

Этап II. Оценка уровня знаний по научной специальности, по которой предполагается подготовить диссертацию.

Научная специальность – 1.5.8. Математическая биология, биоинформатика

Форма проведения этапа - очно в виде ответа на 3 случайных вопроса из приведенных областей знаний.

Содержание этапа - Данный этап предполагает проверку знаний по ключевым аспектам отрасли науки. Абитуриенту будет предложено три вопроса, отобранных из широкого круга тем, охватывающих фундаментальные положения и современные проблемы выбранной научной специальности.

Примерные темы для вступительного испытания

Общие знания

Математика и информатика

1. Случайные величины, распределения, математическое ожидание и дисперсия, основные распределения.
2. Математическая статистика. Выборка, нулевая гипотеза. Критерии χ -квадрат, Фишера, Стьюдента, Колмогорова. Коэффициент корреляции и регрессия. Непараметрические критерии. Множественное тестирование. Дисперсионный анализ (ANOVA). Байесовский подход.
3. Теоретическая информатика. Основные структуры данных: списки, стек, очередь, бинарное дерево поиска.
4. Алгоритмы на графах, обход в ширину и в глубину, Эйлеров цикл, поиск оптимального пути.
5. Алгоритмы для строк. Конечные автоматы, автомат Ахо – Корасик, суффиксное дерево и суффиксный массив, регулярные выражения.
6. Понятие об NP-полных задачах. Примеры NP-полных задач. Стохастические алгоритмы.
7. Реляционные базы данных, язык SQL.

Биомакромолекулы

8. Нуклеиновые кислоты. Биологическая роль. Роль ДНК в биосинтезе белка. Химическое строение и пространственная структура нуклеиновых кислот (ДНК и РНК). Разнообразие формы структур ДНК и РНК. Вода и нуклеиновые кислоты. Взаимодействие нуклеиновых кислот с белками.
9. Строение и свойства аминокислот. Классификации по свойствам боковых групп. Пептидная связь. Вторичная структура полипептидов. Методы экспериментального обнаружения вторичной структуры.
10. Белки. Физико-химические свойства и системы классификации. Иерархия уровней пространственной организации белков. Регулярные структуры полипептидной цепи: α -спираль, β -тяжи и β -листы, их распространение и роль в структурах белков. Пространственная и четвертичная структура белков. Отличия в структурной организации глобулярных и фибриллярных белков.

11. Биологическая роль белков. Ферменты: классификация; соответствие структуры и функции. Регуляторные белки. Белки, образующие матрикс – эластин, коллаген. Мембранные белки, особенности их строения и функции. Пептиды, их биологическая роль.

Молекулярная биология и генетика

12. Основные процессы передачи информации в клетке. Репликация, транскрипция, сплайсинг, трансляция. Различие и сходство основных процессов в эукариотах и в прокариотах. Посттранскрипционные модификации РНК. Посттрансляционные модификации белков. Характерные времена основных процессов.
13. Структура ДНК в прокариотах и в эукариотах. Хроматин, основные уровни организации, гистоны, модификации хроматина. Роль хроматина в регуляции экспрессии генов.
14. Регуляция экспрессии генов. Основные уровни регуляции экспрессии. Регуляция транскрипции, транскрипционные факторы. Аттенуация. Репрессия и активация. Регуляция литического и лизогенного путей фага лямбда. Сходство и различие регуляции транскрипции в прокариотах и в эукариотах. Понятие о регуляторных модулях.
15. Регуляция трансляции. Регуляция инициации трансляции в прокариотах и в эукариотах. МикроРНК и механизм регуляции с помощью микроРНК. Стабильность РНК. РНК-интерференция.
16. Стабильность генома. Репарация. Транспозоны, вирусы, ретроэлементы.
17. Наследование признаков и изменчивость. Полиморфизмы. Вредные, слабо-вредные и полезные мутации. Признаки стабилизирующего и движущего отбора. Дрейф генов. Видообразование. Медицинская генетика. Моногенные и полигенные заболевания, анализ семей, ассоциации и молекулярные причины заболеваний.
18. Понятие об основных экспериментальных методах молекулярной биологии. ПЦР, секвенирование, второе поколение секвенирования, микрочипы, белок-белковые взаимодействия, иммунопреципитация хроматина, ChIP-chip, ChIP-seq, масс-спектрометрия. Точность данных массовых экспериментов.
19. Основные методы расшифровки пространственных структур биомакромолекул. Особенности моделей, получаемых этими методами. Методы оценки качества пространственной модели белка.

