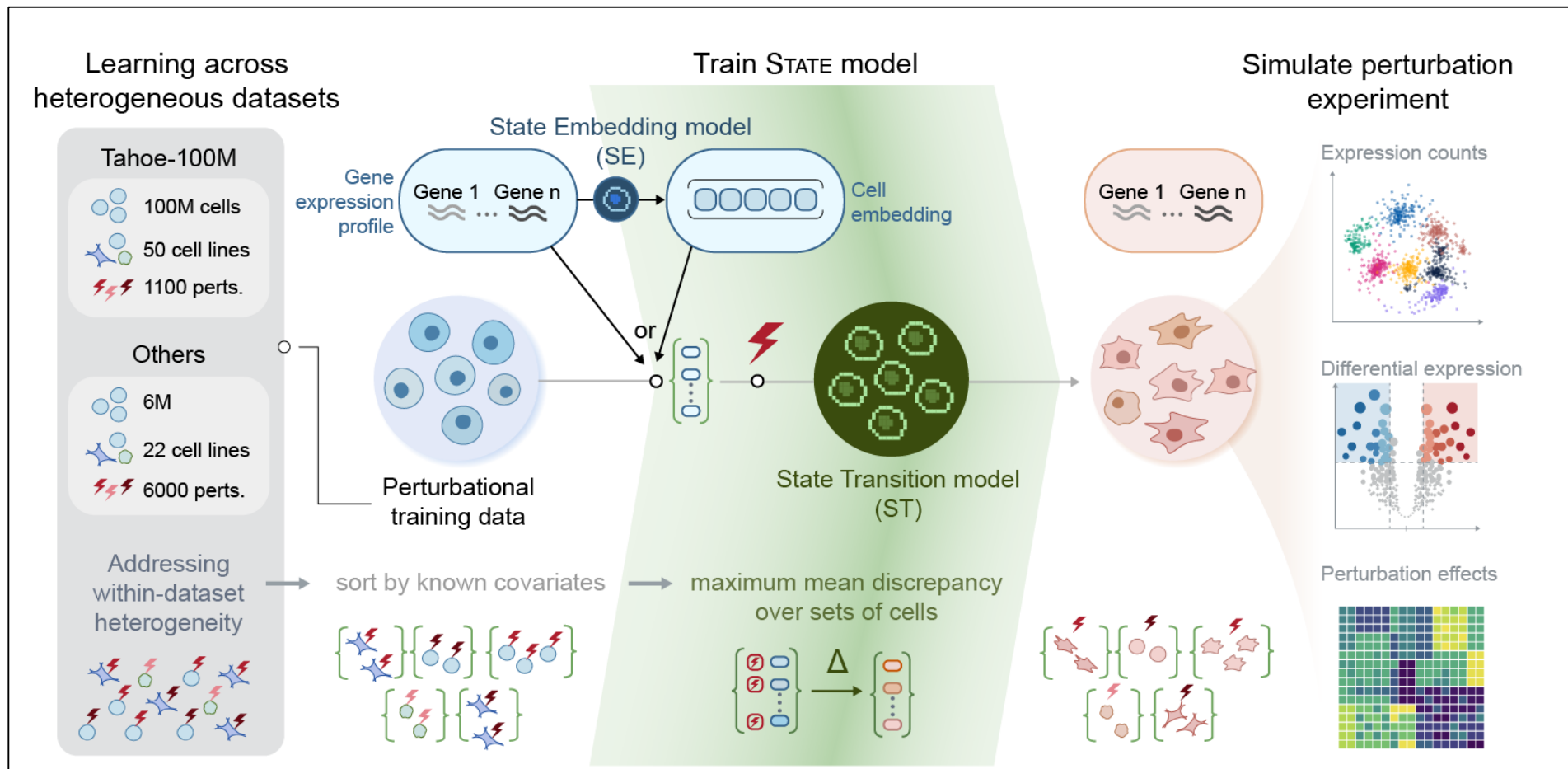


Predicting cellular responses to perturbation across diverse contexts with **State**

Adduri AK, ... Roohani YH. *bioRxiv*. 2025.

Methods

State: A multi-scale machine learning architecture that operates across genes, individual cells, and cell populations

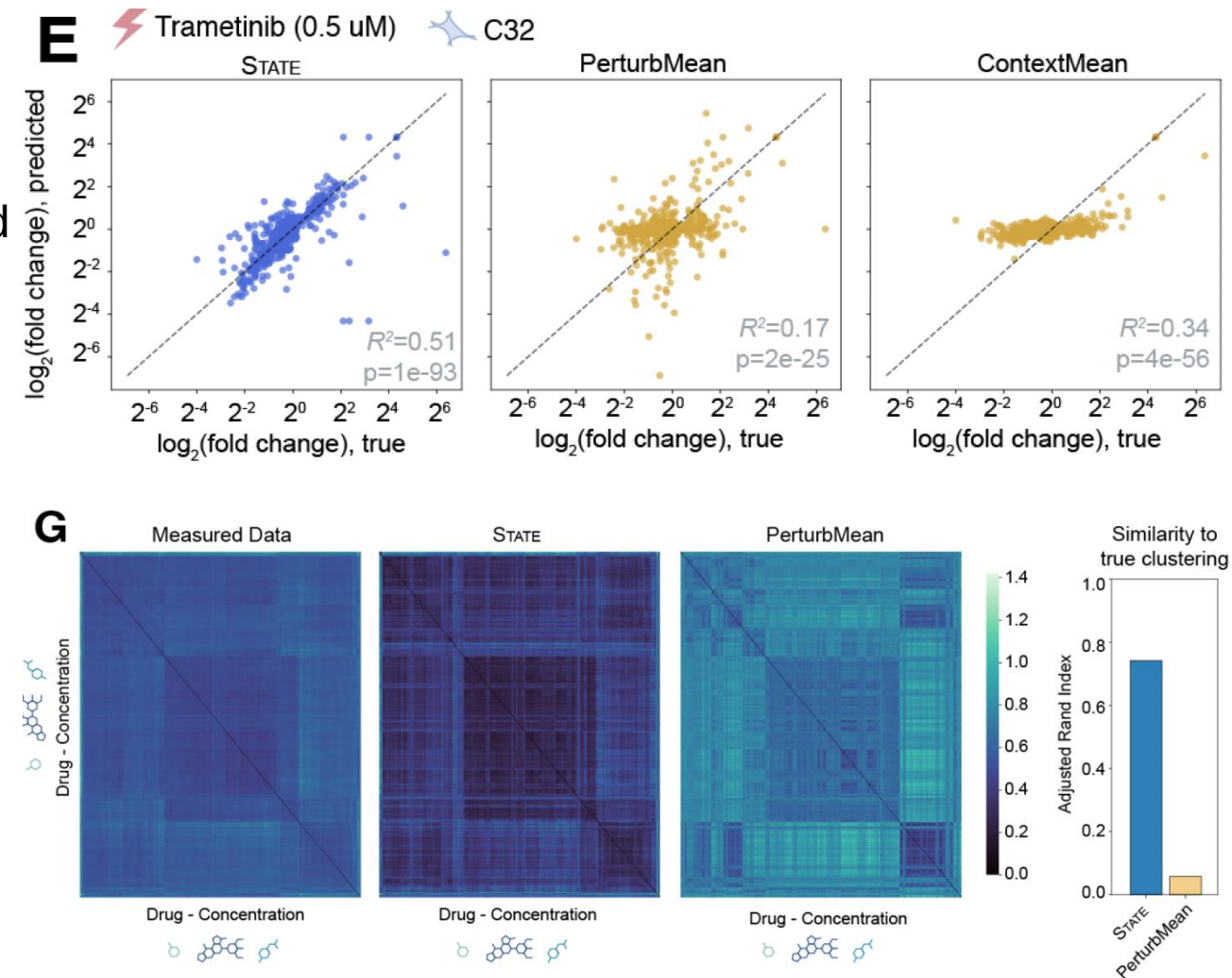


Predicting cellular responses to perturbation across diverse contexts with **State**

Adduri AK, ... Roohani YH. *bioRxiv*. 2025.

Key findings:

- Perturbation Similarity
 - Predicted and observed drug effects showed similar patterns
 - Similarity maps and heatmaps clustered perturbations in comparable ways
 - High alignment with experimental data (ARI > 0.7 vs < 0.1 baseline)
- Improved prediction of perturbation responses across datasets and cell types
- Captures cell type–specific responses and non-transcriptional outcomes (e.g., survival)



Organoid

Bringing AI into organoid research is an emerging area.



Organoid

Definition

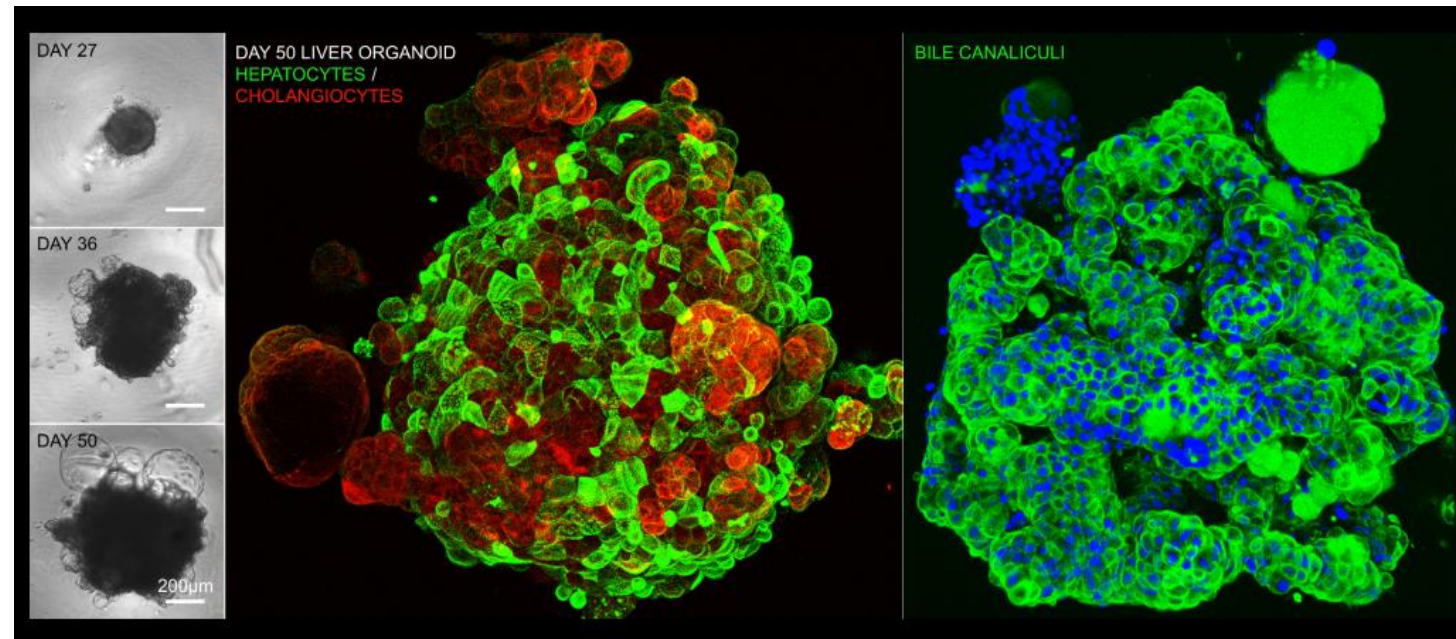
- A 3D structure of cells that self-organize through cell-cell and cell-matrix interaction

Origin

- Primary stem/progenitor cells
- iPSCs

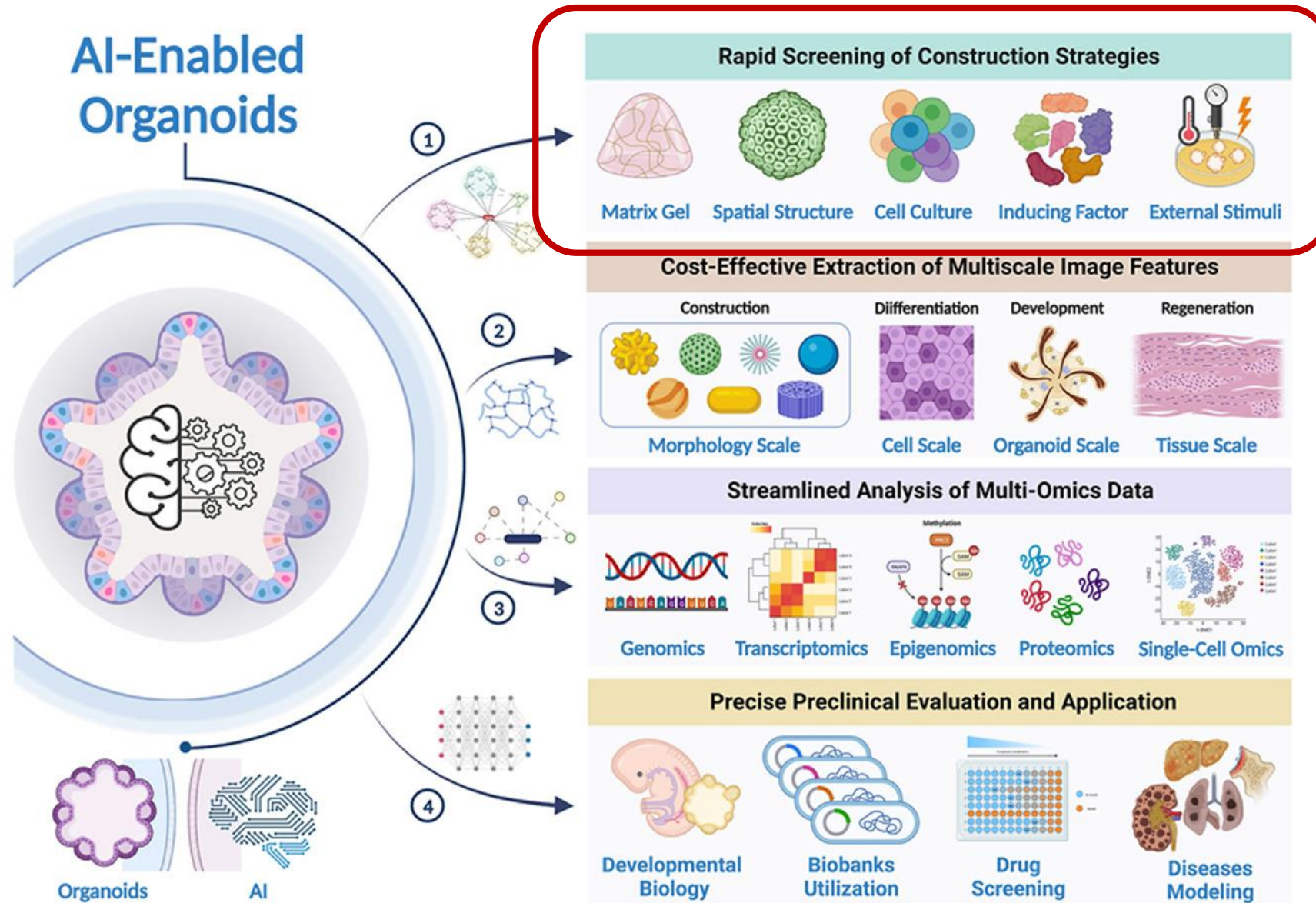
Applications

- Drug discovery, toxicity, mechanistic study

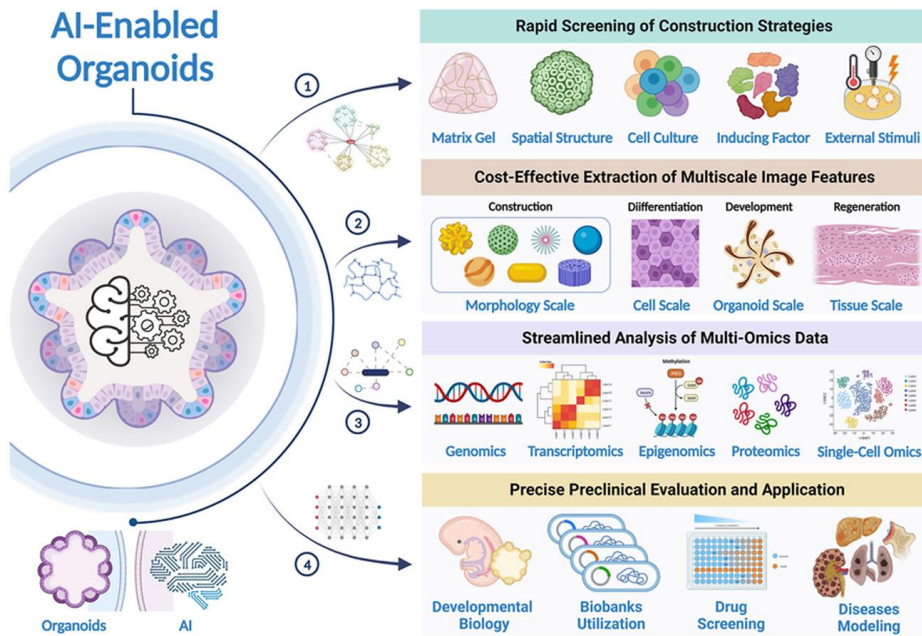


<https://www.a-star.edu.sg/News/astarNews/news/press-releases/singapore-scientists-develop-advanced-human-liver-organoids>

AI-enabled organoids: Construction, Analysis, and Application



AI-enabled organoids:



Bai L, ... Su J. *Bioact Mater.* 2023 (review). PMID: 37746662

1. Rapid screening of construction strategies

→ Business opportunity

Construction strategies:

