

Science Programming Assignment

Snehal Kumar

2019101003

Q2.

DNA sequence used: pUC19

Restriction Endonuclease used: BstUI

Overview

pUC19 is a type of plasmid cloning vectors which has 2686 base pairs. It is a popularly used vector molecule in cloning and research and has a multiple cloning site (MCS) which contains restriction sites for many restriction enzymes.

The RE recognizes the sequence **CGCG** in the DNA and cuts in the middle to give blunt ends.

In pUC19, there are a total of 10 restriction sites recognized by the BstUI enzyme. By arranging the sequence in a circular fashion we get 10 fragments with the largest fragment size of 581 when cut at 853rd base.

Program

The program reads the DNA sequence and for each index, checks the group of 4 next bases and compares to the sequence recognized by the RE. If a match is found, it adds the location of the site to a dictionary.

```
ck=RE_DICT[RE]
length = len(ck)
for i in range(len(DNA)):
    if ck[0] == DNA[i]:
        if DNA[i:i+length] == ck:
            re_map["Location"].append((i,i+length))
            re_map["Cut site"].append(i+2)
return re_map
```

To calculate the fragments and their sizes, the cut sites of the DNA are used which are added to the dictionary and finally printed.

```

check = sites["Cut site"]
for i,val in enumerate(check):
    if i == len(check)-1:
        ps.append((val+1,check[0]))
        ll = len(DNA)-check[i]+check[0]
        ls.append(ll)
    else:
        ps.append((val+1,check[i+1]))
        ls.append(check[i+1]-val)

pos = val
sites["Fragment size"] = ls
sites["Fragment"] = ps

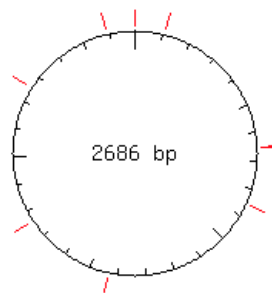
```

Comparison

Results from mapper

Fragment list pUC19 digested with BstUI

[\[Sites w](#)



#	Location	Size [bp]
1	854-1434	581
2	109-653	545
3	1765-2257	493
4	2258-2589	332
5	1435-1764	330
6	656-853	198
7	6-108	103
8	2590-3	100
9	4-5	2
10	654-655	2

Results from program mapper:

```
RE: BstUI
Recognition sequence: CGCG

Restriction sites Map:
  Cut site      Location      Fragment      Fragment size
5      853      (851, 855)      (854, 1434)      581
2      108      (106, 110)      (109, 653)       545
7     1764     (1762, 1766)     (1765, 2257)     493
8     2257     (2255, 2259)     (2258, 2589)     332
6     1434     (1432, 1436)     (1435, 1764)     330
4      655      (653, 657)      (656, 853)       198
1         5         (3, 7)         (6, 108)         103
9     2589     (2587, 2591)     (2590, 3)        100
0         3         (1, 5)         (4, 5)           2
3      653      (651, 655)      (654, 655)       2

Total number of restriction sites: 10
```

The generated restriction map gives the same result as that of the online mapper.

Cut site – the base at which the RE cuts the DNA

Location – the recognized sequence in the DNA

Fragment – the start and end indices of the fragment to the right resulting after the cut

Fragment size - The length of the fragments in bp