12PM

5PM in London (GMT), 2AM in Tokyo (GMT+9)

Welcome and Panel: Multiscale Human: Definitions, Maps, Models

Moderator: Katy Börner, *Indiana University*

Panelists:

- Griffin Weber, *Harvard Medical School (Human Reference Atlas)*
- Karen Miga, *UC Santa Cruz (Pangenome)*
- Clair Walsh, *University College London (Human Organ Atlas)*
- Caterina Strambio, *University of Massachusetts Medical School (4D Nucleome Network)*
- Aviv Regev, *Genentech, Inc. (Human Cell Atlas)*
- Peter Hunter, *Bioengineering Institute New Zealand (SPARC)*
- Maryann Martone, *National Center for Microscopy and Imaging Research (NCMIR)*
- Gary Bader, *University of Toronto, Canada (CIFAR co-director)*

Video: CIFAR MacMillan Multiscale Human

Katy Börner, *Indiana University*
(HuBMAP, SenNet, HRA, CIFAR Co-Director)

Welcome!

To the first hour of 24!

Each hour will introduce a unique topic related to the multiscale mapping of the human body.

We will cover data, maps and models and the role of 2D/3D visualizations in understanding complex multiscale biological systems.

You'll also learn about the program's founders, funding avenues, and the collaborative efforts advancing this research.

We are glad you can join.

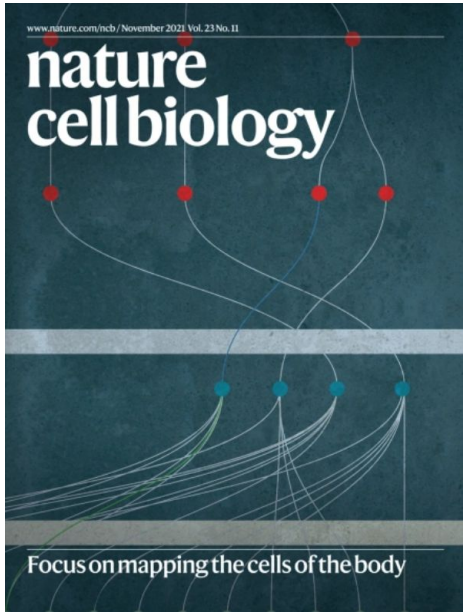
Previous 24h Events: **Science Map 2021** | **Human Reference Atlas 2022** | **Interactive Data Visualizations 2023**

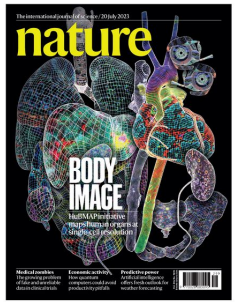
Human Reference Atlas

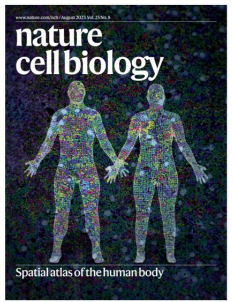
The Human Reference Atlas (HRA)

1. defines the 3D space and shape of anatomical structures and cell types that are of biomedical relevance plus the biomarkers used to characterize them. Anatomical structures, cell types and biomarkers are validated and represented in/added to ontologies (Uberon/FMA, CL, HGNC).
2. defines how new datasets can be mapped to the HRA, e.g., spatially using the Visible Human CCF or Vasculature CCF (or both, see next slide), via ASCT+B ontology terms/IDs, or via gene expression data as in Azimuth.
3. it is
    - authoritative (there exists expert agreement and it was validated by data),
    - computable (supports API queries, UIs),
    - published as LOD (connected to gene, disease, and other ontologies and data),
    - open (anyone can use the HRA data and code), and
    - continuously evolving (e.g., as new technologies become available).

https://www.nature.com/articles/s41556-021-00788-6

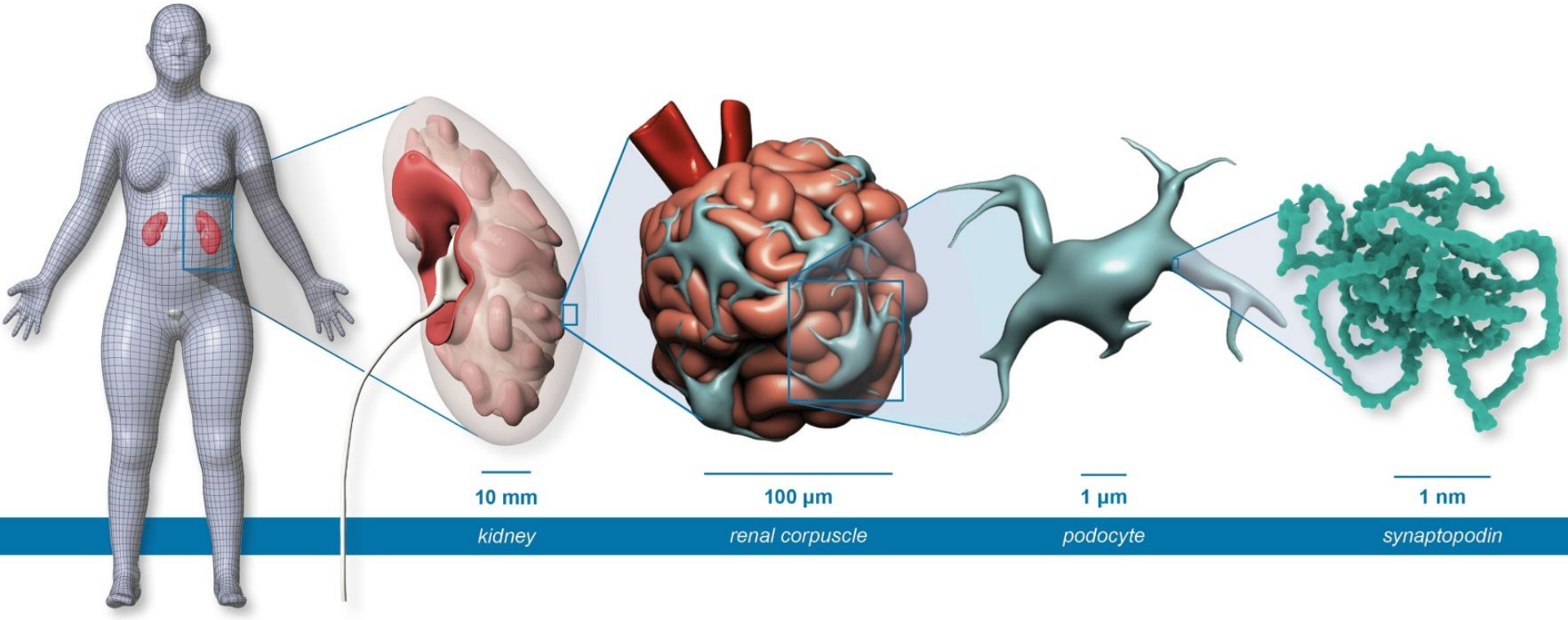






Human Reference Atlas

A multiscale, high-resolution, three-dimensional, ontologically aligned atlas of anatomical structures and cells in the healthy human body



HRA-focused HIVE Marker Paper

Accepted as *Nature Methods* 'Resource' paper

The preprint is at https://www.biorxiv.org/content/10.1101/2024.03.27.587041v3

Thanks go to all 170+ Core and HRA Team authors who made this possible.

It is our hope that this joint paper helps align efforts and optimize data formats, APIs.

Cold Spring Harbor Laboratory

bioRxiv

THE PREPRINT SERVER FOR BIOLOGY

HOME | SUBMIT | FAQ | BLOG | ALERTS / RSS | RESOURCES | ABOUT | CHANNELS

Search

Advanced Search

New Results

Follow this preprint

Previous

Next

Human BioMolecular Atlas Program (HuBMAP): 3D Human Reference Atlas Construction and Usage

Katy Börner, Philip D. Blood, Jonathan C. Silverstein, Matthew Ruffalo, Sarah A. Teichmann, Gloria Pryhuber, Ravi Misra, Jeffrey Purkerson, Jean Fan, John W. Hickey, Gesmira Molla, Chuan Xu, Yun Zhang, Griffin Weber, Yashvardhan Jain, Danial Qaurooni, Yongxin Kong, HRA Team, Andreas Bueckle, Bruce W. Herr II

doi: https://doi.org/10.1101/2024.03.27.587041

This article is a preprint and has not been certified by peer review [what does this mean?].

Posted May 22, 2024.

Download PDF

Print/Save Options

Supplementary Material

Data/Code

Revision Summary

Email

Share

Citation Tools

Get QR code

X Post

Like 0

Abstract | Full Text | Info/History | Metrics | Preview PDF

COVID-19 SARS-CoV-2 preprints from medRxiv and bioRxiv

Abstract

The Human BioMolecular Atlas Program (HuBMAP) aims to construct a reference 3D structural, cellular, and molecular atlas of the healthy adult human body. The HuBMAP Data Portal (https://portal.hubmapconsortium.org) serves experimental datasets and supports data processing, search, filtering, and visualization. The Human Reference Atlas (HRA) Portal (https://humanatlas.io) provides open access to atlas data, code, procedures, and instructional materials. Experts from more than 20 consortia are collaborating to construct the HRA's Common Coordinate Framework (CCF), knowledge graphs, and tools that describe the multiscale structure of the human body (from organs and tissues down to cells, genes, and biomarkers) and to use the HRA to understand changes that occur at each of these levels with aging, disease, and other

Subject Area

Bioinformatics

Subject Areas

All Articles

Animal Behavior and Cognition

Biochemistry

Bioengineering

Bioinformatics

Biophysics

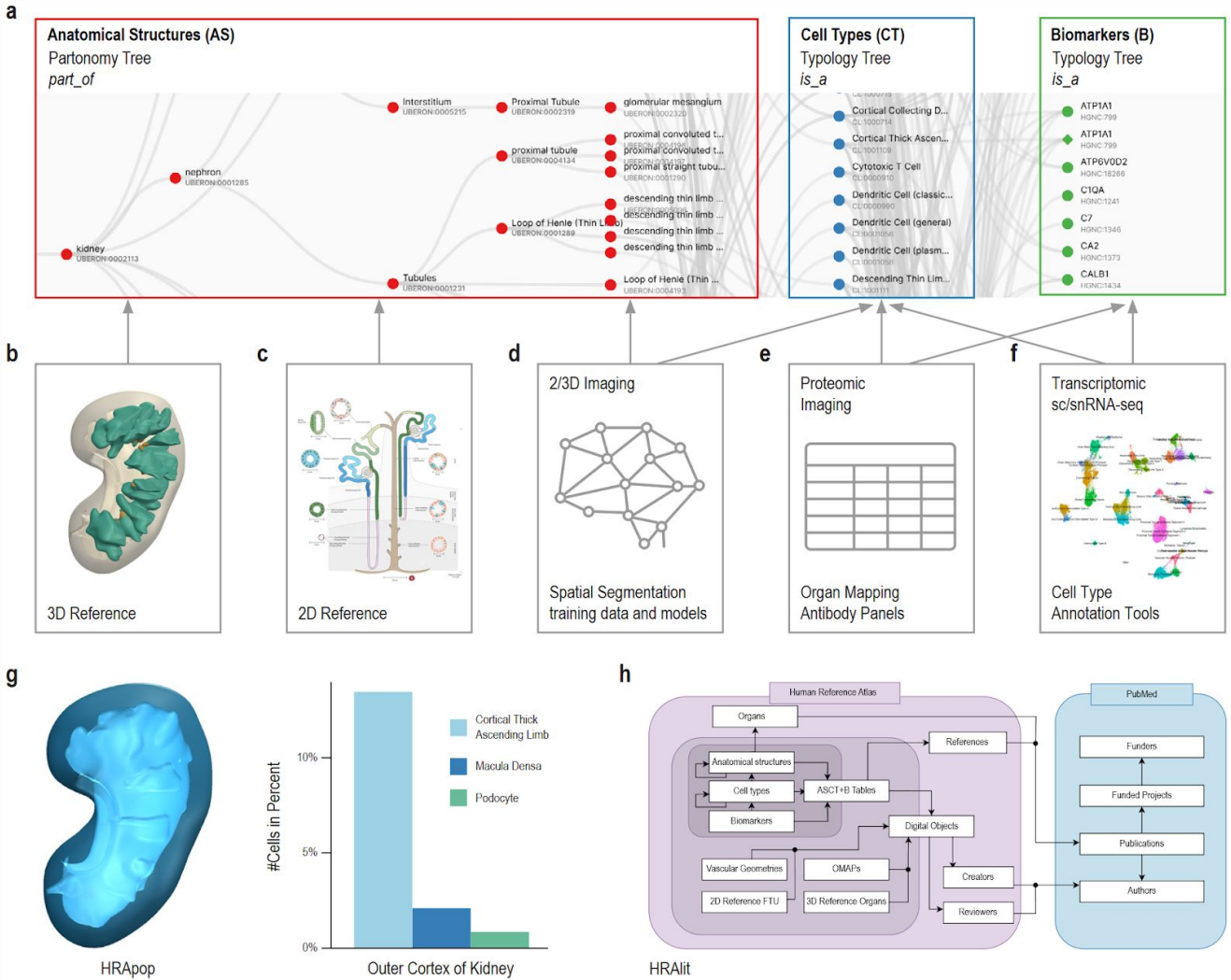


Figure 1: Human Reference Atlas (HRA) components and linkages. a. The anatomical structures, cell types and biomarkers (ASCT+B) tables document the nested *part_of* structure of organs (e.g., cells that make up functional tissue units, successively larger anatomical structures, an entire organ such as the kidney, which is *part_of* the body). The cells that make up (are *located_in*) each of the anatomical structures are organized in a multi-level cell type typology with 'cell' at the root and more and more specialized child nodes. The biomarkers used to *characterize* cell types might have one of five types: genes, proteins, metabolites, proteoforms, and lipids organized in a biomarker typology. Gray arrows indicate crosswalks that connect other HRA DOs to ASCT+B tables. **b.** The HRA 3D reference objects represent the shape, size, location, and rotation of 1,218 3D anatomical structures of 356 types for 65 organs with crosswalks to ASCT+B tables. Shown are 'renal papilla' and 'renal pyramid' in the kidney. **c.** 2D reference illustrations document the shape, size, and spatial layout of 3,726 2D cells of 131 types for 22 FTUs in 10 organs with crosswalks to ASCT+B tables. Shown is the kidney nephron. **d.** Labeled training data exist for FTUs in five organs with crosswalks (gray arrows) to anatomical structures and cell types in the ASCT+B tables. **e.** 13 Organ Mapping Antibody Panels (OMAPs) are linked to 197 Antibody Validation Reports (AVRs) and there exist crosswalks to cell types and biomarkers in ASCT+B tables. **f.** 10 Azimuth references for healthy adult organs plus crosswalks to cell types and biomarkers in ASCT+B tables. **g.** Cell type populations from single cell experimental data exist for 74 3D anatomical structures across 23 organs with 13 unique UBERON IDs in the HRA. Shown is the 'outer cortex of kidney' on left and a bar graph that plots the percentage of cells for three cell types in this anatomical structure on right. **h.** The HRAlit database links HRA DOs to existing ontologies (e.g., Uberon, CL), expert ORCID, publication evidence, funding, and experimental data used for HRApop computation.

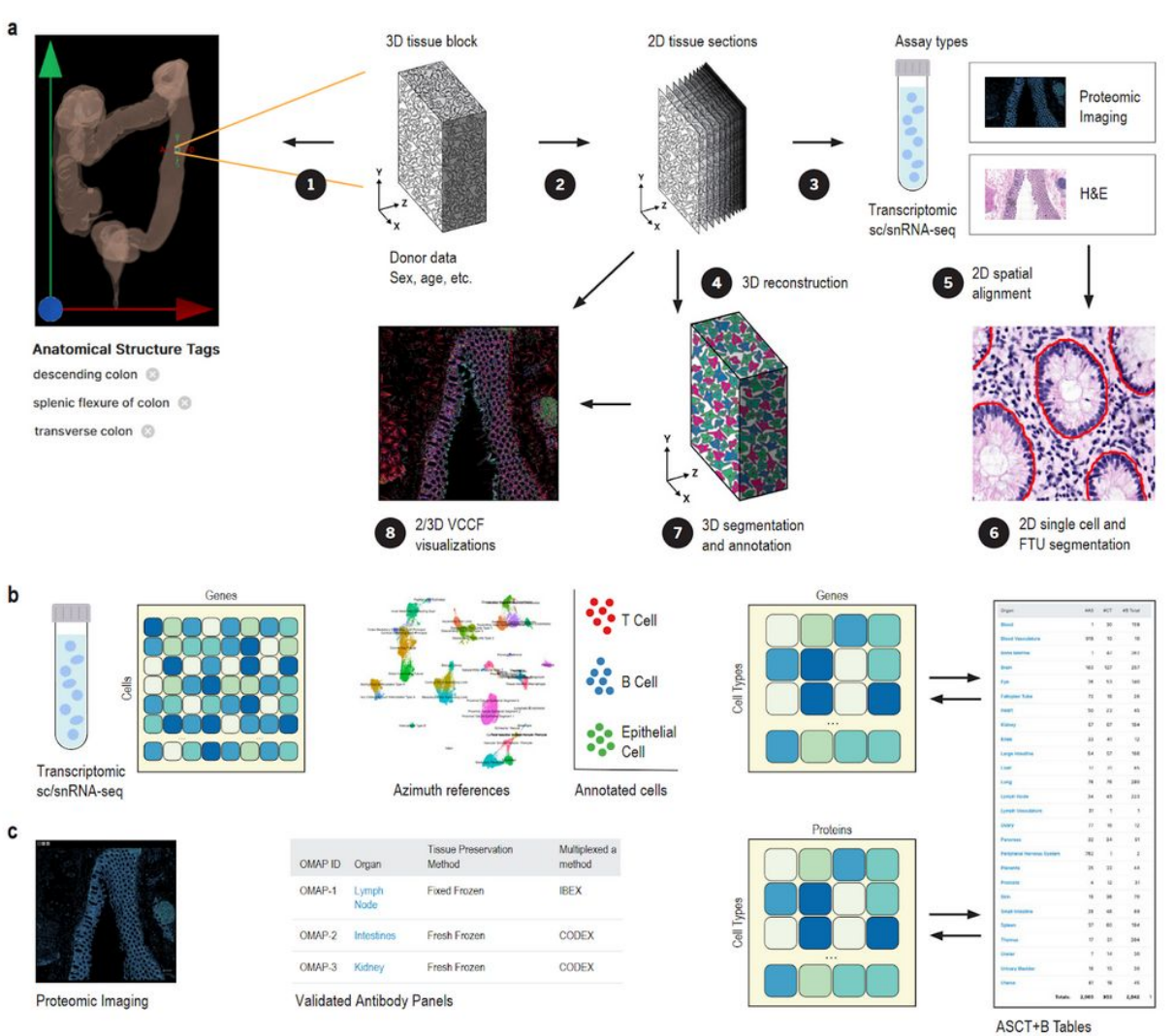


Figure 2: Mapping experimental data to the HRA. a. A 3D tissue block is spatially registered and semantically annotated using the Registration User Interface or the Millitome, see (1). A smaller part of the tissue block might be used for sc/snRNA-seq analysis (not shown) or cut into tissue sections (2). Tissue sections are analyzed using the very same or different assay types (3). Shown are single cell transcriptomics (e.g., sc/snRNA-seq), OMAP-aligned spatial proteomics (e.g., CODEX, Cell DIVE), and high resolution hematoxylin and eosin (H&E) stained histology images. Spatial alignment of different assay types for the very same or different tissue sections is non-trivial (5). H&E data is used to segment functional tissue units (FTUs) using trained machine learning models (6). 3D reconstruction of tissue volumes is accomplished by aligning data from multiple serial tissue sections computationally (4) followed by 3D segmentation and annotation (7). 2D or 3D data is analyzed to identify the distance of different cell types to the vasculature (VCCF Visualizations) as a multiscale common coordinate framework from which no other cell is very distant (8). **b.** Single cell/nucleus data (sc/snRNA-seq) is stored as a cell by gene matrix; cell types are annotated using Azimuth or other cell type annotation tools; results are aggregated to cell type by gene biomarker expression value matrices that are aligned with the ASCT+B tables; and are used in diverse HRA user interfaces (e.g., Exploration User Interface and FTU Explorer). **c.** OMAP-aligned spatial data generated using validated antibody panels linked to AVRs are analyzed to compute cell type by protein biomarker expression value matrices that are aligned with the ASCT+B tables using semi-automated workflows.

