

МИНОБРНАУКИ РОССИИ
Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Астраханский государственный университет»

Центр развития современных компетенций детей
«Дом научной коллаборации им. В.К. Третьяковского»

(ДНК им. В.К. Третьяковского)

СОГЛАСОВАНО

Руководитель ДНК
им В.К. Третьяковского

 Д.Ю. Матвеев



УТВЕРЖДАЮ

Проректор по учебной работе

 А.М. Трещев

« 01 » февраля 2021 г.

ДОПОЛНИТЕЛЬНАЯ ОБЩЕОБРАЗОВАТЕЛЬНАЯ ПРОГРАММА –
ДОПОЛНИТЕЛЬНАЯ ОБЩЕРАЗВИВАЮЩАЯ ПРОГРАММА

БИОИНФОРМАТИКА

Направленность программы – естественнонаучная
Для учащихся 10-11 классов
Составители: д.б.н., доц. Ломтева Н.А.,
к.б.н., с.н.с. Кузина Т.В.

г. Астрахань
2021

1. Пояснительная записка

Осмысливание молекулярно-генетических данных невозможно без привлечения современных информационных технологий, создания и развития эффективных методов и алгоритмов анализа данных и моделирования биологических систем и процессов. На стыке биологии, математики, информатики возникла новая наука – информационная биология (биоинформатика). Дисциплина основана на методах организации информации; методах вычислительной математики и статистики, адаптированных к анализу последовательностей и пространственных структур биополимеров.

Исходя из этого, Биоинформатика представляет собой одно из популярных направлений, возникшее на стыке ряда дисциплин биологической, математической и компьютерной направленностей. В то время как профессия биоинформатик – одной из самых востребованных.

Программа «Биоинформатика» направлена на содействие научно-техническому развитию детей и подростков и создание научной платформы для обучения школьников современным научным методам исследования. Актуализируется вопрос о приобщении учащихся к инновационной, практико-ориентированной деятельности в сфере информационных технологий, развитии у них творческих, интеллектуальных способностей, удовлетворении индивидуальных потребностей, организации свободного времени. Программа предназначена для учащихся 10-11 классов в рамках образовательного проекта «Малая академия» и позволит уменьшить оторванность детей школьного возраста от естественнонаучных исследований, привить им интерес к научным исследованиям, что позволит им лучше понимать и осваивать школьную программу, получать более глубокие знания в области биологии. При этом программа поможет школьникам познавать не только теоретические научные знания, но и практически осваивать методы биоинформатики, работать с современными программами обработки данных.

Дополнительная общеразвивающая программа является нормативным документом, содержащим максимально полную информацию о предлагаемом дополнительном образовании по определенному виду деятельности, имеющим конкретные образовательные цели и диагностируемые образовательные результаты

Перечень документов, на основе которых разработана дополнительная общеразвивающая программа – дополнительная общеразвивающая программа:

- Конституция РФ;
- Федеральный закон от 29.12.2012 № 273-ФЗ «Об образовании в Российской Федерации»;
- Конвенция о правах ребенка;
- СанПиН 2.4.4.3172-14;
- Приказ Министерства просвещения Российской Федерации от 09.11.2018 № 196 «Об утверждении Порядка организации и осуществления образовательной деятельности по дополнительным общеобразовательным программам»;
- Распоряжение Правительства РФ от 04.09.2014 № 1726-р «Об утверждении концепции развития дополнительного образования детей»;
- Распоряжение Правительства РФ от 29.05.2015 № 996-р «Об утверждении стратегии развития воспитания на период до 2025 года»;
- «Примерные требования к программам дополнительного образования детей», предложенные в приложении к письму Департамента молодежной политики, воспитания и социальной поддержки детей Минобрнауки России от 11.12.2006 № 06-1844 и требованиями, содержащимися в письмах МО и ВШ РК от 12.08.2003 № 07-18/94, от 11.01.2007 № 07-18/2 на основании типовых (примерных) программ;
- Приказ от 17 декабря 2010 г. № 1897 «Об утверждении федерального государственного образовательного стандарта основного общего образования» (в

ред. Приказа Министерства образования и науки Российской Федерации от 29.12.2014 № 1644);

- Федеральный перечень учебников, утвержденных, рекомендованных к использованию в образовательном процессе в образовательных организациях, реализующих программы основного общего образования, утвержденный приказом Министерства образования и науки Российской Федерации от 31 марта 2014 г. № 253 (с изменением на 26 января 2016 г.).

