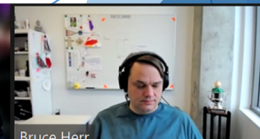
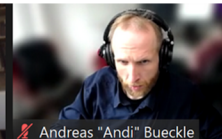
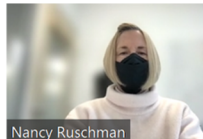
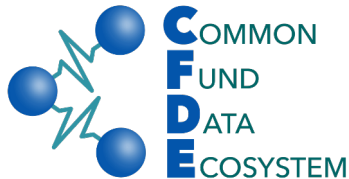


# Cloud-Based Common Coordinate Framework to Enable Integration and Analyses Across GTEx and HuBMAP Data

**Presenting:**

Kristin Ardlie & Jared Nedzel (GTEx) and  
Katy Börner, Bruce W. Herr II & Andreas Bueckle (HuBMAP)

March 1, 2022



# The Challenge

Disparate Common Fund projects such as GTEx and HuBMAP (and others including MoTrPAC, SPARC, Kids First...etc.) have generated (or are generating) large molecular and spatial datasets on human tissues (contributing to even larger, similar data generation efforts).

For these data to be maximally “Interoperable and Reusable”, we need to accurately map and compare data from the same organs, tissues, and cell types across these projects.

We want to map:

- ❑ GTEx samples - collected across MANY individuals at the same tissue site per individual (using SOPS)
- ❑ HuBMAP samples - collected for fewer individuals at MANY specific sites within a tissue (mapped to tissue locations using ASCT+B details).

# Goals

Spatially map samples collected under GTEx biospecimen SOPs to the CCF to which HuBMAP samples are mapped

Generate and make available spatially, semantically, and ontologically explicit FAIR metadata for GTEx and HuBMAP data housed on several cloud environments.

Extend and serve as web components data visualization and exploration user interfaces that were initially developed for HuBMAP but are valuable for other consortia.

Make a combination of anatomically mapped and ontologically linked human tissue data available across HuBMAP (hybrid cloud) and GTEx portal (Google Cloud Platform).

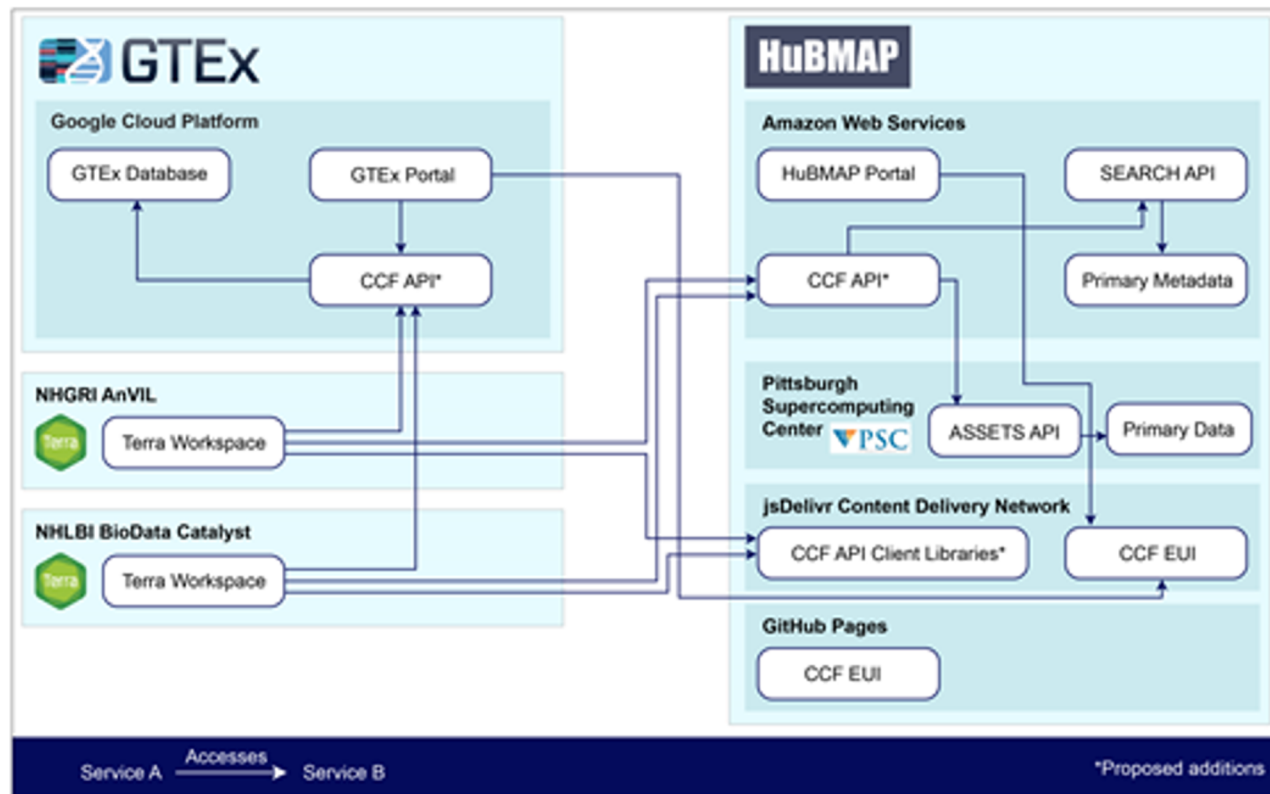
# Approach

Design and deploy a FAIR CCF-API to map GTEx tissue data to HuBMAP CCF.

Develop libraries to enable the use of FAIR CCF-API for visualization of GTEx data in HuBMAP or GTEX portal and cross-search for ASCT+B indexed data across cloud platforms.

Develop training materials and conduct a user survey to support wider usage and adoption of the pilot cloud setup and CCF cross-search.

