



HTAN Data Coordinating Center (DCC)

Ethan Cerami, PhD
Dana-Farber Cancer Institute (DFC)
NCAB Meeting, June 13-15, 2022

**Introduction to
HTAN and DCC**

**Data Availability
and Visualization**

1

2

3

**Data Standards and
Data Integration**

Introduction to HTAN

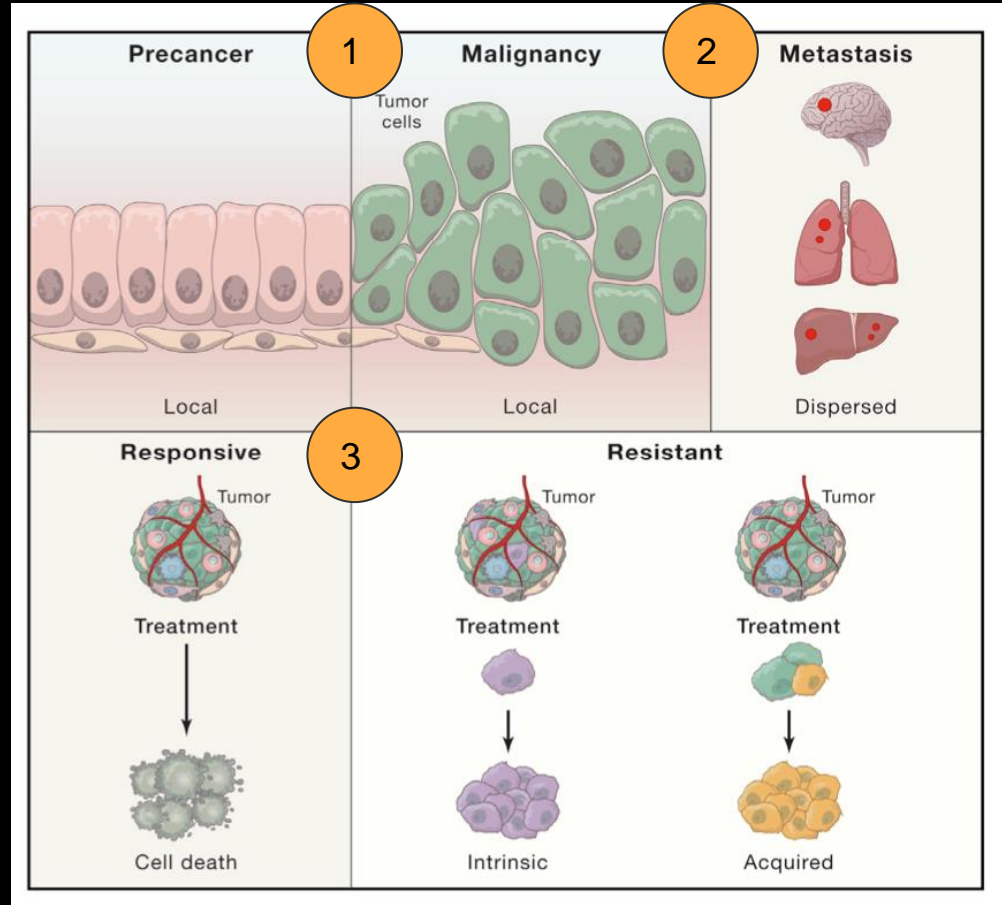
01

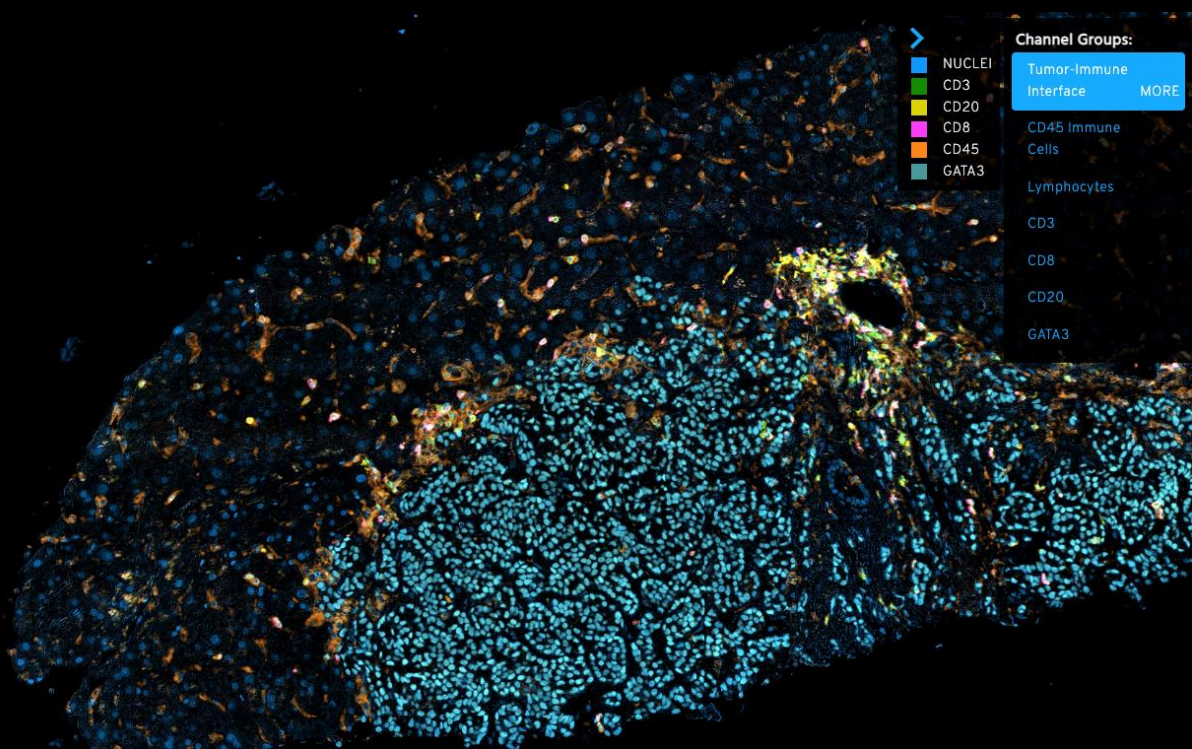
A **collaborative network** of:

