



VERSION 2

APR 16, 2024

OPEN ACCESS



DOI:

[dx.doi.org/10.17504/protocols.io.q26g71rpkgwz/v2](https://dx.doi.org/10.17504/protocols.io.q26g71rpkgwz/v2)

**Protocol Citation:** Kyu Sang Han, Pei-Hsun Wu, Joel Sunshine, Sashank Reddy, Ashley Kiemen, Denis Wirtz 2024. CODA (part 1): setting up environment and preparing sample dataset | HuBMAP | JHU-TMC. **protocols.io** <https://dx.doi.org/10.17504/protocols.io.q26g71rpkgwz/v2> Version created by [Kyu Sang Han](#)

## 🌐 CODA (part 1): setting up environment and preparing sample dataset | HuBMAP | JHU-TMC V.2

Kyu Sang Han<sup>1</sup>, Pei-Hsun Wu<sup>1</sup>, Joel Sunshine<sup>2</sup>, Sashank Reddy<sup>2</sup>, Ashley Kiemen<sup>2</sup>, Denis Wirtz<sup>1,2</sup>

<sup>1</sup>Johns Hopkins University; <sup>2</sup>Johns Hopkins Medicine

Human BioMolecular Atlas Program (HuBMAP) Method Development Community

TMC - Johns Hopkins University



Kyu Sang Han  
Johns Hopkins University

### ABSTRACT

CODA workflow part 1. setting up environment and preparing dataset

**MANUSCRIPT CITATION:**

A.M. Braxton, A.L. Kiemen, M.P. Grahn, A. Forjaz, J. Parksong, J.M. Babu, J. Lai, L. Zheng, N. Niknafs, L. Jiang, H. Cheng, Q. Song, R. Reichel, S. Graham, A.I. Damanakis, C.G. Fischer, S. Mou, C. Metz, J. Granger, X.-D. Liu, N. Bachmann, Y. Zhu, Y.Z. Liu, C. Almagro-Pérez, A.C. Jiang, J. Yoo, B. Kim, S. Du, E. Foster, J.Y. Hsu, P.A. Rivera, L.C. Chu, D. Liu, E.K. Fishman, A. Yuille, N.J. Roberts, E.D. Thompson, R.B. Scharpf, T.C. Cornish, Y. Jiao, R. Karchin, R.H. Hruban, P.-H. Wu, D. Wirtz, and L.D. Wood, “3D genomic mapping reveals multifocality of human pancreatic precancers”, *Nature* (2024)

A.L. Kiemen, A. Forjaz, R. Sousa, K. Sang Han, R.H. Hruban, L.D. Wood, P.H. Wu, and D. Wirtz, “High-resolution 3D printing of pancreatic ductal microanatomy enabled by serial histology”, *Advanced Materials Technologies* 9, 2301837 (2024)

T. Yoshizawa, J. W. Lee, S.-M. Hong, D.J. Jung, M. Noe, W. Sbijewski, A. Kiemen, P.H. Wu, D. Wirtz, R.H. Hruban, L.D. Wood, and K. Oshima. “Three-dimensional analysis of ductular reactions and their correlation with liver regeneration and fibrosis”, *Virchows Archiv* (2023).

A.L. Kiemen, A.I. Damanakis, A.M. Braxton, J. He, D. Laheru, E.K. Fishman, P. Chames, C. Almagro Perez, P.-H. Wu, D. Wirtz, L.D. Wood, and R. Hruban, “Tissue clearing and 3D reconstruction of digitized, serially sectioned slides provide novel insights into pancreatic cancer”, *Med* 4, 75-91 (2023)

A. Kiemen, Y. Choi, A. Braxton, C. Almagro Perez, S. Graham, M. Grahm, N., N. Roberts, L. Wood, P. Wu, R. Hruban, and D. Wirtz, “Intraparenchymal metastases as a cause for local recurrence of pancreatic cancer”, *Histopathology* 82: 504-506 (2023)

(2022)

