

Project 2

Project 2 focuses on a single-cell sequencing study of the *Drosophila* brain following acute cocaine exposure; [paper](#). Flies were exposed to cocaine, which impaired locomotor activity and increased the incidence of seizures and compulsive grooming. To investigate the specific cell populations responding to cocaine, single-cell transcriptional responses were analyzed in duplicate samples of flies that consumed sucrose or sucrose supplemented with cocaine. The study utilized the 10x Genomics Chromium platform for single-cell RNA sequencing.

Unsupervised clustering of transcriptional profiles from 86,224 cells revealed 36 distinct clusters, representing all major cell types (neuronal and glial) and neurotransmitter types from most brain regions. Differential expression analysis within individual clusters indicated cluster-specific responses to cocaine, with Kenyon cells of the mushroom bodies and glia showing particularly large transcriptional changes. The study highlighted profound sexual dimorphism in brain transcriptional responses to cocaine, with males exhibiting more pronounced changes than females.

Cluster-specific coexpression networks and global interaction networks revealed diverse cellular processes affected by acute cocaine exposure, providing an atlas of sexually dimorphic cocaine-modulated gene expression in the *Drosophila* brain.

Available data

Data has been downloaded and prepared for you from [GEO GSE152495](#).

In order to download the data, run:

Code starts here:

```
wget https://single-cell-transcriptomics.s3.eu-central-1.amazonaws.com/projects/data/project2.tar.gz
tar -xzf project2.tar.gz
```

Code ends here

After extracting, a directory project2 appears with the following format:

```
.
├── data
│   ├── Female_Cocaine_1
│   │   ├── barcodes.tsv.gz
│   │   ├── features.tsv.gz
│   │   └── matrix.mtx.gz
│   └── Female_Cocaine_2
│       ├── barcodes.tsv.gz
│       ├── features.tsv.gz
│       └── matrix.mtx.gz
```

```

├── Female_Sucrose_1
│   ├── barcodes.tsv.gz
│   ├── features.tsv.gz
│   └── matrix.mtx.gz
├── Female_Sucrose_2
│   ├── barcodes.tsv.gz
│   ├── features.tsv.gz
│   └── matrix.mtx.gz
├── Male_Cocaine_1
│   ├── barcodes.tsv.gz
│   ├── features.tsv.gz
│   └── matrix.mtx.gz
├── Male_Cocaine_2
│   ├── barcodes.tsv.gz
│   ├── features.tsv.gz
│   └── matrix.mtx.gz
├── Male_Sucrose_1
│   ├── barcodes.tsv.gz
│   ├── features.tsv.gz
│   └── matrix.mtx.gz
├── Male_Sucrose_2
│   ├── barcodes.tsv.gz
│   ├── features.tsv.gz
│   └── matrix.mtx.gz
└── paper.pdf

```

9 directories, 25 files

Showing us that we have two replicates per treatment, and two treatments:

- Male_Sucrose: controls
- Female_Sucrose: controls
- Male_Cocaine: treatments
- Female_Cocaine: treatments

Now create a new project in the project2 directory (Project (None) > New Project ...), and create Seurat object from the count matrices:

Code starts here:

```

library(Seurat)

# vector of paths to all sample directories
datadirs <- list.files(path = "data", full.names = TRUE)

# get the sample names
# replace underscores with hyphen to correctly extract sample names later on
names(datadirs) <- basename(datadirs) |> gsub("_", "-", x = _)

```

```
# for now, we take only one sample per treatment
datadirs <- datadirs[c("Female-Cocaine-1", "Female-Sucrose-1",
                      "Male-Cocaine-1", "Male-Sucrose-1")]

# create a large sparse matrix from all count data
sparse_matrix <- Seurat::Read10X(data.dir = datadirs)

# create a seurat object from sparse matrix
seu <- Seurat::CreateSeuratObject(counts = sparse_matrix,
                                  project = "CocaineStudy")
```

Code ends here

Project exercise

With this dataset, go through the steps we have performed during the course, and try to reproduce the results provided in the paper. Pay specific attention to quality control, clustering and annotation.

Tips

- For mitochondrial genes and ribosomal genes you can use the following patterns: "^mt:" and "^Rp[SL]".
- Work iterative; meaning that based on results of an analysis, adjust the previous analysis. For example, if clustering is not according to cell types, try to adjust the number of components or the resolution.
- Please read the methods section of the paper.
- If the code for data analysis is available, try to adapt it (for specific parameters).
- Check the supplementary figures.
- Try to understand if they used some other tools for the data analysis.