- 10 Research Centers (5 Pre-Cancer Atlases + 5 Cancer Atlases)
- 1 pilot Pre-Cancer project (PCAPP)
- 1 pilot Cancer Atlas project (HTAPP)

What is the Human Tumor Atlas Network?

HTAN is focussed on **transitions** in cancer





Lymphocytes in breast cancer liver metastasis biopsy (OHSU HTAN Atlas).

Most HTAN groups have a **1** focus on **single cell** and **multi-plex imaging** modalities

2 Longitudinal Data

The Human Tumor Atlas Network: Charting Tumor Transitions across Space and Time at Single-Cell Resolution

Orit Rozenblatt-Rosen,^{1,35} Aviv Regev,^{1,2,3,35,36,*} Philipp Oberdoerffer,^{4,35} Tal Nawy,^{5,35} Anna Hupalowska,¹ Jennifer E. Rood,¹ Orr Ashenberg,¹ Ethan Cerami,⁶ Robert J. Coffey,⁷ Emek Demir,⁸ Li Ding,⁹ Edward D. Esplin,¹⁰ James M. Ford,^{10,11} Jeremy Goecks,¹² Sharmistha Ghosh,¹³ Joe W. Gray,¹⁴ Justin Guinney,^{15,16} Sean E. Hanlon,¹⁷ Shannon K. Hughes,⁴ E. Shelley Hwang,^{18,19} Christine A. Iacobuzio-Donahue,²⁰ Judit Jané-Valbuena,¹

(Author list continued on next page)

¹Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA

²Howard Hughes Medical Institute, Chevy Chase, MD 20815, USA

³Koch Institute for Integrative Cancer Research, Department of Biology, MIT, Cambridge, MA 02139, USA

⁴Division of Cancer Biology, National Cancer Institute, NIH, Rockville, MD 20850, USA

⁵Computational and Systems Biology Program, Sloan Kettering Institute, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA

⁶Department of Data Sciences, Dana-Farber Cancer Institute, Boston, MA 02215, USA

Cell Paper or go to <https://humantumoratlas.org/>

Dana-Farber Cancer Institute: Ethan Cerami (co-PI)

Sage Bionetworks: James Eddy (co-PI)

Institute for Systems Biology: Vésteinn Thorsson (co-PI)

Memorial Sloan Kettering Cancer Center: Niki Schultz (co-PI)

+ Imaging Collaboration with Peter Sorger, HMS.

DCC Partner Institutions

- 1 Develop **data standards** and enable **data integration**.
- 2 Enable **data sharing** and **data visualization** of all HTAN data.

Data Coordinating Center (DCC)



