

What is sequence alignment?

Sequence alignment attempts to match two sequences based on their similarity. There are two approaches: (A) local alignment, and (B) global alignment.

(A) Local alignment

Local alignment matches similar regions in both strings and tries to make each similar region as long as possible. In general, we use local alignment to find similar regions in two dissimilar strings.

(B) Global alignment

Global alignment algorithms attempt to find the global optimal match of two strings. Typically, we use global alignment to evaluate the similarity of two strings.

For example, given two strings as shown below

String 1: XXOOTATAYYTATACGCGY00CG00XX
String 2: XOTATACGCGXX

(^ : indicates two characters are matched)

(a) Local alignment

```

XXOOTATAYYTATACGCGY00CG00XX
-XO-----TATACGCG-----XX
=====
  ^^          ^^^^^^^^          ^^
    
```

NOTE:

1. (a) is a local alignment result since it attempts to find the best continuous matches locally. The "XO", "TATACGCG", and "XX" are three matched regions in both strings.
2. This is also one of the possible global alignment results in this case.

(b) Global alignment

```

XXOOTATAYYTATACGCGY00CG00XX
-XO-TA----TA--CG-----CG--XX
=====
  ^^ ^^      ^^  ^^      ^^  ^^
    
```

(c) Global alignment

```

XXOOTATAYYTATACGCGY00CG00XX
X--O--TA----TA--CG---CG--XX
=====
  ^  ^  ^^      ^^  ^^      ^^  ^^
    
```

NOTE:

1. (a), (b), and (c) are all possible global alignment results.
2. All of them are equally good since they all have 12 matches.