Цель и задачи программы

Цель программы – способствовать развитию творческих способностей детей, формированию и развитию учебно-познавательных, информационно-компьютерных компетенций, развитию навыков soft skills (решать проблемы, уметь взаимодействовать в команде, критически мыслить, вести диалог).

Задачи программы:

1. Развить у обучающихся техническое мышление, изобретательность, образное и пространственное мышление.
2. Развить управленческие навыки (управление временем, саморазвитием, решение проблем, тайм-менеджмент).
3. Развить навыки работы с информацией (поиск и анализ информации, креативное мышление).
4. Формировать учебную мотивацию и мотивацию к творческому поиску; развивать волю, терпение, самоконтроль, внимание, память, фантазию; развивать способности осознанно ставить перед собой конкретные задачи и добиваться их выполнения.
5. Воспитывать дисциплинированность, ответственность, самоорганизацию; трудолюбие, уважение к труду; чувство коллективизма и взаимопомощи; способствовать раскрытию внутреннего мира обучающихся.
6. Воспитывать самостоятельность в приобретении дополнительных знаний и умений.
7. Создать систему научно-технического просвещения через привлечение детей и подростков к изучению и практическому применению методов биоинформатики.
8. Способствовать освоению на практике базовых методов биоинформатики, включая работу с молекулярными базами данных, выравнивание последовательностей и молекулярную визуализацию.
9. Обеспечить подготовку профессионально ориентированного подрастающего поколения – будущих специалистов в области биоинформатики для Астраханского региона и Российской Федерации в целом.
10. Развитие мотивации к научно-исследовательской и проектной деятельности, потребности в саморазвитии, самостоятельности, ответственности, активности, аккуратности.
11. Научить создавать и презентовать проекты.

2. Направленность программы: естественнонаучная

3. **Новизна:** предоставление высокотехнологичных образовательных услуг широкому кругу школьников Астраханского региона и Прикаспия с перспективой обеспечения подготовки высококвалифицированных конкурентоспособных на рынке труда кадров в сферах общего и специального биологического образования, здравоохранения и социальной защиты.

4. Актуальность программы: систематические занятия школьников позволят сформировать учебно-познавательные и информационные компетенции, которые помогут сформировать профессионально ориентированного специалиста в области биоинформатики

5. Педагогическая целесообразность: программа выстроена на основе целесообразности освоения учащимися глубокого и полного содержания учебного материала в предметной области «Биология», IT-творчества, умений проектирования, исследования, решения интеллектуальных задач, профессионального самоопределения и профориентации.

6. Практическая значимость программы: учащиеся смогут продолжить образование по выбранному профилю после завершения курса обучения в организациях профессионального и высшего образования по биологическим и медицинским специальностям.

Характеристика программы

Вид – дополнительная общеобразовательная программа – дополнительная общеразвивающая программа.

Адресат программы: учащиеся 10 – 11 классов

Объем и срок освоения программы: 72 часа, 3 месяца

Формы обучения – очная, очно-заочная или заочная форма с применением дистанционных образовательных технологий.

Режим занятий в соответствии с санитарно-эпидемиологическими требованиями к образовательной организации дополнительного образования.

Уровневая дифференциация программы:

Программа имеет «Стартовый уровень» и предполагает минимальную сложность предлагаемого для освоения содержания программы. Реализация стартового уровня предполагает наличие программы не более чем на один год обучения в количестве до 72 часов. Данная программа может быть использована: – как самостоятельный курс освоения определенного вида деятельности; – как первая ступень - переход к базовой общеразвивающей программе обучения; состав учащихся (слушателей) может быть сменным, как одновозрастным, так и разновозрастным, при этом рекомендуемая сменяемость за весь период освоения программы составляет не более 50 %.