Специальные знания

Биоинформатика

20. Типы и качество данных. Биологические базы данных. Первичные (архивные), курируемые и производные базы данных.
21. Выравнивание. Биологический смысл выравнивания. Понятие о «Золотом стандарте». Алгоритмы динамического программирования. Статистическая значимость выравнивания. Линейное и логарифмическое поведение веса выравнивания. Методы быстрого поиска сходства BLAST, FASTA.
22. Скрытые Марковские модели. Алгоритмы оптимального и апостериорного декодирования (Витерби и вперед-назад). Определение параметров моделей. Скрытые Марковские модели для выравнивания. Субоптимальные выравнивания.
23. Множественное выравнивание последовательностей. Динамическое программирование для множественного выравнивания. Прогрессивное выравнивание. Улучшение выравнивания.
24. Вторичные структуры РНК. Методы предсказания оптимальных структур.

Вычисление статистических сумм. Субоптимальные структуры. Поиск консервативных структур.

25. Реконструкция эволюции по последовательностям. Укоренённые и неукоренённые филогенетические деревья. Основные методы реконструкции филогении.
26. Основы анализа пространственной структуры макромолекул. Поверхность макромолекулы, алгоритмы её вычисления. Гидрофобное ядро молекулы белка, алгоритмы его нахождения. Структурные домены белков, алгоритмы их нахождения. Пространственное выравнивание структур белков. Структурные классификации доменов.

Геномика, транскриптомика, протеомика, системная биология

27. Геномы, размер геномов бактерий и эукариот. Метагеномы. Контиги. Расшифровка геномов и сборка контигов. Структура геномов прокариот. Особенности бактериальных геномов. Особенности геномов эукариот. Геном человека и млекопитающих. Полиморфизмы человека.
28. Аннотация геномов. Предсказание генов. Функциональная аннотация. Использование сходства. Сравнительный анализ геномов.
29. Доменные перестройки. Семейства доменов. Методы идентификации доменов в последовательности. Гомологи, ортологи и паралоги. Методы определения ортологичности.
30. Метаболическая реконструкция. Совместная представленность генов в геномах, колокализация, корегуляция, коэкспрессия. Базы данных метаболических путей.
31. Транскриптом. Методы определения транскриптомов. Методы анализа транскриптомов. Тканевая специфичность транскриптомов. Состав транскриптома, анализ сплайсинга. Приложения к исследованию заболеваний и диагностике.
32. Протеом. Методы определения протеома. Пост-трансляционные модификации белков. Определение посттрансляционных модификаций. Участие модификаций белков в регуляторных каскадах.
33. Эпигеномика. Методы определения эпигенома. Роль эпигенома в регуляции экспрессии генов.
34. Типы регуляторных взаимодействий. Регуляторные каскады. Системная биология. Построение и анализ регуляторных сетей. Роль системной биологии в поиске мишеней для лекарственных средств.

Пространственная организация биомacroмолекул

35. Физические взаимодействия, определяющие пространственную структуру биомолекул. Конформации и конформационная подвижность биомacroмолекул. Понятие эмпирического силового поля. Параметризация валентных и невалентных взаимодействия в биополимерах. Роль растворителя в структурной организации биополимеров. Гидрофобные взаимодействия в биомолекулярных системах. Шкалы гидрофобности. Методы учета влияния растворителя в расчетах энергии биомолекулярных систем.
36. Самоорганизация пространственной структуры биополимеров. Парадокс Левинталя. Переход клубок-глобула. Расплавленная глобула. Отличие белковой цепи от случайного сополимера. Динамика конформаций. Проблема сворачивания (фолдинга) биополимеров.
37. Диффузия лигандов в активный центр. Ферментативный катализ химических реакций. Понятие молекулярного докинга. Докинг в разработке лекарственных средств.
38. Молекулярная динамика био- и наноструктур. Подготовка системы к

моделированию молекулярной динамики. Типы силовых полей. Моделирование динамики при постоянной энергии и постоянной температуре. Равновесная и направленная (управляемая) молекулярная динамика. Возможности и ограничения моделирования молекулярной динамики.

ПРИМЕР БИЛЕТА

Вопрос 1. Случайные величины, распределения, математическое ожидание и дисперсия, основные распределения.

Вопрос 2. Регуляция трансляции. Регуляция инициации трансляции в прокариотах и в эукариотах. МикроРНК и механизм регуляции с помощью микроРНК. Стабильность РНК. РНК-интерференция.

