

```
!pip install -q condacolab
import condacolab
condacolab.install()
```

```
from google.colab import drive
drive.mount('/content/drive')
```

```
mcool = 'drive/MyDrive/hse_minor_bioinfo/HiC1.dm3.mapq_30.1000.mcool'
```

WARNING: Running pip as the 'root' user can result in broken permissions and conflicting behaviour with the system package manager. It is recommended to use a virtual environment instead:

<https://pip.pypa.io/warnings/venv>

🍷🍰🍷 Everything looks OK!

Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/drive", force_remount=True).

```
%%bash
```

```
pip install cooler
pip install hicrep
```

Looking in indexes: <https://pypi.org/simple>, <https://us-python.pkg.dev/colab-wheels/public/simple/>

Requirement already satisfied: cooler in

/usr/local/lib/python3.10/site-packages (0.9.1)

Requirement already satisfied: asciitree in

/usr/local/lib/python3.10/site-packages (from cooler) (0.3.3)

Requirement already satisfied: cytoolz in

/usr/local/lib/python3.10/site-packages (from cooler) (0.12.1)

Requirement already satisfied: multiprocessing in

/usr/local/lib/python3.10/site-packages (from cooler) (0.70.14)

Requirement already satisfied: simplejson in

/usr/local/lib/python3.10/site-packages (from cooler) (3.19.1)

Requirement already satisfied: pyfaidx in

/usr/local/lib/python3.10/site-packages (from cooler) (0.7.2.1)

Requirement already satisfied: h5py>=2.5 in

/usr/local/lib/python3.10/site-packages (from cooler) (3.8.0)

Requirement already satisfied: scipy>=0.16 in

/usr/local/lib/python3.10/site-packages (from cooler) (1.10.1)

Requirement already satisfied: pandas>1.0 in

/usr/local/lib/python3.10/site-packages (from cooler) (2.0.1)

Requirement already satisfied: pyyaml in

/usr/local/lib/python3.10/site-packages (from cooler) (6.0)

Requirement already satisfied: numpy>=1.9 in

/usr/local/lib/python3.10/site-packages (from cooler) (1.24.3)

Requirement already satisfied: click>=7 in

/usr/local/lib/python3.10/site-packages (from cooler) (8.1.3)

Requirement already satisfied: tzdata>=2022.1 in

/usr/local/lib/python3.10/site-packages (from pandas>1.0->cooler) (2023.3)

Requirement already satisfied: python-dateutil>=2.8.2 in

/usr/local/lib/python3.10/site-packages (from pandas>1.0->cooler)
(2.8.2)
Requirement already satisfied: pytz>=2020.1 in
/usr/local/lib/python3.10/site-packages (from pandas>1.0->cooler)
(2023.3)
Requirement already satisfied: toolz>=0.8.0 in
/usr/local/lib/python3.10/site-packages (from cytoolz->cooler)
(0.12.0)
Requirement already satisfied: dill>=0.3.6 in
/usr/local/lib/python3.10/site-packages (from multiprocessing->cooler)
(0.3.6)
Requirement already satisfied: six in /usr/local/lib/python3.10/site-
packages (from pyfaidx->cooler) (1.16.0)
Requirement already satisfied: setuptools in
/usr/local/lib/python3.10/site-packages (from pyfaidx->cooler)
(65.6.3)
Looking in indexes: <https://pypi.org/simple>, [https://us-
python.pkg.dev/colab-wheels/public/simple/](https://us-python.pkg.dev/colab-wheels/public/simple/)
Requirement already satisfied: hicrep in
/usr/local/lib/python3.10/site-packages (0.2.6)
Requirement already satisfied: scipy in
/usr/local/lib/python3.10/site-packages (from hicrep) (1.10.1)
Requirement already satisfied: Deprecated in
/usr/local/lib/python3.10/site-packages (from hicrep) (1.2.13)
Requirement already satisfied: h5py in /usr/local/lib/python3.10/site-
packages (from hicrep) (3.8.0)
Requirement already satisfied: pandas in
/usr/local/lib/python3.10/site-packages (from hicrep) (2.0.1)
Requirement already satisfied: cooler in
/usr/local/lib/python3.10/site-packages (from hicrep) (0.9.1)
Requirement already satisfied: numpy>=1.17.0 in
/usr/local/lib/python3.10/site-packages (from hicrep) (1.24.3)
Requirement already satisfied: cytoolz in
/usr/local/lib/python3.10/site-packages (from cooler->hicrep) (0.12.1)
Requirement already satisfied: multiprocessing in
/usr/local/lib/python3.10/site-packages (from cooler->hicrep)
(0.70.14)
Requirement already satisfied: simplejson in
/usr/local/lib/python3.10/site-packages (from cooler->hicrep) (3.19.1)
Requirement already satisfied: click>=7 in
/usr/local/lib/python3.10/site-packages (from cooler->hicrep) (8.1.3)
Requirement already satisfied: asciitree in
/usr/local/lib/python3.10/site-packages (from cooler->hicrep) (0.3.3)
Requirement already satisfied: pyyaml in
/usr/local/lib/python3.10/site-packages (from cooler->hicrep) (6.0)
Requirement already satisfied: pyfaidx in
/usr/local/lib/python3.10/site-packages (from cooler->hicrep)
(0.7.2.1)
Requirement already satisfied: tzdata>=2022.1 in
/usr/local/lib/python3.10/site-packages (from pandas->hicrep) (2023.3)