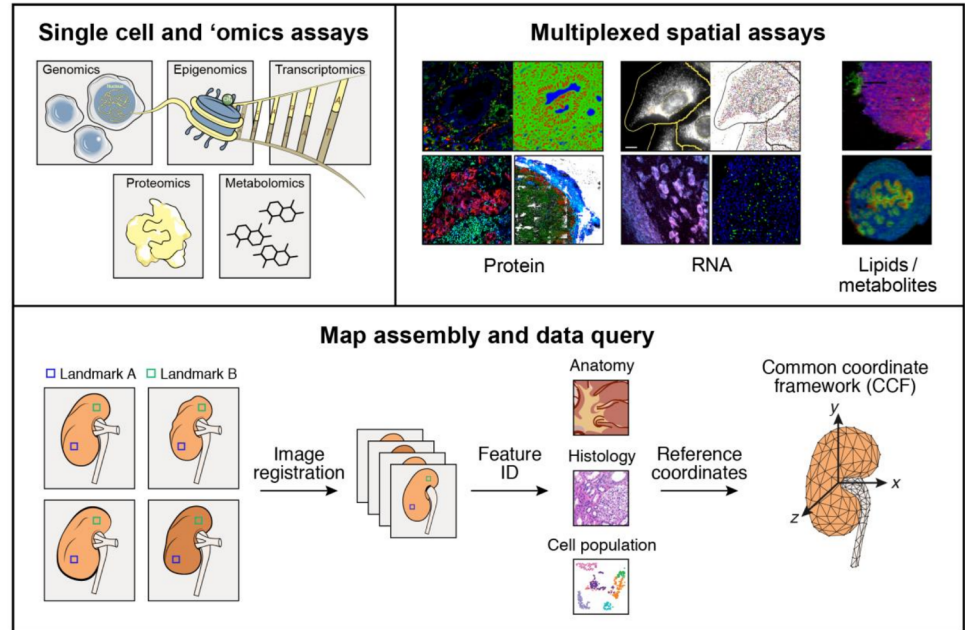
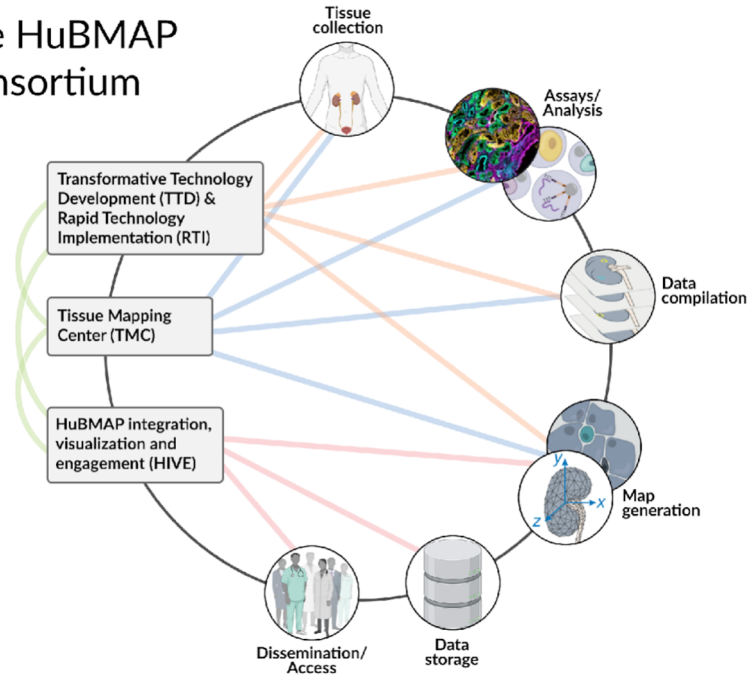
# System Architecture



# The Human Body at Cellular Resolution: The NIH Human Biomolecular Atlas Program.

Snyder et al. *Nature*. 574, p. 187-192.

## The HuBMAP Consortium



# Common Coordinate Framework (CCF): ASCT+B Tables & 3D Reference Object Library

Anatomical Structures (AS), Cell Types (CT), and Biomarkers (B) or ASCT+B tables aim to capture the partonomy of anatomical structures, cell types, and major biomarkers (e.g., gene, protein, lipid or metabolic markers).

ASCT Table

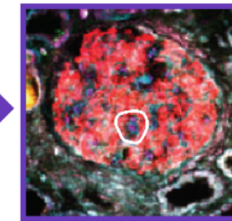
Structure/Region	Sub structure/Sub region	Cell Type
Renal Corpuscle	Bowman's (glomerular) Capsule/parietal layer	Parietal epithelial Cell
	Bowman's (glomerular) Capsule/visceral layer	Podocyte
	Glomerular Tuft	Capillary Endothelial Cell Mesangial Cell
Tubules	Proximal Tubule	Proximal Tubule Epithelial Cell (general)
		Proximal Convoluted Tubule Epithelial Cell Segment 1
		Proximal Tubule Epithelial Cell Segment 2
		Proximal Tubule Epithelial Cell Segment 2
		Proximal Tubule Epithelial Cell Segment 2
	Loop of Henle, Thin Limb	Descending Thin Limb Cell (general)
		Ascending Thin Limb Cell (general)
	Loop of Henle, Thick Limb	Thick Ascending Limb Cell (general)
		Cortex-TAL Cell
		Medulla-TAL Cell
	Distal Convolution	TAL-Macula Densa Cell
		Distal Convoluted Tubule Cell (general)
	Connecting Tubule	DCT Type 1 Cell
DCT Type 2 Cell		
Connecting Tubule Cell (general)		
		CNT-Principal Cell

