OSC (Common Fund)

Concept Clearance: Two additional initiatives to support the Production Phase of HuBMAP

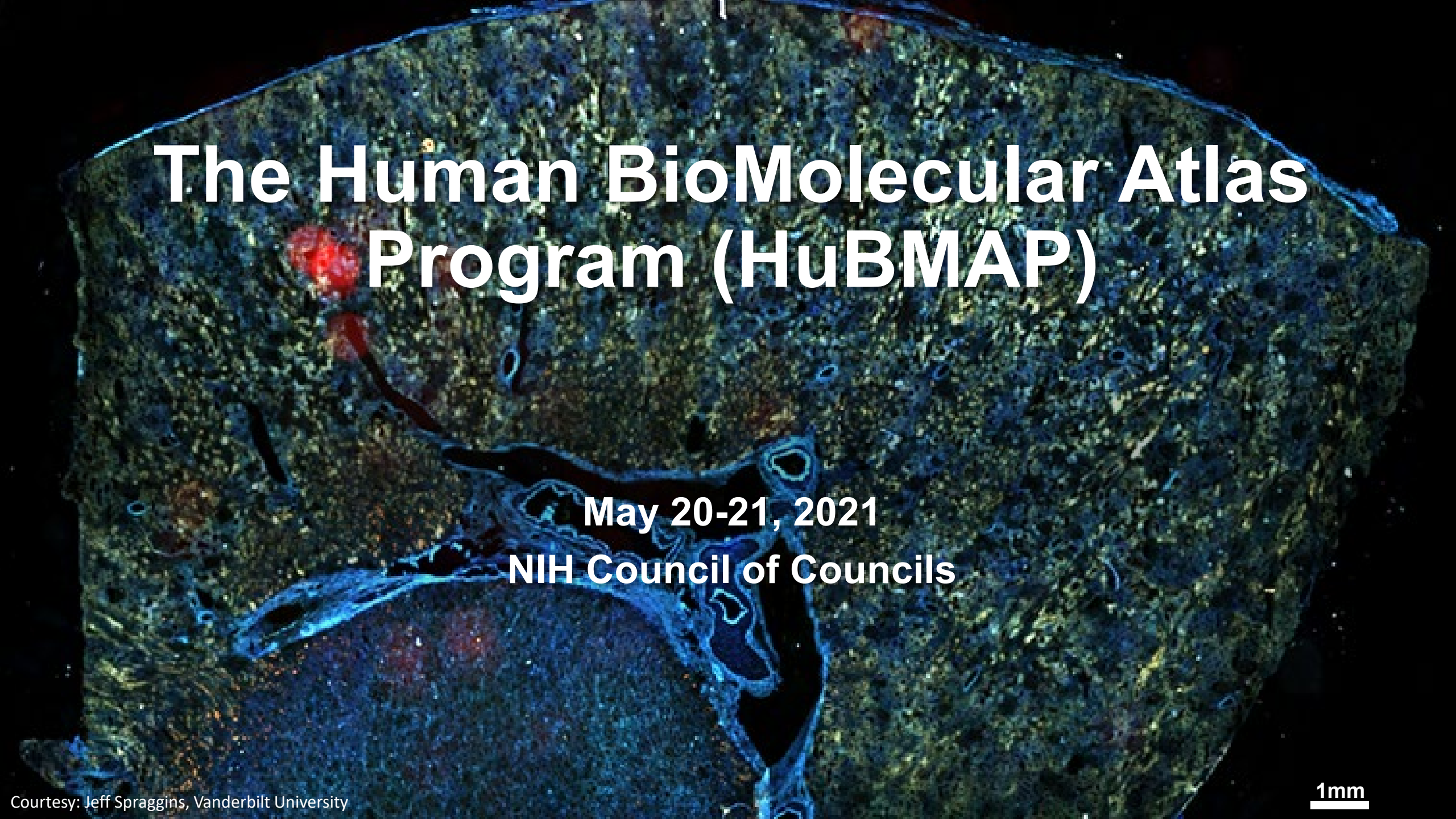**PROGRAM TITLE:** Human BioMolecular Atlas Program (HuBMAP)

Program Objective: To catalyze the development of a framework for mapping the human body at single cell resolution.

Council concurrence to issue two additional Production Phase Initiatives:
1. Reissue the HuBMAP Integration, Visualization, and Engagement (HIVE) RFA [$10.5M / year; 5-7 awards]
2. Issue a Data Mining and Demonstration Projects RFA [$3M / year; 5-9 awards]

Period: 4 years

Council Action: Vote on support for both RFAs.

The Human BioMolecular Atlas Program (HuBMAP)

May 20-21, 2021

NIH Council of Councils

1mm

"While we broadly know how cells are organized in most tissues, a comprehensive understanding of the cellular and molecular states and interactive networks resident in the tissues and organs [...] is lacking."

"The specific three-dimensional organization of different cell types, together with the effects of cell–cell and cell–matrix interactions in their natural milieu, have a profound impact on normal function, natural ageing, tissue remodelling, and disease progression in different tissues and organs."
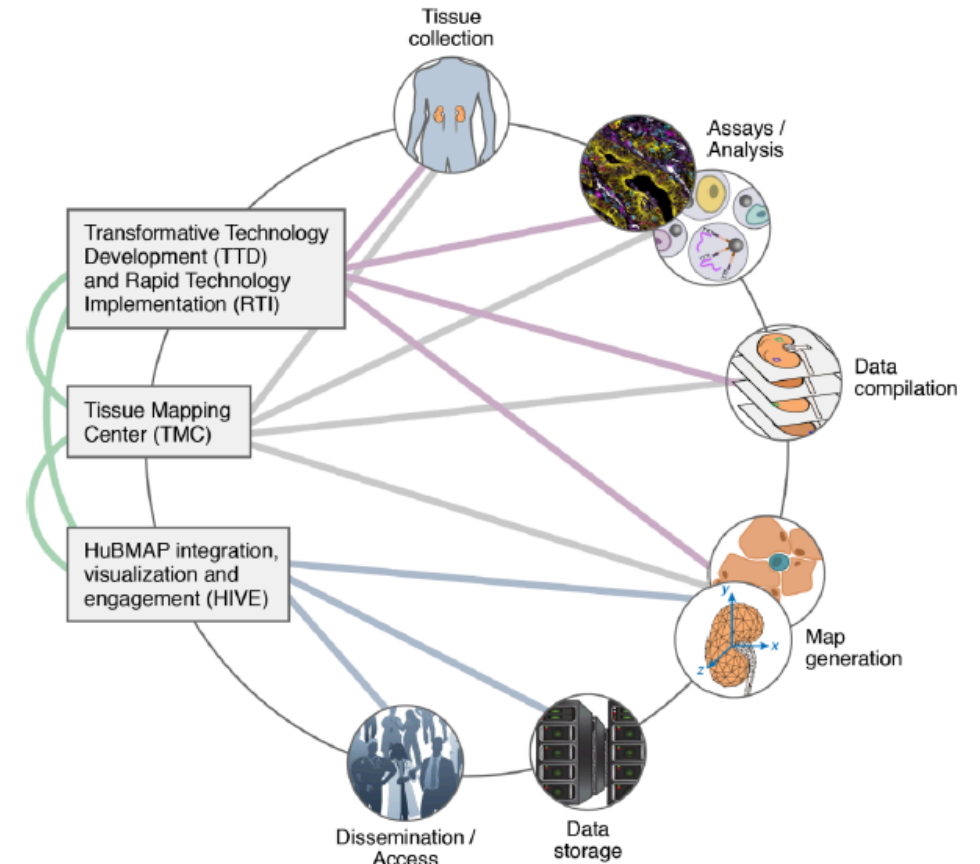
"HuBMAP is an important part of the international mission [and] aspires to help to build a foundation by generating a high-resolution atlas of key organs [...] as well as acting as a key resource for new contributions in the growing fields of tissue biology and cellular ecosystems."

