**4. Compararea secventelor**

 Aliniamentele globale și locale sunt două metode diferite de aliniere a unei secvențe. Pentru a decide ce metodă se alege depinde de scopul alinierii. Aliniamentele globale încearca sa compare fiecare reziduu din fiecare secvență și sunt utilizate cel mai bine atunci când secvențele sunt similare și sunt de aceeași dimensiune, pentru că secvențele care au dimensiuni diferite produc nepotriviri la capetele unui aliniament. Cu toate acestea, atunci când se încearcă să se alinieze fiecare element din secvențe diferite, se vor produce multe gaps-uri(lacune) din cauza nepotrivirilor dintre cele două secvențe, așa cum se vede în figura urmatoare ce reprezinta un exemplu de aliniere globala si locala:



 La compararea a două secvențe lungi, aceste goluri(gaps) pot deveni greu de analizat. Aliniamentele locale sunt cel mai utilizate pentru secvențe diferite care pot avea regiuni similare. Aliniamente locale sunt foarte utile pentru identificarea unui anumit model care există pe ambele secvențe, deoarececa acel model poate avea o funcție similară. Dacă ambele secvențe sunt foarte similare, nu ar trebui sa se faca vreo diferență indiferent de metoda folosita, deoarece alinierile ar produce rezultate similare.

Algoritmii fundamentali de aliniere globale și locale se bazează pe programarea dinamică. Algoritmul Needleman – Wunsch este bazat pe programare dinamică și rezolvă problema de aliniere globală, în timp ce algoritmul Smith – Waterman este de asemenea bazat pe programare dinamică și rezolvă problema de aliniere locală.

Aliniere globală (Needleman – Wunsch) se întinde de la un capăt la celalalt al fiecărei secvențe. Alinierea locală găseste potrivirea optima a regiunilor dintre două secvențe ("subsecvente"). Aliniere locală este aproape întotdeauna folosită pentru căutări de baze de date, cum ar fi BLAST. Este util să se găsească domenii (sau regiuni limitate de omologie) în termen de secvențe. Smith și Waterman (1981) au rezolvat problema efectuării alinierii locale optime a unei secvențe. Alte metode (BLAST, FASTA), sunt mai rapide, dar mai puțin aprofundate.

**Diferente intre algoritmii Needleman – Wunsch si Smith – Waterman**

 Algoritmul Needleman – Wunsch :

1. Este un algoritm de aliniere globala.
2. Este necesar ca scorul de aliniere pentru o pereche de reziduuri sa fie mai mare sau ega cu zero.
3. Nu este necesara penalizarea gap-urilor.
4. Scorul dintre doua celule de pe acelai rand(de pe aceeasi cale) nu poate scadea .

Algoritmul Smith – Waterman:

1. Este un algoritm de aliniere locala.
2. Scorul de aliniere al reziduurilor poate fi pozitiv sau negativ.
3. Pentru o munca eficienta este necesara penalizarea gap-urilor.
4. Scorul dintre doua celule de pe acelai rand(de pe aceeasi cale) poate sa creasca, sa descreasca sau sa ramana acelasi.

Algoritmii de aliniere bazati pe tehnica programarii dinamice (Needleman-Wunsch si Smith-Waterman) genereaza alinieri optime insa sunt costisitori din punct de vedere computational: pentru doua secvente de lungimi m respectiv n ordinul de complexitate al implementarilor traditionale este O(mn) atat din punct de vedere al timpului de calcul cat si din punctul de vedere al spatiului de memorie necesar.

Repere in alinierea secventelor:

* 1970 – Needleman si Wunsch gasesc alinierea globala optima prin programarea dinamica
* 1981 – Smith si Waterman optimizeaza alinierea locala prin prograrea dinamica
* 1985 – Lipman si Pearson pun in aplicare alinierea in format standard in FASTA
* 1990 – ALTCHUL si colaboratorii creaza algoritmul BLAST, accelerand foarte mult potrivirea secventelor.

Putem gasi secvente la urmatoarele adrese:

 • ExPASy Proteomics (UniProt): <http://ca.expasy.org/>

 • Nat’l Center for Biotechnology Information: <http://www.ncbi.nlm.nih.gov/>

 • European Bioinformatics Institute: <http://www.ebi.ac.uk/>

 • DNA Databank of Japan: <http://www.ddbj.nig.ac.jp/>

Algoritmul Needleman – Wunsch

* Gaseste alinierea globala optima intre secvente pentru un anumit set de reguli
* Forma simpla: pentru potrivirea rezidurilor avem scorul 1, iar pentru nepotrivire avem scorul 0
* Prin programarea dinamica, mai multe ainieri posibile pot fi eliminate

Mai multe alinieri posibile:

Fie secventa1: ASLVNDK si secventa 2: ALVNKDK

Alinierile posibile sunt: ASLVNDK – – ASLVN–DK A– – – – SLVNDK

 A–LVN –K D K A–LVNKDK ALVNK – – – –DK

Secventele sunt de obicei formate din sute de reziduuri, si gradul de potrivire poate fi destul de variabil.

Configurarea Algoritmului Needleman – Wunsch

Pasul 1. Pornind din coltul din dreapta jos, schimbam spatiile in zerouri si cele galbene cu

 unu.

Pasul 2. Pentru fiecare celula, gasim cea mai mare valoare din dreapta si de jos.

Pasul 3. In caz de potrivire, adaugam 1.

Pasul 4. Repetam pasii 2 si 3 pana cand sunt umplute toate celulele.

Pasul 5. Ne intoarcem pe cele mai mari numere, in jos si la dreapta.

Pasul 6. Interpretarea caii



Pasul 7. Scriem alinierea optima.

Secventa orizontala: ASLVN–DK

 Secventa verticala: A–LVNKDK

* Potrivire, V gap, potrivire, potrivire, potrivire, H gap, potrivire, potrivire
* Uneori, exista legaturi

Concluzii:

1. Compararea secventelor este utila pentru a descoperi relatiile dintre secvente.
2. Programarea dinamica imbunatateste eficienta alinierii secventelor.
3. Alinierea locala evidentiaza conservarea domeniilor, dar alinierea globala cauta numarul total din cele mai bune potriviri.