Ожидаемый (прогнозируемый) результат освоения программы

Деятельность обучающихся в рамках программы должна быть направлена на достижение обучающимися следующих личностных результатов:

1) формирования теоретической базы знаний в области биологической информатики: содержание, возможности, методы и алгоритмы получения, представления и анализа данных в биоинформатике для решения прикладных задач в области биомедицины, экологии, филогении и т.д.; 2) побуждение интереса у школьников к изучению молекулярной биологии, информатике, программированию, математическому анализу; 3) сформированность познавательных интересов и мотивов, направленных на изучение живой природы; интеллектуальных умений (доказывать, строить рассуждения, анализировать, сравнивать, делать выводы и др.); эстетического отношения к живым объектам.

Метапредметные результаты:

1) овладение составляющими исследовательской и проектной деятельности, включая умения видеть проблему, ставить вопросы, выдвигать гипотезы, давать определения понятиям, классифицировать, наблюдать, проводить эксперименты, делать выводы и заключения, структурировать материал, объяснять, доказывать, защищать свои идеи; 2) умение работать с разными источниками биологической информации: находить биологическую информацию в различных источниках (тексте учебника, научно-

популярной литературе, биологических словарях и справочниках), анализировать и оценивать информацию, преобразовывать информацию из одной формы в другую; 3) способность выбирать целевые и смысловые установки в своих действиях и поступках по отношению к живой природе, здоровью, своему и окружающим; 4) умение адекватно использовать речевые средства для дискуссии и аргументации своей позиции, сравнивать разные точки зрения, аргументировать свою точку зрения, отстаивать свою позицию.

Предметные результаты:

1) знание и соблюдение правил работы в лаборатории; 2) соблюдение правил работы с биологическими программами; 3) знание характеристик основных молекулярно-генетических методов; 4) знание общих принципов работы с последовательностями нуклеотидов, с базами данных; 5) понимание сферы применения биоинформатики (сельское хозяйство, медицина, пищевая промышленность, энергетика и т.п.); 6) умение работать с различными источниками информации.

2. Условия реализации программы

Описание материально-технического и информационно-методического обеспечения программы. Оборудование: компьютеры, программное обеспечение.

3. Учебно-тематический план

№ п/п	Раздел	Количество часов			Формы аттестации/ контроля
		Всего	лекция	ПЗ	
1	Введение в биоинформатику	6	4	2	
2	Базы данных				
2.1	Поиск заданных последовательностей белков и нуклеиновых кислот в базах данных	12	6	6	Отчет по работе
3	Визуализация биомолекул с использованием компьютерных программ RasMol и ACD ChemSketch	12	6	6	Отчет по работе
4	Выравнивание последовательностей				
4.1	Попарное выравнивание заданных последовательностей	10	4	6	Отчет по работе
4.2	Множественное выравнивание заданных последовательностей	10	4	6	Отчет по работе
5	Построение филогенетических деревьев	10	4	6	Отчет по работе
6	Применение биоинформатики в медицинских исследованиях, сельском хозяйстве и пищевой промышленности	10	4	6	
	Зачетное задание – защита проекта	2		2	Выполнение проекта и его защита
	ИТОГО	72 часа			

4. Содержание изучаемого курса

Тема 1 Введение в биоинформатику

Теория. Наука биоинформатика. Цели и задачи биоинформатики. Предмет биоинформатики. Развитие методов расшифровки последовательностей биополимеров – исторический аспект. Работы Ф. Сэнгера и Эдмана. Реакции обрыва цепи и химического расщепления. Полимеразная цепная реакция. Динамика накопления информации в базах данных последовательностей. Проект «Геном человека». Прикладное значение биоинформатики: анализ гомологичности последовательностей; анализ экспрессии генов; разработка лекарственных препаратов. История возникновения биоинформатики как науки. Современные взгляды на биоинформатику, ее возможности и перспективы

Практика.

Знакомство учащихся друг с другом и с педагогом. Игра на знакомство и сплочение. Самопрезентация.