Вопрос 3. Множественное выравнивание последовательностей. Динамическое программирование для множественного выравнивания. Прогрессивное выравнивание. Улучшение выравнивания.

РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА

1. ОСНОВНАЯ

1. Дурбин и др. Анализ биологических последовательностей. – Москва: НИЦ «Регулярная и хаотическая динамика», 2006

2. Леск. Введение в биоинформатику. - НИЦ «Регулярная и хаотическая динамика». – Москва:Бином. Лаборатория знаний, 2013

3. Eugene V. Koonin and Michael Y Galperin. Sequence – Evolution –Function. Computational Approach in Comparative Genomics. – Boston^ Kluwer Academic, 2003

4. Ж. Сетубал, Ж.Мейданис. Введение в вычислительную молекулярную биологию. – Ижевск: Регулярная и хаотическая динамика, 2007

5. С. Игнасимуту. Основы биоинформатики. - Ижевск: Регулярная и хаотическая динамика, 2007

6. Бернхард Хаубольд, Томас Вие. Введение в вычислительную биологию. - Ижевск: Регулярная и хаотическая динамика, 2011

7. Jonathan Pevsner. Bioinformatics and Functional Genomic. – Hoboken: Wiley-Blackwell, 2009

8. Иващев-Мусатов О.С. Теория вероятностей и математическая статистика. М: «ФИМА», 2003

2. ДОПОЛНИТЕЛЬНАЯ

1. Бордовский, Екишева. Задачи и решения по анализу биологических последовательностей. – Москва: НИЦ «Регулярная и хаотическая динамика», 2008

2. Бландел, Джонсон. Кристаллография белка. – Москва: Мир, 1979

3. М. Бородовский, С. Екишева. Задачи и решения по анализу биологических последовательностей. – Ижевск: Регулярная и хаотическая динамика, 2008

4. Alberts B. et al. Molecular Biology of the Cell, 2008, 2014

5. Lewin et al. Lewin`s GENES IX-XI, Jones&Bartlett Learning, 2007-2014

6. Манита А.Д. Теория вероятностей и математическая статистика. Интернет-учебник. <http://teorver-online.narod.ru>

7. <https://ncbi.nlm.nih.gov/>

Критерии и показатели оценивания ответа

	0	Отсутствуют ответы на три заданных теоретических вопроса, отказ от ответа.
Минимальный уровень знаний	1	Отсутствуют ответы на два заданных теоретических вопроса, фрагментарный ответ третий теоретический вопрос, значительные трудности в сопоставлении и анализе сведений из различных разделов программы.
	2	Отсутствует ответ на один из заданных теоретических вопросов, фрагментарный ответ на второй и третий теоретический вопросы, значительные трудности в сопоставлении и анализе сведений из различных разделов программы.
Низкий уровень знаний	3	Отсутствует ответ на один из заданных теоретических вопросов, фрагментарный ответ на второй и неполный ответ на третий теоретический вопросы, значительные трудности в сопоставлении и анализе сведений из различных разделов программы.
	4	Отсутствует ответ на один из заданных теоретических вопросов, неполный ответ на второй заданный теоретический вопрос, полный ответ на третий заданный теоретический вопрос, значительные трудности в сопоставлении и анализе сведений из различных разделов программы.
Средний уровень знаний	5	Отсутствует ответ на один из заданных теоретических вопросов, полный ответ на два заданных теоретический вопроса, значительные трудности в сопоставлении и анализе сведений из различных разделов программы.
	6	Полный ответ на один заданный теоретический вопрос, неполные ответы на два заданных теоретических вопроса, значительные трудности в сопоставлении и анализе сведений из различных разделов программы.
Достаточный уровень знаний	7	Полные ответы на два заданных теоретических вопроса, неполный ответ на один заданный теоретический вопрос, незначительные трудности в сопоставлении и анализе сведений из различных разделов программы.
	8	Полные ответы на три заданных теоретических вопроса, незначительные трудности в сопоставлении и анализе сведений из различных разделов программы.
Высокий уровень знаний	9	Исчерпывающие ответы на все заданные вопросы, свободное владение материалом, имеются недочеты при сопоставлении и анализе сведений из различных разделов программы.
	10	Исчерпывающие ответы на все заданные вопросы, свободное владение материалом, грамотные сопоставление и анализ сведений из различных разделов программы.