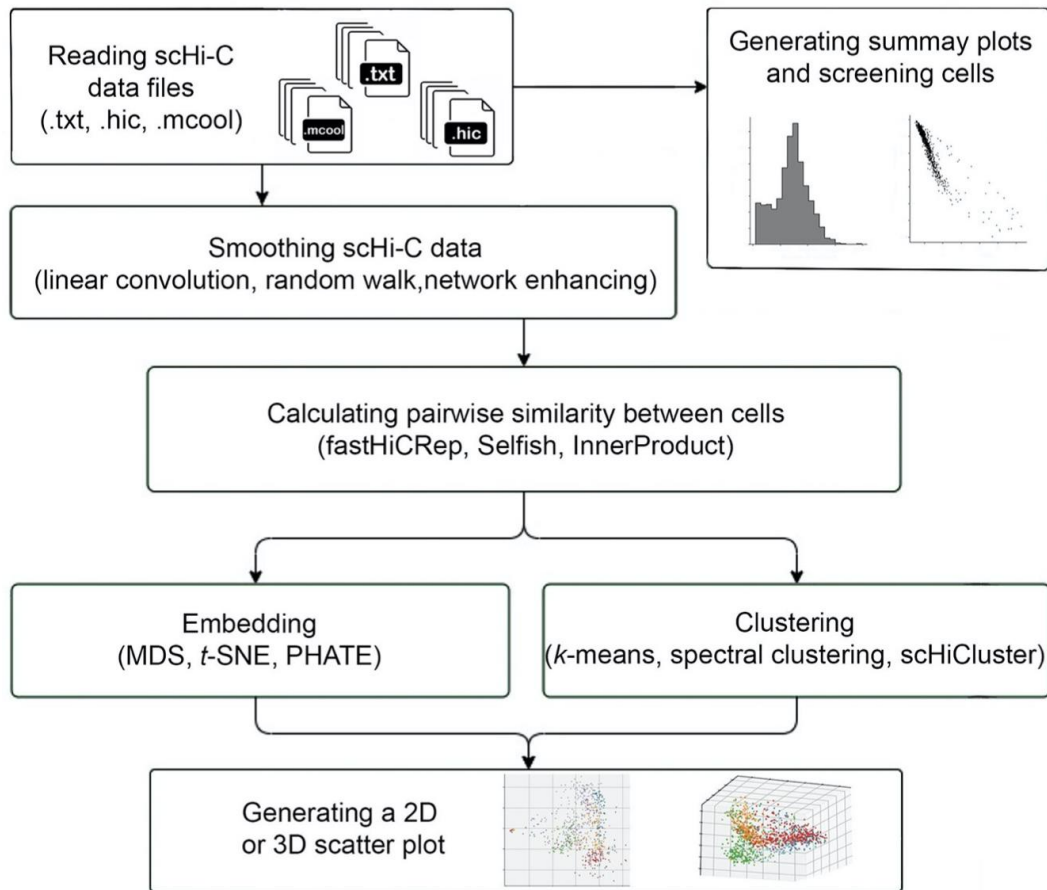# scHiCTools: A computational toolbox for analyzing single-cell Hi-C data

資科碩一 109753102 黃渝庭

Workflow

Install scHiCTools

Install from GitHub

You can install the package with following command:

```
$ git clone https://github.com/liu-bioinfo-lab/scHiCTools.git
$ cd scHiCTools
$ python setup.py install
```

Install from PyPI
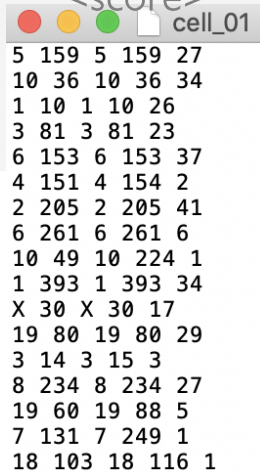
```
$ pip install scHiCTools
```

Loading scHiC Data

.txt   .hic   .mcool

□□□□□□□□□□□□□ ⌐□□□□□□
<chr1> <pos1> <chr2> <pos2> <score>

```
cell_01
5 159 5 159 27
10 36 10 36 34
1 10 1 10 26
3 81 3 81 23
6 153 6 153 37
4 151 4 154 2
2 205 2 205 41
6 261 6 261 6
10 49 10 224 1
1 393 1 393 34
X 30 X 30 17
19 80 19 80 29
3 14 3 15 3
8 234 8 234 27
19 60 19 88 5
7 131 7 249 1
18 103 18 116 1
```

```
files = ['test/data/cell_01', 'test/data/cell_02', 'test/data/cell_03']
loaded_data = scHiCs(
files, reference_genome='mm9',
    resolution=50000,
    max_distance=4000000,
    format='shortest_score',
    adjust_resolution=True,
    chromosomes='except Y',
    operations=[smoothing_method],
    kernel_shape=3,
    keep_n_strata=10,
    store_full_map=True
)
```