Methodologies and techniques to create organoids

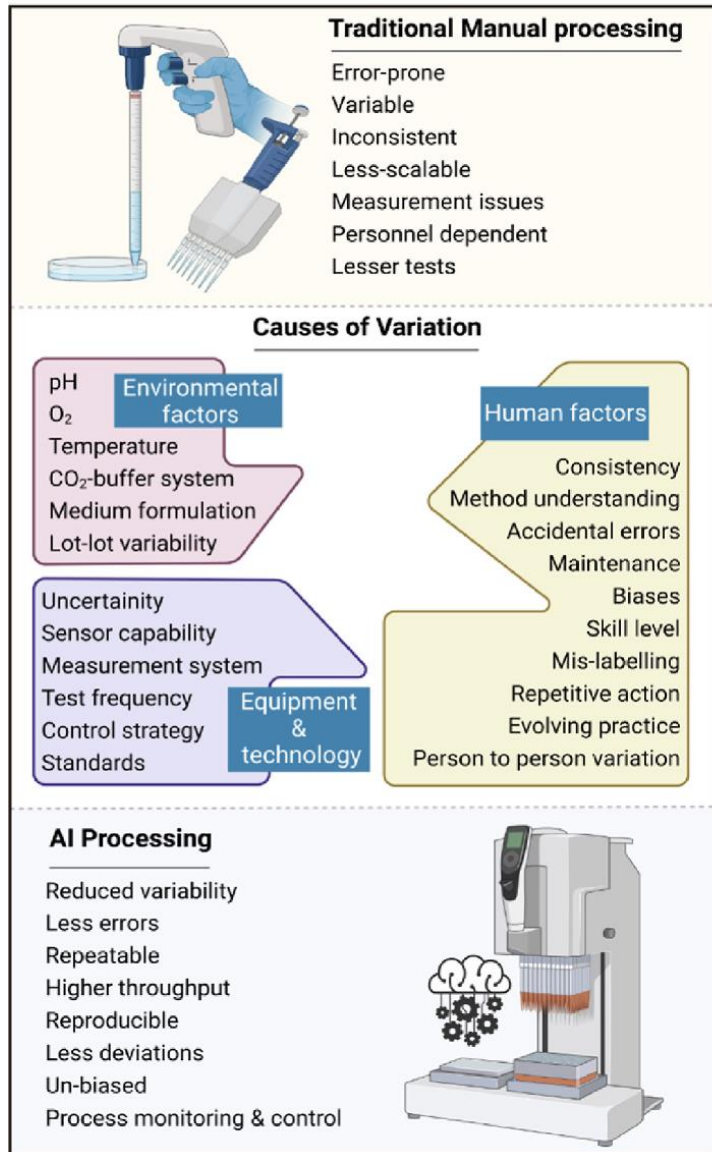
AI for optimization of these strategies:

- Methodologies and techniques to create organoids
- ML → identifying the most effective methods for synthesis and guiding structure of matrix gel, cell culture conditions and assessing external stimuli

Why AI is important?

- Can process a large volume of experimental conditions → optimal protocols

AI-enabled organoids:



1. Rapid screening of construction strategies

→ Big business opportunity!

Construction strategies:

Methodologies and techniques to create organoids

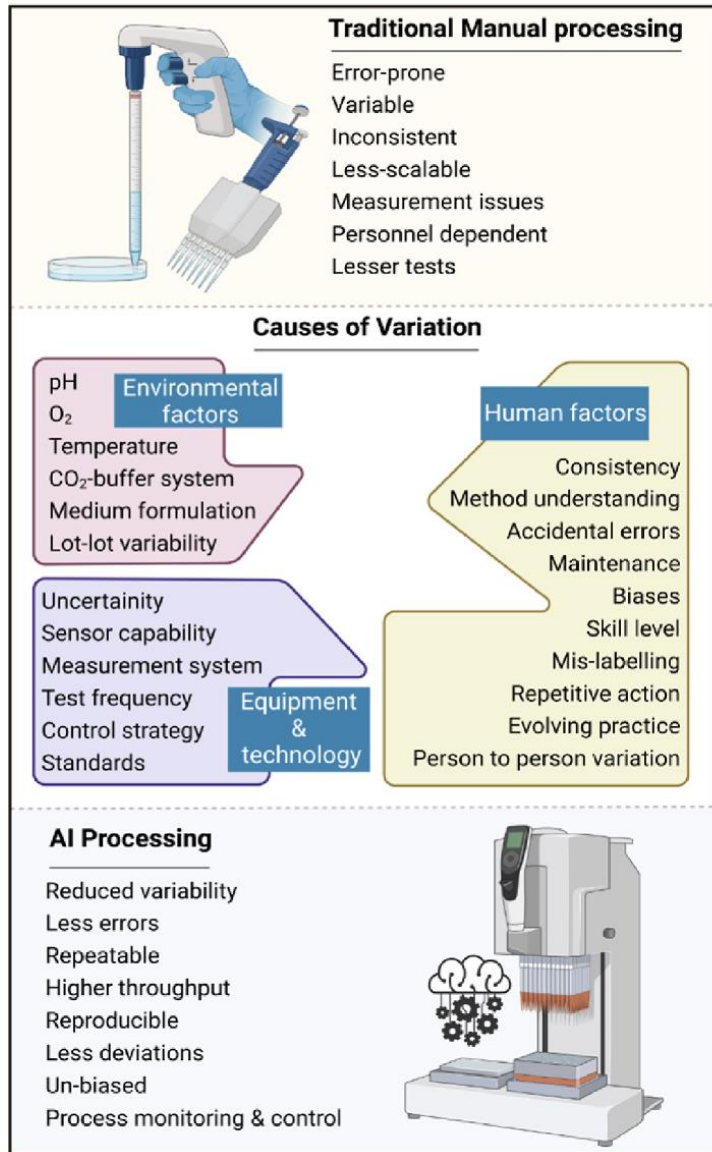
AI for optimization of these strategies:

- Methodologies and techniques to create organoids
- ML → identifying the most effective methods for synthesis and guiding structure of matrix gel, cell culture conditions and assessing external stimuli

Why AI is important?

- Can process a large volume of experimental conditions → optimal protocols

AI-enabled organoids:



1. Rapid screening of construction strategies

→ Big business opportunity!

Construction strategies:

Methodologies and techniques to create organoids

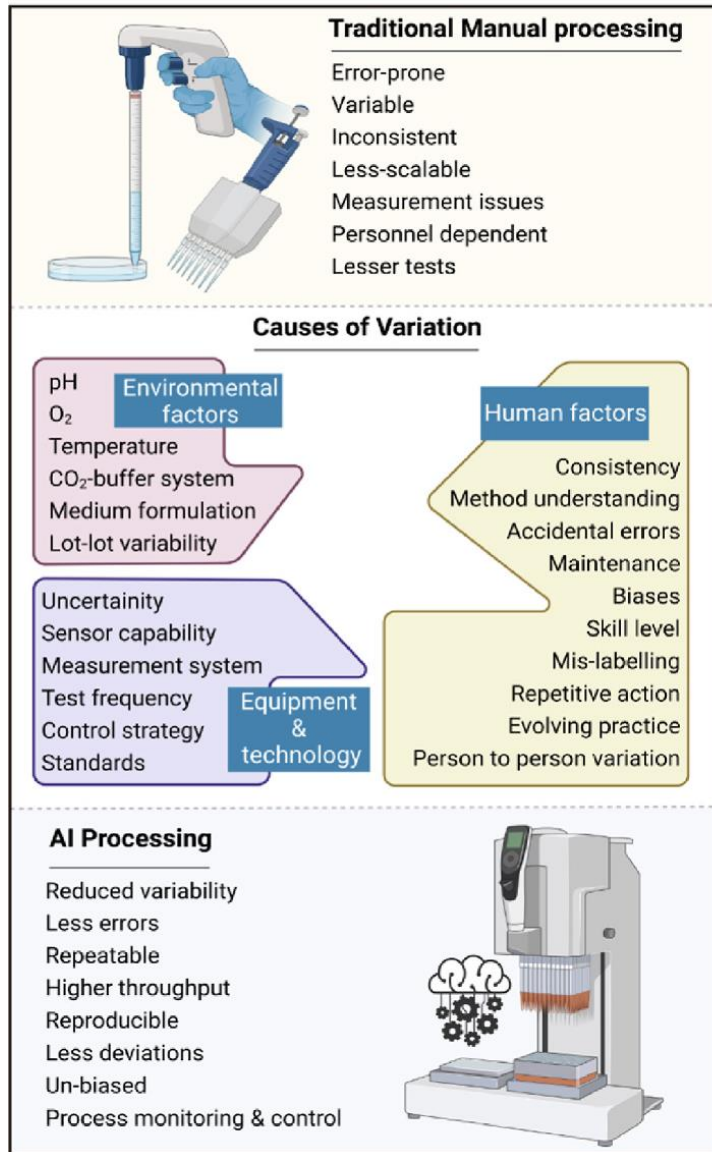
AI for optimization of these strategies:

- Methodologies and techniques to create organoids
- ML → identifying the most effective methods for synthesis and guiding structure of matrix gel, cell culture conditions and assessing external stimuli

Why AI is important?

- Can process a large volume of experimental conditions → optimal protocols

AI-enabled organoids:



1. Rapid screening of construction strategies

AI-enabled culture condition optimization

- Predict successful formation: early stop and refine the failures

Matsumoto R, ... Yamamoto T. Cell Rep Methods. 2025.

PMID: 40763741

- Predict optimal stimulus and factors used for culture

Shin J, ... Kim K. Adv Sci. 2025. PMID: 40287843

AI-assisted quality control

- Bright-field tracking: growth and drug-response analysis — **Organoid**

Matthews JM, ... Tay S. PLoS Comput Biol. 2022. PMID: 36350878

- 3D quantification of organoid culture condition and quality — **Cellos**

Mukashyaka P, ... Chuang JH. Nat Commun. 2025. PMID: 38114489

- Time-lapse imaging of organoid formation and maturation —

TransOrga-plus

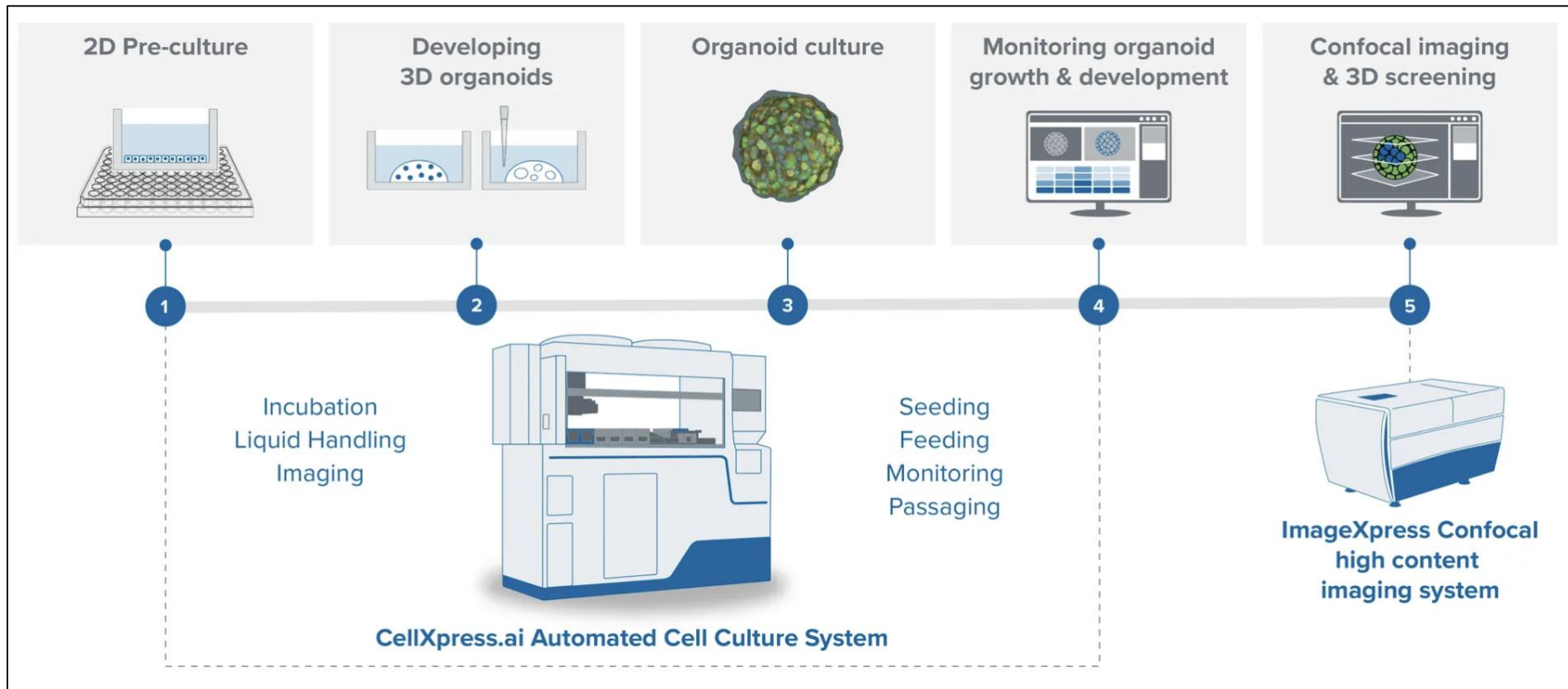
Qin Y, ... Tan F. BMC Biol. 2025. PMID: 41121276

AI-enabled organoids:

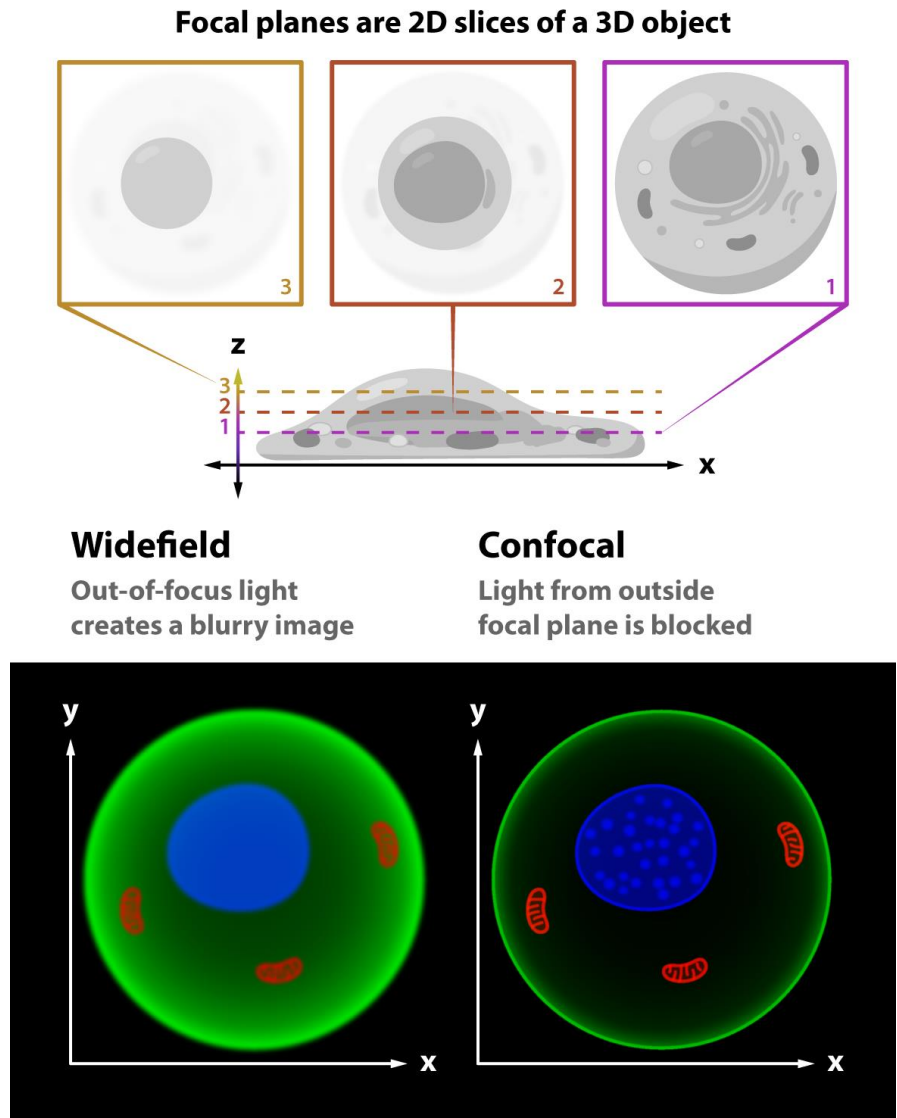
AI-powered automated culture system

CellXpress.ai by Molecular Devices, USA

- ML-assisted schedule of feeding and passaging
- ML-based culture quality evaluation



AI-enabled organoids:

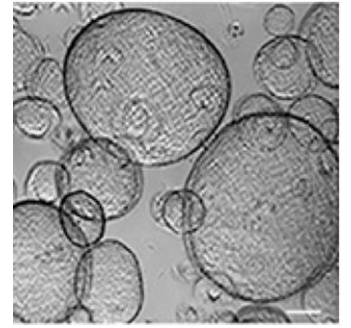


2. Cost-effective extraction of multiscale image features

Challenges of image extraction:

- A collection of images is at a single focal plane
- Differences in size and shape

Organoids



AI for automating image extraction process:

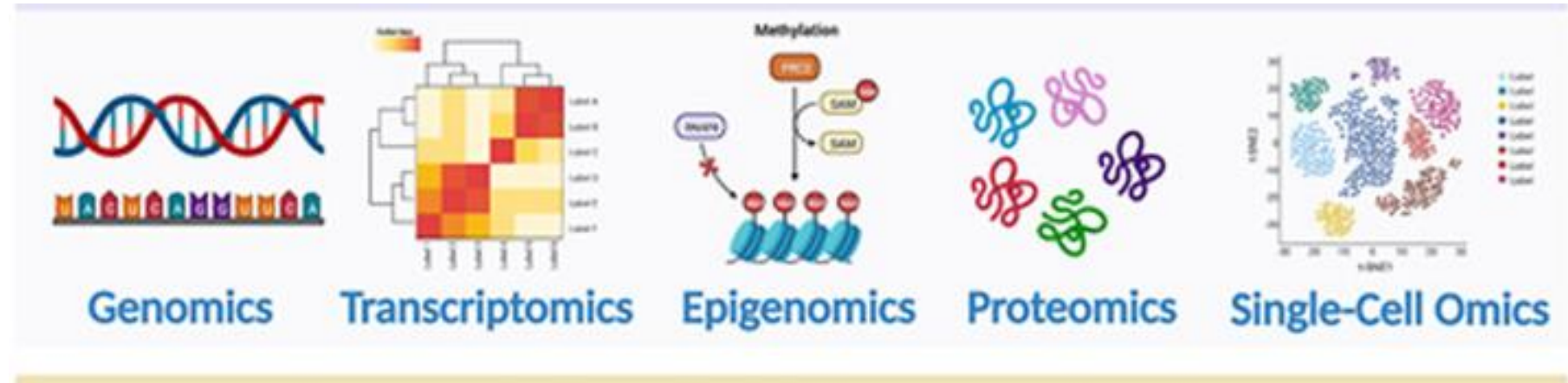
- ML needs to be developed for processing organoid images without changing their characteristics
- Enable adding functional features (e.g., shape changes)

Why AI is important?