A.L. Kiemen, A.M. Braxton, M.P. Grahm, K.S. Han, J.M. Babu, R. Reichel, A.C. Jiang, B. Kim, J. Hsu, F. Amoa, S. Reddy, S.-M. Hong, T.C. Cornish, E.D. Thompson, P. Huang, L.D. Wood, R.H. Hruban, D. Wirtz and P.H. Wu, "CODA: quantitative 3D reconstruction of large tissues at cellular resolution", Nature Methods 19: 1490-1499 (2022)

K.S.Han, I. Sander, J. Kumer, E. Resnick, C. Booth, B. Starich, J. Walston, A.L. Kiemen, S. Reddy, C. Joshu, J. Sunshine, D. Wirtz, P.-H. Wu "The microanatomy of human skin in aging." bioRxiv (2024): 2024-04.

**License:** This is an open access protocol distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited

**Protocol status:** Working

We use this protocol and it's working

**Created:** Apr 16, 2024

**Last Modified:** Apr 16, 2024

**PROTOCOL integer ID:** 98298

**Keywords:** CODA, deeplearning, tissuelabeling, segmentation

**Funders Acknowledgement:**

Institute of Arthritis and Musculoskeletal and Skin Diseases

Grant ID: U54AR081774

National Cancer Institute

Grant ID: U54CA143868

## Software requirements

- 1 MATLAB [MATLAB \(mathworks.com\)](https://mathworks.com)  
Image processing toolbox [Image Processing Toolbox - MATLAB \(mathworks.com\)](https://mathworks.com)  
Deep learning toolbox [Deep Learning Toolbox - MATLAB \(mathworks.com\)](https://mathworks.com)  
MATLAB - Resnet50 model [MATLAB resnet50 \(mathworks.com\)](https://mathworks.com)
- 2 Aperio ImageScope [Aperio ImageScope \(leicabiosystems.com\)](https://leicabiosystems.com)
- 3 FIJI ImageJ [Fiji Downloads \(imagej.net\)](https://imagej.net)

## Download Source code

- 4 Codes are available at the following GitHub: [CODA Github Repository](#)

## Download Sample dataset

- 5 Here, we discuss application to a sample dataset "lungs", containing 150 serial histological images. Download the sample dataset (serial images and sample annotations) here: [Lung Sample Dataset on Google Drive](#)
- 6 Images are .ndpi format and were scanned at 20x magnification (approximately 0.5 micron / pixel resolution), spaced 10 micron apart. Save the images in a local drive folder (e.g. `\\Users\\Ashley\\Documents\\lungs`).
- 7 Filenames for each image should be created such that tissue sections are read consecutively by Matlab. Therefore, include zero-padding in numerical indices.  
  
CORRECT FILENAMES: `lungs_001.ndpi, lungs_002.ndpi, ..., lungs_011.ndpi`  
INCORRECT FILENAMES (no zero padding): `lungs_1.ndpi, lungs_2.ndpi, ..., lungs_11.ndpi`

## Create downsampled copies of high-resolution images

8 The function **create\_downsampled\_tif\_images** will create downsampled copies of the .ndpi files by directly loading each high-resolution images in tiles and down sampling it to the desired pixel resolutions.

9 First, decide the resolution of the images you want to create. Here, we create images of 1 micron / pixel, 2 microns / pixel, and 10 micron / pixel resolution:

```
ds=[1 2 10];
```

10 Next, decide on the name of the output folders for each of the downsampled images you create. Here, we will save the images downsampled to 1 micron / pixel in a folder named "10x", the images downsampled to 2 micron / pixel in a folder named "5x," and the images downsampled to 10 micron / pixel in a folder named "1x."

```
subfolders=["10x" "5x" "1x"];
```

11 Finally, call the function: **create\_downsampled\_tif\_images(pth,ds,subfolders);**

12 Using this function, you will make two subfolders within the original folder containing the .ndpi images. One subfolder named "10x" containing the 20x images downsampled by a factor of 2. The other subfolder named "1x" containing the 20x images downsampled by a factor of 20. Most calculations will be performed on these tif images. Note: here we use 10x and 1x for example, but other resolutions could be created as desired.

```
pth10x=[pth,'10x'];
```

```
pth1x=[pth,'1x'];
```

**\*\*Note: If this code fails due to memory constraints on your computer, try python Openslide.**