SPATIAL TRANSCRIPTOMICS DATA ANALYSIS: THEORY AND PRACTICE

PRACTICAL SESSION 1

DR SIMON J COCKELL

ELEFTHERIOS (LEFTERIS) ZORMPAS

BIOSCIENCES INSTITUTE,

FACULTY OF MEDICAL SCIENCES,

NEWCASTLE UNIVERSITY

23/07/2023

COURSE DOCUMENTATION

https://bookdown.org/sjcockell/ismb-tutorial-2023

In this practical session, you will familiarise yourself with some example spatial transcriptomics (STx) data and the common features of such data.

Log in to the Posit Cloud Space:

bit.ly/ISMB-IP2-Posit

➢Required packages are pre-installed.
➢Data files uploaded.
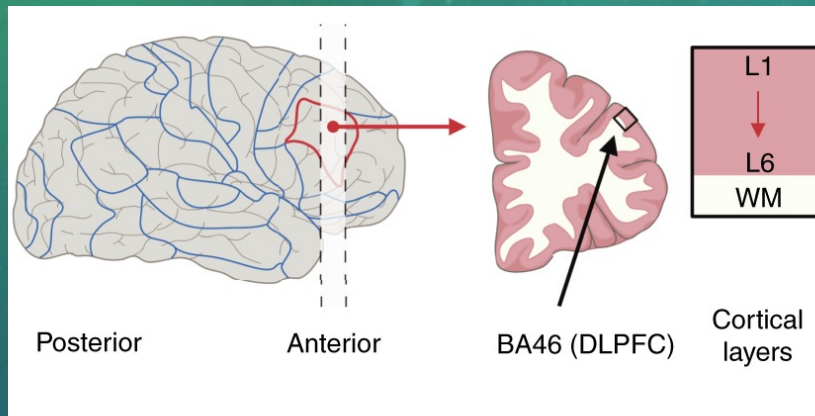➢You get your own copy of the workspace.

Log in to the Posit Cloud Space

➢ Required packages are pre-installed.
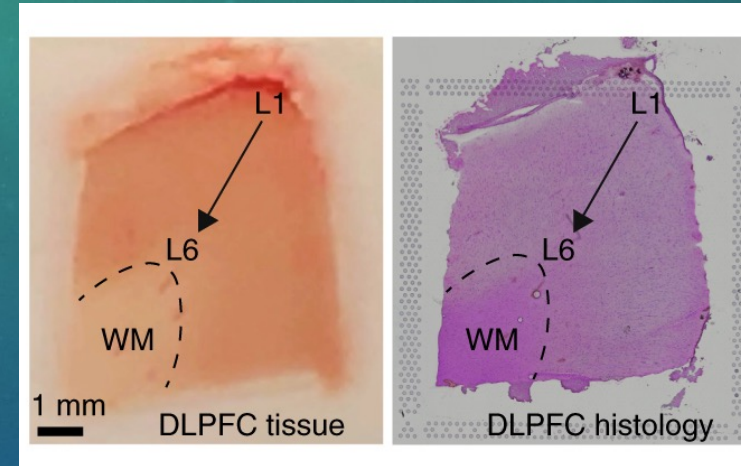➢ Data files uploaded.
➢ You get your own copy of the workspace.

1.3 Import 10X Visium data

In this tutorial session, we will be using data from the STexampleData package.

The dataset:

DLPFC: Dorso-Lateral Pre-Frontal Cortex

For the purpose of this tutorial, we will be using 1 of the 12 samples: sample 151673

➢To Do: Load the dataset and work on the practical's code!

1.4 Explore data types

SpatialExperiment class ➡ A specialized object class that supports the storing of spatially-resolved transcriptomics datasets within the Bioconductor framework.

1. **assays:** gene expression counts

2. **rowData:** information about features, usually genes

3. **colData:** information on spots (non-spatial and spatial metadata)