- Can automate organoid image extraction, manage single-plane images and size variability, and still preserve key features so functional changes can be measured.

AI-enabled organoids:

3. Streamlined analysis of multi-omics data



Challenges:

- HT multi-omics data are large, diverse and hard to integrate.
- Relationships across omics layers are complex.

What AI/ML can do:

- Combine multi-modal data and detect patterns missed by traditional tools.
- Automate tasks like cell-type labeling and multi-modal predictions
- Extract key signals across molecular, cellular and organoid scales.

Why AI is important?

- Makes complex, high-dimensional data usable.
- Reveals biological relationships that are difficult to see manually.
- Supports more interpretable and efficient organoid analysis

AI-enabled organoids:

4. Precise preclinical evaluation and application



Mechanistic insight: Analyzes organoid growth and aligns organoid data with human tissues to reveal key developmental and disease pathways.

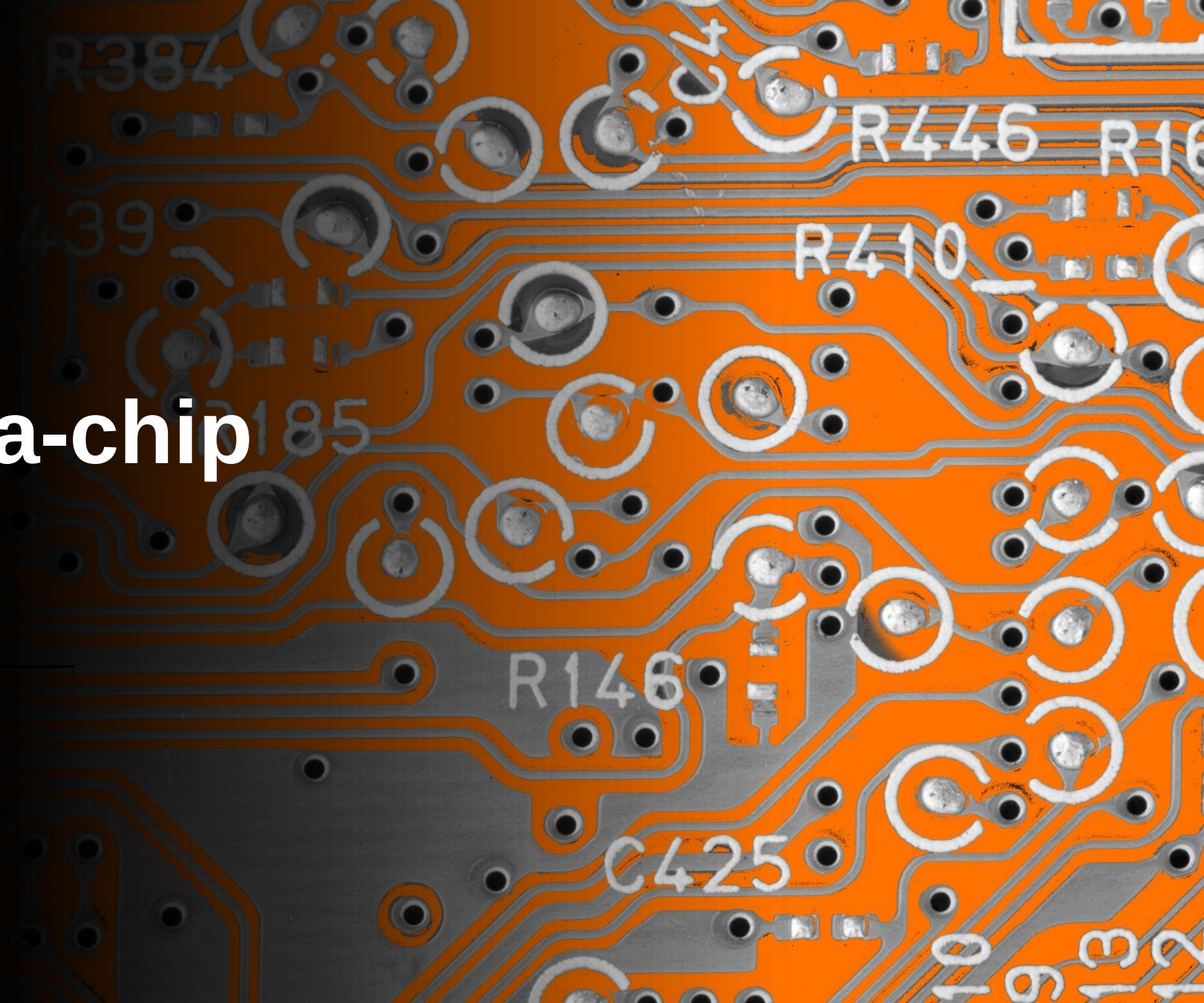
Biobank integration: Handles large, heterogeneous datasets and automates pattern detection across organoid and clinical resources.

Drug response prediction: Uses organoid data to predict drug efficacy, identify stronger biomarkers, and guide treatment selection.

Disease modeling: Builds more accurate, patient-specific organoid models and improves prediction of disease progression and treatment response.



Organo-on-a-chip



Organ-on-a-chip: advanced *in vitro* culture system

Definition

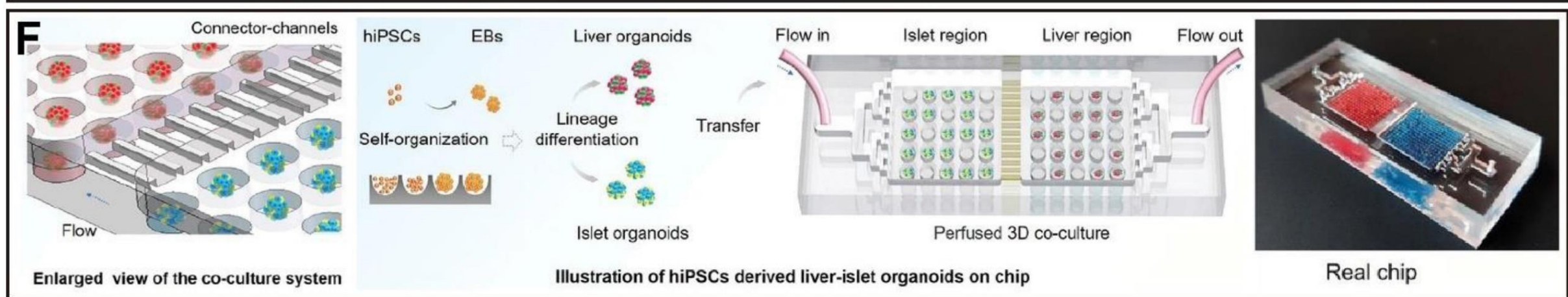
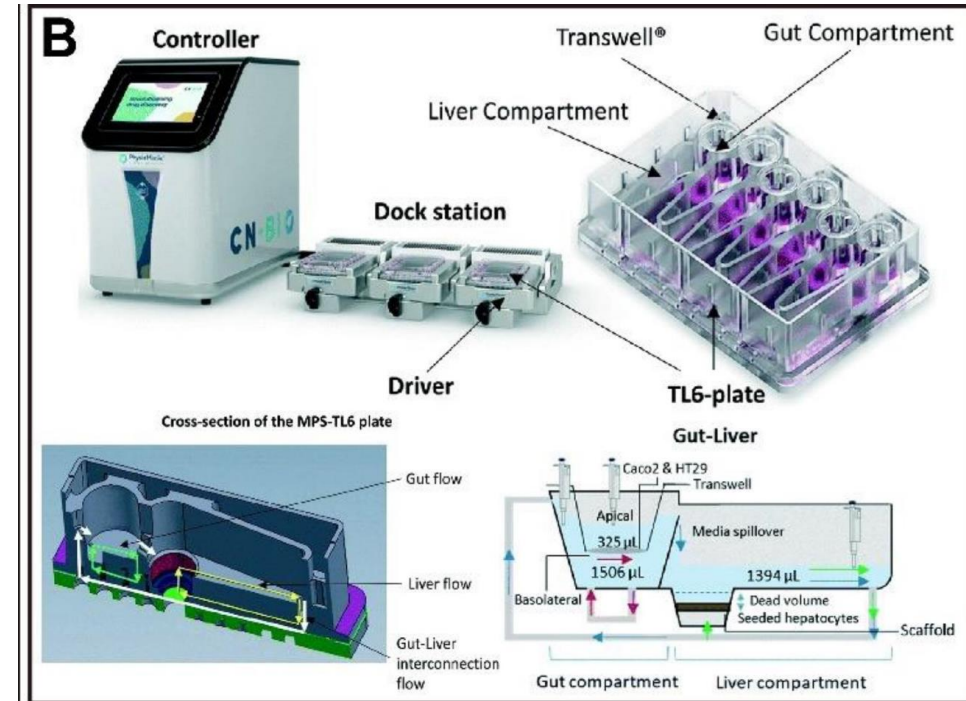
- An microfluidic device used for mimicking the human body environment, representing a simplified but realistic model of its organ level.

Applications

- Drug testing,
- disease modeling,
- personalized medicine, etc

Cell sources

- Primary cells or cell lines
- organoid
- iPSC





Organ-on-a-chip (OoC)

- 1. The challenges faced by OoC technology**
- 2. How AI can help**

Organ-on-a-chip (OoC)

1. Challenges

- **Cell lifespan:** Primary cells and cell lines have limited longevity, reducing long-term model performance.
- **Complexity & scaling:** Adding cell diversity, metabolites, vascularization, and multi-organ scaling remains difficult.
- **Low throughput:** Most OoCs are not high-throughput; automation is needed for drug-discovery workflows.
- **Data demands:** Large datasets require better systems for data handling and analysis.

Organ-on-a-chip (OoC)

2. How AI can help?

- **Data analysis:** ML and deep learning extract patterns from large OoC datasets.
- **Design & optimization:** AI helps select materials and refine device structures for specific applications.
- **Automation & control:** Supports automated workflows, microfluidic control, and stable culture conditions.
- **Predictive modeling:** Builds models of drug response and disease progression to guide evaluation.
- **Robotics integration:** AI-enabled robotics can automate culture maintenance, fluid handling, and sampling.

Organ-on-a-chip (AI applications)

AI-assisted image analysis

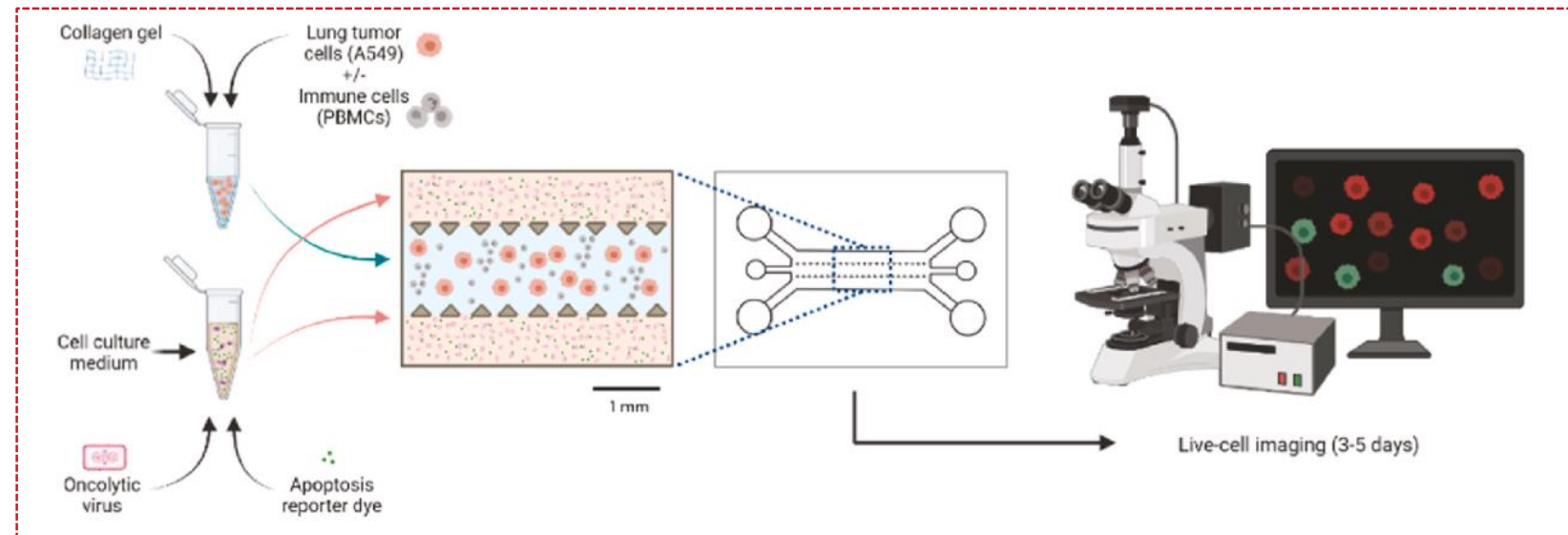
- Quality control: underseeding, channel blockage, poor barrier formation
 - George RM, Kenry. *Chem & Bio Engineering*. 2022. doi: 10.1021/cbe.5c00087
- Non-label viability estimation: drug screening
 - Chiang CC, ... Chen YC. *Lab Chip*. 2024. PMID: 38804084

AI-assisted multimodal data analysis

- Predict biological function of microvascular organ-on-a-chip
 - Based on vascularization morphology and oxygen transport potential
 - Tronolone JJ, ... Jain A. *Ann Biomed Eng*. 2023. PMID: 36913087

AI-assisted cell tracking

- Cell migration and interaction tracking — **CellHunter**
 - Mencattini A, ... Martinelli E. *Biosens Bioelectron*. 2022. PMID: 35932554



Interim



The Value of Imperfection

*“Nothing lasts, nothing is finished,
and nothing is perfect.”*

— **wabi-sabi** (侘び寂び)

The idea that there's real beauty in things
that aren't perfect or complete.



”kintsugi”

Instead of hiding the crack, you highlight it, and it actually makes it more special. So the piece becomes even more beautiful because of its imperfections.