Snyder et al. Nature. 2019; 574(7777)

Goal: Catalyze the development of a framework for mapping the human body at single cell resolution

By:

1. Accelerating development of next generation tools and techniques

2. Generating foundational 3D human tissue maps

3. Establishing an open data platform

4. Collaborating with the research community

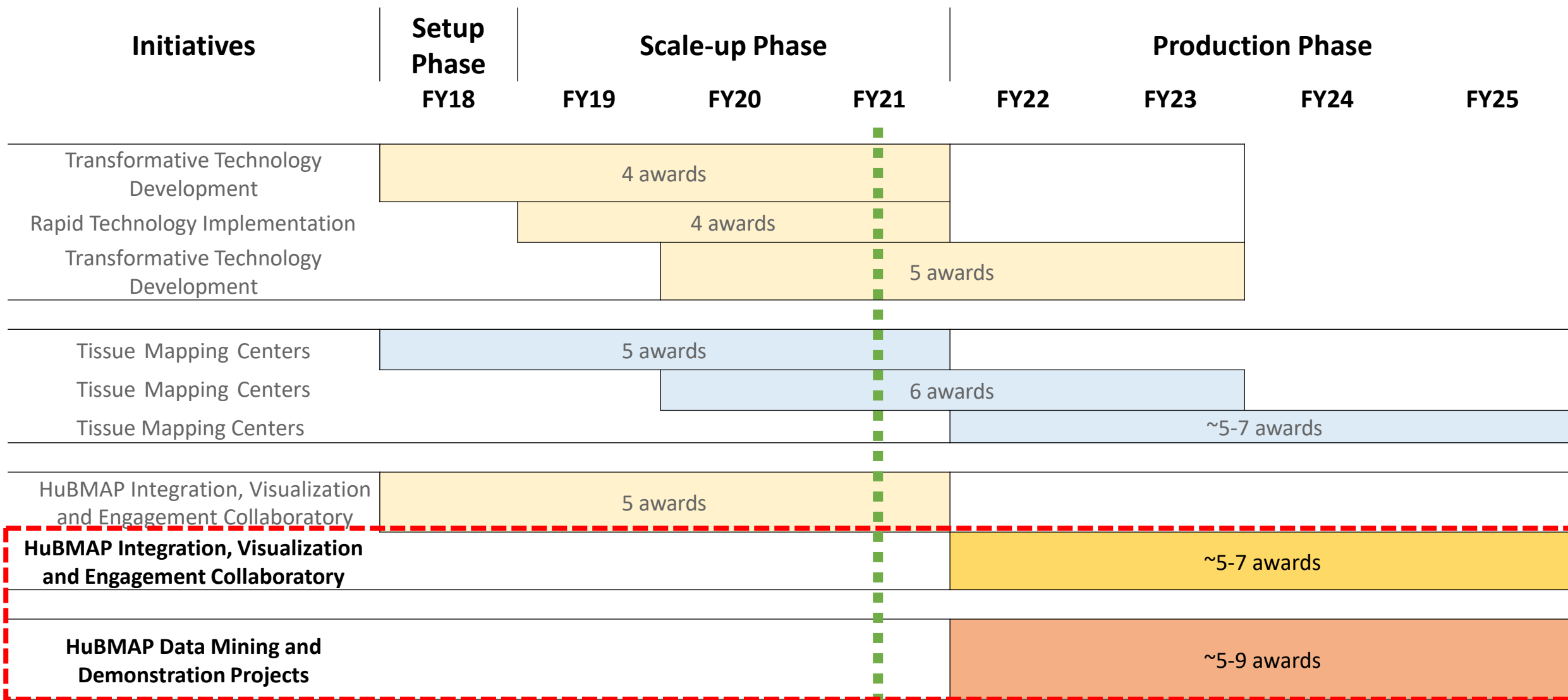5. Supporting pilot projects that demonstrate value of HuBMAP resources

Snyder et al. Nature. 2019; 574(7777)

Program Timeline

Initiatives	Setup Phase	Scale-up Phase			Production Phase			
	FY18	FY19	FY20	FY21	FY22	FY23	FY24	FY25
Transformative Technology Development	4 awards							
Rapid Technology Implementation		4 awards						
Transformative Technology Development			5 awards					
Tissue Mapping Centers	5 awards							
Tissue Mapping Centers			6 awards					
Tissue Mapping Centers					~5-7 awards			
HuBMAP Integration, Visualization and Engagement Collaboratory	5 awards							
HuBMAP Integration, Visualization and Engagement Collaboratory					~5-7 awards			
HuBMAP Data Mining and Demonstration Projects					~5-9 awards			

Focus of the Production Phase

Over the winter, we held 4 focus group meetings with PIs, External Consultants, Related Consortia, and NIH Staff and identified the following priorities:

- **Scaling up high-quality data generation** from proven assays on tissues where progress has already been made

- **Data quality, reproducibility, and robustness** are crucial for wider community use

- **Data integration across scales and molecular classes** are key challenges in the production phase

- **Thinking broadly about what constitutes a map** and share the components (tools, methods and data) to allow others to build their own maps

- **Engage the community to determine needs** and key biological questions that can be addressed

HuBMAP Integration, Visualization and Engagement Center
(requesting support for reissue)

Data Mining and Demonstration Projects
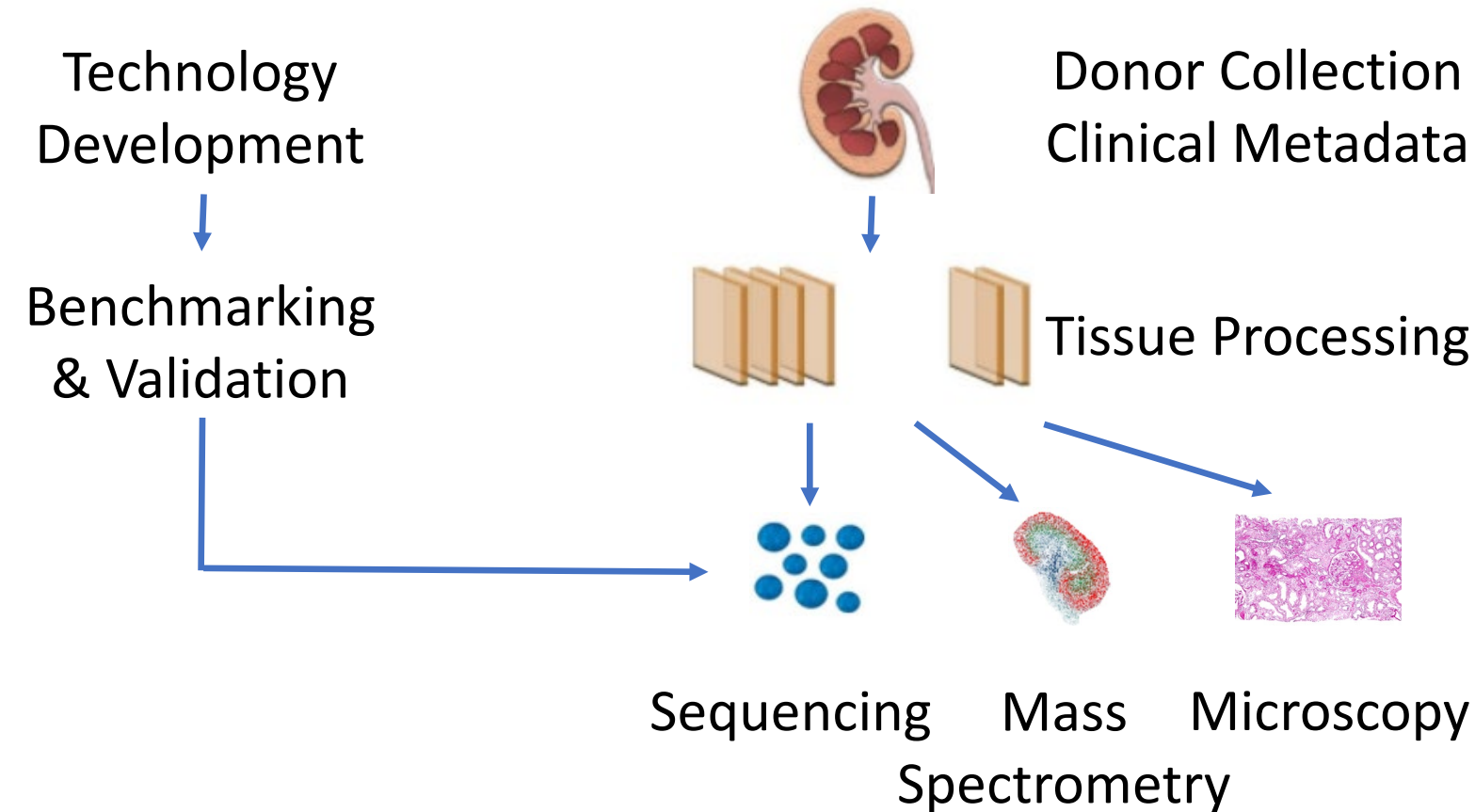(requesting support for new RFA)

