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

Kristin Ardlie (GTEx) and Katy Börner (HuBMAP)

# 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).

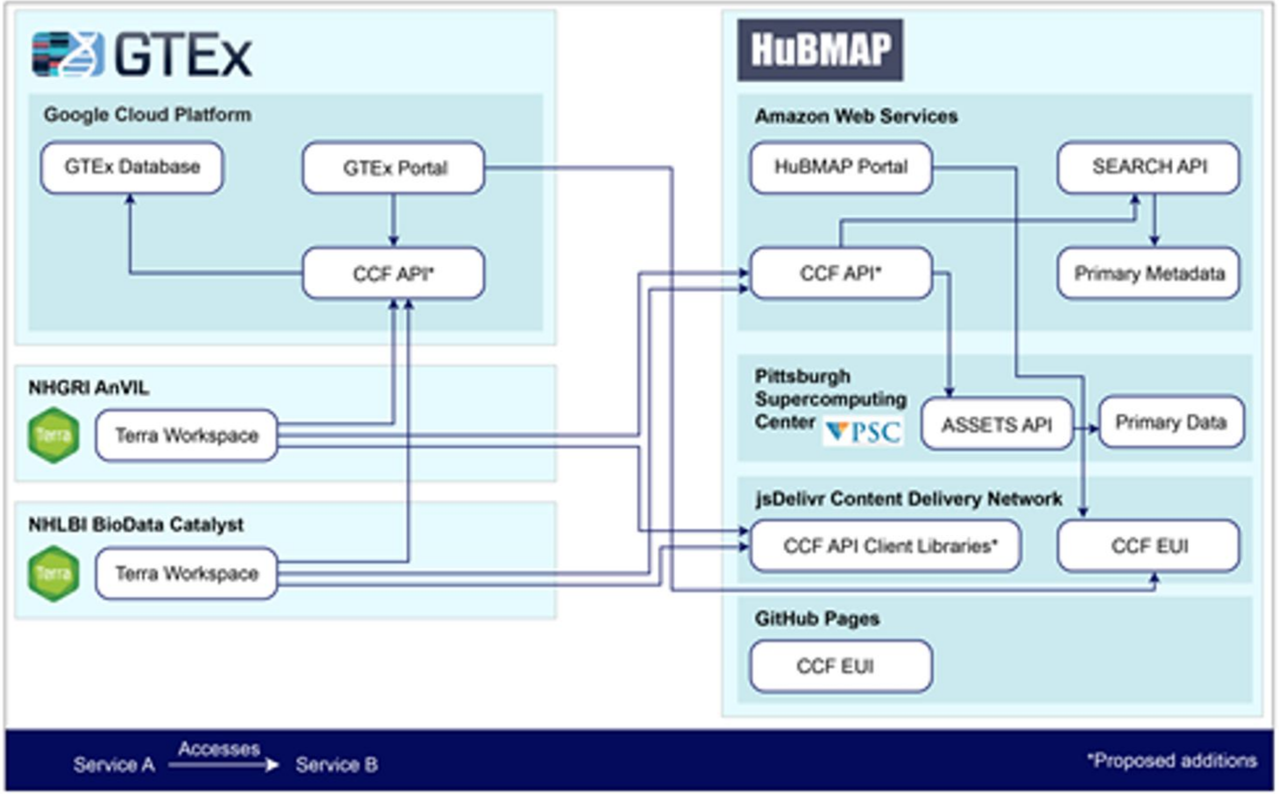
# Approach

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

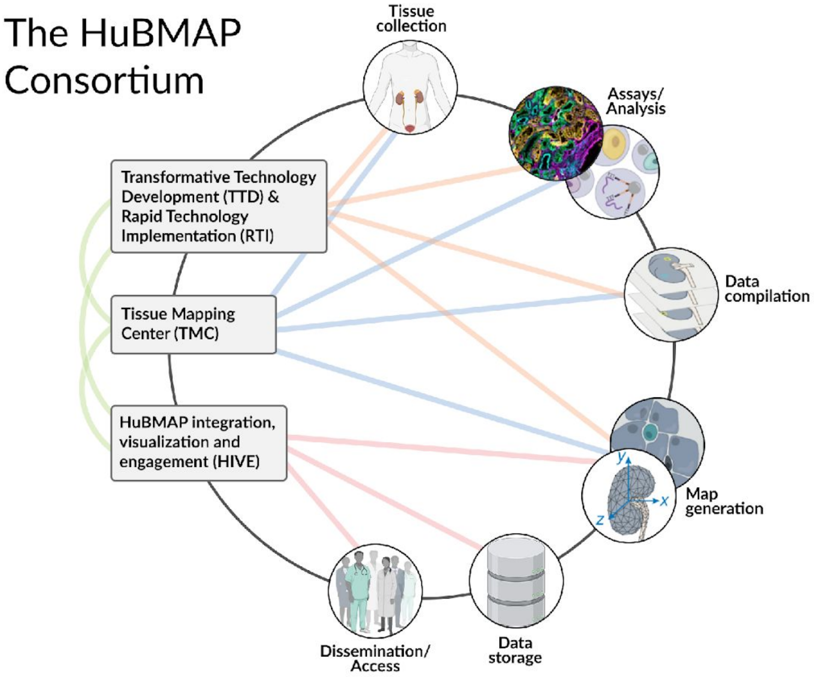
Develop libraries to enable the use of FAIR CCF-API for visualization of GTEx data in HuBMAP or GTEX portal and cross-search for ASCT+B indexed data across cloud platforms.