Section III

AI-enhanced image analysis



Cell analysis

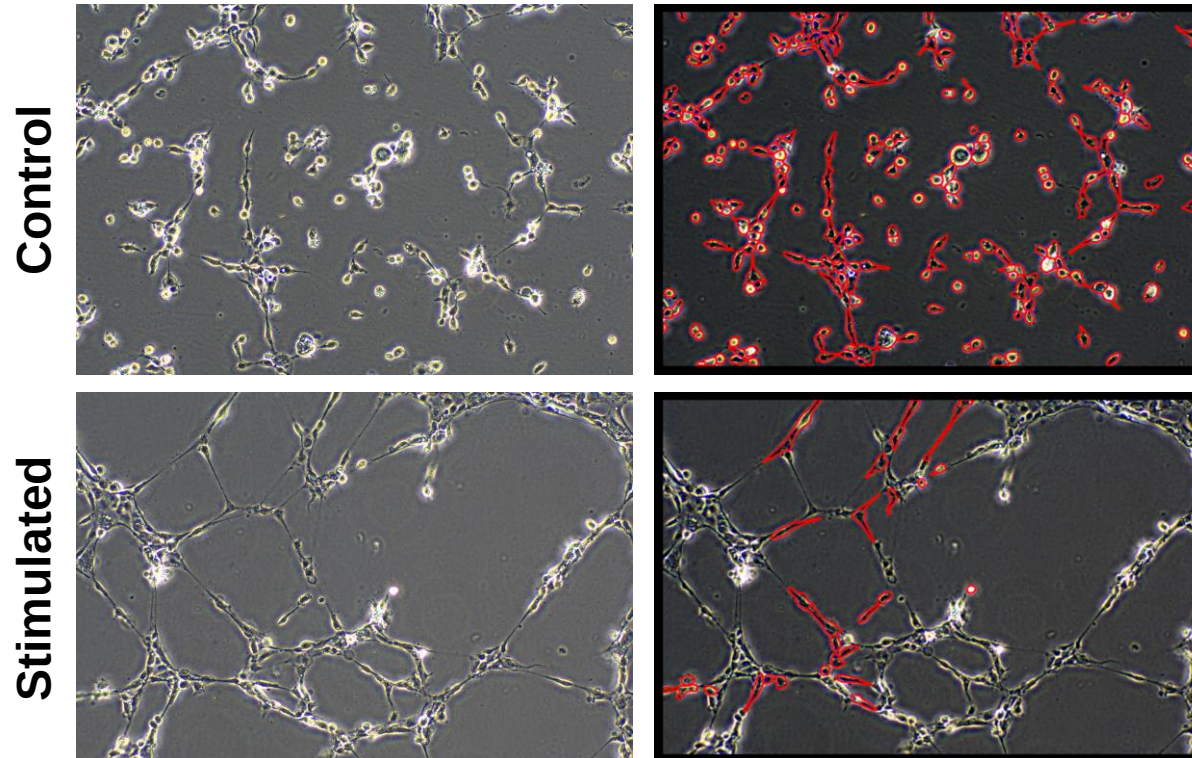


AI-enhanced cell image analysis

—— *Cellpose* as example of cellular segmentation

- Analysis based on free web-service of pre-trained Cellpose (Cellpose-SAM)

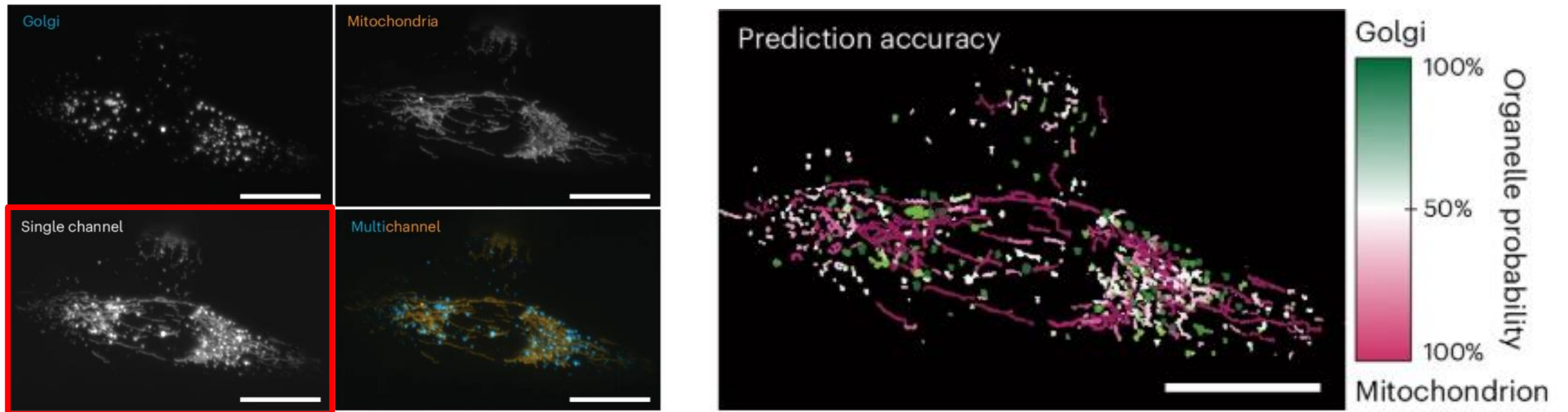
Brightfield of endothelial cell tube formation



AI-enhanced cell imaging analysis

— *Nellie* for organelle segmentation

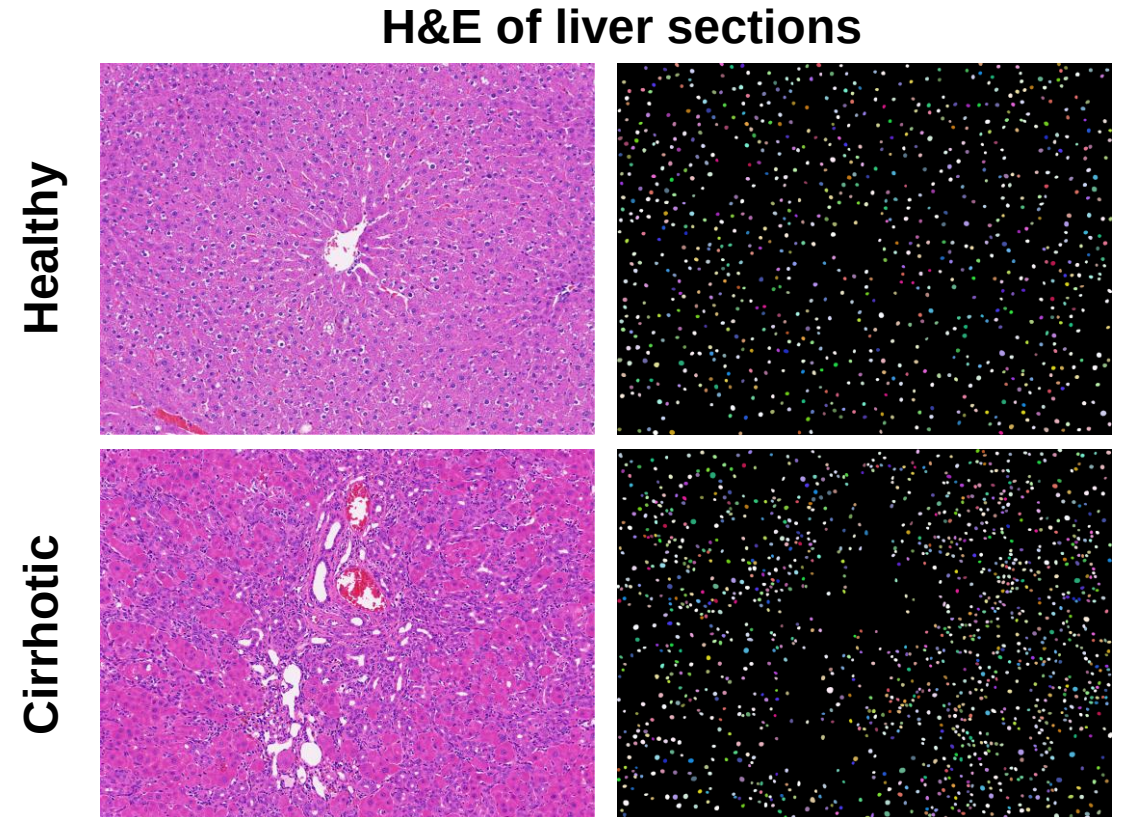
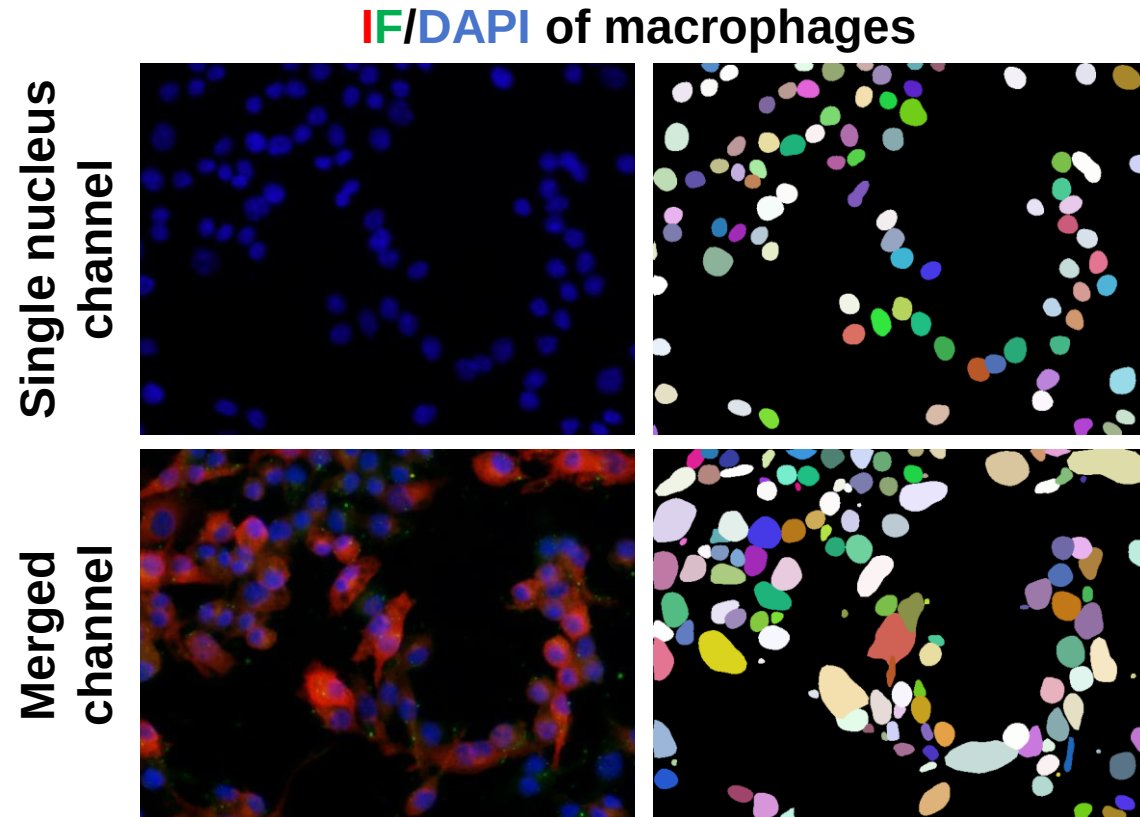
Nellie: Single-channel based separation of Golgi vs. Mitochondria



AI-enhanced liver cell analysis

—— *StarDist* for nucleus segmentation

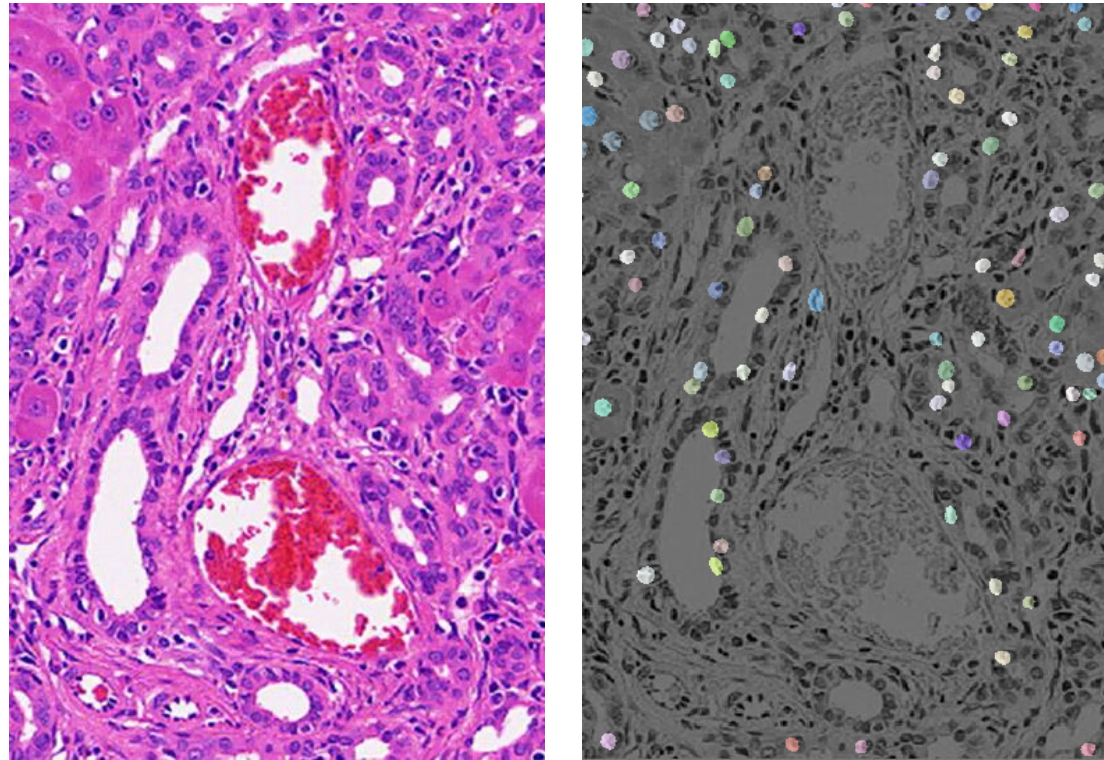
- Based on StarDist Fiji Plugin, pre-trained with segmentation of fluorescent and H&E nuclei



AI-enhanced liver image analysis

—— *StarDist* for nucleus segmentation

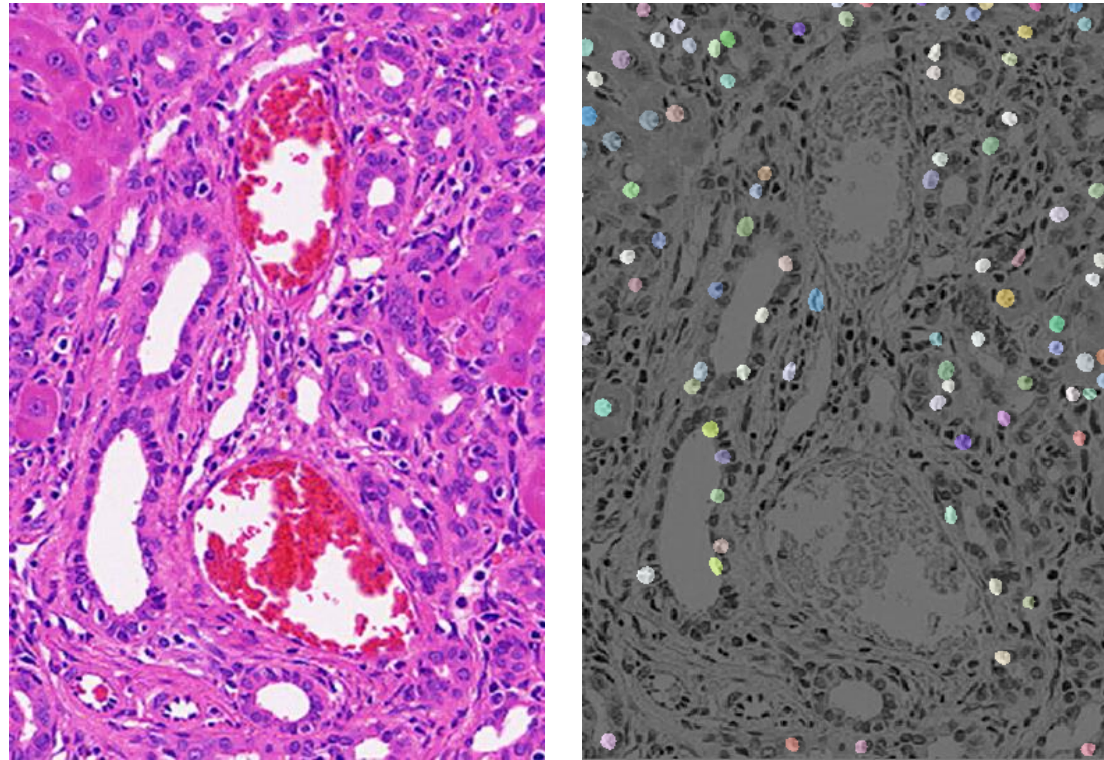
- Based on StarDist Fiji Plugin, pre-trained with segmentation of fluorescent and H&E nuclei



AI-enhanced liver image analysis

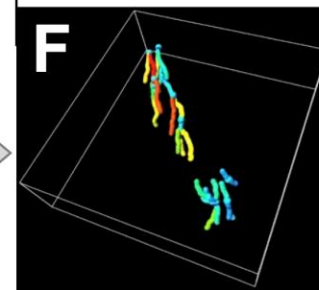
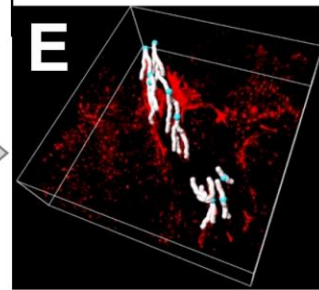
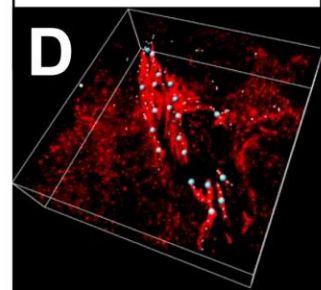
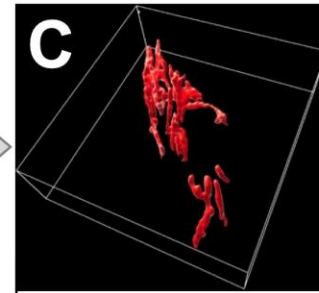
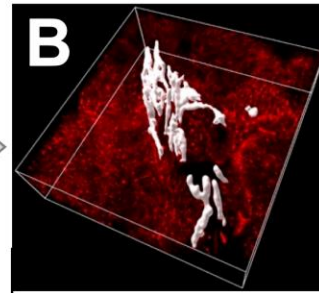
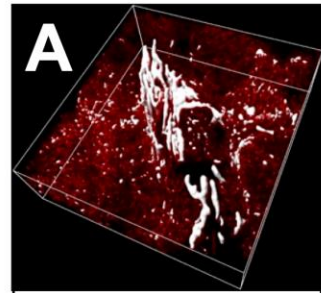
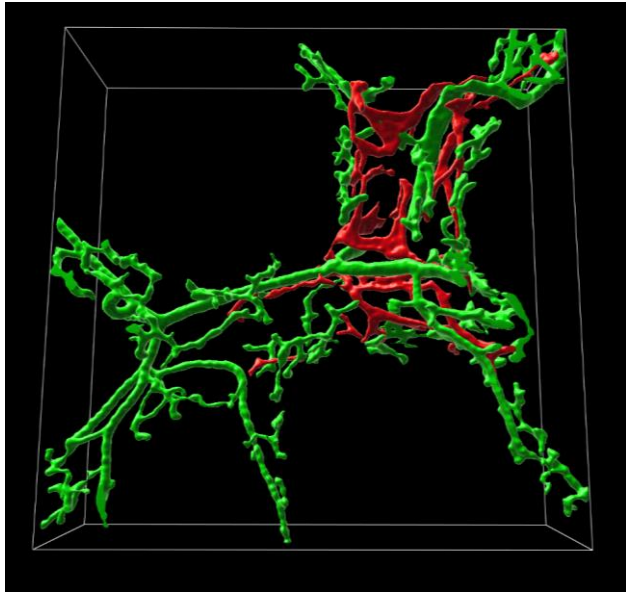
—— *StarDist* for nucleus segmentation

