

NeMO

the Neuroscience Multi-Omic Archive

Biol 367
Dean Symonds and Hailey Ivanson

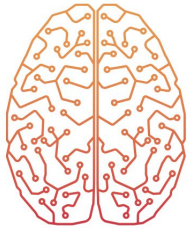
Department of Biology
Loyola Marymount University
February 8, 2024

Outline

- NeMO is a Neuroscience database with sequence data from a variety of mammals
- The data contained on the site is extensive and very comprehensive, but almost excessive
- The database has a nice appearance and has potential to function well, but it is not very easy to use and is even slightly misleading
-

NeMO, or Neuroscience Multi-Omic Archive, is a repository for single-cell genomic and transcriptomic data of various brain cells.

Figure 1. NeMO logo.



NeMO
The Neuroscience Multi-omic Data Archive

From NeMO landing page (2024). Retrieved from NeMOarchive.org.

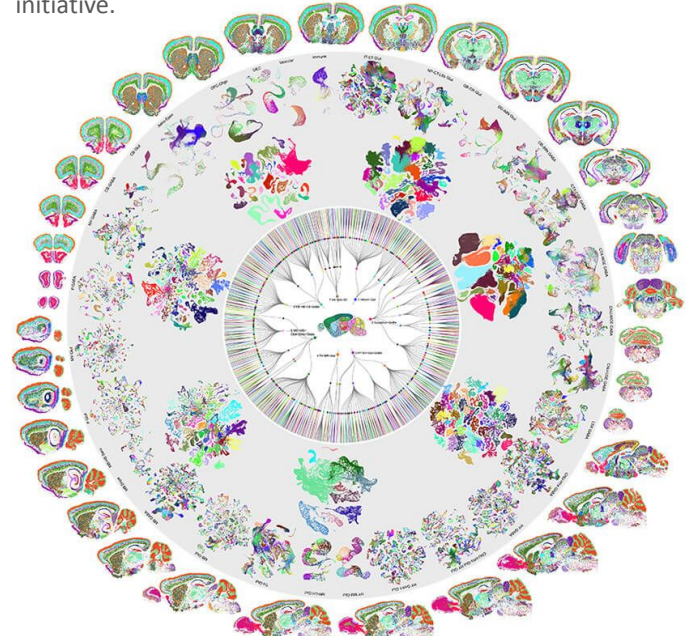
- NeMO's goal \Rightarrow help map all cell types in mammalian brains
- Genomic and transcriptomic data with morphology can help identify cell type and function
- More data in repository = more accurate maps!

(Ament et al., 2023)

The NIH has funded NeMO to increase data FAIRness and expand the access of the BRAIN initiative.

- FAIRness = Findable, accessible, interoperable, reusable
- NeMO for BICCN [BRAIN (Brain Research Through Advancing Innovative Neurotechnologies) Initiative Cell Census Network] projects for access by researchers everywhere
 - BICCN Goal: map human and mouse brains

Figure 2. Cellular map of mouse brain, funded by NIH BRAIN initiative.

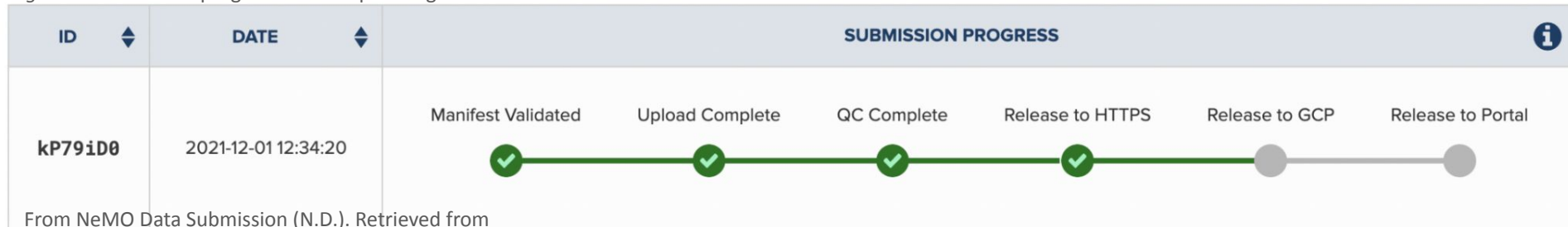


From Whole_mouse_brain_classification By Jake Siegel (2023). Retrieved from alleninstitute.org/news/scientists-unveil-first-complete-cellular-map-of-adult-mouse-brain/