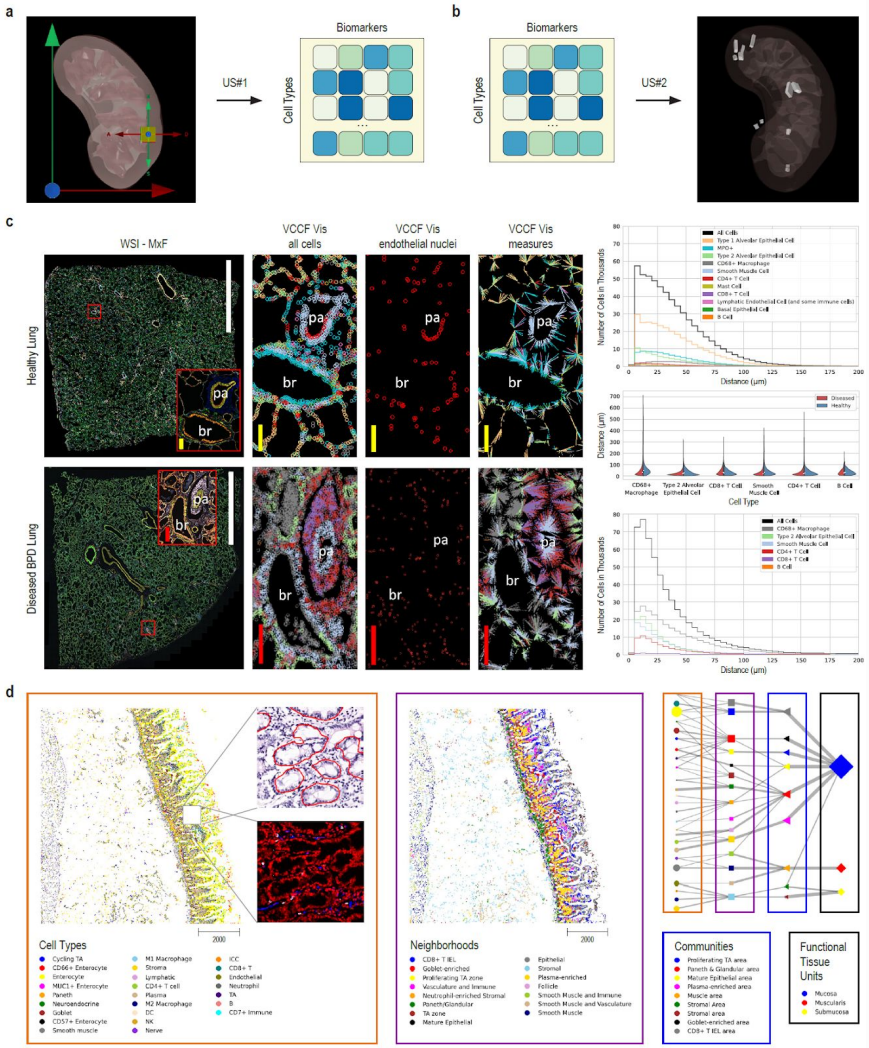
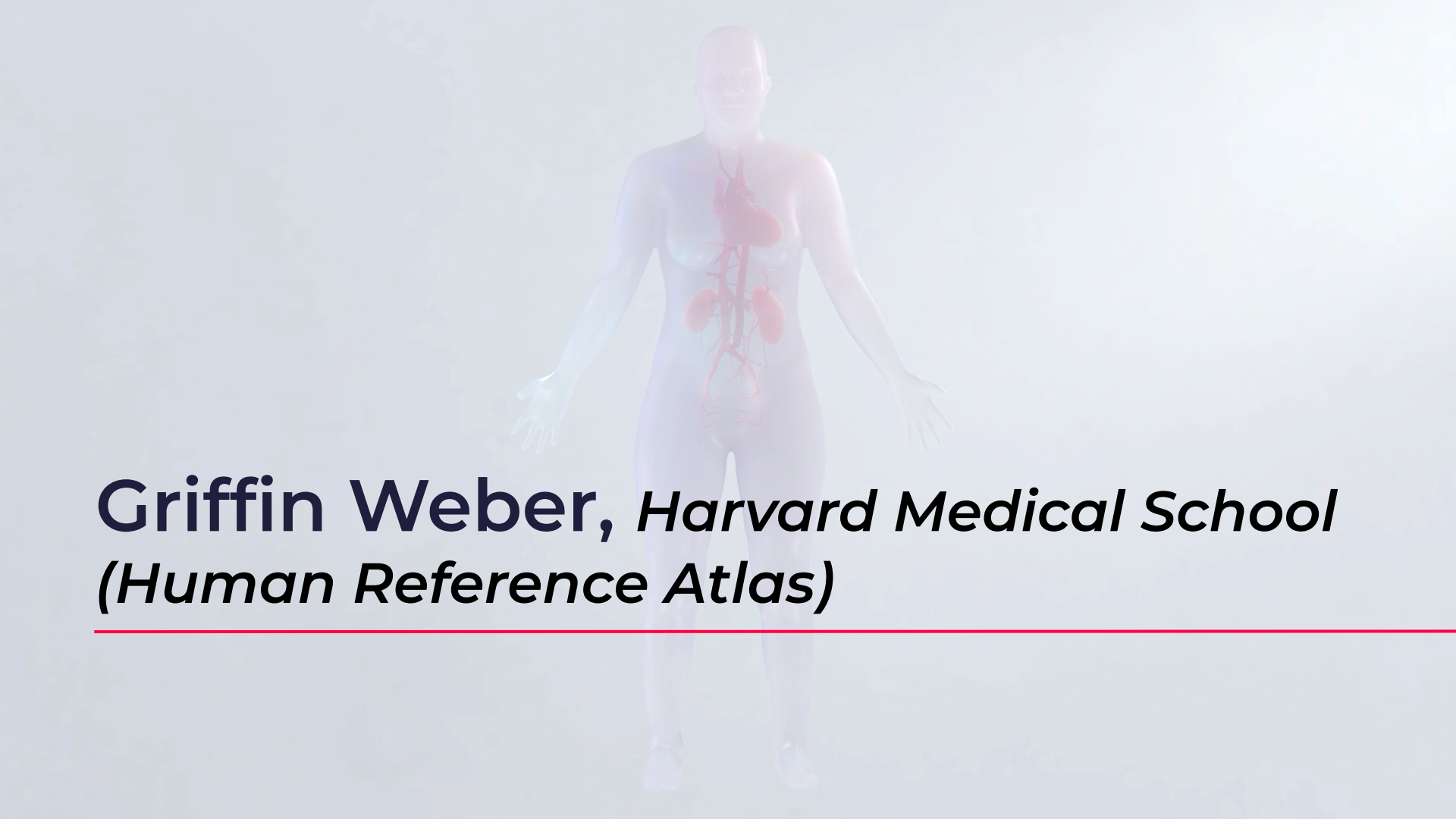


Figure 3: Human Reference Atlas Usage. a. User story #1 (US#1) lets a user define a 3D volume inside of the HRA reference body using the RUI and it predicts cell type populations and mean expression values for cell types in that volume. **b.** User story #2 (US#2) reads cell type population data and predicts the 3D origin of tissue, shown as a collection of extraction sites that have a similar cell type population. **c.** HRA can be used to compare the distribution of parenchymal cells including endothelial, epithelial, and muscle that compose the blood vessels, airways and gas exchanging functional lung structures, and resident immune cells including macrophages, to local vasculature (VCCF Visualizations) in healthy (top) and diseased (bottom) lung using multiplexed immunofluorescence microscopy images with bronchiole (br) and an accompanying small pulmonary artery (pa). Scale bar legend: white: 5 mm, red: 200 μm, yellow: 100 μm. The graphs on the right show distance distributions for cell types present in the healthy lung (top) and diseased BPD lung (bottom); the violin plot (middle) shows a comparison between distance distributions for cell types common in both datasets. **d.** Multi-level cell neighborhoods can be computed to analyze and communicate the structure and function of FTUs; tissue image with cell type annotations and zoom into H&E with FTU segmentations (red outlines) and zoom into the multiplexed image (CODEX) is shown in left, neighborhoods are given in the middle; hierarchy of FTUs, neighborhoods, communities, and cell types are shown on the right.



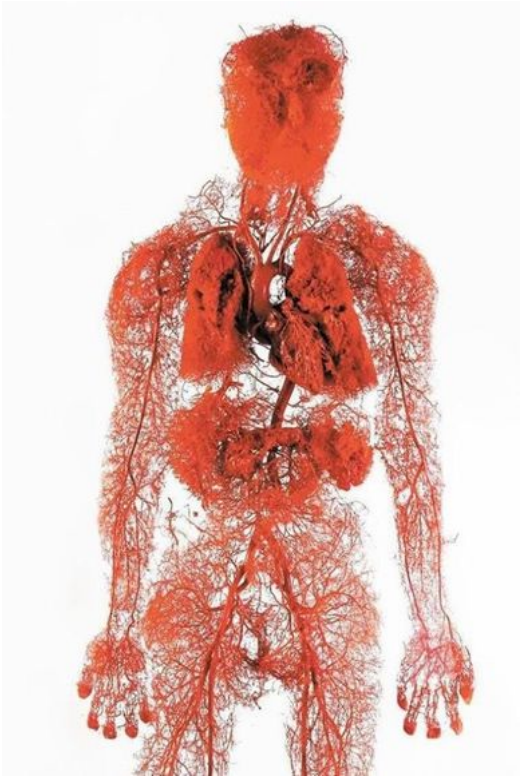
Griffin Weber, *Harvard Medical School (Human Reference Atlas)*

Vasculature Common Coordinate Framework

Griffin M Weber, MD, PhD
Associate Professor of Biomedical Informatics
Harvard Medical School
weber@hms.harvard.edu

Katherine S Gustilo
Sujin Lee
Rajeev Malhotra
Marc Halushka
Ellen M Quardokus

Avinash Boppana
Bruce W Herr II
Ushma Patel
Zorina Galis
Katy Börner

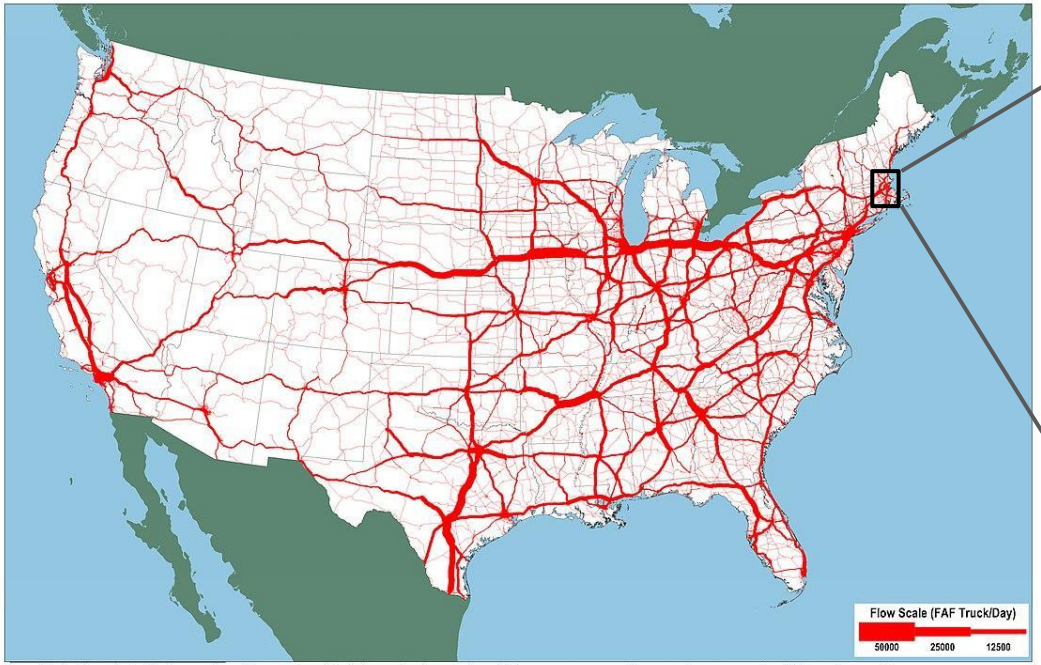


https://bodyworlds.com/

Multiscale Maps of Roads

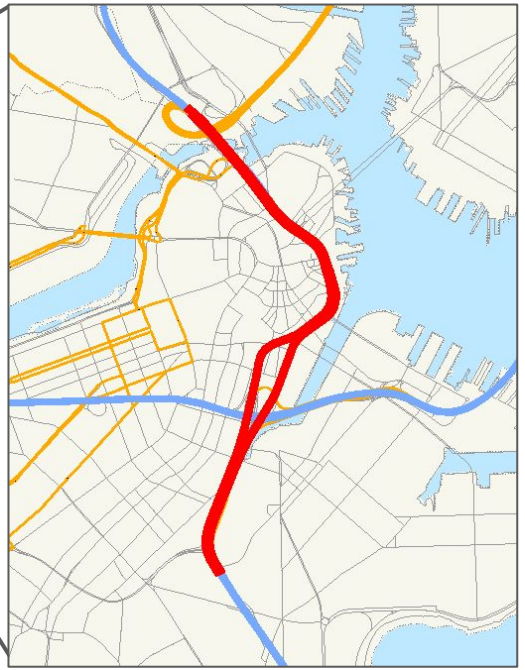
(similar to multiscale maps of blood vessel pathways through the body)

Daily traffic, U.S. National Highway System



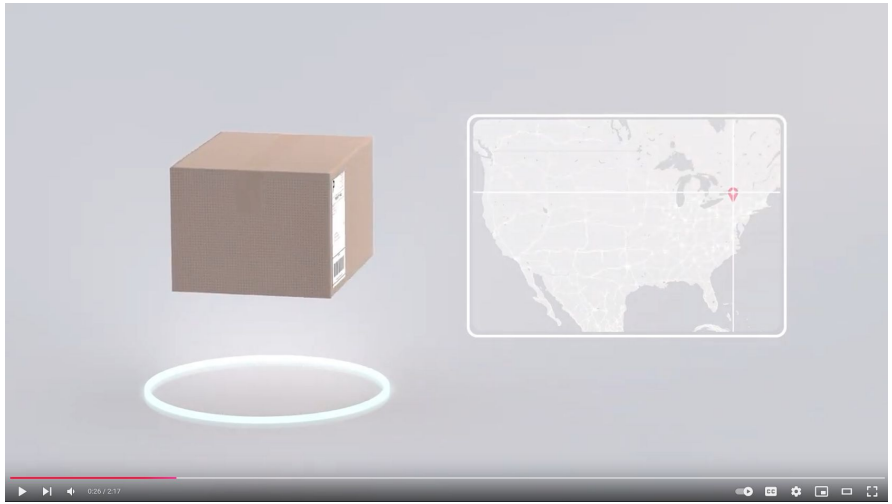
Note: Major flows include domestic and international freight moving by truck on highway segments with more than twenty five FAF trucks per day and between places typically more than fifty miles apart.
Source: U.S. Department of Transportation, Federal Highway Administration, Office of Freight Management and Operations, Freight Analysis Framework, version 4.3, 2017.

Boston “Central Artery”

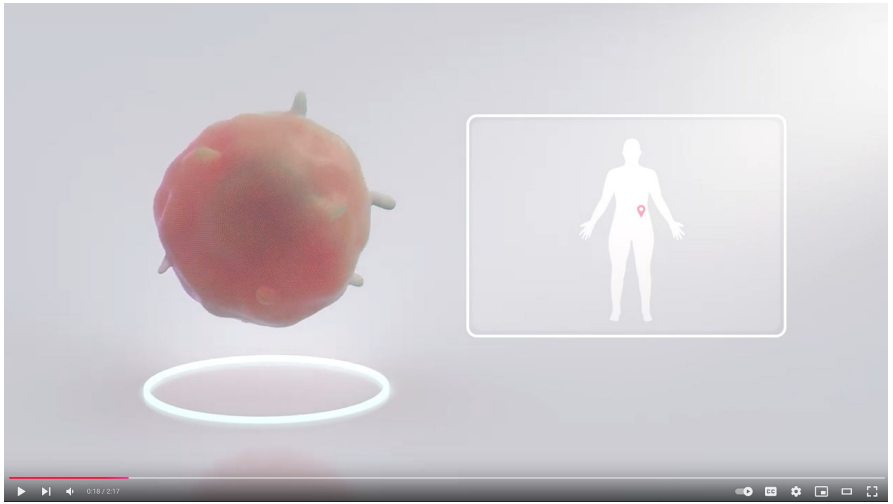


John F. Fitzgerald Expressway, By Sswonk, Public Domain, https://commons.wikimedia.org/w/index.php?curid=4538754

Trucks follow roads to deliver a package to a house



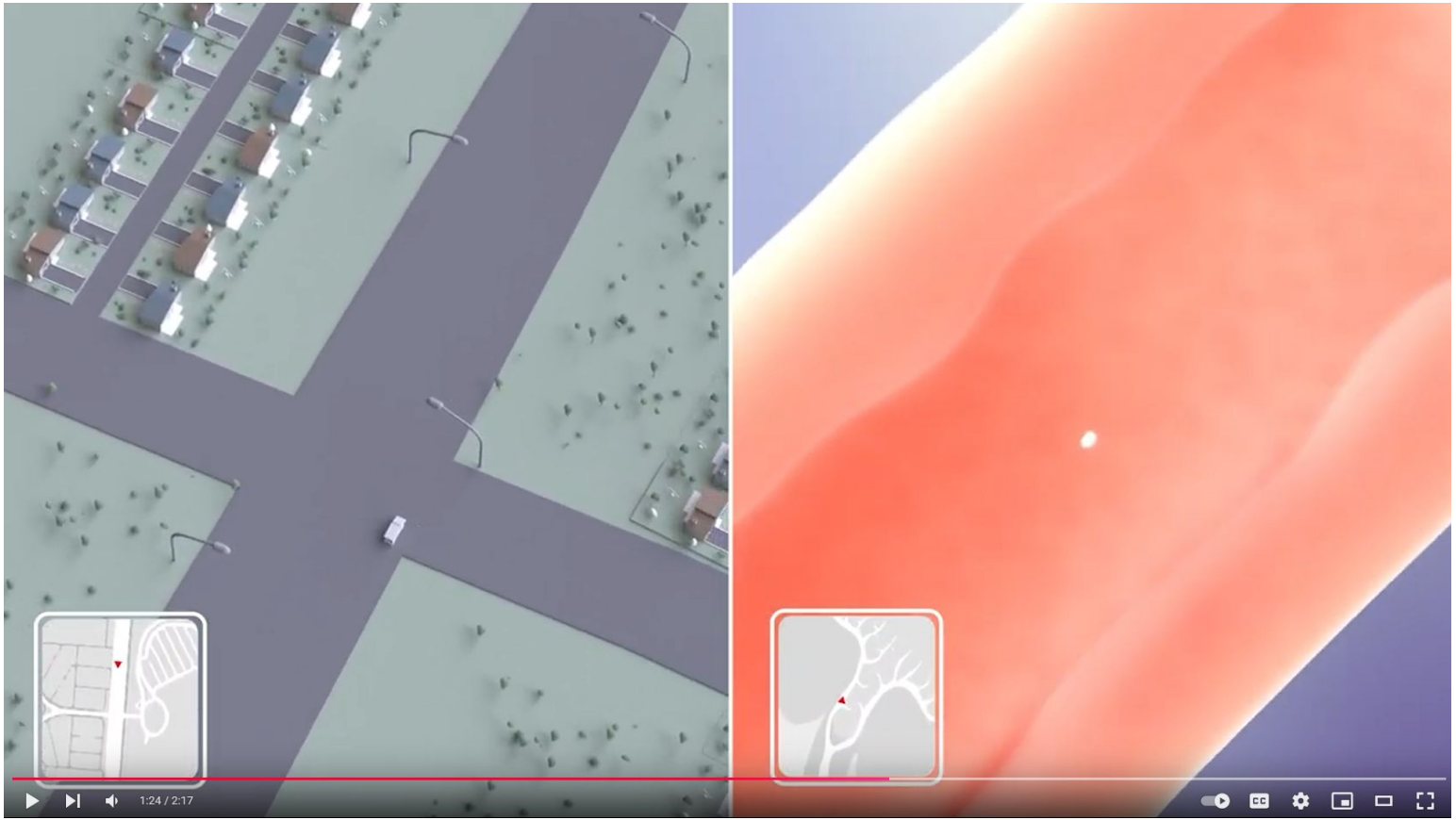
Blood cells follow vessels to deliver oxygen to organs



https://www.youtube.com/watch?v=zQeMgxo8n_U

Highway (1000 km)

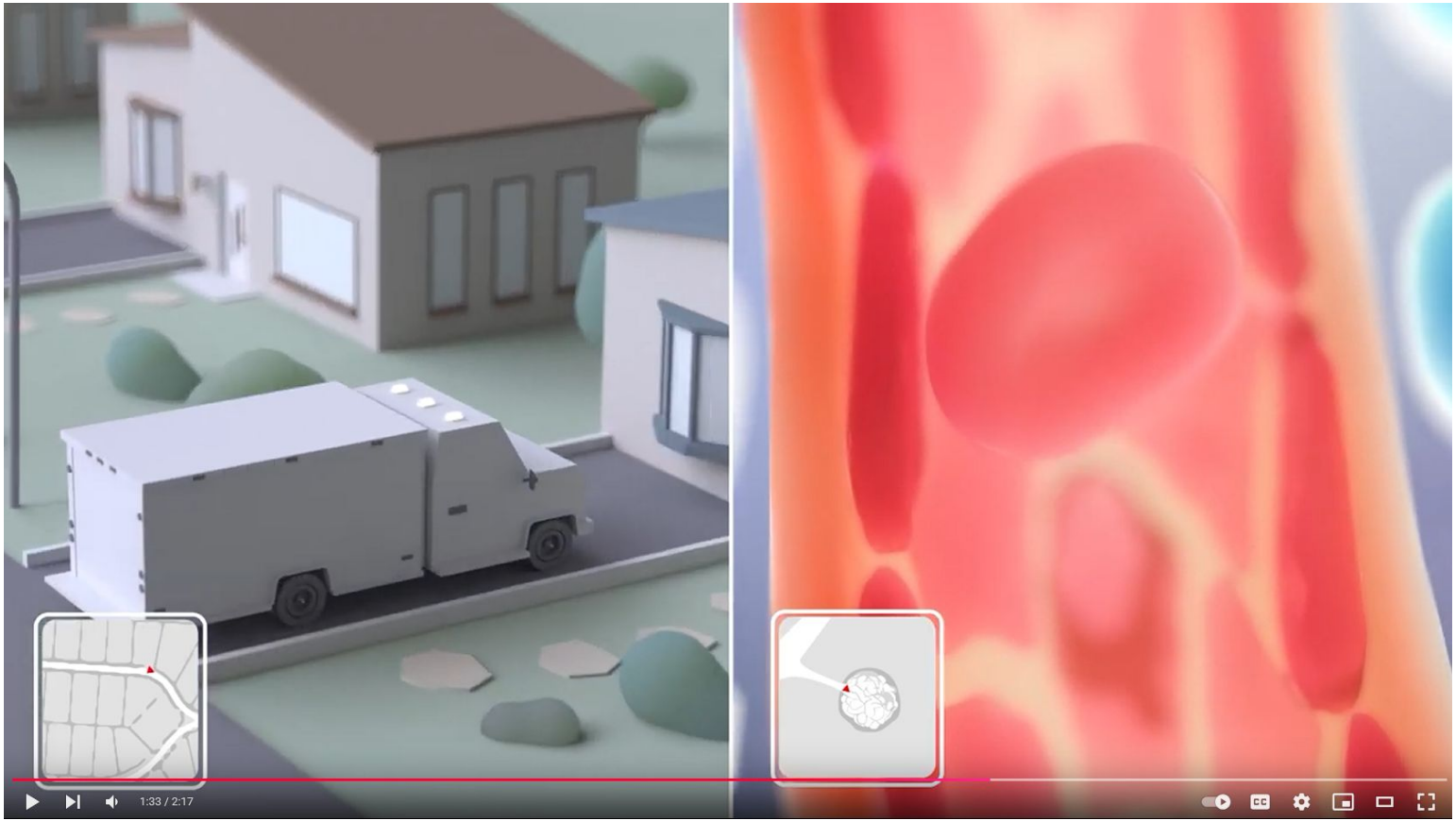
Artery (1 m)



https://www.youtube.com/watch?v=zQeMgxo8n_U

Street (1 km)

Arteriole (0.5 cm)



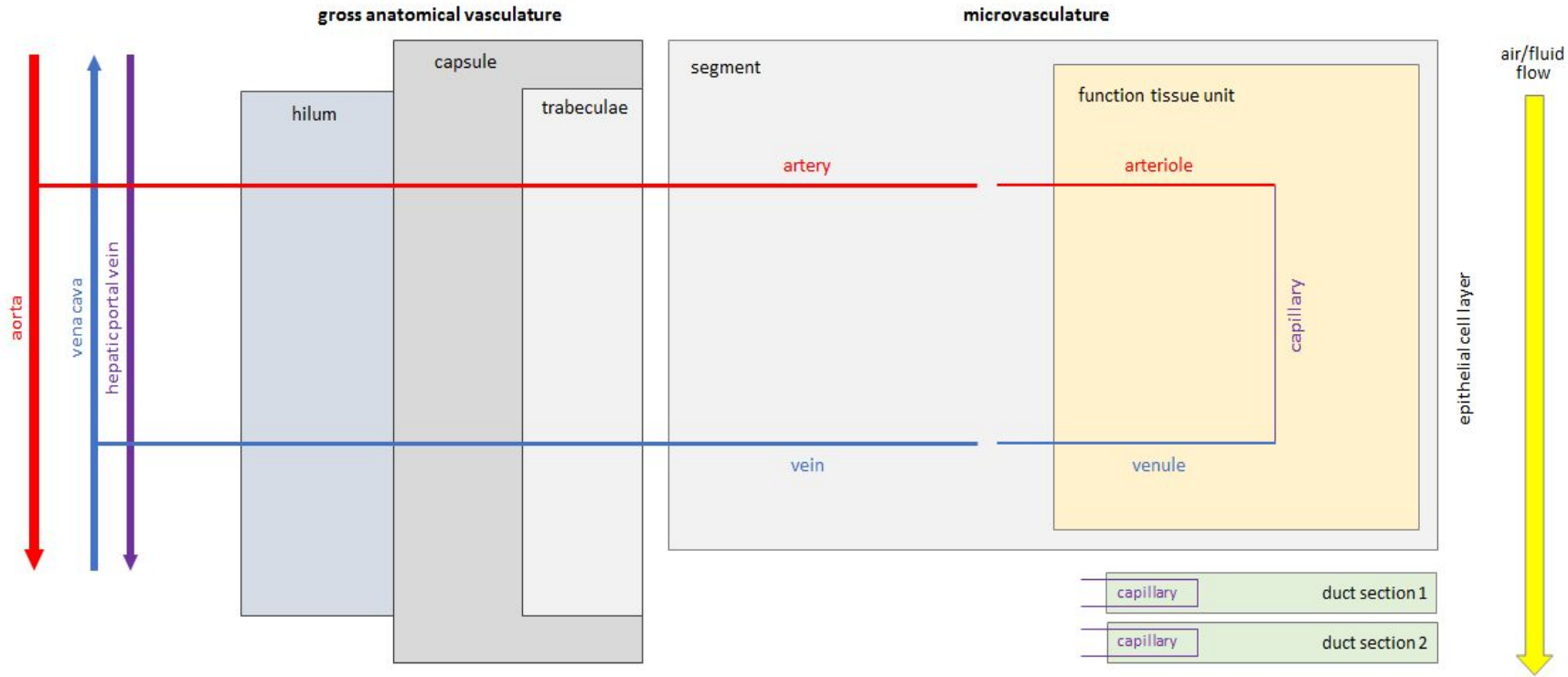
https://www.youtube.com/watch?v=zQeMgxo8n_U

Driveway (10 m) Capillary (0.5 mm x 0.01 mm)

https://www.youtube.com/watch?v=zQeMgxo8n_U

Blood Vasculature to Organ Crosswalk Diagrams

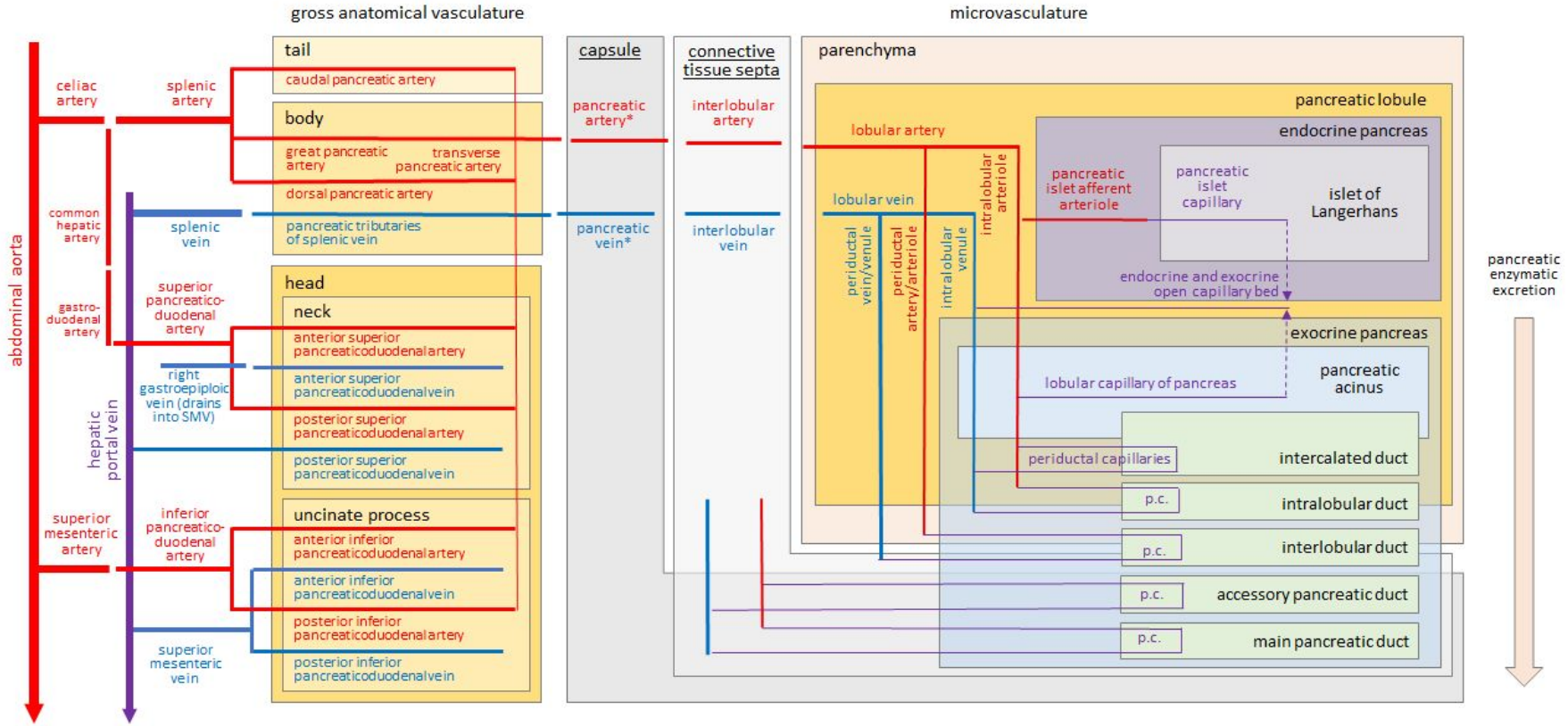
Template



Boppana A, Lee S, Malhotra R, Halushka M, Gustilo KS, Quardokus EM, Herr BW 2nd, Börner K, Weber GM. Anatomical structures, cell types, and biomarkers of the healthy human blood vasculature. Sci Data. 2023 Jul 19;10(1):452. doi: 10.1038/s41597-023-02018-0.

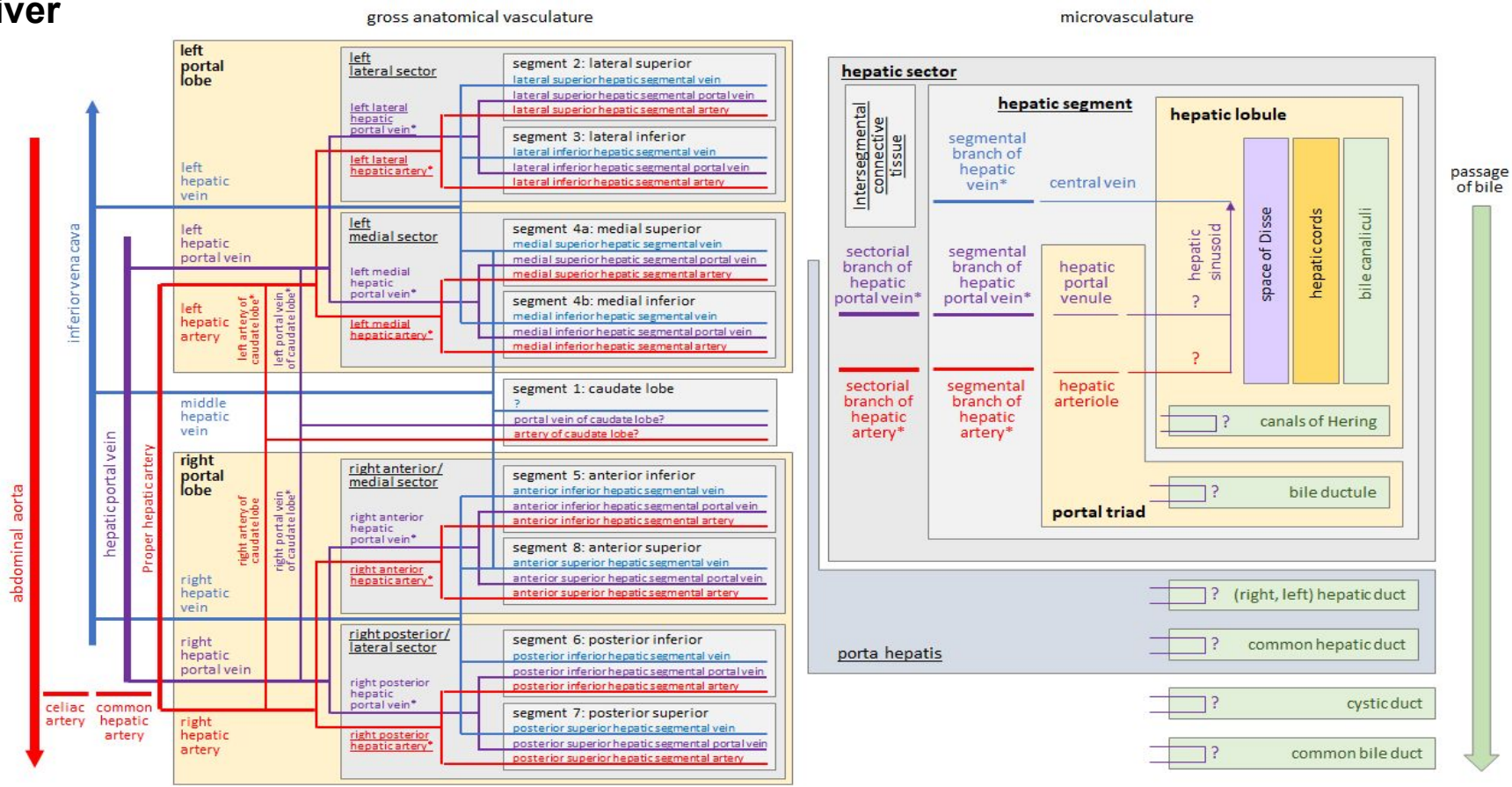
Blood Vasculature to Organ Crosswalk Diagrams

Pancreas



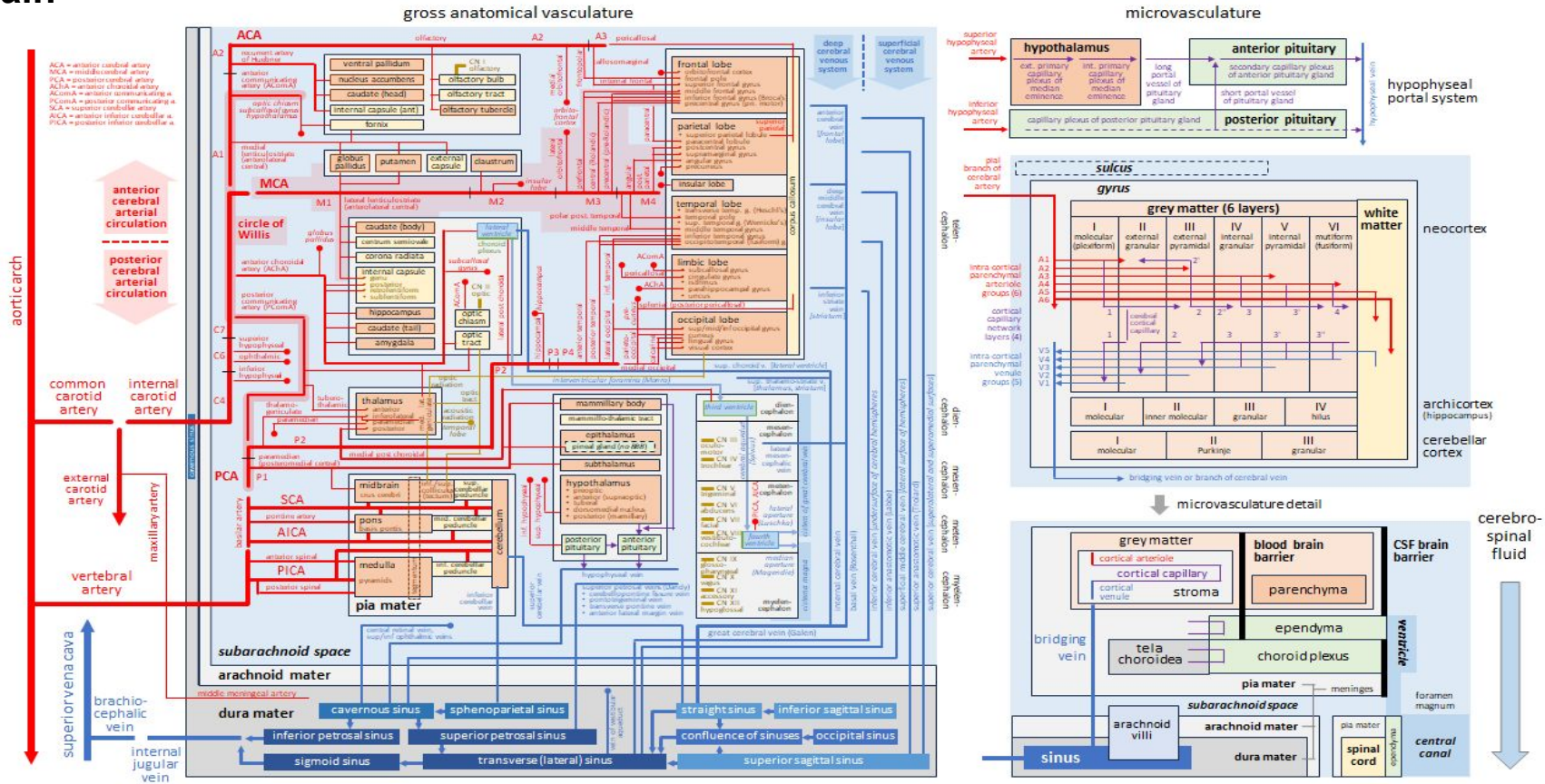
Blood Vasculature to Organ Crosswalk Diagrams

Liver



Blood Vasculature to Organ Crosswalk Diagrams

Brain



Karen Miga, *UC Santa Cruz (Pangenome)*

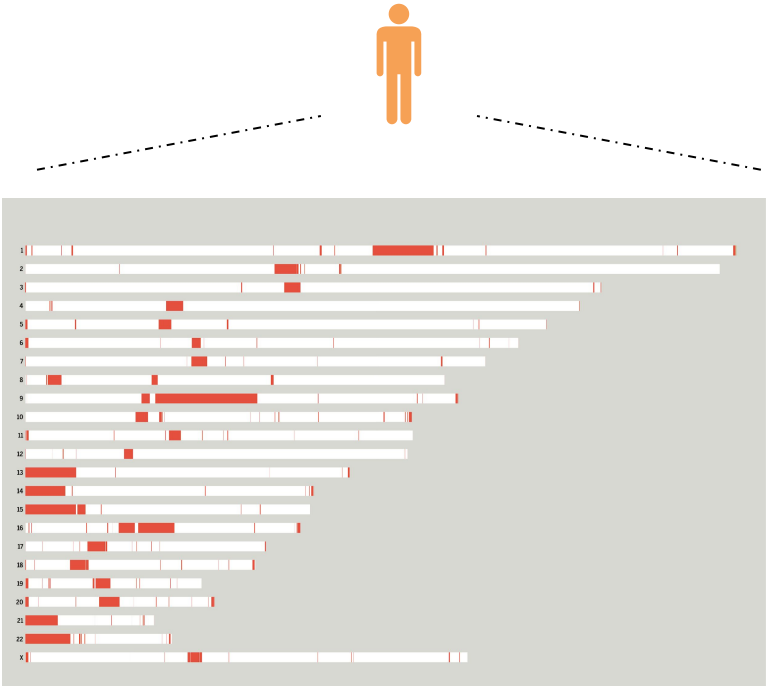
A Need to Modernize the Human Reference Genome

- The human reference genome is a foundational resource in human genetics and **like most technology-driven resources, is overdue for an upgrade.**

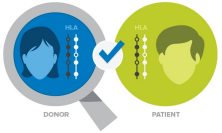
- Improvements in long-read sequencing and assembly methods allow the **production of high-quality genomes**

- The current structure is a **linear haplotype, largely representing a single individual.** This introduces biases and excludes sequence variation.