Tissue Mapping Centers RFA
(reissue previously approved)

Program Update

- **HuBMAP has grown significantly over the scale-up phase**

- **57 publications, 4600+ citations to date; first publication package this summer**

National Institutes of Health
Office of Strategic Coordination - The Common Fund

Consortium Data Releases

Technology
Development

↓

Benchmarking
& Validation

Donor Collection
Clinical Metadata

↓

Tissue Processing

Sequencing   Mass   Microscopy
            Spectrometry

First data release
(July 2020)
- 33 donors
- 7 organs
- 451 datasets

Second data release
(Planned July 2021)
- +42 donors
- +5 organs
- + ~400 datasets

- HuBMAP is generating spatial and compositional reference datasets that capture "normal" variance

- Mapping patterns at multiple spatial scales is central to HuBMAP vision

  - molecular (e.g. collagen), cellular (e.g. B cells), functional tissue units (glands)

- Establishing a normal reference complements goals of other programs - e.g. Gut Cell Atlas (GCA), Human Tumor Atlas Network (HTAN)

Normal colon

Adenocarcinoma

CD3 CD19 CD31 Cytokeratin Collagen IV

100μm

100μm
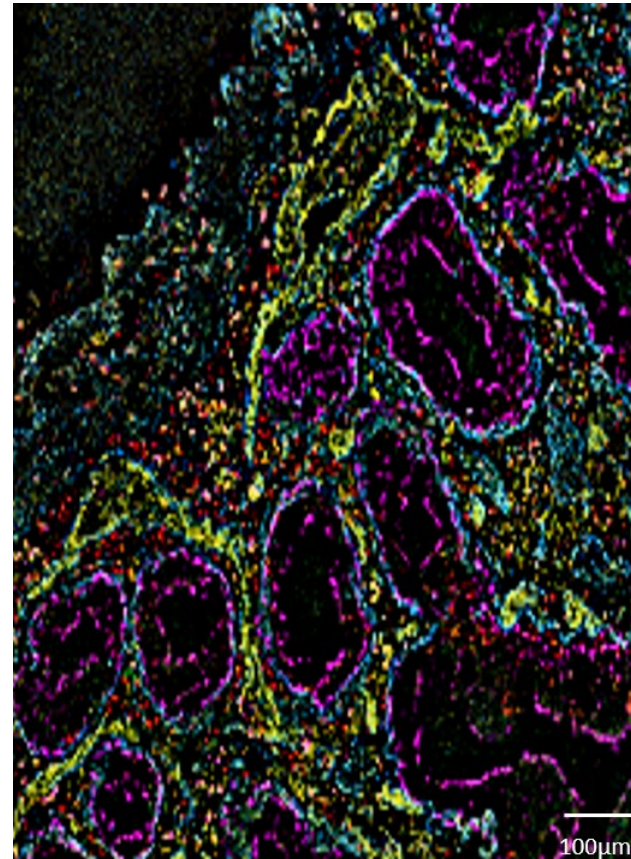
HuBMAP Data
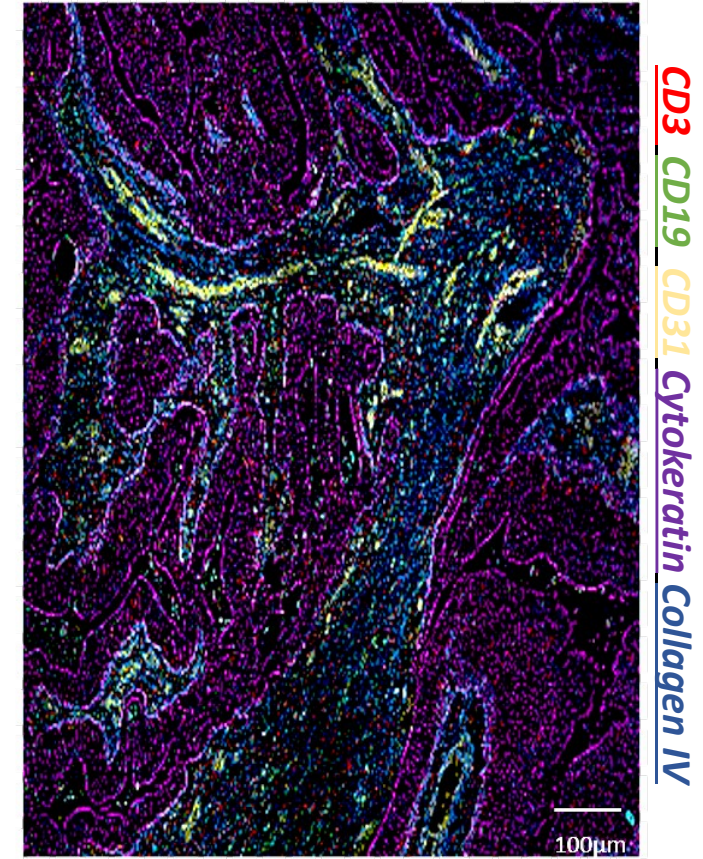
HTAN Data

Courtesy: Tuhin Guha, Ph.D. Stanford University

- The consortium has established the processes for building reference datasets:

  - Generating large, high quality datasets

  - Framework for integration

  - Expert annotation

  - Curation & feedback

- First drafts of reference RNAseq datasets and tools to use them are now available

- HuBMAP has benefited significantly from working closely with expert communities – e.g. Kidney Precision Medicine Program
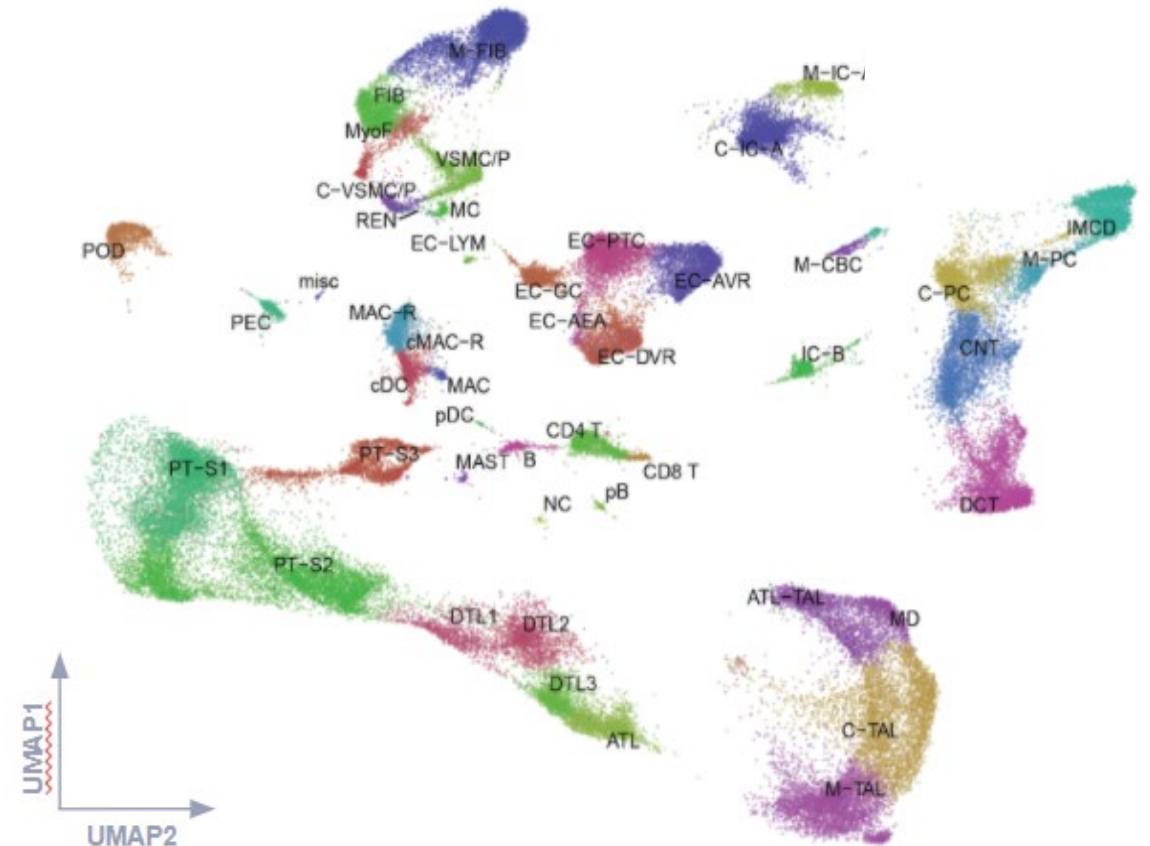

First Draft of Kidney Reference RNA-seq Dataset

Role of the HIVE in HuBMAP

The HuBMAP Integration, Visualization & Engagement (HIVE) Collaboratory:

1. manages the data generated by the HuBMAP Consortium,

2. coordinates internal and external Consortium activities,

3. develops novel tools for visualizing, searching and modelling data, and

4. is building an atlas of tissue maps.