- Based on StarDist Fiji Plugin, pre-trained with segmentation of fluorescent and H&E nuclei



Name	Function	Pros & Cons	Ease to use	Cost
Cellpose	Cell segmentation on 2D/3D cellular images	Pros: Fit various image types; Recognize complex cell shapes; 3D segmentation with 2D training Cons: Difficulty in highly concave shapes; Difficulty in heterogeneous size distributions	Beginnnger to Intermediate	Free
StarDist	Nucleus segmentation on 2D/3D images; Cellular & Histopathology	Pros: Official plugins available; Separate touching/crowded nuclei Cons: Difficulty in non-star-convex shapes; Difficulty in membrane/cytoplasm-only stain	Beginner	Free
FAMOUS	Organelle ultrastructure segmentation	Pros: Work on single consumer-grade GPU; Across SEM and TEM images Cons: Difficulty in complex connected structures	Beginner	Not in public
Nelie	Organelle ultrastructure segmentation; Extract ultrastructure features; 2D/3D live-cell microscopy	Pros: Detailed feature extraction; Easy and visual GUI Cons: Difficulty in large smooth structures like nuclei; High hardware and computational requirments	Beginnnger to Intermediate	Free
Trackmate	2D/3D live-cell imaging tracking	Pros: GUI and modular plugin; Cons: Limited built-in quantitative analysis; Difficulty in high-throughput tracking	Beginner	Free
DeepCell tracking	Time-lapse 2D cell imaging tracking Distinguish different cell types	Pros: Fit high-throughput tracking; Semantic segmentation of co-cultured cells; Cons: Need high-quality annotated training data; Need hyperparameter tuning; High hardware and computational requirements	Intermediate to Advanced	Free

Current situation..... Manual quantification

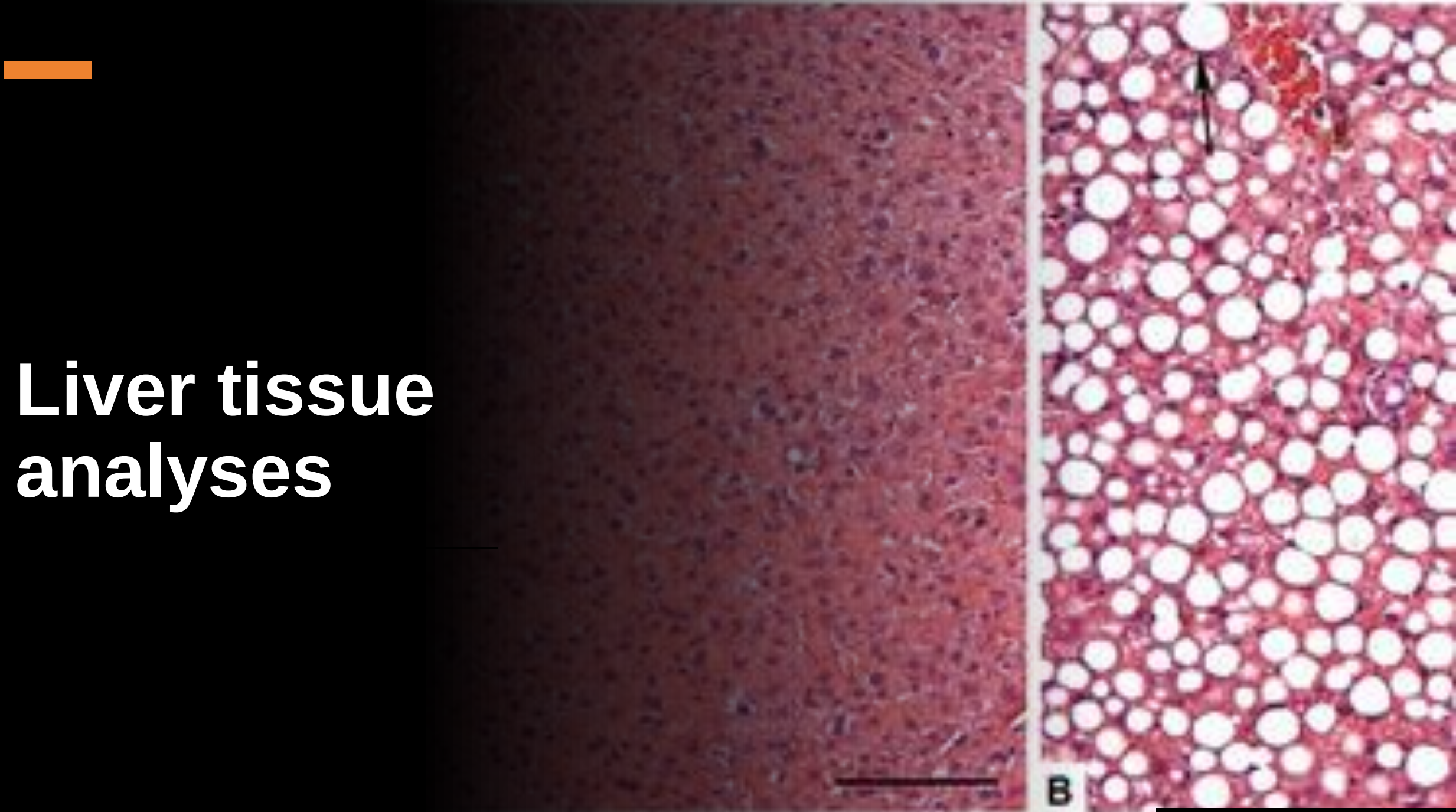


- Total volume

**Data collection
and
quantification**

- Diameter
- Branch number

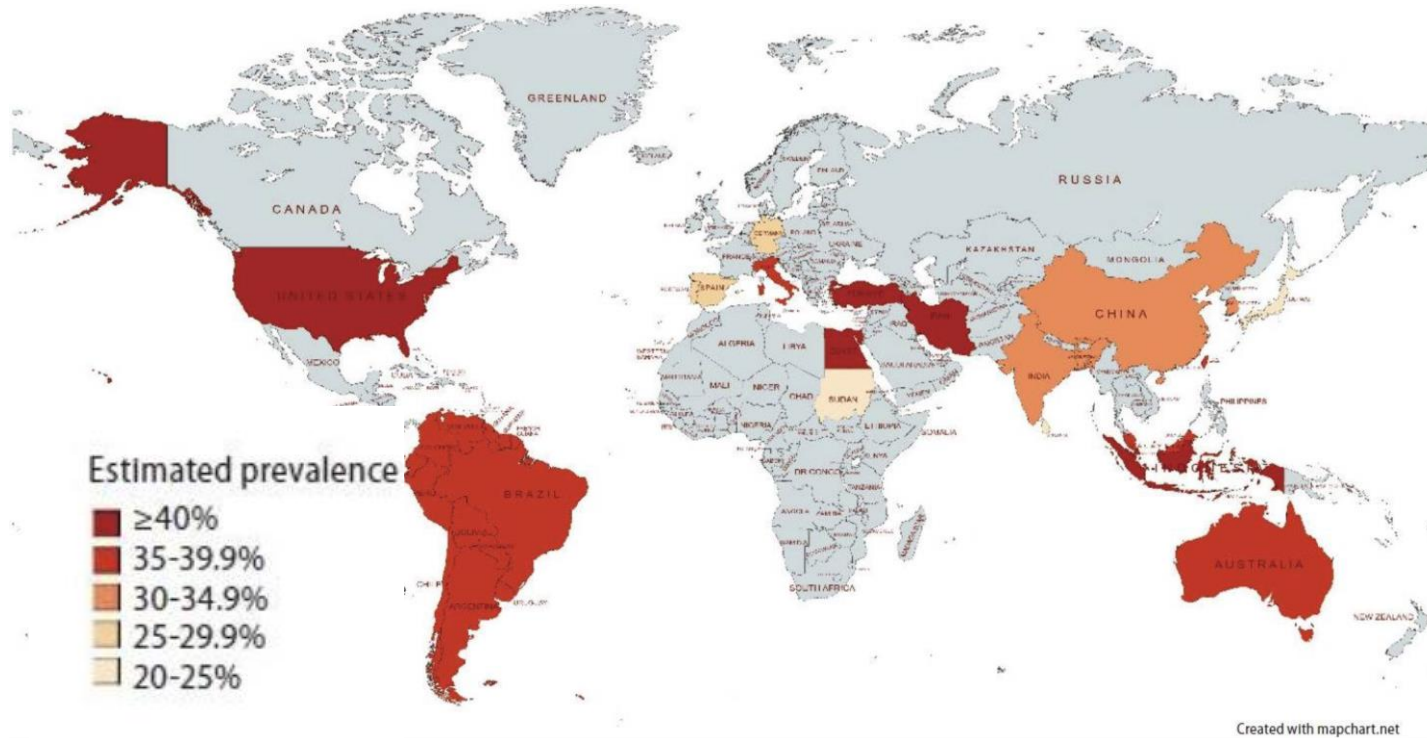
Liver tissue analyses



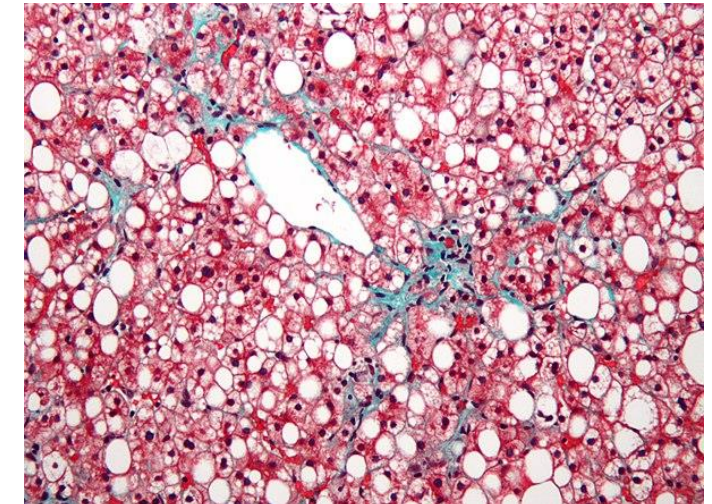
MAFLD (Metabolic dysfunction-associated fatty liver disease)

- Most common form of chronic liver disease globally

MASH: Metabolic dysfunction-associated steatohepatitis



Complex pathological features



[This Photo](#) by Unknown Author is licensed under [CC BY-SA](#)

AI-enhanced image analysis in liver diseases (#1)

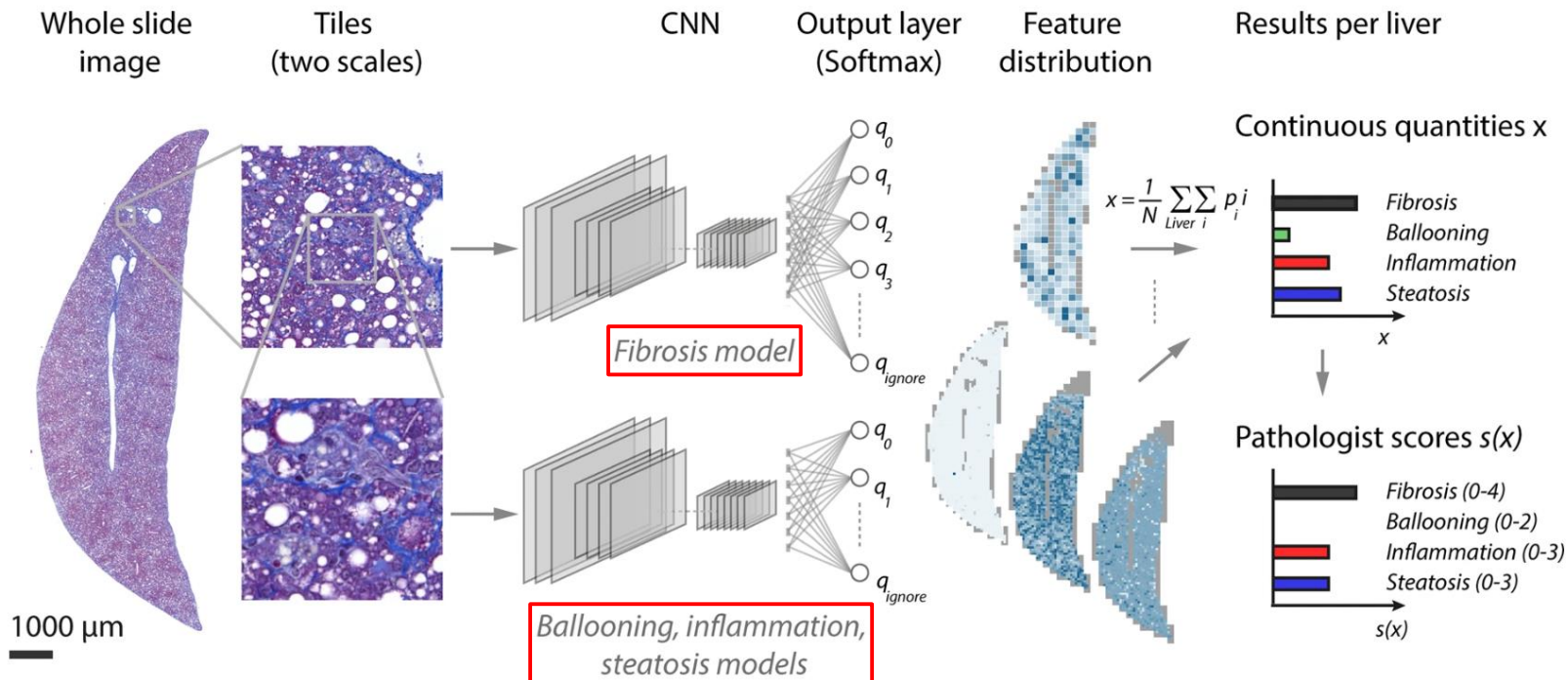
“Deep learning enables pathologist-like scoring of NASH models”

by Heinemann F, ... Stierstorfer B. *Sci Rep.* 2019. PMID: 31804575

Automated Kleiner scoring of **ballooning, steatosis, inflammation, and fibrosis**

Methods

- Deep learning model with whole-slide images of **Masson** staining from MASH mice and rats
- Fibrosis model and model for other three features trained separately