One genome cannot represent the genetic diversity of the human species



HLA-A Typing:

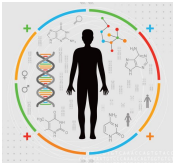


Autoimmunity Allergy, Transplant (allogenic stem cell transplant, solid organ, and blood marrow)

SMA Locus

Leading cause of early infant death (1:6,000-10,000 live births)

CYP2D6



Responsible for the metabolism of around 25% of clinically used drug

Human Pangenome Reference Consortium

- Call to action from NHGRI to improve representation of **global genomic diversity** (common alleles, ~1% AF)
- **Organized a team of researchers with expertise in long read technologies, complex variation, and T2T assemblies.**

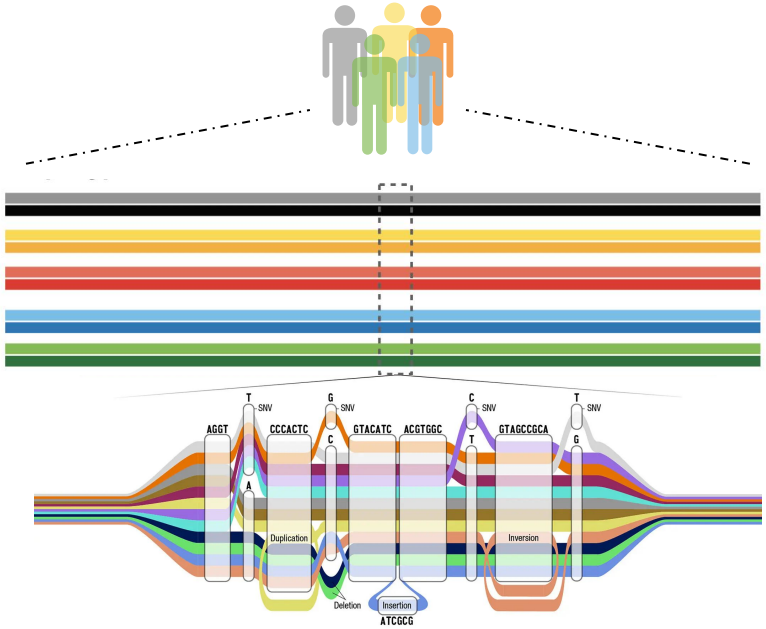
TELOMERE-TO-TELOMERE CONSORTIUM

VERTEBRATE GENOMES PROJECT
A PROJECT OF THE G10K CONSORTIUM

HGSVC

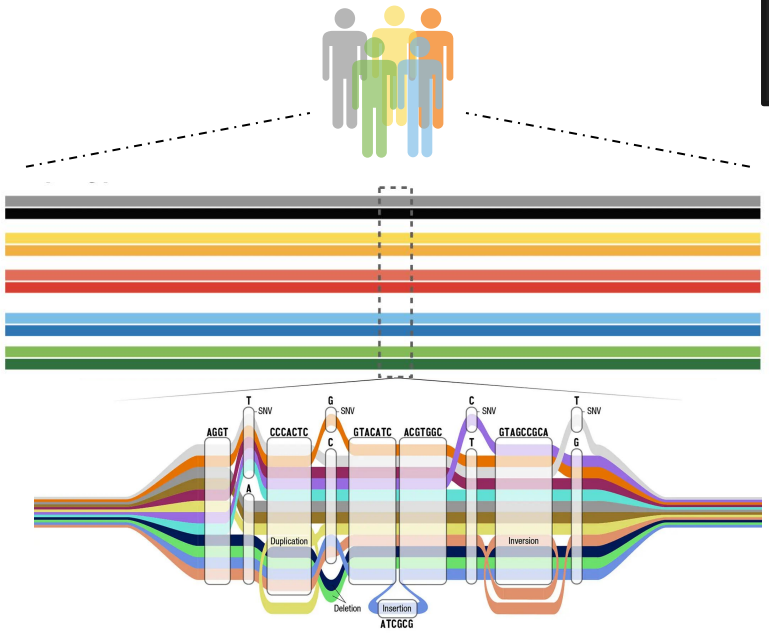
- **Develop a new, non-linear reference data structure** and foster an innovative ecosystem of pangenomic tools

Pangenome reference resource will better represents and serve humanity



One genome can introduce bias in genomics medicine initiatives

Pangenome reference resource will better represents and serve humanity



~1000 haplotypes

Assemblies

Alignments

Annotation

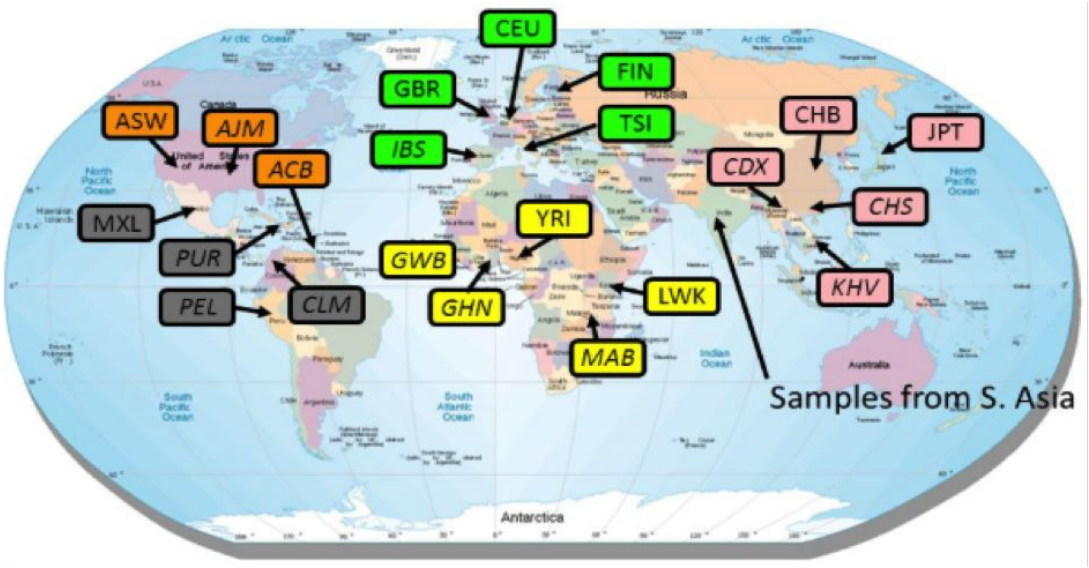
OPEN ACCESS

New Tooling Ecosystem

- Pangenome construction: T2T Assemblies, QC, Genome Annotation
- Read/assembly alignment and variant calling

Who is currently represented in the Pangenome?

Pangenome should comprehensively capture most common variants, defined as variants at **>1% frequency**, in human populations globally



Phase I: Use of 1000 Genomes Cell Lines

Advancement in automated T2T assemblies

TELOMERE-TO-TELOMERE CONSORTIUM

Rautiainen, M., Nurk, S., Walenz, B.P. et al. Telomere-to-telomere assembly of diploid chromosomes with Verkko. Nat Biotechnol (2023)

Yale

HUMAN PANGENOME

Cheng, H., Jarvis, E.D., Fedrigo, O., Koepfli, K.P., Urban, L., Gemmell, N.J., Li, H. Haplotype-resolved assembly of diploid genomes without parental data. Nat Biotechnol (2022)

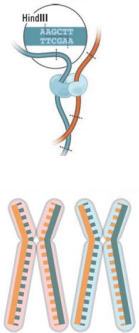
Assembly methods: **verkko** and **hifiasm-UL**

Both methods rely on a combination of
long accurate reads (**PacBio HiFi**) + ultra-long data (**ONT-UL**) +
Phasing data (**Illumina** HiC or Strand-Seq/verkko only)

PacBio

Oxford NANOPORE Technologies

illumina



Advancement in automated T2T assemblies

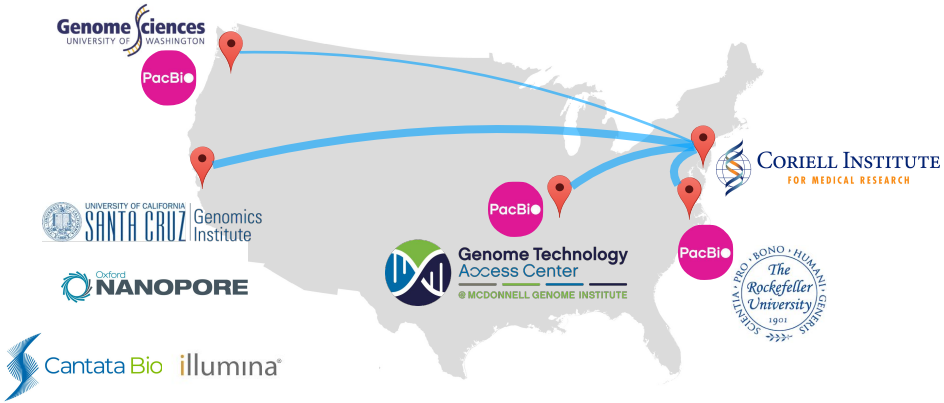
TELOMERE-TO-TELOMERE CONSORTIUM

Rautiainen, M., Nurk, S., Walenz, B.P. et al. Telomere-to-telomere assembly of diploid chromosomes with Verkko. Nat Biotechnol (2023)

Yale

HUMAN PANGENOME

Cheng, H., Jarvis, E.D., Fedrigo, O., Koepfli, K.P., Urban, L., Gemmell, N.J., Li, H. Haplotype-resolved assembly of diploid genomes without parental data. Nat Biotechnol (2022)



PacBio		illumina
~60x HiFi	~30x 100kb+	~60x Omni-C

Multi-Platform Approach

Jarvis*, Formenti*, *et al. Nature 2022*

Draft Release of a Human Pangenome

nature

Explore content | About the journal | Publish with us

nature > articles > article

Article | Open Access | Published: 10 May 2023

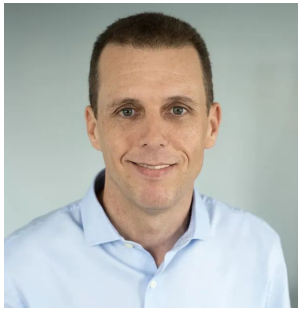
A draft human pangenome reference

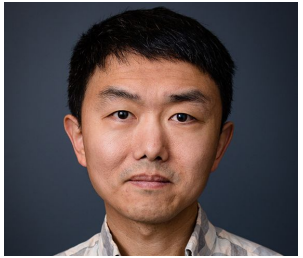
Wen-Wei Liao, Mobin Asri, Jana Ebler, Daniel Doerr, Marina Haukness, Glenn Hickey, Shuangjia Lu, Julian K. Lucas, Jean Monlong, Haley J. Abel, Silvia Buonaiuto, Xian H. Chang, Haoyu Cheng, Justin Chu, Vincenza Colonna, Jordan M. Eizenga, Xiaowen Feng, Christian Fischer, Robert S. Fulton, Shilpa Garg, Cristian Groza, Andrea Guarracino, William T. Harvey, Simon Heumos, ... Benedict Paten

+ Show authors

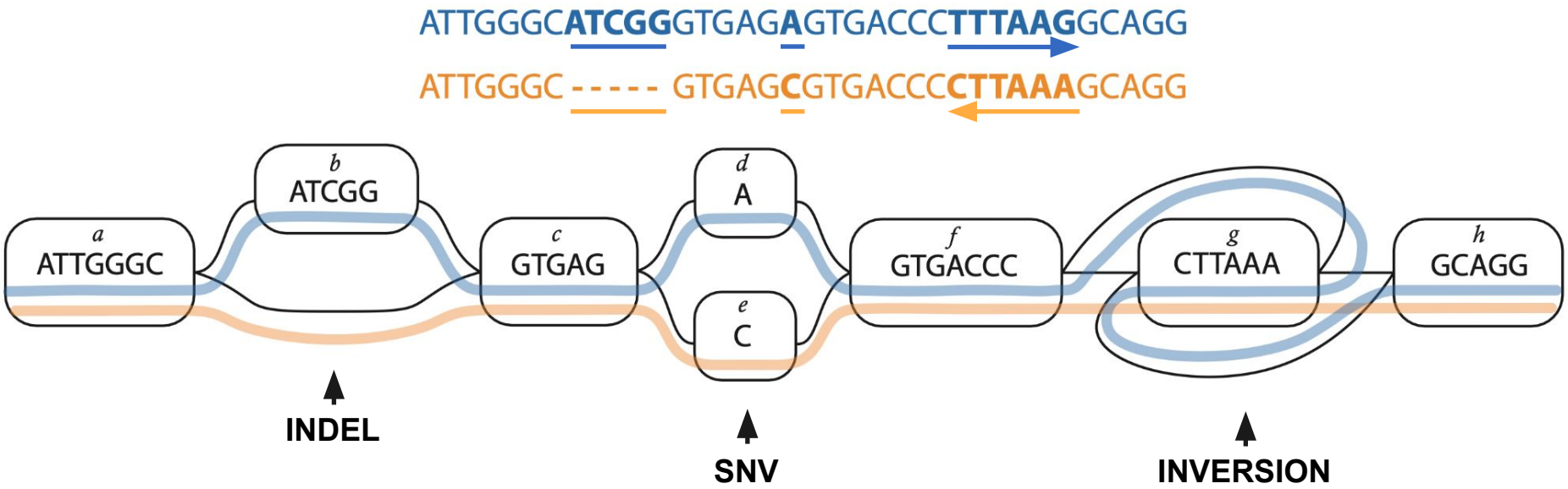
Nature **617**, 312–324 (2023) | Cite this article

5 Citations | **2985** Altmetric | Metrics

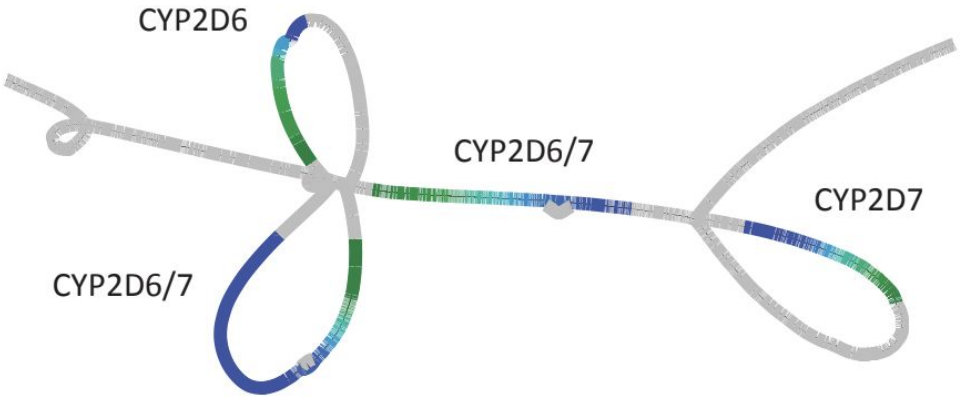




Genome Graphs



CYP2D6/7 genes: cytochrome P450 family of enzymes

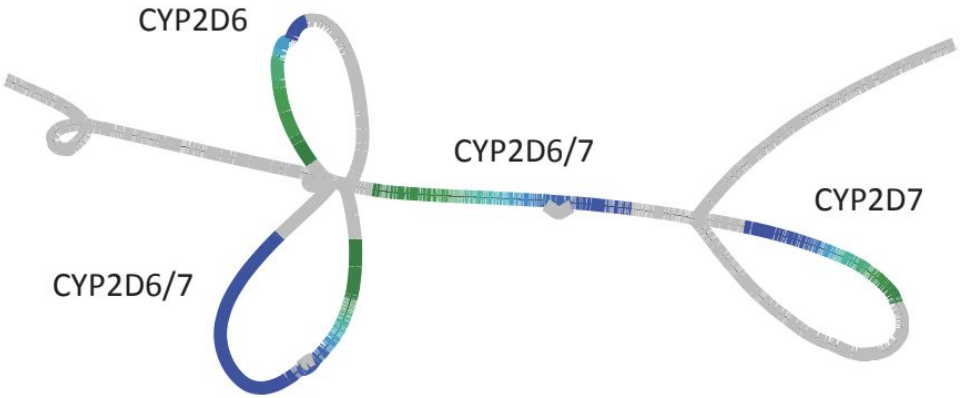


Variations in the CYP2D6 gene can greatly affect how an individual metabolizes these drugs.

CYP2D6 is particularly important because it is responsible for the metabolism of around 25% of clinically used drugs, including antidepressants, antipsychotics, analgesics, and beta-blockers.

Figure Credit:
Shuangjia Liu (Yale)

CYP2D6/7 genes: cytochrome P450 family of enzymes



Count	Frequency	Haplotype name
78	0.87	CYP2D6-CYP2D7
6	0.07	CYP2D7
2	0.02	CYP2D6-CYP2D6-CYP2D7
1	0.01	spacer-CYP2D6-CYP2D7
1	0.01	CYP2D6-CYP2D7-CYP2D7
1	0.01	CYP2D6-spacer-CYP2D6-CYP2D7
1	0.01	CYP2D6-CYP2D7-CYP2D7-CYP2D7

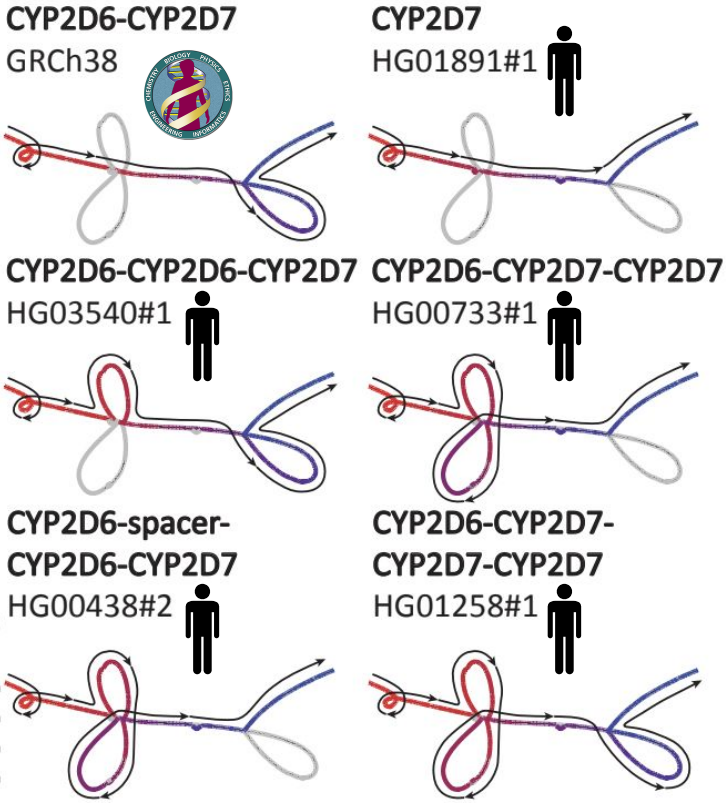
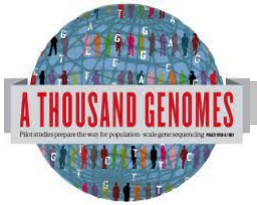


Figure Credit: Shuangjia Liu.