Clinical



Biospecimens



Molecular Profiling



Imaging

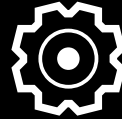


FAIR Data Principles

Findable



Accessible

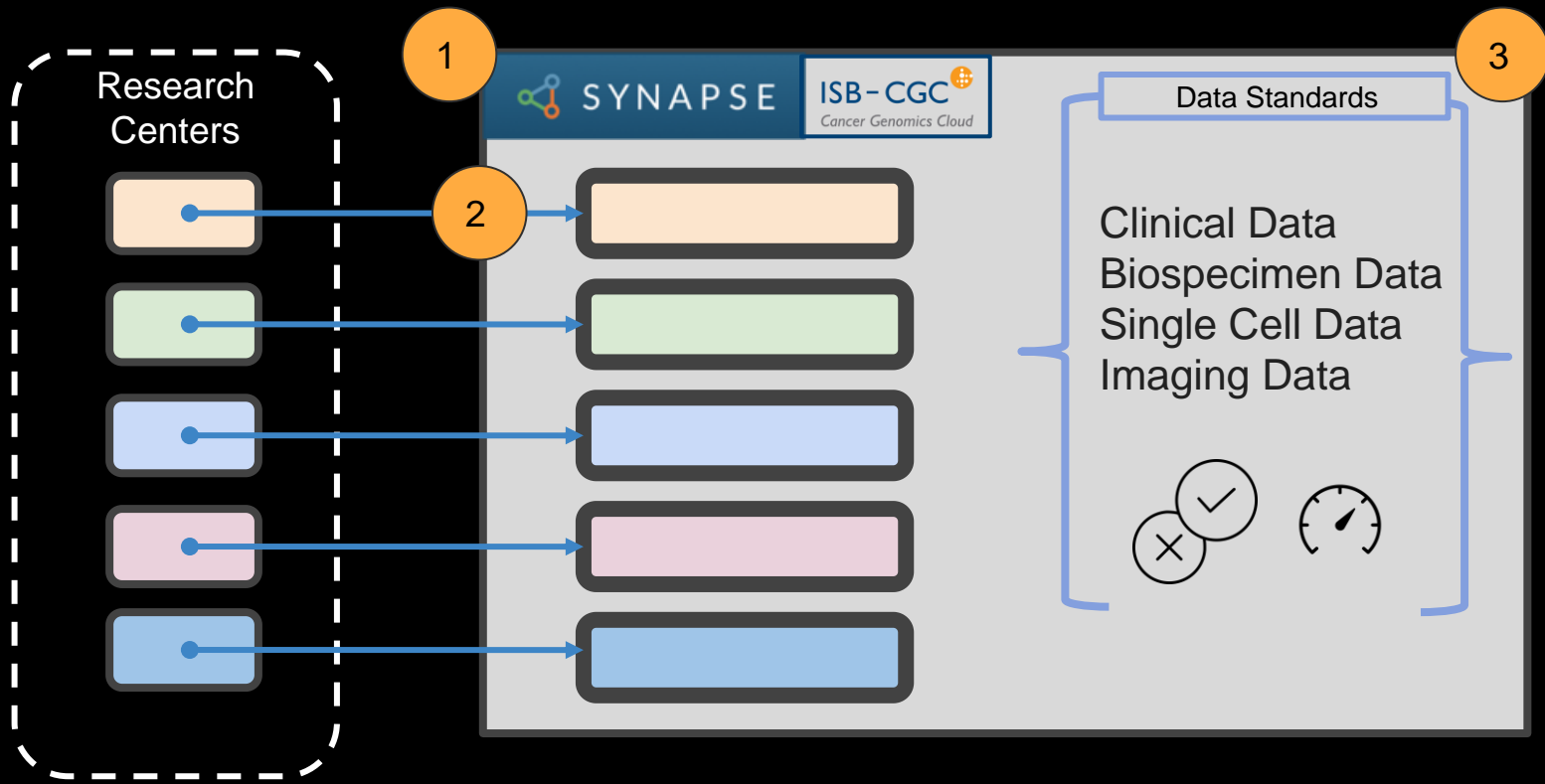


Interoperable

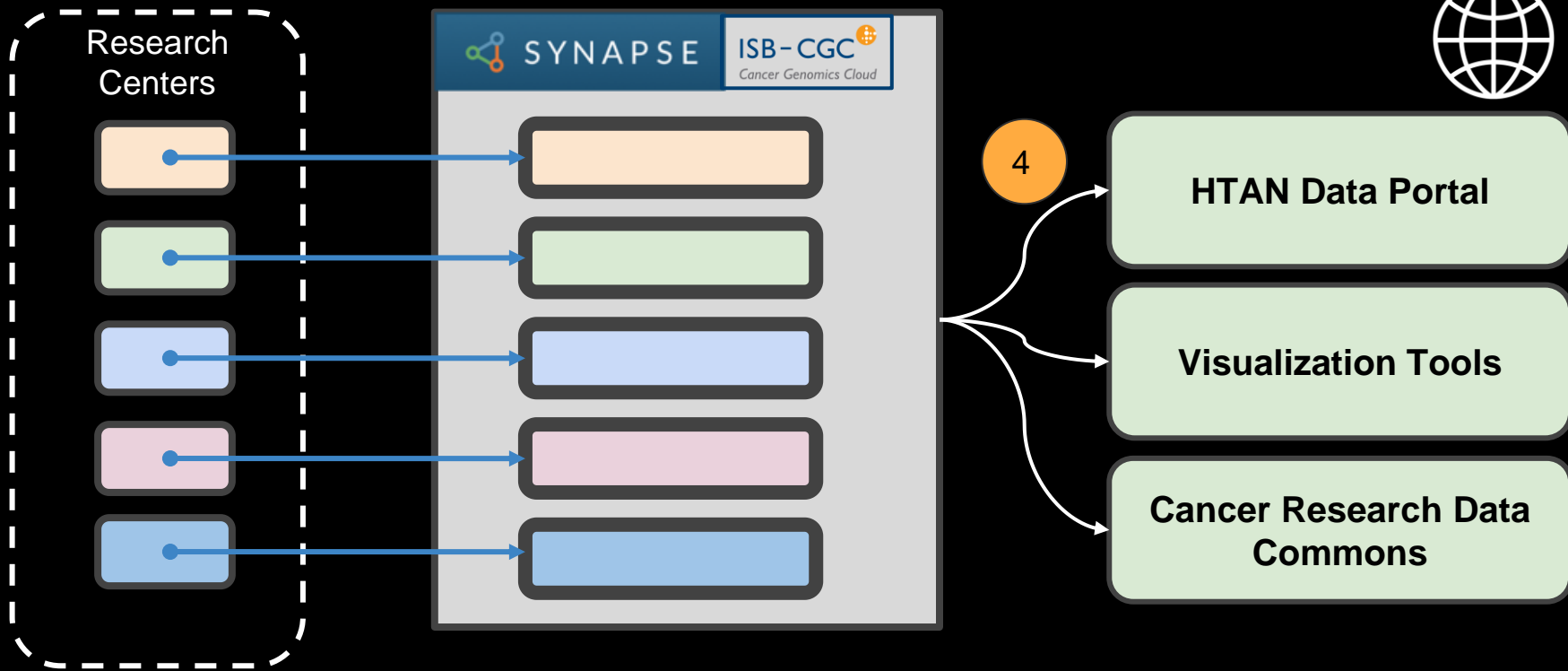


Reusable

What exactly is an Atlas?