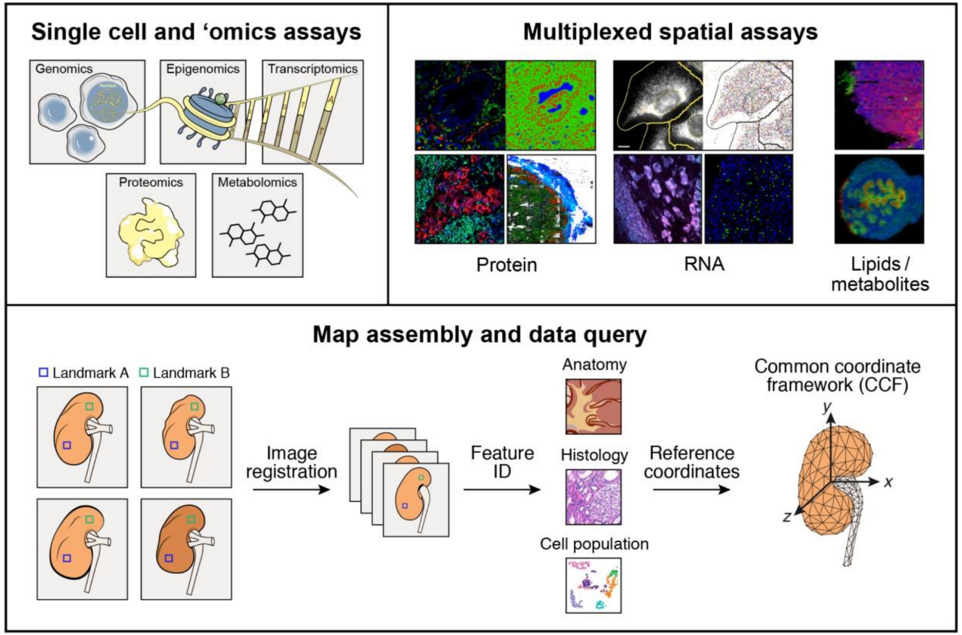
Develop training materials and conduct a user survey to support wider usage and adoption of the pilot cloud setup and CCF cross-search.

# System Architecture



**The Human Body at Cellular Resolution: The NIH Human Biomolecular Atlas Program.**
Snyder et al. *Nature*. 574, p. 187-192.



# 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 (glomular) Capsule/parietal layer | Parietal epithelial Cell |
| | Bowman's (glomular) 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 |
| | 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 |
| | | TAL-Macula Densa Cell |
| | Distal Convolution | Distal Convoluted Tubule Cell (general) |
| | | DCT Type 1 Cell |
| | | DCT Type 2 Cell |
| | Connecting Tubule | Connecting Tubule Cell (general) |
| | | CNT-Principal Cell |

Ontology

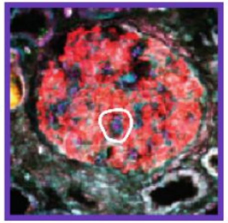
*Anatomical Structures Partonomy*

kidney
- kidney capsule
- cortex of kidney
  - outer cortex of kidney
- renal medulla

*Cell Types Ontology*

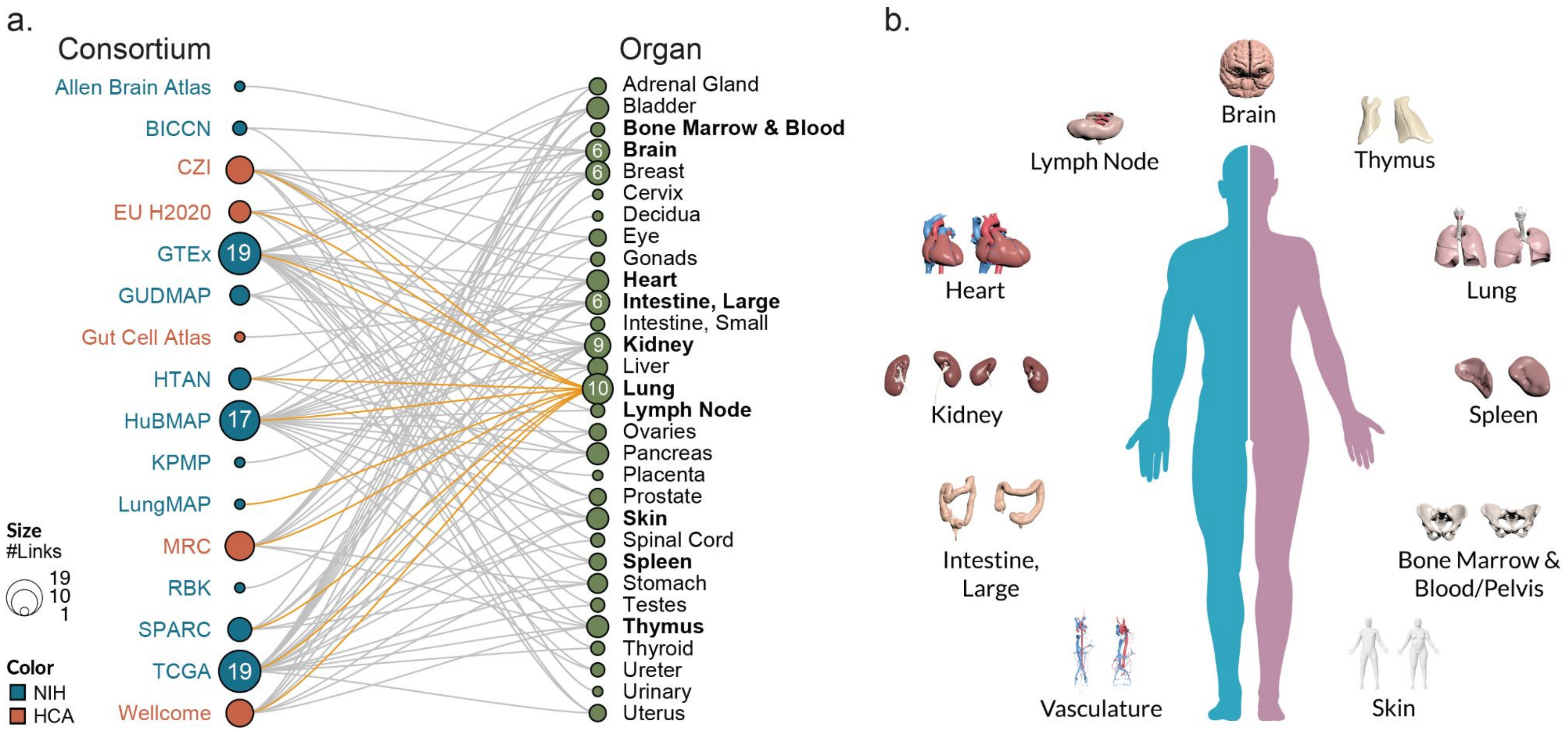
connective tissue cell
- pericyle cell
  - mesangial cell
    - extraglomerular mesangial cell
    - glomerular mesangial cell