HPRC Pangenome Release Roadmap



"Alpha" - **Draft**
47 Genomes
May 2023

- HiFi (30x)
- ONT-UL (30x) - R9
- OmniC (30x)
- Illumina (30x)

"Beta" - **Improvement**
>170 Genomes
Early 2025

- HiFi (60x), Deep Consensus
- ONT-UL (30x) - R9, R10
- OmniC (30x)
- Illumina (30x)
- Long read RNA (Kinnex, 10M)

1.0 - **Stable "T2T"**
>350 Genomes
2026

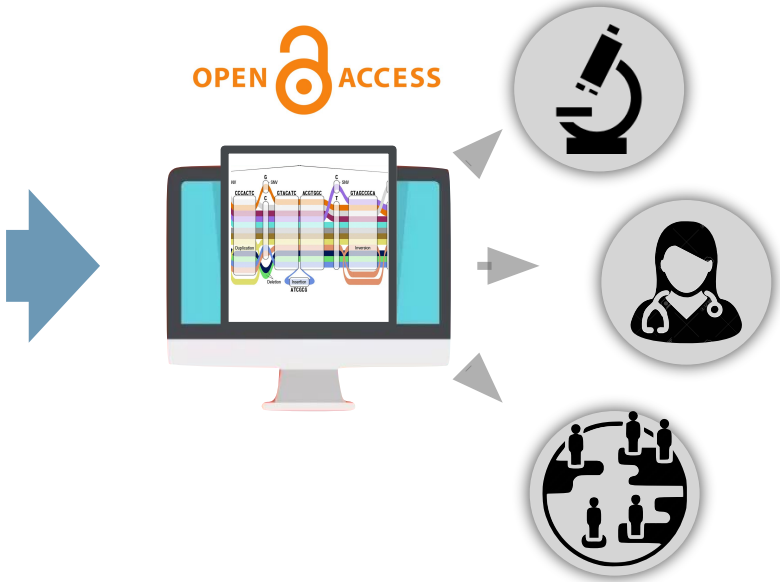
- HiFi (60x), Deep Consensus
- ONT-UL (30x) - R10
- OmniC (30x)
- Illumina (30x)
- Long read RNA (Kinnex, 10M)
- FiberSeq (40x)
- Element (30x)

1.1 - **Refinement**
>550 Genomes
2028

- HiFi (60x), Deep Consensus
- ONT-UL (30x) - R10
- OmniC (30x)
- Illumina (30x)
- Long read RNA (Kinnex, 10M)
- FiberSeq (40x)
- Element (30x)

HUMAN PANGENOME PROJECT

A Global Human Pangenome Resource



Caterina Strambio De Castillia,

UMass Chan Medical School (4D Nucleome Network, BioImaging North America, QUAREP-LiMi)

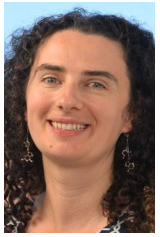
Integration of the 4D Nucleome Nuclear CCF with the HuBMAP Human Reference Atlas (HRA) CCF

Frank Alber
UCLA

Susanne Rafelski
AICS

Yin Shen
UCSF

4DN Integrating and Imaging and Omics WG

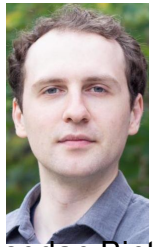


Lacra Bintu
Stanford

Caterina Strambio DC
UMass Chan

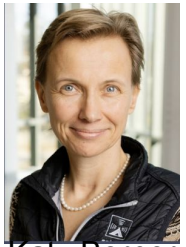
4DN Imaging WG

Quan Zhu
UCSD



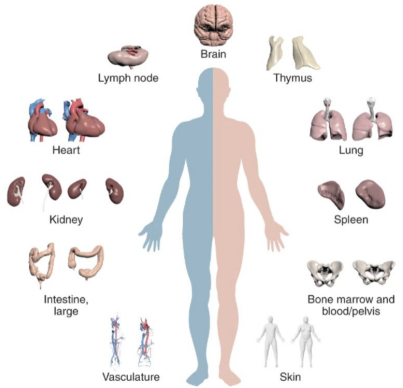
Bogdan Bintu
UCSD

Benchmarking Datasets

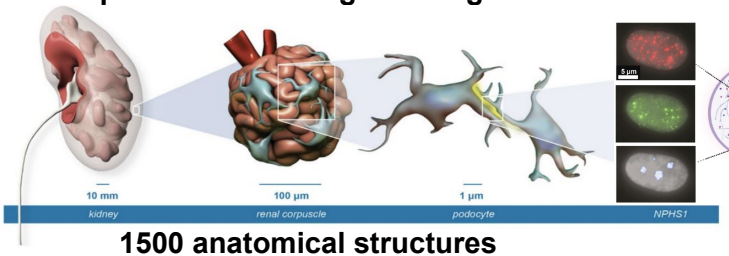


Katy Borner
Indiana U.

HuBMAP HRA CCF

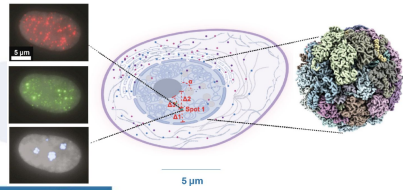


Common Coordinate Framework for Spatial and Ontological Integration

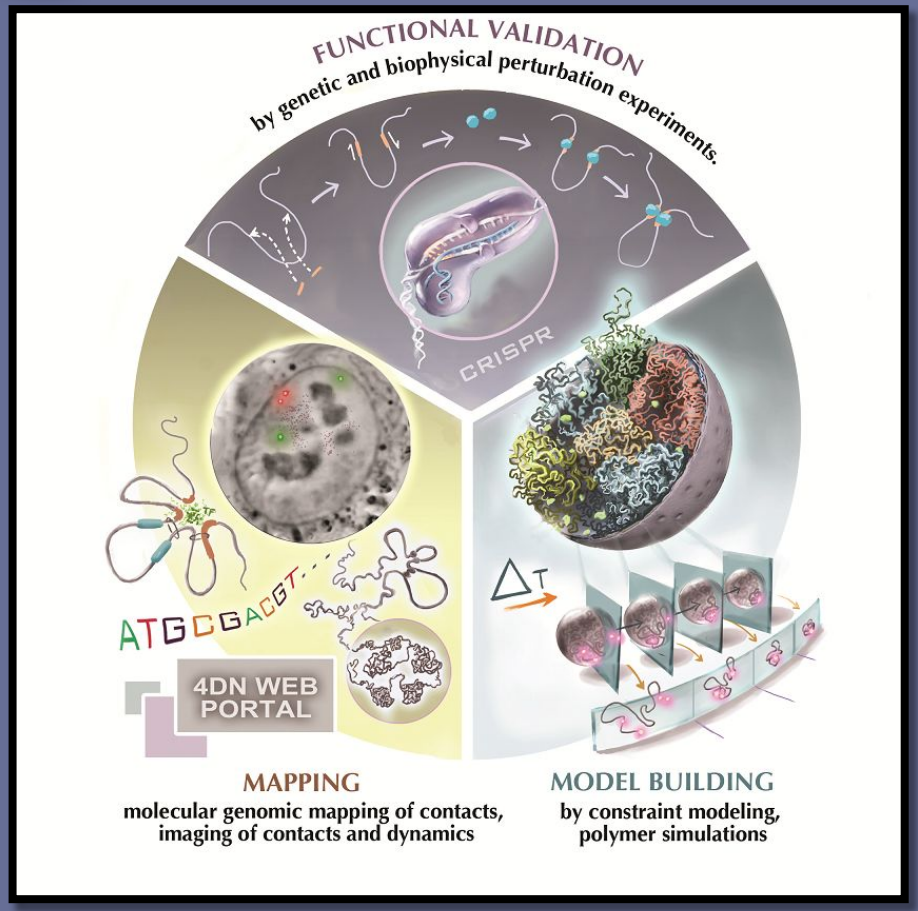


1500 anatomical structures

Integration of 4DN FISH Omics



NIH Common Fund 4D Nucleome Initiative

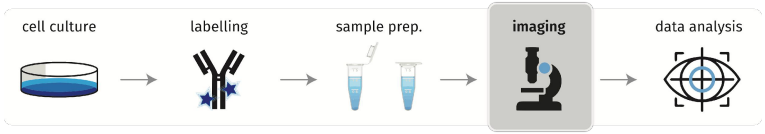


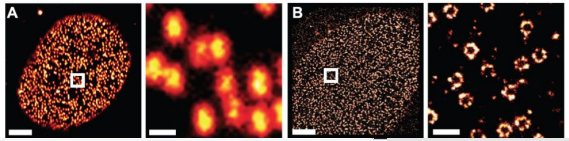
- Phase 1: 2015-19
- Phase 2: 2020-25
- 4DN Data Portal https://data.4dnucleome.org/
- Dekker et al. Current state and future aims of the 4D nucleome project. *Molecular Cell*. https://doi.org/10.1016/j.molcel.2023.06.018

Production and Utilization of FAIR Imaging Data via Inter-Consortia Partnerships

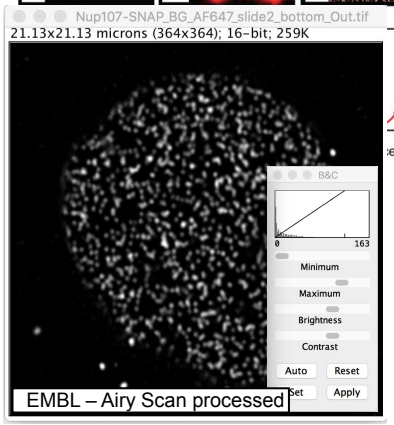
1. **Microscopy Metadata** specifications and **Micro-Meta App** initiated by 4DN to expand the OME Data Model are forming the basis for a **global effort to standardize image acquisition metadata**
2. The 4DN developed **FISH-OMICS Format for Chromatin Tracing (FOF-CT) for the exchange of results of multiplexed DNA and RNA FISH data**
3. The **4DN Nuclear Common Coordinate Framework (CCF)** developed by the IOWG and IWG is being **integrated with the HuBMAP Human Reference Atlas (HRA) CCF**

Data comparability

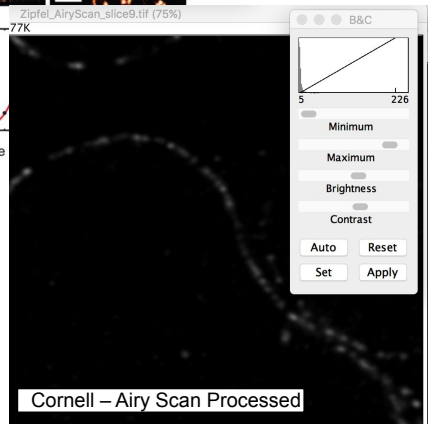




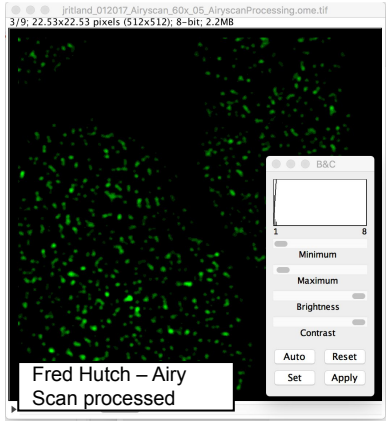
Rob Singer



Jonas Ries

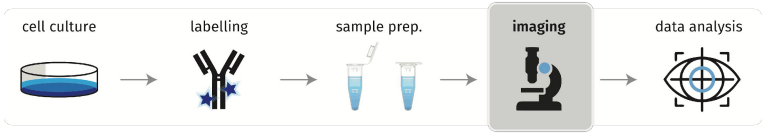


Richard Conroy

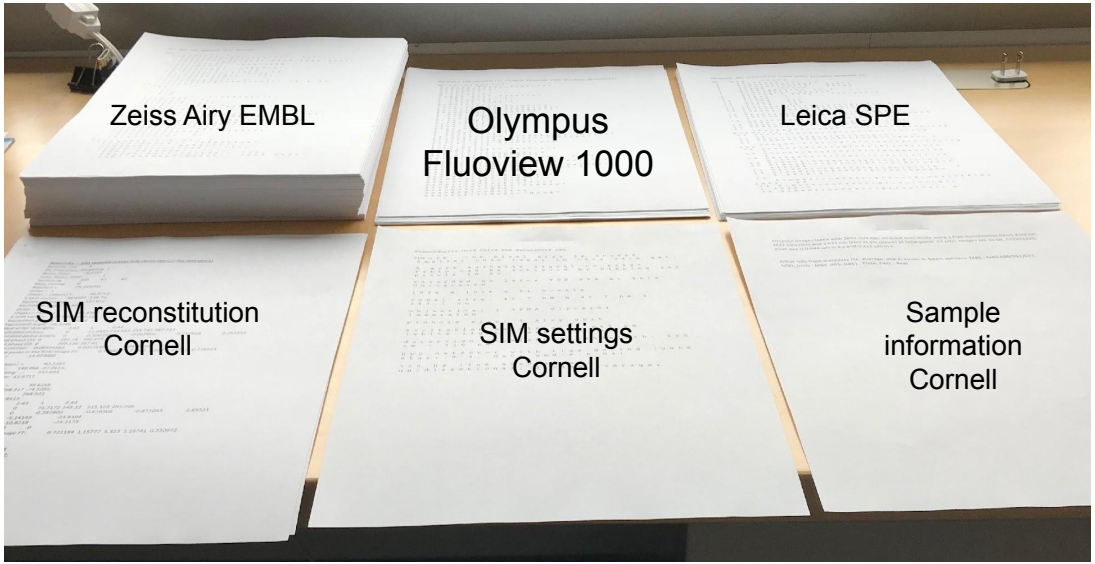


Joan Ritland-Politz

Metadata comparability



Example: Metadata documentation



The 4DN has partnered with a global networks of imaging scientists to build consensus around standardization

Image credit: https://quarep.org

Community standards: Microscopy Metadata Specifications to expand the OME-data model

Comment

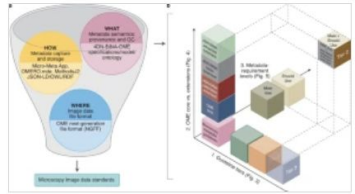
3 Dec 2021

Nature Methods

Towards community-driven metadata standards for light microscopy: tiered specifications extending the OME model

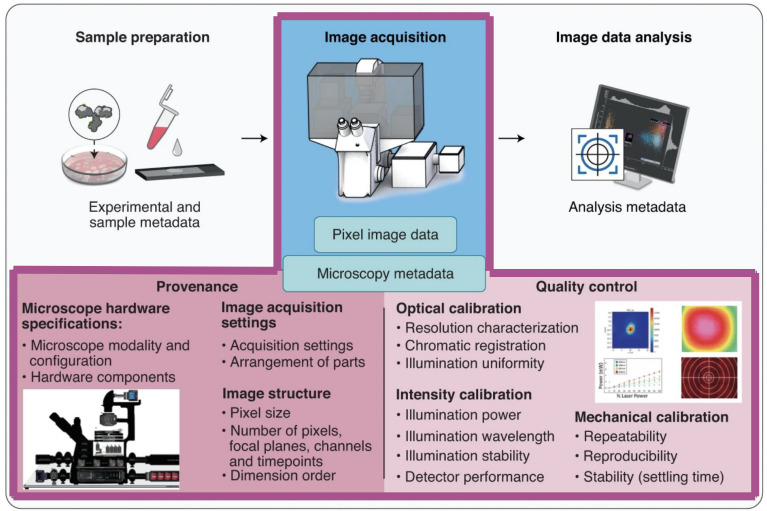
Rigorous record-keeping and quality control are required to ensure the quality, reproducibility and value of imaging data. The 4DN Initiative and BINA here propose light Microscopy Metadata Specifications that extend the OME Data Model, scale with experimental intent and complexity, and make it possible for scientists to create comprehensive records of imaging experiments.

Mathias Hammer, Maximiliaan Huisman ... Caterina Strambio-De-Castillia



4D Nucleome

BioImaging North America



OME

QUAREP-LiMi

Hammer et al. (2021) *Nat Methods*;
https://doi.org/10.1038/s41592-021-01327-9

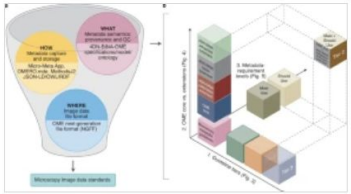
Community standards: Microscopy Metadata Specifications to expand the OME-data model

Comment
3 Dec 2021
Nature Methods

Towards community-driven metadata standards for light microscopy: tiered specifications extending the OME model

Rigorous record-keeping and quality control are required to ensure the quality, reproducibility and value of imaging data. The 4DN Initiative and BINA here propose light Microscopy Metadata Specifications that extend the OME Data Model, scale with experimental intent and complexity, and make it possible for scientists to create comprehensive records of imaging experiments.

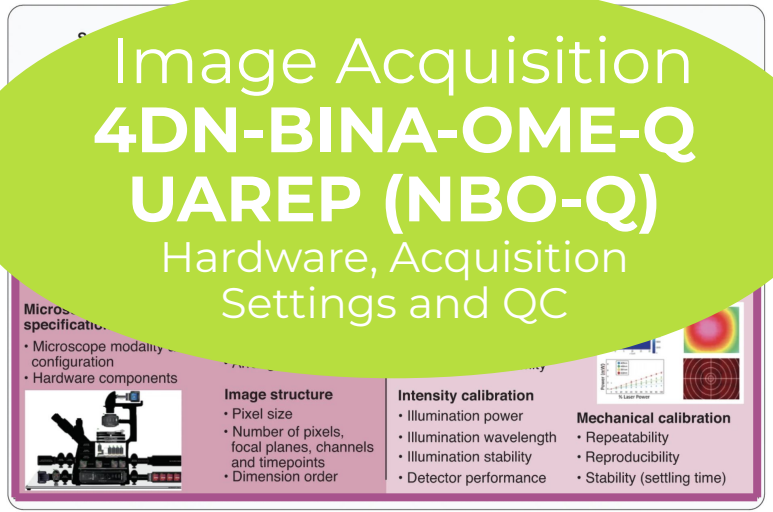
Mathias Hammer, Maximiliaan Huisman ... Caterina Strambio-De-Castillia



4D Nucleome

BioImaging North America

Image Acquisition
4DN-BINA-OME-Q
UAREP (NBO-Q)
Hardware, Acquisition Settings and QC



OME

QUAREP-LiMi

Hammer et al. (2021) *Nat Methods*;
https://doi.org/10.1038/s41592-021-01327-9

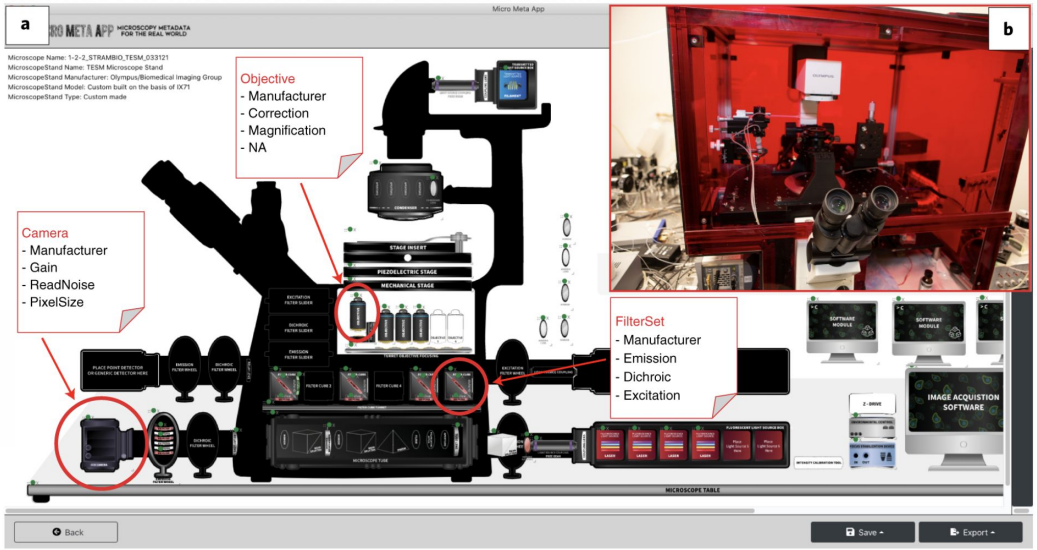
Community standards: Micro-Meta App implements the NBO-Q Microscopy Metadata Model

Brief Communication
Open Access
3 Dec 2021
Nature Methods

Micro-Meta App: an interactive tool for collecting microscopy metadata based on community specifications

Micro-Meta App is an intuitive, highly interoperable, open-source software tool designed to facilitate the extraction and collection of relevant microscopy metadata as specified by recent community guidelines.

Alessandro Rigano, Shannon Ehmsen ... Caterina Strambio-De-Castillia



4D Nucleome

BioImaging North America

OME

QUAREP-LiMi

Partnership with manufacturers to develop community metadata specifications

The making of microscope camera standards

Cameras are a crucial part of microscopes and are also built into many kinds of instruments. To make their output comparable takes standards.

Vivien Marx

Cameras are a crucial part of microscopes and imaging systems. Agreeing on standards to provide defined descriptions for aspects such as gain or readout speed is tricky. Credit: W. Bulgar/Science Photo Library

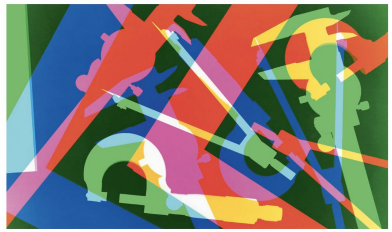
technology feature

Check for updates

Imaging standards to ease reproducibility and the everyday

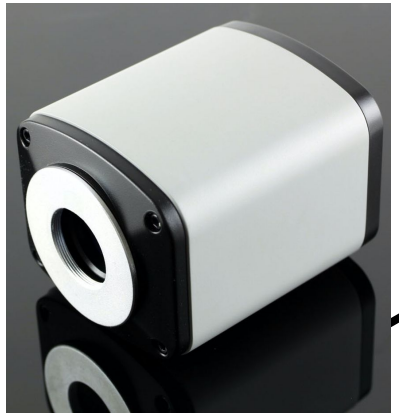
Imaging and microscopy technology advances in leaps and bounds. To address accumulated pain points, academics and companies are making headway on standards.