Тема 2 Базы данных

Теория. Виды биологических последовательностей. Источники. Биоинформатические базы данных. Поиск информации в системе Entrez. Проект RefSeq. Поиск научной литературы в PubMed. Генная онтология (GO Ontology) и функциональный анализ генов в системе AmiGO.

Практика. Поиск заданных последовательностей белков и нуклеиновых кислот в базах данных. Знакомство с банком данных NCBI. Реализация простых и сложных запросов. Поиск последовательностей в БД по гомологии.

Поиск белков в базах данных.

1. Найдите в БД UniProtKB документ, содержащий информацию о белке, указанном в таблице 1 напротив вашей фамилии. Для этого откройте главную страницу (<http://www.uniprot.org/uniprot/>). Наверху страницы появится поле для введения запроса для поиска по БД UniProtKB. Введите в него ID номер вашего белка.
2. Внимательно изучите полученный документ и заполните табл. 2.
3. Создайте текстовый файл с последовательностью Вашего белка в формате FASTA.
4. Определите, сколько документов в БД UniProtKB содержит белки с тем же кратким описанием, что и Ваш белок.

Работа с системой PubMed

1. Зайдите на сайт NCBI (<http://www.ncbi.nlm.nih.gov/>) и перейдите по гиперссылке "PubMed" (слева экрана в меню Popular Resources) Найдите ссылки на статьи, в которых упоминается полное название Вашего белка. Приведите в протоколе это название и количество найденных статей. Сохраните результаты поиска в виде текстового файла (используйте меню "Send to", в котором выберите "File").
2. Из списка найденных статей выберите одну (укажите в протоколе, какую). Сохраните её аннотацию ("Abstract") в текстовом формате.
3. Найдите ссылки на все публикации последнего автора этой статьи за последние 3 года. В протоколе укажите число статей. Чтобы найти публикации человека, который подписывает статьи как A.I. Petrov, следует написать в строке запроса: petrov ai [Author]. Для ограничения поиска последними тремя годами воспользуйтесь сервисом "Limits".

Тема 3 Визуализация биомолекул с использованием компьютерных программ RasMol и ACD ChemSketch

Теория. Понятие о белках и нуклеиновых кислотах как биологических макромолекулах: открытие, строение, пространственная структура, физико-химические

свойства, функции, методы выделения и анализ. Структура белка (вторичная, третичная, четвертичная). Методы получения трехмерной структуры белка. PDB. Структура PDB файла.

Практика.

Визуализация молекул с помощью программы ACD ChemSketch

1. В редакторе ACD нарисуйте химическую структуру пятой аминокислоты вашего белка (табл. 1). Изобразите как L, так и D форму аминокислоты, сделайте подписи. Проведите трехмерную оптимизацию, переведите рисунок в 3D формат и скопируйте полученные структуры в документ Microsoft Word.

2. В редакторе ACD создайте новую страницу и скопируйте туда L-форму вашей аминокислоты. Нарисуйте реакцию образования пептидной связи между четвертой и пятой аминокислотами вашего белка. Скопируйте полученный рисунок в файл *.doc.

3. На новой странице нарисуйте трипептид, образованный третьей, четвертой и пятой аминокислотами вашего белка. Выделите красным остов пептида и скопируйте рисунок в файл *.doc. В редакторе ACD сохраните файл со всеми полученными структурами.

Визуализация молекул с помощью программы RasMol

1. Найдите в БД изображение пространственной структуры своего белка (табл. 1). Для этого нужно войти на сервер NCBI/Structure (<http://www.ncbi.nlm.nih.gov/Structure/index.shtml>) и ввести в поисковой строке PDB_ID белка. В появившемся окне нажмите на иконку с изображением белка. Затем выберите кнопку «Structure View in RasMol» и загрузите файл на компьютер. Скопируйте в протокол образец описания состава PDB-структуры (табл. 5) и отредактируйте заголовок в соответствии со своим белком.

2. Откройте загруженный pdb-файл с помощью программы RasMol. Для этого необходимо запустить программу и через меню File загрузить изображение. Обратите внимание, что появляется два окна: с изображением белка и командное: "RasMol command Line". В командное окно можно вводить команды, позволяющие работать как со всей молекулой, так и с отдельными множествами атомов, входящих в ее структуру. Подробное описание команд представлено в приложении 2.