HTAN Data Integration



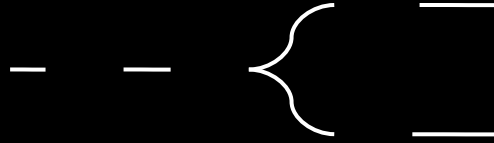
HTAN Data Distribution

Data Standards and Integration

02

1 Community-driven
Request for
Comments

2 Unified schema providing
validation and provenance



3 Data levels:
(Inspired by TGCA)

? HTAN IDs:

HTA9_04_3329



HTAN Data Model Process

Finalized

Participant

Demographics

Diagnosis

Family history

Exposure

Follow up

Therapy

Molecular test

Biospecimen

Single cell/nucleus sequencing

sc/snRNA-seq

sc/snATAC-seq

sc/snMethyl-C-seq

Bulk sequencing

Bulk DNA-seq

Bulk RNA-seq

Multiplex Imaging

CyCIF

MxIF

mIHC

CODEX

MIBI

IMC

SABER

In Progress

Spatial Transcriptomics

Mass Spectrometry

Publications

Tools

Single Cell RNA-Seq

1

Metadata attributes harmonized with Human Cell Atlas metadata.

2

Level 1: Unaligned sequenced reads

FASTQs or unaligned BAMs.

Level 2: Aligned sequence reads

Aligned BAMs

Level 3: Gene expression values

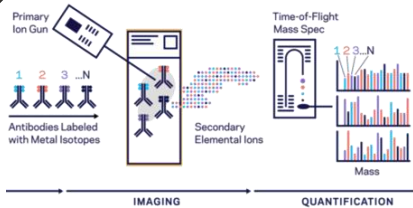
MTX and CSV file format. Genes are specified with Ensembl IDs (GENCODE v. 33)

Level 4: Cell type assignments

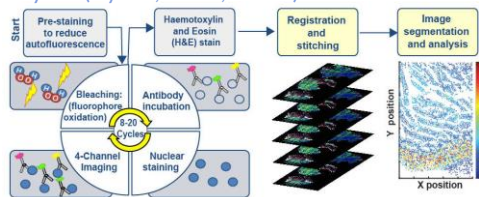
Anndata and CSV file format. Can be used to assign cell type annotations, tSNE coordinates, etc.

1

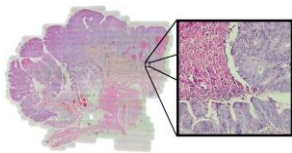
MIBI & IMC



Cyclic (CyCIF, MxIF, mIHC)

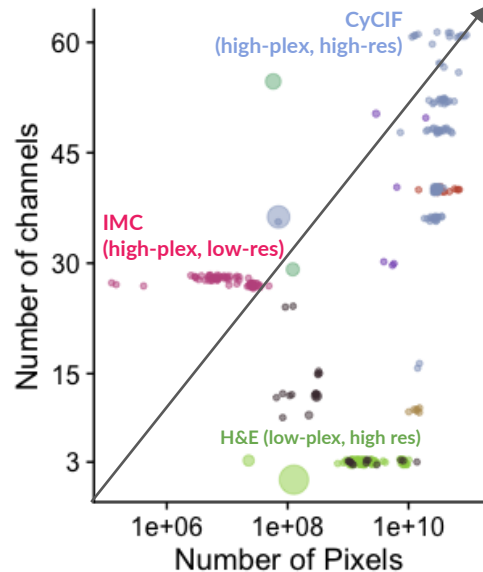
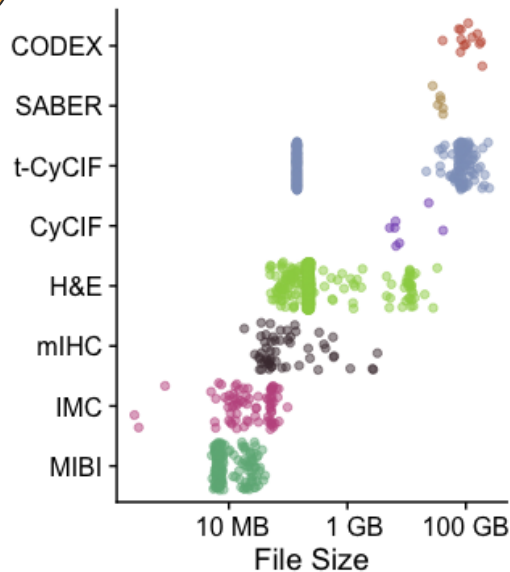


Digital pathology (H&E, IHC)