Vivien Marx



Emerging standards in microscopy are being set up to address many pain points in the field. Credit: TEK Image/Science Photo Library

- **January – August 2022**: 10+ focused feedback sessions to build consensus
- **Completed first parsing of camera hardware specifications and image acquisition settings!**
- **Due Summer 2023:** Revision of **4DN-BINA-OME-QUAREP** Camera Metadata model + Terms definitions



Camera
- Manufacturer: *Xyz*
- Catalog Nr: *0000*
- Mount: *C-mount*
- FrameRate: *20 fps*
- ReadOutRate: *30 MHz*

EVIDENT OLYMPUS

ZEISS

Leica MICROSYSTEMS

Nikon

pco.

Scientifica

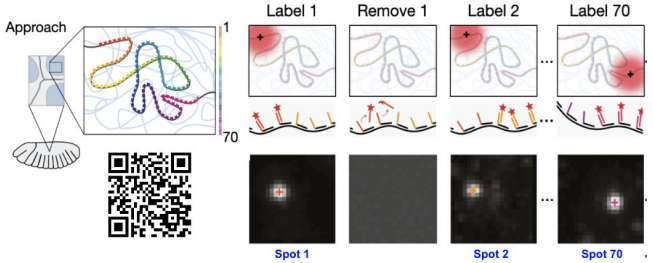
OXFORD INSTRUMENTS ANDOR

TELEDYNE PHOTOMETRICS

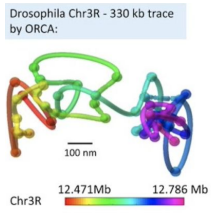
HAMAMATSU PHOTON IS OUR BUSINESS

Why do we need a common format for Chromatin Tracing?

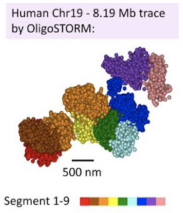
MULTIPLEXED FISH CHROMATIN TRACING



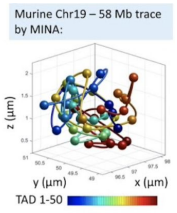
Mateo et al.,
https://doi.org/10.1038/s41586-019-1035-4



(Mateo et al. 2019)

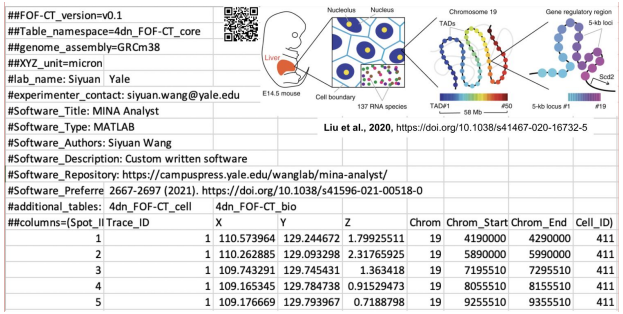


(Nir et al. 2018)



(Takei et al. 2021)

FOF-CT Data and Metadata Exchange Format



##FOF-CT_version=v0.1								
##Table_namespace=4dn_FOF-CT_core								
##genome_assembly=GRCm38								
##XYZ_unit=micron								
#lab_name: Siyuan Yale								
#experimenter_contact: siyuan.wang@yale.edu								
#Software_Title: MINA Analyst								
#Software_Type: MATLAB								
#Software_Authors: Siyuan Wang								
#Software_Description: Custom written software								
#Software_Repository: https://campuspress.yale.edu/wanglab/mina-analyst/								
#Software_Preferre	2667-2697 (2021). https://doi.org/10.1038/s41596-021-00518-0							
#additional_tables:	4dn_FOF-CT_cell	4dn_FOF-CT_bio						
##columns=(Spot_II	Trace_ID	X	Y	Z	Chrom	Chrom_Start	Chrom_End	Cell_ID)
1	1	110.573964	129.244672	1.79925511	19	4190000	4290000	411
2	1	110.262885	129.093298	2.31765925	19	5890000	5990000	411
3	1	109.743291	129.745431	1.363418	19	7195510	7295510	411
4	1	109.165345	129.784738	0.91529473	19	8055510	8155510	411
5	1	109.176669	129.793967	0.7188798	19	9255510	9355510	411

Lacra Bintu
Stanford

Rahi Navelkar
HMS

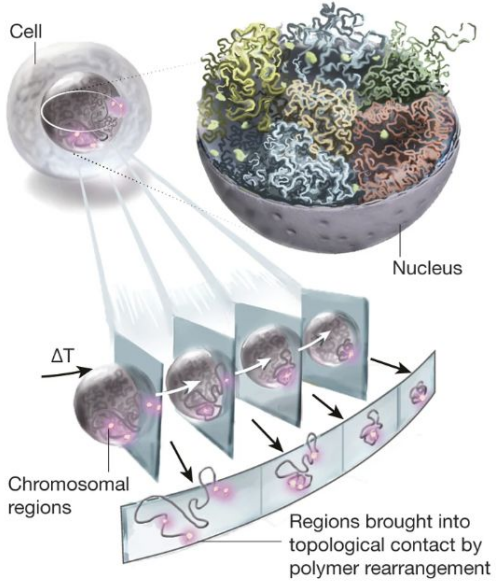
Alistair Boettiger
Stanford

Caterina Strambio DC
UMass Chan

Siyuan Wang
Yale

Sarah Aumfkolk
HMS

PREDICTIVE MODELING/ AND MECHANISMS

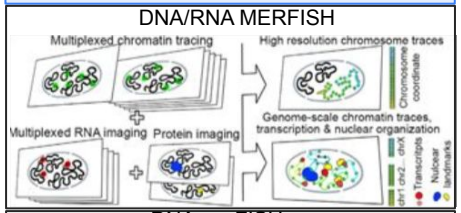


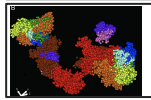
https://doi.org/10.1038/nature23884

Development of a Nuclear Common Coordinate Framework

IWG FISH-Omics Format for Chromatin Tracing

Imaging FISH-Omics (multiplexed FISH) and other imaging

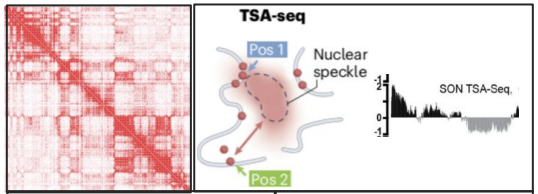




OligoSTORM volumetric tracing

⋮ etc

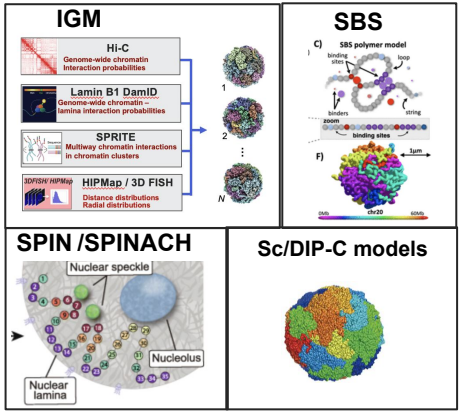
Single cell and **bulk genomics** data
Cytological mapping data



NAD-seq, SPRITE, etc.

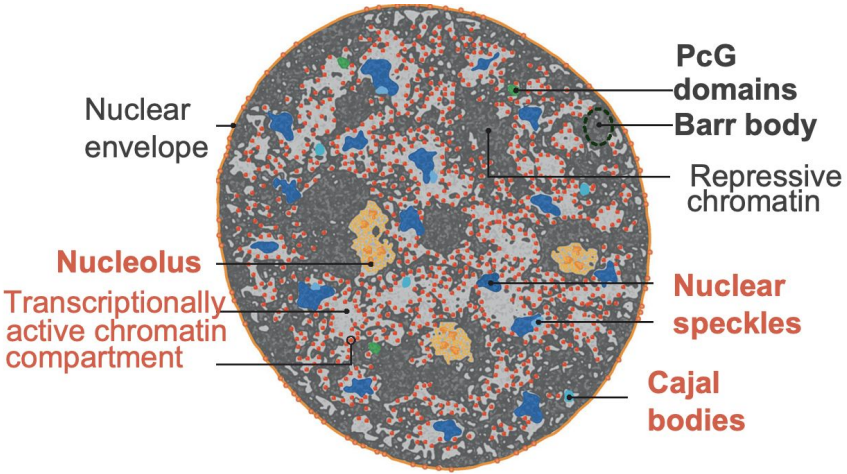
⋮ etc

Computational models and methods



⋮ etc

Why do nuclear coordinates matter?

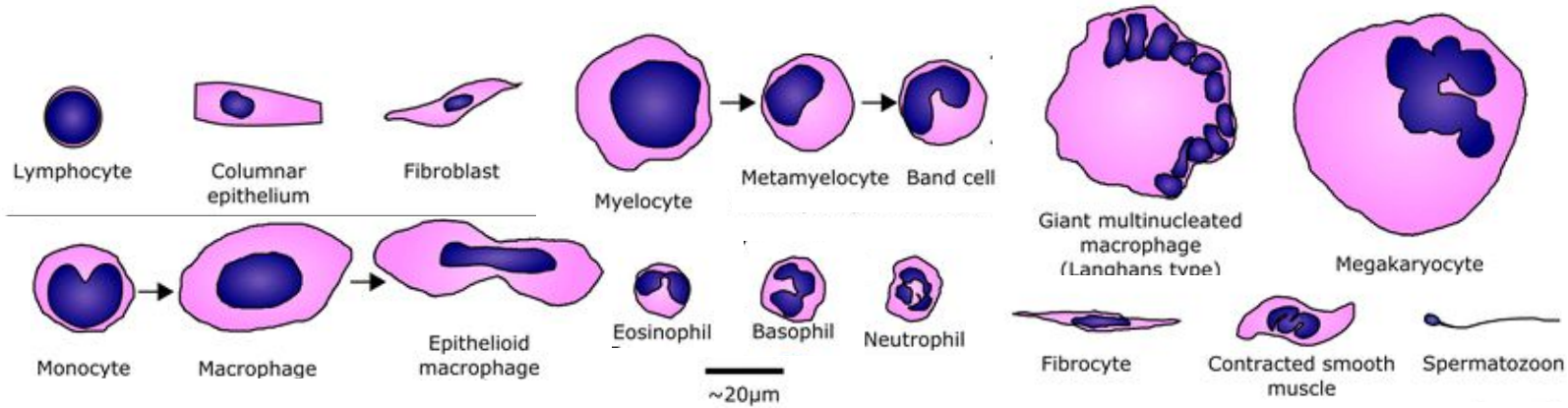


Caudron-Herger et al., Curr.Op.Gen.&Dev. 2012

The nucleus is organized into functional compartments. The spatial proximity to nuclear compartments and nuclear bodies matters.

Why do nuclear coordinates matter?

The nucleus varies widely across cell types, tissues, and differentiation states.
This affects the overall shape, size, and internal nuclear compartment organization.



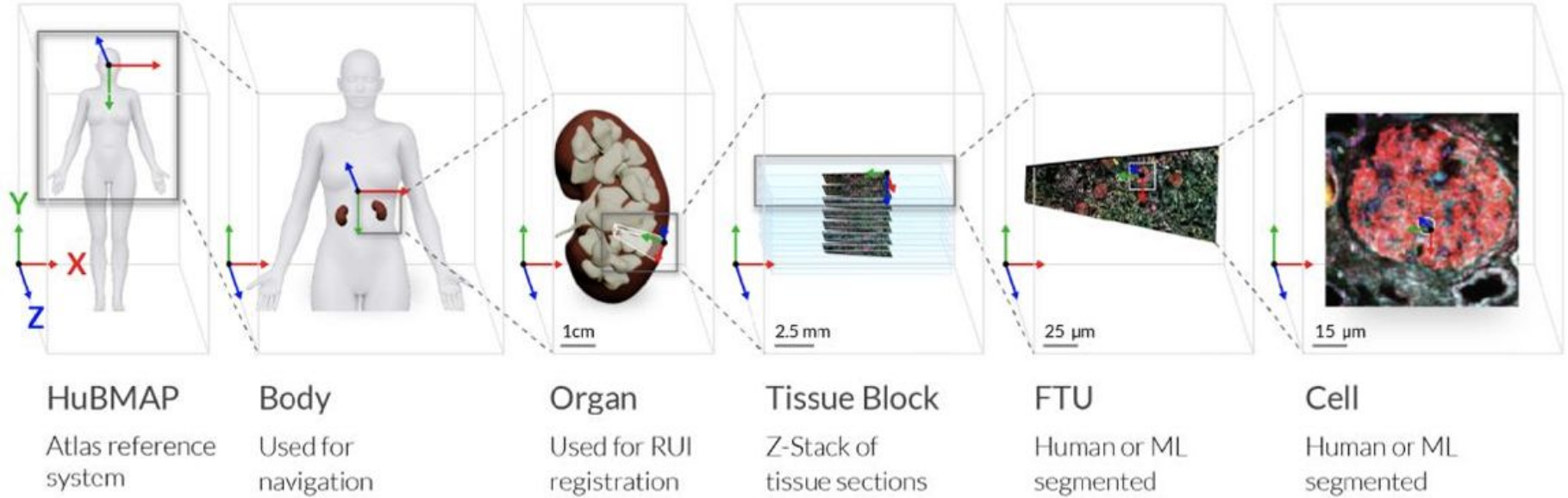
Skinner et al. Chromosoma 2017

1) How do we compare data across different cell types and conditions?
   - We need a **shared and systematic way to describe nuclear topography**
   - We need a **reference system to quantitatively characterize describe the nuclear landscape**

2) What are the **minimum requirements for a nuclear landscape reference system** (e.g., with respect to the location of nuclear compartments and landmarks)?

What is a Common Coordinate Framework?

A standard spatial coordinate system is used to integrate spatial and molecular data across different laboratories, bio-samples, specimens, and conditions and **move past the use of single standardized samples.** The HuBMAP Human Reference Atlas effort developed a recent example.



Borner et al. (2020)
https://doi.org/10.48550/arXiv.2007.14474

Borner et al., (2022)
https://doi.org/10.1038/s42003-022-03644-x

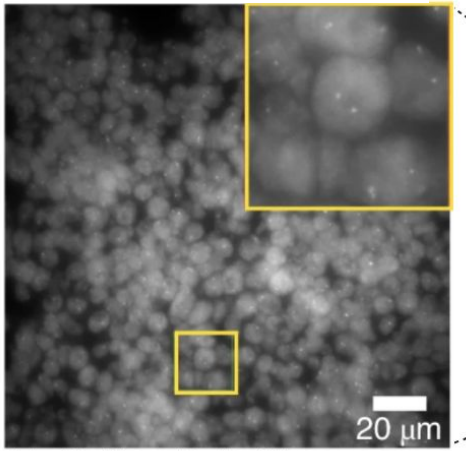
Herr et al., (2023)
https://doi.org/10.1038/s41597-023-01993-8

Borner et al., (2024)
https://doi.org/10.1101/2024.03.27.587041

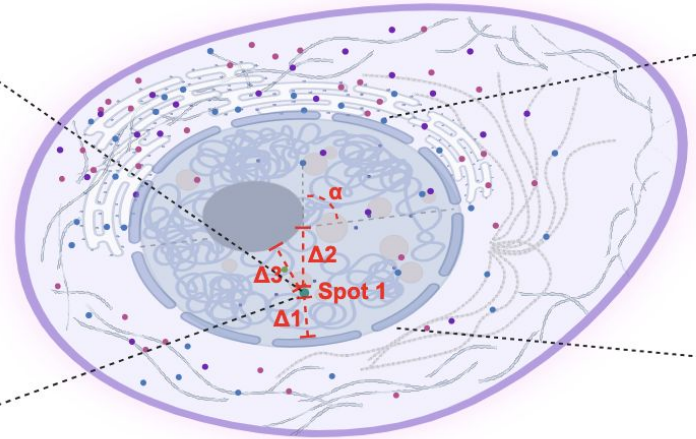
4DN Nuclear Common Coordinate Framework (CCF)

The goal is to have a set of recommendations how to measure and store nuclear information together with locations of loci.

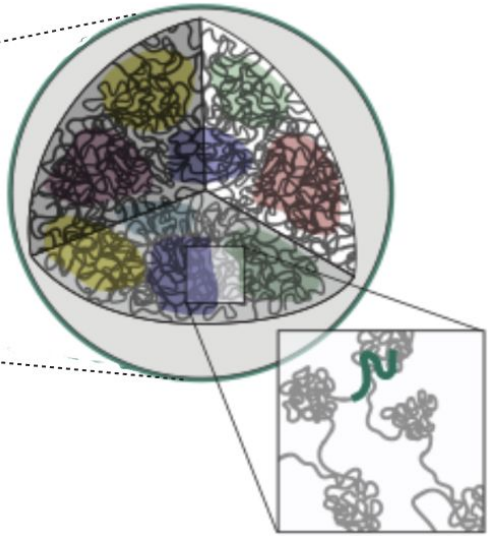
Multiplexed FISH



Liu, M., Yang, B., Hu, M. *et al.* (2021). 10.1038/s41596-021-00518-0



Models from genomics data (Hi-C)



Spot 1 (Δ1, Δ2, Δ3, α)

An **underlying common 'language' for describing and indexing the data in a spatially explicit and semantically consistent way to integrate knowledge** from diverse data types (i.e., multiplexed FISH and 3C methods) and sources and build coherent predictive models of 4D Nucleome structure and function

Recommendations for Nuclear CCF best practices: should be minimally intrusive and widely applicable

- A **nuclear boundary marker is necessary (but not sufficient)**:
  - Examples: DAPI, Nucleoporin, Lamina (LaminA/C, LaminB)
- **Other markers are required for triangulation and breaking symmetry**:
  - Examples: Nucleoli, Nuclear speckles, Histone epigenetic markers, RNA polymerase, mRNA transcripts
- **Key requirements**:
  - Easy-to-use,
  - Different options for different experimental designs
  - Consider methods that do not require the use of fluorescence markers and use Machine Learning to predict nuclear markers localization:
    - Brightfield images
    - Fluorescence background (from DNA/RNA FISH-Omics probes)
    - Autofluorescence

Recommendations for Nuclear CCF best practices: should be minimally intrusive and widely applicable

- A **nuclear boundary marker is necessary (but not sufficient)**:
  - Examples: DAPI, Nucleoporin, Lamina (LaminA/C, LaminB)
- **Other markers are ... reaking symmetry**:
  - Examples ... markers, RNA polym...
- **Key requ...**
  - Easy-t...
  - Different o...
  - Consider method... fluorescence markers and use Machine Learning to predict nuclear markers localization:
    - Brightfield images
    - Fluorescence background (from DNA/RNA FISH-Omics probes)
    - Autofluorescence

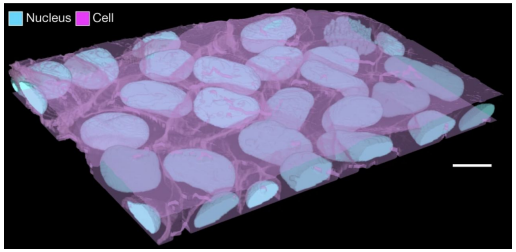
Different methods have to be tested to develop best practices for Nuclear CCF

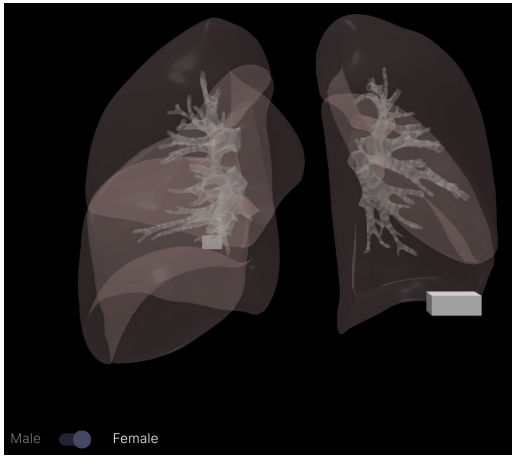
Ongoing: acquisition of benchmarking datasets @ Center for Epigenomics, UCSD

Quan Zhu
UCSD

Bogdan Bintu
UCSD

- **Optimization**
  - Tested different nuclear markers
  - Tested Chromatin Tracing probe library
- **Model System**
  - WTC-11 hiPSC
  - Human adult lung sections (from HuBMAP)
- **Ongoing experiments**:
  - RNA and DNA MERFISH
  - Multiple markers: DAPI, Lamina, NPC, Nucleoli, RNA Polymerase II, SC35, Histone epigenetic markers, PCNA
  - Brightfield image
- **Questions**
  - What combination of markers are necessary and sufficient?
  - What protocol should be used for
  - Can we use machine learning approaches to predict the position of nuclear landmarks in stain-free transmitted light, background fluorescence, or autofluorescence images?