Plot number of contacts and select
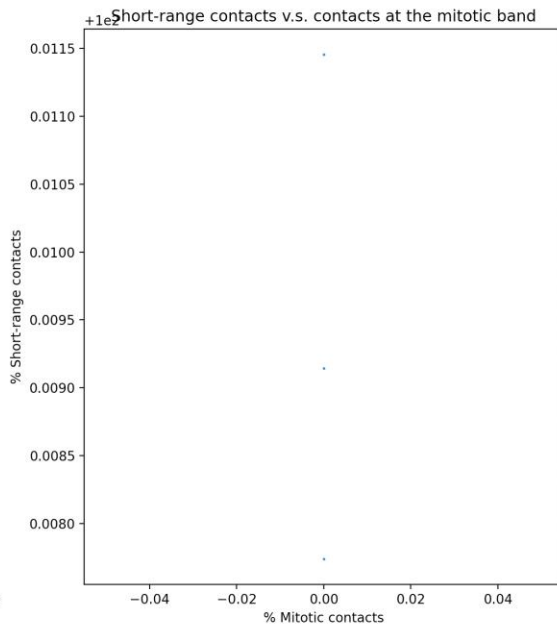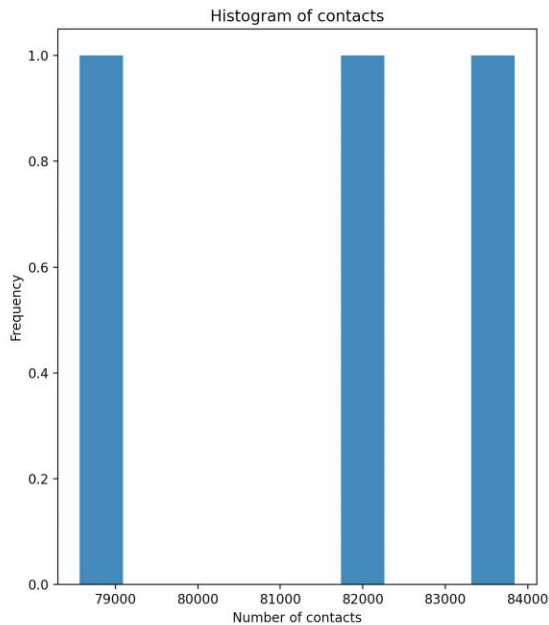

```
loaded_data.plot_contacts(hist = True,
                          percent = True)
```

Smoothing & Embedding Data and ualize
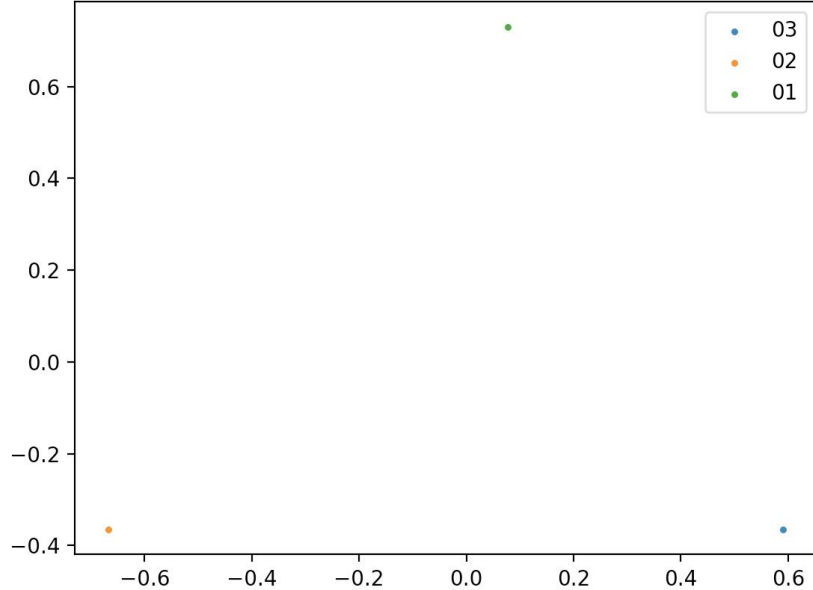

```python
embs = loaded_data.learn_embedding(
dim = 2, similarity_method = similarity_method,
embedding_method = embedding_method,
n_strata = None,aggregation = 'median',
return_distance = False)
scatter(embs,label = ['01','02','03'],point_size = 5)
```

Compare With Seurat

- Both of them do for analyzing single-cell

- scHiCTools is only for HiC data and Seurat is only for RNA-Seq

- The workflow of two tools is similar, but scHicTools have on more step for calculating pairwise similarity

The End