And more: incl EM & clinical imaging

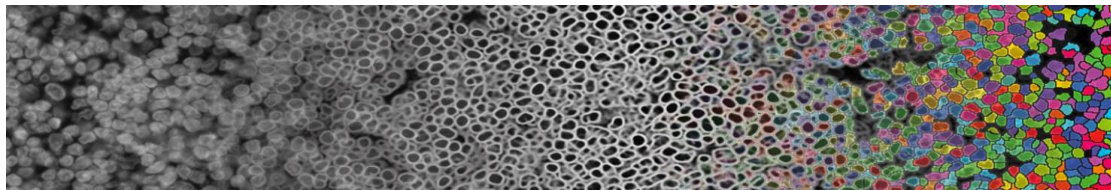
2



Increasingly larger, higher dimension and higher resolution data

Attribution: Adam Taylor, HTAN DCC

Explosion in imaging data size and diversity

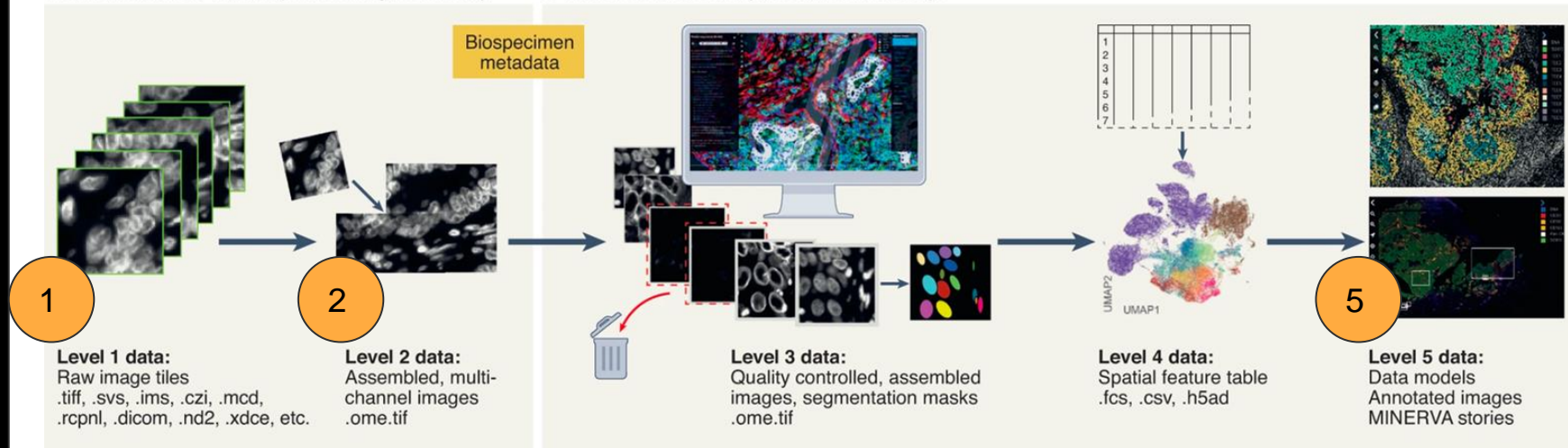


MCMICRO/C. Yapp

Schapiro D, et al. MITI minimum information guidelines for highly multiplexed tissue images. *Nat Methods*. 2022 Mar;19(3):262-267.

a MITI data levels 1 and 2 (access may be limited):

b MITI data levels 3–5 (unrestricted access):



From pixels to cell annotations

Data Availability and Visualization

03

1

HTAN Data
Portal

2

Visualization
Tools

3



SYNAPSE

4



NATIONAL CANCER INSTITUTE
Cancer Research Data Commons

Just released!

5



Google BigQuery

Coming soon!

Five Modes of Data Access

Data	Data Access	Status
1 De-identified clinical and biospecimen data	Open	Available now via the HTAN Portal and Synapse. Available via BigQuery soon.
Level 1-2 Sequencing Data	Access Controlled	Available via Cancer Data Service (CDS) and Seven Bridges. 2
1 Level 3-4 Sequencing Data	Open	Available now via the HTAN Portal, Synapse and Cell x Gene. Available via BigQuery soon.
1 Image Browsing	Open	Available Now via the HTAN Portal and Synapse.
Image Processing	Open	Available <u>soon</u> via the Cancer Data Service (CDS) and the Imaging Data Common (IDC). 3

We currently do **data releases every 3-4 months.**

Most recent release on **May 18, 2022.** See <https://data.humantumoratlas.org/data-updates>

10

Atlases

1,223

Cases

2,752





















Biospecimens

Data Releases

Search all filters ▾ Atlas ▾ Organ ▾ Disease Type ▾ Demographics ▾ Assay Type ▾ File Type ▾

Atlases **Cases** Biospecimens Files **Summary:** 10 Atlases 45 Organs 36 Cancer Types 1223 Cases 2750 Biospecimens 14 Assays 23116 Files

Columns (7/7) Search 

<input type="checkbox"/>	Atlas Name	Lead Institution	Atlas Description	Metadata	Cases	Biospecimens	Assays
<input type="checkbox"/>	HTAN BU	Boston University, University of California, Los Angeles (UCLA)	The Lung Pre-Cancer Atlas				
<input type="checkbox"/>	HTAN HTAPP	Broad Institute, Dana-Farber Cancer Institute, Harvard Medical School	Human Tumor Atlas Pilot Project (HTAPP)				
<input type="checkbox"/>	HTAN CHOP	Children's Hospital of Philadelphia	Center for Pediatric Tumor Cell Atlas				3 
<input type="checkbox"/>	HTAN Duke	Duke University	Breast Pre-Cancer Atlas		680	994	107 
<input type="checkbox"/>	HTAN HMS	Harvard Medical School, Brigham and Women's Hospital	Pre-Cancer Atlases of Cutaneous and Hematologic Origin		105	620	348 
<input type="checkbox"/>	HTAN MSK	Memorial Sloan Kettering Cancer Center	Transition to Metastatic State: Lung Cancer, Pancreatic Cancer and Brain Metastasis		87	94	11 
<input type="checkbox"/>	HTAN OHSU	Oregon Health & Science University (OHSU)	Omic and Multidimensional Spatial Atlas of Metastatic Breast Cancers		1	65	1  1 
<input type="checkbox"/>	HTAN Stanford	Stanford University	Multi-omic Characterization of Transformation of Familial Adenomatous Polyposis		15	181	
<input type="checkbox"/>	HTAN Vanderbilt	Vanderbilt University	Colon Molecular Atlas Project		90	193	1  20 
<input type="checkbox"/>	HTAN WUSTL	Washington University	Washington University Human Tumor Atlas Research Center		21	313	149 

List of all atlases

1

Search all filters ▾ Atlas ▾ Organ ▾ Disease Type ▾ Demographics ▾ Assay Type ▾ File Type ▾

Atlases **Biospecimens** Files **Summary:** 10 Atlases 45 Organs 36 Cancer Types 1223 Cases 2750 Biospecimens 14 Assays 23116 Files