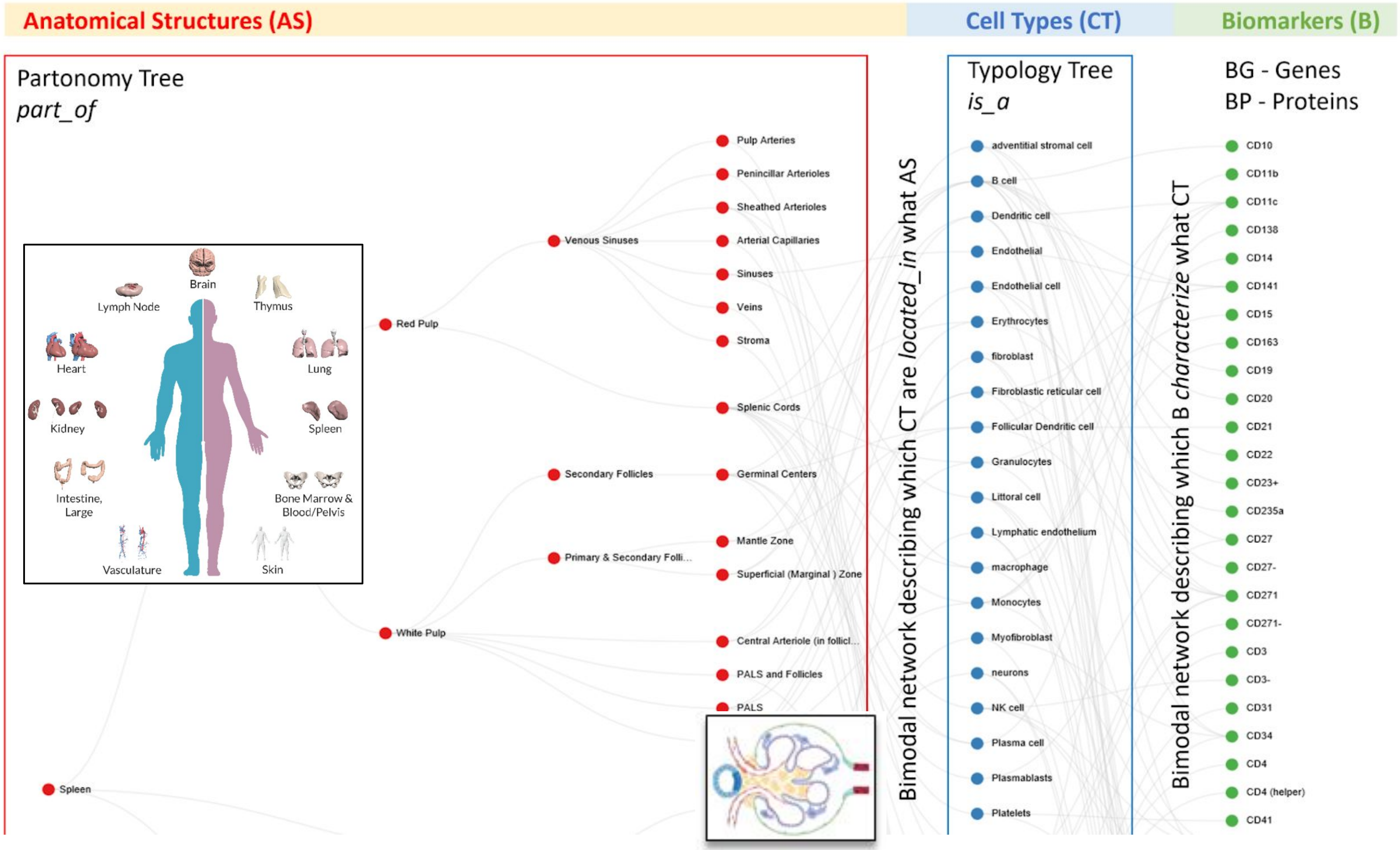
3D Reference Object Library



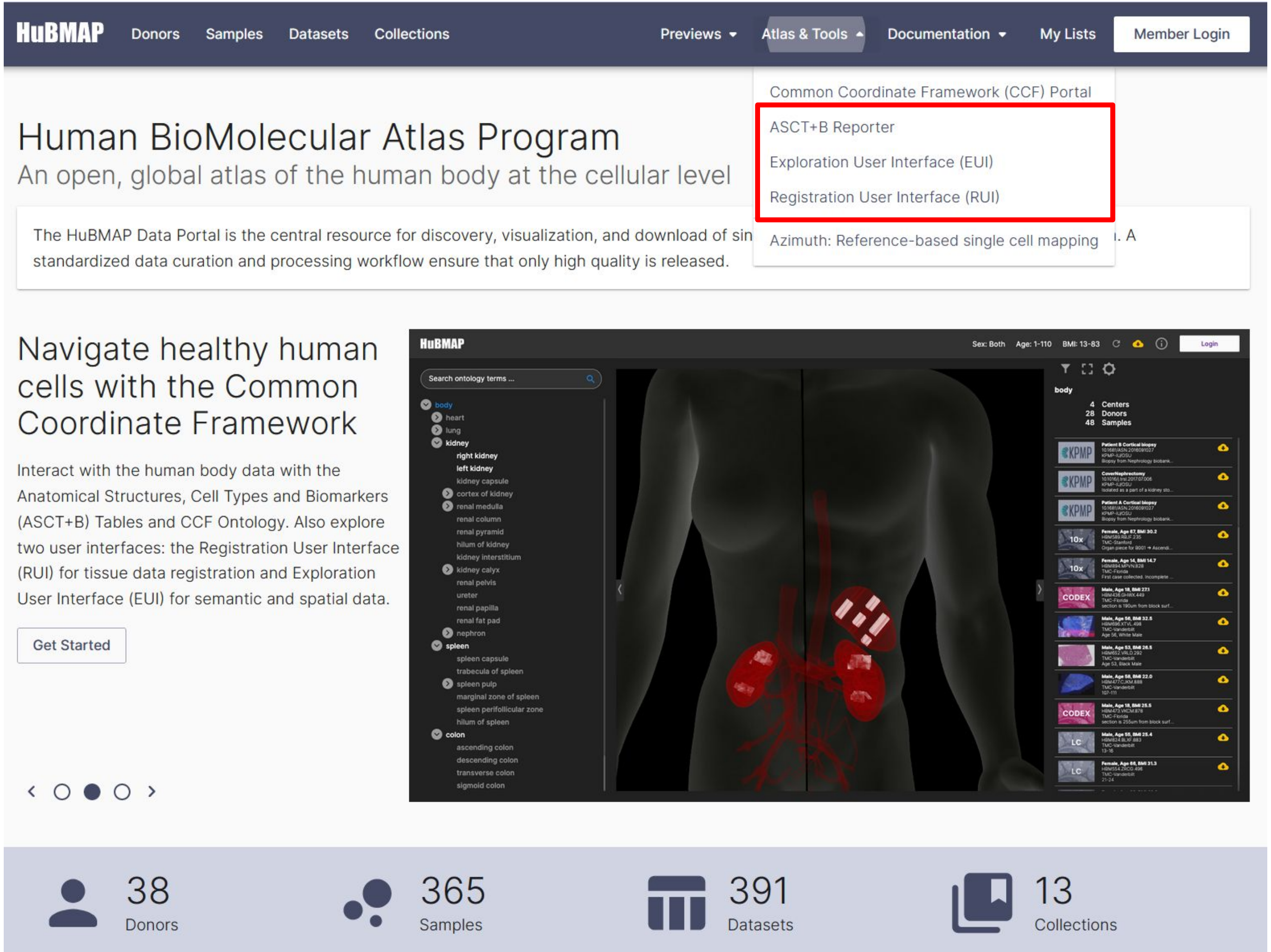
https://www.biorxiv.org/content/10.1101/2021.05.31.446440v2