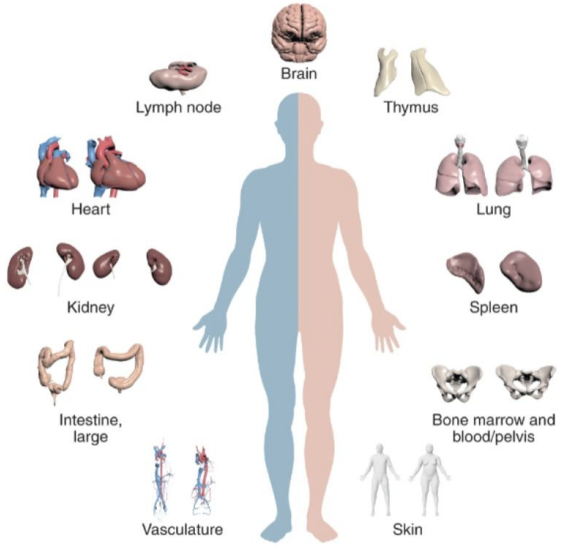




Next Steps: Integration of Nuclear CCF with the HuBMAP Human Reference Atlas

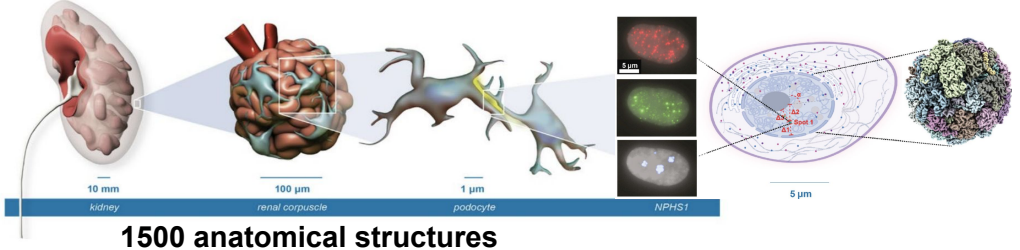
Questions:

- What is the role of the chromatin organization and gene expression regulation in determining the organization of microenvironments in healthy and diseased Functional Tissue Units (FTUs)
- Determine how the molecular and cellular functions for a given cell type compare across organs (for example, genes essential in water transport across the kidneys, intestines, and lungs)



Common Coordinate Framework for Spatial and Ontological Integration

Integration of 4DN FISH Omics



1500 anatomical structures

267	2358	3080	31	18
Donors	Samples	Datasets	Organs	Collections

Aviv Regev, *Genentech, Inc. (Human Cell Atlas)*

Claire Walsh, *University College London (Human Organ Atlas)*

The Human Organ Atlas (HOA)

Claire Walsh
c.walsh.11@ucl.ac.uk
@hip_ct

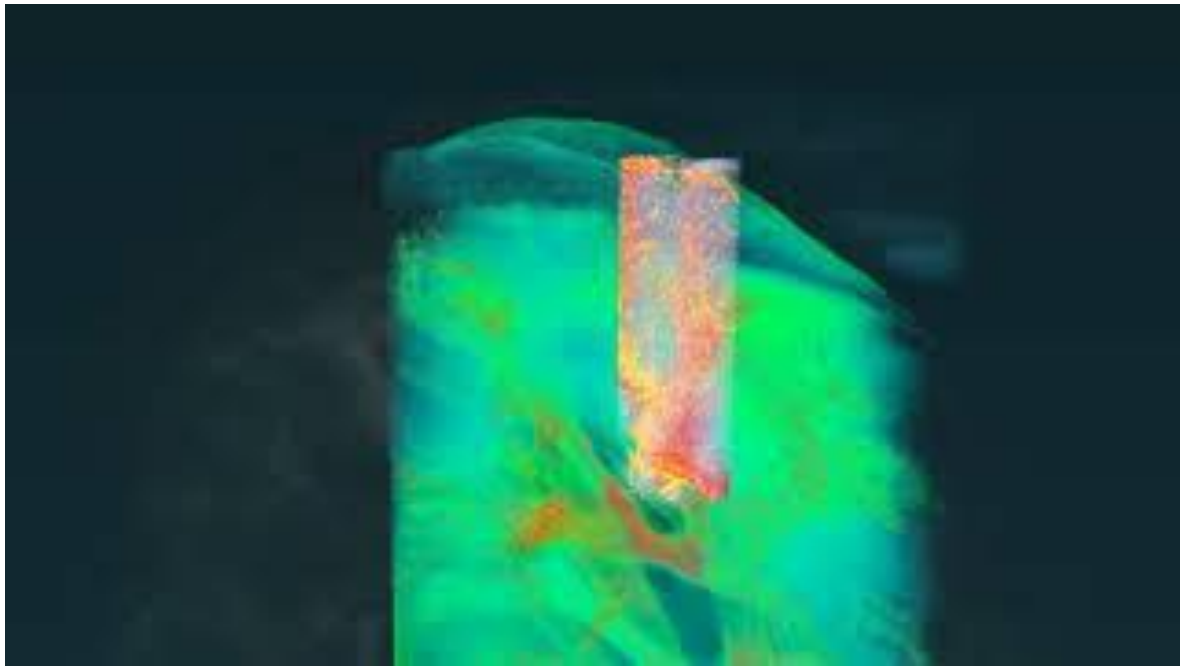
UCL

ESRF

CHAN ZUCKERBERG INITIATIVE

Hierarchical Phase-contrast tomography

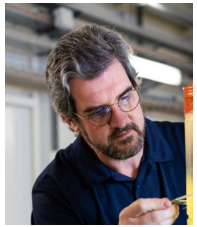
- **Intact human placenta and organs can be scanned ex vivo at 25-7.8µm/voxel**
- **Regions of interest anywhere can then be scanned at higher resolution (down to 0.7 µm/voxel)**
- **We can reach single cell resolution in an intact human organs**



Video Credit: Paul Tafforeau data credit UCL lead ESRF beamtime md1252

Walsh, C. L., et al. *Nature methods* (2021)

HiP-CT at the ESRF



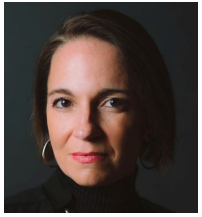
Paul Tafforeau
Beamline responsible
BM18

The Human Organ Atlas HUB (HOAHub)

Peter Lee
Executive Co-Chair
Co-PI
UCL

Max Ackermann
Executive Co-Chair
Co-PI
Aachen Medical School

Claire Walsh
Director
Co-PI
UCL



Anastasia Yendiki
Member
Co-PI
Harvard Medical School

Danny Jonigk
Member
Co-PI
Aachen Medical School

Bernadette de Bakker
Member
Co-PI
Amsterdam UMC

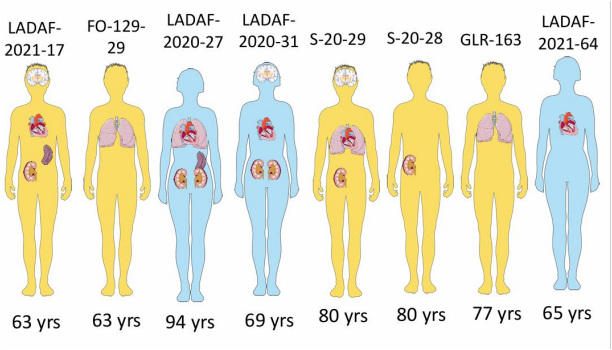
Stijn Verleden
Member
Co-PI
Antwerp University

Paul Tafforeau
Beamline responsible
BM18

Allexandre Bellier
Member
Co-PI
LADAF

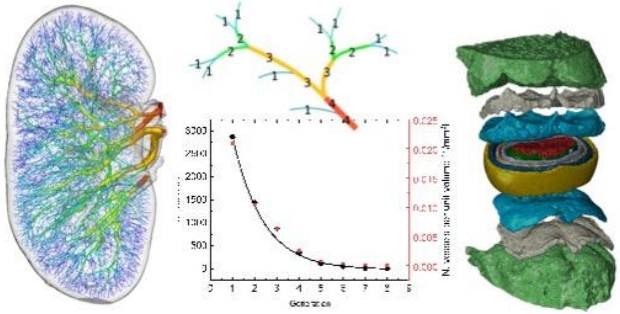
https://mecheng.ucl.ac.uk/HOAHub/

1. Human Organ Atlas

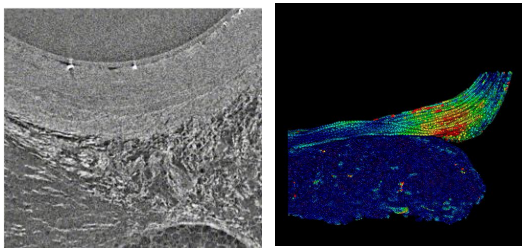


human-organ-atlas.esrf.eu

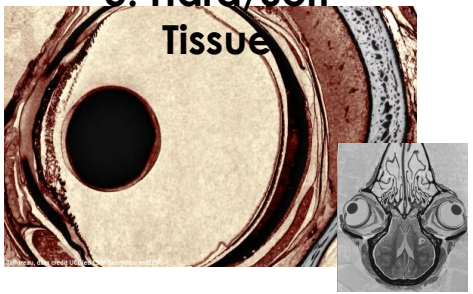
3. Quantifying & Modelling



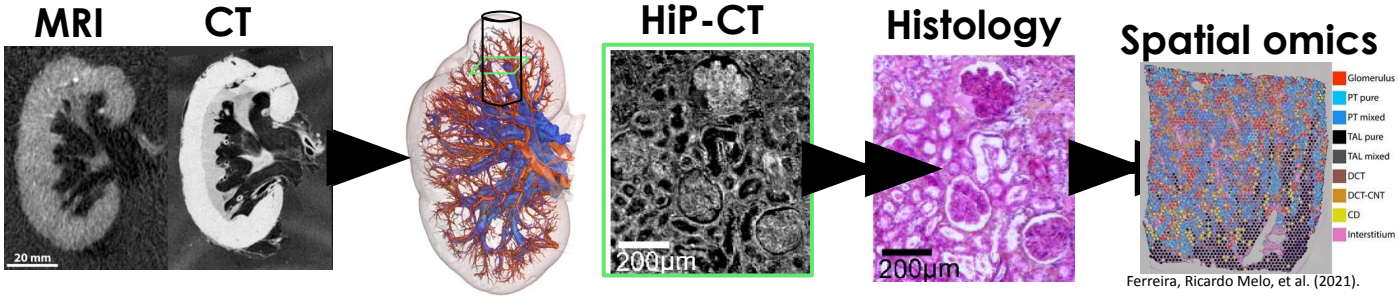
4. Dynamics



5. Hard/Soft Tissue



2. Correlation

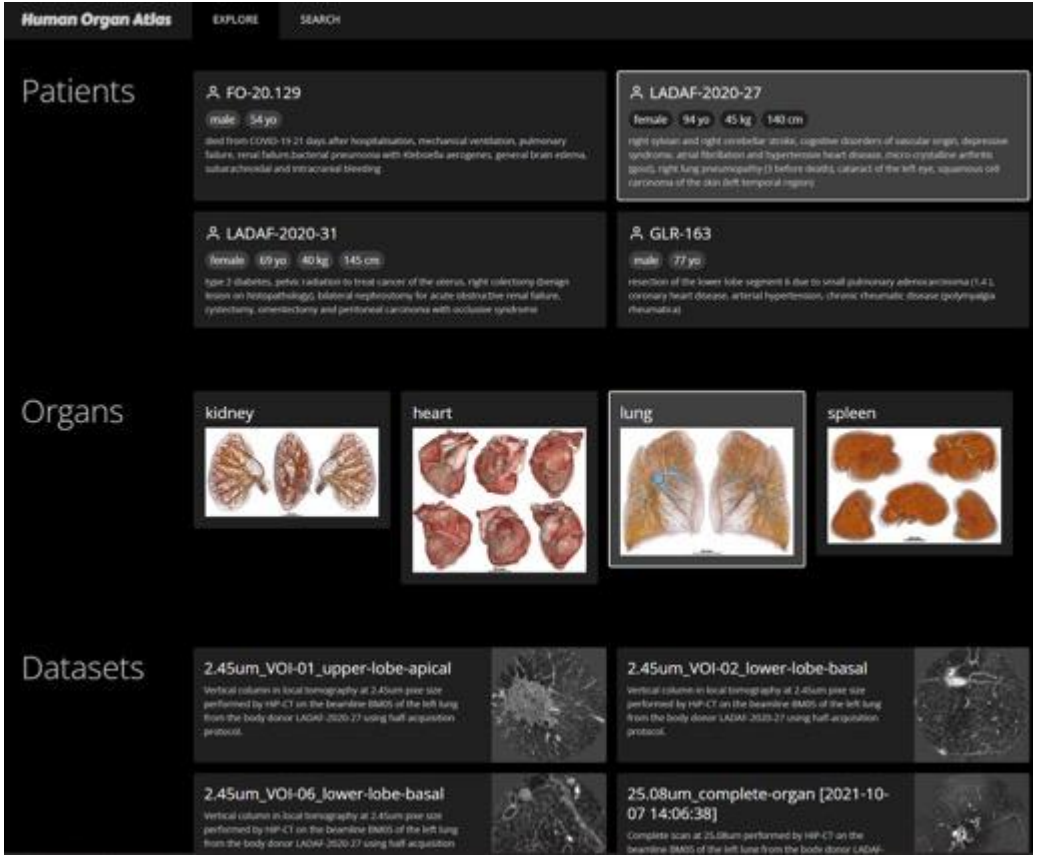


Ferreira, Ricardo Melo, et al. (2021).

The Human Organ Atlas

human-organ-atlas.esrf.fr

Public database with complete organs imaged by HiP-CT in health and disease



ESRF The European Synchrotron

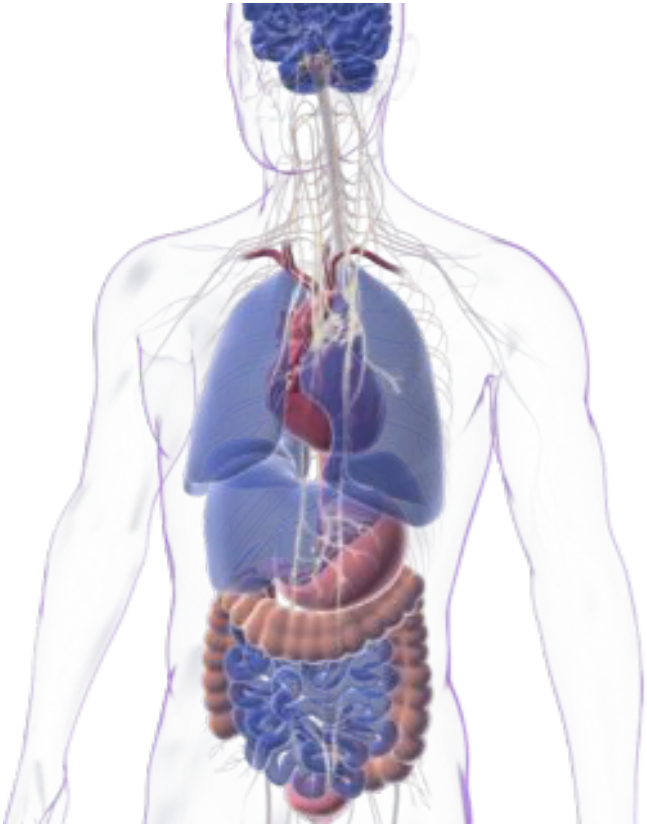
UCL

HOA Hub

école de Chirurgie LADAF Laboratoire d'Anatomie Des Alpes Françaises

Maryann Martone, ***University of California, San Diego***

SPARC: Bridging the body and brain



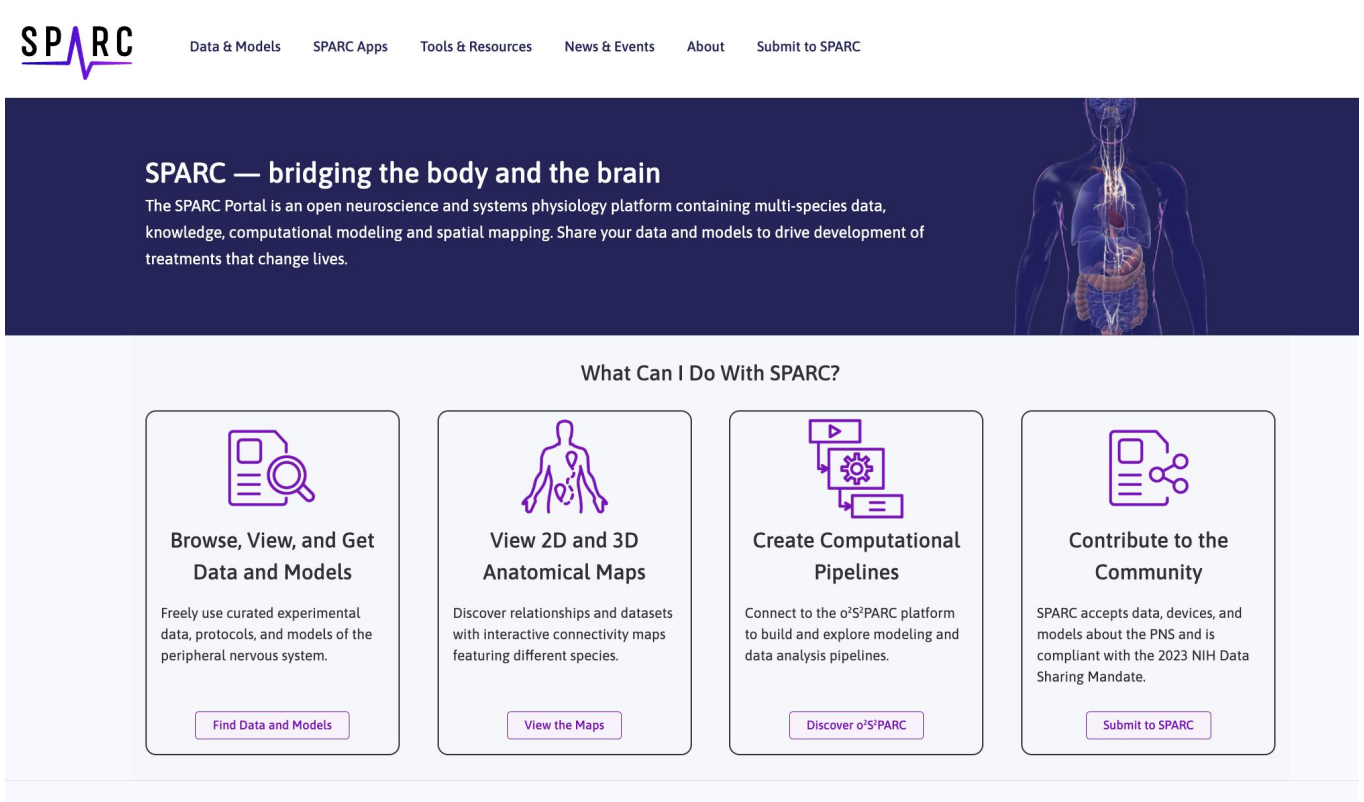
Opportunity: Neuromodulation of end-organ function holds promise in treating many diseases/conditions.

Challenge: Mechanisms of action remain poorly understood. Many neuromodulation trials have failed to reach clinical endpoints.