2 Filter data

Columns (7/7) ▾ Search

<input type="checkbox"/>	Atlas Name	Lead Institution	Atlas Description	Metadata	Cases	Biospecimens	Viewers
<input type="checkbox"/>	HTAN BU	Boston University, University of California, Los Angeles (UCLA)	The Lung Pre-Cancer Atlas		6	15	1
<input type="checkbox"/>	HTAN HTAPP	Broad Institute, Dana-Farber Cancer Institute, Harvard Medical School	Human Tumor Atlas Pilot Project (HTAPP)		188	243	
<input type="checkbox"/>	HTAN CHOP	Children's Hospital of Philadelphia	Center for Pediatric Tumor Cell Atlas		30	34	3
<input type="checkbox"/>	HTAN Duke	Duke University	Breast Pre-Cancer Atlas		680	994	107
<input type="checkbox"/>	HTAN HMS	Harvard Medical School, Brigham and Women's Hospital	Pre-Cancer Atlases of Cutaneous and Hematologic Origin		105	620	348
<input type="checkbox"/>	HTAN MSK	Memorial Sloan Kettering Cancer Center	Transition to Metastatic State: Lung Cancer, Pancreatic Cancer and Brain Metastasis		87	94	11
<input type="checkbox"/>	HTAN OHSU	Oregon Health & Science University (OHSU)	Omic and Multidimensional Spatial Atlas of Metastatic Breast Cancers		1	65	1 1
<input type="checkbox"/>	HTAN Stanford	Stanford University	Multi-omic Characterization of Transformation of Familial Adenomatous Polyposis		15	181	
<input type="checkbox"/>	HTAN Vanderbilt	Vanderbilt University	Colon Molecular Atlas Project		90	193	1 20
<input type="checkbox"/>	HTAN WUSTL	Washington University	Washington University Human Tumor Atlas Research Center		21	313	149

Search all filters ▾ Atlas ▾ Organ ▾ Disease Type ▾ Demographics ▾ Assay Type ▾ File Type ▾

Atlases **Cases** Biospecimens Files **Summary:** 10 Atlases 45 Organs 36 Cancer Types 1223 Cases 2750 Biospecimens 14 Assays 23116 Files

Columns (7/7) ▾ Search

<input type="checkbox"/>	Atlas Name	Lead Institution	Atlas Description	Metadata	Cases	Biospecimens	Vir
<input type="checkbox"/>	HTAN BU	Boston University, University of California, Los Angeles (UCLA)	The Lung Pre-Cancer Atlas				
<input type="checkbox"/>	HTAN HTAPP	Broad Institute, Dana-Farber Cancer Institute, Harvard Medical School	Human Tumor Atlas Pilot Project (HTAPP)		18		
<input type="checkbox"/>	HTAN CHOP	Children's Hospital of Philadelphia	Center for Pediatric Tumor Cell Atlas		3		
<input type="checkbox"/>	HTAN Duke	Duke University	Breast Pre-Cancer Atlas		680	994	107
<input type="checkbox"/>	HTAN HMS	Harvard Medical School, Brigham and Women's Hospital	Pre-Cancer Atlases of Cutaneous and Hematologic Origin		105	620	348
<input type="checkbox"/>	HTAN MSK	Memorial Sloan Kettering Cancer Center	Transition to Metastatic State: Lung Cancer, Pancreatic Cancer and Brain Metastasis		87	94	11
<input type="checkbox"/>	HTAN OHSU	Oregon Health & Science University (OHSU)	Omic and Multidimensional Spatial Atlas of Metastatic Breast Cancers		1	65	1 1
<input type="checkbox"/>	HTAN Stanford	Stanford University	Multi-omic Characterization of Transformation of Familial Adenomatous Polyposis		15	181	
<input type="checkbox"/>	HTAN Vanderbilt	Vanderbilt University	Colon Molecular Atlas Project		90	193	1 20
<input type="checkbox"/>	HTAN WUSTL	Washington University	Washington University Human Tumor Atlas Research Center		21	313	149





















Download metadata

Search all filters ▾ Atlas ▾ Organ ▾ Disease Type ▾ Demographics ▾ Assay Type ▾ File Type ▾

Atlases **Cases** Biospecimens Files **Summary:** 10 Atlases 45 Organs 36 Cancer Types 1223 Cases 2750 Biospecimens 14 Assays 23116 Files

Columns (7/7) ▾ Search 

<input type="checkbox"/>	Atlas Name	Lead Institution	Atlas Description	Metadata	Cases	Biospecimens	Viewers
<input type="checkbox"/>	HTAN BU	Boston University, University of California, Los Angeles (UCLA)	The Lung Pre-Cancer Atlas			15	1 
<input type="checkbox"/>	HTAN HTAPP	Broad Institute, Dana-Farber Cancer Institute, Harvard Medical School	Human Tumor Atlas Pilot Project (HTAPP)			243	
<input type="checkbox"/>	HTAN CHOP	Children's Hospital of Philadelphia	Center for Pediatric Tumor Cell Atlas		30	34	3 
<input type="checkbox"/>	HTAN Duke	Duke University	Breast Pre-Cancer Atlas		680	994	107 
<input type="checkbox"/>	HTAN HMS	Harvard Medical School, Brigham and Women's Hospital	Pre-Cancer Atlases of Cutaneous and Hematologic Origin		105	620	348 
<input type="checkbox"/>	HTAN MSK	Memorial Sloan Kettering Cancer Center	Transition to Metastatic State: Lung Cancer, Pancreatic Cancer and Brain Metastasis		87	94	11 
<input type="checkbox"/>	HTAN OHSU	Oregon Health & Science University (OHSU)	Omic and Multidimensional Spatial Atlas of Metastatic Breast Cancers		1	65	1  1 
<input type="checkbox"/>	HTAN Stanford	Stanford University	Multi-omic Characterization of Transformation of Familial Adenomatous Polyposis		15	181	
<input type="checkbox"/>	HTAN Vanderbilt	Vanderbilt University	Colon Molecular Atlas Project		90	193	1  20 
<input type="checkbox"/>	HTAN WUSTL	Washington University	Washington University Human Tumor Atlas Research Center		21	313	149 