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

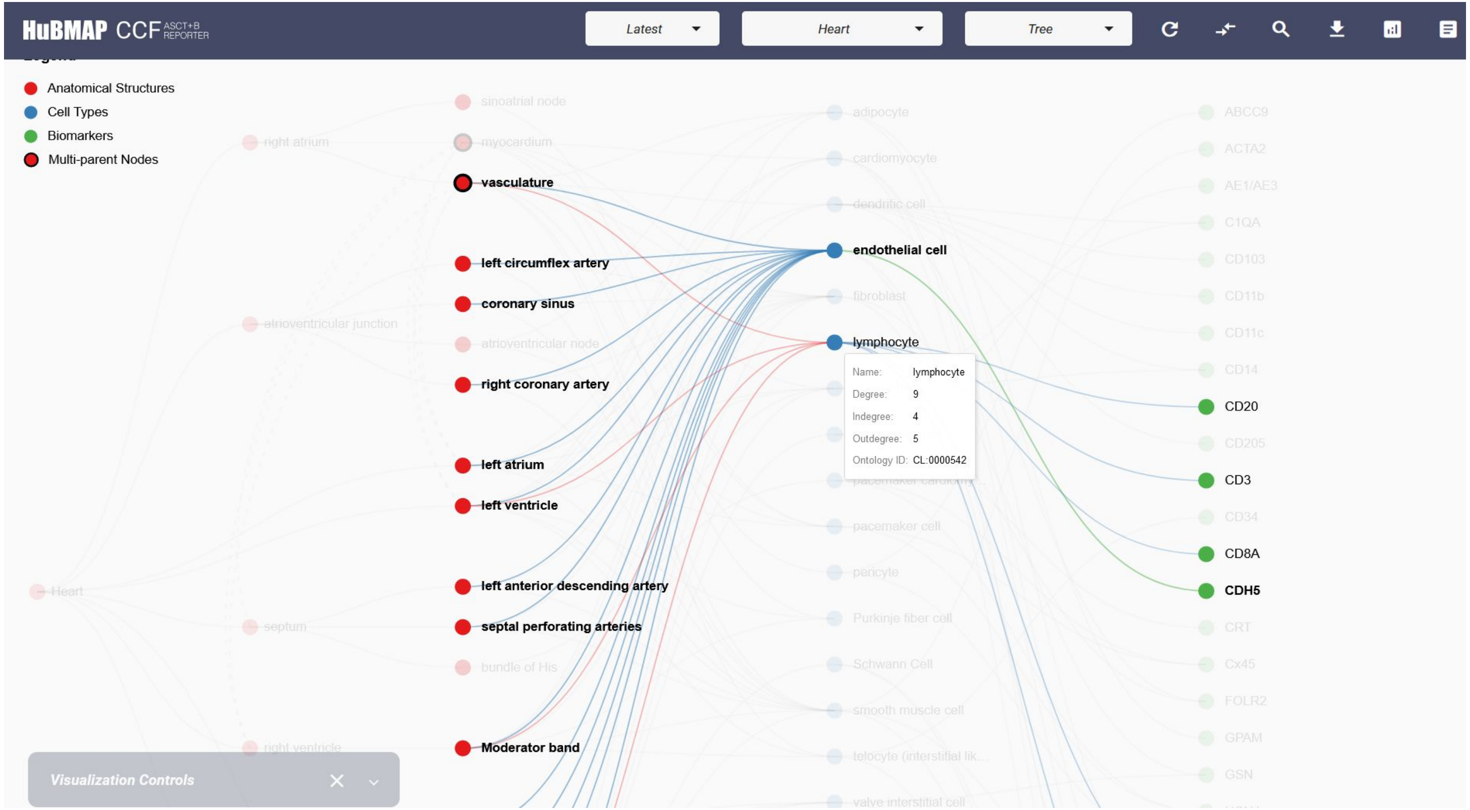



Anatomical Structures (AS)
Cell Types (CT)
Biomarkers (B)
Partonomy Tree
part_of
Brain
Lymph Node
Thymus
Heart
Lung
Kidney
Spleen
Intestine, Large
Bone Marrow & Blood/Pelvis
Vasculature
Skin
Spleen
Red Pulp
White Pulp
Venous Sinuses
Secondary Follicles
Primary & Secondary Folli...
Pulp Arteries
Penincillar Arterioles
Sheathed Arterioles
Arterial Capillaries
Sinuses
Veins
Stroma
Splenic Cords
Germinal Centers
Mantle Zone
Superficial (Marginal ) Zone
Central Arteriole (in follicl...
PALS and Follicles
PALS
Bimodal network describing which CT are located_in what AS
Typology Tree
is_a
adventitial stromal cell
B cell
Dendritic cell
Endothelial
Endothelial cell
Erythrocytes
fibroblast
Fibroblastic reticular cell
Follicular Dendritic cell
Granulocytes
Littoral cell
Lymphatic endothelium
macrophage
Monocytes
Myofibroblast
neurons
NK cell
Plasma cell
Plasmablasts
Platelets
Bimodal network describing which B characterize what CT
BG - Genes
BP - Proteins
CD10
CD11b
CD11c