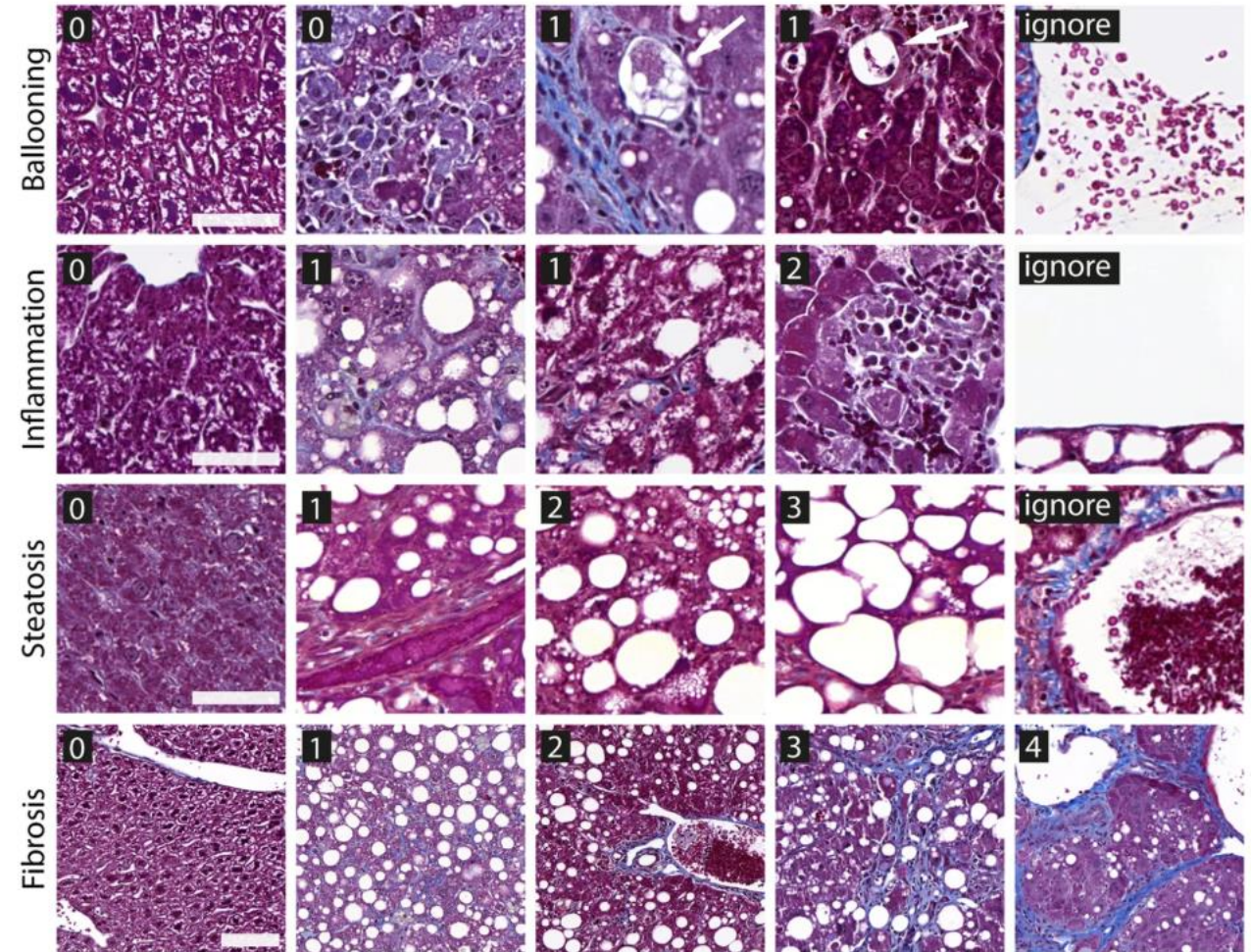


Kleiner score: histological scoring system for MAFLD

Scoring criteria

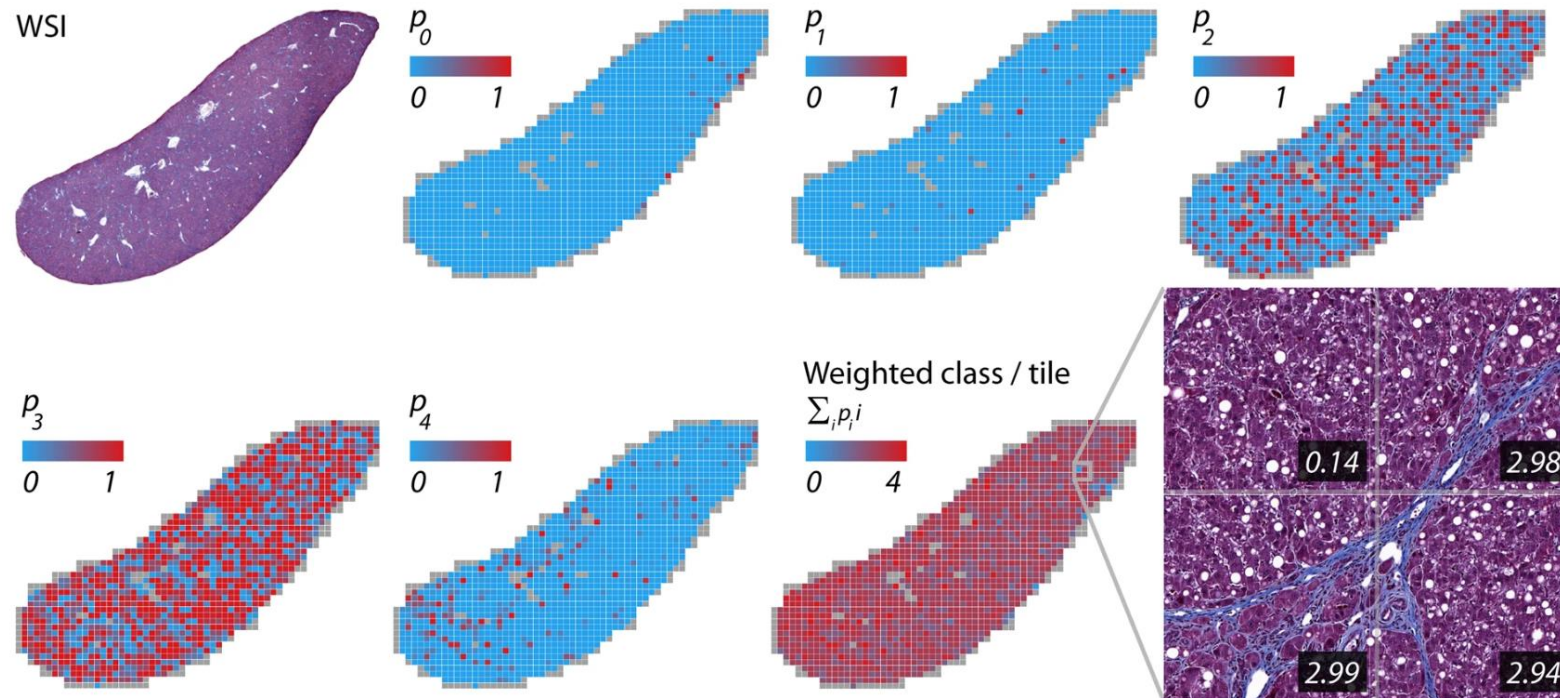
NASH activity grade: grade = total score: S + L + B (range 0–8)					
Steatosis	S score	Lobular inflammation	L score	Hepatocyte ballooning	B score
< 5%	0	None	0	None	0
5–33%	1	< 2	1	Few ballooned cells	1
34–66%	2	2–4	2	Many ballooned cells	2
> 66%	3	> 4	3		
NASH fibrosis stage			Stage		
None			0		
Mild, zone 3 perisinusoidal fibrosis			1a		
Moderate, zone 3 perisinusoidal fibrosis			1b		
Portal/periportal fibrosis only			1c		
Zone 3 perisinusoidal and portal/periportal fibrosis			2		
Bridging fibrosis			3		
Cirrhosis			4		

Example tiles



AI-enhanced image analysis in liver diseases (#1)

Example of MASH: “Deep learning enables pathologist-like scoring of NASH models”



Results: using trained **Fibrosis model** as an example

- Tiles were analyzed and predicted possibility at every fibrosis score (0-4)
- The majority of tiles are classified as fibrosis score 3

AI-enhanced image analysis in liver diseases (#2)

Example of MASH-Fibrosis: “Deep learning-based method for grading histopathological liver fibrosis in rodent models of metabolic dysfunction-associated steatohepatitis”

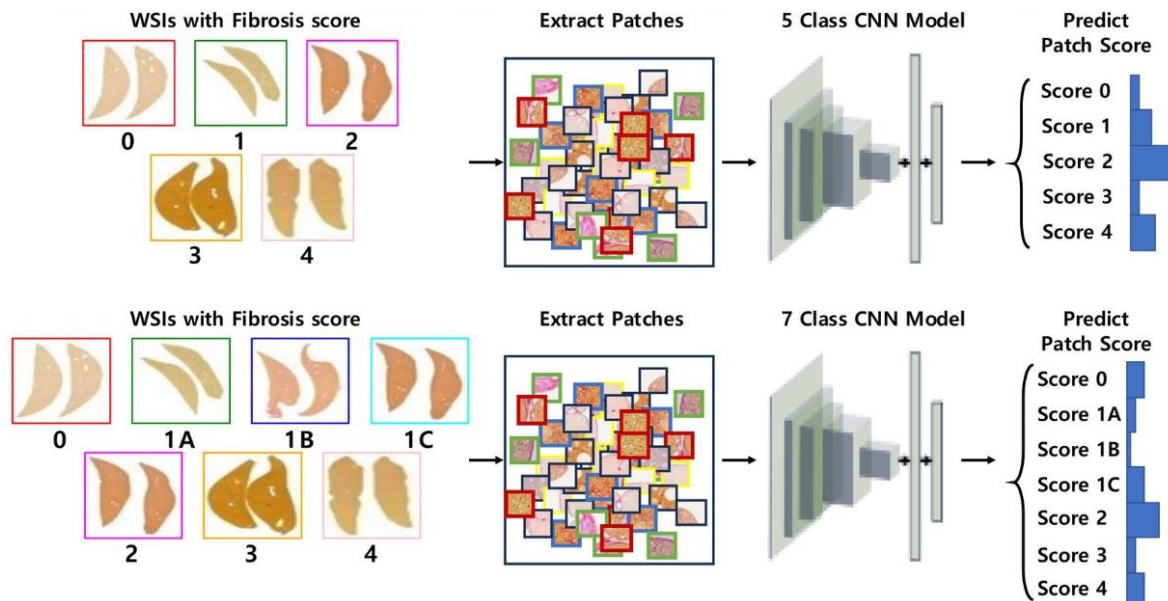
Purpose

- Automated and standardized Kleiner scoring of **fibrosis**, especially distinguishing Kleiner 1A/B/C

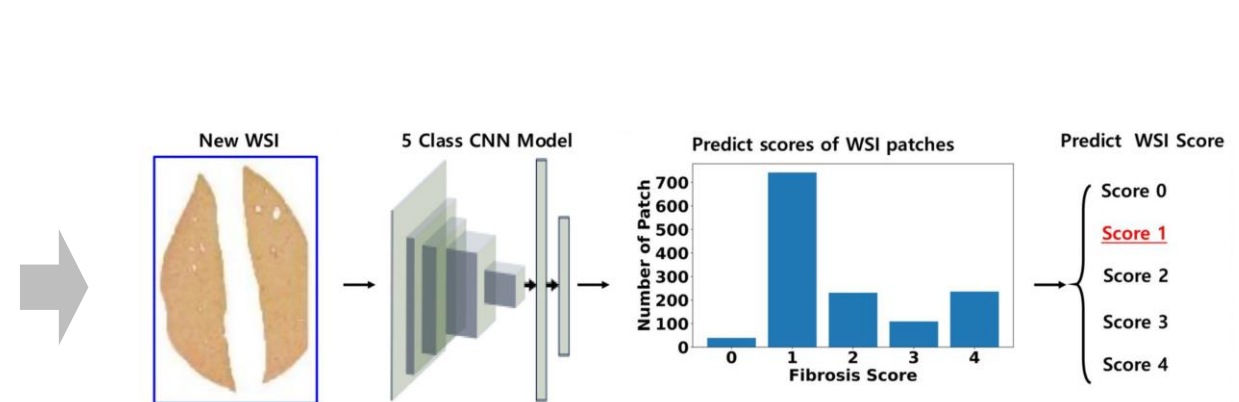
Methods

- Deep learning model with whole-slide images of **Sirius red** staining from MASH mice and rats

Training

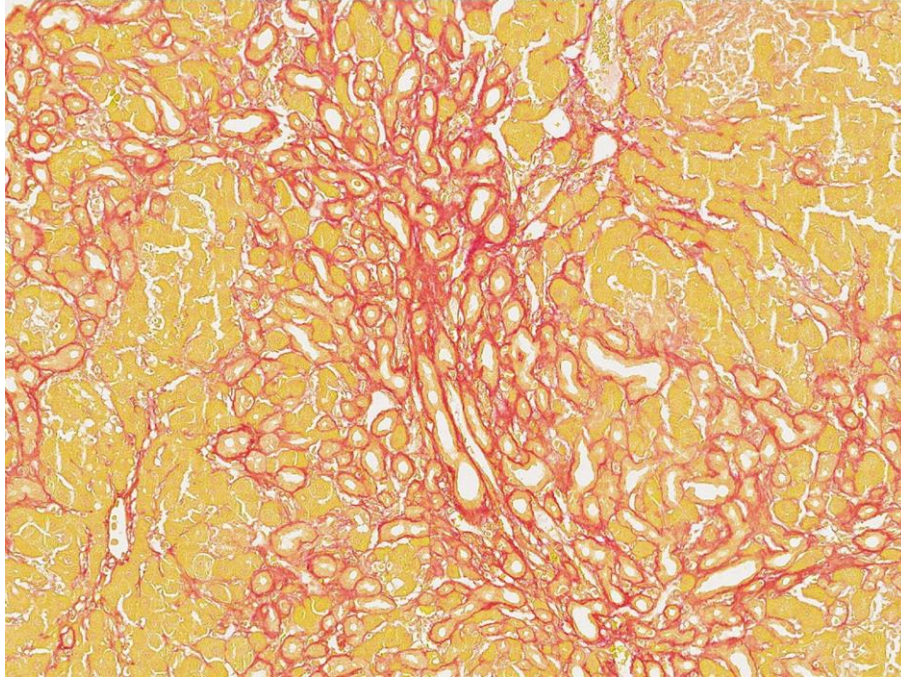


Application



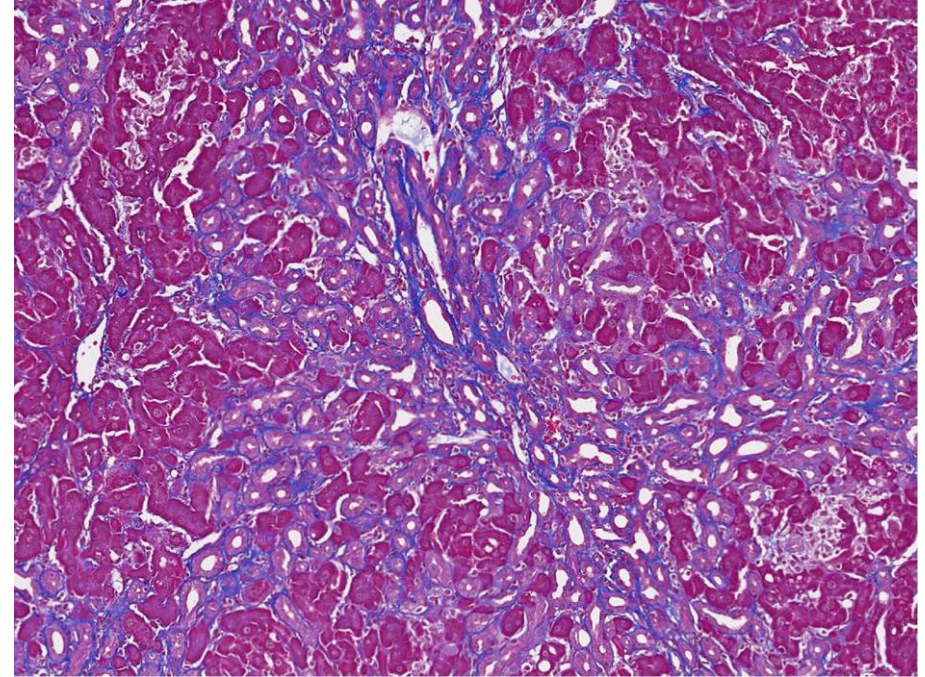
Histological staining for fibrosis

Sirius red



- Background marked **yellow-brown**
- Collagen marked **red**

Masson's trichrome

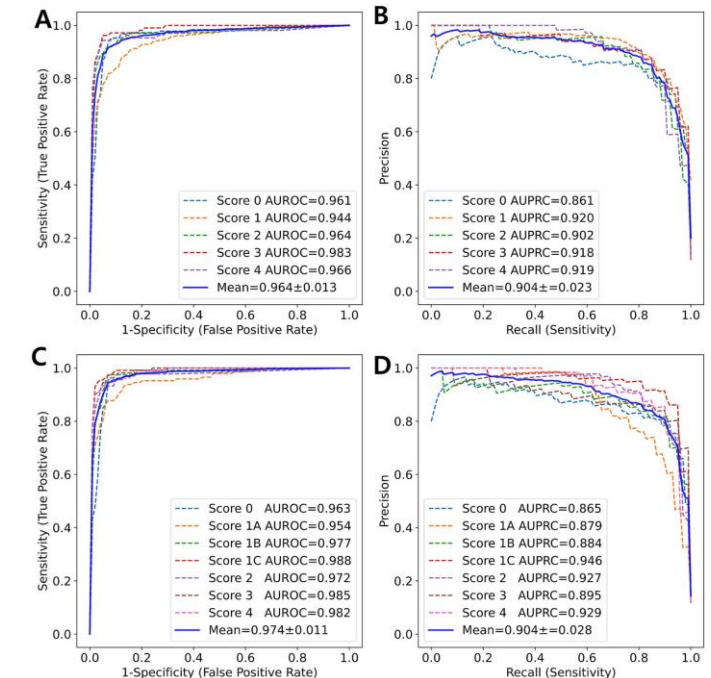
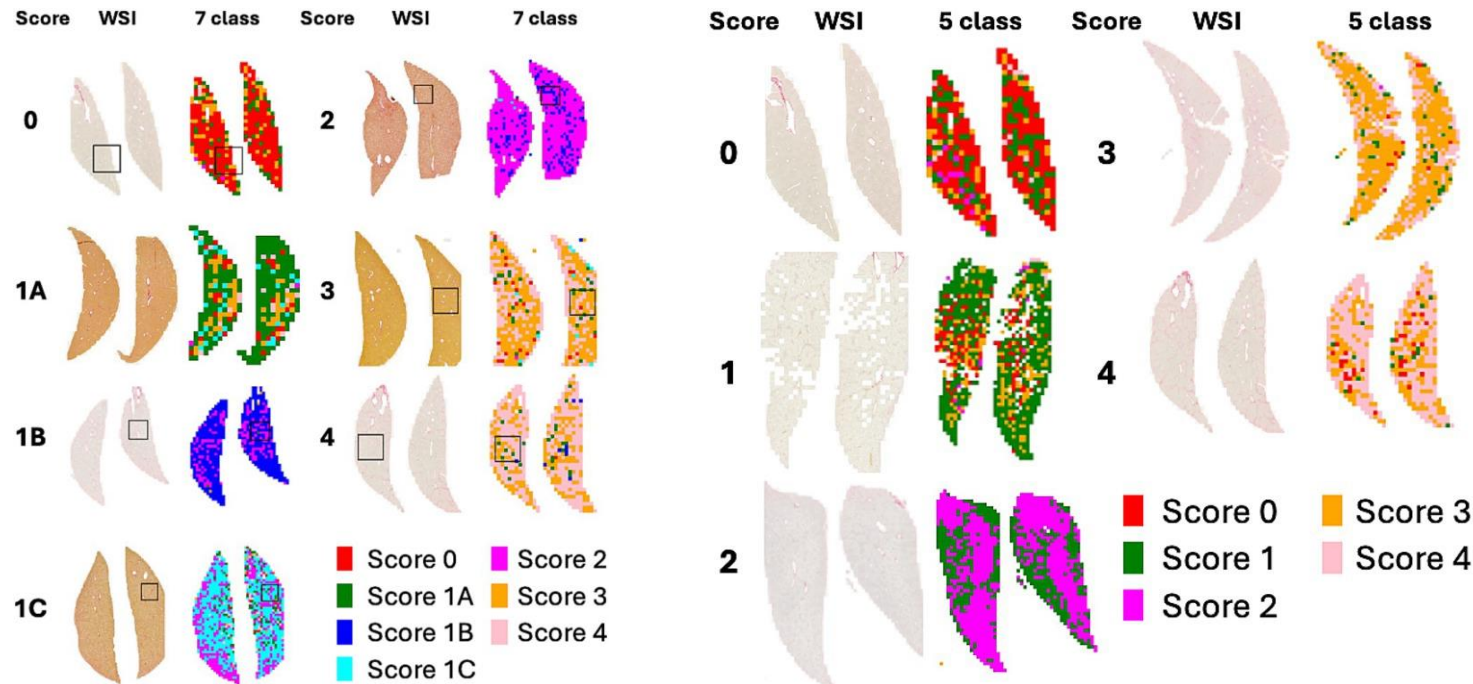


