

Шагин Дмитрий Алексеевич

**Термостабильная дезоксирибонуклеаза из
Paralithodes camtschaticus –
новый инструмент исследования сложных геномов**

Специальность 1.5.3 – «молекулярная биология»

диссертация на соискание ученой степени
доктора биологических наук

Научный консультант
доктор биологических наук
академик РАН
Лукьянов Сергей Анатольевич

ОГЛАВЛЕНИЕ

ОГЛАВЛЕНИЕ.....	2
ВВЕДЕНИЕ.....	5
ОСНОВНАЯ ЧАСТЬ.....	15
Глава 1. Обзор литературы.....	15
1.1. Общая характеристика и разнообразие ферментов, расщепляющих нуклеиновые кислоты.....	15
1.1.1 Реакции, катализируемые нуклеазами.....	16
1.1.2 Экзо- и эндонуклеазы.....	20
1.1.3 Субстратная специфичность нуклеаз.....	21
1.1.4 Механизмы катализа.....	29
1.1.5 Нуклеазы, катализирующие расщепление субстрата с участием двух или более ионов металла.....	32
1.1.5.1 DEDD-мотив содержащие (или ДнкQ-подобные) нуклеазы.....	36
1.1.5.2 LAGLIDADG-содержащие хоуминг-эндонуклеазы.....	42
1.1.5.3 Суперсемейство PD-(D/E)XK.....	45
1.1.5.4 FEN-подобные нуклеазы.....	46
1.1.6 Нуклеазы, катализирующие расщепление субстрата с участием одного иона металла.....	47
1.1.6.1 Нуклеазы His-Me.....	48
1.1.6.2 HUH эндонуклеазы (релаксазы).....	53
1.1.6.3 Хоуминг-эндонуклеазы, содержащие His-Cys мотив.....	55
1.1.6.4 ДНКазы I подобные нуклеазы.....	55
1.1.6.5 GIY-YIG эндонуклеазы.....	56
1.1.7 Суперсемейство TOPRIM — нуклеазы, катализирующие расщепление нуклеиновых кислот с участием одного или двух ионов металла.....	57
1.1.8 Нуклеазы, осуществляющие катализ, не зависящий от ионов металла.....	60
1.1.8.1 Нуклеазы суперсемейства фосфолипазы D.....	61
1.1.8.2 РНКазы-колицины.....	62
1.1.8.3 Суперсемейство РНКазы А.....	63
1.1.8.4 Топоизомеразы IB.....	63
1.1.8.5 Сиквенс-специфические рекомбиназы.....	64
1.1.8.6 Другие металлонеинзависимые нуклеазы.....	66
1.1.9 Каталитические механизмы рибозимов и дезоксирибозимов.....	67

1.2. Использование нуклеаз в биотехнологии и медицине.....	70
1.3. Некоторые методы анализа нуклеиновых кислот, требующие выделения одноцепочечной ДНК.....	78
1.3.1 Нормализация кДНК и геномной ДНК.....	78
1.3.2 Удаление нецелевых транскриптов из библиотек кДНК.....	94
1.3.3 Анализ генетических полиморфизмов и выявление целевых последовательностей НК в биологических образцах.....	97
1.3.3.1 Технологии на основе кинетики гибридизации НК.....	99
1.3.3.2 Технологии на основе реакции удлинения праймера.....	102
1.3.3.3 Методы, основанные на использовании нуклеазной активности.....	106
1.3.3.4 Технологии на основе лигирования.....	108
1.3.3.5 Методы на основе анализа конформационных изменений.....	109
Глава 2. Материалы и методы.....	112
2.1. Олигонуклеотиды.....	112
2.2. Выделение и очистка НК.....	112
2.3. Секвенирование ДНК.....	113
2.4. Синтез первой цепи кДНК.....	113
2.5. Амплификация и клонирование ДНК.....	114
2.6. ПЦР в реальном времени.....	124
2.7. Сайт-направленный мутагенез.....	125
2.8. Рестрикционный анализ.....	126
2.9. Гибридизация НК с радиоактивно мечеными пробами.....	127
2.10. Получение рекомбинантных белков.....	127
2.11. Получение поликлональных антител.....	128
2.12. Приготовление ацетонового порошка из гепатопанкреаса камчатского краба.....	128
2.13. Выделение Par_DSN из камчатского краба.....	128
2.14. Определение активности Par_DSN.....	129
2.15. Исследование параметров, влияющих на активность Par_DSN.....	130
2.16. Определение субстратной специфичности.....	131
2.17. Масс-спектрометрический анализ.....	133
2.18. Анализ последовательностей НК и аминокислот.....	134
2.19. ДСНП-анализ.....	135
2.20. Статистическая обработка данных.....	137
2.21. Выявление целевых молекул РНК.....	139