CD138
CD14
CD141
CD15
CD163
CD19
CD20
CD21
CD22
CD23+
CD235a
CD27
CD27-
CD271
CD271-
CD3
CD3-
CD31
CD34
CD4
CD4 (helper)
CD41

# Atlas & Tools


HuBMAP
Donors
Samples
Datasets
Collections
Previews
Atlas & Tools
Documentation
My Lists
Member Login
Common Coordinate Framework (CCF) Portal
ASCT+B Reporter
Exploration User Interface (EUI)
Registration User Interface (RUI)
Azimuth: Reference-based single cell mapping
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 sin
. A
standardized data curation and processing workflow ensure that only high quality is released.
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
Sex: 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
trabecula 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 Centers
28 Donors
48 Samples
38
Donors
365
Samples
391
Datasets
13
Collections

# CCF ASCT+B Reporter UI

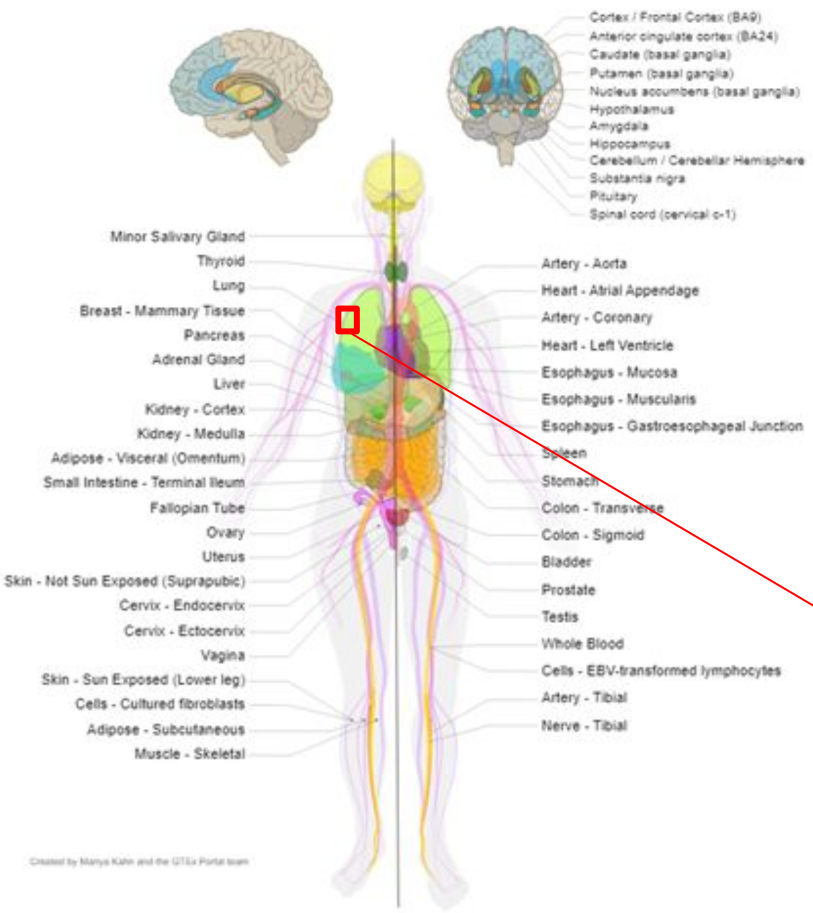


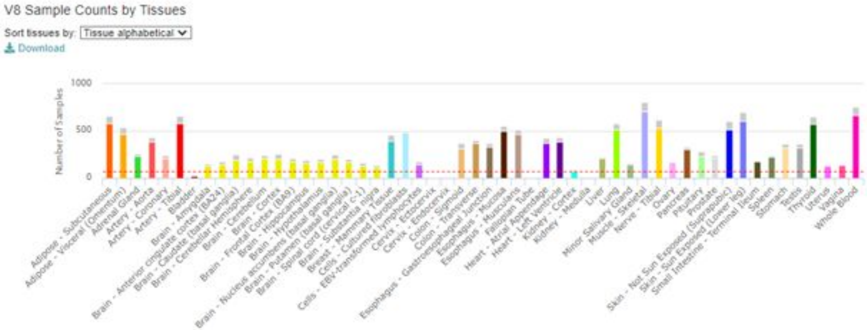

https://hubmapconsortium.github.io/ccf-asct-reporter/

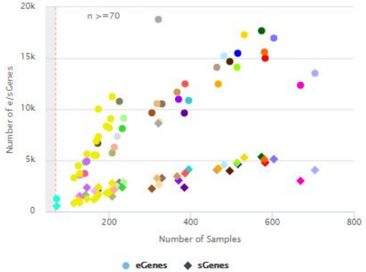
# CCF Registration User Interface (RUI)

