

Вопросы к экзамену по биоинформатике

1. Парное выравнивание. Виды, авторы алгоритмов, цели, значение. Глобальное выравнивание.
2. Вторичные структуры белков, их характеристики и предсказание. ПО и сервисы.
3. Локальное выравнивание. Цели, значение. Алгоритм локального выравнивания.
4. Биоинформатика. Объекты биоинформатики. Задачи, решаемые этой наукой. Методы биоинформатики.
5. Матрицы сравнения последовательностей. PAM, BLOSUM.
6. По приведенной матрице расстояний постройте филогенетическое дерево (Neighbor Joining method, UPGMA). Опишите процесс построения.

	1	2	3	4
1	0	0.3	0.5	0.6
2		0	0.6	0.5
3			0	0.9
4				0

7. Основные алгоритмы построения филогенетических деревьев – их достоинства и недостатки. UPGMA и NJ (их отличия), максимальной бережливости (maximal parsimony), максимального правдоподобия, минимальной эволюции.
8. Биоинформатика и филогенез. Молекулярные часы. Клада, OTU, ветвь, лист, корень. Ультраметрическое и неультраметрическое дерево. Ортологи, паралоги, гомологи, ксенологи.
9. Редакционное расстояние между двумя последовательностями. Сложность наивного алгоритма его определения.
10. Дано: последовательности WATER и WINE. Скоринг: match- 5, mismatch- -5, вставка промежутка (gap insertion)- -1. Построить таблицу выравнивания и найти по ней путь для него.
11. Локальное выравнивание, задачи, примеры.
12. Множественное выравнивание.
13. Третичная структура белка. Фолдинг.
14. Предсказание третичной структуры белка. Моделирование гомологов. Методы, ПО и сервисы
15. Предсказание третичной структуры белка. Распознавание фолда. ПО, сервисы.

16. Динамическое программирование и выравнивание последовательностей. Способы оптимизации поиска – FASTA, BLAST
17. Классификации белков. Базы данных Pfam, SCOPE, CATH
18. NCBI, ENTREZ и BLAST – назначение, инструменты, задачи.
19. Штрафы за вставку промежутка, схемы, различия.
20. Профиль и консенсус. Сходство и различия.
21. Выравнивание и его статистическая достоверность. Bootstrap.
22. Докинг – цель и задачи. Трудности.
23. Жёсткий докинг. Методы, применение.
24. Гибкий докинг.
25. Экспериментальное определение структуры белка. Оценка качества полученной структуры.
26. Hamming distance и Edit distance – отличия.
27. Метод GOR и Chou-Fasman. Их применение.
28. Дана следующая матрица скоринга ДНК:

	A	C	G	T
A	10	2	5	2
C	2	10	2	5
G	5	2	10	2
T	2	5	2	10

Какова максимально возможная оценка выравнивания AATAAT и AAGG, при условии цены промежутка -5?

29. Допустим, нам даны 4 последовательности: S1=act, S2=agct, S3=aact, and S4=acct. Парные выравнивания этих последовательностей следующие:

a-ct
agct

a-ct
aact

a-ct
acct

По ним были построены 2 варианта MSA (в зависимости от параметров\применения алгоритма). Какой из этих вариантов Вы предпочтёте и почему?

a-ct	a---ct
agct	ag--ct
aact	a-a-ct
acct	a--cct