3. Создайте светлый фон (команда background white или по системе RGB: background [230,250,200] (числа не должны превышать 255)). Изобразите цепи белка в остовой модели, выбрав подходящую толщину линий: backbone [30, 40 и т.д.]

Раскрасьте разные цепи белка в разные цвета: color chain

4. Определите имена полипептидных цепей, входящих в состав молекулы белка. Щелкая по концевым C -атомам (в остовой модели C -атомы расположены в вершинах ломаной), получите в командной строке информацию об имени цепи, типе и номере остатка и др. Внесите информацию об именах цепей, типах и номерах концевых остатков в протокол.

5. Изобразите N-концевые азоты большими (400) синими шариками:

```
select <имя атома>:цепь (например, select arg58:a.n , где a – цепь, n – азот)
cpk 400
color blue
```

Тип остатка, номер остатка и имя цепи прочитайте в собственном протоколе. Аналогичным образом изобразите C-концевые кислороды красными шариками. Имейте в виду, что их по 2 на каждую цепь (-COOH !). Чтобы узнать как называется второй кислород из карбоксильной группы выделите C-концевой остаток целиком (например, select 327A), изобразите его в проволочной модели (wireframe 50), покрасьте по типам атомов (color cpk) и щелкая по атомам кислорода, определите их имена в командном окне.

6. Измените окраску цепей так, чтобы α -цепи гемоглобина отличались от β -цепей. Чтобы покрасить цепь с именем X:

```
select *:X
color <цвет>
```

Например, А-цепь зеленая (green), В-цепь - оранжевая (orange). Имена цепей — в вашем протоколе.

7. Определите, какие небелковые компоненты присутствуют в структуре молекулы. Для этого их необходимо выделить (select not protein или select hetero), увеличить размер и покрасить по типам атомов. После их выделения, небелковые компоненты в командной строке будут иметь ID-шифр. Чтобы узнать, какое соединение зашифровано под условным обозначением, можно так же воспользоваться сервисом NCBI / Structure.

Заполните таблицу 5 в протоколе.

8. Вращая молекулу, подберите наиболее удачный ракурс и сохраните полученное изображение в графическом файле формата gif (меню Export). Творческий подход к выполнению заданий (подбор цветов объектов, толщины линий и способов изображения и т.п.) приветствуется!

Тема 4 Выравнивание последовательностей

Теория 1. Попарное выравнивание заданных последовательностей

Выравнивание, его цели. Общие принципы выравнивания. Основные понятия и определения. Цели и типы выравниваний. Методы изучения подобий. Попарное выравнивание. Попарное выравнивание заданных последовательностей с помощью программы GeneDoc. Попарное выравнивание заданных последовательностей с использованием программы BLAST

Теория 2. Множественное выравнивание заданных последовательностей

Множественное выравнивание заданных последовательностей. Выравнивание последовательностей с использованием программы CLUSTAL W.

Практика.

Выравнивание последовательностей с помощью программы GeneDoc

1. Скопируйте пару коротких последовательностей из таблицы 3 (против своей фамилии) в текстовый файл. Запустите программу GeneDoc и импортируйте этот файл. Для этого в меню File выбираете Import. В открывшемся окошке выбираете формат "Fasta (Pearson)" и нажимаете "Import". Выбираете файл, затем "Open" и "Done".

2. Выводите последовательности, стараясь, чтобы было сопоставлено максимальное число одинаковых букв. Редактировать выравнивание можно двумя способами: а) Включите режим "Grab and drag sequences". Это можно сделать либо в меню Arrange, либо комбинацией <Ctrl+A>. В этом режиме можно "цеплять" мышью участки последовательности и сдвигать их. б) Режимы "Insert gap into sequence" и "Delete gap from sequence" в меню Arrange. В этих режимах можно вставлять/убирать гэпы щелчком мыши. (Более детальное описание работы с программой представлено ниже).