2.22. ДСН-детекция.....	140
2.23. Аллель-специфическая ПЦР, совмещенная с ДСН-детекцией.....	142
2.24. ДСН-нормализация кДНК.....	142
2.25. ДСН-нормализация геномной ДНК.....	144
2.26. ДСН-деплеция.....	145
2.27. Вычитающая гибридизация.....	147
Глава 3. Результаты и обсуждение.....	148
3.1. Идентификация термостабильной нуклеазы, селективно расщепляющей двухцепочную ДНК, и её характеристика.....	148
3.2. Идентификация семейства нуклеаз, специфически расщепляющих дц ДНК, и его структурно-функциональная характеристика.....	169
3.3. Методы анализа мутаций и выявления ДНК-мишеней в сложных смесях нуклеиновых кислот.....	195
3.4. Нормализация кДНК	223
3.5. Нормализация геномной ДНК	242
3.6. Селективное удаление нецелевых последовательностей из образцов кДНК. Удаление рибосомальной РНК перед полномасштабным секвенированием транскриптомов прокариот.....	251
ЗАКЛЮЧЕНИЕ.....	266
ВЫВОДЫ.....	270
Список сокращений и условных обозначений.....	271
СПИСОК ЛИТЕРАТУРЫ.....	273

ВВЕДЕНИЕ

Актуальность темы. Развитие современной биологии и медицинской науки тесно связано с исследованиями молекулярных механизмов, лежащих в основе нормальных и патологических процессов в живых организмах, а также с изучением генетических основ, определяющих эти механизмы. Для анализа геномов и транскриптомов в последние десятилетия был создан мощный технологический аппарат, включающий высокопроизводительные методы секвенирования ДНК и РНК, методы исследования генома и транскриптома единичных клеток и методы масштабного анализа генной экспрессии на микрочипах [1-7].

Многочисленные технологические прорывы в молекулярной биологии часто связаны с открытием и внедрением в практику ферментов, обладающих новыми активностями или ранее не описанными комбинациями свойств. Так, обнаружение бактериальных эндонуклеаз рестрикции и ДНК-лигазы в 1960-70-е годы сделало возможным создание химерных нуклеиновых кислот и открыло эру технологий рекомбинантных ДНК [8-10]. Выявление термостабильной ДНК-полимеразы из бактерий *Thermophilus aquaticus* (Taq-полимераза) привело к широкому использованию в биологических науках, медицинской диагностике и криминалистике метода полимеразной цепной реакции (ПЦР) [11]. А обнаружение систем CRISPR-Cas, ставшее одним из значимых событий последних десятилетий, сделало реальностью редактирование геномов [12-16].

Нуклеазы являются группой ферментов, которые широко используются в биотехнологии и медицине. Однако среди нуклеаз, применяемых на момент инициации данной работы, отсутствовали клонированные термостабильные ферменты, способные избирательно разрушать только двухцепочечную (дц) ДНК, сохраняя одноцепочечную (оц) ДНК и РНК неповрежденными. В то же время фермент с такой активностью был высоко востребован для разработки ряда технологий анализа сложных смесей нуклеиновых кислот (НК) [17]. Вот лишь несколько примеров.

Значительная часть геномной ДНК эукариот представлена последовательностями, которые не несут смысловую нагрузку и не представляют особого интереса для исследователей. К таким последовательностям относятся повторяющиеся некодирующие элементы генома: ретроэлементы, сателлитные повторы, простые двух-, трех-, четырех нуклеотидные повторы и т.д. Повторяющиеся последовательности составляют более половины генома человека [18, 19], а в геноме кукурузы и пшеницы их количество достигает 85% и 87% соответственно [20, 21, 22]. С учетом того, что для большинства эукариотических геномов все еще стоит задача определения первичной нуклеотидной последовательности, актуальной является разработка методических подходов, позволяющих значительно снизить содержание повторяющихся последовательностей в образцах геномной ДНК перед секвенированием.

Другая проблема обусловлена широким диапазоном вариаций в концентрации различных мРНК в клетках эукариот. Так, в большинстве тканей эукариот около 85% всей мРНК является продуктом небольшого числа так называемых «мажорных» генов [23]. Для примера, в тканях сердца человека около 60% общей мРНК представлены всего пятью типами РНК, кодирующими белки митохондрий [24]. Из-за этого выявление редких и крайне редких транскриптов оказывается нетривиальной задачей, требующей либо многократного увеличения объемов секвенирования, либо применения специальных подходов для выравнивания концентраций разнопредставленных транскриптов в образце.

Еще одна проблема, стоящая перед исследователями, связана с функциональным скринингом кодирующих последовательностей, продукты которых обладают определенной активностью. Кроме разной представленности таких транскриптов, выявление целевых молекул затрудняет функциональная избыточность биологических систем, в которых одна и та же активность может быть присуща нескольким белкам [25, 26].

Эффективным решением этих проблем может быть нормализация образцов геномной ДНК/кДНК (выравнивание уровня представленности различных последовательностей) или удаление определенных заданных последовательностей из популяции исследуемой кДНК (деплегция нецелевых последовательностей кДНК). Указанные технологии требуют изолирования оц ДНК из смеси НК – задача, которая легко решается при использовании фермента, обладающего такой активностью, и трудно разрешима при его отсутствии.

Таким образом, отсутствие в широкой палитре нуклеаз, используемых в биотехнологии и молекулярной биологии, ферментов, избирательно расщепляющих дц ДНК существенно ограничивает развитие методов, связанных с необходимостью выделения оц ДНК-фракции из сложных смесей НК. Выявление и клонирование таких ферментов с последующим созданием на их основе новых технологий является актуальной задачей.