- Background marked **pink-purple**
- Collagen marked **blue**

Sirius red and Masson's trichrome staining of liver sections from the same BDL cirrhotic rat

AI-enhanced image analysis in liver diseases (#2)

Example of MASH-Fibrosis:



Results

- Automated staging based on patches from WSIs: “Spatial Heatmap” of Kleiner’s score
- Both 5-class and 7-class showed good sensitivity and specificity

AI-enhanced image analysis in liver diseases (#3)

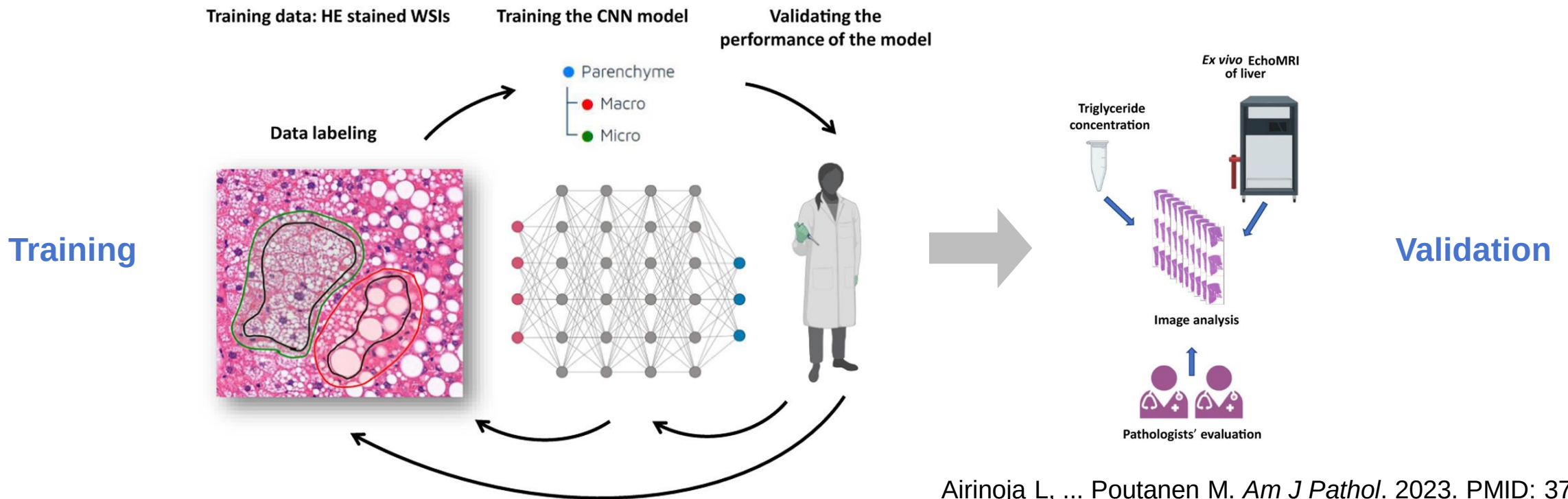
Example of MASH-Steatosis: “Deep Learning-Based Image Analysis of Liver Steatosis in Mouse Models”

Purpose

- Automated analysis of **steatosis**, especially quantifying microvesicular and macrovesicular steatosis

Methods

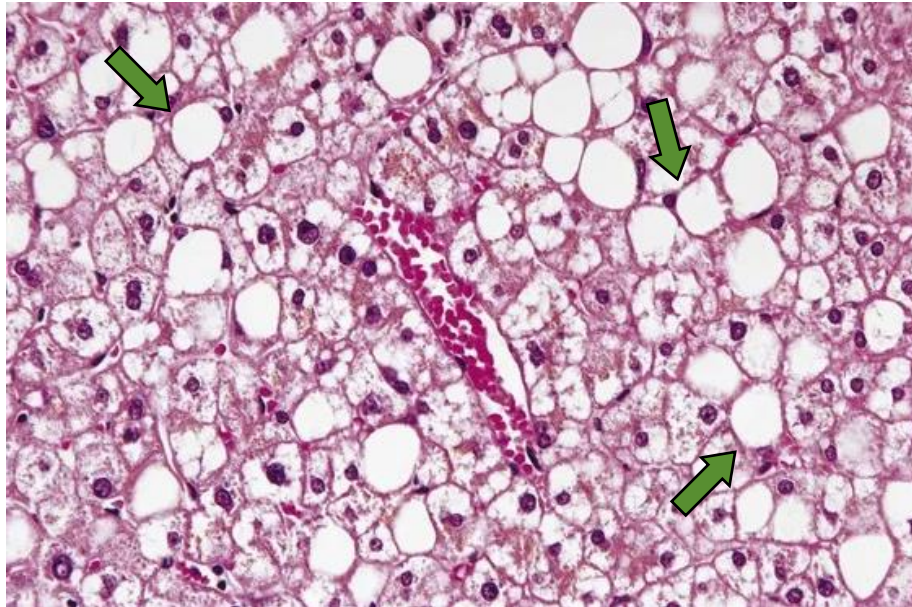
- Deep learning model with whole-slide images of **H&E** staining from MAFLD mice



Steatosis: macrovesicular and microvesicular

Macrovesicular

Metabolic syndrome and alcohol exposure

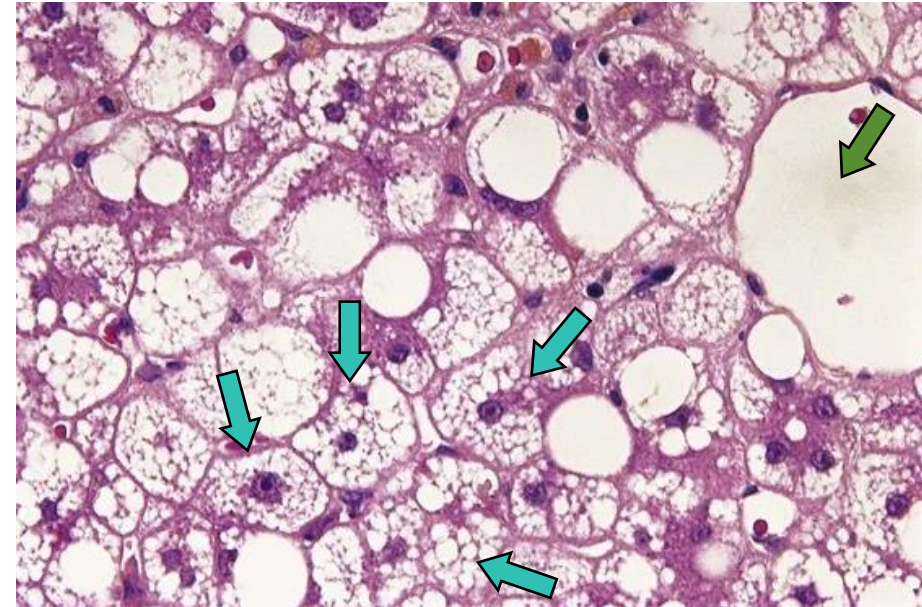


- **Large vacuoles** in hepatocytes
- Displace **nuclei to the edge** of cell

➡ Macrovesicular
➡ Microvesicular

Microvesicular

Mitochondrial damage in serious liver injury



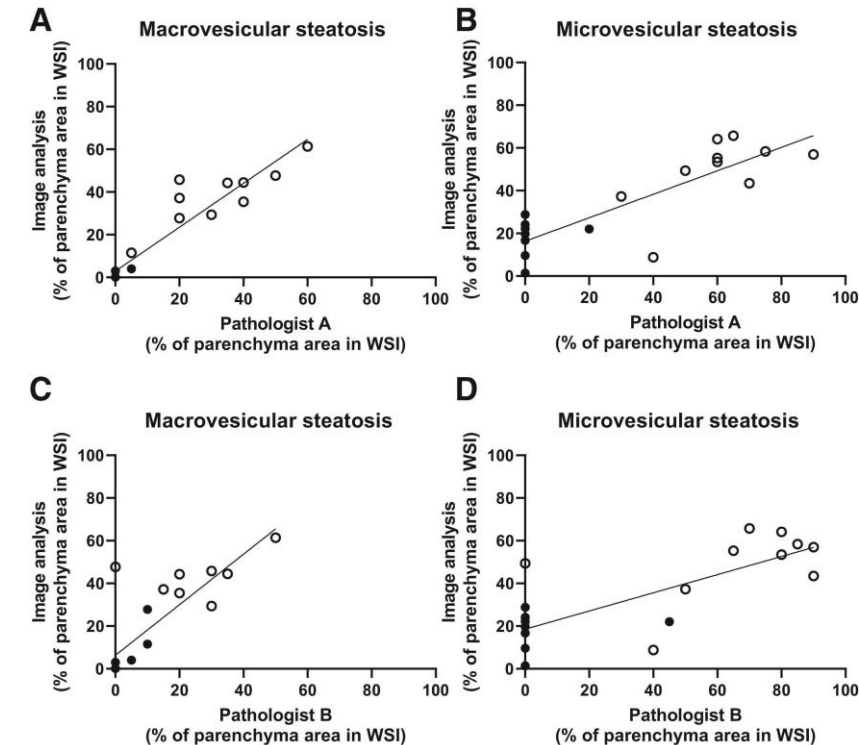
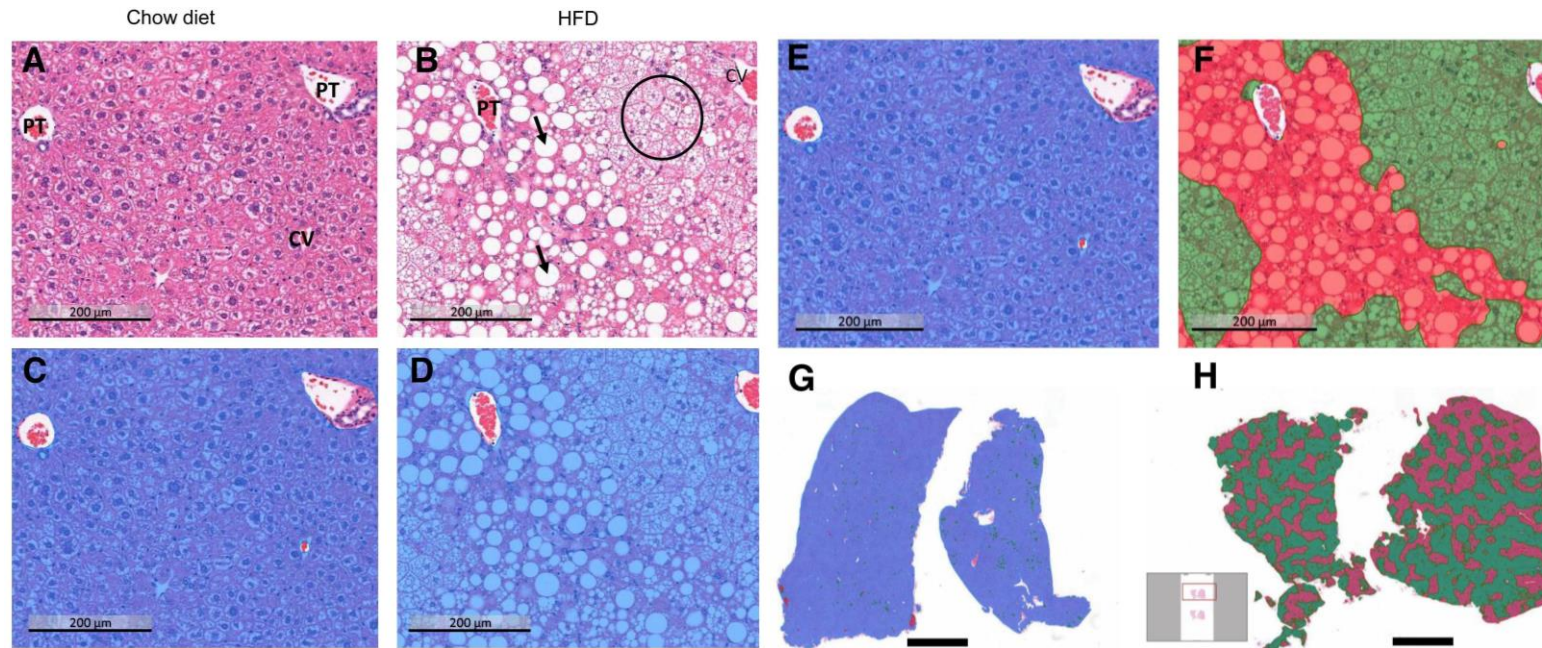
- Numerous finely divided **small vacuoles**
- **Nuclei maintained central** position

AI-enhanced image analysis in liver diseases (#4)

Example of MASH-Steatosis: “Deep Learning-Based Image Analysis of Liver Steatosis in Mouse Models”

Application

Efficiency



Results

- Trained model able to recognize parenchyma, microvesicular, and macrovesicular steatosis
- Good correlations with pathological evaluation, plasma lipid profile, and Echo-MRI

Section IV

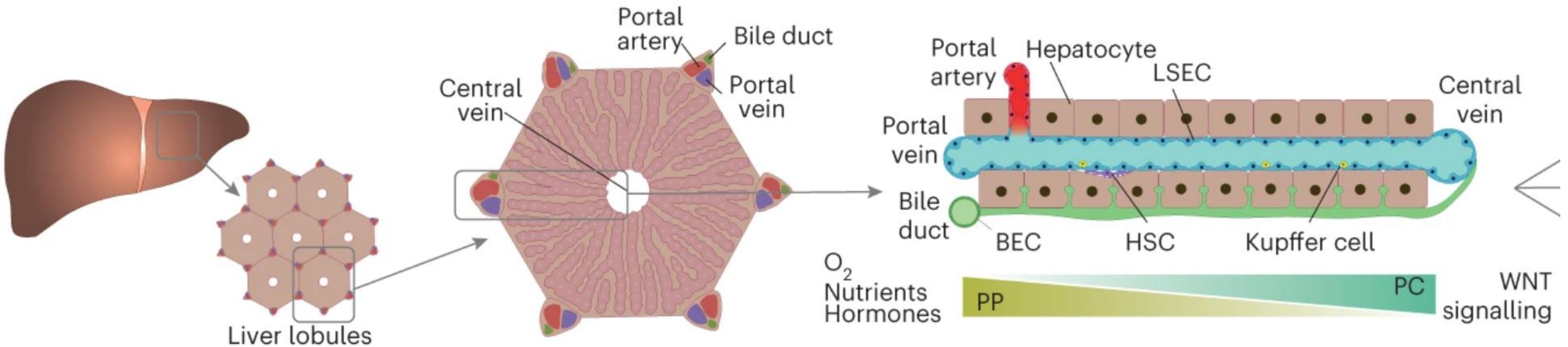
AI-driven integration of diverse data

AI-driven integration of diverse data

—— *DeepLiver*, a multi-omic data analysis

Purpose:

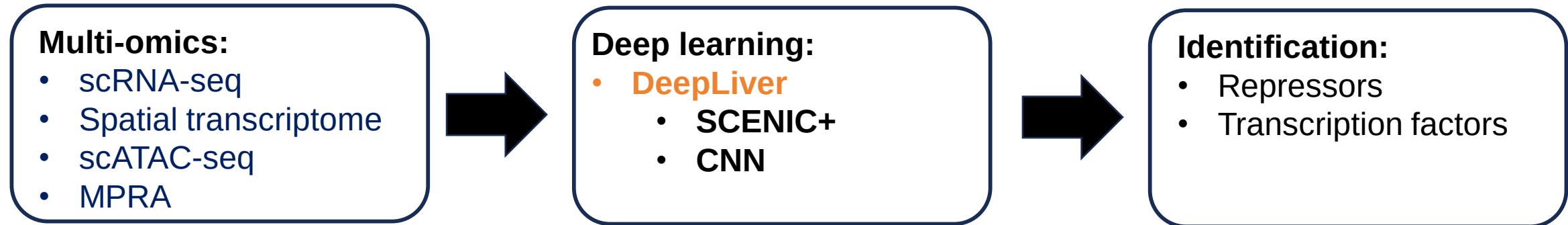
To determine what regulates liver zonation