3. Сохраните выравнивание под именем alignment1.msf. Посчитайте и занесите в протокол процент идентичности и процент сходства двух последовательностей. Процентом идентичности считается отношение числа колонок выравнивания, в которых стоят одинаковые буквы, к общему числу колонок (включая "гэповые"), умноженное на 100%. Процентом сходства будем считать отношение числа колонок со сходными буквами к общему числу, умноженное на 100%.

Выравнивание последовательностей с помощью программы BLAST

Последовательность seq1 из табл. 3 представляет собой фрагмент последовательности вашего белка (табл. 1). Чтобы определить, какой именно это фрагмент, выровняйте его с полной последовательностью белка. Для этого используйте программу "bl2seq", которая предназначена в первую очередь для выравнивания одинаковых или очень сходных последовательностей.

Зайдите на сайт "<http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi>" (web-интерфейс к программе "bl2seq"). Определите с помощью этого сервиса координаты последовательности seq1 в полной последовательности вашего белка.

Чтобы выравнивать аминокислотные последовательности, надо в меню Program выбрать "blastp". В поле "Sequence 1" скопируйте либо последовательность вашего белка в fasta-формате, либо просто его первый AC или ID. В поле "Sequence 2" положите фрагмент, координаты которого Вы хотите найти. Нажмите кнопку Blast и дождитесь результата. Занесите ответ в протокол.

2. Пользуясь тем же сервисом, выровняйте ваш белок с другим белком, указанным против вашего имени в таблице 4. Занесите в протокол: какие белки из каких организмов Вы выравнивали, проценты идентичности ("Identities") и сходства ("Positives"), число гэпов, координаты выровненного участка (участков, если выдано несколько выравниваний) в обеих последовательностях. Сохраните карту локального сходства в виде gif-файла; желательно, чтобы в имени файла фигурировали идентификаторы последовательностей.

Редактирование результатов выравнивания

1. Импортируйте последовательности, которые Вы выравнивали в предыдущем задании, в GeneDoc. Попробуйте вручную воспроизвести выравнивание, полученное программой bl2seq. Поскольку программа выдала одно или несколько частичных выравниваний, а вам нужно сделать полное, можно против всех остатков, не вошедших в частичные выравнивания, поставить гэпы. Ну а можно попытаться найти что-нибудь разумное и вне выданных выравниваний.

2. Поэкспериментируйте с программой bl2seq — попробуйте поменять матрицу или штрафы за гэпы (гиперссылка "Algorithm parameters"). Опишите результаты в протоколе.

Поиск белка по его гомологу

При выполнении задания пользуйтесь web-интерфейсом к BLASTP на сервере NCBI: <http://www.ncbi.nlm.nih.gov/blast/>. В разделе Basic BLAST необходимо перейти по гиперссылке "protein blast".

1. Подайте на вход программе BLASTP (Enter Query Sequence) последовательность своего белка из табл. 4. Проведите поиск гомологичных последовательностей в БД Swissprot. Внимание! Для этого надо изменить значение параметра database – по умолчанию стоит банк "nr". Чтобы получить результат, надо нажать кнопку BLAST и подождать одну-две минуты.

2. В выдаче программы отыщите поданный на вход белок и ваш белок из таблицы 1. Занесите в протокол их порядковые номера в выдаче, Score, E-value.

Письменно ответьте на вопросы: Что за белок является последним в выдаче программы? Каков его порядковый номер, Score, E-value? Объясните, почему программа не выдала больше находок. Какие параметры программы в данном случае надо изменить, чтобы находок стало больше? Проверьте своё предположение.

Повторите поиск с той же входной последовательностью, указав в качестве банка pdb. Для первой последовательности в списке находки укажите: PDB-коды и идентификаторы, Score, E-value, начало и конец выравнивания во входной последовательности (Query) и в находке (Subject), процент совпадений (Identity).

3. Повторите поиск по Swiss-Prot, подав на вход не всю последовательность, а любую её треть. Ответьте на вопросы: Является ли исходная последовательность первой в списке находок? Как изменились Score и E-value? Объясните наблюдаемые изменения.