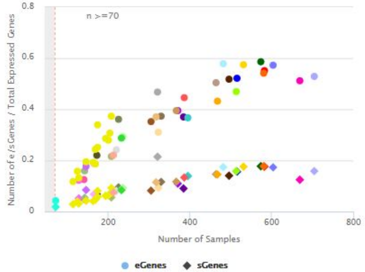

https://hubmapconsortium.github.io/ccf-ui/rui/

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

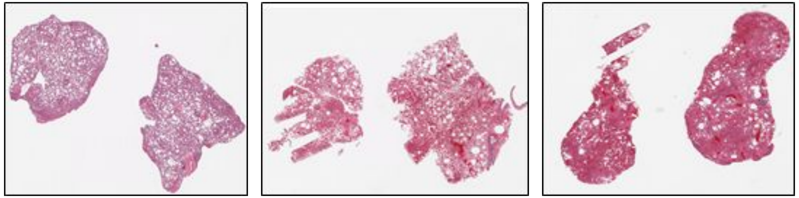



**LUNG**



**GTEx Consortium** (2020) Science 369:1318

# GTEx SOPs designed for repeatable anatomical sampling

| NIH NATIONAL CANCER INSTITUTE<br>BBRB Biorepositories and Biospecimen Research Branch | | GTEx Tissue Harvesting Work Instruction | |
|---|---|---|---|
| PR-0004-W1 | VER. 03.05 | Effective Date: mm/dd/yyyy | Page 2 of 21 |

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 ≥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**

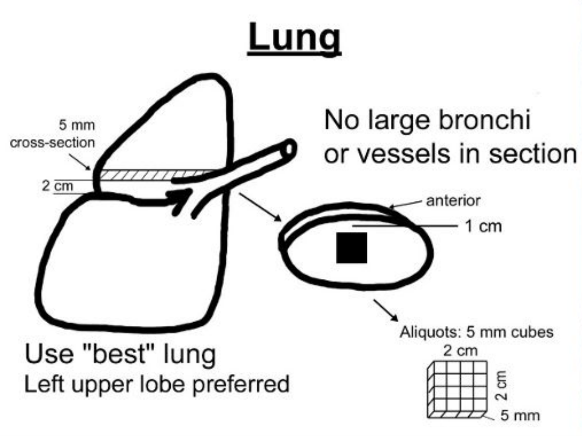
The aliquot size depends upon the organ and is specified in the organ-specific sections below.

A ruler or the cutting board marker should be used to measure the aliquot size. It is important to follow the required aliquote 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 ≤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.




# GTEx 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 ≤8 mm slice divided into two 10 mm x 10 mm x ≤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 ≤4 mm thick slice, divided evenly across the long (20 mm) axis.** Each cassette should contain two 10 mm x 10 mm x ≤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

| NIH NATIONAL CANCER INSTITUTE BBRB Biorepositories and Biospecimen Research Branch | | GTEx Tissue Harvesting Work Instruction | |
|---|---|---|---|
| PR-0004-W1 | VER. 03.05 | Effective Date: mm/dd/yyyy | Page 13 of 21 |

4.3.6.20 **Colon**

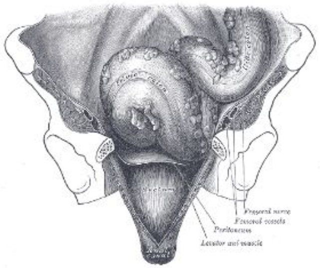
4.3.6.20.1 **Preferred Location: Transverse colon.** Gently rinse mucosa with normal saline before aliquot preparation. Aliquots should contain the full thickness of the colonic wall, i.e., **mucosa and muscularis propria. Trim adjacent adipose tissue.**

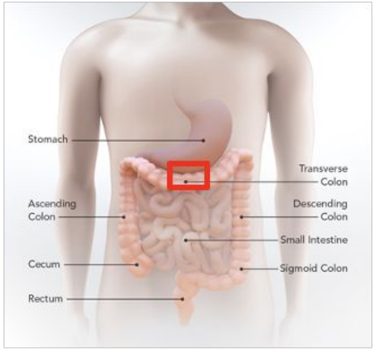
4.3.6.20.2 **Preferred Aliquot:** 20 mm x 10 mm x thickness (≤4 mm), divided into two adjacent 10 mm x 10 mm x thickness aliquots. Each cassette should contain two 10 mm x 10 mm x thickness aliquots.

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.

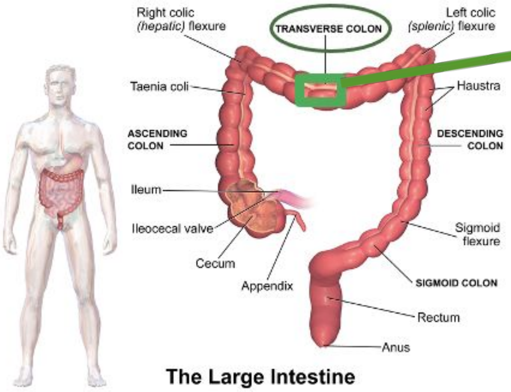
4.3.6.20.4 **Preferred Aliquot:** 20 mm x 10 mm x thickness (≤4 mm), divided into two adjacent 10 mm x 10 mm x thickness aliquots. Each cassette should contain two 10 mm x 10 mm x thickness aliquots.