4

Visualize data



https://data.humantumoratlas.org/selectedFilters=MSB%7B"value"%3A"1-CyCIF"%2C"label"%3A"1-CyCIF"%2C"group...

The HTAN Data Portal is in the alpha release phase. Only level 3 and 4 data can be downloaded. Feedback is welcome!

HTAN EXPLORE DATA STANDARDS DATA TRANSFER ANALYSIS TOOLS HTAN MAIN SITE

Search all filters Atlas Organ Disease Type Demographics Assay Type File Type

Assay: 1-CyCIF Clear all filters

Summary: 1 Atlas 6 Organs 2 Cancer Types 16 Cases 0 Biospecimens 1 Assay 16 Files

Columns (9/55) Search

Filename	Atlas Name	Biospecimen	Assay	Level	Organ	Diagnosis	Details	View
<input type="checkbox"/> CRC02.ome.tif	HTAN HMS	HTA7_926_2	1-CyCIF	Level 2	Rectum NOS	Rectum NOS	View Details	
<input type="checkbox"/> CRC03.ome.tif	HTAN HMS	HTA7_927_2	1-CyCIF	Level 2	Sigmoid colon	Sigmoid colon	View Details	
<input type="checkbox"/> CRC04.ome.tif	HTAN HMS	HTA7_932_2	1-CyCIF	Level 2	Sigmoid colon	Sigmoid	View Details	
<input type="checkbox"/> CRC05.ome.tif	HTAN HMS	HTA7_934_2	1-CyCIF	Level 2	Rectosigmoid junction	Rectosigmoid junction	View Details	

Click to view image in Minerva
Thumbnail generated with Miniature

https://absynpharm.github.io/HTA-CRCATLAS-1/Minerva/crc04-overview.html

https://data.humantumoratlas.org/explore?tab=file&thumbnail&selectedFilters=MSB%7B"value"%3A"OME-TIFF"%2C"label"%3A"OME-TIFF"%2C"group...

The HTAN Data Portal is in the alpha release phase. Only level 3 and 4 data can be downloaded. Feedback is welcome!

HTAN EXPLORE DATA STANDARDS DATA TRANSFER ANALYSIS TOOLS HTAN MAIN SITE

Search all filters Atlas Organ Disease Type Demographics Assay Type File Type

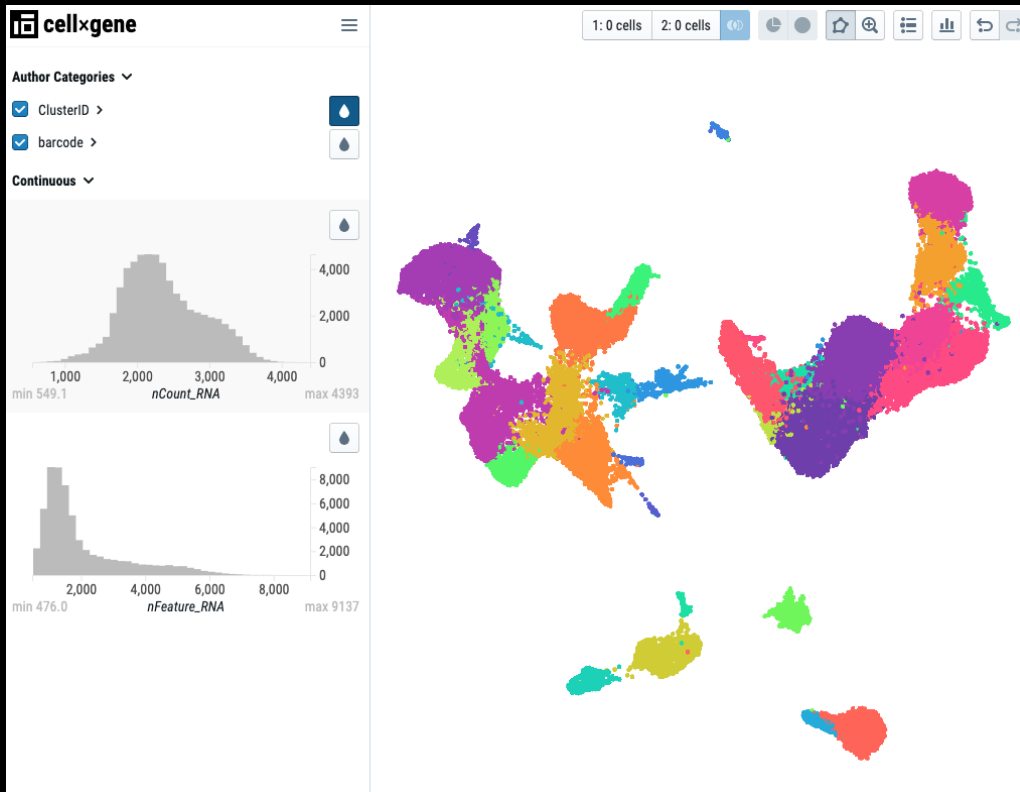
File Format OME-TIFF Atlas HTAN OHSU Assay miHC Clear all filters