Ontology

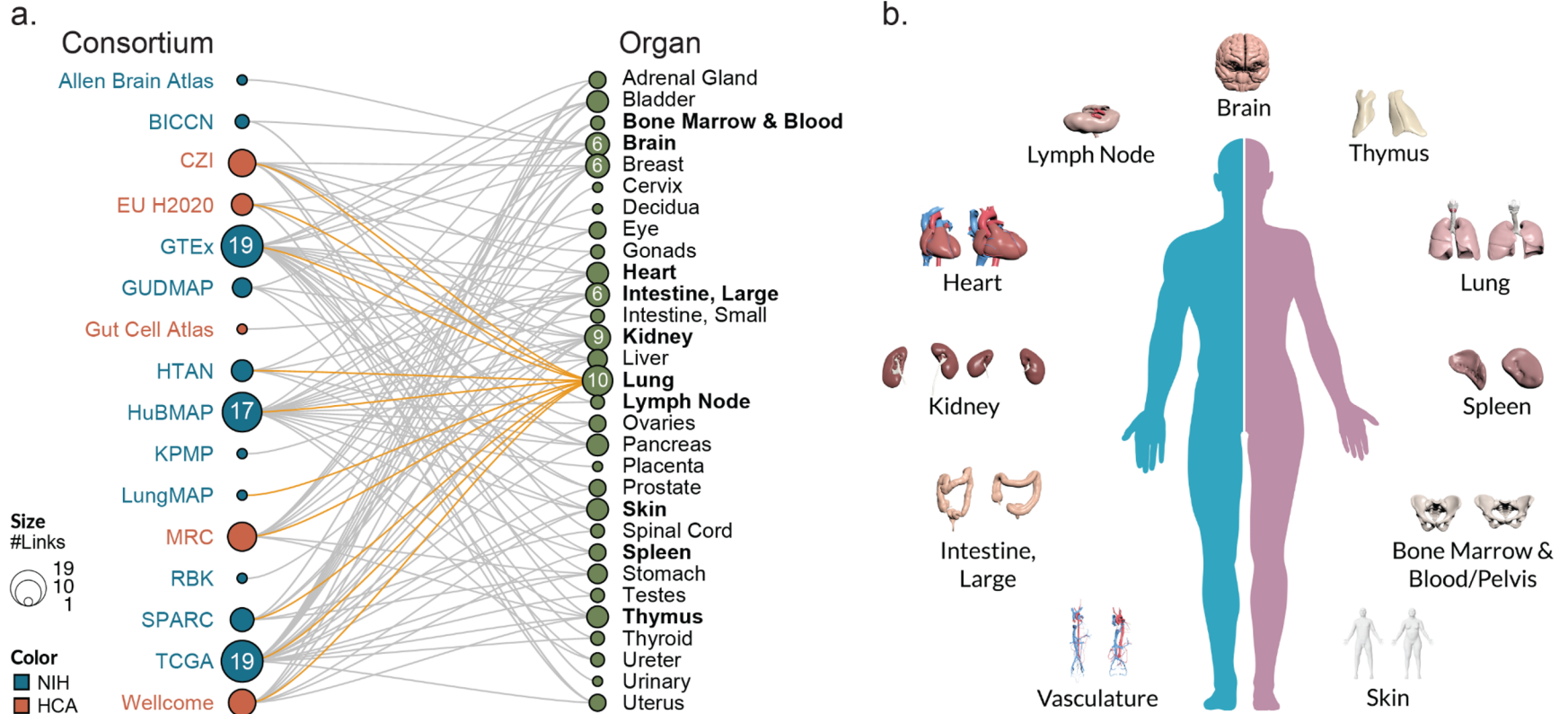
*Anatomical Structures Partonomy*  
 kidney  
 kidney capsule  
 cortex of kidney  
 outer cortex of kidney  
 renal medulla

*Cell Types Ontology*  
 connective tissue cell  
 pericyte cell  
 mesangial cell  
 extraglomerular mesangial cell  
 glomerular mesangial cell

3D Reference Object Library



ASCT+B Tables & 3D Reference Objects are compiled across 16 consortia since the NIH-HCA Joint Meeting in March 2020, <https://hubmapconsortium.org/nihhca2020>



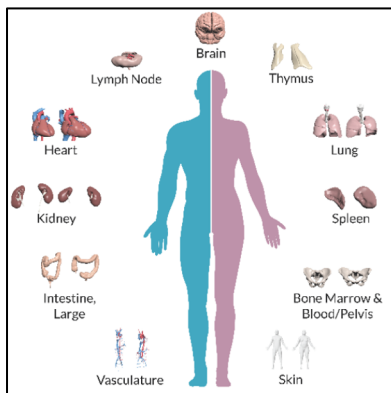


## Anatomical Structures (AS)

## Cell Types (CT)

## Biomarkers (B)

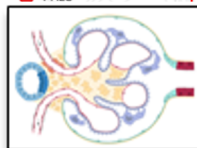
Partonomy Tree  
*part\_of*



Red Pulp

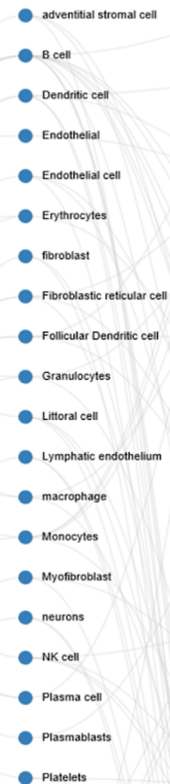
White Pulp

Spleen



Bimodal network describing which CT are located\_in what AS

Typology Tree  
*is\_a*



Bimodal network describing which B characterize what CT

BG - Genes  
BP - Proteins



# Atlas & Tools

**HuBMAP** Donors Samples Datasets Collections Previews Atlas & Tools Documentation My Lists Member Login

## Human BioMolecular Atlas Program

An open, global atlas of the human body at the cellular level

The HuBMAP Data Portal is the central resource for discovery, visualization, and download of single-cell data. Standardized data curation and processing workflow ensure that only high quality is released.

- Common Coordinate Framework (CCF) Portal
- ASCT+B Reporter
- Exploration User Interface (EUI)
- Registration User Interface (RUI)
- Azimuth: Reference-based single cell mapping

## Navigate healthy human cells with the Common Coordinate Framework

Interact with the human body data with the Anatomical Structures, Cell Types and Biomarkers (ASCT+B) Tables and CCF Ontology. Also explore two user interfaces: the Registration User Interface (RUI) for tissue data registration and Exploration User Interface (EUI) for semantic and spatial data.

Get Started

**HuBMAP** Ser: Both Age: 1-110 BMI: 13-83 Login

Search ontology terms ...

