Basic Single Cell Lab

Jenny Drnevich, PhD

Assistant Director, HPCBio

Roy J. Carver Biotechnology Center

June 26, 2024

Learning objectives

1. Examine 10x Genomics's web_summary.html report from the output of Cell Ranger
2. Walk through the steps of a standard Seurat analysis
    1. Read explanations of why each step done
    2. Run codes to do each step
    3. Explore how to interact with a Seurat object
3. Try both manual and computational cell type calling
4. Get links for examples of more complex single cell and spatial analyses

Web_summary.html file

This public one is from cellranger multi not cellranger count (example) but all the same information is in there:

Cells tab on left:

1. Any warnings at top?
2. Cell metrics:
    1. **# called** - If >> targeted, high background. If << targeted, poor-quality cells
    2. **mean reads** - If >> targeted, few good cells. If << targeted,
    3. **median genes** - variable between tissue/cell types but ideally > 1000
    4. **total genes** - summed over all genes; 20K+ for mammalian cells.
    5. **median UMIs** - compare with mean reads (PCR dups) and median genes (shows ~UMIs per gene)
    6. **% mapped reads in cells** - if >> 90% can indicate ambient background RNA
3. t-SNE
    1. On left shaded by UMI counts and show strong correlation due to no normalization
    2. right shows first clusters

10k Mouse E18 Combined Cortex, Hippocampus and Subventricular Zone Cells, Chromium GEM-X Single Cell 3' Gene Expression Dataset by Cell Ranger 8.0.0, 10x Genomics, Inc, (2024, Apr 15). Used with permission of 10x Genomics, Inc.

Web_summary.html file

This public one is from cellranger multi not cellranger count but all the same information is in there:

Library tab on left:

1. Cell Statistics (same as Cells tab)
2. Sequencing and Mapping metrics: click on ? for explanations
3. Metrics Per Physical Library
    1. **Sequencing saturation** - "The fraction of reads originating from an already-observed UMI." Indicates whether more sequencing would result in substantially more reads.
4. Plots
    1. **GEX Barcode Rank Plot** - should see sharp "cliff"
    2. **Sequencing Saturation & Median Gene per Cell** - show downsampled values to observe curve and extrapolate what extra sequencing would gain.

Log on to Biocluster's Jupyter Hub

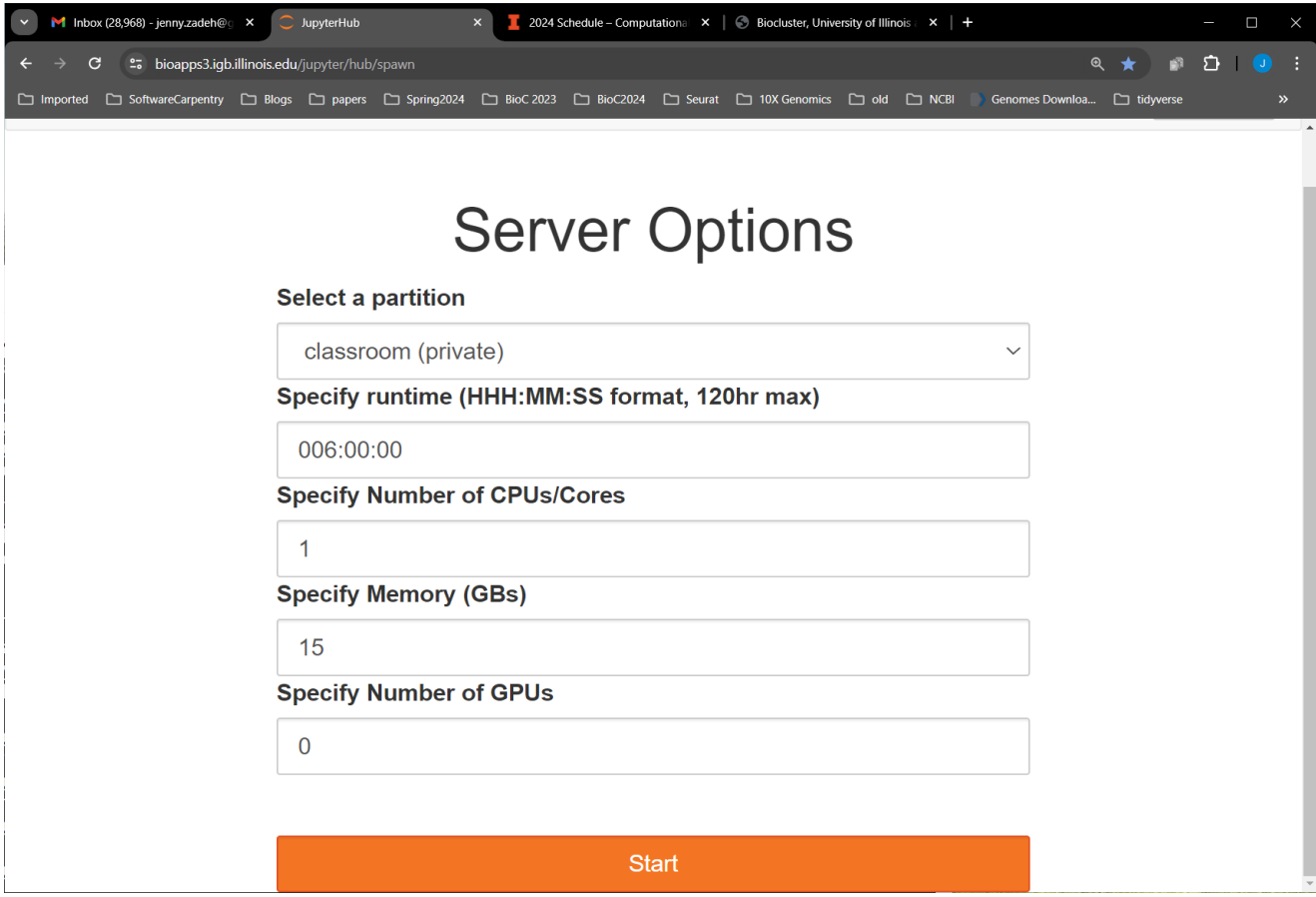
Server options:

Classroom partition

Runtime: **6:00**

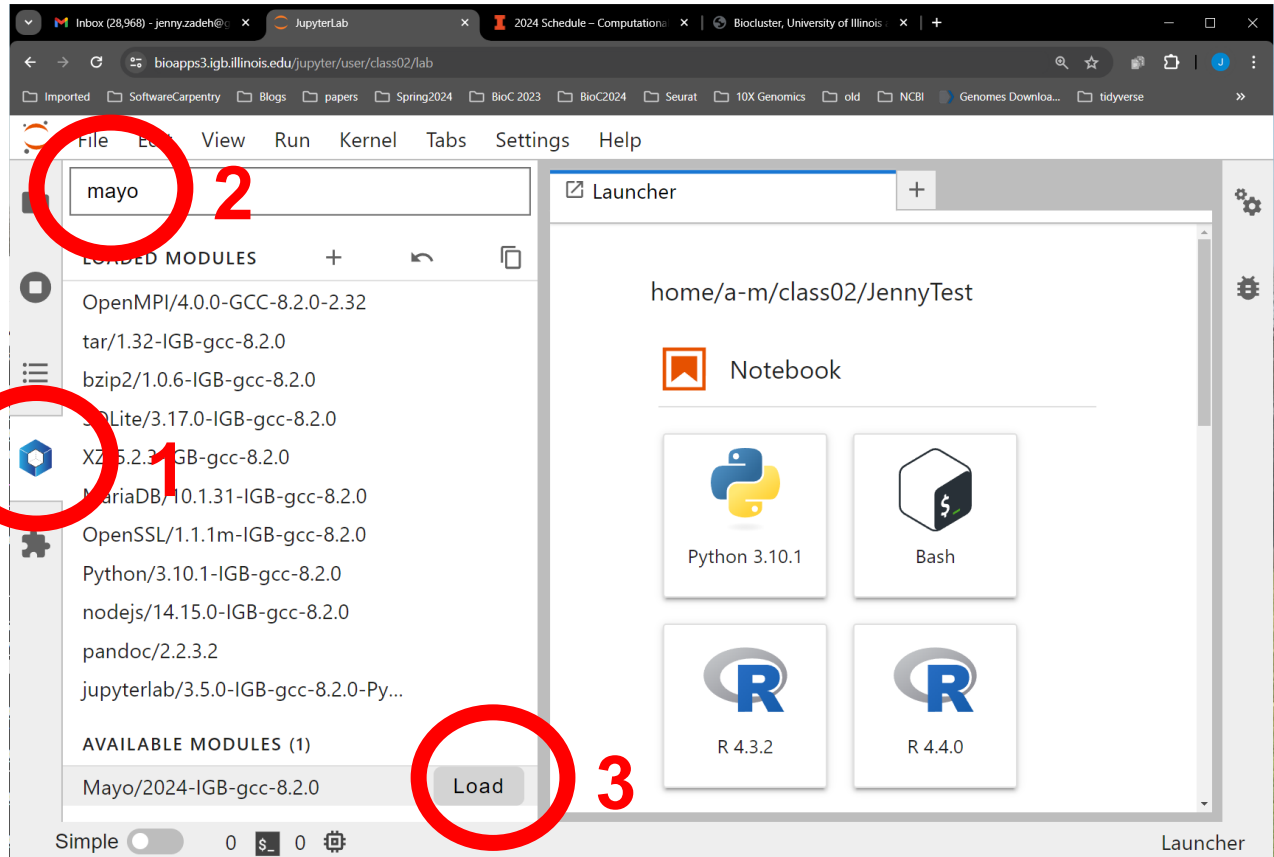
Other defaults fine

Click "Start"



