

Zadanie 1.

Wygeneruj losowo lub pobierz bazy danych dowolną sekwencję DNA o długości pomiędzy 10 a 20. Zapisz ją w dwóch kopiach.

- (a) oblicz „podobieństwo” sekwencji do samej siebie za pomocą **SmithWatermanSimilarity**, **NeedlemanWunschSimilarity** i **EditDistance**
- (b) w jednej z nich zamień jeden z nukleotydów na inny i oblicz odległości pomiędzy sekwencją oryginalną a zmienioną
- (c) jak wyżej, ale dokonaj większej ilości zmian
- (d) wstaw w dowolnym miejscu oryginalnego DNA losowo kilka-kilkanascie nukleotudow i oblicz odleglosci
- (e) jak wyżej, ale tym razem skasuj fragment oryginalnej sekwencji
- (f) powtórz dwie powyższe czynności zmieniając opcję **GapPenalty**
- (g) jak zmienia się odległość pomiędzy ciągiem oryginalnym i zmienionym w przypadku gdy liczba modyfikacji dąży do ∞ ?

Zadanie 2.

Wykonaj powyższe czynności dla wybranej sekwencji aminokwasowej.

Zadanie 3.

Dla czwórki genów: HOXB1, HOXC12, HOXD4, HOXD12, odpowiedzieć na następujące pytania:

- (a) jak jest długość DNA każdego z nich
- (b) jaka jest odległość pomiędzy parami tych genów wyznaczona przez **SmithWatermanSimilarity**, **NeedlemanWunschSimilarity**, **EditDistance**
- (c) które z 4 genów są do siebie najbardziej podobne wg. ww kryteriów

Zadanie 4.

Powtórz zadanie 4 dla sekwencji białkowych kodowanych przez ww. geny.

Zadanie 5.

Napisać funkcję replikującą sekwencje DNA z błędami. Jako parametry powinna przyjmować prawdopodobieństwo wystąpienia mutacji punktowych, delecji i insercji.