- body
  - heart
  - lung
  - kidney
    - right kidney
      - left kidney
        - kidney capsule
        - cortex of kidney
        - renal medulla
        - renal column
        - renal pyramid
        - hilum of kidney
        - kidney interstitium
        - kidney calyx
        - renal pelvis
        - ureter
        - renal papilla
        - renal fat pad
        - nephron
      - spleen
        - spleen capsule
        - tabecula of spleen
        - spleen pulp
          - marginal zone of spleen
          - spleen perifollicular zone
          - hilum of spleen
        - colon
          - ascending colon
          - descending colon
          - transverse colon
          - sigmoid colon

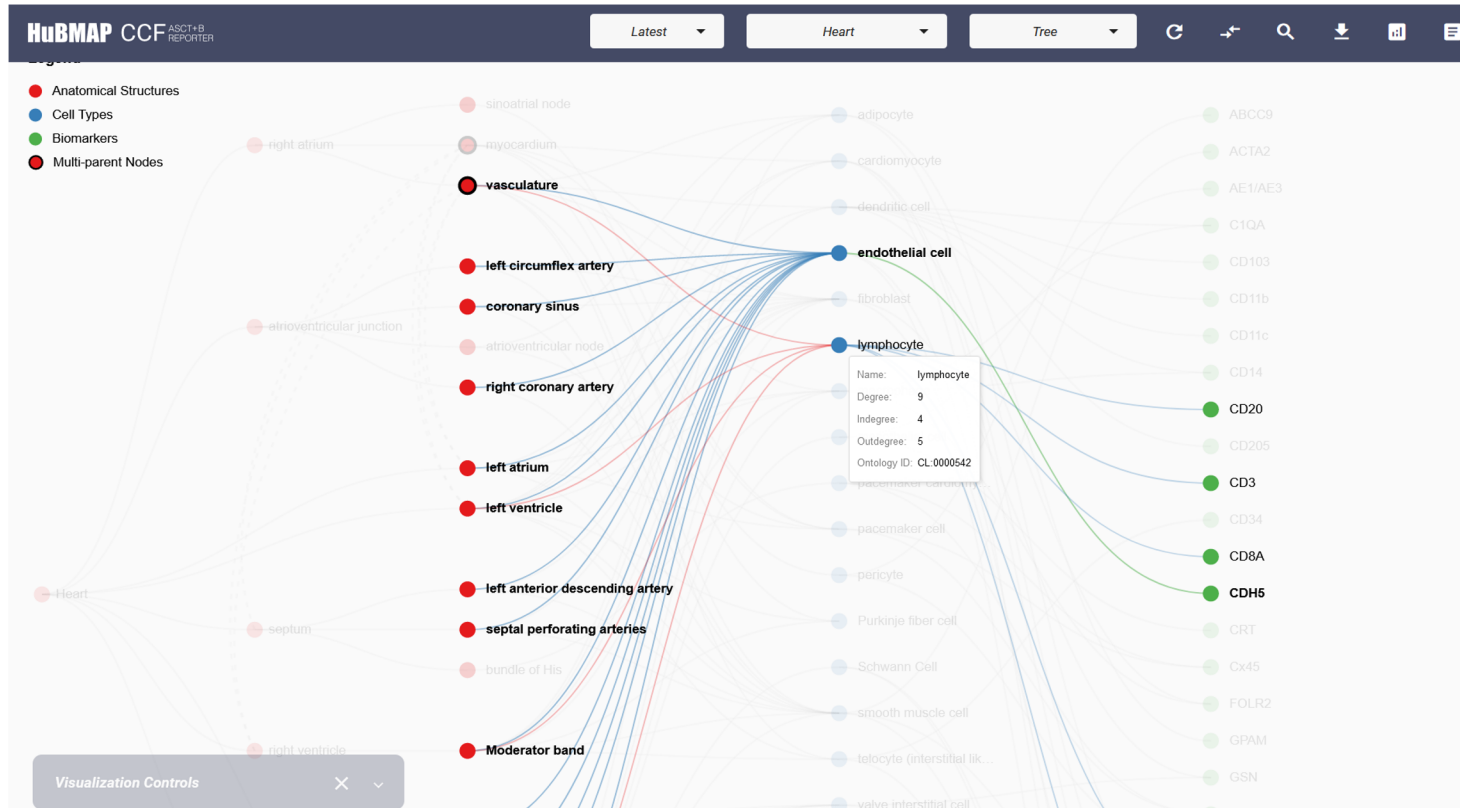
body

- 4 Contours
- 26 Donors
- 48 Samples

Sample ID	Method	Age	Sex	Contour	Donor	Sample
KMP	Cellular Atlas	18	M	1	1	1
KMP	Cellular Atlas	18	M	1	1	1
KMP	Cellular Atlas	18	M	1	1	1
10x	Cellular Atlas	18	M	1	1	1
10x	Cellular Atlas	18	M	1	1	1
CODX	Cellular Atlas	18	M	1	1	1
CODX	Cellular Atlas	18	M	1	1	1
CODX	Cellular Atlas	18	M	1	1	1
LC	Cellular Atlas	18	M	1	1	1
LC	Cellular Atlas	18	M	1	1	1

38 Donors    365 Samples    391 Datasets    13 Collections

# CCF ASCT+B Reporter UI



<https://hubmapconsortium.github.io/ccf-asct-reporter/>

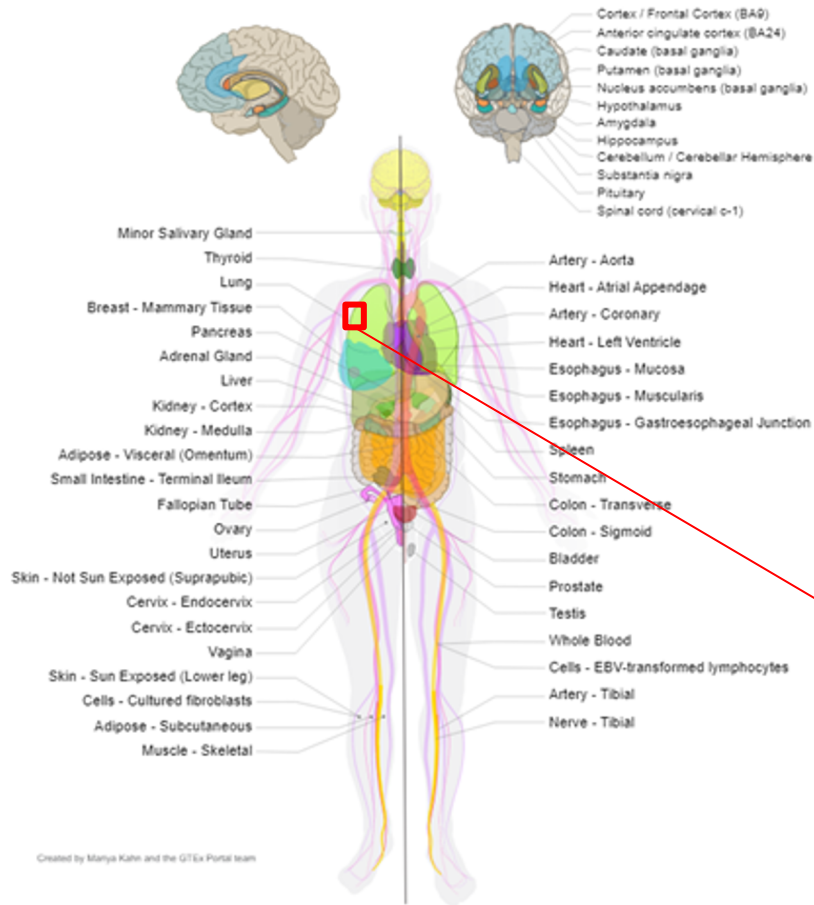
# CCF Registration User Interface (RUI)

