

waterNeedle

April 5, 2019

Smith-Waterman local alignment, and Needleman-Wunsch global alignment using Biopython

Author: Deepak Khatri

[Online demo\(1\) for local and global alignment](#)

```
In [1]: # Biopython modules for pairwise alignment
        from Bio import pairwise2
        from Bio import SeqIO
        from Bio.SubsMat.MatrixInfo import blosum62
```

1 Smith-Waterman local alignment

[Online Demo\(2\) for Smith-Waterman local alignment](#)

```
In [2]: # alignment with no match score and gap penalty
        local_alignment = pairwise2.align.localxx('AGGTTGCA', 'AGGTCA')
        # formatted output
        for i in range(len(local_alignment)):
            print(pairwise2.format_alignment(*local_alignment[i]))
```

```
AGGTTGCA
||| | ||
AGG-T-CA
Score=6
```

```
AGGTTGCA
||||  ||
AGGT--CA
Score=6
```

```
In [3]: # local alignment with opengap= -10, extendgap= -0.5
        local_alignments = pairwise2.align.localds('AGGTTGCA', 'AGGTCA', blosum62, -10, -0.5)
        # formatted output
        for i in range(len(local_alignments)):
            print(pairwise2.format_alignment(*local_alignments[i]))
```

```

AGGTTGCA
||||  ||
AGGT--CA
    Score=23.5

```

2 Needleman-Wunsch global alignment

[Online Demo\(3\) for Needleman-Wunsch global alignment](#)

```

In [4]: # alignment with no match score and gap penalty
        global_alignmnet = pairwise2.align.globalxx('ATTAC', 'AATTC')
        # formatted alignment
        for i in range(len(global_alignmnet)):
            print(pairwise2.format_alignment(*global_alignmnet[i]))

```

```

A-TTAC
|  |  |
AATT-C
    Score=4

```

```

-ATTAC
|  |  |
AATT-C
    Score=4

```

```

In [5]: # global alignment using blosum62 substitution matrix, opengap= -1, extendgap= -1
        global_alignmenttds = pairwise2.align.globalds('ATTAC', 'AATTC', blosum62, -1, -1)
        # formatted output
        for i in range(len(global_alignmenttds)):
            print(pairwise2.format_alignment(*global_alignmenttds[i]))

```

```

A-TTAC
|  |  |
AATT-C
    Score=21

```

```

-ATTAC
|  |  |
AATT-C
    Score=21

```

```

In [6]: # global alignmnet with match= 1, mismatch= -1, opengap= -1, endgap= -1
        global_alignmenttds = pairwise2.align.globalms('ATTAC', 'AATTC', 1, -1, -1, -1)

```

```
# formatted output
for i in range(len(global_alignments)):
    print(pairwise2.format_alignment(*global_alignments[i]))
```

```
A-TTAC
|  |  |
AATT-C
Score=2
```

```
-ATTAC
|  |  |
AATT-C
Score=2
```

xx in localxx and globalxx can have these values
 (x)x The match score parameters are:

1. x No parameters. Identical characters have score of 1, otherwise 0.
2. m A match score is the score of identical chars, otherwise mismatch score.
3. d A dictionary returns the score of any pair of characters.
4. c A callback function returns scores.

x(x) The gap penalty parameters are: 1. x No gap penalties. 2. s Same open and extend gap penalties for both sequences. 3. d The sequences have different open and extend gap penalties. 4. c A callback function returns the gap penalties.

3 References:

1. [wikipedia waterman algorithm](#)
2. [wikipedia Wunsch algorithm](#)
3. [Biopython pairwise2 module](#)
4. [Biopython website](#)