Snyder et al. Nature. 2019; 574(7777)

# HIVE Progress
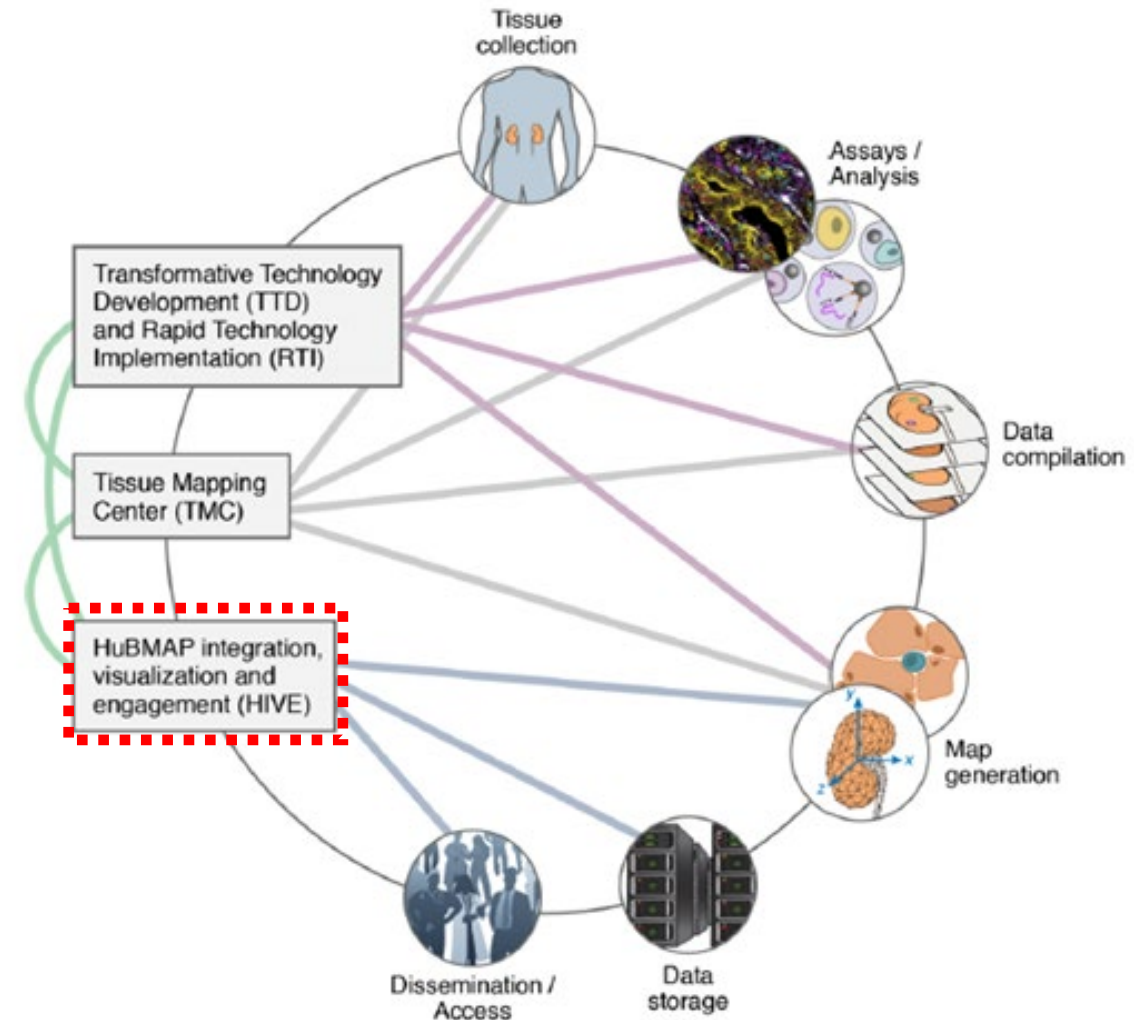
1) Data Management

- Hybrid, microservices architecture

- 6 common data analysis pipelines

- First data release (2020): 23.4 TB data, 451 datasets

- 90 code contributors, 46 in-use repositories, 22 in-use Docker containers

- Rapid dissemination: bioRxiv, Protocols.io, GitHub
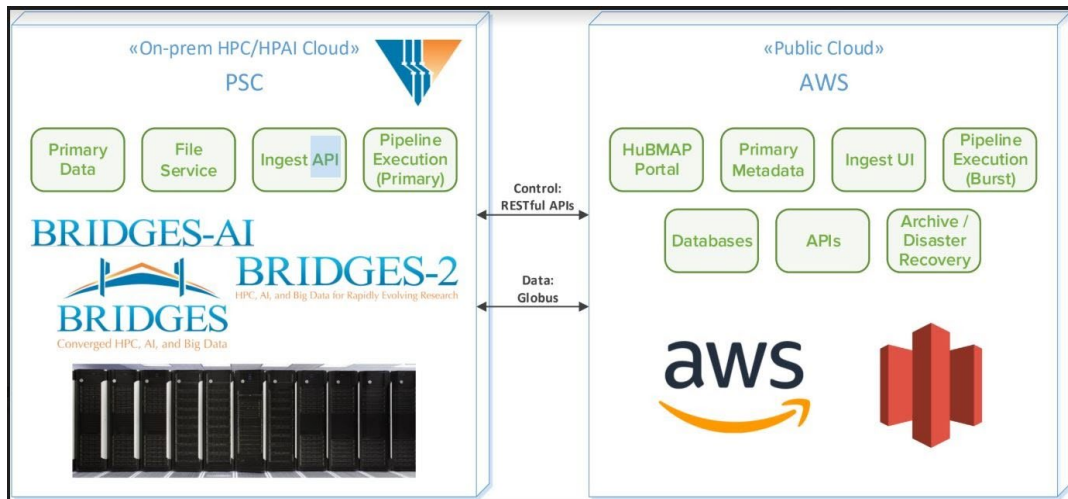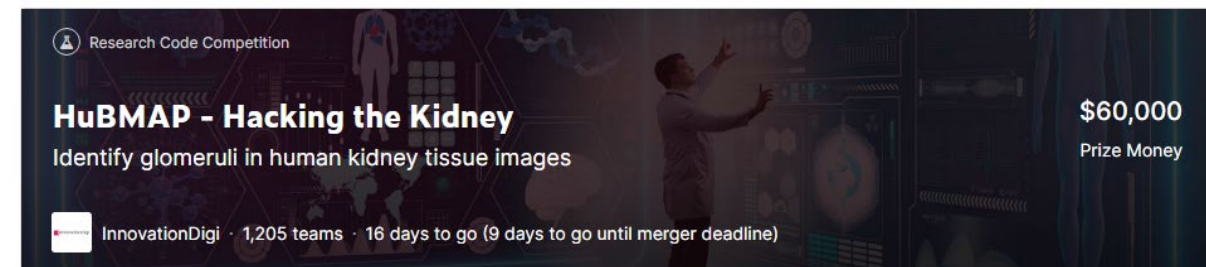

