

Межрегиональная олимпиада школьников

«Будущие исследователи — будущее науки»

Анализ информации касательно, полученных
поиском информации, результатов - анализ, выявление
данных SARS-CoV-2 Omicron B.1.1.529

Секция: Биология

Научный руководитель
(ученая степень, звание, должность) Курь Курьев А.В.
(подпись) (расшифровка подписи)

Количество баллов,
полученных на защите 95
(заполняется председателем жюри)

Председатель жюри Курь А.В.
(подпись) (расшифровка подписи)

Работу выполнил(а)
учащий(ая)ся 22.01 класса

Университетской гимназии
г. Саров
(полное наименование учебного заведения)

г. Саров
(название населенного пункта)

Тихомирова Елена Евгеньевна
(Ф.И.О. учащегося ПОЛНОСТЬЮ)

Саров
2025 год

ВВЕДЕНИЕ.....	3
1. Теоретические основы исследования.....	4
1. 1. Нанотела, их структура, сравнение с антителами; IgNAR и VNAR.....	4
.....	5
1.2. Библиотеки нанотел	6
1.3. Дисплейные технологии	6
1.4. Метод секвенирования.....	9
1.5. Биоинформатические методы	9
1.6. Подготовка и конструкция VNAR библиотек для дисплея.....	10
2. Материалы и методы исследования	10
3. Практическая часть	12
ЗАКЛЮЧЕНИЕ	15
ЛИТЕРАТУРА.....	16

ВВЕДЕНИЕ

Актуальность. Иммунные библиотеки являются важным ресурсом для получения новых молекул, нацеленных на биологические мишени, в том числе для создания новых лекарственных препаратов. Библиотека нанотел, полученная из акульих белков представляет из себя набор кДНК, каждая из которых кодирует последовательность нанотела — антиген-связывающего домена, обладающего аффинностью к определенному антигену. Большое разнообразие генов, представленных в библиотеках, делает возможным выбор нанотел, обладающих наибольшей аффинностью к молекуле-мишени, что увеличивает эффективность борьбы с патогеном.

Данная работа посвящена анализу библиотеки нанотел, полученной в ходе иммунизации рецептор связывающим доменом SARS-CoV-2 Omicron BA.2, который относится к оболочечным РНК-содержащим вирусам и несет на поверхности клетки особый S-белок, использующийся вирусом для прикрепления к клеткам-мишеням, имеющим рецепторы ангиотензинпревращающего фермента II типа (АПФ2) [1]. Данный вариант вируса характеризуется высоким числом мутаций по сравнению с более ранними штаммами SARS-CoV-2, что обуславливает его контагиозность и патогенность [2]. Появление новых вариантов вируса SARS-CoV-2, таких как Omicron BA.2, обуславливает необходимость разработки новых методов диагностики и лечения, которые будут действовать эффективно [2]. Исследование библиотек нанотел, специфичных к SARS-CoV-2 Omicron BA.2 может иметь значительное влияние на будущие разработки в области вакцин и терапий.

Проблема исследования. Для того чтобы провести скрининг с помощью фагового дисплея библиотеки необходимо подготовить кДНК для клонирования последовательностей в фагмиды. Это в свою очередь требует выявления константных участков среди большого числа клонов и подбор соответствующих праймеров. Выявление таких участков проводится на основе биоинформатического анализа библиотек после секвенирования.

Для решения указанной проблемы была поставлена цель исследования: провести анализ последовательностей и получить информацию о константных участках библиотеки кДНК.

Для достижения данной цели были поставлены следующие задачи:

- 1) провести иммунизацию кошачьих акул для получения мРНК из лимфоцитов с дальнейшим осуществлением обратной транскрипции и секвенирования последовательностей кДНК
- 2) провести биоинформатическую обработку сиквенсов, полученных в результате секвенирования
- 3) практически выявить константные участки библиотеки кДНК
- 4) провести подбор праймеров по константным участкам библиотеки

Гипотеза: возможно выявить константные участки библиотеки кДНК, полученной в результате иммунизации акул, которая после подбора праймеров будет использоваться для успешного клонирования последовательностей в фагмиды и дальнейшего проведения скрининга с помощью фагового дисплея, если:

- 1) получить кДНК из лимфоцитарной фракции крови кошачьих акул;
- 2) провести RACE-секвенирование полученных последовательностей кДНК;
- 3) проанализировать полученные последовательности с использованием биоинформатической программы UGENE и базы данных GenBank.

1. ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ИССЛЕДОВАНИЯ

1. 1. Нанотела, их структура, сравнение с антителами; IgNAR и VNAR

Нанотела представляют собой антигенсвязывающие фрагменты антител, состоящие только из тяжелых цепей [3]. История открытия данных молекул началась в 1993 году, когда группа бельгийских ученых обнаружила в крови животных семейства Верблюдовых (Camelidae), помимо обычных иммуноглобулинов, уникальные антитела, которые сильно отличались по строению и размеру от иммуноглобулинов других млекопитающих: в их составе были обнаружены только константные участки CH2 и CH3, переменный домен тяжелой цепи (VH), шарнирная область, соединяющая переменный домен и второй константный участок (CH2) [4]. В состав переменного домена входят каркасные области (FR1, FR2, FR3, FR4), гиперпеременные участки, распознающие антиген (CDR1, CDR2, CDR3), в то время как обычные антитела (IgG) являются гетеродимерами и имеют в своем составе, помимо двух тяжелых цепей (VH), две легкие цепи (VL), антигенсвязывающий фрагмент (Fab), также Fc-фрагмент, состоящий из константных доменов C2 и C1, благодаря которому осуществляется связывание молекулы иммуноглобулина с Fc-рецептором на поверхности клетки-эффектора (например, на поверхности фагоцита или тучной клетки) или на поверхности белков системы комплемента [5]. Антигенсвязывающий фрагмент в своем составе имеет константный домен легкой цепи (CL), константный домен тяжелой цепи (C1), переменный домен легкой цепи (VL) и переменный домен тяжелой цепи (VH).

Нанотела обладают сравнительно низкой молекулярной массой (до 15 кДа, по сравнению с массой в 150-170 кДа, характерной для антител), что обеспечивает их быстрое проникновение в ткани и короткий период полураспада в крови за счет легкого прохождения почечного фильтра [6]. Для нанотел характерна высокая растворимость и стабильность за счет дисульфидной связи между CDR1 и CDR3 и четырех гидрофильных аминокислотных остатков в FR2 [7]. CDR3 в составе нанотела образует протяженный петлевой участок, обладающий высокой аффинностью к небольшим скрытым эпитопам на поверхности антигенов [7]. Из-за вышеперечисленных характеристик данные молекулы обладают большим потенциалом при использовании их в качестве инструментов в таких областях биотехнологии, как диагностика и терапия.