Requirement already satisfied: python-dateutil<=2.8.2 in
/usr/local/lib/python3.10/site-packages (from pandas->hicrep) (2.8.2)
Requirement already satisfied: pytz<=2020.1 in
/usr/local/lib/python3.10/site-packages (from pandas->hicrep) (2023.3)
Requirement already satisfied: wrapt<2,>=1.10 in
/usr/local/lib/python3.10/site-packages (from Deprecated->hicrep)
(1.15.0)
Requirement already satisfied: six<=1.5 in
/usr/local/lib/python3.10/site-packages (from python-dateutil<=2.8.2->pandas->hicrep) (1.16.0)
Requirement already satisfied: toolz<=0.8.0 in
/usr/local/lib/python3.10/site-packages (from cytoolz->cooler->hicrep)
(0.12.0)
Requirement already satisfied: dill<=0.3.6 in
/usr/local/lib/python3.10/site-packages (from multiprocessing->cooler->hicrep) (0.3.6)
Requirement already satisfied: setuptools in
/usr/local/lib/python3.10/site-packages (from pyfaidx->cooler->hicrep)
(65.6.3)

WARNING: Running pip as the 'root' user can result in broken permissions and conflicting behaviour with the system package manager. It is recommended to use a virtual environment instead:

<https://pip.pypa.io/warnings/venv>

WARNING: Running pip as the 'root' user can result in broken permissions and conflicting behaviour with the system package manager. It is recommended to use a virtual environment instead:

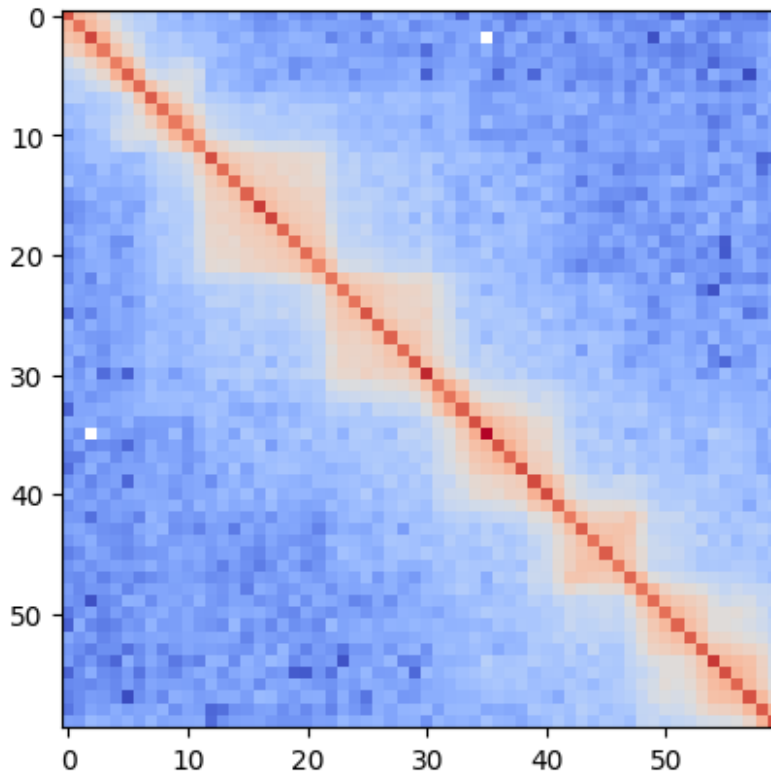
<https://pip.pypa.io/warnings/venv>

```
import matplotlib.pyplot as plt
import cooler
import numpy as np
import seaborn as sns
import pandas as pd
import hicrep
from hicrep import hicrepSCC
from hicrep.utils import readMcool
from scipy.cluster.hierarchy import dendrogram, linkage

resolution = 20000
clr = cooler.Cooler(f'{mcool}::resolutions/{resolution}')
m=clr.matrix(balance=True).fetch('chrX')
pix=clr.pixels(join=True)[: ]
bins=clr.bins()[ : ]
plt.imshow(np.log(m[120:180,120:180]), cmap='coolwarm')

<ipython-input-4-b0c383a86f07>:6: RuntimeWarning: divide by zero
encountered in log
  plt.imshow(np.log(m[120:180,120:180]), cmap='coolwarm')

<matplotlib.image.AxesImage at 0x7fc22b34c130>
```



Итерируемся по матрице диагонально, сохраняем степени и получаем вектор средних

```
z=np.zeros(len(m))
for i in range(len(m)):
    z[i]=np.nanmean(np.diagonal(m,i))
```

<ipython-input-5-e76e9ce83b1f>:5: RuntimeWarning: Mean of empty slice
z[i]=np.nanmean(np.diagonal(m,i))