Research groups can go through NeMO's process of uploading data to be included in the repository.

- Lab PI must have an ORCID ID to make a NeMO Aspera account
- Must fill out form for manifest file
- Electronically validated
 - If error, internal NeMO team member will step in
- Defaults to “restricted” once submitted

Figure 3. Submission progress bar for uploading to NeMO.



From NeMO Data Submission (N.D.). Retrieved from nemoarchive.org/resources/data-submission

(NeMO, 2024)

NeMO employs three categories of data post-submission:

1. *Public*

- Immediate free access for public browsing; no account required
 - This is the goal category for all data by NeMO and the NIH

2. *Embargo*

- Will be made available for public browsing at a later date

3. *Restricted*

- May contain sensitive (human) information. Consent will be obtained at a later date or not at all. Potentially made public at a later date

NeMO is constantly being improved by its team.

- Intended features for the future include:
 - Annotating cell types and other metadata for entries
 - Integrate epigenomes + transcriptomes with their function + morphology
 - Multi-modal cell model integration (physiology + imaging)

NeMO has a navigable and well-organized landing page.

- Page is clean, containing multiple different sections
- Summary of the site is informative in its coverage
- Different tabs at the top for site-wide navigation
- Data portal is also easy to access and has a clear location on the site

Figure 4. NeMO landing page.

The Neuroscience Multi-omic Archive

The Neuroscience Multi-omic Archive (NeMO Archive) is a data repository specifically focused on the storage and dissemination of omic data generated from the BRAIN Initiative, SCORCH consortium and other brain research projects. [NeMO Data Submission \(N.D.\). Retrieved from nemoarchive.org/resources/data-submission](https://nemoarchive.org/resources/data-submission)

Welcome

NeMO data will include:

1. Genomic regions associated with brain abnormalities and disease
2. Transcription factor binding sites and other regulatory elements
3. Transcription activity
4. Levels of cytosine modification
5. Histone modification profiles and chromatin accessibility

To search the NeMO Archive please visit the [NeMO Archive Data Portal](#).

Data Volume	416.3 TB
Files	1,225,515
Downloads	2.7 PB

NeMO Analytics NeMO Analytics enables web-based visualization and analysis of multi-omic data. [Learn more →](#)

From NeMO landing page (2024). Retrieved from NeMOarchive.org.

NeMO has an extensive data portal... almost too extensive for its level of organization.

- Search engine has several different filters:
 - Grants, organism, brain region, brain sub-region, sub specimen species, format, data type...
- Certain filters have a lot of different options
 - Brain regions has over 800+ options
 - Brain sub-regions has 600+ options
- The database has a total of 2,964,588 files

Samples Files << Hide Filters

[Add a filter](#)

▼ Brain Region ↓^A/_Z

<input type="checkbox"/> NT	510,666
<input type="checkbox"/> NP	510,666
<input type="checkbox"/> Br	510,662
<input type="checkbox"/> F	476,468
<input type="checkbox"/> FGM	450,601

[798 More...](#)

▼ Brain Sub- Region ↓^A/_Z

<input type="checkbox"/> BNST	17,903
<input type="checkbox"/> M1C	17,876
<input type="checkbox"/> AMY	17,835
<input type="checkbox"/> SI	17,596
<input type="checkbox"/> CaB	17,595

[672 More...](#)

NeMO Data Portal filters.

Image via Portal.NEMOarchive.org

Figure 4. NeMO landing page.



NeMO

The Neuroscience Multi-omic Data Archive

LOGIN 

HOME

ABOUT

DATA

RESOURCES

CONTACT

The Neuroscience Multi-omic Archive

The Neuroscience Multi-omic Archive (NeMO Archive) is a data repository specifically focused on the storage and dissemination of omic data generated from the BRAIN Initiative, SCORCH consortium and other brain research projects.






Welcome

NeMO data will include:

1. Genomic regions associated with brain abnormalities and disease
2. Transcription factor binding sites and other regulatory elements
3. Transcription activity
4. Levels of cytosine modification
5. Histone modification profiles and chromatin accessibility

To search the NeMO Archive please visit the [NeMO Archive Data Portal](#).

From NeMO landing page (2024). Retrieved from NeMOarchive.org.

	Data Volume	416.3 TB
	Files	1,225,515
	Downloads	2.7 PB

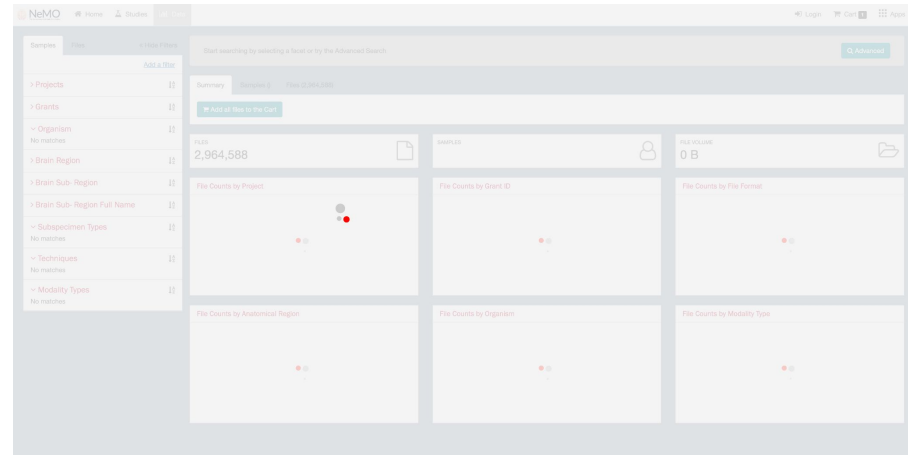
NeMO
Analytics

NeMO Analytics enables web-based visualization and analysis of multi-omic data. [Learn more](#) →

Searching on NeMO is time-consuming.

- Adding a filter \Rightarrow +multiple minutes to load
 - +time with each additional filter
- Possible reason: NeMO may query the database with every user query variation.
 - Possible solution: Caching database periodically
 - Sacrifices up-to-dateness of display

Figure 5. NeMO Data Portal display while loading.



From NeMO Data Portal (2024). Retrieved from Portal.NeMOarchive.org.

NeMO does have standard data formats, although they are often quite large

- The most common file types are FASTQ, TSV, and BAM
 - FASTQ files contain the biological sequence as well as the score associated with that sequence
 - TSV is a text based file
 - BAM is a biological sequence in binary
- Getting to these files also has the same problems that running a search has
- Once you have accessed a file, it is easy to download it

Figure 6. NeMO Data Portal showing downloadable files' IDs and names.

