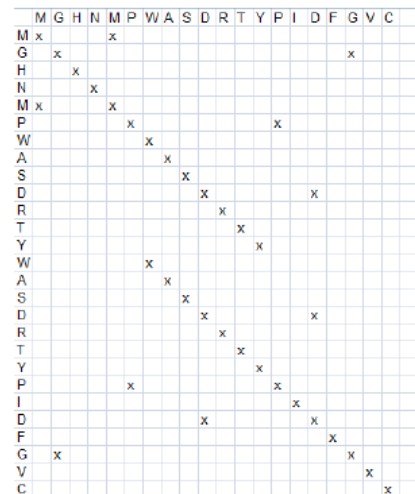
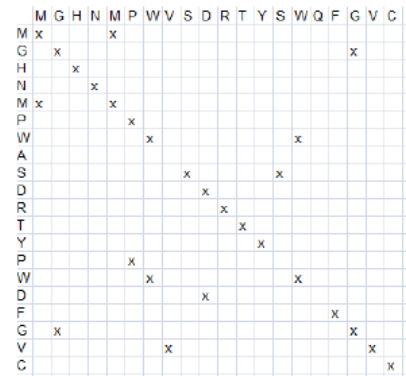
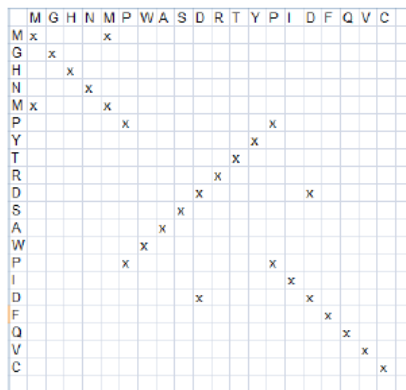


PODSTAWY BIOINFORMATYKI LISTA ZADAŃ 4

DOPASOWANIE SEKWENCJI

- Przyrównano sekwencje aminokwasowe za pomocą macierzy kropkowych. Co możesz wywnioskować z podanych poniżej wizualizacji? Jeśli stwierdzisz mutację przyrównaj sekwencje (jedna pod drugą) i wskaż (np. zaznacz kolorem) zmutowane fragmenty.



- Na podstawie macierzy substytucji aminokwasów **BLOSUM62** oblicz score poniższego przyrównania. Przyjmij, że kara za otwarcie przerwy wynosi -10, a za jej rozszerzenie -1.

C	M	D	A	E	G	F	T	Y	H	N	S	D	E	W	Q	A	A	G	S	G	H	K	L
C	M	Q	A	E	G	F	-	-	-	-	T	E	E	W	N	-	-	-	-	A	H	H	L

- W bazie UniProt odnajdź sekwencje białka FOXP2 dla *Homo sapiens*, *Bos taurus* oraz *Rattus norvegicus*. Sekwencje przyrównaj w programie Clustal. Jaki jest procent identyczności sekwencji pomiędzy
 - Homo sapiens* a *Bos taurus*?
 - Rattus norvegicus* a *Bos taurus*?
 - Rattus norvegicus* a *Homo sapiens*?

W którym przypadku procent identyczności jest najwyższy?

4. Kluczową rolą białka SHH (Sonic hedgehog protein) jest regulacja organogenezy u kręgowców. W bazie danych Uniprot odnajdź to białko dla człowieka oraz dla myszy domowej. Sekwencje w formacie fasta wprowadź do programu dotlet (<https://dotlet.vital-it.ch/>). Czy na podstawie otrzymanej wizualizacji jesteś w stanie stwierdzić, że sekwencje są podobne?
5. Rozważmy dwie sekwencje DNA: TAGCAT i TGACAT. Za pomocą programowania dynamicznego wyznaczono dla obu sekwencji dopasowanie globalne. Ustalono następujący system punktacji: dopasowanie=1, niedopasowanie=0, kara za przerwę o jednostkowej długości=-0,1 pkt. Rozpisz możliwe sposoby dopasowania sekwencji i odpowiedz na poniższe pytania.
 - a. Czy w **optymalnym*** dopasowaniu nie ma przerw?
 - b. Jaka jest punktowa ocena optymalnego dopasowania?
 - c. Czy istnieje kilka alternatywnych, optymalnych dopasowań? Jak wiele ich jest?

*optymalne przyrównanie ma najwyższy score.