

Aluno: Leonardo Oliveira Wellausen

1) Código:

```
def find_motif(seqs, l, m):
    seq_trg = seqs[0]
    ret = [-1 for i in range(len(seqs) + 1)]
    for i in range(len(seq_trg) - l + 1):
        candi1 = seq_trg[i:i+l]
        seqs_ok = 0
        ret = [-1 for i in range(len(seqs) + 1)]
        for k, seq in enumerate(seqs):
            for j in range(len(seq_trg) - l + 1):
                candi2 = seq[j:j + l]
                if mut_cmp(candi1, candi2) <= m:
                    seqs_ok += 1
                    ret[k] = j
                    break
            if seqs_ok == len(seqs):
                ret[k + 1] = candi1
                return ret
    return ret
```

Código principal para encontrar um motivo de tamanho **l** dentro das sequências em **seqs**, permitindo até **m** mutações.

Pegamos todas as subsequências de tamanho **l** dentro da primeira sequência e a procuramos em todas as outras sequências. Se uma subsequência **candi1** é encontrada em todas as sequências (permitindo uma diferença de **m** caracteres) então **candi1** é o motivo que procuramos. O retorno é uma lista com o motivo encontrado e os índices onde ele aparece em cada sequência.

```
def mut_cmp(str1, str2):
    diffs = 0
    for i in range(len(str1)):
        if str1[i] != str2[i]:
            diffs += 1
    return diffs
```

Função para a comparação de subsequências de mesmo tamanho que retorna o número de mutações entre as strings. É utilizada para encontrar os motivos permitindo mutações.

2) Resultados:

2a) Resultados para o primeiro conjunto de teste:

```
Motivo encontrado para tamanho 8 com 2 mutações: acgtacat PARA O CONJUNTO 1
Presente na sequência 1 na posição 27 como acgtacat
Presente na sequência 2 na posição 9 como gtgtacat
Presente na sequência 3 na posição 59 como acgaaaat
Presente na sequência 4 na posição 57 como acgtacgt
Presente na sequência 5 na posição 61 como acgtaggt

Motivo encontrado para tamanho 5 com 3 mutações: cccct PARA O CONJUNTO 1
Presente na sequência 1 na posição 0 como cccct
Presente na sequência 2 na posição 3 como ttact
Presente na sequência 3 na posição 6 como gtccg
Presente na sequência 4 na posição 0 como agccc
Presente na sequência 5 na posição 10 como caacg

Motivo encontrado para tamanho 3 com 1 mutações: ccc PARA O CONJUNTO 1
Presente na sequência 1 na posição 0 como ccc
Presente na sequência 2 na posição 31 como cac
Presente na sequência 3 na posição 7 como tcc
Presente na sequência 4 na posição 1 como gcc
Presente na sequência 5 na posição 13 como cgc
```

2b) Resultados para o segundo conjunto de teste:

```
Motivo encontrado para tamanho 3 com 1 mutações: gtc PARA O CONJUNTO 2
Presente na sequência 1 na posição 0 como gtc
Presente na sequência 2 na posição 0 como atc
Presente na sequência 3 na posição 8 como gtt
Presente na sequência 4 na posição 2 como atc
Presente na sequência 5 na posição 0 como gac
Presente na sequência 6 na posição 3 como atc
Presente na sequência 7 na posição 0 como gcc
Presente na sequência 8 na posição 1 como gtc
Presente na sequência 9 na posição 3 como gtc
Presente na sequência 10 na posição 6 como gta

Motivo encontrado para tamanho 5 com 2 mutações: gtcac PARA O CONJUNTO 2
Presente na sequência 1 na posição 0 como gtcac
Presente na sequência 2 na posição 0 como atccc
Presente na sequência 3 na posição 8 como gttcc
Presente na sequência 4 na posição 9 como gccac
Presente na sequência 5 na posição 3 como ggcac
Presente na sequência 6 na posição 3 como atcat
Presente na sequência 7 na posição 11 como gcctc
Presente na sequência 8 na posição 1 como gtccg
Presente na sequência 9 na posição 3 como gtctc
Presente na sequência 10 na posição 9 como gtgat
```