

MBI, 24.06.2023, czas: 60 min

Imię i nazwisko:

Proszę wpisywać odpowiedzi w ramkach

Zadanie 1 (3 pkt)

Badamy podobieństwo globalne algorytmem Needlemana-Wunscha dla sekwencji ABABB i ABBA. Stosujemy liniową karę za przerwę $d = -2$. Macierz podobieństwa pokazano obok, n jest liczbą liter Twojego nazwiska podzieloną na 3, dzielenie całkowite, bez reszty.

	A	B
A	$3 * n$	$-n$
B	$-n$	$2 * n$

Uzupełnij macierz podobieństwa, macierz pomocniczą. Podaj najlepsze dopasowanie, jeżeli jest kilka, podaj jedno z nich.

	A	B
A		
B		

		A	B	A	B	B
A	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
B	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
B	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
A	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>

Rozwiązanie:

Zadanie 2 (3 pkt)

Sekwencje ABBAABBA, BABBBBB, BAABAB użyto do asemblacji pod-grafem de Bruijna (bez wag w krawędziach) 4 rzędu (w wierzchołkach są sekwencje o długości 3 symboli).

Narysuj graf.

Odpowiedz na pytanie dotyczące wyniku asemblacji, wpisując a, b albo c:

- a) wynik nie istnieje,
- b) istnieje dokładnie jedna sekwencja wyjściowa,
- c) istnieje więcej niż jedna sekwencja wyjściowa.

Jeżeli istnieje sekwencja wyjściowa (jedna lub kilka) jest nią:

Zadanie 3 (3 pkt)

Zbuduj drzewo filogenetyczne metodą UPGMA (średnich połączeń) dla 5 taksonów. Macierz odległości podano obok.

	B	C	D	E
A	4	8	8	12
B		16	4	12
C			20	8
D				8

Zadanie 4 (3 pkt)

W wyniku częściowego strawienia jednym enzymem uzyskano odcinki DNA o długościach: 20, 40, 40, 50, 60, 70, 90, 110, 110 i 150. Podaj pozycje sekwencji (markerów), w których tnie enzym (wypełnij tyle kratek, ile potrzebujesz).

--	--	--	--	--	--	--	--	--	--

Podaj ile jest tych pozycji?:

miejsce na obliczenia

Pytanie (2 pkt)

Jakie zagadnienie / zagadnienia warto byłoby dodać do MBI

Notatki lub uwagi do Roberta Nowaka:

Imię i nazwisko:

Zadanie 5 (3 pkt)

Poniżej podano fragment sekwencji referencyjnej oraz pliku SAM (kolumny POS i SEQ, odpowiadające pozycji oraz sekwencji) dla czterech kolejnych odczytów.

sekwencja referencyjna:
>chr1:1-15
agtcgcttgctcgac

plik SAM:
1 agtcgctg
3 tcgctgca
5 gcttgcacg
7 ttgcacga

- Określ sekwencję aminokwasów, które są kodowane przez tę sekwencję zakładając, że na pozycji 1 zaczyna się nowy kodon

- Czy dane z pliku SAM pozwalają wnioskować o wystąpieniu wariantu(ów)?

Jeżeli tak:

- określ typ wariantu(ów) (SNV / Indel ?),
- jaki jest jego (ich) efekt na białko,
- czy jest (są) to wariant(y) hetero czy homo-zygotyczny(e)?

- Dodaj dwa odczyty (podaj ich lokalizacje i sekwencje), które wprowadzą heterozygotyczny wariant na pozycji 3.

Tablica kodonów:

		Second base					
		U	C	A	G		
First base	U	UUU } Phenylalanine F UUC } UUA } Leucine L UUG }	UCU } Serine S UCC } UCA } UCG }	UAU } Tyrosine Y UAC } UAA } Stop codon UAG } Stop codon	UGU } Cysteine C UGC } UGA } Stop codon UGG } Tryptophan W	U	C
	C	CUU } Leucine L CUC } CUA } CUG }	CCU } Proline P CCC } CCA } CCG }	CAU } Histidine H CAC } CAA } Glutamine Q CAG }	CGU } Arginine R CGC } CGA } CGG }	C	A
	A	AUU } Isoleucine I AUC } AUA } AUG } Methionine start codon M	ACU } Threonine T ACC } ACA } ACG }	AAU } Asparagine N AAC } AAA } Lysine K AAG }	AGU } Serine S AGC } AGA } Arginine R AGG }	A	G
	G	GUU } Valine V GUC } GUA } GUG }	GCU } Alanine A GCC } GCA } GCG }	GAU } Aspartic acid D GAC } GAA } Glutamic acid E GAG }	GGU } Glycine G GGC } GGA } GGG }	G	
						U	C
						A	G

Notatki lub uwagi do Tomasza Gambina:

Zadanie 6 (3 pkt)

MBI, 24.06.2023, czas: 60 min

Imię i nazwisko:

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Dany jest dedukcyjny system wnioskowania opisany następującą bazą wiedzy (baza reguł i baza faktów).

- Baza reguł:

```
IF jest ładna pogoda THEN pójdę na spacer  
IF pada deszcz THEN siedzę w domu  
IF jest mroźna pogoda THEN pójdę na narty
```

- Baza faktów:

Jest ładna pogoda

Polecenia:

1. zaproponować system dedukcyjny in vitro na nitkach DNA implementujący powyższą bazę (jaka metoda reprezentacji wiedzy będzie wykorzystana?).
2. Opisać działanie systemu.
3. Pokazać, jaki będzie wynik wnioskowania.

Notatki lub uwagi do J. Mulawki: