

Zadanie 1.

Dane są dwie sekwencje aminokwasowe:

"PEPAKSAPAPKKGSKKAVTKTQKKDGGKRRKTRKESYAIYVYKVLKQVHPDTGISSKAMSIMNSFVNDV
FERIAGEASRLAHYNKRSTITSREIQTAVRLLLPGELAKHAVSEGTKAVTKYTSK" oraz
"PEPAKSAPAPKKGSKKAVTKTQKKGDGKRRKRSRKESYSIYVYKVLKQVHPDTGISSKAMGIMNSFVNDIF
ERIAGEASRLAHYNKRSTITSREIQTAVRLLLPGELAKHAVSEGTKAVTKYTSSK"

Obliczyć długości podanych sekwencji i podać ile razy pojawiają się w niej poszczególne aminokwasy.

Znaleźć optymalne dopasowanie tych sekwencji. Podać pozycje insercji, delecji i substytucji oraz najdłuższego wspólnego fragmentu. W jaki sposób zmieni się dopasowanie jeżeli użyjemy metody globalnej z macierzą podobieństwa PAM30?

Zadanie 2.

Pobrać z Protein Data Bank sekwencje białkowe (pierwsze na liście składników kompleksu) zawarte w rekordach o identyfikatorach: 3P9P, 1F4J, 1TGU, 3N3N, 1GWE. Utworzyć macierz 5x5 której elementami będzie odległość **EditDistance** pomiędzy podanymi sekwencjami. Na podstawie zawartości macierzy i rekordów PDB podać nazwy dwóch najbardziej „spokrewnionych” organizmów z listy.

Zadanie 3.

W katalogu <http://ribes.if.uj.edu.pl/bioinfo/kolokwium/> każdy znajdzie plik tekstowy zawierający sekwencję DNA opisany imieniem i nazwiskiem. Proszę utworzyć następujące pliki, na podstawie zawartości podanego pliku:

- (a) Plik nr. 2 zawierający w każdej linii po 3 kolejne nukleotydy
- (b) Zamienić wszędzie T na U w pliku 2 i zapisać do pliku nr. 3

- (c) Każdą trójkę nukleotydów (kodon) zamienić na trójliterowe oznaczenie aminokwasu kodowanego przez nią. Wynik zapisać do pliku nr 4.
- (d) W pliku nr. 4 zamienić oznaczenia trójliterowe na standardowe jednoliterowe i zapisać w pliku nr. 5
- (e) Usunąć znaki końca wiersza z pliku nr 5, tak aby powstał nieprzerwany ciąg znaków (jeden wiersz). Zapisać wynikową sekwencję aminokwasową do pliku nr. 6

Sześć plików skonstruowanych wg. powyższej procedury (wliczając treść) stanowi rozwiązanie zadania.

UWAGA: końcowe odpowiedzi do zadań 1 i 2 należy udzielić na podpisanych kartkach. Dodatkowo, pliki stanowiące rozwiązanie Zad. 3 oraz pliki zawierające użyte polecenia/programy należy wysłać (nie później niż do momentu zakończenia kolokwium) e-mailowo na adres: andrzej.odrzywolek@uj.edu.pl

W trakcie kolokwium można używać **własnych** programów napisanych w celu rozwiązania zadań przerabianych na pracowni. Korzystanie z notatek, plików pomocy oraz internetu jest dozwolone, *za wyjątkiem bezpośredniej komunikacji* interpersonalnej (e-mail, komunikatory, serwisy WWW itp.).