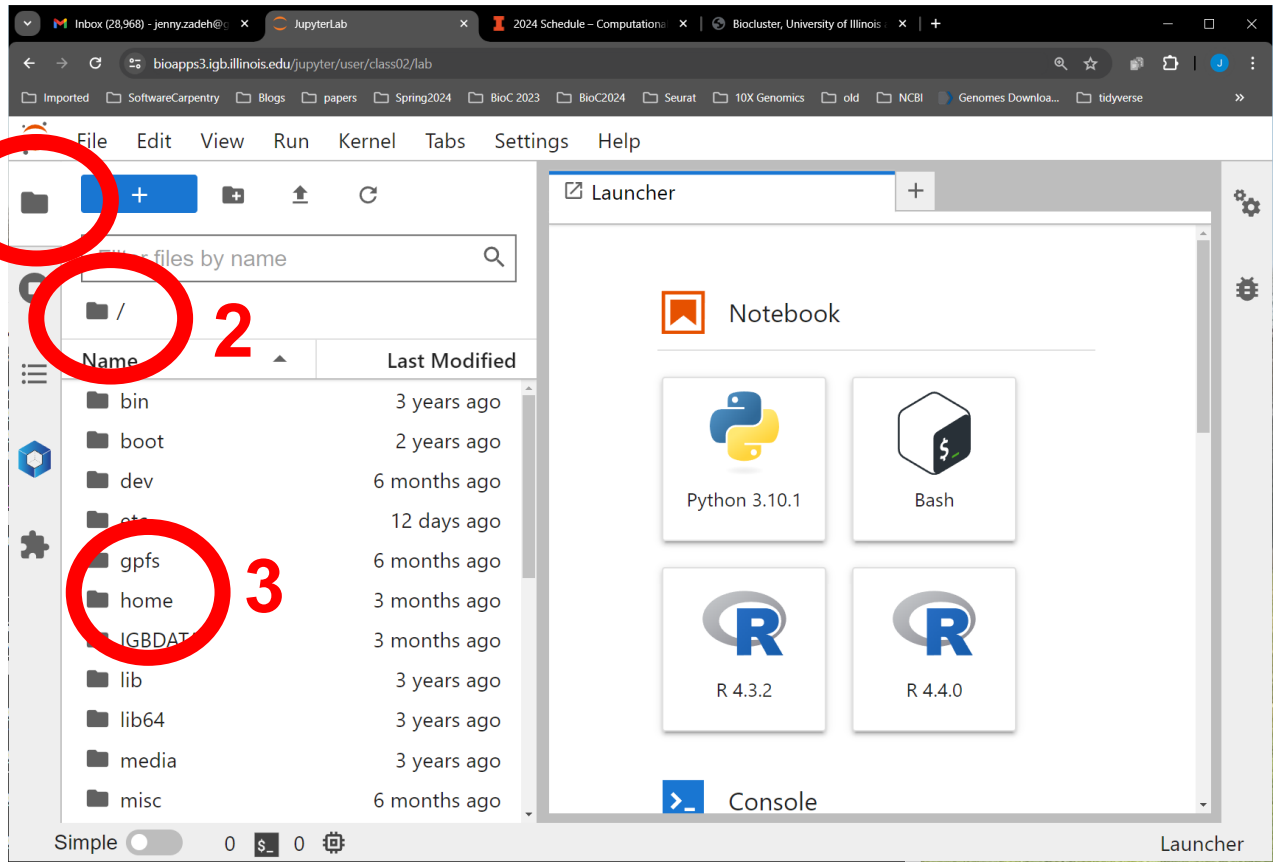
Load Mayo module

1. Click on blue hexagon
2. Search for Mayo at top
3. Click on "Load" by Mayo module at bottom.



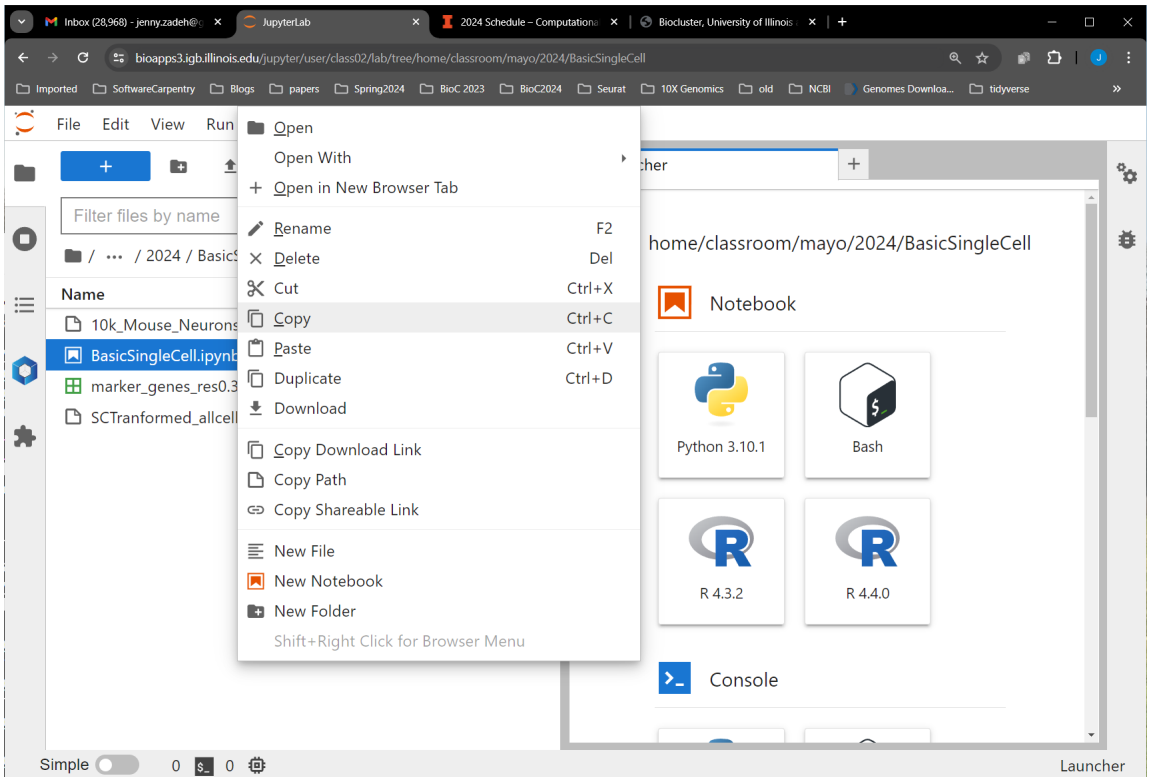
Navigate to proper directory

1. Click on gray folder on very left
2. Click on second gray folder under search box to get to root directory
3. Double click on "home" directory



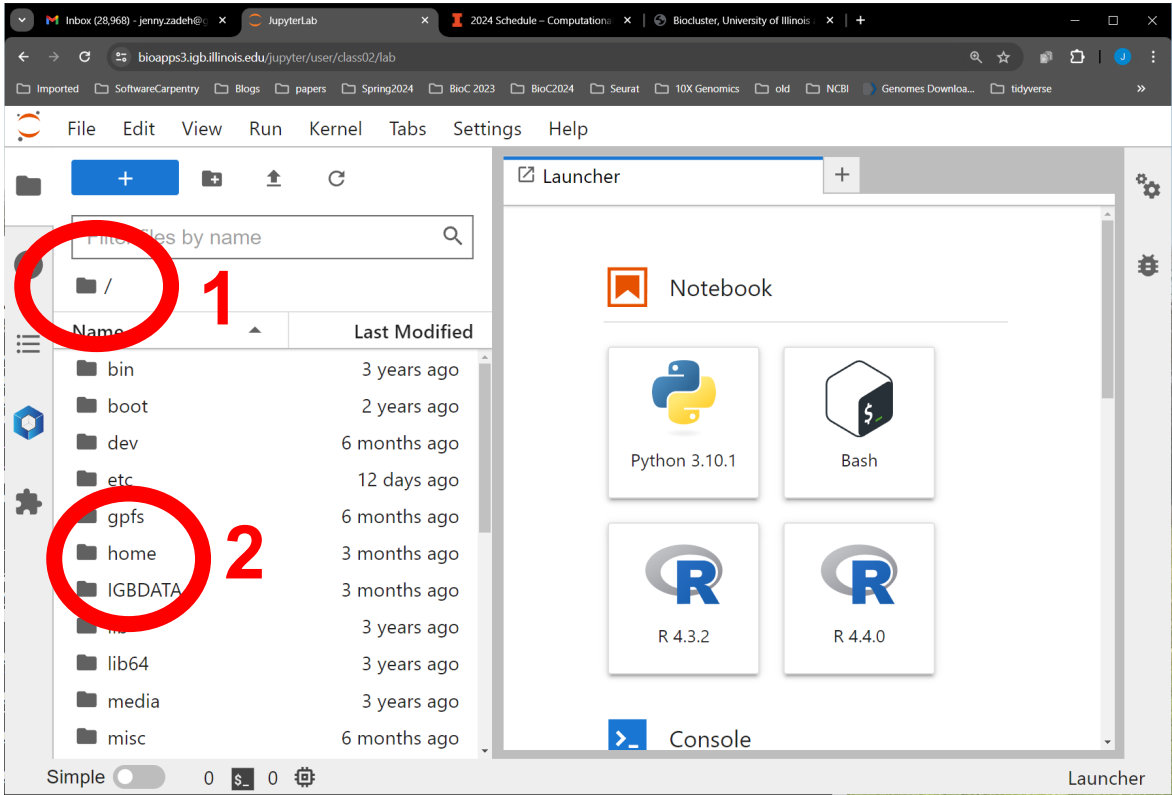
Continue navigating by double clicking:

- Classroom -> mayo -> 2024 -> BasicSingleCell
- Right click on "BasicSingleCell.ipynb" and click "Copy"