The screenshot displays the CCF Registration User Interface (RUI) for the Human Brain Atlas (HuBMAP). The interface is organized into several key sections:

- Header:** "HuBMAP CCF REGISTRATION" is displayed in the top left corner.
- Navigation Bar:** A horizontal bar at the top contains icons for various anatomical structures, including Blood Vessels, Brain, Eye, Fallopian Tube, Heart, Kidney, Liver, Lung, Lymph Node, Ovary, Pancreas, Pelvis, Prostate, Skin, Small Intestine, and Spleen.
- Donor Sex:** A toggle switch is set to "Male".
- Anatomical Structures:** A list on the left includes: renal pyramid, kidney calyx, major calyx, minor calyx, cortex of kidney, renal column, outer cortex of kidney, and renal pelvis.
- Landmarks:** A list on the left includes: bisection line.
- Buttons:** "Previously Registered Blocks" and "UPLOAD PREVIOUS REGISTRATION DATA" are located at the bottom left.
- Central Viewport:** A large 3D visualization of a kidney is shown. The view is set to "Anterior" (selected), with other options being "Left", "Right", "Posterior", "Register", and "3D Preview". A small human figure icon is visible in the top right of the viewport. Coordinates are displayed as X: 95, Y: 53, Z: 41.
- Right Panel:** Contains registration and visualization controls:
  - Tissue Block Dimensions (mm):** Width (X) 10, Height (Y) 10, Depth (Z) 10.
  - Tissue Sections:** A table with columns for Thickness and # Sections.
  - Tissue Block Rotation:** Sliders for X, Y, and Z axes, all set to 0.
  - Anatomical Structure Tags:** A search box "Add Anatomical Structures ..." and a legend for "Assigned" (black dot) and "Added" (red dot).
- Bottom Right:** A "REVIEW AND DOWNLOAD" button is located at the bottom right.

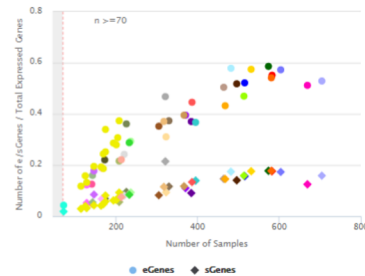
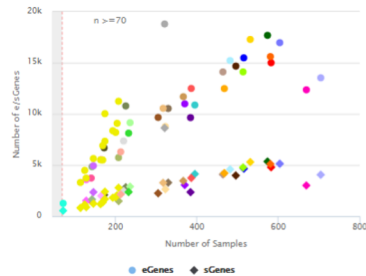
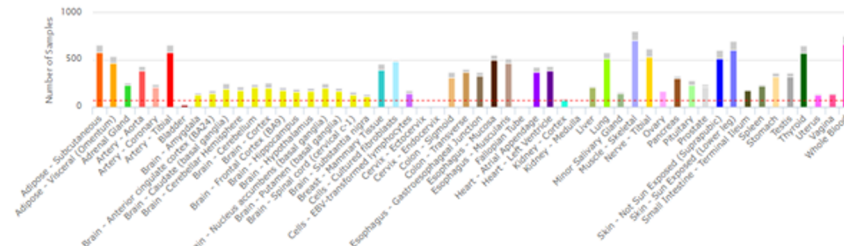
<https://hubmapconsortium.github.io/ccf-ui/rui/>

# GTEX - Breadth across tissue types, less depth within, but large # donors

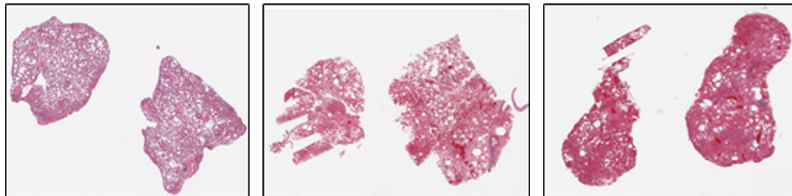


V8 Sample Counts by Tissues

Sort tissues by:    
[Download](#)



## LUNG



# GTEx SOPs designed for repeatable anatomical sampling

### 4.3 TISSUE PROCUREMENT

#### 4.3.1 General

For non-brain donors, tissue collection must be started AND the first tissue must be placed into fixative within 8.0 hours of cardiac cessation or recorded time of death (observed or presumed). For brain donors, all tissues must be collected and placed into fixative within 24.0 hours of cardiac cessation (observed or presumed).

**NOTE: The brain should NOT be collected if the donor was on a ventilator for  $\geq 24.00$  hrs.**

**NOTE: In the event that the GTEx donor was a transplant recipient (either human or xenotransplant, as noted in question #15 of the Donor Eligibility Form), tissue should not be collected from the transplanted organ/tissue or the native organ/tissue of the same type.**

#### 4.3.2 Documentation

Capture biospecimen-related data on the **GTEx Tissue Recovery Case Report Form, PM-0003-F5**.

#### 4.3.3 Organ Priority

The order of organ removal is left to the discretion of the individual BSSs, with TWO important distinctions:

- The brain **must** be removed last.
- If there is difficulty dissecting the coronary artery, it should be removed after the brain.