Степень разработанности темы. В разработку методов анализа сложных смесей НК, в частности методов нормализации и деплеции ДНК и кДНК, вовлечено большое количество исследовательских групп, результаты деятельности которых подробно рассмотрены в обзоре литературы. Хотя показано, что некоторые ДНКазы обладают предпочтительностью к дц ДНК по сравнению с оц [27-30], эти ферменты не отличаются термостабильностью и/или не имеют достаточной специфичности и избирательности к дц ДНК, что серьезно ограничивает их использование при работе со сложными смесями НК *in vitro*. Отсутствие термостабильных ферментов, способных удалять дц ДНК, оставляя интактными одноцепочечные молекулы, привело к созданию многочисленных обходных путей, которые отличаются высокой трудоемкостью и многостадийностью, требуют больших количеств биологического материала и высокого профессионализма исследователя, что делает их неудобными, а в ряде случаев неприменимыми для работы [17].

Цель исследования. Целью настоящей работы было выявление, выделение и характеристика термостабильной нуклеазы, обладающей субстратной специфичностью к дц ДНК; и разработка с использованием этой нуклеазы технологий анализа сложных смесей нуклеиновых кислот.

Задачи исследования:

- клонирование кодирующей последовательности нуклеазы из гепатопанкреаса камчатского краба, специфичной к дц ДНК;
- выявление гомологов нуклеазы камчатского краба и структурно-функциональный анализ аминокислотных последовательностей для обнаружения ключевых участков, определяющих их специфическую ферментную активность;
- разработка новых высокоэффективных технологий, направленных на структурно-функциональные и сравнительные исследования геномов и транскриптомов как эукариот, так и прокариотических организмов;
- валидация разработанных технологий в модельных экспериментах;
- применение разработанных технологий для решения исследовательских задач и внедрение их в лабораторную практику.

Научная новизна исследования. В рамках настоящей работы была клонирована полноразмерная кодирующая последовательность нового фермента из гепатопанкреаса камчатского краба – дуплекс-специфическая нуклеаза (Par_DSN), которая демонстрировала уникальную комбинацию свойств, включающих термостабильность, высокую оптимальную температуру катализа, способность гидролизовать только ДНК в составе ДНК-ДНК и ДНК-РНК дуплексов.

Был идентифицирован ряд новых гомологов Par_DSN из других членистоногих и впервые охарактеризовано новое семейство дуплекс-специфических нуклеаз.

Были предложены новые высокоэффективные технологии для выявления однонуклеотидных различий в молекулах ДНК, идентификации целевых молекул ДНК и РНК в биологических образцах, нормализации кДНК и геномной ДНК, деплеции нецелевых последовательностей в образцах кДНК и ДНК.

Теоретическая и практическая значимость работы. В рамках работы был охарактеризован новый фермент, обладающий уникальной селективностью по отношению к дц ДНК и термостабильностью.

Был проведен структурно-функциональный анализ Par_DSN и ее гомологов и выявлены аминокислотные участки, определяющие субстратную специфичность и термостабильность нуклеазы камчатского краба.

Сравнительный и структурно-функциональный анализа Par_DSN и ее гомологов позволил выделить дуплекс-специфические нуклеазы в отдельное семейство внутри суперсемейства His-Me нуклеаз.

Уникальные свойства Par_DSN послужили толчком для разработки целого спектра молекулярно-биологических технологий анализа сложных геномов.

Данные настоящей работы легли в основу ряда коммерчески доступных продуктов, что привело к широкому внедрению полученных в исследовании результатов в мировую лабораторную практику. Нормализация и деплеция с использованием Par_DSN стали методами выбора при подготовке биологических образцов для высокопроизводительного секвенирования – NGS (от англ. next generation sequencing, секвенирование нового поколения). На основе уникальных свойств Par_DSN и разработанных нами технологий рядом западных лабораторий были предложены новые методы анализа микро РНК [31-35]. В настоящее время число статей, упоминающих дуплекс-специфическую нуклеазу камчатского краба, в базе данных Google Scholar превышает 2000.

Методология и методы исследования. Работа выполнялась в период 2002-2012 гг. в Лаборатории молекулярных технологий Института биоорганической химии имени академиков М.М. Шемякина и Ю.А. Овчинникова РАН. В работе использовались наборы реактивов российских и зарубежных производителей, стандартные и оригинальные (разработанные в Лаборатории молекулярных технологий) методы генной инженерии.

Определение последовательностей нуклеиновых кислот осуществлялось секвенированием по методу Сэнгера в Лаборатории молекулярных технологий или с использованием автоматического секвенатора серии 454 GS FLX Standard (Roche,

Швейцария) в центре «Биоинженерия» РАН. ПЦР проводили с использованием амплификатора MJ Research PTC-200 DNA Thermal Cycler (BioRad, США), ПЦР в режиме реального времени – с использованием амплификатора MX3005P QPCR Systems (Stratagene, США). Для электропорации использовали микропульсатор Biorad Micropulser (Biorad, США). Масс-спектрометрический анализ проводили с использованием MALDI-TOF/TOF с возможностью LIFT (Ultraflex TOF/TOF, Bruker Daltonik GmbH, Бремен, Германия) на базе ЦКП ИБХ РАН при участии Р.Х. Заганшина. Фрагментацию ДНК осуществляли с использованием ультразвукового процессора Cole-Parmer CP750 (Vernon Hills, IL, США).