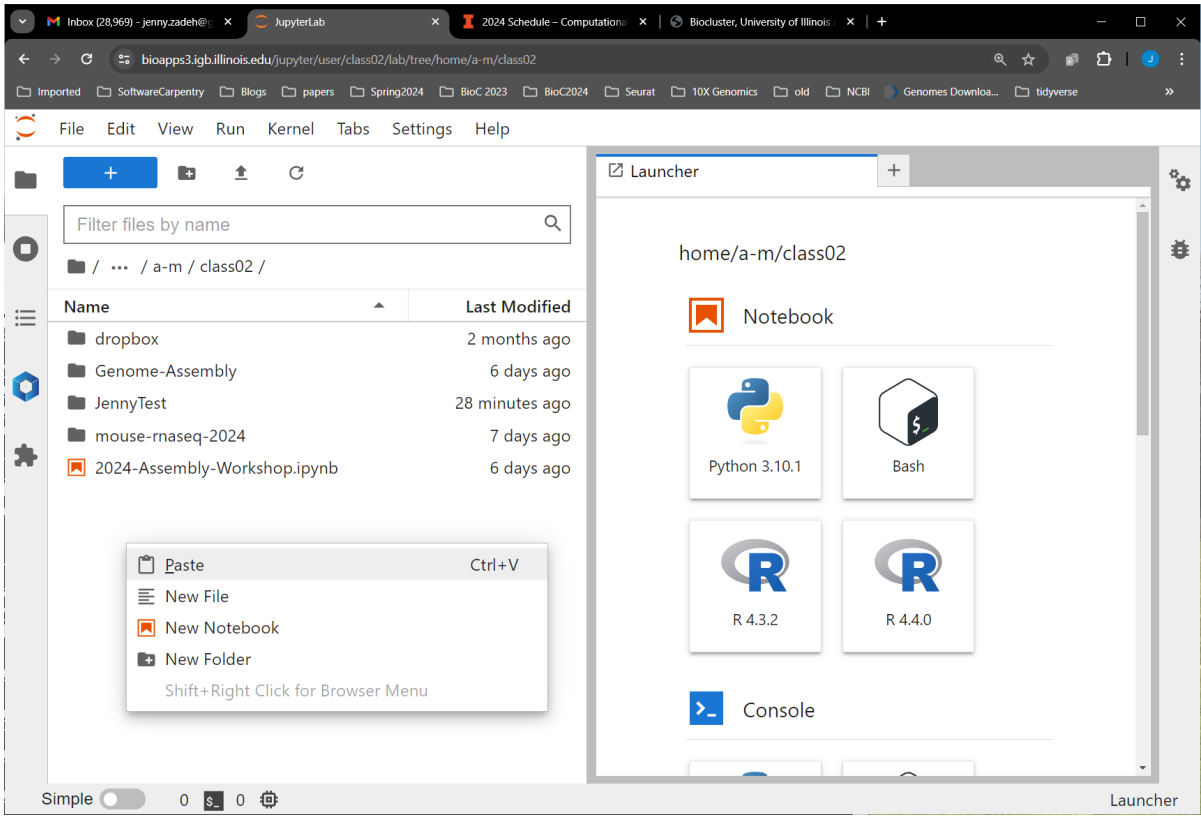
Navigate back to your home directory

1. Click on second gray folder under search box to get to root directory
2. Double click on "home" directory



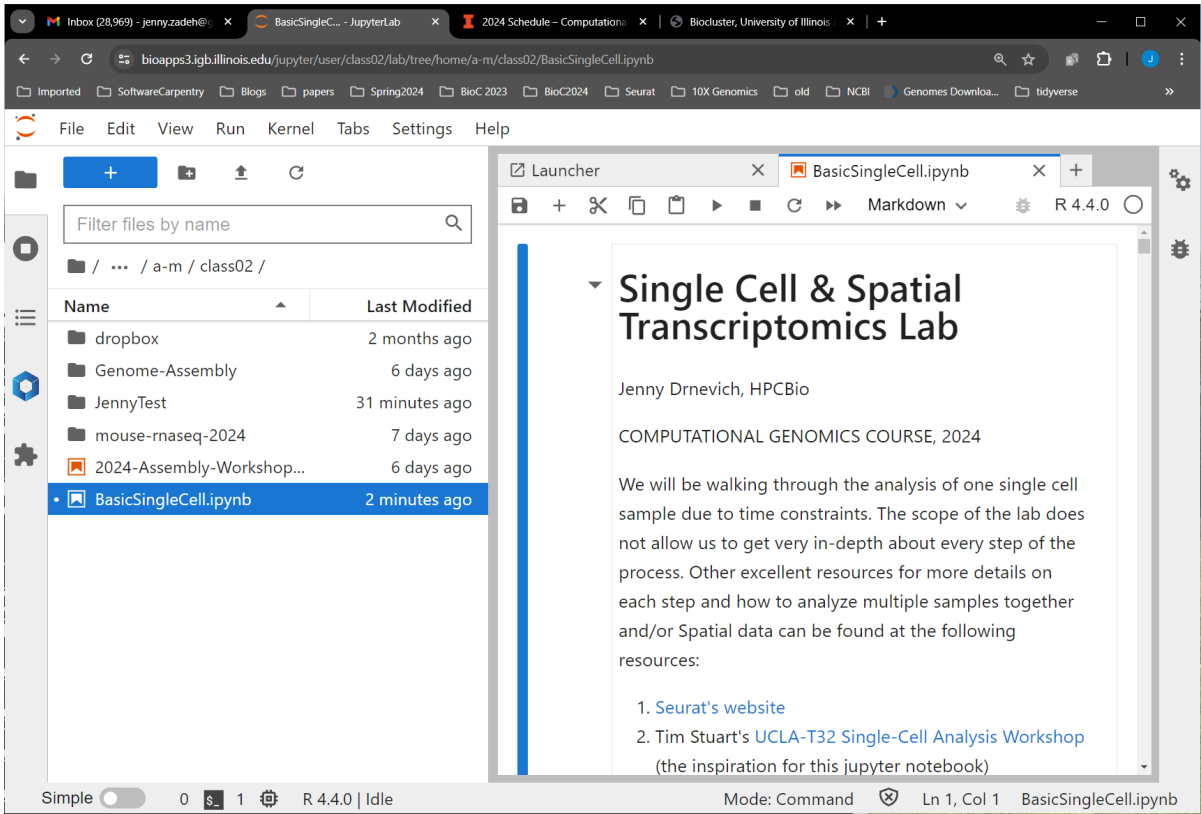
Continue navigating by double clicking:

- a-m -> classXX
- Right click below any files/directories and select "Paste"
- You should then see the "BasicSingleCell.ipynb"



Open BasicSingleCell.ipynb

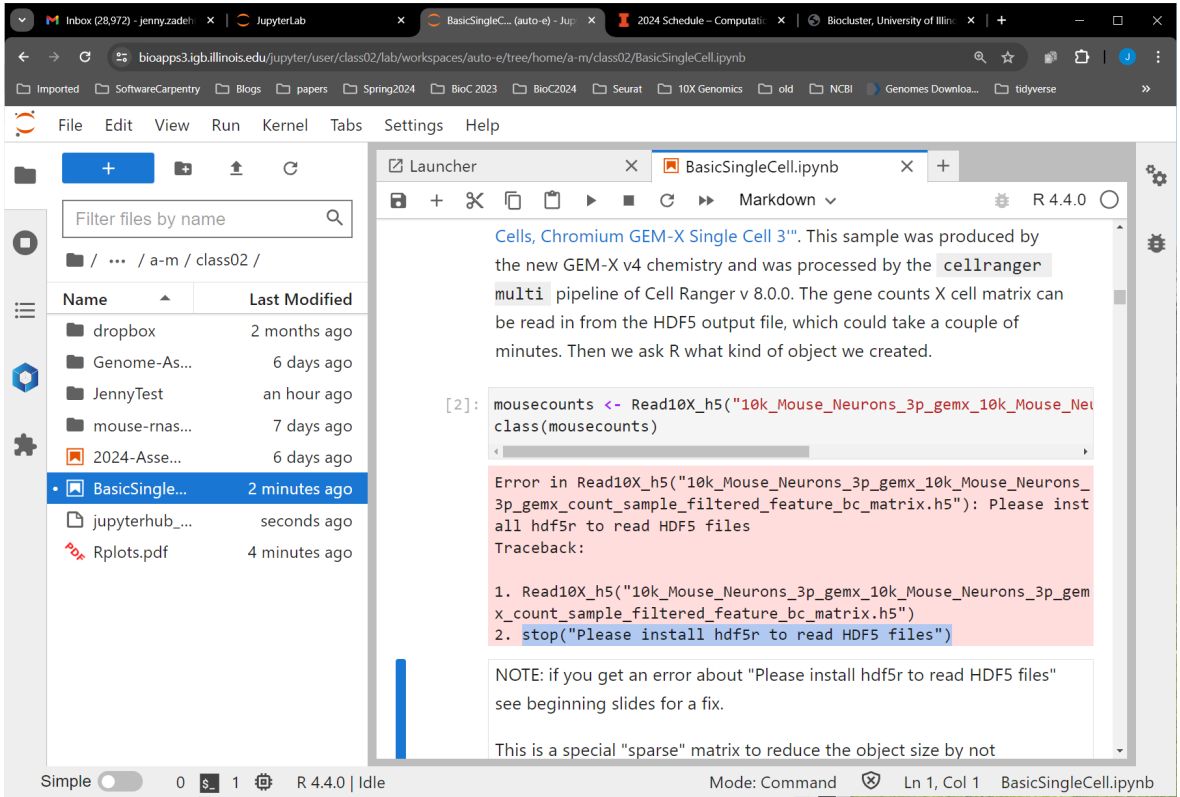
- Double click on the file to open in right pane
- This is a self-paced lab with explanations and runnable code boxes
- If the second code box gives you an error, see next slide



stop("Please install hdf5r to read HDF5 files")