#### 4.3.4 Aliquot Location

Any deviation from the preferred tissue location of collected aliquots must be documented on the **GTEx Tissue Recovery Case Report Form, PM-0003-F5**. This should be done by noting the actual location either by checking one of the listed locations or manually entering it into the "comment" field.

#### 4.3.5 Aliquot Preparation

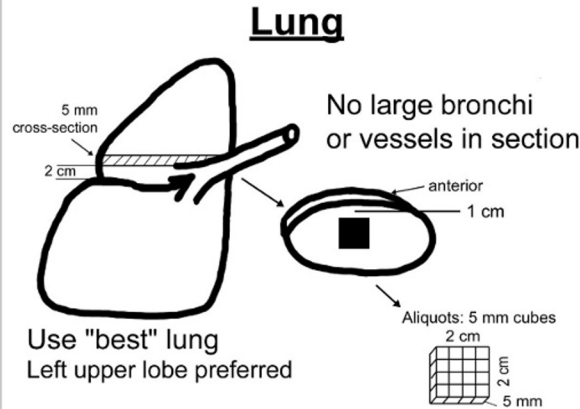
The aliquot size depends upon the organ and is specified in the organ-specific sections below.

A ruler or the cutting board marker should be used to measure the aliquot size. It is important to follow the required aliquot size for tissues to ensure that they are properly fixed. Any deviation to the aliquot size should be documented on the **GTEx Tissue Recovery Case Report Form, PM-0003-F5**. This should be done by noting the deviation in the "comment" field.

##### 4.3.5.1 Preferred Aliquot Size

In general, contiguous aliquots should be obtained per organ/tissue site.

- 4.3.5.1.1 For tissue to be preserved in the PAXgene® Tissue fixative, the preferred aliquot size is 10 mm x 10 mm x  $\leq 4$  mm; two aliquots per cassette; one cassette for histology (CBR) and one cassette for molecular studies (LDACC). The preferred thickness range is 3 to 4 mm.



# GTEEx SOPs - But limited anatomical details for some



## 4.3.6.18 Kidney

4.3.6.18.1 Preferred Location **Left cortex**

4.3.6.18.2 Preferred Aliquot (**Cortex**): 10 mm x 10 mm x  $\leq 8$  mm slice divided into two 10 mm x 10 mm x  $\leq 4$  mm contiguous aliquots. **If cortex is too thin to obtain an 8 mm thick slice, prepare aliquots from a 20 mm x 10 mm x  $\leq 4$  mm thick slice, divided evenly across the long (20 mm) axis.** Each cassette should contain two 10 mm x 10 mm x  $\leq 4$  mm aliquots.

**But many had no visual reference, and/or were limited in anatomical definitions of where to sample.**

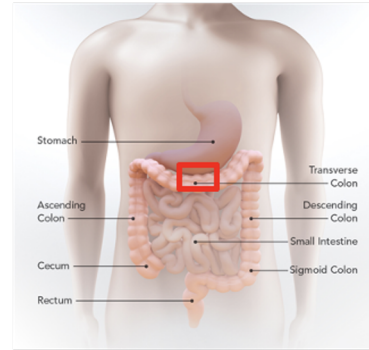
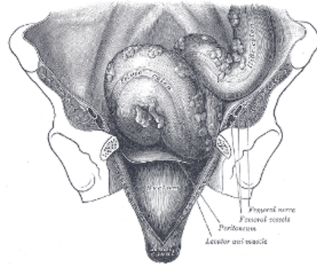
# Example: Mapping GTEx Colon to HuBMAP RUI

 <b>NATIONAL CANCER INSTITUTE</b> 		<b>GTEx Tissue Harvesting Work Instruction</b>	
PR-0004-W1	VER. 03.05	Effective Date: mm/dd/yyyy	Page 13 of 21

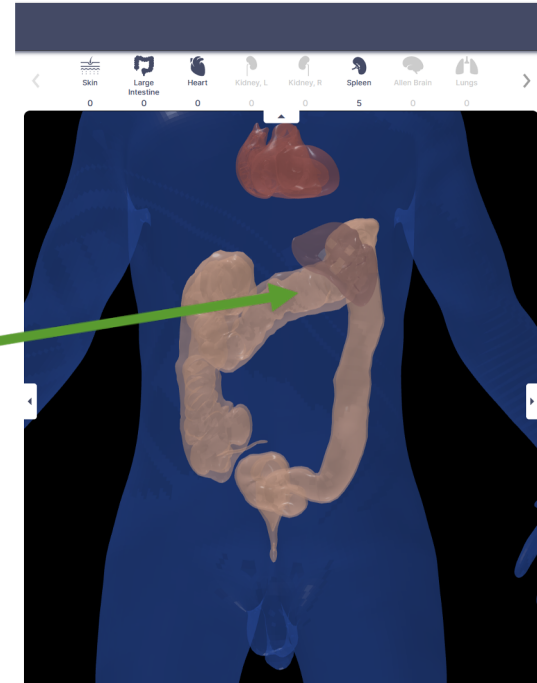
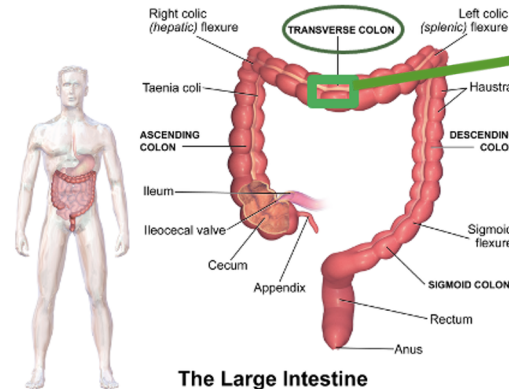
## 4.3.6.20 Colon

- 4.3.6.20.1 **Preferred Location: Transverse colon.** Gently rinse mucosa with normal saline before aliquot preparation. Aliquots should contain the full thickness of the colonic wall, i.e., **mucosa and muscularis propria. Trim adjacent adipose tissue.**
- 4.3.6.20.2 **Preferred Aliquot: 20 mm x 10 mm x thickness (≤4 mm),** divided into two adjacent 10 mm x 10 mm x thickness aliquots. Each cassette should contain two 10 mm x 10 mm x thickness aliquots.
- 4.3.6.20.3 **Preferred Location: Sigmoid colon.** Preferred Location: Sigmoid colon. Gently rinse mucosa with normal saline before aliquot preparation. **Obtain only muscularis propria;** discard mucosa and any serosal adipose tissue.
- 4.3.6.20.4 **Preferred Aliquot: 20 mm x 10 mm x thickness (≤4 mm),** divided into two adjacent 10 mm x 10 mm x thickness aliquots. Each cassette should contain two 10 mm x 10 mm x thickness aliquots.