2) Internal & External Coordination
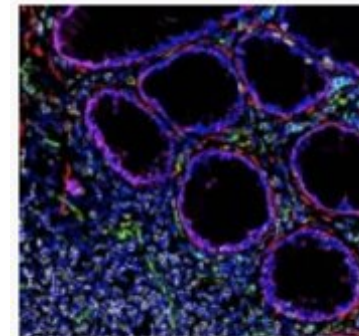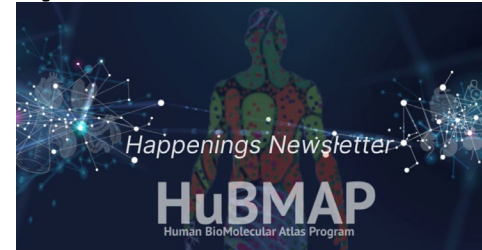
This image uses CODEX imaging to map 7 proteins in a section of colon tissue. Courtesy of John Hickey of the Stanford Tissue Mapping Center.
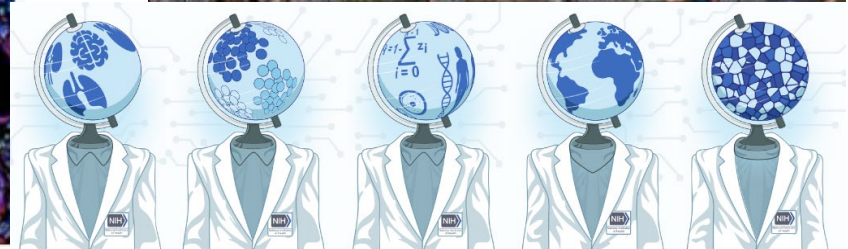
Pre-Meeting Webinars and Schedule │ Meeting Agenda   Breakout summaries

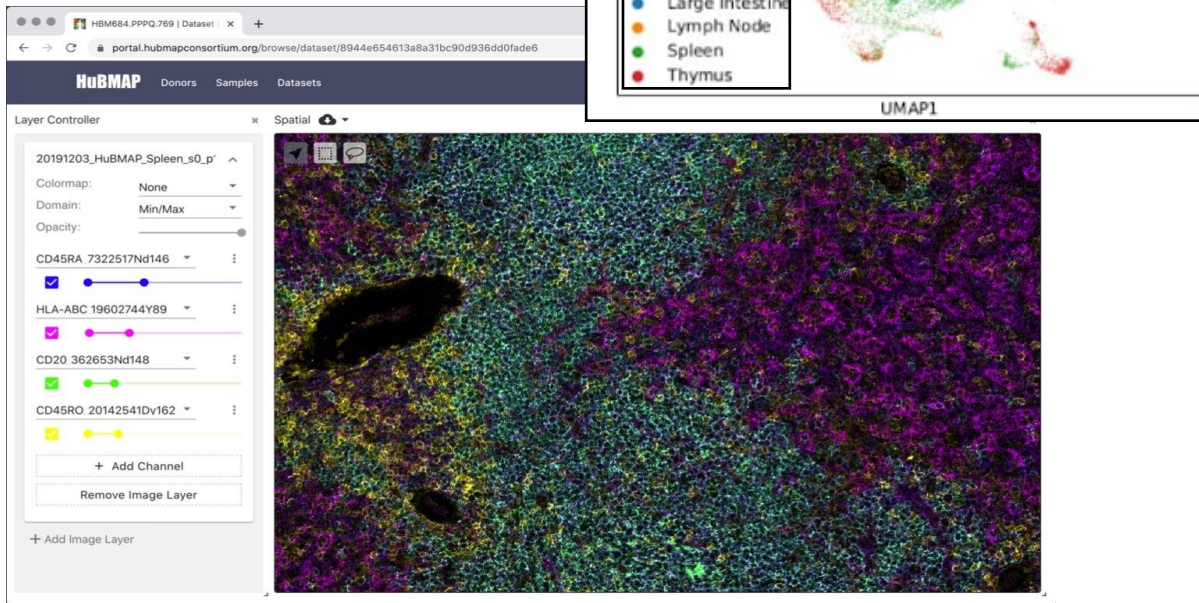NIH-HCA 2020 Joint Meeting

Research Code Competition

HuBMAP – Hacking the Kidney
Identify glomeruli in human kidney tissue images

$60,000
Prize Money

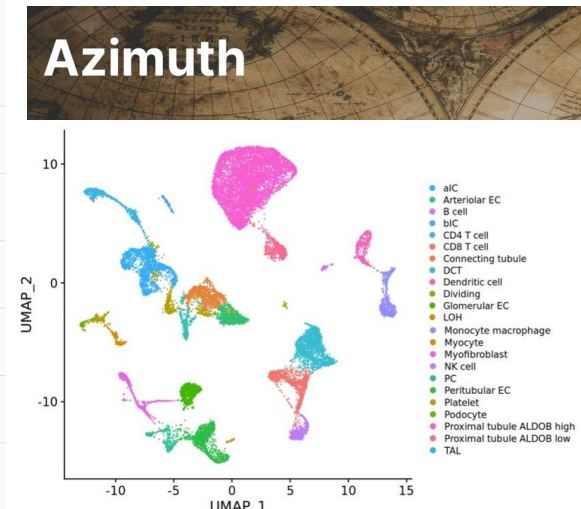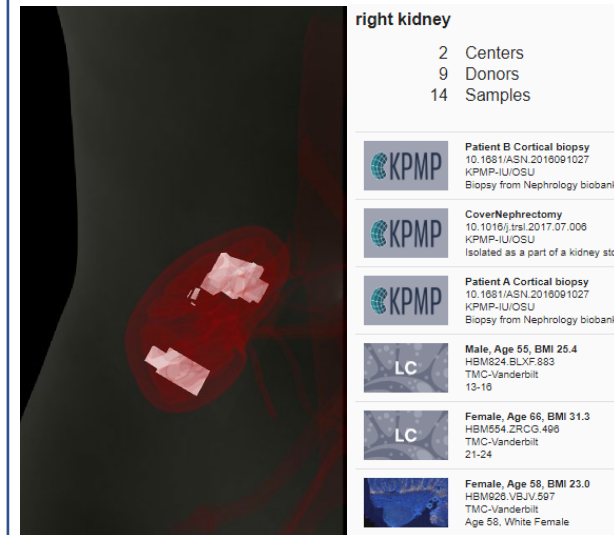InnovationDigi · 1,205 teams · 16 days to go (9 days to go until merger deadline)

HIVE Progress
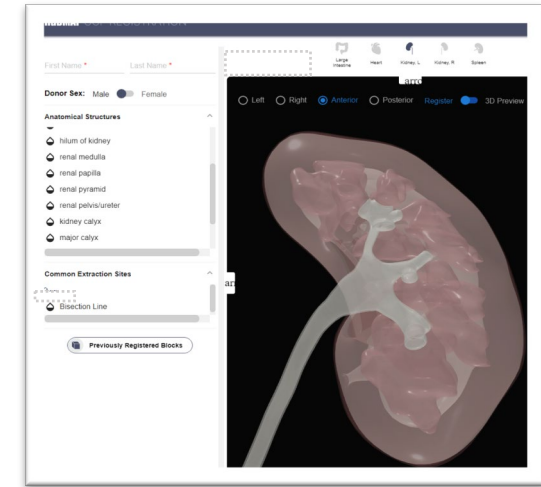
3) Visualization, Search, & Modeling Tools