Для анализа последовательностей нуклеиновых кислот и аминокислот применяли пакеты программ и программные продукты BLAST (<http://www.ncbi.nlm.nih.gov/BLAST>), SMART (<http://smart.embl-heidelberg.de>), SignalP (версии 3.0, 5.0 и 6.0; <https://services.healthtech.dtu.dk/service.php/SignalP>), TMHMM-2.0 (<https://services.healthtech.dtu.dk/service.php?TMHMM-2.0>) и PredictProtein (<https://predictprotein.org/>), FGESH (<http://sun1.softberry.com>), Repeat-Masker (www.repeatmasker.org/cgi-bin/WEBRepeatMasker), TREE-PUZZLE, версия 5.0, ClustalX, GeneDoc. Анализ макрочипов осуществляли с помощью пакета программ AtlasImage™ Software. Статистическую обработку данных осуществляли с использованием моделей дисперсионного анализа ANOVA, Systat Software SigmaPlot 11. Для анализа данных масс-спектрометрии применяли пакеты программ FlexAnalysis версии 2.4, BioTools версии 3.0 (Bruker Daltonics GmbH), Mascot (<http://www.matrixscience.com/>).

Сотрудники ЗАО Евроген оказали значительную помощь в синтезе олигонуклеотидов для экспериментальной работы.

Подробное описание методик приведено в разделе Материалы и методы (глава 2) настоящей работы.

Основные положения, выносимые на защиту:

1. Открыто новое семейство нуклеаз, специфически разрушающих двухцепочечную ДНК, и получившее название дуплекс специфические нуклеазы (ДСН).
2. Проведен филогенетический анализ белков данного семейства, а также исследованы функциональные и структурные особенности ферментов данного семейства на примере дуплекс специфической нуклеазы Par_DSN из камчатского краба (*Paralithodes camtschaticus*).
3. Определены границы нуклеазного домена Par_DSN, а также проведен анализ сайтов, ответственных за каталитическую активность данного фермента.
4. Разработана технология мутационного анализа в генах эукариот с использованием Par_DSN.
5. Разработана технология выявления ДНК-мишеней в комплексном продукте ПЦР с использованием Par_DSN.
6. Разработана технология селективного удаления транскриптов из популяций кДНК с использованием Par_DSN.
7. Разработана технология создания нормализованных библиотек кДНК из клеток и тканей эукариотических организмов с использованием Par_DSN.
8. Разработана технология создания нормализованных библиотек геномной ДНК из клеток и тканей эукариотических организмов с использованием Par_DSN.
9. Разработана технология удаления кДНК рРНК последовательностей при создании библиотек кДНК прокариотических организмов с использованием Par_DSN.

Личный вклад автора. Работа по клонированию и очистке нуклеазы камчатского краба Par_DSN, создание технологий с её использованием осуществлялась в Лаборатории молекулярных технологий Института биоорганической химии имени академиков М.М. Шемякина и Ю.А. Овчинникова РАН рабочей группой в составе: Д.А. Шагина, Д.В. Ребрикова, А.С. Щеглова, В.Е. Анисимовой, Е.А. Богдановой и П.А. Жулидова, под общим руководством д.б.н.

С.А. Лукьянова. Автором диссертации был разработан общий дизайн исследования и предложена стратегия клонирования дуплекс-специфической нуклеазы и ее гомологов, ему принадлежит ведущая роль в систематизации и обработке результатов экспериментов, координации и продвижении исследований, выполнении основных частей экспериментальной работы. Автор диссертации является автором всех предложенных технологий с использованием Par_DSN и их модификаций. Общая стратегия нормализации кДНК является совместной разработкой с С.А. Лукьяновым. Разработка протокола нормализации кДНК выполнялась совместно с П.А. Жулидовым. Первая модификация метода мутационного анализа была разработана совместно с Д.В. Ребриковым на основании идеи автора диссертации. Основной объем экспериментальной работы по клонированию нуклеаз из членистоногих был выполнен А.С. Щегловым, В.Е. Анисимовой и Д.В. Ребриковым. Характеристика физико-химических свойств нового фермента осуществлялась диссертантом совместно с А.С. Щегловым и П.А. Жулидовым. Протокол препаративного выделения нуклеазы был разработан А.С. Щегловым при участии диссертанта. Процедура очистки и активации рекомбинантных белков была разработана В.Е. Анисимовой и А.С. Щегловым. Филогенетические и структурные исследования нового семейства нуклеаз проводились диссертантом совместно с В.Е. Анисимовой и Е.А. Богдановой. В выполнении отдельных этапов экспериментальной работы участвовали сотрудники Лаборатории молекулярных технологий ИБХ РАН.

Сотрудники Дальневосточного института биоорганической химии В.А. Рассказов, В.Б. Кожемяко, Н.И. Мензорова принимали участие в обсуждении результатов исследования свойств Par_DSN.

Работа по изучению аллельных вариантов гена COMT, ассоциированных с высокой болевой чувствительностью, выполнялась диссертантом в рамках международного научного консорциума по исследованию хронической боли.

Работа по исследованию aberrантно-слитых транскриптов в клеточной линии клеток рака молочной железы MCF-7 проводилась диссертантом в

сотрудничестве с группой лабораторий, входящих в международную программу по исследованию рака молочной железы (Breast Cancer Research Program).

Работа по выявлению маркеров старения у рыбы *Nothobranchius furzeri* выполнялась диссертантом в рамках сотрудничества с группой лабораторий Германии в рамках проекта по изучению механизмов старения.

Соответствие диссертации паспорту научной деятельности. Научные положения диссертации соответствуют паспорту специальности 1.5.3 – Молекулярная биология (биологические науки).

Степень достоверности и апробация результатов. Результаты исследования были доложены и обсуждены на всероссийских и международных конференциях: «Третий съезд биотехнологов России им. Ю.А. Овчинникова» (Москва, 2005), «Генетика в России и мире» (Москва, 2006), «Human Genome Meeting» (Bethesda, 2006), «Ломоносов 2007» (Москва, 2007), «Биология – наука XXI века» (Пушино, 2007), «The 12th biological sciences graduate congress» (Kuala Lumpur, 2007), «Молекулярная диагностика» (Москва, 2017), «Высокопроизводительное секвенирование в геномике» (Новосибирск, 2017), научном семинаре Отдела молекулярной диагностики и эпидемиологии ФБУН ЦНИИ Эпидемиологии (Москва, 2017).

Результаты исследования были использованы при разработке учебного курса «Основы биологии» для студентов кафедры Молекулярной физики факультета Молекулярной и биологической физики ГОУ ВПО Московского физико-технического института; учебных курсов: «Современные методы в биологии» и «Биоинженерия» для студентов кафедр биохимии, молекулярной биологии, вирусологии Московского государственного университета им. М.В. Ломоносова; учебного пособия для студентов Московского инженерно-физического института «Применение современных молекулярно-биологических методов для поиска и клонирования полноразмерных нуклеотидных последовательностей кДНК»

Объективность и достоверность полученных результатов проверена в ряде модельных экспериментов и подтверждена многочисленными исследованиями,

проведенными во множестве научных лабораторий мира. Результаты исследования внедрены в международную лабораторную практику.

На основании результатов работы созданы наборы реактивов:

- *Duplex-specific nuclease, cat.# EA001; EA002; EA003; EA008 (Evrogen),*
- *Trimmer-2 cDNA normalization kit, cat.# NK003 (Evrogen)*

Работа прошла апробацию на открытом заседании Отдела геномики и постгеномных технологий ФГБУН Института биоорганической химии им. академиков М.М. Шемякина и Ю.А. Овчинникова РАН 17.05.2023 г.

Публикации. По результатам исследования опубликована 21 работа, включая 15 исследовательских и одну обзорную статью в рецензируемых журналах, представленных в международных базах данных, а также три главы в книгах международных издательств. Получено два патента на изобретения США.

Структура и объем диссертации. Диссертация изложена на 342 страницах машинописного текста, включает 32 таблицы, 78 рисунков. Диссертация состоит из введения, обзора литературы, материалов и методов, результатов исследования и их обсуждения, заключения, выводов, списка сокращений, а также списка использованной литературы. Список литературы включает 669 работ.

ОСНОВНАЯ ЧАСТЬ

ГЛАВА 1. ОБЗОР ЛИТЕРАТУРЫ

1.1. Общая характеристика и разнообразие ферментов, расщепляющих нуклеиновые кислоты

Нуклеазы являются частью более крупной группы ферментов, фосфодиэстераз, которые способны катализировать расщепление фосфорноэфирных связей. Участие ферментов в расщеплении НК было впервые описано в 1903 г., тогда же Л. Иванофф и предложил термин «нуклеазы» [36, 37].

Нуклеазы вовлечены практически во все жизнеобеспечивающие процессы клетки. Так в ходе репликации ДНК 5'-3'-экзонуклеазы и эндонуклеазы необходимы для удаления РНК-затравок [38, 39], а 3'-5'-экзонуклеаза – для точности считывания [40, 41]. Такие важные процессы метаболизма ДНК, как рекомбинация и репарация, также инициируются нуклеазами [42-44]. Нуклеазная активность необходима и для структурных изменений НК, например, топоизомеризации [45-47], сайт-специфической рекомбинации [48] и процессинге и сплайсинге РНК [49-53], РНК интерференции [54]. Деградация ДНК и РНК – один из важнейших защитных механизмов бактерий против фаговой активности [55-57]. В последние годы появились данные, что РНКазы млекопитающих, кроме непосредственной атаки на РНК-субстраты, могут выступать в качестве опсонизирующих агентов, обладают цитотоксической и иммуномодулирующей активностями, направленными на защиту организма от бактерий и вирусов [58]. Нуклеазы также играют значительную роль в процессе запрограммированной смерти клеток [59-61]. Нарушение ДНКазной или РНКазной активности ассоциировано с развитием серьезных патологий человека. Так снижение способности восстанавливать поврежденную ДНК увеличивает нестабильность генома и ведет к развитию онкологических заболеваний, а неполное удаление

эндогенных нуклеиновых кислот может быть ассоциировано с развитием аутоиммунных состояний [62, 63].

Нуклеазы включают в себя как белки, так и каталитические РНК (рибозимы). Для классификации этих ферментов используют несколько критериев [64-66]:

- характер гидролизуемой связи и природа образовавшихся продуктов гидролиза (т.е. моно- или олигонуклеотиды, оканчивающиеся на 3'- или 5'-фосфат);
- тип нуклеолитической атаки (экзонуклеаза или эндонуклеаза);
- субстратная специфичность;

Схематически эта система показана на рисунке 1.1.

Еще один способ классификации нуклеаз основан на организации их активного центра с участием ионов двухвалентных металлов и каталитическом механизме реакций. Соответственно выделяют безкатионные, однокатионные и двухкатионные нуклеазы. В активных сайтах некоторых нуклеаз обнаружены три иона металлов, однако этот вариант можно рассматривать как разновидность двухкатионного катализа [67, 68].

1.1.1 Реакции, катализируемые нуклеазами

Нуклеазы катализируют расщепление одной из двух фосфорноэфирных связей ($5'O-P$ или $3'O-P$) в НК (рисунок 1.2).

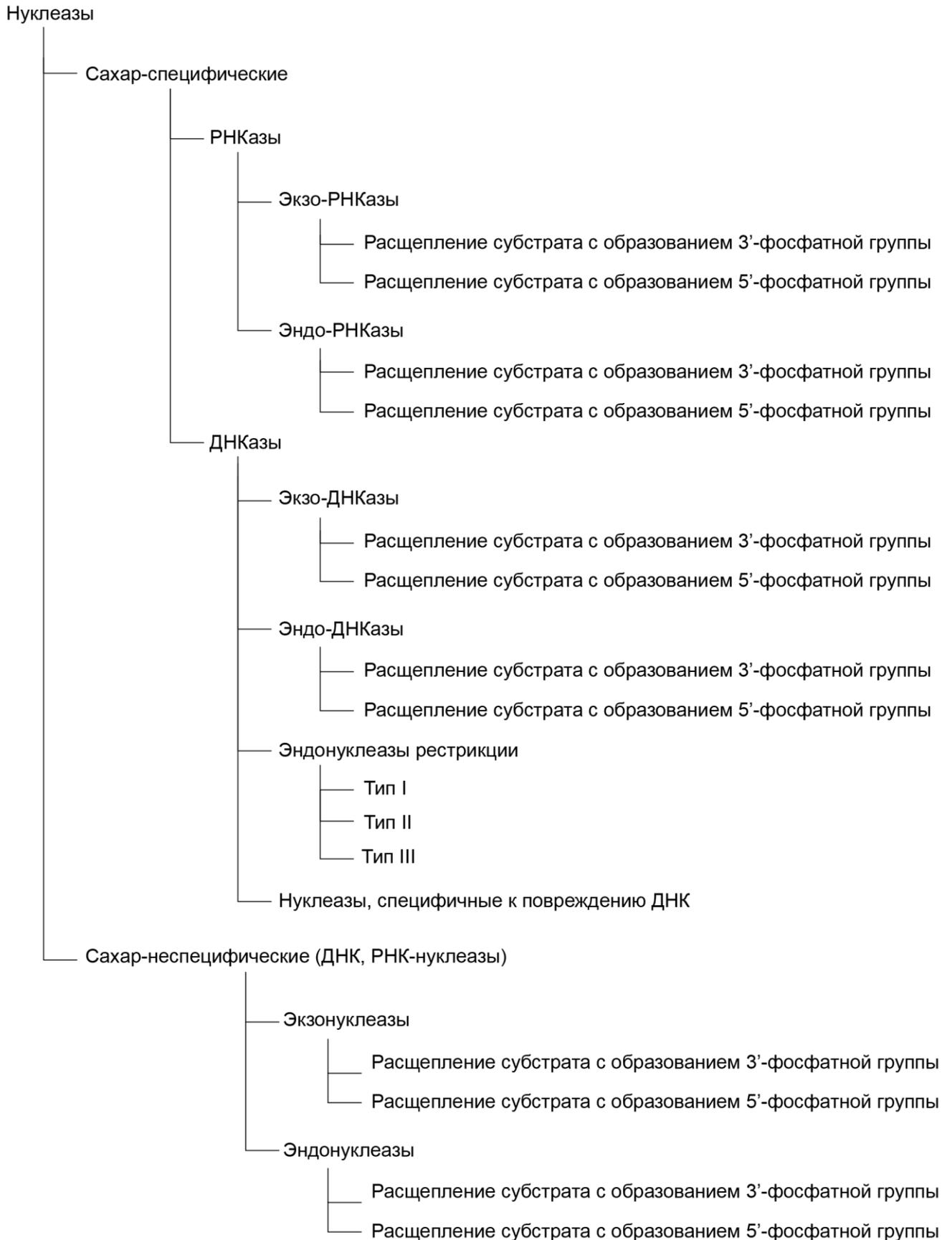
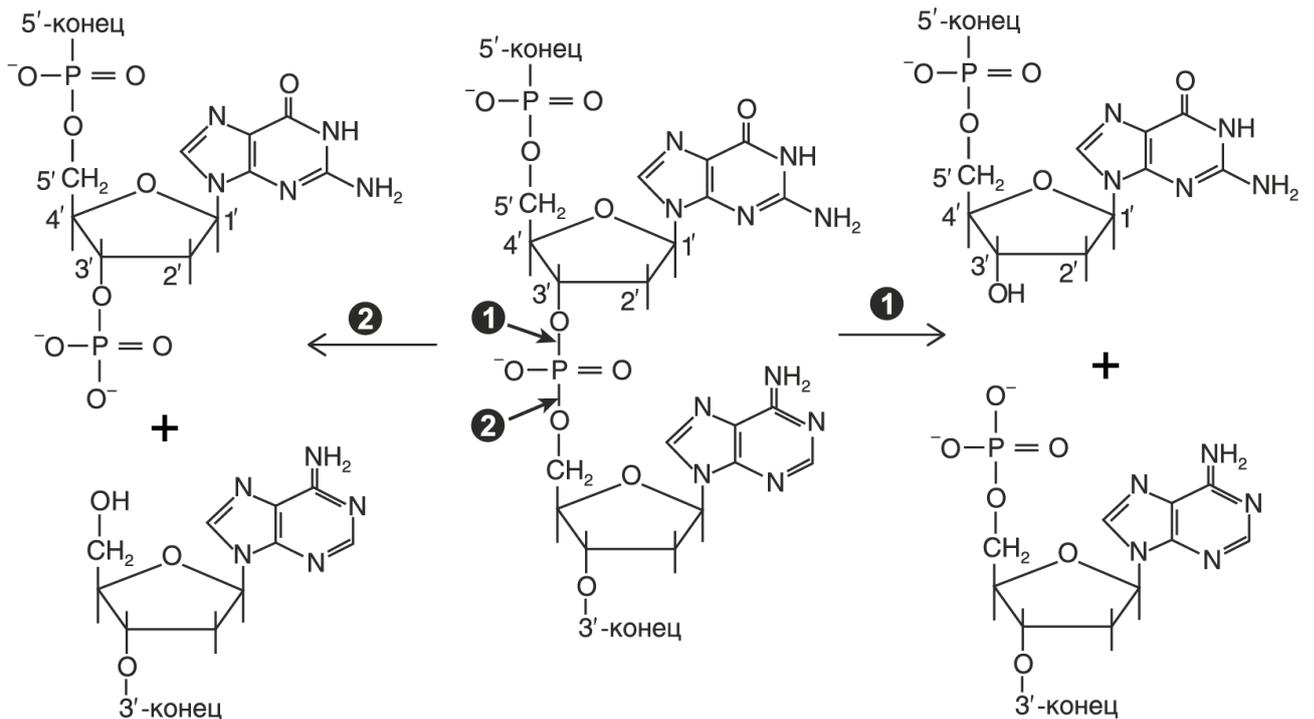


Рисунок 1.1. Классификация нуклеаз (адаптировано из Linn, 1982) [66]



Наиболее часто продукты расщепления имеют концы с 3'-гидроксильной и 5'-фосфатной группами (стрелка 1), но некоторые нуклеазы образуют продукты расщепления с 3'-фосфатной и 5'-гидроксильной концевыми группами (стрелка 2).

Рисунок 1.2. Расщепление фосфорноэфирных связей в полинуклеотидах нуклеазами (адаптировано из Zhang & Reha-Krantz, 2013) [62]

Для расщепления связи разделительного фосфата нуклеазы используют различные нуклеофилы. Наиболее распространенным нуклеофилом является молекула воды. Однако в качестве нуклеофила также может вступать и гидроксильная группа на 3'-конце ДНК или РНК во время РНК-сплайсинга, переноса ДНК-цепи и формирования шпильки [69, 70]. Кроме того, роль нуклеофила могут выполнять боковые цепи серина, тирозина и гистидина, при этом образуется промежуточный продукт «ДНК-фосфат-белок», который впоследствии распадается с возвращением фосфата на ДНК в ходе рекомбинации или топоизомеризации, или расщепляется в результате гидролиза [48]. РНКазы могут использовать в качестве нуклеофилов 2'-ОН РНК или свободные рибонуклеотиды. В качестве нуклеофила часто выступает соседний с расщепляемым фосфатом 2'-ОН рибозы. В результате гидролиза на концах продуктов расщепления образуется

5'-гидроксил и 2',3'-циклический фосфата, который затем превращается в 3'-монофосфат [71, 72]. Наконец, РНКазы H и полинуклеотидфосфорилазы, чтобы деградировать одноцепочечную РНК путем фосфоролита и получить нуклеозид-5'-дифосфаты, используют в качестве нуклеофила неорганический фосфат [73].

По положению расщепляемой фосфорноэфирной связи, которая обеспечивает различные типы образующихся продуктов гидролиза, нуклеазы делятся на две группы. Ферментов, которые могли бы расщеплять межнуклеотидную связь с любой стороны, не идентифицировано. В большинстве случаев продукты гидролитического расщепления полинуклеотида содержат 5'-фосфатную и 3'-гидроксильную группы. Такой тип расщепления наиболее предпочтителен, чем образование 3'-фосфатной и 5'-гидроксильной групп. Это связано с тем, что, во-первых, 3'-гидроксильная группа может в дальнейшем использоваться в качестве нуклеофила для реакций, катализируемых, например, РНК- и ДНК- полимеразы, ДНК лигазы, транспозазы и т.д. Во-вторых, 5'-фосфат представляет собой готовый субстрат для ДНК лигирования в конце репликации, репарации и рекомбинации.