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

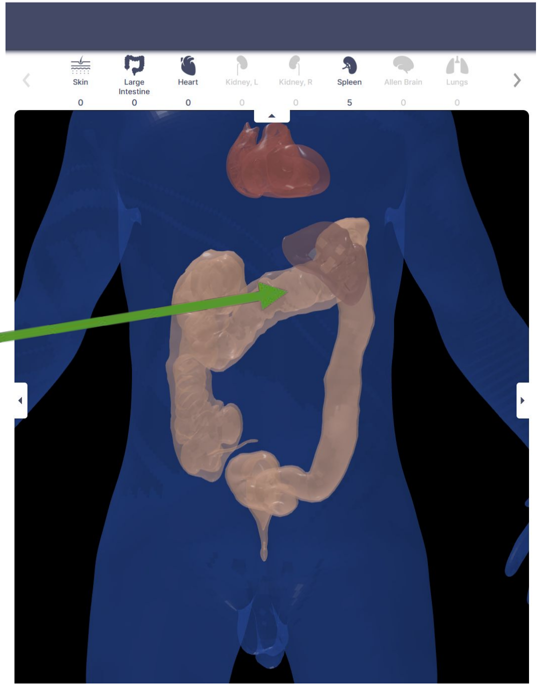





Recover the transverse colon starting 10 cm back from the right colic (hepatic) flexure.




The Large Intestine



# CCF Registration User Interface (RUI) - GTEx samples

The RUI was used to capture 14 extraction sites. RUI usage is well-defined via an SOP and video demo. These extraction sites can be associated with up to **2,433** tissue blocks (and GTEx summary information, such as eQTLs).

**Heart Atrial Appendage** | 253 Blocks Male | 119 Blocks Female

**Heart Left Ventricle** | 264 Blocks Male | 122 Blocks Female

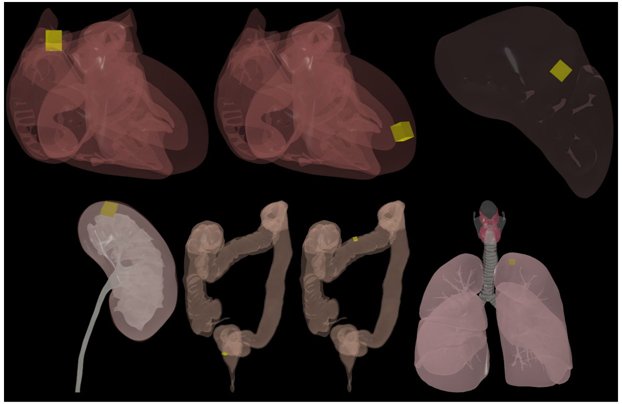
**Spleen** | 141 Male | Female 86 Female

**Kidney Cortex Male** | 55 Male | 18 Female

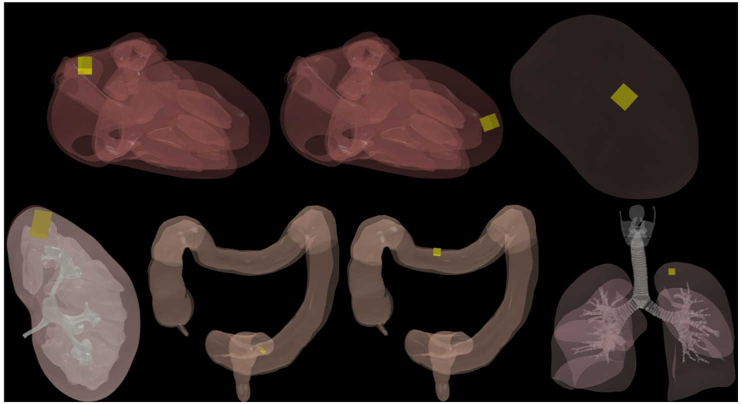
**Colon Sigmoid** | 205 Male | 113 Female

