

Zadanie 1.

Dane są dwie sekwencje DNA:

s1=TATGGCAGGTTAGACGGAGG

s2=CTATGGCAGCAGTAGCTGTTAGACGGACG

Znaleźć optymalne dopasowanie podanych sekwencji korzystając z **SequenceAlignment**.

Zinterpretować uzyskany wynik i przedstawić go w postaci dwóch ciągów identycznej długości jeden nad drugim. Przerwy wypełnić odpowiednią ilością elementów neutralnych, zgodnie z wynikiem dopasowania.

Zadanie 2.

Wypisać jedno nad drugim dopasowanie sekwencji białek NAT1 i NAT2. Podać ilość delecji/insercji, pozycji identycznych oraz różnych. Powtórzyć używając macierzy podobieństwa PAM250.

Zadanie 3.

Napisać funkcję konwertującą format wyjściowy polecenia **SequenceAlignment** do postaci dwóch sekwencji identycznej długości poprzez wprowadzenie elementów neutralnych.

Przetestować funkcję na przynajmniej kilkunastu przykładach par sekwencji DNA oraz aminokwasowych różnego stopnia podobieństwa i długości, zarówno generowanych losowo jak i pobranych z baz danych **GenomeData** i **ProteinData**.

*Wskazówka: użycie opcji **MergeDifferences**→**False** produkuje bardziej regularny output, co może zmniejszyć nieznacznie ilość koniecznych do rozwiązania przypadków.*

PRZYKŁAD DZIAŁANIA:

s1=TATCCAAGGCTCTCTATCATC

s2=TCCACGGCTCTCTATCATC

SequenceAlignment[s1,s2]

{{TA, }, TCCA, {A, C }, GGCTCTCTATCATC}

Oczekiwany rezultat konwersji:

TATCCAAGGCTCTCTATCATC

--TCCACGGCTCTCTATCATC