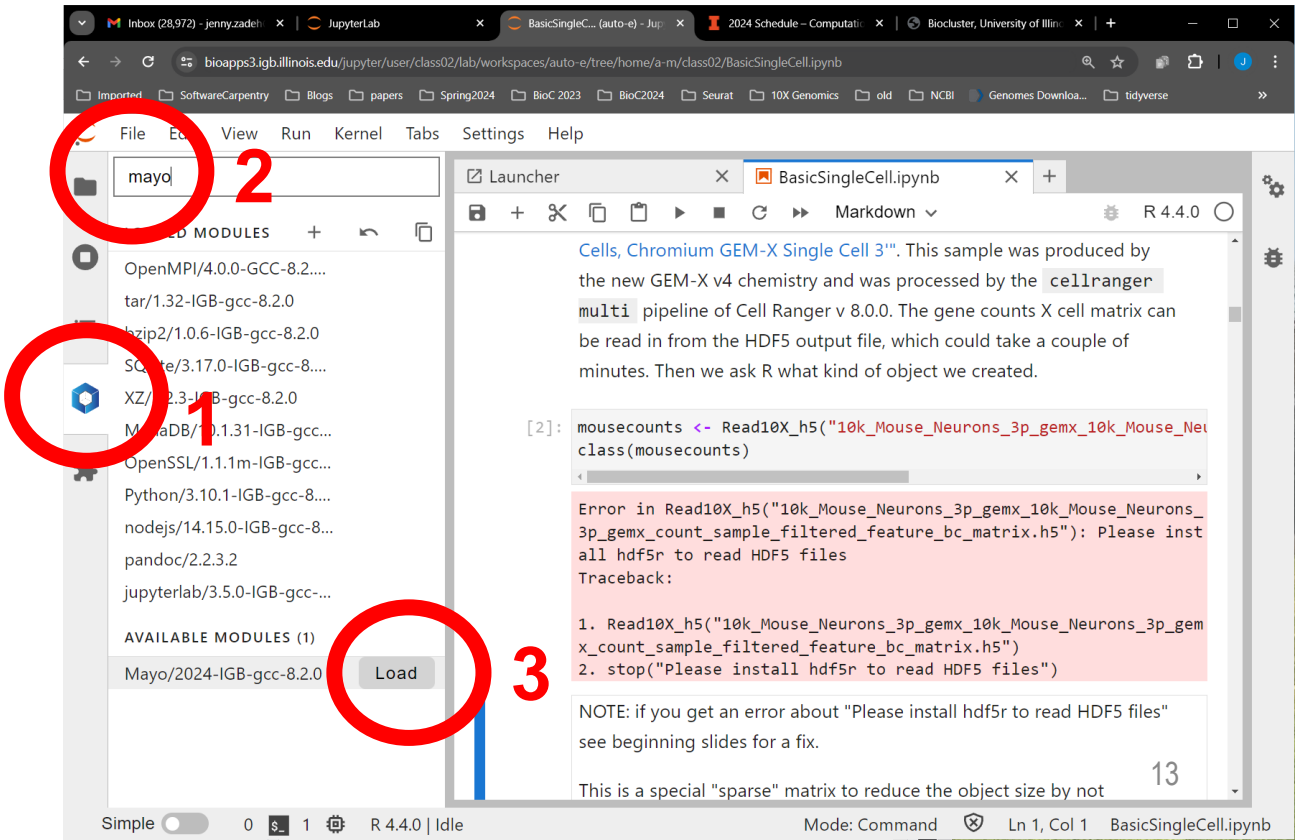
If you get this error in the second code box then you did not load the Mayo module before opening the .ipynb.