Помимо семейства Верблюдовых, нанотела были также обнаружены в организме хрящевых рыб. Они были названы иммуноглобулинами с новым антигеновым рецептором, или IgNAR. Данные молекулы существуют в виде димеров идентичных тяжелых цепей, между константными доменами C3 и C4 находится шарнирная область [8]. Каждая цепь состоит из пяти константных доменов (C1,C2,C3,C4,C5) и одного переменного домена [8]. VNAR является однодоменным антителом, обнаруженным в переменном домене IgNAR, в его антигенсвязывающем участке имеется 8 β-тяжей вместо 10, как в переменных доменах иммуноглобулинов млекопитающих, что делает его самой маленькой среди обнаруженных на данный момент молекул фрагментов антител с молекулярной массой около 12 кДа [8,9]. У молекулы VNAR отсутствует CDR2, при этом она обладает длинным фрагментом CDR3, который облегчает антигенное распознавание и доступ к различным эпитопам [7].

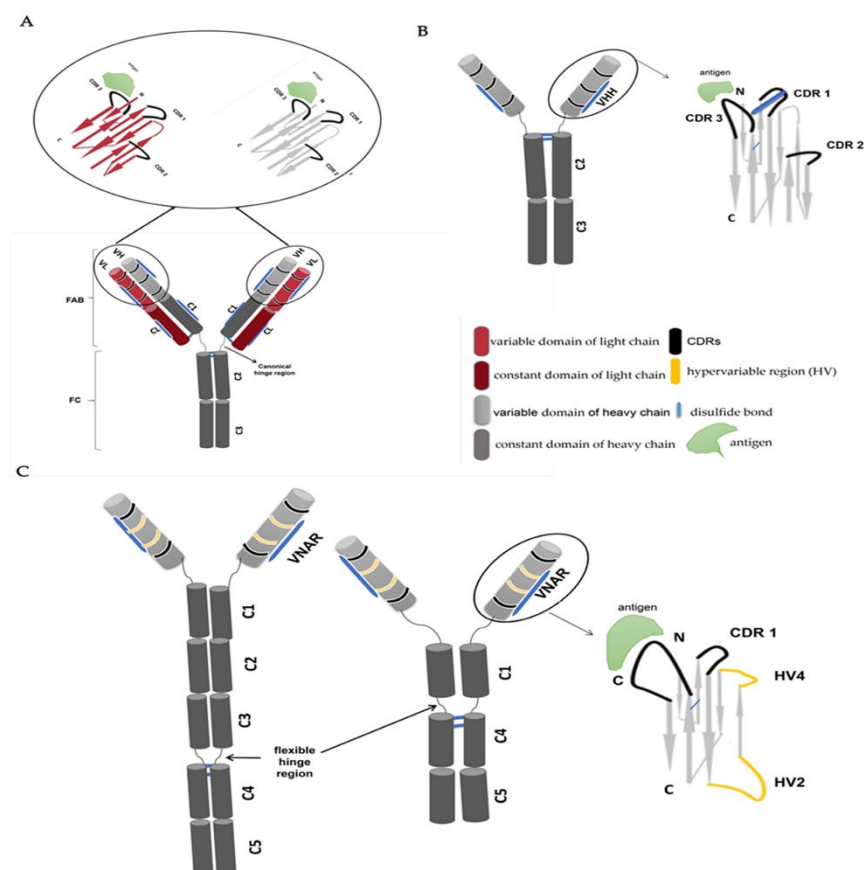


Рисунок 1. Структуры антител (<https://www.mdpi.com/2073-4409/10/5/1140>)

А - структура антитела IgG, состоящего из двух тяжелых и двух легких цепей. В - структура антитела животных семейства Верблюдовые. С - структура молекулы IgNAR акул, которая состоит только из тяжелых цепей с пятью константными доменами, альтернативная форма молекулы IgNAR, подвергшаяся сплайсингу, содержит три константных домена (C5, C4, C1). Вариабельный домен IgNAR (VNAR) включает четыре гипервариабельных петли: CDR1, CDR3, HV2 и HV4

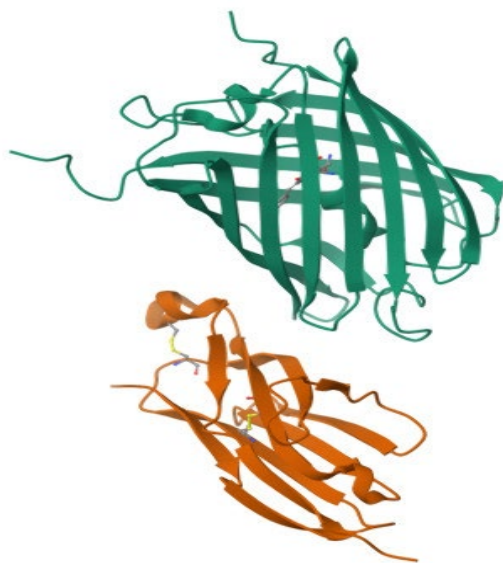


Рисунок 2. Crystal structure of VNAR aGFP14 in complex with GFP (<https://www.rcsb.org/structure/8HGI>)

1.2. Библиотеки нанотел

Отбор нанотел с желаемыми свойствами и специфичностью можно производить с помощью создания библиотек нанотел. Существует несколько типов библиотек в зависимости от метода их конструирования и происхождения: натуральные и синтетические. Натуральные библиотеки нанотел подразделяются на иммунные и наивные, они конструируются на основе мРНК, которую выделяют из периферической крови, селезенки или костного мозга [10].

1) Иммунные библиотеки

Данные библиотеки конструируются на основе мРНК периферических лимфоцитов животных или людей, иммунизированных представляющим интерес антигеном [11]. После проведения иммунизации из крови получают фракцию В-лимфоцитов и амплифицируют содержащиеся в них последовательности кДНК и мРНК, которые кодируют переменные домены тяжелых цепей (VНН) [11]. После амплификации последовательностей методом ПЦР производится их лигирование в вектор и трансформация для осуществления экспрессии, обычно для процесса трансформации используют клетки *E.coli* [11].

2) Наивные библиотеки

Для создания данного вида библиотек используется лимфоцитарная фракция неиммунизированных животных [12]. Такой метод позволяет получить широкий спектр различных нанотел, обладающих специфичностью ко множеству антигенов [12].

3) Синтетические библиотеки

Эти библиотеки создаются с применением мутагенеза путем встраивания искусственно синтезированных молекул ДНК в последовательности, кодирующие переменные домены [13]. С помощью встраивания ДНК производится изменение переменных участков, что увеличивает разнообразие библиотеки за счет увеличения повышения количества переменных участков. Синтетические библиотеки позволяют работать с токсичными или неиммуногенными молекулами (не вызывающими иммунный ответ из-за малой молекулярной массы) [13].