SPARC program goals:

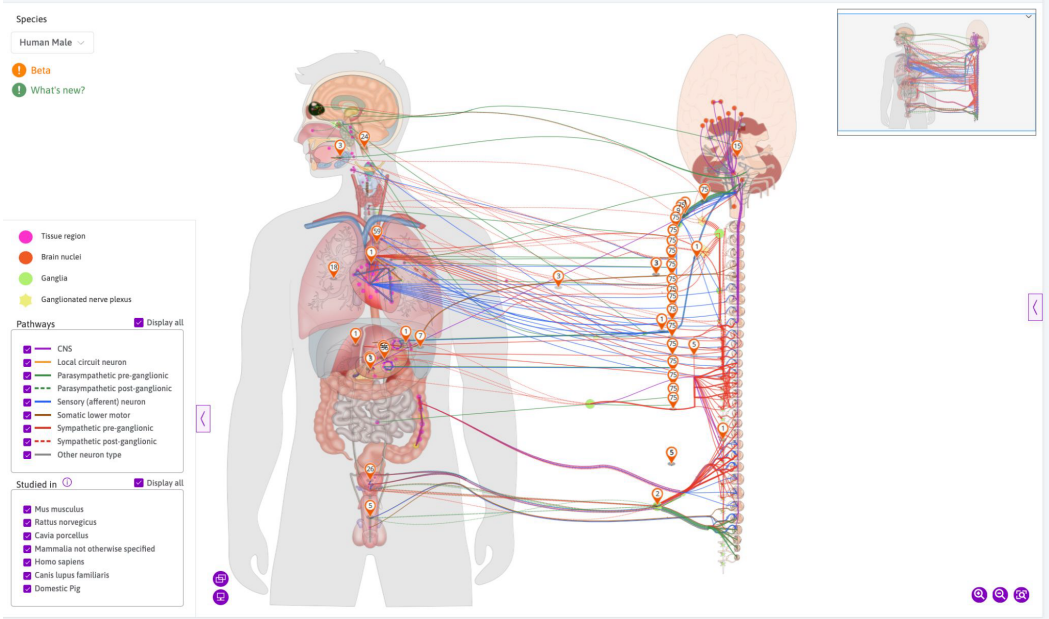
- Deliver detailed, integrated functional and anatomical neural circuit maps for organs and technologies to improve neuromodulation studies
- Provide the scientific foundation necessary to translate advanced and more effective neuromodulation protocols into clinical

Start exploring at SPARC.science



SPARC Maps and Connectivity KB

- Explore SPARC's interactive 2D and 3D maps of the autonomic nervous system
- These maps are drawn automatically from a knowledge base that contains detailed information about how nerves connect different parts of the body



https://sparc.science/apps/maps?type=ac

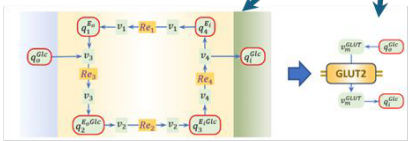
Peter Hunter, *Bioengineering Institute New Zealand (SPARC)*

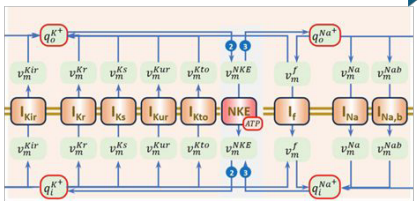
Physics-based multiscale modelling

1. BG protein templates

(full) (reduced)

- Metabolic enzymes
- SLC transporters
- ATPase pumps
- ABC transporters
- Ion channels
- Contractile proteins
- Receptors: GPCR, TRK
- Signal transduction
- Immune cytokines
- Cell cycle
- Gene transcription & translation



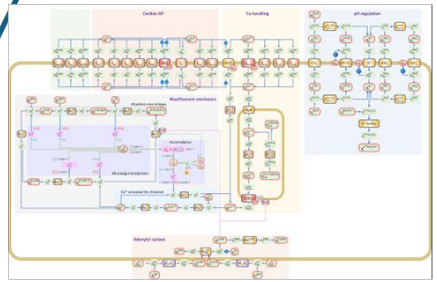


2. FCUs

- Metabolism
- Action potentials
- Calcium handling
- pH regulation
- Myofilament mechs
- Receptors & signaling
- Glucose transport
- Na/K regulation
- Cell cycle
- Cell adhesion
- Gene regulation
- ...

3. Cells

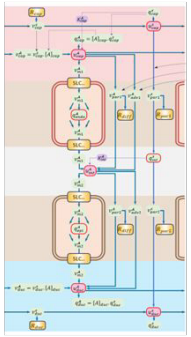
- Endothelial cell
- Epithelial cell
- Cardiomyocyte
- Muscle fiber
- Fibroblast
- Smooth muscle cell
- Osteocyte
- Enterochromaffin cell
- Hepatocyte
- Goblet cell
- Red blood cell
- ...



4. FTUs

(including 3D structure)

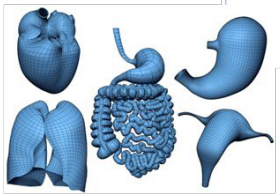
- Glomerulus, ... → nephron
- Lung alveoli → acinus
- Myocardial sheet → block
- Liver lobule → lobules
- Bone osteon → osteon
- Fibre, etc → fascicle
- ... → ...



5. Organs

=surrogate

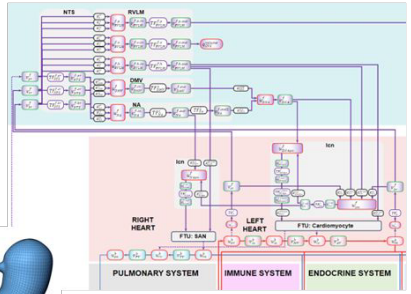
- nephron → Kidney
- acinus → Lung
- block → Heart
- lobules → Liver
- osteon → Bone
- fascicle → Muscle
- ... → ...



6. Systems

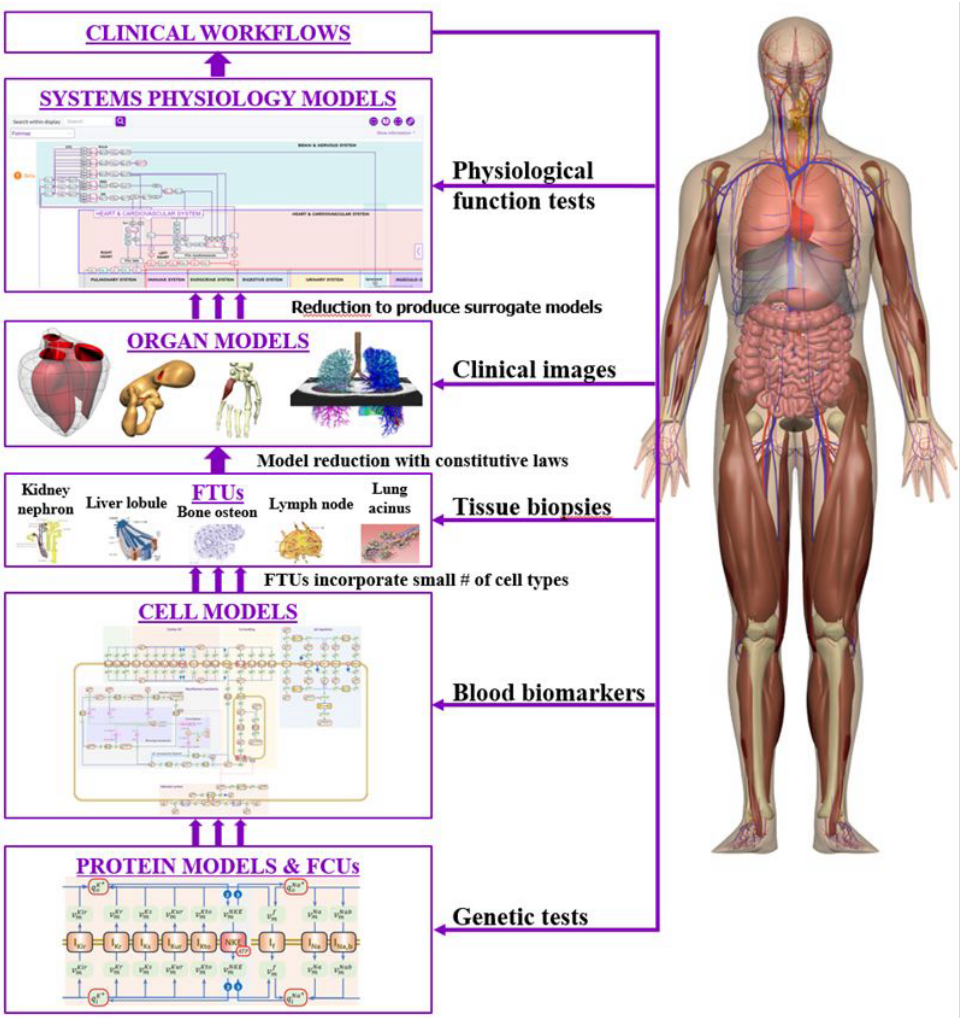
(3D anatomical model of the body)

- Regulation of fluid volume
- Regulation of arterial blood pressure
- Regulation of pH
- Regulation of electrolyte balance
- Regulation of glucose
- Regulation of O_2 and CO_2
- Regulation of calcium
- Regulation of phosphate
- Regulation of body temperature
- ...



Finite element models that are subsequently reduced (with AI methods) to surrogate models.

Hunter et al The Physiome Project and Digital Twins 2024 *IEEE Reviews in Biomedical Engineering*



Constrained by observed behaviour

Needs AI based on physics

Model parameters

Constrained by physics and genetics

Gary Bader, *University of Toronto, Canada (CIFAR co-director)*

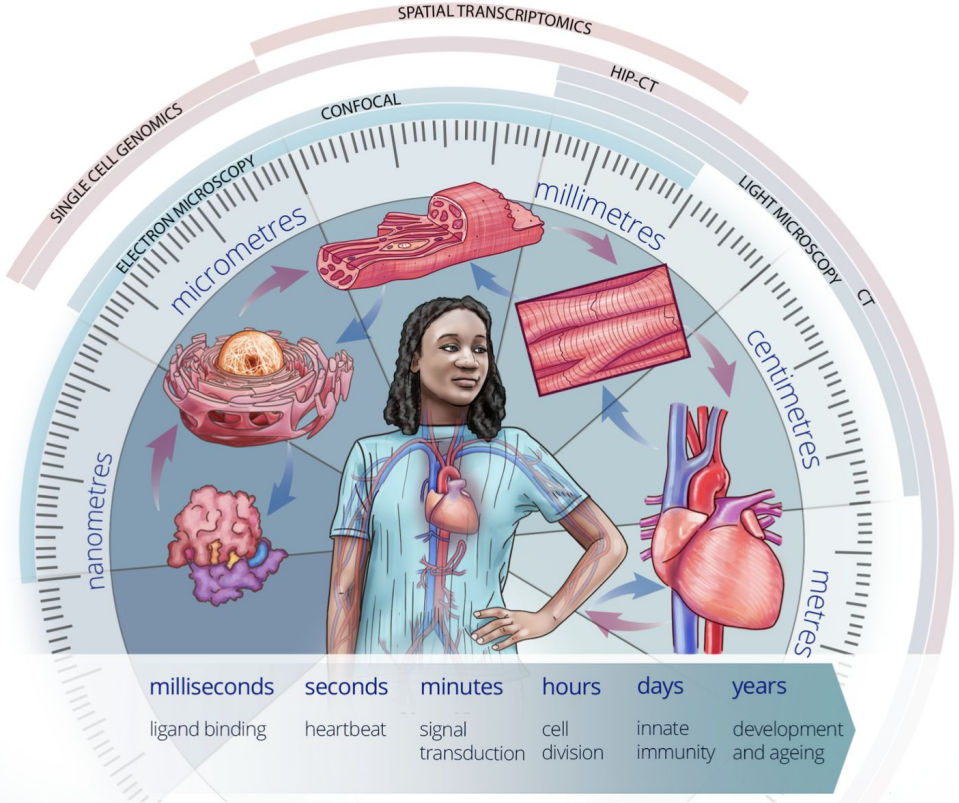
The CIFAR Multiscale Human Program

Gary Bader, Katy Börner, Sarah Teichmann, Alain Chédotal, Barbara Engelhardt, Ali Ertürk, Ferdia Gallagher, Sidhartha Goyal, Muzlifah Haniffa, Peter Lee, Ed Lein, Dana Pe'er, Aviv Regev, Nozomu Yachie, Peter Zandstra, Mei Zhen
Guests: Fabian Theis, Maria Brbić, Stefan Bauer et al.

https://cifar.ca/research-programs/cifar-macmillan-multiscale-human

Major goal

To understand how the human body works across scales, from molecules to organs to the whole body to revolutionise our understanding, treatment and prediction of major diseases



Silos across scales: Molecular biology vs. physiology

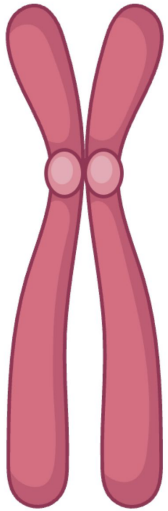
Focus on understanding mechanism at separate scales and rarely integrate

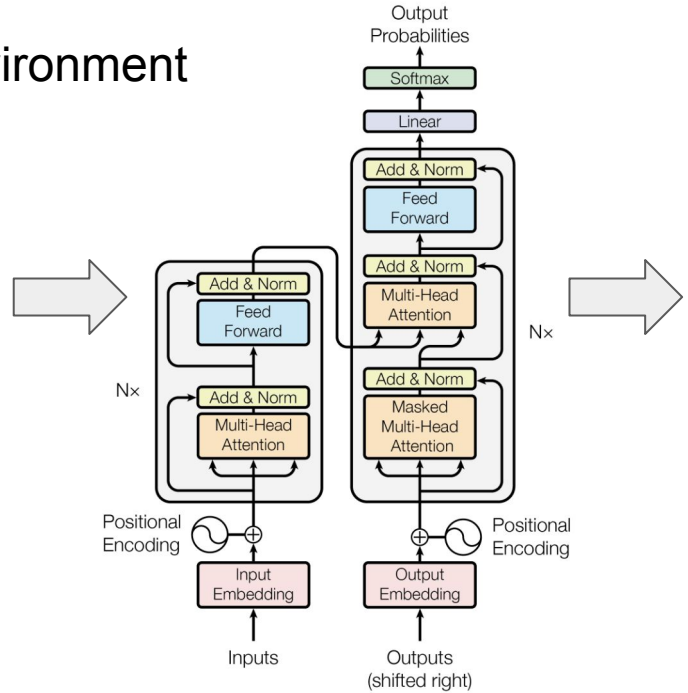
Single cell and spatial genomics creates a bridge via multicellular tissues

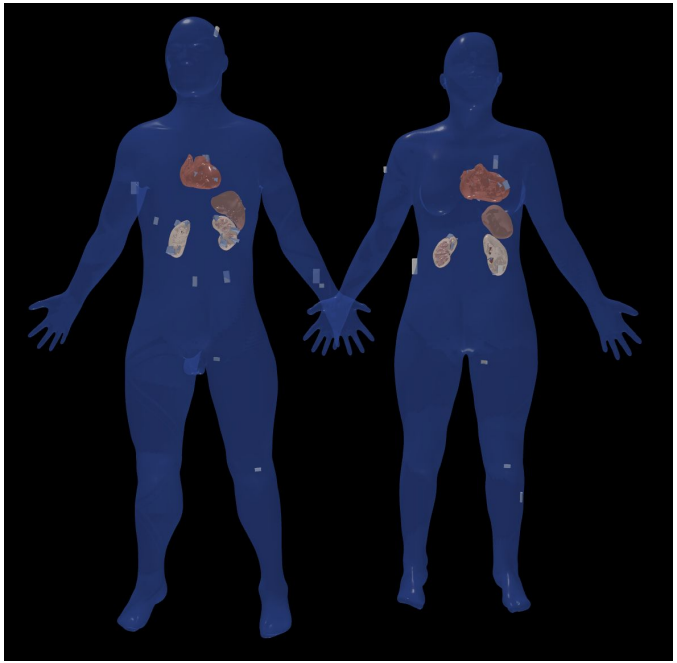
Can we develop a unified field that considers how the whole body works across scales?

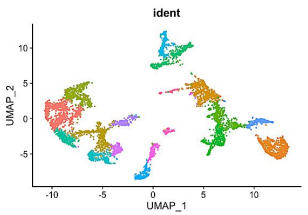
The genome as the ultimate generative model

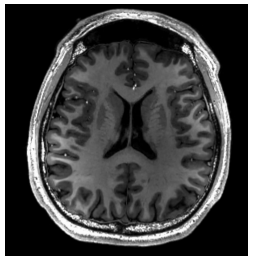
+environment











Correlated

Biorender.com; Attention Is All You Need, 2017; wikipedia
https://portal.hubmapconsortium.org/ccf-eui, 10x Genomics; DALL-E

Is machine learning a good approach for understanding the body?

Evolution: copy and mutate (+ memorize)

Result: redundancy, variation

Perfect for ML: redundancy helps with pattern recognition, variation helps link data measurements (e.g. regression)

Will require mechanistic insight, multiscale thinking

(Rare events will require mechanism-based interpretation)

Genetics can link scales

Genome to phenotype relationship works across scales

Useful to link scales: SNP, protein, complex, pathway, cell, tissue, organ, body

Large biobanks help us map biology

However, mechanistic insight is challenging to get

May need to combine genetics and mechanistic modeling

Mapping the human body (structure)

HUMAN CELL ATLAS

HuBMAP
Human BioMolecular Atlas Program

3D Multiscale Biomolecular Human Reference Atlas

Generating the virtual human (function)

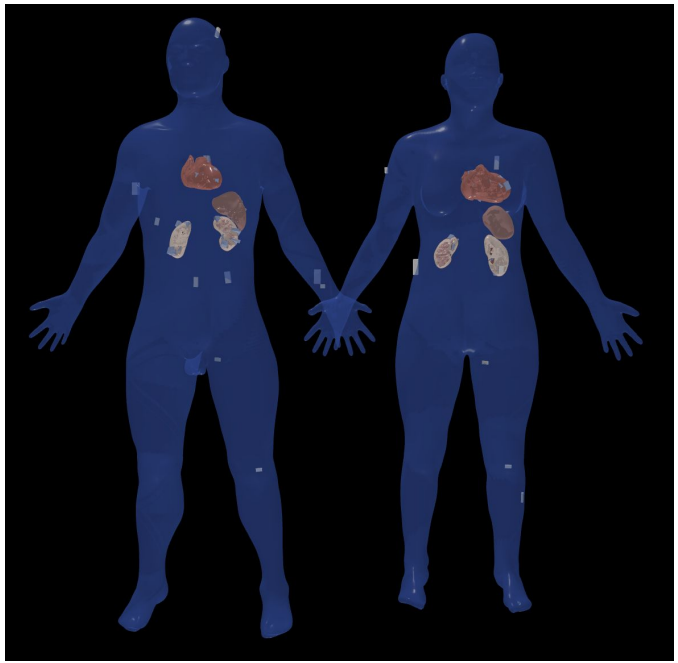
Generative model

Mechanistic

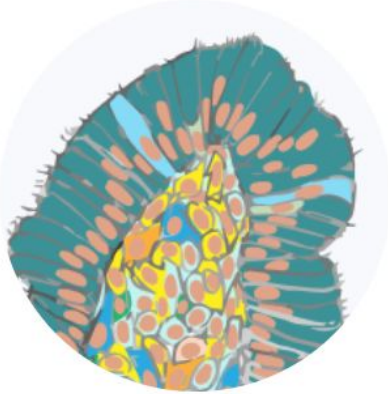
Predictive (e.g., response to perturbation)

Multiscale - how are scales connected?

Medical applications (e.g., digital twin)



https://portal.hubmapconsortium.org/ccf-eui

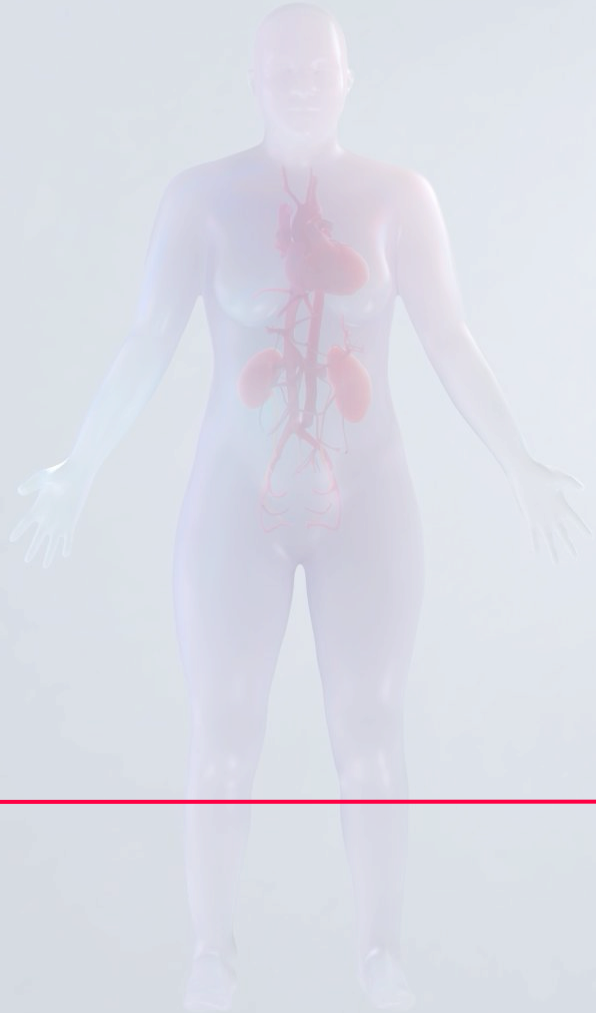


5AM

10AM in London (GMT), 7PM in Tokyo (GMT+9)

VIDEOS: Human Atlas Insights

- Mapping the Multiscale Human by Gary Bader, *University of Toronto, Canada (CIFAR co-director)*



Q&A

https://humanatlas.io/events/2024-24h

Questions

How do we define a Multiscale Human?

How do we map a Multiscale Human?

How do we model a Multiscale Human?

How can LLMs or RAGs be used to advance science and clinical practice?

Thank you