To fix, follow next slides



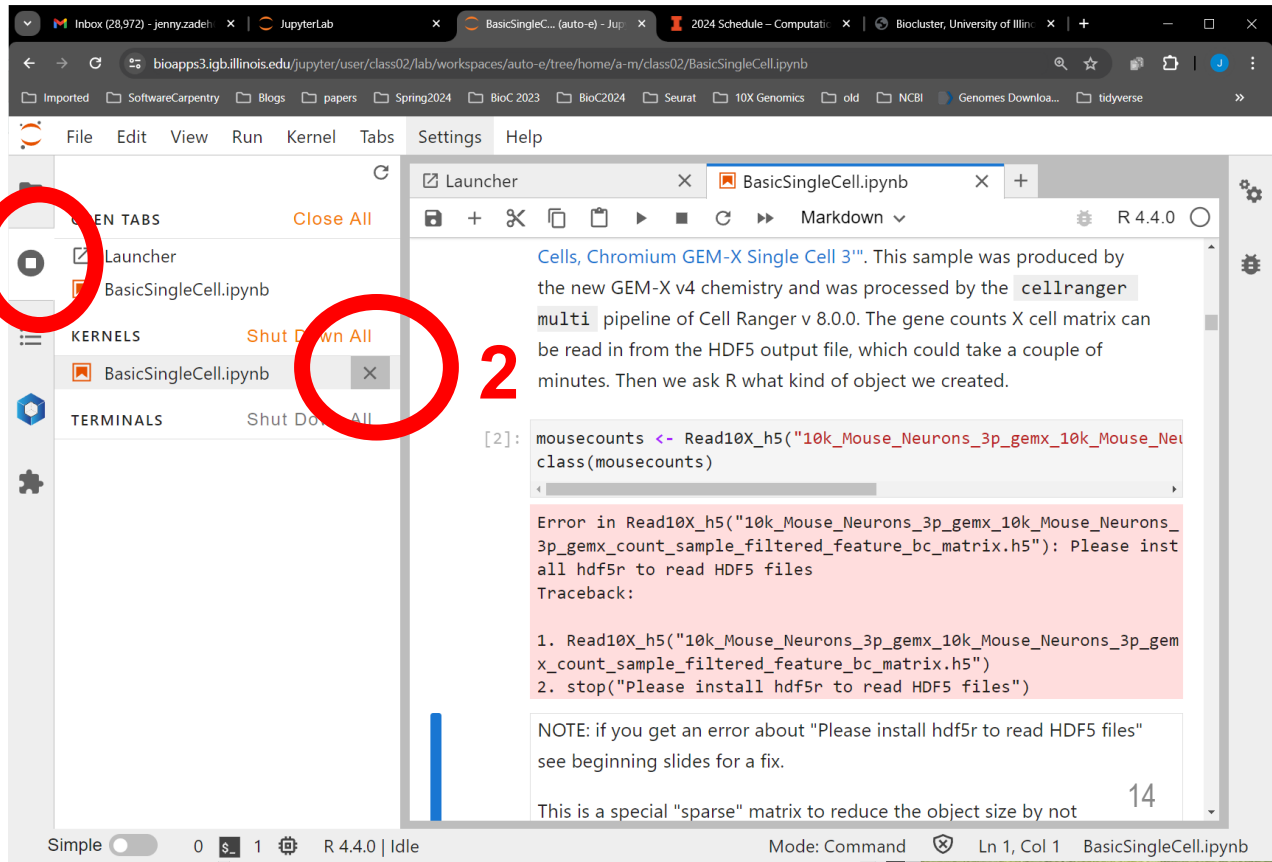
Load Mayo module

1. Click on blue hexagon
2. Search for Mayo at top
3. Click on "Load" by Mayo module at bottom.



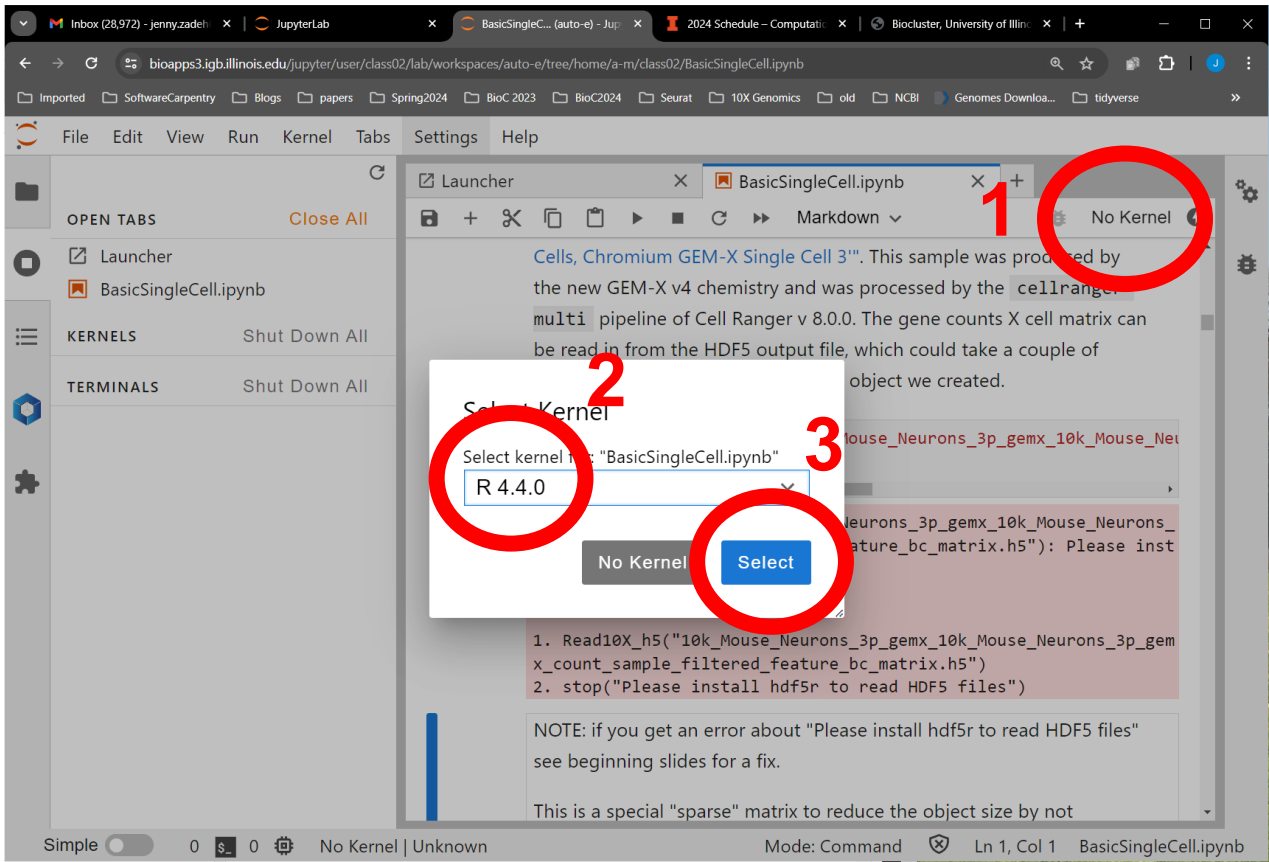
Stop Kernel

1. Click on gray circle with white square on left
2. Under KERNELS click on the X by BasicSingleCell.ipynb



Re-start Kernel

1. Click on the "No Kernel" in the top right of the .ipynb to get the pop-up
2. Pick R 4.4.0 from the drop down menu
3. Click on "Select"



Re-run the code boxes

You need to re-run the **first code box** to load the packages from the library, then you can successfully run the second code box