- Data Portal
- Vitessce Framework
- 8+ APIs
- Integrated Analysis

4) Tissue Atlas

- Registration User Interface
- ASCT+B Tables
- Exploration User Interface
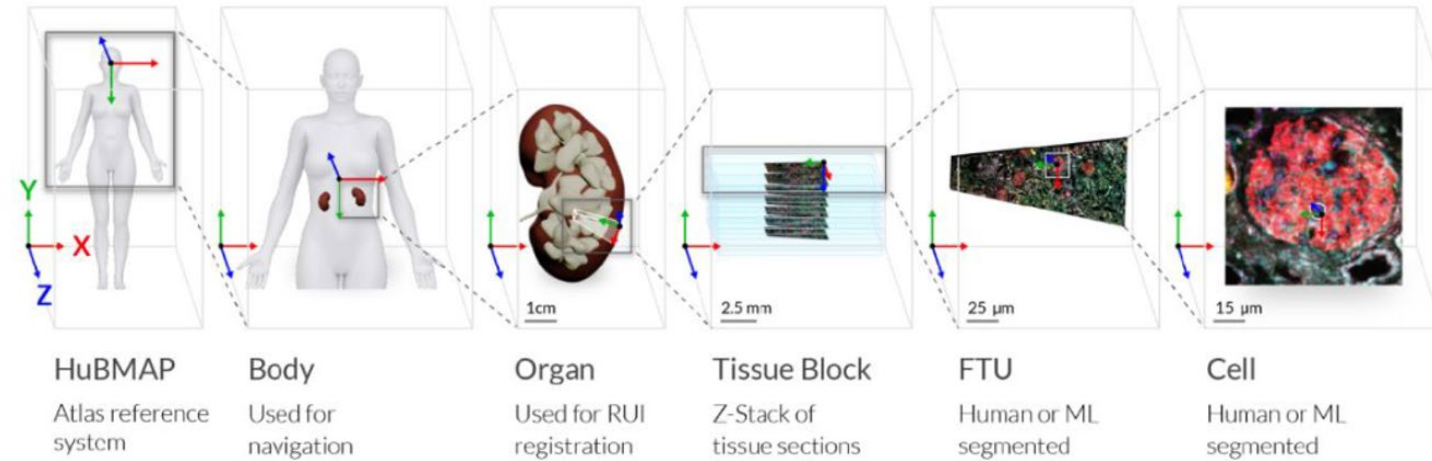- References Tools (Azimuth)

Production Phase Opportunities

Two key challenges identified from Focus Groups:

1) Spatial and molecular integration of data into 3D maps

2) Building richer references and use cases in collaboration with the wider community

National Institutes of Health
Office of Strategic Coordination - The Common Fund

Opportunity: Spatial and Molecular Integration of HuBMAP Data

NIH
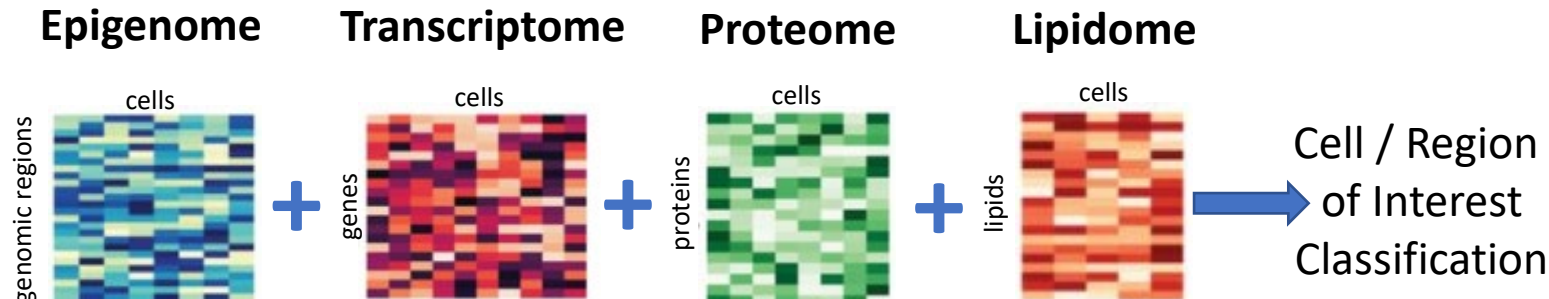*The Common Fund*

Spatial Integration:

- Progress: draft of Common Coordinate Framework (CCF)

- Future Focus: Community engagement; Functional Tissue Unit (FTU) pipelines; visualization tools

HuBMAP	Body	Organ	Tissue Block	FTU	Cell
Atlas reference system	Used for navigation	Used for RUI registration	Z-Stack of tissue sections	Human or ML segmented	Human or ML segmented

Börner et al. arXiv:2007.14474

Molecular Integration:

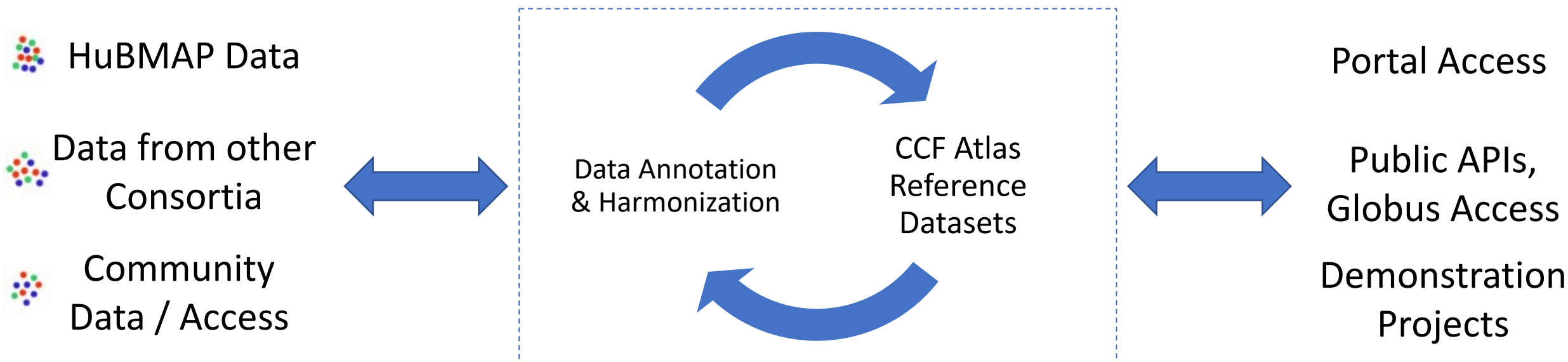**Epigenome** + **Transcriptome** + **Proteome** + **Lipidome** → Cell / Region of Interest Classification

- Progress: new techniques (Gluer, TANGRAM, Cellar)

- Future Focus: High quality datasets; agile testing & comparison; visualization tools

Request: Concurrence with Two Additional Production Phase Initiatives

NIH
The Common Fund

HuBMAP Integration, Visualization and Engagement (HIVE) Collaboratory

- $10.5M / year for 4 years; 5-7 awards
- Extend CCF to additional HuBMAP organs
- Build richer reference datasets and tools to annotate, share, and query them
- Demonstrate querying and integrated analysis with other consortia, and community-contributed datasets
- Enhance awareness and engagement

Demonstration Projects

- $3M / year for 4 years; 5-9 awards
- Address well-defined biological challenges utilizing HuBMAP-generated resources
- Highlight the utility of a Common Coordinate Framework and normal reference datasets
- Develop push-pull relationships with the other consortium members
- Contribute new data, models, and tools to HuBMAP

NIH HuBMAP Working Group

Co - Chairs
- Gary Gibbons (NHLBI)
- Jill Heemskerk (NIBIB)
- Robert Star (NIDDK)

OD
- Richard Conroy
- Tyler Best
- Tony Casco
- Dena Procaccini
- Ananda Roy

CSR
- David Balasundaram

NCI
- Katarzyna Bourcier
- Jennifer Couch
- Allen Dearry
- Shannon Hughes
- Lillian S. Kuo
- Jerry Li
- Stefanie Nelson
- Philipp Oberdoerffer
- Pothur Srinivas
- Wendy Wang
- Dan Xi

NEI
- Grace Shen

NHGRI
- Dan A. Gilchrist
- Julie Kim
- Ajay Pillai

NIDA
- John Satterlee

NHLBI
- Bishow Adhikari
- Brian Bai
- Marishka Brown
- Marc Charette
- Zorina Galis
- Lucy Hsu
- Cashell Jaquish
- Sara Lin
- Margaret Ochocinska
- Asif Rizwan