Для эффективной работы с библиотеками используются дисплейные технологии, которые требуются для отбора высокоспецифичных молекул к нужному антигену из множества последовательностей.

1.3. Дисплейные технологии

Для эффективного отбора наиболее аффинных молекул VNAR, связывающихся с целевыми антигенами, используются дисплейные технологии. Они позволяют осуществлять скрининг больших библиотек молекул, из которых можно выделить молекулы с желаемыми характеристиками и свойствами [11]. За последние десятилетия были разработаны различные методы в связи с быстрым развитием молекулярной биологии и протеомики. Развитие этих дисциплин и связанных с ними методов позволило создать библиотеки антител, полипептидов и нуклеиновых кислот, отличающиеся широким разнообразием. Методы отображения в сочетании с библиотеками предоставили возможности для выделения специфических молекул с высокой эффективностью действия. Эти методы отображения в основном делятся на две категории: методы на основе клеток и бесклеточные методы. Универсальный принцип дисплейных технологий заключается в том, что они создают связь между фенотипом (отображаемой связывающей молекулой) и генотипом (соответствующей генетической информацией, кодирующей эту связывающую молекулу) [14].

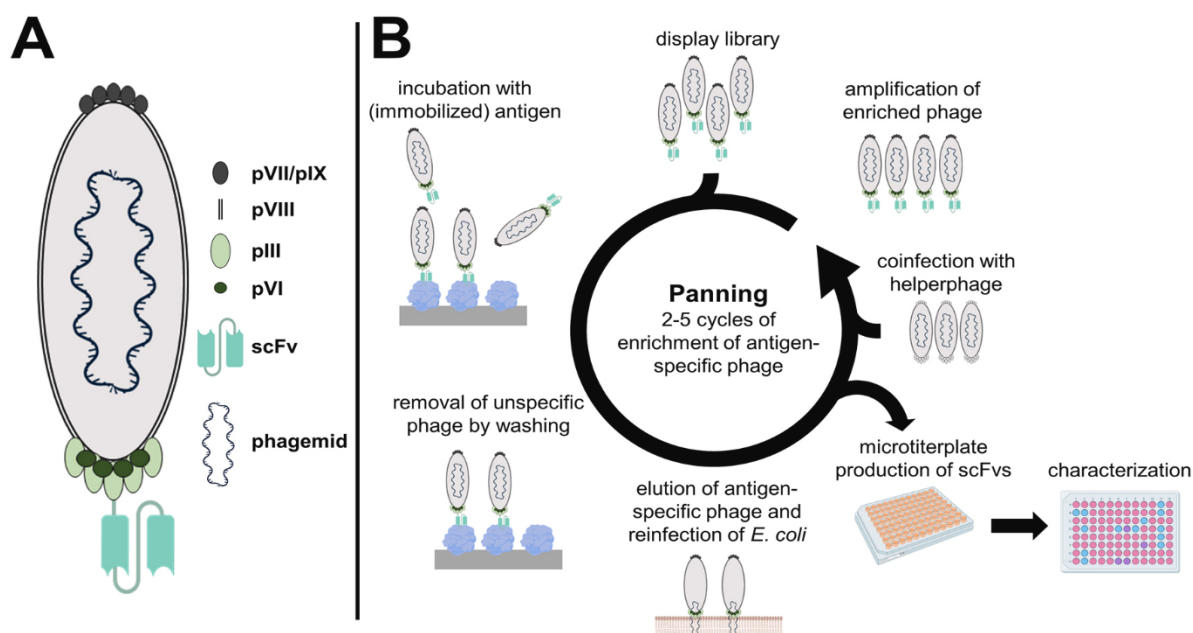


Рисунок 3. А - изображение бактериофага. pVII, pIX, pVIII, pIII, pVI - белки, образующие капсид фага. В - метод фагового дисплея. Созданная библиотека инкубируется с иммобилизованным антигеном, затем происходит отмывка фагов, несущих неспецифические к антигену лиганды, и элюция антиген-специфичных фагов. Для обогащения библиотеки происходит заражение клеток кишечной палочки отобранными фагами. (<https://www.degruyter.com/document/doi/10.1515/hsz-2020-0377/html>)

Фаговый дисплей является ведущей технологией для получения рекомбинантных молекул, которые могут найти как терапевтическое применение, так и служить объектом для изучения. Изначально в основу принципа работы фагового дисплея легло экспрессирование фрагментов эндонуклеазы рестрикции *EcoRI* на поверхности капсида нитчатого бактериофага M13 наряду с другими белками, входящими в состав вирусной оболочки [15]. Для этого было произведено встраивание генов, кодирующих фрагменты эндонуклеаз, в геном бактериофага в единую рамку считывания с белком g3P оболочки бактериофага M13, с помощью этого стало возможным отображение экспонированного на поверхность фага полипептида через встраивание соответствующего гена в вирусную частицу, что позволяет отбирать различные варианты белков по отображающимся с помощью фагового капсида свойствам [15].

Фаговый дисплей позволяет проводить анализ множества фагов, несущих различные варианты белков на поверхности, одновременно, большой пул фаговых частиц позволяет эффективно отбирать образцы, наиболее специфичные по отношению к нужному антигену. Это происходит в ходе этапа биопэннинга-процедуры аффинного обогащения, которая позволяет отбирать фаговые частицы с последовательностями, обладающими сродством к определенной мишени [16]. Сначала проводится инкубация фагмидных частиц библиотеки с иммобилизованным антигеном, фаги, несущие молекулы, способные связываться с мишенью, задерживаются на поверхности, а фаговые частицы, не связавшиеся с антигеном, отмываются [15]. После этого связавшиеся фаги элюируются, ими заражают клетки бактерий-хозяев, в которых происходит увеличение количества вирусных частиц [17]. Таким образом получают более крупную фаговую библиотеку, которая содержит меньшее количество неспецифических молекул. Повторяя циклы биопэннинга (скрининга), элюции и амплификации несколько раз, удастся провести селекцию фагов, несущих молекулы с высокой аффинностью к мишени. Из отобранных вирусных частиц извлекают ДНК, с помощью метода секвенирования

определяют нуклеотидные последовательности целевых молекул, которые клонируются в бактериальные или иные векторы для получения и наработки белковых молекул [17].

Скрининг молекул возможно осуществлять с помощью технологии рибосомного дисплея. Рибосомный дисплей (Ribosome Display) – это технология селекции белков, которая позволяет идентифицировать и выделять белки с желаемыми свойствами из огромных библиотек. Рибосомный дисплей не требует проведения клеточной трансформации, что значительно ускоряет процесс и позволяет работать с более широким спектром белков, включая токсичные и нестабильные [18]. Рибосомный дисплей позволяет работать с очень большими библиотеками (до 10^{13} вариантов), что повышает вероятность обнаружения особо специфичных молекул с желаемыми свойствами [19]. Этот метод основан на физической связи между мРНК, рибосомой и синтезируемым полипептидом, что позволяет проводить селекцию непосредственно на уровне мРНК [14]. Технология дисплея заключается в трансляции *in vitro*, при этом предотвращается выход вновь синтезированного белка и кодирующей его мРНК из рибосомы [20].

Постановка рибосомного дисплея включает следующие этапы.

1. Создание библиотеки последовательностей мРНК, кодирующих разнообразные варианты целевого белка. Библиотеку мРНК транслируют *in vitro* (например, на основе лизата *E. coli*) в полипептид [20]. Образуется комплекс, состоящий из синтезированных белков, рибосомы и мРНК, в котором мРНК связана с рибосомой, а рибосома удерживает новосинтезированный полипептид. В нормальных условиях трансляционный комплекс существует короткое время, высвобождение полипептида и мРНК обеспечивается факторами высвобождения (RF) и факторами рециркуляции рибосом (RRF) [21]. RF1 и RF2 относятся к факторам высвобождения первого класса, RF3 относится к факторам высвобождения второго класса, работа данного фактора обеспечивает отделение факторов первого класса от рибосомы. При проведении рибосомного дисплея используется мРНК, не содержащая в своем составе стоп-кодон, что гарантирует отсутствие высвобождения синтезированного полипептида и кодирующей его мРНК из рибосомы [18]. Данный комплекс может дополнительно стабилизироваться охлаждением раствора и добавлением ионов Mg^{2+} [20]. Считается, что высокое содержание ионов магния «конденсирует» рибосому, связываясь с рРНК и затрудняя диссоциацию пептидил-тРНК [20].

2. Проведение биопаннинга. Полученный комплекс инкубируется с иммобилизованным поверхностно-связанным лигандом [14]. Комплексы, содержащие наиболее специфичные к лигандам пептиды, иммобилизуются [14].

3. Элюция и обогащение. Несвязанные комплексы смываются, а комплексы, специфически связавшиеся с лигандом, элюируются с помощью хелатирующих агентов, высоких концентраций солей или мобильных лигандов, которые образуют комплекс со связывающей частью белка и позволяют мРНК диссоциировать [20]. Извлеченная мРНК подвергается обратной транскрипции и амплификации с помощью ПЦР, что позволяет сохранить генотип для следующего цикла селекции или для клонирования и секвенирования [22].

Для отбора молекул может также использоваться технология дрожжевого дисплея, которая по своему принципу схожа с методом фагового дисплея. В основе дрожжевого дисплея лежит генетическая инженерия клеток дрожжей, в основном таких видов, как *Saccharomyces cerevisiae*, *Pichia pastoris* [23]. В геном клеток встраивается генетическая конструкция, кодирующая целевой белок, который соединен с якорным белком. В качестве якорных белков используются белки клеточной стенки дрожжей, такие как Aga2p, Flo1p, Cwp2p или Sed1p [23]. Якорный белок обеспечивает связывание слитого белка с клеточной стенкой, представляя целевой белок на клеточной поверхности. Целевой белок, представленный на поверхности дрожжевой клетки, может взаимодействовать с лигандами, иммобилизованными на твердой подложке или

находящимися в растворе. Отобранные клетки культивируются, и процесс может быть повторен для обогащения клонов с желаемыми свойствами [18].

1.4. Метод секвенирования

Секвенирование— это процесс определения точного порядка нуклеотидов в молекуле ДНК или РНК [24]. С помощью данного метода происходит расшифровка генетических конструкций, что открывает широкий спектр применений в медицине и биотехнологии [24]. Существует большое число способов определения последовательностей молекул нуклеиновых кислот, однако мы рассмотрим два следующих метода: NGS (Next Generation Sequencing) и RACE (Rapid amplification of cDNA ends).

NGS-секвенирование (Next Generation Sequencing)

С помощью данного метода возможно одновременное проведение секвенирования большого количества молекул ДНК или РНК. Вплоть до миллионов последовательностей могут обрабатываться параллельно, что снижает время и стоимость, которые требуются для секвенирования транскриптомов [25].

RACE-секвенирование (Rapid amplification of cDNA ends)

Метод, используемый для определения полных последовательностей мРНК с помощью известных 5' или 3' концов цепи нуклеиновой кислоты. RACE бывает двух основных типов: 5'-RACE и 3'-RACE, в зависимости от того, последовательность какого конца необходимо исследовать [26] (таблица 1)

Таблица 1 - Сравнение методов секвенирования

Характеристика	RACE-секвенирование	NGS-секвенирование
Цель	Определение последовательности 5' или 3' концов РНК	Параллельное секвенирование большого количества молекул ДНК или РНК
Объем данных	Ограниченный, секвенируются конкретные концы РНК	Могут секвенироваться миллионы последовательностей
Применение	Изучение нетранслируемых областей последовательности РНК, получение информации о концах мРНК	Метагеномное секвенирование, геномное секвенирование, секвенирование регуляторных и кодирующих молекул РНК
Объект	Применяется для получения полной длины последовательности транскрипта РНК	Широкий спектр применения, используется для секвенирования как последовательностей ДНК, так и последовательностей РНК

1.5. Биоинформатические методы

В последние годы технологии секвенирования стали важным инструментом в молекулярной биологии и биоинформатике. Важной областью применения секвенирования является анализ библиотек последовательностей vNAR (variable New Antigen Receptor), которые представляют собой уникальные фрагменты антител, полученные из организма хрящевых рыб. Эти молекулы характеризуются высокой специфичностью и стабильностью, что делает их перспективными для использования в сфере биомедицины.

С учетом большого объема данных, получаемых в результате секвенирования последовательностей vNAR, важным этапом является их биоинформатическая обработка. Одним из мощных инструментов для анализа и визуализации данных является программа Unipro UGENE. UGENE (Universal Genomics Nucleotide Edition) — это многофункциональная биоинформатическая программа, предназначенная для анализа

данных геномики и протеомики, которая широко применяется в различных областях биологии [27]. UGENE предоставляет широкий спектр инструментов для визуализации, аннотирования, сборки, выравнивания и анализа последовательностей ДНК, РНК и белков, а также для работы с данными из разных источников, включая геномные последовательности и метагеномные данные [28]. Основные преимущества UGENE заключаются в его понятном интерфейсе, поддержке различных форматов данных, включая FASTA, FASTQ, SAM/BAM и GFF [27], что позволяет работать с различными биологическими данными. Данная программа позволяет пользователям эффективно обрабатывать большие объемы данных, что особенно актуально для анализа библиотек vNAR, которые характеризуются сложностью в обработке из-за высокого разнообразия последовательностей, что вызывает определенные трудности у исследователей.

1.6. Подготовка и конструкция VNAR библиотек для дисплея

Для создания иммунных библиотек требуется проведение иммунизации акул интересующим антигеном, в ответ на который иммунной системой будут вырабатываться VHH-антитела. При проведении иммунизации важно убедиться, что выбранный антиген не является токсичным для акул. Успешность проведенной иммунизации можно проверить титром антител против антигена в сыворотке крови с помощью ELISA-метода [17]. На успешную иммунизацию указывает высокий титр антител. Для получения VNAR производится забор В-лимфоцитов из периферической крови акул. Плазму крови центрифугируют для выделения молекул РНК, которые подвергаются обратной транскрипции с использованием праймеров, специфичных для последовательности VNAR, и обратной транскриптазы [29]. С помощью метода ПЦР производится амплификация последовательностей переменных доменов VNAR. Продукты ПЦР подвергаются электрофорезу в агарозном геле для разделения на фрагменты. С помощью ДНК-лигазы полученные фрагменты подвергаются лигированию. Далее продукты лигирования очищаются от ферментов и трансформируются в компетентные клетки для создания библиотеки.

2. Материалы и методы исследования

Материалом для создания иммунной библиотеки являлась лимфоцитарная фракция, полученная из крови иммунизированного животного.

Далее представим методы и технологии, применяемые для решения конкретных задач.

1) Для получения последовательностей нуклеотидов в молекулах мРНК, выделенных из крови иммунизированных животных, использовалась технология RACE-секвенирования, выбор данного метода был обусловлен наличием последовательности лишь одного конца нуклеотида, которая начиналась с константного домена.

Существуют разные способы проведения RACE-секвенирования:

5'-RACE:

Позволяет получить последовательность 5' конца мРНК, включая 5'-нетранслируемую область. На матрице мРНК синтезируется кодирующая ДНК (кДНК) в ходе реакции обратной транскрипции, для этого используются специфические праймеры, комплементарные известной последовательности мРНК. Далее исходная мРНК-матрица удаляется с помощью РНКазы Н. Фермент терминальная дезоксирибонуклеотидилтрансфераза (TdT) добавляет поли(А)-хвост к 3' концу кДНК [30]. Амплификация кДНК с помощью метода ПЦР происходит с использованием олиго-dT праймера и специфического праймера, комплементарного известной последовательности в мРНК [30]. Для определения последовательности 5' конца производится секвенирование продукта ПЦР. Для получения секвенсов в работе был использован данный вид секвенирования.

3'-RACE:

Используется для определения последовательности 3' конца мРНК, включая поли(А)-хвост и 3'-нетранслируемую область (3'-UTR). С помощью олиго-dT адаптерного праймера и обратной транскриптазы происходит синтез цепи кДНК. Адаптерный праймер содержит поли(Т)-участок для связывания с поли(А)-хвостом мРНК. Для проведения ПЦР с использованием кДНК в качестве матрицы используются следующие праймеры: праймер, комплементарный адаптерной последовательности, и праймер, комплементарный известной последовательности в мРНК. Таким образом возможно проведение амплификации 3' конца кДНК. Далее продукт ПЦР секвенируется с целью определения последовательности 3' конца мРНК.

2) Для работы с большим числом последовательностей, полученных в ходе секвенирования, была проведена их биоинформатическая обработка в программе Unipro UGENE.

1. Инструмент BLAST

BLAST (Basic Local Alignment Search Tool) позволяет находить участки сходства между последовательностями путем сравнения заданных последовательностей нуклеотидов или белков с последовательностями, находящимися в базах данных [31].

2. Инструмент UGENE Genome Aligner

Данный инструмент позволяет выравнивать (сопоставлять нуклеотидные или аминокислотные последовательности для достижения максимального сходства между друг другом) последовательности ДНК различной длины на референсную (эталонную) последовательность с настраиваемой частотой несовпадений [32].

3. Инструмент CAP3

С помощью данного инструмента возможна сборка последовательностей без использования референсного генома. На основе фрагментов, полученных в результате секвенирования, создается отдельная последовательность. В данном исследовании это было необходимо для осуществления дальнейшего поиска референса, наиболее схожего с полученным нами консенсусом [33].

4. Инструмент Trim and Map Sanger Reads

Данный инструмент позволяет считывать наборы последовательностей, полученных в результате секвенирования, обрезать концы прочтений и осуществлять их картирование на референсную последовательность [34]. В данной работе этот инструмент был применен для осуществления множественного выравнивания последовательностей и картирования библиотек.

Картирование- это процесс, в ходе которого определяется соответствие нуклеотидных последовательностей частям эталонного генома [35]. С помощью картирования имеющиеся прочтения ставятся в соответствие определенным участкам в геноме [35].

5. Множественное выравнивание с использованием MUSCLE

С помощью данного инструмента происходит множественное выравнивание большого числа последовательностей между собой без использования референса [36].

3. Практическая часть

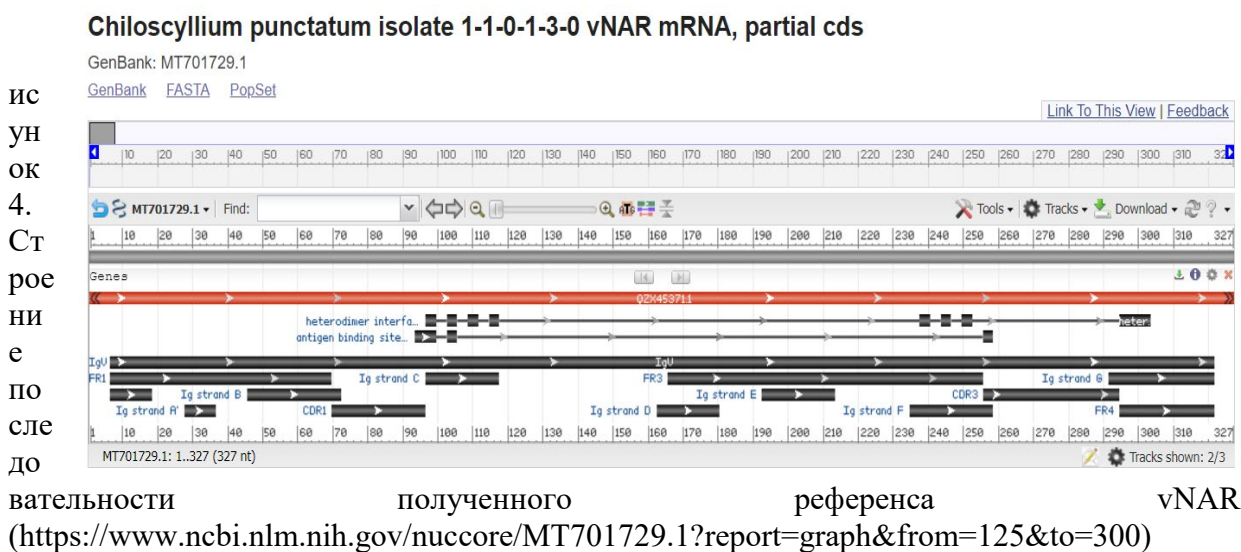
3.1 Этапы исследования

1. Для создания библиотеки последовательностей vNAR проводилась иммунизация кошачьих акул RBD-доменом SARS-CoV-2 Omicron. Процесс иммунизации проходил в 4 раунда в центре океанографии и морской биологии «Москвариум». После проведенной иммунизации из крови акул была получена лимфоцитарная фракция, из которой выделили тотальную РНК. Работы с кровью производились специалистами Роспотребнадзора в лаборатории института системной биологии и медицины Роспотребнадзора. В ходе проведения обратной транскрипции с помощью набора Mint kit по протоколу Mint [37] из РНК была синтезирована кДНК.

2. На синтезированной кДНК были выбраны участки, по которым проводилось RACE-секвенирование с контантного домена CH1 молекул IgNAR [38] специалистами биотехнологической компании «Евроген», таким образом были определены полные последовательности мРНК [39].

3. Полученные последовательности кДНК были загружены в программу UGENE для проведения глобального выравнивания с помощью сборки прочтений de novo. В результате множественного выравнивания был получен пробный консенсус. Полученный консенсус далее анализировали с использованием базы данных нуклеотидных последовательностей Национального центра биотехнологической информации (NCBI, США) с помощью инструмента BLAST в программе UGENE. В результате анализа был найден наиболее валидный референс vNAR *Chiloscyllium punctatum* isolate 1-1-0-1-3-0 vNAR mRNA, partial cds [40].

Через Workflow Designer в UGENE был сконструирован путь, по которому последовательности экспортировались, далее проводилось множественное выравнивание без использования референса с помощью MUSCLE. Далее было проведено картирование полученной библиотеки на референсную последовательность, найденную ранее с помощью инструмента BLAST. На основе данных, расположенных на сайте Национального центра биотехнологической информации (NCBI, США) [41] были определены положения участков CDR1, FR1, FR3, CDR3, FR4 на референсной последовательности.



С помощью найденных положений CDR1, FR1, FR3, CDR3, FR4 было произведено картирование библиотеки, полученной в ходе множественного выравнивания.

По аннотированным у выбранного референса переменным участкам, наиболее схожих с участками в библиотеке, которые являются началом переменного домена, была произведена обрезка библиотеки для удаления константных доменов из последовательностей ввиду того, что они не были важны для поиска константных участков. Далее было проведено повторное глобальное выравнивание полученных библиотек с целью минимизации возникновения ошибок выравнивания по переменному домену. В ходе выравнивания были выявлены наиболее различающиеся между собой последовательности. Данные последовательности были разделены на подгруппы, исходя из их нуклеотидного состава.



Рисунок 5.

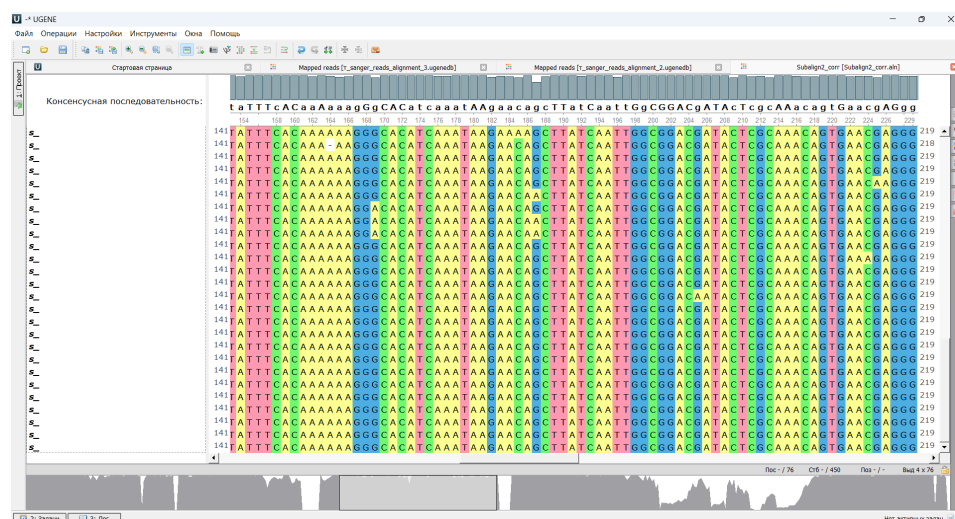


Рисунок 6.

Сверху представлены снимки экрана, сделанные при работе в ранее упомянутой программе UGENE. Производилось множественное выравнивание последовательностей с использованием MUSCLE, в результате чего были выделены две подгруппы на основании наиболее различающихся между собой участков

4. После картирования стали известны места нахождения FR1, FR3 и FR4 - константных участков библиотеки vNAR, которые использовались для последующего подбора праймеров. С помощью BLAST были определены аминокислотные последовательности участков FR1, FR3 и FR4, которые сопоставлялись с аминокислотными последовательностями нашей библиотеки.

Аминокислотная последовательность FR1 в библиотеке: IEQTPRTATKEAGESLTINCV

Аминокислотная последовательность FR3 в библиотеке:
SYEKGAGTVLTETVNSGSKSFSLRINDLTVEDGGTYRCGVC
Аминокислотная последовательность FR4 в библиотеке: YYEGSGTTTLTV

С помощью найденных последовательностей были подобраны праймеры:

5'- AGA-AGG-CTG-TAT-GCC-AGG - 3' - праймер CH1revPCR

5' - GCC-TGG-TTA-CCA-AAT-GTC-TTC-ACT - 3' - праймер Leader 1F

5' - GCC-TGG-TTA-CCA-AAT-GTC-TTC-ACW - 3' - вырожденный праймер Leader 1F, который представляет собой смесь олигонуклеотидов, кодирующих одинаковую аминокислотную последовательность. Данный праймер состоит из двух, имеющих идентичную нуклеотидную последовательность, за исключением последней позиции в последовательности, которая соответствует как адениновому, так и тимидиновому нуклеотиду. Несмотря на высокую консервативность константных участков библиотеки, они могут отличаться друг от друга на один или несколько нуклеотидов в различных позициях, что объясняет необходимость создания вырожденного праймера, который будет комплементарен одновременно нескольким последовательностям.

Обсуждение результатов

В результате исследования, включающего биоинформатическую работу, был проведен анализ библиотеки сиквенсов, полученных в результате 5'-RACE секвенирования. С использованием биоинформатических инструментов обработки стали известны нуклеотидные последовательности константных FR-участков, на основе которых были подобраны праймеры для клонирования молекул в фагмиды, в том числе был найден вырожденный праймер, который представляет собой смесь двух праймеров, различающихся одной нуклеотидной позицией в конце последовательности. Создание такого праймера позволит сохранить генетическое разнообразие библиотеки нанотел за счет его специфичности одновременно к нескольким разным последовательностям. Дальнейшей задачей исследования является постановка фагового дисплея, в результате чего станет возможен отбор наиболее специфичных нанотел, обладающих наибольшей аффинностью к требуемому антигену. Это позволит выявить нанотела, которые в будущем могут быть использованы исследователями в качестве основы для создания эффективных вакцин.

ЗАКЛЮЧЕНИЕ

Проведение исследований в области библиотек нанотел, специфичных к SARS-CoV-2 Omicron, может иметь значительное влияние на будущие разработки в области вакцин и терапий, так как библиотеки позволяют выбирать из множества представленных в них молекул те, которые будут обладать наиболее нужными свойствами, за счет чего будет обеспечиваться их эффективная борьба с антигеном. В ходе исследования была выявлена последовательность применения методов и технологий, а также подтверждена выдвинутая гипотеза: возможно выявить константные участки библиотеки кДНК, полученной в результате иммунизации акул, которая после подбора праймеров будет использоваться для успешного клонирования последовательностей в фагмиды и проведения фагового дисплея.

В работе использовались следующие методы: RACE-секвенирование, инструменты биоинформатической программы UGENE, такие как сборка прочтений *de novo*, выравнивание последовательностей, осуществление поиска с помощью инструмента BLAST, UGENE Workflow Designer.

Работа включала в себя биоинформатическую обработку большого объема данных, полученных в результате RACE-секвенирования кДНК, анализ последовательностей производился в биоинформатической программе UGENE с использованием базы данных Национального центра биотехнологической информации (NCBI, США). Для дальнейших манипуляций с полученной библиотекой кДНК, исходя из последовательностей найденных константных участков библиотеки, были подобраны праймеры. Данный метод в перспективе может использоваться для создания препарата против SARS-CoV-2 широкого действия на линию Omicron.

ЛИТЕРАТУРА

1. Министерство здравоохранения РФ. Временные методические рекомендации. //Профилактика, диагностика и лечение новой коронавирусной инфекции (COVID-19). - 2023.
2. Chen K. et al. Analysis of the protective efficacy of approved COVID-19 vaccines against Omicron variants and the prospects for universal vaccines // Front Immunol. Frontiers Media SA, 2023. Vol. 14. P. 1294288.
3. Salvador J.P., Vilaplana L., Marco M.P. Nanobody: outstanding features for diagnostic and therapeutic applications // Anal Bioanal Chem. Springer Verlag, 2019. Vol. 411, № 9. P. 1703–1713.
4. Alekseevich A.A. et al. “CAMEL” ANTIBODIES: FEATURES OF THEIR STRUCTURE, PREPARATION AND APPLICATION // Siberian Journal of Life Sciences and Agriculture. Science and Innovation Center Publishing House, 2020. Vol. 12, № 2. P. 25–40.
5. Bannas P., Hambach J., Koch-Nolte F. Nanobodies and nanobody-based human heavy chain antibodies as antitumor therapeutics // Front Immunol. Frontiers Media S.A., 2017. Vol. 8, № NOV. P. 309808.
6. Alexander E., Leong K.W. Discovery of nanobodies: a comprehensive review of their applications and potential over the past five years // Journal of Nanobiotechnology 2024 22:1. BioMed Central, 2024. Vol. 22, № 1. P. 1–36.
7. Mei Y. et al. Research progress and applications of nanobody in human infectious diseases // Front Pharmacol. Frontiers Media S.A., 2022. Vol. 13. P. 963978.
8. Ho M. et al. Construction and next-generation sequencing analysis of a large phage-displayed V NAR single-domain antibody library from six naïve nurse sharks // Antib Ther. 2019. Vol. 2, № 1. P. 1–11.
9. Juma S.N. et al. Shark New Antigen Receptor (IgNAR): Structure, Characteristics and Potential Biomedical Applications // Cells 2021, Vol. 10, Page 1140. Multidisciplinary Digital Publishing Institute, 2021. Vol. 10, № 5. P. 1140.
10. Alexander E., Leong K.W. Discovery of nanobodies: a comprehensive review of their applications and potential over the past five years // Journal of Nanobiotechnology 2024 22:1. BioMed Central, 2024. Vol. 22, № 1. P. 1–36.
11. Romão E. et al. Construction of high-quality camel immune antibody libraries // Methods in Molecular Biology. Humana Press Inc., 2018. Vol. 1701. P. 169–187.
12. Hust M. et al. Construction of human naïve antibody gene libraries // Methods in Molecular Biology. Humana Press Inc., 2012. Vol. 907. P. 85–107.
13. Creating a Nanobody Library with PCR. URL: <https://www.azolifesciences.com/article/Creating-a-Nanobody-Library-with-PCR.aspx> (дата обращения: 09.02.2025).
14. Zahnd C., Amstutz P., Plückthun A. Ribosome display: selecting and evolving proteins in vitro that specifically bind to a target // Nature Methods 2007 4:3. Nature Publishing Group, 2007. Vol. 4, № 3. P. 269–279.
15. Vikhrova M.A. et al. Combinatorial libraries of phage antibodies: Application for analysis of gene segments encoding variable domains of autoantibodies to tumor necrosis factor // Mol Biol. 2011. Vol. 45, № 1. P. 82–92.
16. Фаговый дисплей на основе нитчатых бактериофагов: применение для отбора рекомбинантных антител // Акта Naturae (русскаяязычная версия). Общество с ограниченной ответственностью Парк-медиа, 2009. Vol. 1, № 3.
17. Feng M. et al. Construction and next-generation sequencing analysis of a large phage-displayed VNAR single-domain antibody library from six naïve nurse sharks // Antib Ther. Oxford Academic, 2019. Vol. 2, № 1. P. 1–11.
18. Valldorf B. et al. Antibody display technologies: selecting the cream of the crop // Biol Chem. De Gruyter Open Ltd, 2022. Vol. 403, № 5. P. 455–477.