Sigmoid Colon ('pelvic colon') Dissection Guide (Diagram 4)



Recover the transverse colon starting 10 cm back from the right colic (hepatic) flexure.





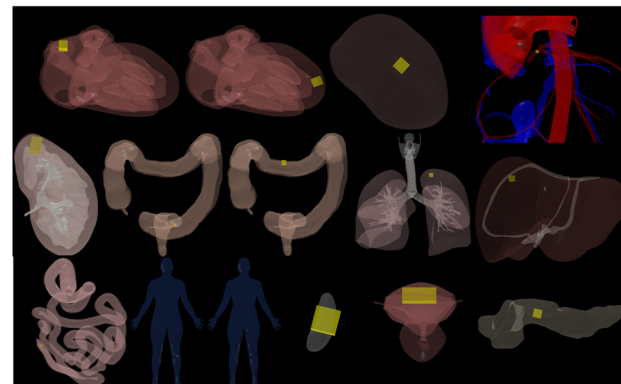
# CCF Registration User Interface (RUI) - GTEx samples

As of March 2022, the RUI was used to capture 29 extraction sites. RUI usage is well-defined via an [SOP](#) and [video demo](#). These extraction sites can be associated with **4,951** tissue blocks (3,095 male/1,856 female) and GTEx summary information, such as eQTLs.

<b>Heart Atrial Appendage</b>	253 Blocks Male   119 Blocks Female
<b>Heart Left Ventricle</b>	253 Male   122 Female
<b>Spleen</b>	141 Male   86 Female
<b>Kidney Cortex Male</b>	55 Male   18 Female
<b>Colon Sigmoid</b>	55 Male   113 Female
<b>Colon Transverse</b>	232 Male   136 Female
<b>Lung</b>	349 Male   166 Female
<b>Liver</b>	161 Male   65 Female
<b>Ovary</b>	180 Female
<b>Pancreas</b>	207 Male   121 Female
<b>Skin (not sun exposed)</b>	411 Male   193 Female
<b>Skin (sun exposed)</b>	467 Male   234 Female
<b>Small Intestine</b>	120 Male   67 Female
<b>Uterus</b>	142 Female
<b>Coronary Artery</b>	146 Male   94 Female
<b>Prostate</b>	245 Male



Extraction sites (male)



Extraction sites (female)

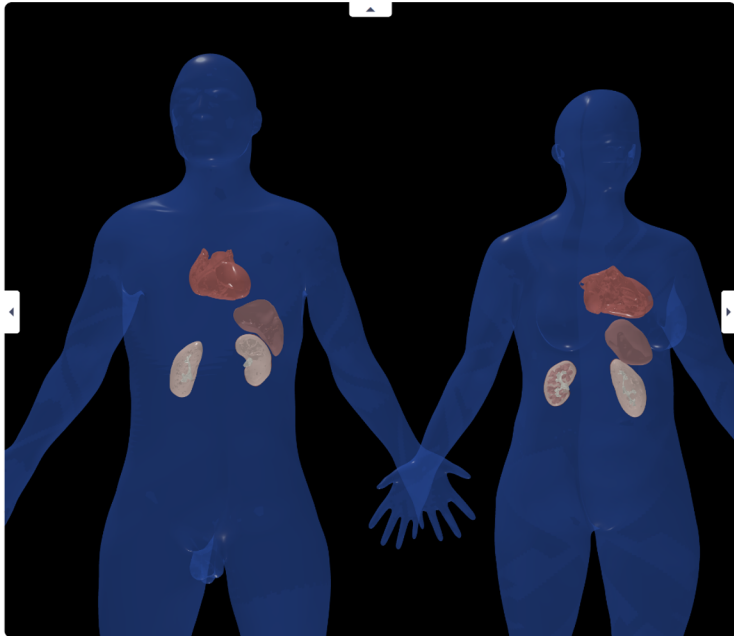
# CCF Exploration User Interface (EUI)

**HuBMAP CCF EXPLORATION** LOGIN

Sex: Both Age: 1-110 BMI: 13-83

Search ontology terms ...

- body 29
  - blood vasculature 2
  - brain 0
  - eye 0
  - fallopian tube 0
  - heart 4
  - kidney 2
  - knee 0
  - large intestine 4
  - liver 2
  - lung 2
  - lymph node 0
  - ovary 1
  - pancreas 2
  - pelvis 0
  - prostate gland 1
  - skin 5
  - small intestine 1
  - spleen 2
  - thymus 0
  - ureter 0
  - urinary bladder 0
  - uterus 1



1 Tissue Data Providers  
29 Donors  
29 Tissue Blocks  
0 Tissue Sections  
29 Tissue Datasets

- Females (n=94), Mean Age 53.5 (range 21-7...  
Entered 1/26/2022, Kristin Ardlie, GTEx Project
- Males (n=146), Mean Age 54.3 (range 20-70)  
Entered 1/26/2022, Kristin Ardlie, GTEx Project
- Females (n=113), Mean Age 51.1 (range 21-7...  
Entered 5/18/2021, Kristin Ardlie, GTEx Project
- Males (n=55), Mean Age 55.8 (range 27-70)  
Entered 5/18/2021, Kristin Ardlie, GTEx Project
- Females (n=136), Mean Age 48.9 (range 21-...  
Entered 5/18/2021, Kristin Ardlie, GTEx Project
- Males (n=232), Mean Age 50.1 (range 21-70)  
Entered 5/18/2021, Kristin Ardlie, GTEx Project
- Females (n=119), Mean Age 54.7 (range 21-...  
Entered 5/18/2021, Kristin Ardlie, GTEx Project
- Males (n=253), Mean Age 55.6 (range 20-70)  
Entered 5/18/2021, Kristin Ardlie, GTEx Project
- Females (n=122), Mean Age 52.9 (range 21-...  
Entered 5/18/2021, Kristin Ardlie, GTEx Project
- Males (n=253), Mean Age 55.6 (range 20-70)  
Entered 5/18/2021, Kristin Ardlie, GTEx Project
- Females (n=18), Mean Age 56.8 (range 30-6...  
Entered 9/17/2021, Kristin Ardlie, GTEx Project
- Males (n=55), Mean Age 55.8 (range 27-70)

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



# GTE<sub>x</sub>-API

- OpenAPI [spec](#) and [Implementation](#), on GTE<sub>x</sub> side to provide data to Hubmap
- Have designed the API, added the data to our database, and implemented an interim API.
- Are working to implement a revised FAIR version of the API this month.