AI-driven integration of diverse data

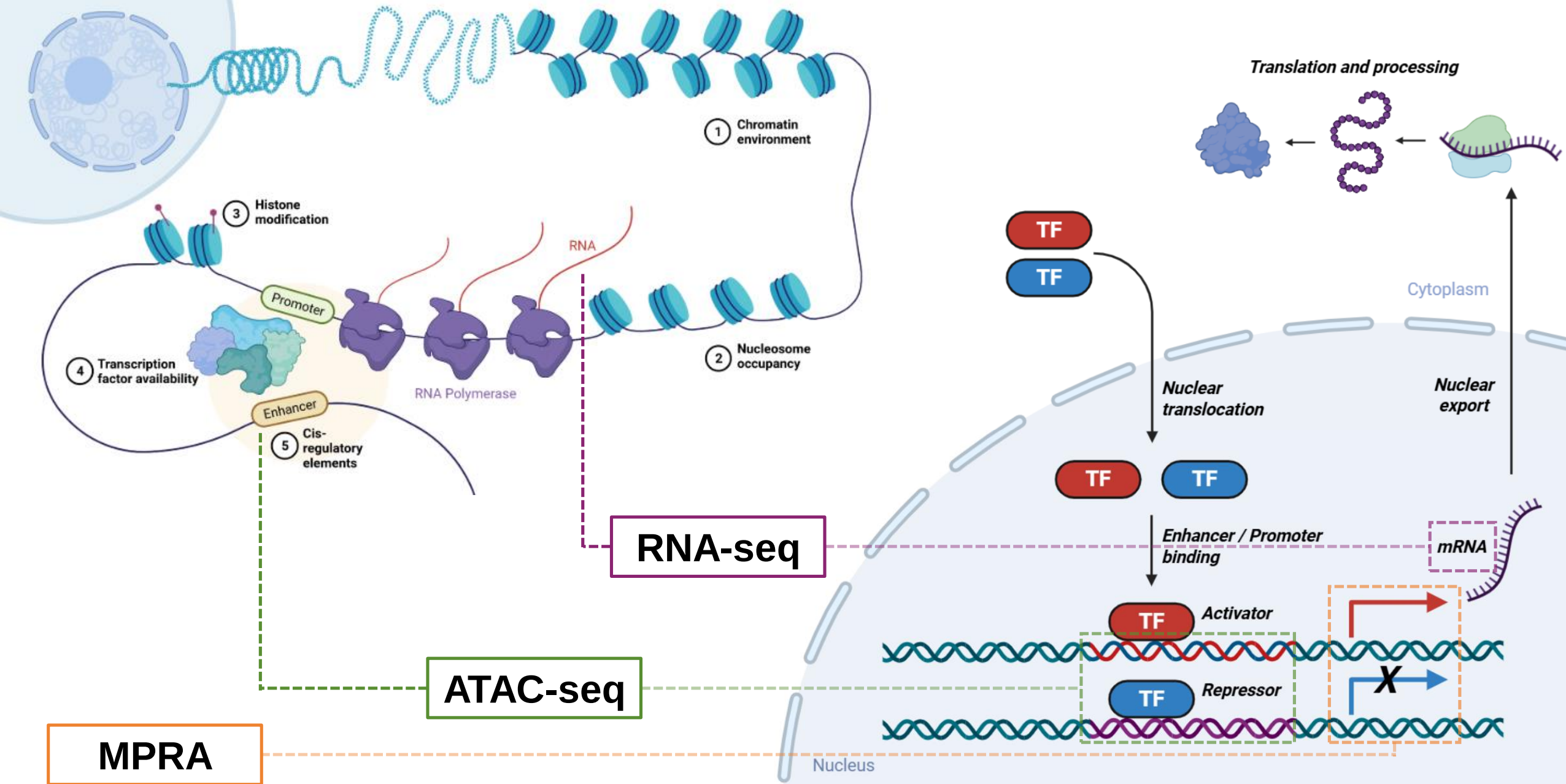
—— *DeepLiver*, a multi-omic data analysis

Methods



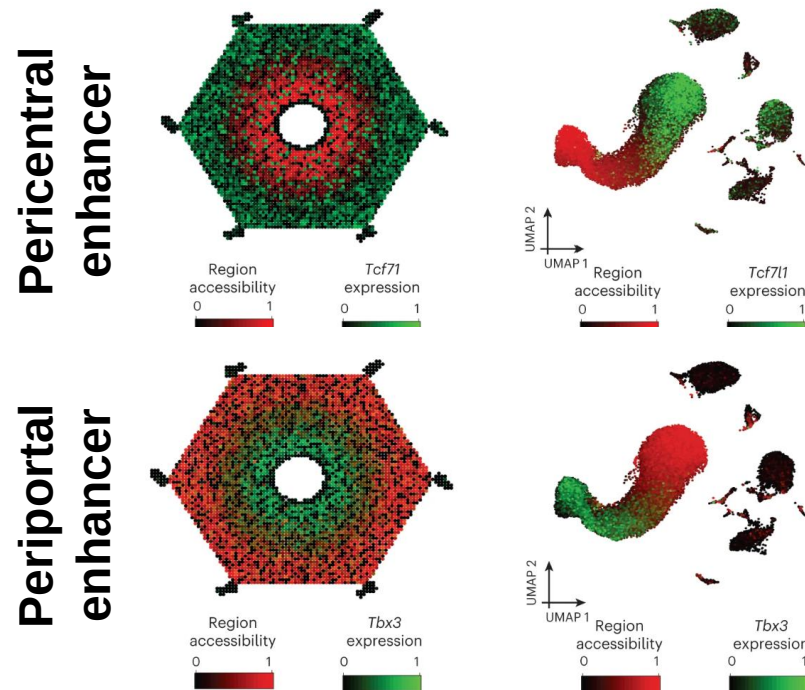
Massively parallel reporter assay (**MPRA**): regulatory target (enhancer) activity

Regulation of gene expression in eukaryocytes

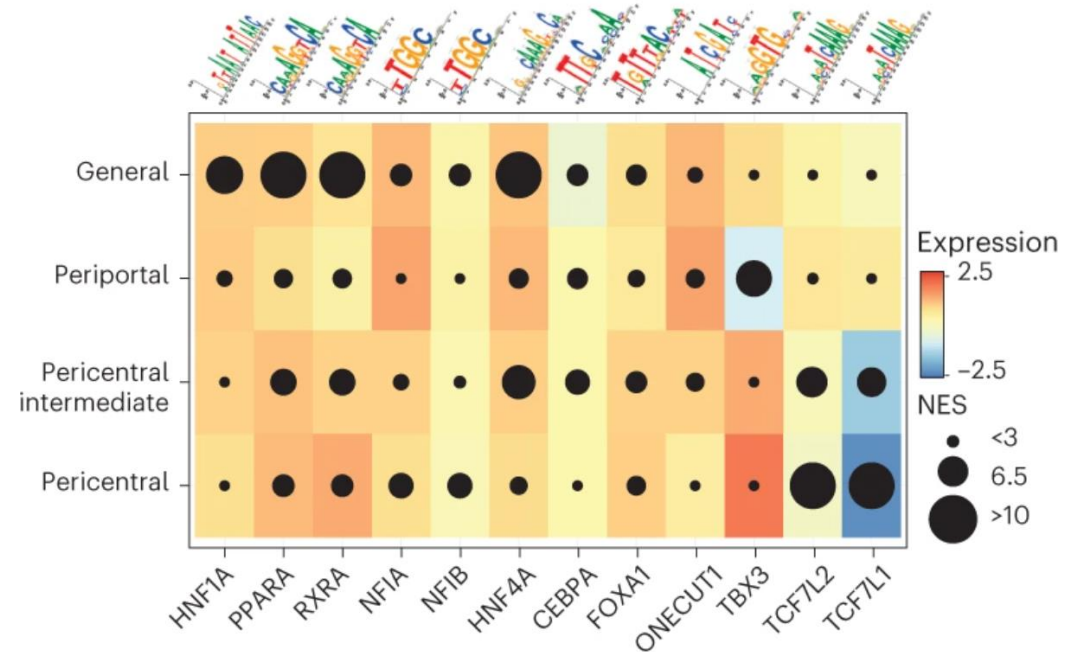


AI-driven integration of diverse data

— *DeepLiver* as example of multi-omic data analysis



Motif accessibility of TFs from periportal to pericentral



Results: Spatially heterogeneous TF-enhancer regulatory network

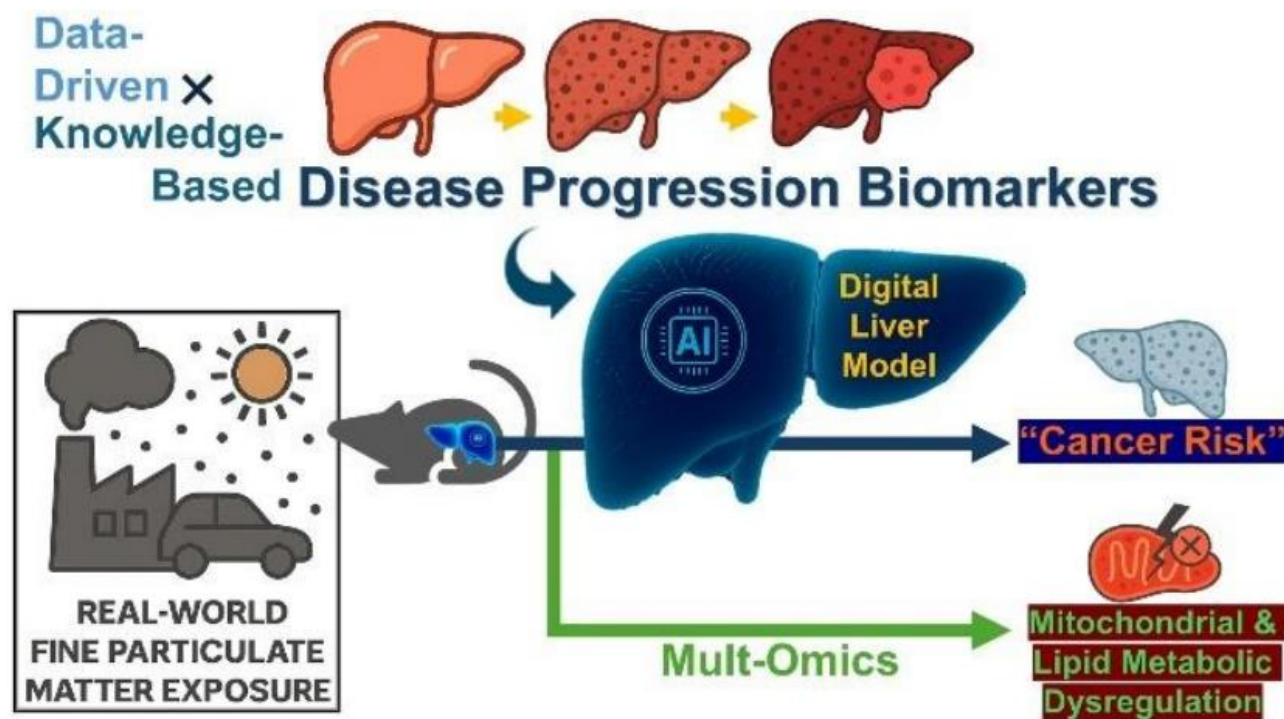
- Tcf71 and Tbx3 as key transcriptional repressors: antagonize each other and create spatial domains
 - Manipulating them switches hepatocyte identity
- Identified enhancer motifs controlling zonation: driven by TFs including Tcf71/Tbx3

Digital Liver model - for evaluating the risk of long-term PM_{2.5} exposure in liver disease progression

“Two-Hit Hypothesis: Digital Liver Modeling and Multi-Omics Reveal Mitochondrial and Lipid Metabolic Dysregulation in PM_{2.5}-Induced Liver Fibrosis” by Hiu-Lok Ngan, ... Zongwei Cai. ChemRxiv. 2025.

Purpose

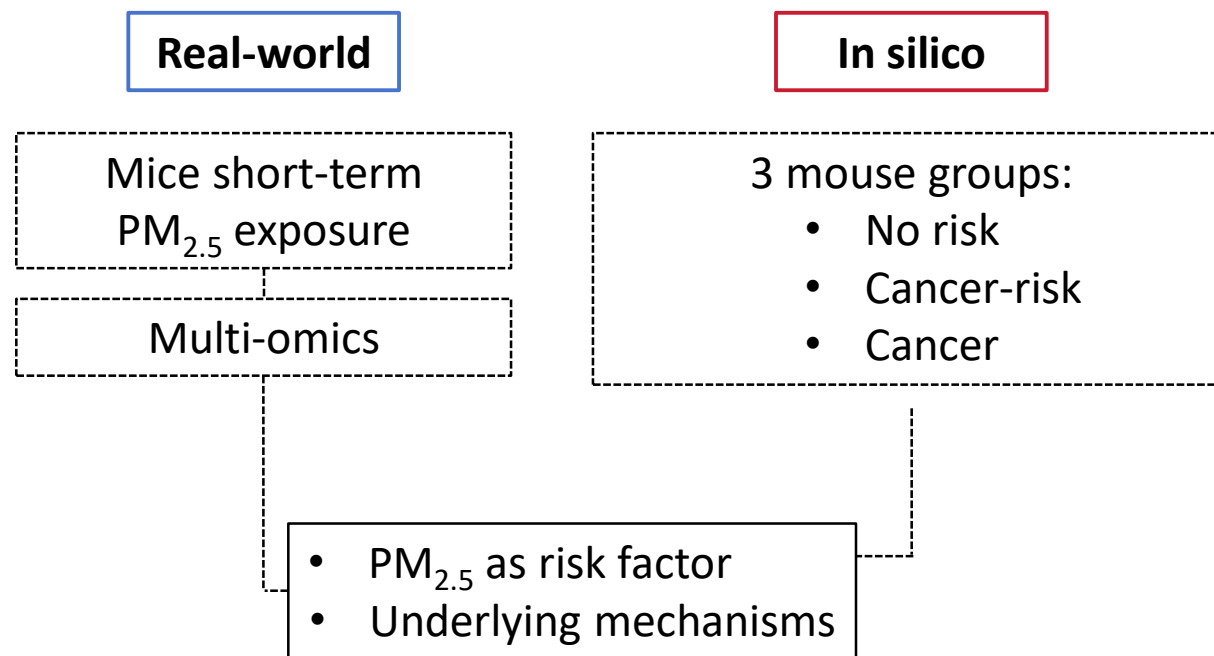
- Evaluate environmental pollute as cancer risk by developing digital rodent liver model



Digital Liver model - for evaluating the risk of long-term PM_{2.5} exposure in liver disease progression

“Two-Hit Hypothesis: Digital Liver Modeling and Multi-Omics Reveal Mitochondrial and Lipid Metabolic Dysregulation in PM_{2.5}-Induced Liver Fibrosis” by Hiu-Lok Ngan, ... Zongwei Cai. ChemRxiv. 2025.

Methods



Section V

AI for drug discovery and tests

AI for drug discovery and tests

AI tools

- DILITracer
- DeepTox
- AnimalGAN
- GPD-based

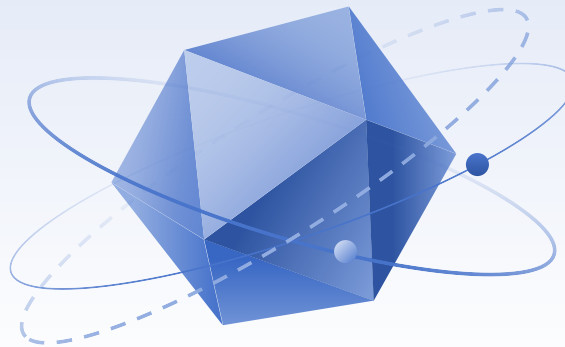
- AtomNet
- GENTRL

Drug discovery

Safety screening

- Human PK profile
- Compound PK parameter

Dosing guidance



AI for drug discovery and tests

Case of Use	AI tools	Function	Ease of Use	Cost / Access
Drug discovery	AtomNet	Predict binding affinity to target by chemical structure; Allow prediction for orphan receptors	Intermediate to Advanced	Not in pulic
Drug discovery	GENTRL	De novo small molecule design: DDR1 kinase inhibitor as example	Intermediate to Advanced	Free
Saftety screening	DILITracer	Stage DILI by brightfield images of liver organoid; Non-invasive - allow continuous surveilliance	Intermediate to Advanced	Free
Saftety screening	DeepTox	High-throughput screening of chemical compounds and toxicity prediction	Intermediate to Advanced	Not in pulic
Saftety screening	AnimalGAN	Predict liver and renal toxicity parameters based on exposure time and dose; Suitable for new compound without in vivo data	Intermediate to Advanced	Free
Saftety screening	GPD-based	Genotype-phenotype difference (GPD)-based; Evaluate difference of importance of gene targets among species; Predict drug toxicity in human based on the gene target and related GPD	Intermediate to Advanced	Free
Dosing guidance	PK prediction	Predict physicochemical and PK parameters from chemical structures; Use the parameters to predict plasma concentration-time profiles in human	Intermediate to Advanced	Not in pulic

Section VI

Key Takeaways

Key Takeaways



Human-centered, AI-enabled experimental approaches are growing, but still early in development



AI helps interpret images, interpret complex data, and link information that is usually scattered



Integration across datasets remains the core challenge, and AI can support that effort



Progress depends on collaboration among basic scientists, computational and engineering teams, bioinformatic groups, and clinicians