**Global Alignment Example:**

Sequence 1: GATCCAG

Sequence 2: GATCAC

**Scoring Scheme:**

**S(i,j) ->** Match= +1

Mismatch=-1

**W** -> Gap-penality= -1

**M(i,j)** -> Score of the cell

**Procedure**:

* Draw a matrix and fill the first row(j=0) and column(i=1) starting with zero and incrementing with gap scores moving further.
* The second step of the algorithm is matrix filling starting from the upper left hand corner of the matrix. To find the maximum score of each cell, it is required to know the neighbouring scores (diagonal, left and right) of the current position.
* From the assumed values, add the match or mismatch score to the diagonal value.
* Similarly add the gap score to the other neighbouring values.
* In terms of matrix positions, it is important to know
  + M(i-1,j-1)+S(i,j),
  + M(i,j-1)+w
  + M(i-1,j)+w
* Thus, we can obtain three different values, from that take the maximum among them and fill the ith and jth position with the score obtained.
* Replace all the negative values with zero.
* Traceback is the final step where you start from the cell with the cell in the matix and follow the arrows backwards to the first cell of the matrix.
* The diagonal movement indicates a match or a mismatch in the alignment, whereas the upward or sideward movement indicated a gap in the alignment.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | G | A | T | C | C | A | G |
|  | 0 | -1 | -2 | -3 | -4 | -5 | -6 | -7 |
| G | -1 | 1 | 0 | -1 | -2 | -3 | -4 | -5 |
| A | -2 | 0 | 2 | 1 | 0 | -1 | -2 | -3 |
| T | -3 | -1 | 1 | 3 | 2 | 1 | 0 | -1 |
| C | -4 | -2 | 0 | 2 | 4 | 3 | 2 | 1 |
| A | -5 | -3 | -1 | 1 | 3 | 3 | 4 | 3 |
| C | -6 | -4 | -2 | 0 | 2 | 2 | 3 | 3 |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | G | A | T | C | C | A | G |
|  | 0 | -1 | -2 | -3 | -4 | -5 | -6 | -7 |
| G | -1 | 1 | 0 | -1 | -2 | -3 | -4 | -5 |
| A | -2 | 0 | 2 | 1 | 0 | -1 | -2 | -3 |
| T | -3 | -1 | 1 | 3 | 2 | 1 | 0 | -1 |
| C | -4 | -2 | 0 | 2 | 4 | 3 | 2 | 1 |
| A | -5 | -3 | -1 | 1 | 3 | 3 | 4 | 3 |
| C | -6 | -4 | -2 | 0 | 2 | 2 | 3 | 3 |

Pairwise alignment:

GATCCAG

|||| ||

GATC- AG

OR

GATCCAG

||| |||

GAT- CAG

**Local Alignment Example:**

**Scoring Scheme:**

**S(i,j) ->** Match= +1

Mismatch=-1

**W** -> Gap-penality= -1

**M(i,j)** -> Score of the cell

**Procedure**:

1. Draw a matrix and fill the first row(j=0) and column(i=1) with zeros.
2. The second step of the algorithm is matrix filling starting from the upper left hand corner of the matrix. To find the maximum score of each cell, it is required to know the neighbouring scores (diagonal, left and right) of the current position.
3. From the assumed values, add the match or mismatch score to the diagonal value.
4. Similarly add the gap score to the other neighbouring values.
5. In terms of matrix positions, it is important to know
   * M(i-1,j-1)+S(i,j),
   * M(i,j-1)+w
   * M(i-1,j)+w
6. Thus, we can obtain three different values, from that take the maximum among them and fill the ith and jth position with the score obtained.
7. Replace all the negative values with zero.
8. Traceback is the final step where you start from the cell with the highest score and follow the arrows backwards and end at zero.
9. The diagonal movement indicates a match or a mismatch in the alignment, whereas the upward or sideward movement indicated a gap in the alignment.

**Alignment 1: Basic Local Alignment**

Sequence 1: ACTAGACT

Sequence 2: AGAC

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | ***i*** | **A** | **C** | **T** | **A** | **G** | **A** | **C** | **T** |
| ***j*** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **A** | 0 | 1 | 0 | -1 | 1 | 0 | 1 | 0 | -1 |
| **G** | 0 | 0 | 0 | -1 | 0 | 2 | 1 | 0 | -1 |
| **A** | 0 | 1 | 0 | -1 | 0 | 1 | 3 | 2 | 1 |
| **C** | 0 | 0 | 2 | 1 | 0 | 0 | 2 | **4** | 3 |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | A | C | T | A | G | A | C | T |
|  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| A | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 0 |
| G | 0 | 0 | 0 | 0 | 0 | 2 | 1 | 0 | 0 |
| A | 0 | 1 | 0 | 0 | 0 | 1 | 3 | 2 | 1 |
| C | 0 | 0 | 2 | 1 | 0 | 0 | 2 | **4** | 3 |