4. **spatialCoords:** spatial coordinates

5. **imgData:** image data
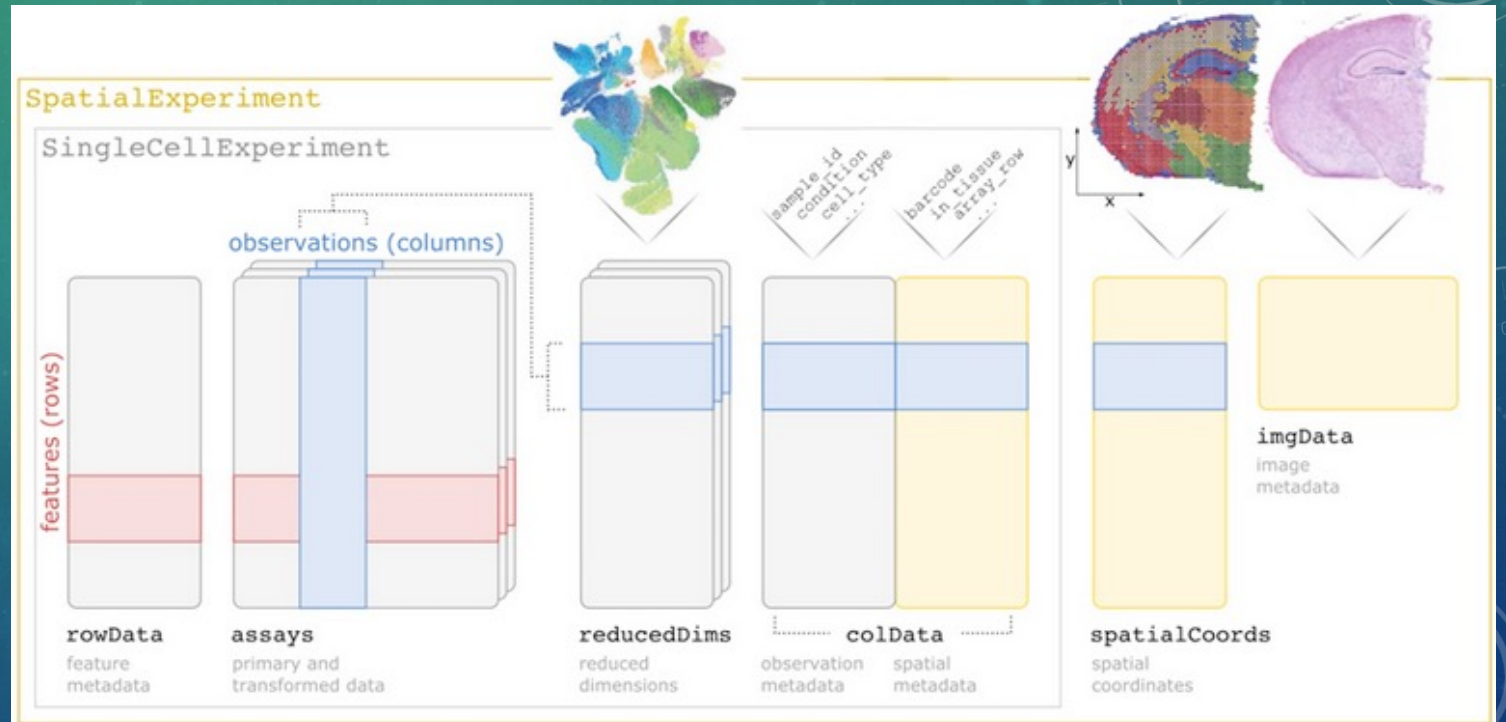
1.4 Explore data types

```
## class: SpatialExperiment
## dim: 33538 4992
## metadata(0):
## assays(1): counts
## rownames(33538): ENSG00000243485 ENSG00000237613 ... ENSG00000277475
##   ENSG00000268674
## rowData names(3): gene_id gene_name feature_type
## colnames(4992): AAACAACGAATAGTTC-1 AAACAAGTATCTCCCA-1 ...
##   TTGTTTGTATTACACG-1 TTGTTTGTGTAAATTC-1
## colData names(7): barcode_id sample_id ... ground_truth cell_count
## reducedDimNames(0):
## mainExpName: NULL
## altExpNames(0):
## spatialCoords names(2) : pxl_col_in_fullres pxl_row_in_fullres
## imgData names(4): sample_id image_id data scaleFactor
```

1.4 Explore data types

Sparse gene expression data

```
assay(spe)[33488:33508, 2000:2010]
```

```
## 21 x 11 sparse Matrix of class "dgTMatrix"
##
## ENSG00000160294  .   .    .   .    .   .    .   .   .    .   .
## ENSG00000228137  .   .    .   .    .   .    .   .   .    .   .
## ENSG00000239415  .   .    .   .    .   .    .   .   .    .   .
## ENSG00000182362  .   .    .   .    .   .    .   .   .    1   .   .
## ENSG00000160298  .   .    .   .    .   .    .   .   .    .   .
## ENSG00000160299  .   .    1   .    1   .    .   .   .    .   .
## ENSG00000160305  .   .    .   .    .   2    .   .   .    .   .
## ENSG00000160307  1   3    1   1    4   5    1   1   .    2   1
## ENSG00000160310  .   .    .   .    1   .    .   .   .    2   .
## ENSG00000198888 17  44   71  16  154  97   12  14  32  167   6
## ENSG00000198763 16  59   64  11  116  63   11  12  18  123   6
## ENSG00000198804 37  85  155  25  252 176   24  27  38  335  12
## ENSG00000198712 23  79  120  23  214 170   22  25  48  242  10
## ENSG00000228253  2   .    3   .    1   .    .   1   1    6   .
## ENSG00000198899 20  39   93   9  136 108   20  18  25  165   7
## ENSG00000198938 27  59  133  20  216 120   22  26  43  232   9
## ENSG00000198840  5  27   33   5   71  39    8  11  12   78   .
## ENSG00000212907  2   .    4   2    7   5    .   1   1    9   .
## ENSG00000198886 15  65   95   9  183  98   18  19  33  178   7
## ENSG00000198786  2  10   10   3   20  14    1   2   2   25   4
## ENSG00000198695  1   1    3   .    2   2    .   .   .    1   .
```

