

Ćwiczenie 5 – Metody przyrównywania sekwencji

W tym ćwiczeniu poznasz:

- metodę programowania dynamicznego, która pozwala uzyskać najlepsze (z punktu widzenia matematycznego) dopasowanie między dwiema sekwencjami.

Słowa kluczowe: programowanie dynamiczne, dopasowanie lokalne i globalne, algorytm Smith-Watermana, algorytm Needleman-Wunsch, matryce substytucji, kara za przerwę, kara za otwarcie przerwy

Zadanie 1.

Przyrównaj sekwencje GATACTA oraz GATTACCA przy pomocy algorytmu programowania dynamicznego Needlemana-Wunscha stosując następujący schemat punktowania:

- Dopasowanie = 1;
- Niedopasowanie = -1;
- Przerwa = -2.

Pytanie 1a. Jaka jest wartość punktacji tego przyrównania?

Pytanie 1b. Ile przyrównań otrzymał(e/a)ś? Wypisz je.

Pytanie 1c. Czy w przyrównaniu występują mutacje? Wypisz je.

Zadanie 2.

Przy pomocy programu needle na stronie narzędzi EMBOSS

(http://www.ebi.ac.uk/Tools/psa/emboss_needle/nucleotide.html) wykonaj przyrównanie pary sekwencji nukleotydowych z pliku needle.fasta:

- (A) s15:s16
- (B) s19:s20
- (C) s17:s18

Pytanie 2. Jakie zmiany sekwencji występują w dopasowaniu między sekwencjami?

Zadanie 3.

Otwórz matrycę BLOSUM62 (<http://www.ncbi.nlm.nih.gov/Class/FieldGuide/BLOSUM62.txt>)

Pytanie 3a. Co oznacza indeks 62 przy nazwie matrycy BLOSUM?

Pytanie 3b. Co oznaczają dodatnie, zerowe i ujemne wartości w tych macierzach?

Pytanie 3c. Jaka jest wartość zamiany asparagianu na glutaminian?

Zadanie 4.

W pliku peptidases.fasta znajdują się sekwencje peptydazy pochodzące z *Bacillus lentus* i człowieka. Bakteryjne białko wykorzystywane jest powszechnie w proszkach czyszczących i służy do usuwania zabrudzeń

typu białkowego w zmywarkach i podczas prania tkanin. Oba białka wywodzą się z sekwencji genu wspólnego

przodka. Otwórz stronę narzędzi EMBOSS (<http://www.ebi.ac.uk/Tools/psa/>) do przyrównywania sekwencji i

odpowiedz na poniższe pytania.

Pytanie 4a. Przyrównaj globalnie (Global alignment -> Needle) i lokalnie (Local alignment -> Water) ich sekwencje. Dlaczego długości tych przyrównań różnią się?

Pytanie 4b. Przyrównaj lokalnie te sekwencje przy użyciu matrycy BLOSUM62 zmniejszając kary za stosowanie przerw: otwarcie przerwy (gap open) = 1 i wydłużenie przerwy (gap extend) = 0.001. W jaki sposób zmniejszenie kar za stosowanie przerw wpłynęło na to dopasowanie?

Pytanie 4c. Wykonaj to samo przyrównanie jak w Pytaniu 4b zwiększając kary za stosowanie przerw: otwarcie przerwy = 50, wydłużenie przerwy = 5. W jaki sposób zwiększenie tych kar wpłynęło na dopasowanie?