Pairwise alignment : ACTAGACT

|| ||

\_\_\_AGAC\_

**Alignment 2: Overlap Alignment**

**Sequence 1:** ACGTAAGGTC

**Sequence 2:** GTCACG

An overlap alignment between two strings and is a local alignment of a suffix of a string/sequence with a prefix of the other string/sequence.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **j** | **A** | **C** | **G** | **T** | **A** | **A** | **G** | **G** | **T** | **C** |
| **i** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **G** | 0 | -1 | -1 | 1 | 0 | -1 | -1 | 1 | 1 | 0 | -1 |
| **T** | 0 | -1 | -2 | 0 | 2 | 1 | 0 | 0 | 0 | 2 | 1 |
| **C** | 0 | -1 | -2 | -1 | 1 | 1 | 0 | -1 | -1 | 1 | 3 |
| **A** | 0 | 1 | 0 | -1 | 0 | 2 | 2 | 1 | 0 | 0 | 2 |
| **C** | 0 | 0 | 2 | 1 | 0 | 1 | 1 | 1 | 0 | -1 | 1 |
| **G** | 0 | -1 | 1 | 3 | 2 | 1 | 0 | 2 | 2 | 1 | 0 |

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **j** | **A** | **C** | **G** | **T** | **A** | **A** | **G** | **G** | **T** | **C** |
| **i** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **G** | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 |
| **T** | 0 | 0 | 0 | 0 | 2 | 1 | 0 | 0 | 0 | 2 | 1 |
| **C** | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 3 |
| **A** | 0 | 1 | 0 | 0 | 0 | 2 | 2 | 1 | 0 | 0 | 2 |
| **C** | 0 | 0 | 2 | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 1 |
| **G** | 0 | 0 | 1 | 3 | 2 | 1 | 0 | 2 | 2 | 1 | 0 |

Pairwise alignment: ACGTAAGGTC

|||

GTCACG

OR

ACGTAAGGTC

|||

GTCACG

**Alignment 3: Suboptimal Alignment**

**Sequence 1:** CAGTCGAGTGC

**Sequence 2:** AGTC

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **j** | **C** | **A** | **G** | **T** | **C** | **G** | **A** | **G** | **T** | **G** | **C** |
| **i** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **A** | 0 | -1 | 1 | 0 | -1 | -1 | -1 | 1 | 0 | -1 | -1 | -1 |
| **G** | 0 | -1 | 0 | 2 | 1 | 0 | 0 | 0 | 2 | 1 | 0 | -1 |
| **T** | 0 | -1 | -1 | 1 | 3 | 2 | 1 | 0 | 1 | 3 | 2 | 1 |
| **C** | 0 | 1 | 0 | 0 | 2 | 4 | 3 | 2 | 1 | 2 | 2 | 3 |

Suboptimal Alignmentis an extension of local alignment that finds multiple similar regions in two given genetic strings; for instance, if we want to find all occurrences of a certain repeat, perhaps one corresponding to a transposon.

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **j** | **C** | **A** | **G** | **T** | **C** | **G** | **A** | **G** | **T** | **G** | **C** |
| **i** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **A** | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| **G** | 0 | 0 | 0 | 2 | 1 | 0 | -1 | 0 | 2 | 1 | 0 | 0 |
| **T** | 0 | 0 | 0 | 1 | 3 | 2 | 1 | 0 | 1 | 3 | 2 | 1 |
| **C** | 0 | 1 | 0 | 0 | 2 | 4 | 3 | 2 | 1 | 2 | 2 | 3 |

Pairwise alignment:

CAGTCGAGTGC

|| ||

\_AGTC\_\_\_\_\_\_\_

OR

CAGTCGAGTGC

||| |

\_\_\_\_\_\_\_AGT- C

**Alignment 4: Tandem repeat Alignment**

Tandem repeats occur in DNA when a pattern of one or more nucleotides is repeated and the repetitions are directly adjacent to each other.

Sequence 1: ACGTCGCGCGACT

Sequence 2: CG

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **j** | **A** | **C** | **G** | **T** | **C** | **G** | **C** | **G** | **C** | **G** | **A** | **C** | **T** |
| **i** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **C** | 0 | -1 | 1 | 0 | -1 | 1 | 0 | 1 | 0 | 1 | 0 | -1 | 1 | 0 |
| **G** | 0 | -1 | 0 | 2 | 1 | 0 | 2 | 1 | 2 | 1 | 2 | 1 | 0 | 0 |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **j** | **A** | **C** | **G** | **T** | **C** | **G** | **C** | **G** | **C** | **G** | **A** | **C** | **T** |
| **i** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **C** | 0 | -1 | 1 | 0 | -1 | 1 | 0 | 1 | 0 | 1 | 0 | -1 | 1 | 0 |
| **G** | 0 | -1 | 0 | 2 | 1 | 0 | 2 | 1 | 2 | 1 | 2 | 1 | 0 | 0 |

Pairwise alignment:

ACGTCGCGCGACT

|| || ||

\_\_\_\_CGCGCG\_\_\_