بسمه تعالى





تمرین سری سوم بیوانفورماتیک

نکته مهم : تمرین سوم در سایت کوئرا قرار داده شده است و شما میتوانند از طریق $\frac{|_{y}|_{y}}{|_{y}|_{y}}$ در کلاس مربوطه در کوئرا با کلید $\frac{|_{y}|_{y}}{|_{y}|_{y}}$ ثبت نام نمایید. تحویل پاسخهای تشریحی نیز از طریق همین صفحه ی کوئرا انجام خواهد شد.

سوال ۱)

سه روش BLAST ،FASTA و داینامیک پروگرمینگ را برای جستجو یک توالی در پایگاه داده با هم مقایسه کنید.

سوال ۲)

الف) با توجه به جدول زیر تمام کلماتی که با کلمه "QIV" با حداقل امتیاز ۱۰ همتراز می شوند را بیابید.

Α	4																			
R	-1	5																		
N	-2	0	6																	
D	-2	-2	1	6																
С	0	-3	-3	-3	9															
Q	-1	1	0	0	-3	5														
E	-1	0	0	2	-4	2	5													
G	0	-2	0	-1	-3	-2	-2	6												
Н	-2	0	1	-1	-3	0	0	-2	8											
I	-1	-3	-3	-3	-1	-3	-3	-4	-3	4										
L	-1	-2	-3	-4	-1	-2	-3	-4	-3	2	4									
K	-1	2	0	-1	-3	1	1	-2	-1	-3	-2	5								
М	-1	-1	-2	-3	-1	0	-2	-3	-2	1	2	-1	5							
F	-2	-3			-2	-3	-3	-3		0	0	-3		6						
P	-1	-2	-2	-1	-3	-1	-1	-2	-2	-3	-3	-1	-2	-4	7					
S	1	-1	1	0	-1	0	0	0	-1	-2	-2	0	-1		-1	4				
T	0	-1		-1	-1	-1	-1	-2	-2	-1	-1	-1	-1	-2	-1	1	5			
W	-3	-3	-4		-2	-2	-3	-2	-2	-3		-3	-1	1	-4	-3		11		
Y	-2	-2	-2	-3	-2	-1	-2	-3	2	-1	-1	-2	-1	3	-3	-2	-2	2	7	
L V	0	-3	-3	-3	-1	-2	-2	-3	-3	3	1	-2	1	-1	-2	-2	0	-3	-1	4
	Α	R	N	D	С	Q	E	G	Н	I	L	K	М	F	Р	S	Т	W	Υ	V

ب) با توجه به کلمات بدست آمده از قسمت قبل، درخت جستجو کلمات را رسم کنید.

سوال ۳)

در هر مورد با استفاده از الگوریتم FASTA، بلندترین زیر رشته مشترک دنبالههای پایگاه داده را با رشته ورودی (Query) پیدا کنید.

الف)

Query: CACGTTGACAT

Database:

- ATGACATTCGAA
- CGATTCGGACA

ktup = 1

ب)

Query: GTTACCACG

Database:

• TACGTCGT

ktup = 2

سوال ۴)

همترازی چندگانه دنبالههای زیر را با استفاده از الگوریتم Star محاسبه کنید.

A: ACGCTAACB: TTGCACATCC: TCGGTAGATCD: TCACTGGAC

برای ماتریس امتیازدهی، از ماتریس DNAfull استفاده کنید.

/https://rosalind.info/glossary/dnafull

و همچنین برای همترازیهای دوتایی میتوانید از لینک زیر استفاده کنید. دقت شود که در مرحله اول باید رشتههای DNA را انتخاب کرده و در مرحله دوم ماتریس امتیازدهی DNAfull را انتخاب کنید.

/https://www.ebi.ac.uk/Tools/psa/emboss_needle

سوال ۵)

روش و عملکرد سه الگوریتم ClustalW ،Star و T-coffee را با یکدیگر مقایسه کنید و مزایا و معایب و محدودیتهای هر کدام را ذکر کنید.

نكات تحويل تمرين

- ۱) تمرینات را به صورت انفرادی انجام دهید.
- ۲) پاسخ های خود را با کیفیت مناسب و خوانا اسکن کرده و یا تایپ شده به صورت یک فایل درآورید و با نام
 "HW3_studentNumber_FirstnameLastname.pdf" در سایت بارگذاری نمایید.
 - ۳) در صورت پیداشدن هرگونه کپی چه در سوالهای تشریحی و چه برنامهنویسی نمرهی هر دو نفر ۱۰۰- درنظر گرفته
 خواهد شد. این نکته را جدی بگیرید تا بعداً به مشکل نخورید.
 - ۴) اشكالات خود را مى توانيد از طريق اين ايميلها بپرسيد. ftmzhrasafaei@gmail.com, aref78.m@gmail.com
- ۵) مهلت ارسال این تمرین تا ۱۴۰۱/۲/۲۴ است و پس از این تاریخ حداکثر تا دو روز به ازای هر ساعت تاخیر ۱ درصد جریمه در نظر گرفته خواهد شد. بعد از این مدت پاسخ سوالات در کانال درس ارسال شده و دیگر نمرهای به تمرین تعلق نخواهد گرفت.