**File download instructions**

**PatchSeq file release**

**June 25, 2020**

Download transcriptomics files from the NeMO archive

The Neuroscience Multi-omic Data Archive (NeMO) hosts the transcriptomics data files. File access options consist of HTTP (web browser), Aspera, NeMO command line client [link: https://github.com/IGS/portal\_client] and Terra. More information about NeMO and each of these options is available at <https://nemoarchive.org/data/>

The paths to individual data files are listed in the [file manifest] in the ‘archive\_uri’ column. Directory paths are available here:

Mouse fastq files (460 GB):

<http://data.nemoarchive.org/other/grant/AIBS_patchseq/transcriptome/scell/SMARTseq/raw/20200611/>

Note that NeMO combines forward and reverse fastq files into a single “.tar” archive file. This will need to be unpacked in order to access each of the fastq files (which themselves are compressed with gzip). The checksum values listed in the file manifest are for the individual gzipped fastq files, after they have been extracted from the .tar bundle. The download link provided in the archive\_uri column of the file manifest will download the .tar bundle containing BOTH the forward and reverse fastq files – so please note that you may only want to download the set a single time.

Mouse BAM files (559 GB):

<http://data.nemoarchive.org/other/grant/AIBS_patchseq/transcriptome/scell/SMARTseq/processed/align/20200611/>

Note that the checksum values provided in the file manifest are for the files after they have been extracted from the .tar bundle.

Mouse summary analysis files (2.5 GB):

<http://data.nemoarchive.org/other/grant/AIBS_patchseq/transcriptome/scell/SMARTseq/processed/analysis/20200611/>

* Exon file <- reads aligned to exons
* Intron file <- reads aligned to introns
* Count file <- Exon + Intron reads
* CPM file <- Normalized Count matrix (Count/million reads)

 Note that the checksum values provided in the file manifest are for the files after they have been extracted from the .tar bundle.

**Human transcriptomics data:**

As noted on the NeMO website, they are still in the process of establishing procedures for accessing restricted data. Until this work is complete on their end, access to the human data files is not possible.

Download electrophysiology files from the DANDI archive

The Distributed Archives for Neurophysiology Data Integration (DANDI) hosts electrophysiology files. DANDI supports data download using HTTP and the DANDI command line client [pip install dandi].

The paths to individual data files are listed in the [file manifest]. Directory paths are available here:

Mouse data (114 GB): [https://dandiarchive.org/dandiset/000020](https://dandiarchive.org/dandiset/000020/draft)

Human data (12 GB): [https://dandiarchive.org/dandiset/000023](https://dandiarchive.org/dandiset/000023/draft)

Example:

**Download all mouse ephys data:**

pip install dandi

dandi download <https://dandiarchive.org/dandiset/000020>

**Download a single ephys experiment file:**

Refer to the file manifest for the value in the ‘archive\_uri’ column. Use that to download with your web browser or with the ‘dandi’ command line tool.

dandi download <https://girder.dandiarchive.org/api/v1/item/5edab04e2dace54b6f9b2a01/download>

Download neuron reconstruction files from the BIL archive

The Brain Image Library (BIL) hosts neuron reconstruction files. BIL supports file downloads using HTTP, SFTP and Globus. Further information on each method is available at <http://www.brainimagelibrary.org/download.html> .

The paths to individual data files are listed in the [file manifest]. Directory paths are available here:

Mouse data (1.3 GB): ftp://download.brainlib.org:8811/biccn/zeng/pseq/morph/200526/

Human data (211 MB): ftp://download.brainlib.org:8811/biccn/lein/pseq/morph/200526/

Linux example:

Get all human reconstruction files

wget -r ftp://download.brainlib.org:8811/biccn/lein/pseq/morph/200526/