Summary: 1 Atlas 1 Organ 1 Cancer Type 1 Case 0 Biospecimens 1 Assay 54 Files

Columns (9/55) Search

Filename	Atlas Name	Biospecimen	Assay	Level	Organ	Diagnosis	Details	View
<input type="checkbox"/> HTA8_1_BA_F.ome.tif	HTAN OHSU	HTA8_1_19	miHC	Level 2	Breast NOS	Breast NOS	View Details	
<input type="checkbox"/> HTA8_1_BA_...1.ome.tif	HTAN OHSU	HTA8_1_19	miHC	Level 2	Breast NOS	Breast NOS	View Details	
<input type="checkbox"/> HTA8_1_BA_...2.ome.tif	HTAN OHSU	HTA8_1_19	miHC	Level 2	Breast NOS	Breast NOS	View Details	
<input type="checkbox"/> HTA8_1_BA_...3.ome.tif	HTAN OHSU	HTA8_1_19	miHC	Level 2	Breast NOS	Breast NOS	View Details	
<input type="checkbox"/> HTA8_1_BA_...4.ome.tif	HTAN OHSU	HTA8_1_19	miHC	Level 2	Breast NOS	Breast NOS	View Details	
<input type="checkbox"/> HTA8_1_BA_L.ome.tif	HTAN OHSU	HTA8_1_17	miHC	Level 2	Breast NOS	Breast NOS	View Details	

Minerva Visualization (collab with Sorger Group)



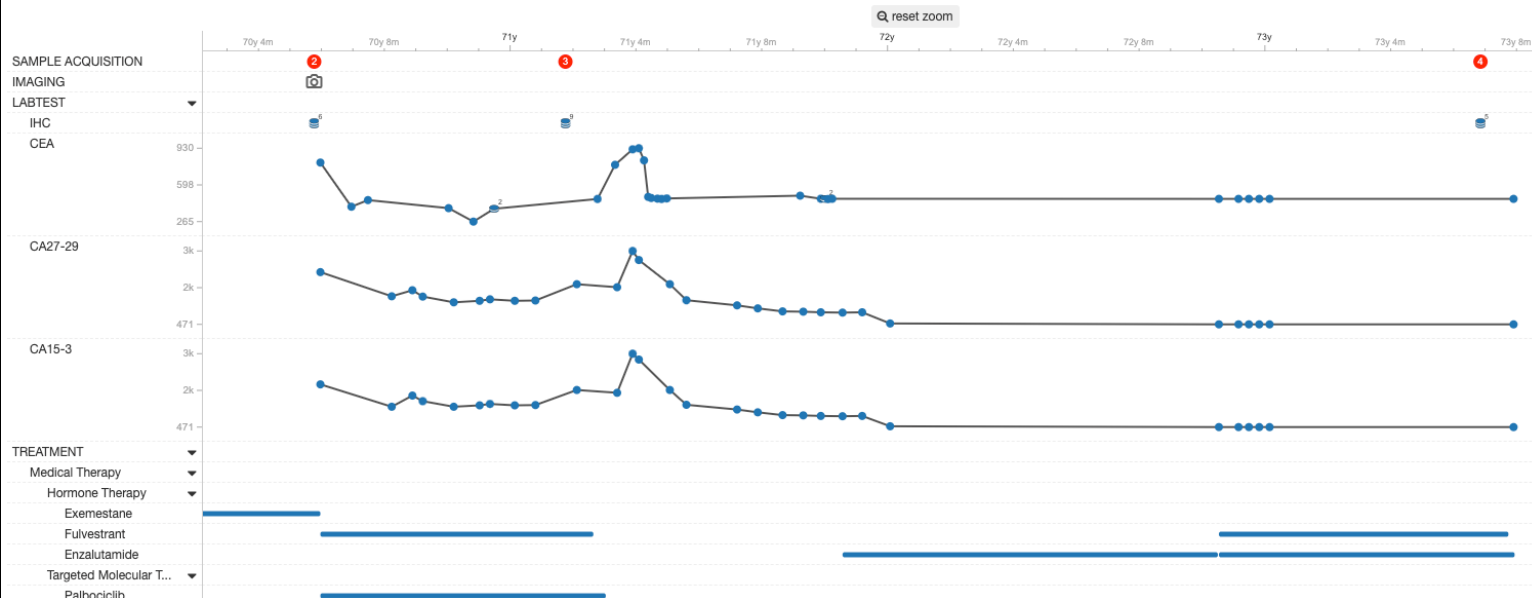
Cell x Gene Visualization (collab with CZI)

Patient: HTA9_1, Female, 64 years old, Breast Cancer (Breast Invasive Ductal Carcinoma), LIVING (128 months)

Breast Cancer

Samples: 1 HTA9_1-PR, Primary (Breast NOS) 2 HTA9_1-Bx1, Metastasis (Liver) 3 HTA9_1-Bx2, Metastasis (Liver) 4 HTA9_1-Bx3, Metastasis (Bone NOS) 5 HTA9_1-Bx4, Metastasis (Liver)

Summary Genomic Evolution Pathways Clinical Data



cBioPortal Visualization

Thank you!

HTAN DCC Team



James



Elyse



Milen



Adam



Yooree

Aaron Lisman

Bruno Grande

Mialy DeFelice



Ethan



Rochelle



Ashley



Niki



Ino

Onur Sumer

Xiang Li

Pierrette Lo

Tiarah Jeudi



Vésteinn



Bill



Clarisse



Dave



Dar'ya