Тема 5 Построение филогенетических деревьев

Теория. Филогения и эволюционные деревья. Подходы к изучению филогенеза, видового разнообразия и эволюционных взаимоотношений на основе геномных и протеомных исследований. Современные принципы биологической таксономии.

Филогенетические модели и анализ данных. Сравнительный анализ геномов в филогенетических исследованиях.

Практика.

Провести идентификацию нуклеотидной последовательности и построить филогенетическое дерево с использованием системы BLAST

Требуется компьютер исходный файл с нуклеотидными последовательностями COI.

Шаг 1. Из файла, содержащего выборку последовательностей COI, выберите последовательность интересующего вас вида и скопируйте ее в новый текстовый файл в формате .doc. Буквенная последовательность не должна содержать информации о названии вида, которому принадлежит последовательность. Обменяйтесь созданными файлами с коллегой по группе. Получив от коллеги файл, содержащий нуклеотидную последовательность неизвестного вам происхождения, попытайтесь установить, какому виду она принадлежит.

Для этого пройдите по ссылке <http://www.ncbi.nlm.nih.gov/> или наберите в любой поисковой системе слово GenBank. На странице NCBI внизу найдите ссылку на страницу системы BLAST. Войдите в систему BLAST. Поскольку вы имеете дело с нуклеотидной последовательностью, выберите программу «nucleotide blast». В открывшемся окне введите вашу буквенную последовательность без каких-либо пробелов в специальное поле «Enter accession number(s)». Убедитесь, что в условиях поиска заданы параметры «Искать в нуклеотидной коллекции» и «искать высоко сходные последовательности» (Highly similar sequences (mega blast)).

Проведите поиск, нажав кнопку BLAST внизу страницы. Определите принадлежность анализируемой последовательности. Оцените уровень сходства анализируемой последовательности с другими последовательностями в базе данных. Загрузите на свой компьютер описание полученных результатов («Taxonomy reports»). Программа представляет результаты как список видов, с указанием процента сходства между анализируемой последовательностью и последовательностью из списка. Последовательности в списке ранжируются в соответствии с уровнем подобия содержащихся в базе последовательностей и анализируемой. Кроме того программа выводит в окне результатов подробный отчет о результатах сравнения со всеми максимально подобными последовательностями в виде попарных выравниваний с указанием как идентичных, так и различающихся сайтов.

Шаг 2. На странице с результатами поиска нажмите на вкладку «Distance tree of results». Программа построит филогенетическое дерево, на котором ваша последовательность будет выделена желтым цветом. Убедитесь, что автоматически выбранный программой метод построения филогенетического дерева вас устраивает. В противном случае проведите повторное построение дерева («Reset»). Выберите тот вариант представления дерева, который кажется вам наиболее удачным («rectangle», «slanted», «radial» или «force»). Построение филогенетического дерева может быть использовано как дополнительный способ идентификации последовательности, но только в том случае, когда речь идет об определении таксономической принадлежности образца, для которого эта последовательность была получена. Идентифицировать гены таким образом нельзя, поскольку объем информации, который потребуется для анализа, превысит возможности любого из существующих компьютеров. Поскольку на эволюционных деревьях последовательности объединяются в кластеры по принципу генетического сходства, в один кластер попадут родственные последовательности, в идеальном случае – последовательности одного вида. В зависимости от того, с какими последовательностями кластеризуется анализируемая последовательность неизвестного вида, можно сделать достаточно точный вывод о ее происхождении. Идентификация по филогенетическому родству может быть использована в том случае, когда важно не только установить видовую принадлежность, но и оценить генетическую близость образца к отдельным выборкам

последовательностей организмов того же вида, например, к конкретной популяции, линии, гаплотипу и т.д. Загрузите полученное дерево в формате «Newick» на свой компьютер.

Тема 6 Применение биоинформатики в медицинских исследованиях, сельском хозяйстве и пищевой промышленности

Теория. Применение биоинформатики в медицинских исследованиях. Поиск генов, генетических полиморфизмов, генов предрасположенности. Применение биоинформатики в сельском хозяйстве, пищевой промышленности.

Практика.