NIA
- Viviana Perez Montes

NIAID
- Halonna Kelly
- Kentner Singleton

NIAMS
- Robert Carter
- Su-Yau Mao

NICHD
- Lisa Halvorson
- Reiko Toyama

NIDDK
- Eric Brunskill
- Patricia Greenwel
- Salvatore Sechi

NIGMS
- Sarah Dunsmore
- Joseph G. Gindhart
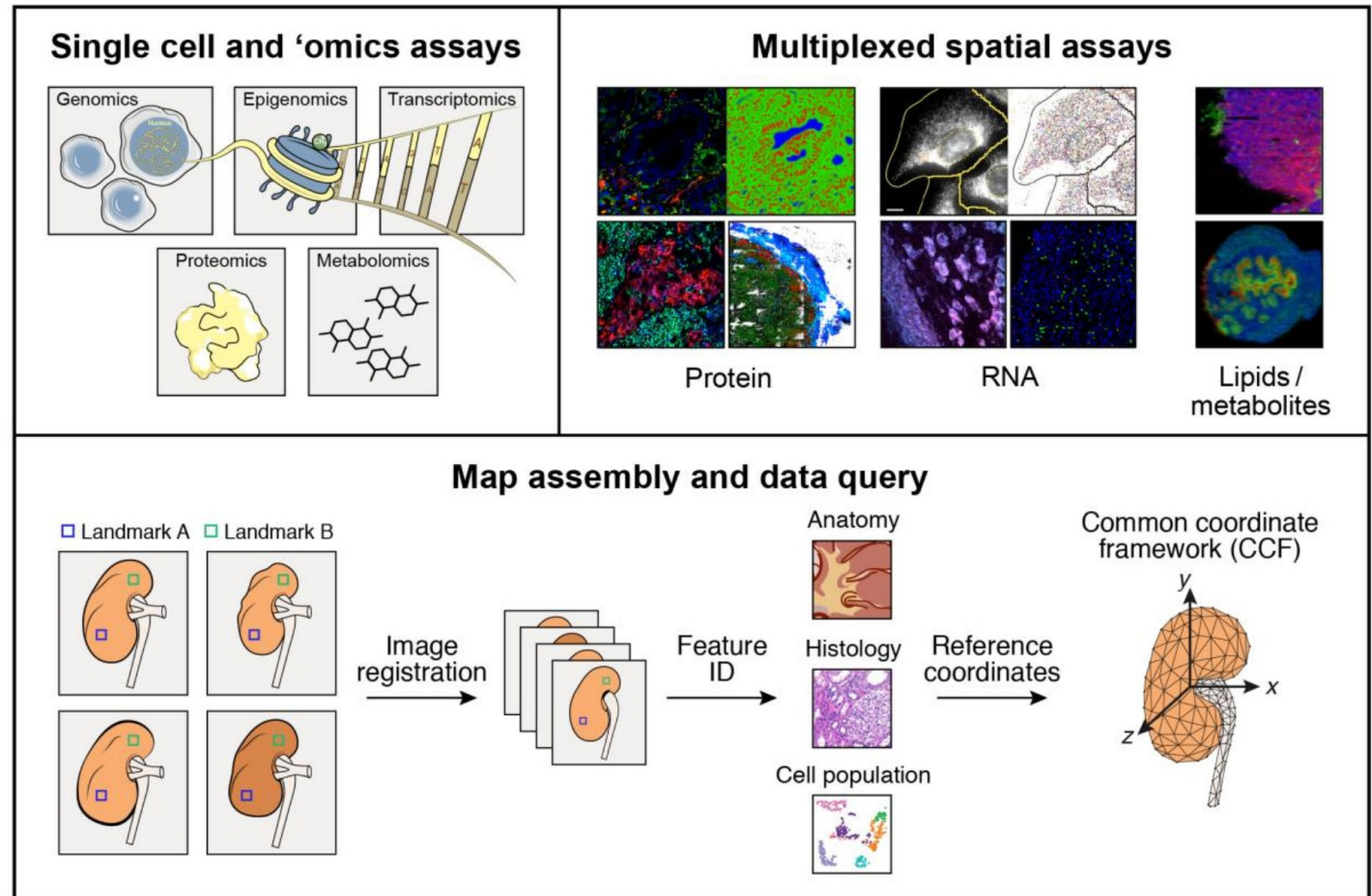
NIMH
- Andrea Beckel-Mitchener
- Yong Yao

NINDS
- Francesca Bosetti
- Timothy LaVaute
- Daniel Miller
- Jill Morris

To catalyze the development of a framework for mapping the human body at single cell resolution.

Single cell and 'omics assays

Genomics   Epigenomics   Transcriptomics

Proteomics   Metabolomics

Multiplexed spatial assays

Protein   RNA   Lipids / metabolites

Map assembly and data query

☐ Landmark A   ☐ Landmark B

Image registration → Feature ID → Reference coordinates

Anatomy
Histology
Cell population

Common coordinate framework (CCF)
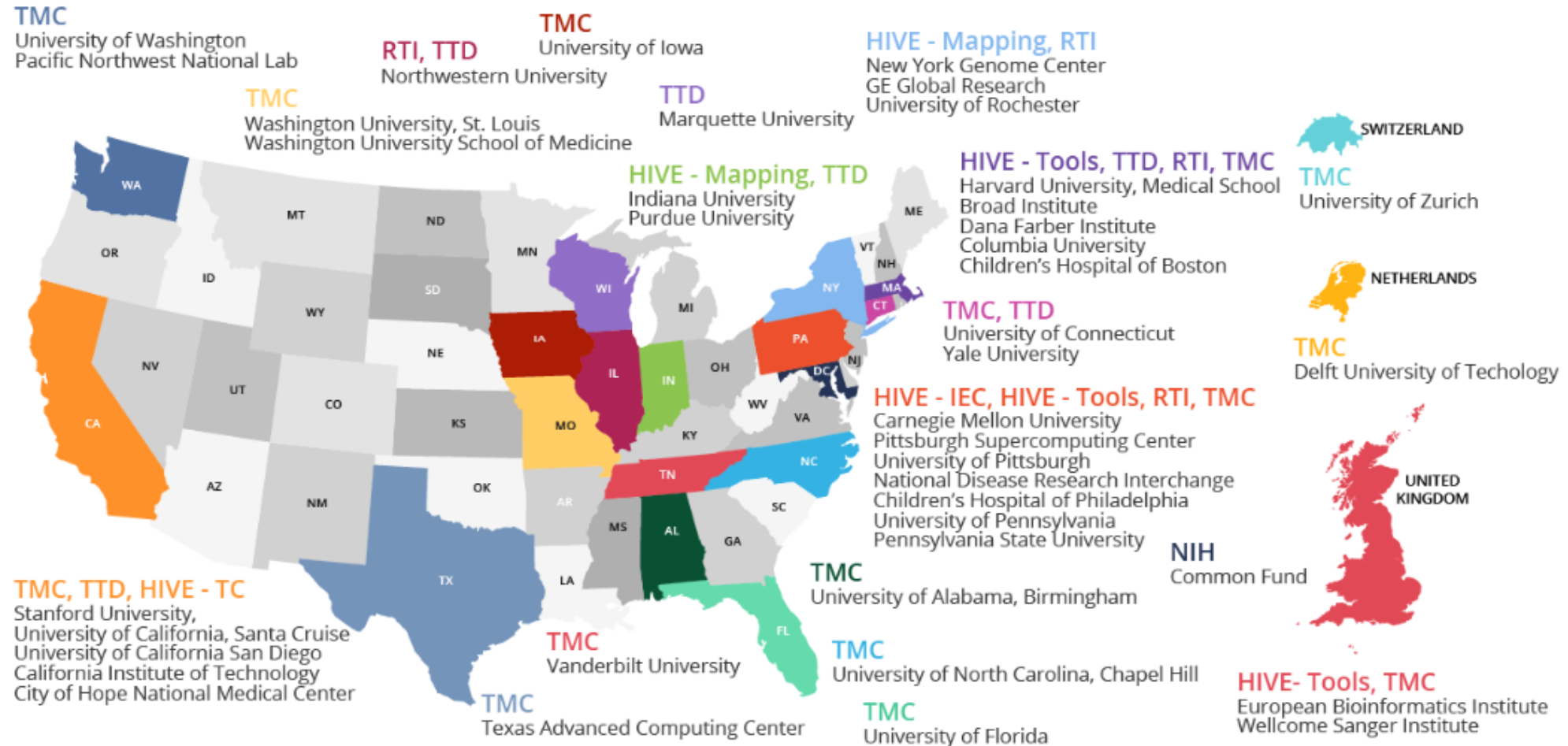
Thank You.

Questions?