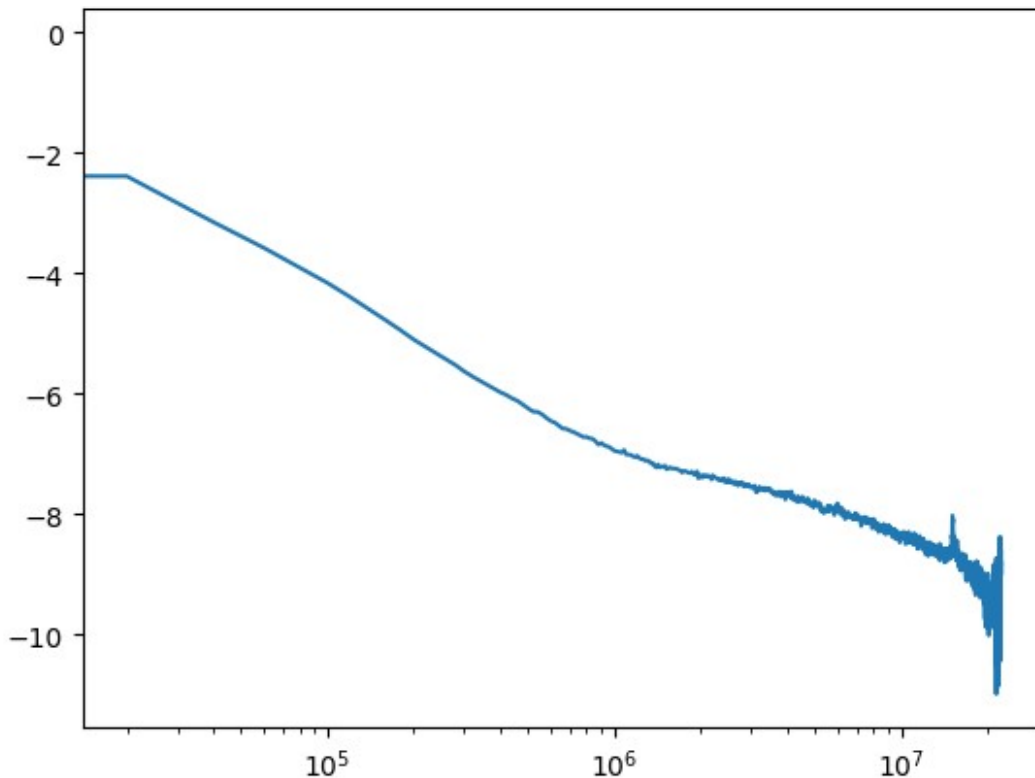
На данном графике вдоль оси абсцисс распределены координаты диагоналей, а вдоль оси ординат логарифмическая шкала по степеням, полученным в прошлой ячейке

!!! why do we paste 20000 below? - потому что resolution = 20000 ;)

```
plt.plot(np.arange(len(m))*20000,np.log(z))
plt.xscale('log')
```

<ipython-input-6-fb803bb700e4>:4: RuntimeWarning: divide by zero encountered in log

```
plt.plot(np.arange(len(m))*20000,np.log(z))
```



```
binSize = 20000 # - Указываем наше разрешение
dBPMaх = 5000000 # - Максимальное допустимое расстояние между
участками генома
bDownSample = True # - Опция, позволяющая включить выравнивание до
одинакового количества контактов
h=0 # - Опция, позволяющая включить фильтр сглаживания
```

```
fmcool1 =
'drive/MyDrive/hse_minor_bioinfo/HiC1.dm3.mapq_30.1000.mcool'
fmcool2 =
'drive/MyDrive/hse_minor_bioinfo/HiC2.dm3.mapq_30.1000.mcool'
cool1, binSize1 = readMcool(fmcool1, binSize)
cool2, binSize2 = readMcool(fmcool2, binSize)
scc=hicrepSCC(cool1, cool2, h, dBPMaх, bDownSample)
```

```
SCC_matrix = np.zeros((4,4))
```

```
for i in range (4):
    for j in range(i,4):
        fmcool1 = f'drive/MyDrive/hse_minor_bioinfo/HiC{i +
1}.dm3.mapq_30.1000.mcool'
        fmcool2 = f'drive/MyDrive/hse_minor_bioinfo/HiC{j +
1}.dm3.mapq_30.1000.mcool'
        cool1, binSize1 = readMcool(fmcool1, binSize)
        cool2, binSize2 = readMcool(fmcool2, binSize)
```

```

scc=hicrepSCC(cool1, cool2, h, dBPMMax, bDownSample)
SCC_matrix[i ,j] = np.nanmean(scc)
SCC_matrix[j ,i] = np.nanmean(scc)

Z=linkage(SCC_matrix, 'single', 'correlation')
plt.figure(figsize=(8,8))
plt.ylabel('distance')

dendrogram(Z, color_threshold=0,labels=[f'Hi{i}' for i in range (1,
5)])
plt.show()

Z=linkage(SCC_matrix, 'single', 'correlation')
plt.figure(figsize=(8,8))
plt.ylabel('distance')

dendrogram(Z, color_threshold=0,labels=[f'Hi{i}' for i in range (1,
5)])
plt.show()

```