19. Lipovsek D., Plückthun A. In-vitro protein evolution by ribosome display and mRNA display // J Immunol Methods. Elsevier, 2004. Vol. 290, № 1–2. P. 51–67.
20. Plückthun A. Ribosome display: A perspective // Methods in Molecular Biology. Humana Press Inc., 2012. Vol. 805. P. 13–28.
21. Светлов М.С. Контроль качества у бактерий, ассоциированный с рибосомами // Биохимия. The Russian Academy of Sciences, 2021. Vol. 86, № 8. P. 1138–1148.
22. Hanes J. et al. Ribosome display efficiently selects and evolves high-affinity antibodies in vitro from immune libraries // Proc Natl Acad Sci U S A. The National Academy of Sciences, 1998. Vol. 95, № 24. P. 14130–14135.
23. Teymennet-Ramírez K. V., Martínez-Morales F., Trejo-Hernández M.R. Yeast Surface Display System: Strategies for Improvement and Biotechnological Applications // Front Bioeng Biotechnol. Frontiers Media S.A., 2022. Vol. 9. P. 794742.
24. Что такое секвенирование? URL: <https://ngseq.ru/news/chto-takoe-sekvenirovanie> (дата обращения: 09.02.2025).
25. Next-Generation Sequencing (NGS)- Definition, Types URL: <https://microbenotes.com/next-generation-sequencing-ngs/> (дата обращения: 09.02.2025).
26. Olivarius S., Plessy C., Carninci P. High-throughput verification of transcriptional starting sites by Deep-RACE.
27. Okonechnikov K. et al. Unipro UGENE: a unified bioinformatics toolkit // Bioinformatics. Oxford Academic, 2012. Vol. 28, № 8. P. 1166–1167.
28. Rose R. et al. Flexible design of multiple metagenomics classification pipelines with UGENE // Bioinformatics. Oxford Academic, 2019. Vol. 35, № 11. P. 1963–1965.
29. Casonato S. et al. Abstract 1675: Parallel genome and transcriptome-wide profiling of PD-L1 expressing tumor and infiltrating immune single cells in NSCLC // Cancer Res. American Association for Cancer Research (AACR), 2017. Vol. 77, № 13_Supplement. P. 1675–1675.
30. When PCR Get RACE-y: Sequencing Unknown mRNA URL: <https://bitesizebio.com/27087/when-pcr-get-race-y-from-unknown-mrna-segments-to-sequenced-cdna/> (дата обращения: 09.02.2025).
31. BLAST - Unipro UGENE User Manual URL: <https://doc.ugene.net/wiki/display/UM/BLAST> (дата обращения: 09.02.2025).
32. UGENE Genome Aligner - Unipro UGENE User Manual URL: <https://doc.ugene.net/wiki/display/UM/UGENE+Genome+Aligner> (дата обращения: 09.02.2025).
33. CAP3 - Unipro UGENE User Manual URL: <https://doc.ugene.net/wiki/display/UM/CAP3> (дата обращения: 09.02.2025).
34. Trim and Map Sanger Reads - Unipro UGENE User Manual. URL: <https://doc.ugene.net/wiki/display/UM/Trim+and+Map+Sanger+Reads> (дата обращения: 09.02.2025).
35. Hands-on: Mapping / Mapping / Sequence analysis. URL: <https://training.galaxyproject.org/training-material/topics/sequence-analysis/tutorials/mapping/tutorial.html> (дата обращения: 09.02.2025).
36. Align sequences with MUSCLE - Workflow Designer Documentation v. 1.14.0 - URL: <https://doc.ugene.net/wiki/display/WDD/Align+sequences+with+MUSCLE> (дата обращения: 09.02.2025).
37. Mint: обратная транскриптаза и наборы для синтеза. URL: <https://evrogen.ru/products/cdna/synthesis/mint> (дата обращения: 09.02.2025).
38. RACE - быстрая амплификация концевых фрагментов кДНК. URL: <https://evrogen.ru/products/cdna/race> (дата обращения: 09.02.2025).
39. De Silva D.P.N. et al. Transcriptomic analysis of immunoglobulin novel antigen receptor (IgNAR) heavy chain constant domains of brownbanded bamboo shark (*Chiloscyllium punctatum*) // Fish Shellfish Immunol. Fish Shellfish Immunol, 2019. Vol. 84. P. 370–376.

40. Chiloscyllium punctatum isolate 1-1-0-1-3-0 vNAR mRNA, partial cds -
Nucleotide — NCBI. URL:
[https://www.ncbi.nlm.nih.gov/nucleotide/MT701729.1?report=genbank&log\\$=nuclalign&blast_rank=1&RID=UEJWA7Y7013](https://www.ncbi.nlm.nih.gov/nucleotide/MT701729.1?report=genbank&log$=nuclalign&blast_rank=1&RID=UEJWA7Y7013) (дата обращения: 08.02.2025).
41. MT701729.1: Chiloscyllium punctatum isolate 1-1-0-1-3-0 vNAR mRNA, partial
cds . URL: <https://www.ncbi.nlm.nih.gov/nuccore/MT701729.1?report=graph> (дата
обращения: 09.02.2025).

Фамилия И.О. Тимофеева
 Город Москва
 Школа СГ МГУ

Шифр

Шифр

ЛИСТ ОТВЕТОВ

на задания теоретического тура олимпиады школьников

г. Саров 2025 г.

Внимание! Образец заполнения матрицы:

Правильный ответ	×	Отмена ответа	⊗	Отмена исправления	■
------------------	---	---------------	---	--------------------	---

Задание А. Один правильный ответ (максимально 60 баллов)

№	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
а																			×	×
б			×	×				×	×			×								
в	×				×	×					×		×		×	×	×	×		
г		×					×			×				×						

№	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
а				×		×				×									×	
б	×											×	×			×	×	×		×
в					×		×		×		×			×	×					
г		×	×					×												

№	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
а	×		×			×	×						×			×				
б		×								×		×					×			×
в					×			×	×		×			×	×			×		
г				×															×	

50

Задание Б. Множественные ответы (максимально 30 баллов)

№	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
а	×		×		×							×		×		×	×	×		×
б		×	×	×	×	×	×		×	×	×	×							×	⊗
в		×	×	×					×	×						×	×	×	×	
г	×	×			×	×	×	×			×		×	×	×	×				
д	×					×	×	×	×			×	×	×	×			×		×
е				×				×		×	×		×		×		×		×	×

26,5

Задание В. Суждения (максимально 10 баллов)

№	1	2	3	4	5	6	7	8	9	10
ДА			×	×		⊗	×			
НЕТ	×	×			×	×		×	×	×

8

Максимальное количество баллов – 100

ИТОГО: 84,5