1.4 Explore data types

Gene-level metadata

```
## Have a look at the genes metadata
head(rowData(spe))
```

```
## DataFrame with 6 rows and 3 columns
##                         gene_id    gene_name    feature_type
##                       <character> <character>    <character>
## ENSG00000243485 ENSG00000243485 MIR1302-2HG Gene Expression
## ENSG00000237613 ENSG00000237613     FAM138A Gene Expression
## ENSG00000186092 ENSG00000186092       OR4F5 Gene Expression
## ENSG00000238009 ENSG00000238009  AL627309.1 Gene Expression
## ENSG00000239945 ENSG00000239945  AL627309.3 Gene Expression
## ENSG00000239906 ENSG00000239906  AL627309.2 Gene Expression
```

1.4 Explore data types

Spot coordinates


```
## Check the spatial coordinates
head(spatialCoords(spe))
```

	pxl_col_in_fullres	pxl_row_in_fullres
## AAACAACGAATAGTTC-1	3913	2435
## AAACAAGTATCTCCCA-1	9791	8468
## AAACAATCTACTAGCA-1	5769	2807
## AAACACCAATAACTGC-1	4068	9505
## AAACAGAGCGACTCCT-1	9271	4151
## AAACAGCTTTCAGAAG-1	3393	7583

1.4 Explore data types

Spot-level metadata

```
## spot-level metadata
head(colData(spe))
```

```
## DataFrame with 6 rows and 7 columns
##                              barcode_id      sample_id in_tissue array_row
##                             <character>    <character> <integer> <integer>
## AAACAACGAATAGTTC-1 AAACAACGAATAGTTC-1 sample_151673         0         0
## AAACAAGTATCTCCCA-1 AAACAAGTATCTCCCA-1 sample_151673         1        50
## AAACAATCTACTAGCA-1 AAACAATCTACTAGCA-1 sample_151673         1         3
## AAACACCAATAACTGC-1 AAACACCAATAACTGC-1 sample_151673         1        59
## AAACAGAGCGACTCCT-1 AAACAGAGCGACTCCT-1 sample_151673         1        14
## AAACAGCTTTCAGAAG-1 AAACAGCTTTCAGAAG-1 sample_151673         1        43
##                     array_col ground_truth cell_count
##                     <integer>  <character>  <integer>
## AAACAACGAATAGTTC-1        16           NA         NA
## AAACAAGTATCTCCCA-1       102       Layer3          6
## AAACAATCTACTAGCA-1        43       Layer1         16
## AAACACCAATAACTGC-1        19           WM          5
## AAACAGAGCGACTCCT-1        94       Layer3          2
## AAACAGCTTTCAGAAG-1         9       Layer5          4
```
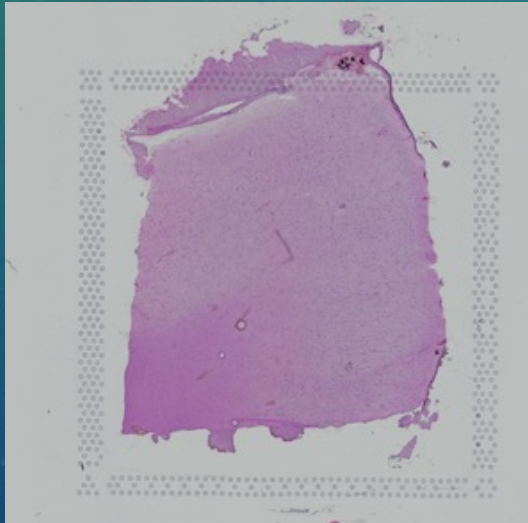
1.4 Explore data types

Image data and metadata

```
## Have a look at the image metadata
imgData(spe)


## DataFrame with 2 rows and 4 columns
##       sample_id    image_id   data scaleFactor
##      <character> <character> <list>   <numeric>
## 1 sample_151673      lowres   ####   0.0450045
## 2 sample_151673       hires   ####   0.1500150
```

Eleftherios Zormpas

Dr Simon J Cockell

Dr Rachel Queen

Prof. Alex Comber

Newcastle University

Newcastle University

Newcastle University

UNIVERSITY OF LEEDS

iSMB feedback form:

Newcastle University

14