**Colon Transverse** | 232 Male | 136 Female

**Lung** | Male + Female 515

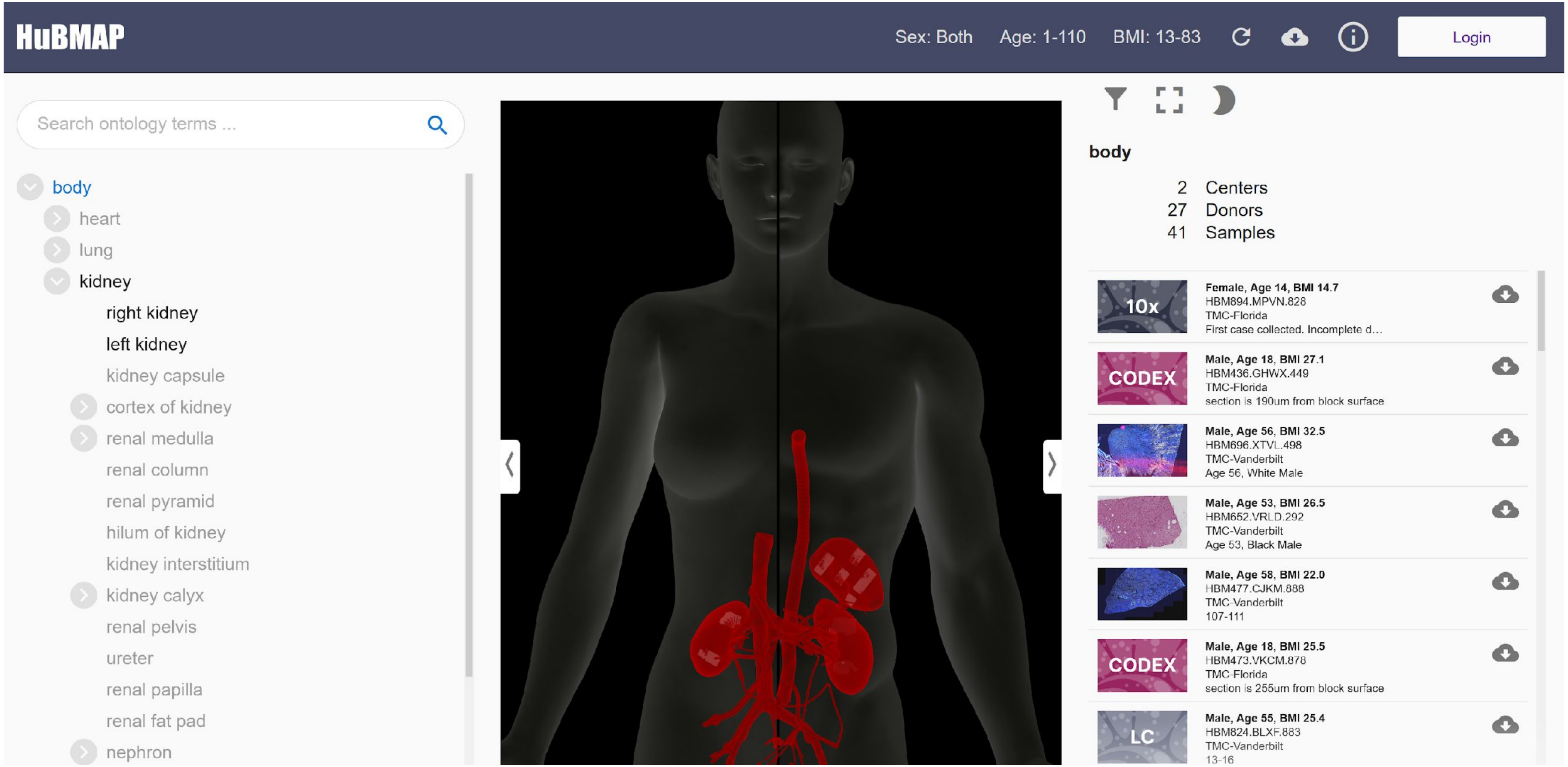


Extraction sites (male)



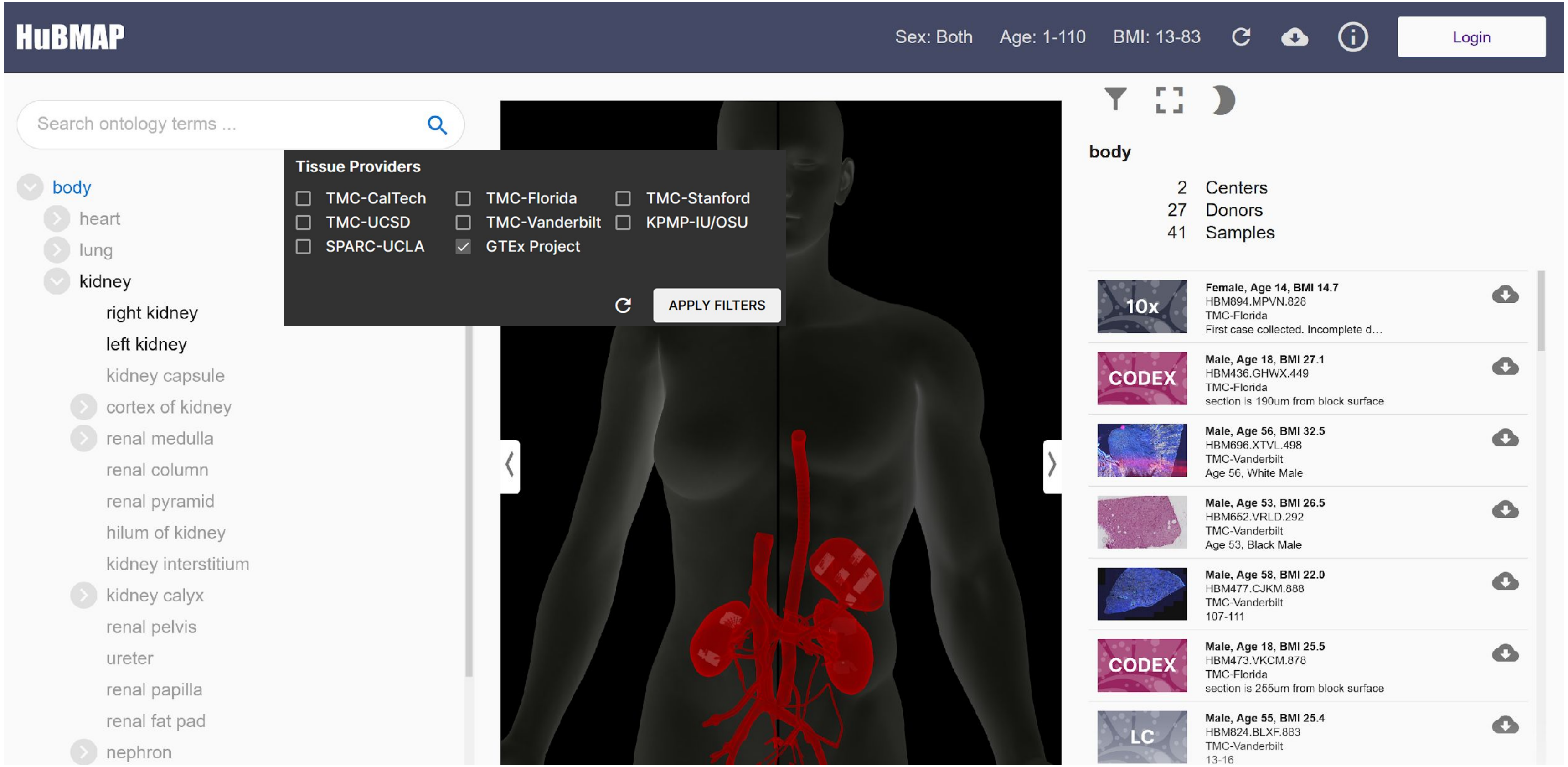
Extraction sites (female)

# CCF Exploration User Interface (EUI)




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

# CCF Exploration User Interface (EUI)




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