National Institutes of Health
Office of Strategic Coordination - The Common Fund

Supplemental Slides

May 20-21, 2021

NIH Council of Councils

National Institutes of Health
Office of Strategic Coordination - The Common Fund

HuBMAP Contributing Sites

HuBMAP Assays

IMAGING

DNA/RNA
DART-FISH
seqFISH
smFISH
MERFISH
Slide-seq
SABER-FISH
GeoMx

Lipids/Metabolites
MALDI Imaging MS
SIMS Imaging
DESI Imaging MS
NanoDESI Imaging MS

Proteins
Multiplexed IF
IHC
Lightsheet
CODEX
Cell DIVE
DART-FISH
CyTOF Imaging
MALDI Imaging MS
nanoPOTS
MIBI
Immuno-SABER

Other
MR Imaging
CT Imaging
Autofluorescence
Stained Microscopy

SEQUENCING

snDropseq
scRNAseq
snRNA-seq
snATAC-seq
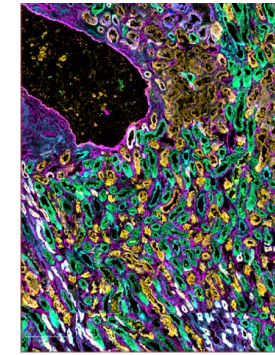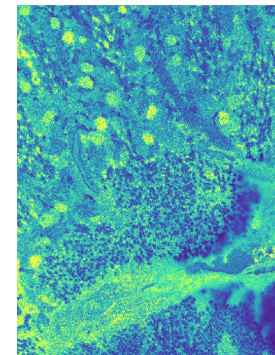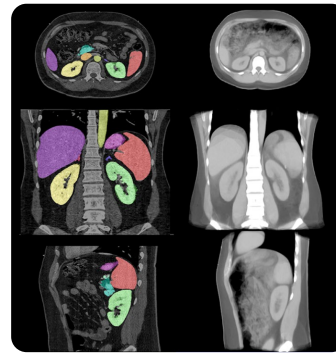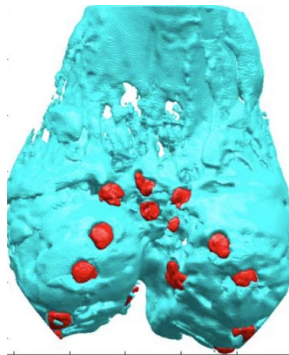sciRNAseq
sciATACseq
scTHSseq
SNAREseq
scATACseq

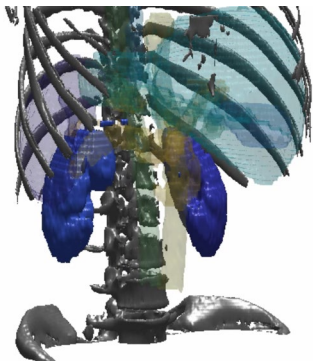## BULK OMICS

Lipids/Metabolites
LC-MS/MS

Proteins
Bottom-up LC-MS/MS
Top-down LC-MS/MS
TMT LC-MS/MS

Consortium Publications

- **57 publications**

- 4600+ citations
- 6 highly cited (100+) papers
- HuBMAP Consortium Marker Paper, *Nature,* 2019

PERSPECTIVE
OPEN
https://doi.org/10.1038/s41586-019-1629-x

The human body at cellular resolution: the NIH Human Biomolecular Atlas Program

HuBMAP Consortium*

nature biotechnology
ARTICLES

Identification of spatially associated subpopulations by combining scRNAseq and sequential fluorescence *in situ* hybridization data

Qian Zhu[1], Sheel Shah[2,3], Ruben Dries[1], Long Cai[2] & Guo-Cheng Yuan[1]

LETTER
https://doi.org/10.1038/s41586-019-1049-y

Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH+

Chee-Huat Linus Eng[1], Michael Lawson[2], Qian Zhu[3], Ruben Dries[3], Noushin Koulena[3], Yodai Takei[3], Jina Yun[3], Christopher Cronin[3], Christoph Karp[3], Guo-Cheng Yuan[3] & Long Cai[2]*

Cell
Resource

Comprehensive Integration of Single-Cell Data

Graphical Abstract

Authors
Tim Stuart, Andrew Butler, Paul Hoffman, ..., Marlon Stoeckius, Peter Smibert, Rahul Satija

Genome Biology
Hafemeister and Satija *Genome Biology*   (2019) 20:296
https://doi.org/10.1186/s13059-019-1874-1

Method
Open Access

Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression

Christoph Hafemeister[1]* and Rahul Satija[1,2]*

nature methods
ARTICLES
https://doi.org/10.1038/s41592-019-0404-0

SABER amplifies FISH: enhanced multiplexed imaging of RNA and DNA in cells and tissues

Jocelyn Y. Kishi[1,2,6], Sylvain W. Lapan[3,6], Brian J. Beliveau[1,2,5,6]*, Emma R. West[3,6], Allen Zhu[1,2], Hiroshi M. Sasaki[1,2], Sinem K. Saka[1,2], Yu Wang[1,2], Constance L. Cepko[3,4]* and Peng Yin[1,2]*

Production Phase Focus Group Meetings

Four focus groups meetings took place between December 2020 – February 2021.

- **External Program Consultants** - Stephen Hewitt (NCI), Jason Swedlow (Dundee), Alexandra Naba (UIC), Janet Iwasa (Utah), Sue Hammoud (UMich)

- **HuBMAP PI's** - Katy Börner, Sheng Zhong, Louise Laurent, Jeff Spraggins, Fiona Ginty, Julia Laskin, Nils Gehlenborg, Neil Kelleher

- **Related Consortia** - Laura Clarke (EBI, HCA), James Monaghan (SPARC), Tim Tickle (HCA, BICCN, LungMAP), Tim Downing (HRHR at UC Irvine), Michael Eadon (KPMP), Yan Xu (LungMAP, SPARC)

- **NIH WG** (see separate list)

Each focus group was asked to share their opinions on the following questions:

- Are high quality biomolecular maps being generated and validated at a sufficient rate? Are the data interoperable?

- How well do you think the components of HuBMAP interact with each other? What can be done to improve this interaction?

- Have HuBMAP collaborations with other consortia been fruitful?

- Where are the opportunities for HuBMAP as it moves to the production phase?

Feedback from the Focus Groups

Common themes from the four meetings include:

- Data integration across scales and molecular classes are key challenges in the production phase

- Thinking broadly about what constitutes a map and share the components (tools, methods and data) to allow others to build their own maps

- Engage the community to determine needs and key biological questions that can be addressed

- Focus on scaling up high-quality data generation from proven assays on tissues where progress has been made

- Data quality, reproducibility, and robustness are crucial for wider community use

Additional themes from the four meetings include:

- As a "top-down" consortium, HuBMAP is in a singular position to create SOPs and standards that would benefit the single-cell atlasing communities

- The production phase is the time to change the focus from technology development to answering biological questions.

- In the production phase, HuBMAP needs to scale-up using an engineering approach like the Human Genome Project.

- Both 2D and 3D maps that are interactive and represent molecular, cellular, and spatial information that will allow analysis of various scales and resolutions are necessary for cell atlases to be a true resource for the community.

**Tissue Mapping Centers
(Reissue approved by CoC, September 2019)**

- $11M/yr; 4 years; 5-7 awards
- Existing organs
  - Multiple tissues from the same individual; body-wide systems; diversity
- Core, common set of assays
- Multiscale approaches
- Adaptive (push-pull) data generation strategies
- End-to-end Quality Assurance; data reproducibility
  - Tissue quality control
  - Validation of affinity reagents
  - Reportable quality metrics

HuBMAP Resources

Consortium Webpage: https://hubmapconsortium.org

- Consortium Publications, BioRxiv channel, and marker paper

- Open Working Groups: Cell Atlas Curation, Affinity Reagent Imaging and Validation, ASCT+B Working Group

- Consortium Policies

- Image of the Week

- Protocols on Protocols.io

HuBMAP Data Portal: https://portal.hubmapconsortium.org

- Anatomical Structures, Cell Types and Biomakers Reporter

- CCF Registration User Interface

- CCF Exploration User Interface

- Data Analysis Pipelines

- GitHub - code repository

- Vitessce - visualization platform

- Azimuth – reference labelling for RNA-seq data