# CCF-API

- OpenAPI [spec](#) and [Implementation](#), allowing querying of LOD data registered to the CCF
- HTTP API implementation is live and in production
- JavaScript and Python APIs becoming available in March 2022

CCF-API	
Overview	
<b>ENDPOINTS</b>	
Get Tissue Block Results	GET
Get technology names (for filtering)	GET
Get tissue provider names (for filtering)	GET
Get aggregate results / statistics	GET
Get number of ontology term occurren...	GET
Get ontology term tree nodes	GET
Get all reference organs	GET
Get all nodes to form the 3D scene of r...	GET
Get all nodes to form the 3D scene for ...	GET
Get all hubmap rui locations (if enabled)	GET
<b>SCHEMAS</b>	
NumberRange	🔗
CommaSeparatedStrings	🔗
TissueBlockResult	🔗
DonorResult	🔗
DatasetResult	🔗
TissueSectionResult	🔗
AggregateResult	🔗
TermOccurrences	🔗
OntologyTreeModel	🔗
OntologyTreeNode	🔗
SpatialEntity	🔗
SpatialObjectReference	🔗
SpatialSceneNode	🔗

🔗 powered by **Stoplight**

## CCF-API

Export ▾

v1.0.0

API Base URL

**CCF-API:** <https://ccf-api.hubmapconsortium.org/v1>

**Production Server (backup):** <https://ccf-api.herokuapp.com/v1>

**Staging Server:** <https://ccf-api-staging.herokuapp.com/v1>

This API provides programmatic access to data registered to the CCF. See the [HuBMAP CCF Portal](#) for details.

# GTE<sub>x</sub> / HuBMAP FAIR API Workspace Pilot



The Common Fund

Overall status	On Track	Project Manager	Lan Nguyen, Nancy Ruschman	Reporting Period	10/01 - 12/31
----------------	----------	-----------------	----------------------------	------------------	---------------

Progress during Current Reporting Period
<ul style="list-style-type: none"> <li>HuBMAP GTE<sub>x</sub> Pilot released, including the OpenAPI specification for the CCF-API. Website with deliverables: <a href="https://hubmapconsortium.github.io/ccf-gtex-pilot/">https://hubmapconsortium.github.io/ccf-gtex-pilot/</a></li> <li>CCF Release v1.1 adding many new reference organs and ASCT+B tables CCF v1.1 Release: <a href="https://hubmapconsortium.github.io/ccf-releases/v1.1/docs/index.html">https://hubmapconsortium.github.io/ccf-releases/v1.1/docs/index.html</a></li> <li>CCF Portal: ASCT+B - <a href="https://hubmapconsortium.github.io/ccf/pages/ccf-anatomical-structures.html">https://hubmapconsortium.github.io/ccf/pages/ccf-anatomical-structures.html</a></li> <li>Updated CCF Portal: Reference Organs - <a href="https://hubmapconsortium.github.io/ccf/pages/ccf-3d-reference-library.html">https://hubmapconsortium.github.io/ccf/pages/ccf-3d-reference-library.html</a></li> </ul>

Priorities for Next Reporting Period
<ul style="list-style-type: none"> <li>Continue working on Table Comparison feature using GTE<sub>x</sub> RNA and snRNAseq data for AS, CT +B table</li> <li>Implement libraries for using the FAIR CCF-API from cloud workspaces. Visualize GTE<sub>x</sub> data within EUI on HuBMAP Portal and GTE<sub>x</sub> Portal. Showcase cross-search of AS, CT, B data in support of increased interoperability and FAIR data access                             <ul style="list-style-type: none"> <li>Implement CCF-API client libraries in Python and JavaScript for use in most cloud workspaces</li> <li>Deploy EUI as a web component compatible with most cloud workspaces (complete)</li> <li>Integrate EUI into GTE<sub>x</sub> Portal</li> <li>Demonstrate cross-search of AS, CT, B data in support of increased interoperability</li> </ul> </li> <li>Develop training materials of cloud setup in support of wider usage and adoption of the pilot cloud setup and CCF cross-search and run user survey                             <ul style="list-style-type: none"> <li>Publish documentation for adopting the FAIR CCF-API for data publication to the cloud, data querying from cloud workspaces, and EUI integration into portals/cloud workspaces. (complete)</li> </ul> </li> </ul>

### Key Milestones, Oct - Dec 2021

Milestone	Due Date	Status
Design cloud workspace compatible FAIR APIs. Extend the CCF to capture GTE <sub>x</sub> data. Use the RUI to map GTE <sub>x</sub> tissue data onto the HuBMAP CCF	12/31/2021	Completed
Extend CCF to capture Anatomical Structures, Cell Types, plus Biomarkers (ASCT+B) information in GTE <sub>x</sub> anatomogram data	12/31/2021	Completed
Spatially register GTE <sub>x</sub> tissue samples that currently overlap with HuBMAP using the CCF Registration User Interface	12/31/2021	Completed
Implement and deploy the FAIR CCF-API for HuBMAP data in the HuBMAP cloud	12/31/2021	Completed
Validate CCF ASCT+B tables using GTE <sub>x</sub> RNA and snRNAseq data and revising the tables as needed	03/31/2022	On Track

Risks, Issues and Corrective Action Strategies		
Risk/Issue Description	Severity	Mitigation Strategy

Dependencies		
Description	Functional Team	Owner(s)