Sample ID	File Name
03_13_2023_Ai75_Negative_Neun_3_Negative	pBICCNsMMrCBFPurKf004d190219_mex.tar.gz
03_13_2023_Ai75_Negative_Neun_3_Positive	pBICCNsMMrCB COPIM006d190313_CellRangerV5_bam.tar
03_13_2023_Ai75_Negative_Neun_8_Negative	pBICCNsMMrCBAN2IM002d190312b_SS_L002_fasta.tar
03_13_2023_Ai75_Negative_Neun_8_Positive	pBICCNsMMrCBAN2IM002d190312b_mex.tar.gz
03_13_2023_Ai75_Positive	pBICCNsMMrCBXGd180108_mex.tar.gz
05_17_2023_Ai75_Negative	pBICCNsMMrCBAN1IM006d190321a_CellRangerV5_bam.tar
05_17_2023_Ai75_Positive	pBICCNsMMrCB SMM002d190312d_CellRangerV5_bam.tar
1028166110	pBICCNsMMrCBAN2IM002d190312b_SS_L001_fasta.tar
1031544486	pBICCNsMMrCBIXIM003d181129c_CellRangerV5_bam.tar
1031667752	pBICCNsMMrCBAN2IM002d190312b_mex.tar.gz
10E_HCV_3C	pBICCNsMMrCBPRMIM002d190312c_mex.tar.gz
10F_DGV_3C	pBICCNsMMrCBIXIM003d181129c_mex.tar.gz
10X120-3	pBICCNsMMrCB COPIM006d190313_mex.tar.gz
10X120-4	pBICCNsMMrCBAN1IM006d190321a_mex.tar.gz
10X128-1	pBICCNsMMrCBPFIM006d190124_mex.tar.gz
10X128-2	pBICCNsMMrCBXGd180108_CellRangerV5_bam.tar
10X131-1	pBICCNsMMrCBPFIM006d190124_CellRangerV5_bam.tar
10X131-2	pBICCNsMMrCBFPurKf004d190219_CellRangerV5_bam.tar
10X131-3	pBICCNsMMrCB SMM002d190312d_mex.tar.gz
10X131-4	pBICCNsMMrCBPRMIM002d190312c_CellRangerV5_bam.tar

From NeMO Data Portal (2024). Retrieved from Portal.NeMOarchive.org.

Overall, NeMO is not a user-friendly database

- NeMO has potential to be user-friendly
 - Appears well-organized, has good quality data
- NeMO's execution and functionality needs improvement
 - NeMO archive also could be narrowed in its focus
- Naive user could use this database just as well as an experienced user
- No tutorial videos on the website
 - Some help pages present on website. Others listed in databases article

NeMO is a step in the right direction to map mammalian brains.

- NeMO's goals are noble.
 - Make data accessible
 - Encourage neurobiological research
 - Advance understanding of mammalian brains
- Anyone can access NeMO's public data, but not everyone can add to it.
- There are a few kinks that should be worked out.
 - Zebrafish in a mammal-focused database?
 - Load times

References

Ament, S. A., Adkins, R. S., Carter, R., Chrysostomou, E., Colantuoni, C., Crabtree, J., Creasy, H. H., Degatano, K., Felix, V., Gandt, P., Garden, G. A., Giglio, M., Herb, B. R., Khajouei, F., Kiernan, E., McCracken, C., McDaniel, K., Nadendla, S., Nickel, L., ... White, O. R. (2023). The Neuroscience Multi-Omic Archive: A BRAIN Initiative resource for single-cell transcriptomic and epigenomic data from the mammalian brain. *Nucleic Acids Research*, 51(D1), D1075–D1085. <https://doi.org/10.1093/nar/gkac962>

BRAIN Initiative Cell Census Network | *BRAIN Initiative*. (n.d.). Retrieved February 7, 2024, from

<https://braininitiative.nih.gov/research/tools-and-technologies-brain-cells-and-circuits/brain-initiative-cell-census-network>

NeMO Archive—Home. (2024). Retrieved February 7, 2024, from <https://nemoarchive.org/>

Neuroscience Multi-Omic (NeMO) Data Portal. (2024). Retrieved February 7, 2024, from <https://portal.nemoarchive.org/>

Siegel, J. (2023, December 13). *Scientists unveil first complete cellular map of adult mouse brain*. Allen Institute.

<https://alleninstitute.org/news/scientists-unveil-first-complete-cellular-map-of-adult-mouse-brain/>



Questions?

NeMO

the Neuroscience Multi-Omic Archive

Biol 367
Dean Symonds and Hailey Ivanson

Department of Biology
Loyola Marymount University
February 8, 2024