

ZADANIE 1: Dopasowanie par sekwencji – algorytm kropkowy

Termin oddania: 16.03.2020

Celem zadania 1 jest napisanie działającego programu do wizualnego porównywania pary sekwencji DNA, RNA lub białek, oraz wykonanie z jego pomocą przykładowej analizy. Program powinien nie tylko dawać poprawne wyniki, ale także być zwięzły, czytelny i zgodny z zasadami stylu języka. Odrębne zadania powinny być wykonywane przez odrębne funkcje.

Program ma spełniać następujące funkcje:

1. Wczytanie sekwencji – możliwe przynajmniej na dwa różne sposoby:
 - ręczne,
 - z pliku w formacie FASTA,
 - zdalne (z bazy danych NCBI),oraz ich sparsowanie i zapis w formie tablicy struktur.
2. Wygenerowanie macierzy kropkowej (ang. dot plot)
 - w postaci tablicy 2-wymiarowej;
3. Filtrowanie macierzy kropkowej dla zadanych
 - rozmiaru okna
 - progu filtra
4. Wyświetlenie filtrowanej macierzy kropkowej w postaci graficznej
 - z opisem macierzy (tytuły wykresu i osi, skala, legenda);
5. Zapis otrzymanej macierzy do pliku graficznego;
6. Opracowanie interfejsu użytkownika
 - z poziomu linii komend (wiersza poleceń), z zestawem testów (przykładów użycia) albo
 - graficznego.

Sprawozdanie ma zawierać:

1. Odnosnik do repozytorium z kodem źródłowym;
2. Analizę złożoności obliczeniowej czasowej i pamięciowej;
3. Porównanie przykładowych par sekwencji ewolucyjnie
 - powiązanych,
 - niepowiązanych;
4. Interpretację macierzy kropkowej dla przykładowych par (określenie rodzaju mutacji).

Witold Dyrka
(w oparciu o zadanie prof. Małgorzaty Kotulskiej)