

# How to use DNA\_Repeated\_Sequence \_Detector

Shintaro Miyazaki

---

 RSD(外部)20200505.ipynb

 RSD(内部)20200505.ipynb

 testSequence.txt

①3つのファイルを同じフォルダ内に置きます

## テキストファイルの読み込み

```
sequenceFileName='testSequence.txt'  
f = open(sequenceFileName, 'r')  
sequenceList = f.readlines()  
f.close()
```

②RSD(外部)を開き、この部分に入力テキストファイルの名前を入力します

```
#パラメータ決定
lengthOfRepeat=2000
NF1=0
#NF2=10
L1 = np.round(lengthOfRepeat/20).astype(np.int)
density = (2*NF2+1+1+1)/L1
L1_R = np.round(lengthOfRepeat/4).astype(np.int)
L2 = np.round(lengthOfRepeat/4).astype(np.int)
minLength = np.round(lengthOfRepeat/2).astype(np.int)
Ratio=1.5
N_fourier=10
```

③RSD(内部)を開き、このセルで必要に応じてパラメータを変更します

lengthOfRepeatには見つけたい反復配列のおおよその長さを入力してください

```
100%|██████████| 2/2 [01:03<00:00, 31.64s/it]
```

④RSD(外部)にたいしてjupyter notebookのRun All Cellsを実行します。一番下のセルの下に進行状況が表示されます

## DNA\_Repeated\_Sequence\_Detector

開始時刻 2020-05-05 19:06:52

Sequenceファイルの名前 testSequence.txt

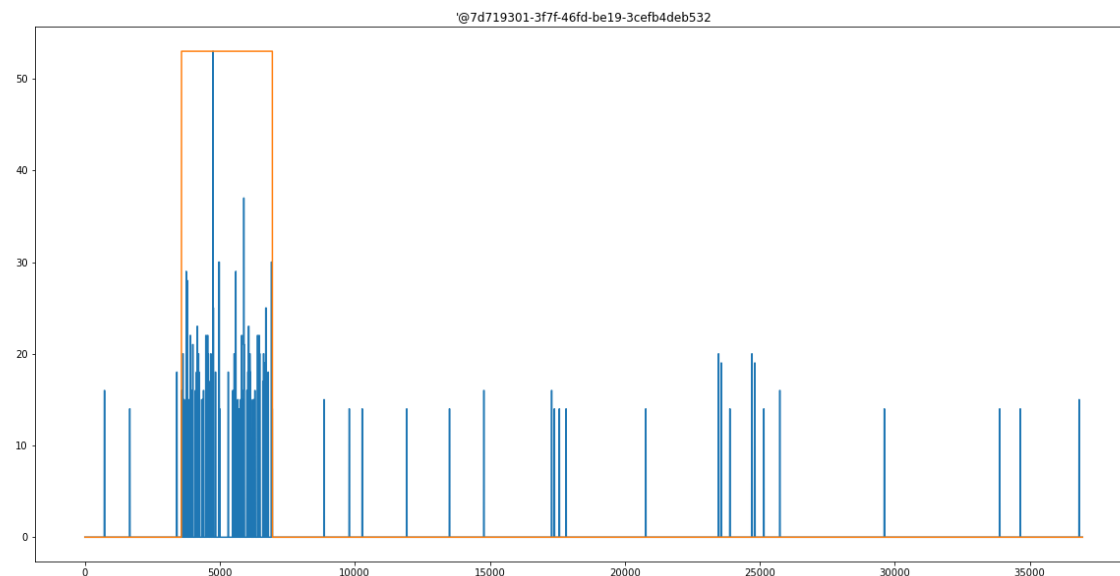
Sequenceファイル中のSequenceの個数 2

元の配列の名前※配列の長さ——※NF2※反復配列開始位置——※反復配列終了位置——※反復配列の個数(推定)※反復配列の個数(フーリエ変換)※第n候補——※反復配列の長さ(フーリエ変換)※領域判定値

'@7d719301-3f7f-46fd-be19-3cefb4deb532—36955—13—3574—6944—2—2—1—1685—2.4834186055841148

'@0d36f7fe-acc8-4496-aae7-8a91b49d6c39—48330—13—33817—39557—3—3—1—1913—2.7546626091724407

終了時刻 2020-05-05 19:07:55



出力はテキストファイル1つと、  
入力テキストファイルに含まれて  
いた配列の個数の図です

出力テキストファイルの分析法についてはDNA\_Sequence\_Detectorと同じです