Поиск в Pubmed публикаций за 2020 год с ключевыми словами: деменция, Fatal familial insomnia, neuro genetics.

5. Методическое обеспечение программы

методы обучения (словесный, наглядный практический; объяснительно-иллюстративный, репродуктивный, частично-поисковый, исследовательский проблемный; игровой, дискуссионный, проектный и др.) и воспитания (убеждение, поощрение, упражнение, стимулирование, мотивация и др.);

формы организации образовательного процесса: индивидуальная, индивидуально-групповая и групповая; выбор той или иной формы обосновывается с позиции профиля деятельности (музыкального, спортивного, художественного и др.), категории обучающихся (дети-инвалиды, дети с ОВЗ) и др.;

формы организации учебного занятия – акция, аукцион, бенефис, беседа, вернисаж, встреча с интересными людьми, выставка, галерея, гостиная, диспут, защита проектов, игра, концерт, КВН, конкурс, конференция, круглый стол, круиз, лабораторное занятие, лекция, мастер-класс, «мозговой штурм», наблюдение, олимпиада, открытое занятие, посиделки, поход, праздник, практическое занятие, представление, презентация, рейд, ринг, салон, семинар, соревнование, спектакль, студия, творческая мастерская, тренинг, турнир, фабрика, фестиваль, чемпионат, шоу, экскурсия, экзамен, экспедиция, эксперимент, эстафета, ярмарка;

педагогические технологии – технология индивидуализации обучения, технология группового обучения, технология коллективного взаимообучения, технология программированного обучения, технология модульного обучения, технология блочно-модульного обучения, технология дифференцированного обучения, технология разноуровневого обучения, технология развивающего обучения, технология проблемного обучения, технология дистанционного обучения, технология исследовательской деятельности, технология проектной деятельности, технология игровой деятельности, коммуникативная технология обучения, технология коллективной творческой деятельности, технология развития критического мышления через чтение и письмо, технология портфолио, технология педагогической мастерской, технология образа и мысли, технология решения изобретательских задач, здоровьесберегающая технология, технология-дебаты и др.

алгоритм учебного занятия – краткое описание структуры занятия и его этапов; **дидактические материалы** – раздаточные материалы, инструкционные, технологические карты, задания, упражнения, образцы изделий и т.п.

6. Список литературы

1. Белькова Н.Л. Большой практикум по биоинженерии и биоинформатике : учеб.-метод. пособие : в 3 ч. / Н. Л. Белькова. – Иркутск : Изд-во ИГУ, 2013.

2. Дурбин Р. Анализ биологических последовательностей / Р. Дурбин, Ш. Эдди, А. Крэг, Г. Митчисон. – М. – Ижевск : РХД, 2006. – 480 с.
3. Игнасимуту С. Основы биоинформатики / С. Игнасимуту ; пер. с англ. А. А. Чумичкин. – Ижевск : Регулярная и хаотическая динамика : Ин-т компьютер. исслед., 2007. – 316 с.
4. Каменская М.А. Информационная биология / М. А. Каменская. – М.: Академия, 2006. – 368 с.
5. Леск А. Введение в биоинформатику : пер. с англ. / А. М. Леск ; ред.: А. А. Миронов, В. К. Швядаса. – М. : Бином. Лаборатория знаний, 2009. – 318 с.
6. Огурцов А.Н. Методы биоинформационного анализа / А.Н. Огурцов. – Х.: НТУ "ХПИ", 2011. – 114 с. 5. Основы биоинформатики : учеб. пособие / А. Н. Огурцов. – Х.: НТУ «ХПИ», 2013. – 400 с.
7. Приставка А.А. Большой практикум по биоинженерии и биоинформатике : учеб.-метод. пособие : в 3 ч. / А.А. Приставка, В. П. Саловарова – Иркутск : Изд-во ИГУ, 2013.
8. Структура и функционирование белков: применение методов биоинформатики / пер. с англ.: В. Н. Новоселецкий, Е. Д. Балицкая, Т. В. Науменкова ; ред. В. Н. Новоселецкий. – М. : УРСС : Ленанд, 2014. – 414 с.