Однако некоторые нуклеазы образуют продукты, на концах которых содержатся 3'-фосфатная и 5'-гидроксильная группы. В частности, к таким случаям относится расщепление РНК с формированием 2',3'-циклического фосфорноэфирного продукта и расщепление ДНК с помощью тирозин-зависимых сайт-специфических рекомбиназ и топоизомераз типа IB с образованием временной 3'-фосфотирозильной связи [74, 48, 71]. В случае РНК 3'-фосфат может являться конечным продуктом, либо молекулы РНК подвергаются дальнейшему гидролизу и дефосфорилированию. Например, 2',3'-О-циклофосфат сформировавшийся при тРНК-сплайсинге гидролизует с образованием 3'-гидроксильной группы и 2'-фосфата [75].

1.1.2 Экзо- и эндонуклеазы

По положению атакуемой связи в полинуклеотидной цепи нуклеазы делят на экзо- и эндонуклеазы. Экзонуклеазы атакуют связи на концах молекул ДНК и РНК, высвобождая фосфомононуклеотиды. Расщепление может происходить либо в направлении от 3' к 5' с высвобождением 5'-фосфомононуклеотидов, либо в направлении от 5' к 3' с образованием 3'-фосфомононуклеотидов [76]. Эндонуклеазы осуществляют гидролиз внутри цепи НК, что приводит к разрушению молекулы и образованию олигонуклеотидных фрагментов. Расщепление эндонуклеазами дц ДНК/РНК-субстратов может происходить по механизму «одиночного удара» или «двойного удара» или по их комбинации. В механизме «двойного удара» разрезы делаются в случайных участках каждой цепи в точках, удаленных друг от друга, и полная фрагментация дц ДНК не происходит пока напротив друг друга не окажутся парные разрывы. При «одиночном ударе» разрезы на любой из цепей делаются в точках, близких или противоположных друг другу, что приводит к полному расщеплению дуплексной ДНК за одно столкновение [76]. Разрыв молекулы будет происходить с относительно большей скоростью в первом случае по сравнению со вторым. Классическим примером эндоДНКазы является панкреатическая ДНКаза (ДНКаза 1, ЕС 3.1.21.1), которая в присутствии Mg^{2+} расщепляет ДНК по механизму двойного удара, тогда как в присутствии Mn^{2+} , Ca^{2+} или Co^{2+} механизм гидролиза переходит в режим однократного удара [77, 78]. В тоже время ионы металлов не влияют на механизм действия нуклеазы Rsn, и она расщепляет двухцепочечную ДНК по однократному механизму в присутствии Mg^{2+} , Mn^{2+} и Co^{2+} [79]. В связи с тем, что химия реакций экзо- и эндонуклеазного расщепления одна и та же, нет ничего необычного в том, что единственный активный сайт может обладать как экзо- так и эндонуклеазной активностями. Например, члены семейства FEN1 в дополнение к эндонуклеазной имеют 5'-3' экзонуклеазную активность [80], а апуриновая/апиримидиновая эндонуклеаза 1 (APE1) млекопитающих, вовлеченная в репарацию ДНК, обладает эндонуклеазной и 3'-5'-экзонуклеазной активностями [81, 82].

1.1.3 Субстратная специфичность нуклеаз

По типу субстратной специфичности выделяют ДНКазы и РНКазы, расщепляющие одно- или двухцепочечные ДНК или РНК соответственно. Большую группу представляют сахаронеспецифические нуклеазы, способные расщеплять как ДНК, так и РНК [83, 84].

В большинстве случаев ДНК – это двойная спираль. Даже РНК-субстраты в пре мРНК-сплайсинге или расщеплении, опосредованном малыми РНК, чаще всего спарены основаниями в конформацию двойной спирали. В случае, когда разделительный фосфат находится в составе двойной спирали, возможности для расположения внешнего нуклеофила ограничены. Нуклеофил может подойти к «остову» НК только снаружи от двойной спирали, потому что внутренняя часть занята сахарами и основаниями. Если же НК – одноцепочечная, ограничения на приближение нуклеофила не настолько строги. Например, расщепление РНК с помощью 2'-ОН, соседнего с расщепляемым фосфатом, может быть реализовано только на оц РНК-субстрате, так как данная реакция несовместима с конформацией двойной спирали [85, 86, 68]. Таким образом, возможных вариантов координации реакции гидролиза дц НК гораздо меньше, чем возможностей для расщепления оц НК [87, 88, 68].

Практически все нуклеазы, расщепляющие дц ДНК, способны также разрушать и оц-ДНК, хотя часто с меньшей эффективностью. Так, ДНКаза I расщепляет в оптимальных условиях дц ДНК примерно в 100-500 раз более эффективно, чем денатурированную ДНК [89]. При этом ДНКаза I не имеет предпочтений к каким-либо последовательностям нуклеотидов, но чувствительна к конформации двойной спирали. Лучший субстрат для ДНКазы I — В-форма ДНК смешанного нуклеотидного состава, при этом протяженные участки А-Т или G-C более устойчивы к нуклеазному расщеплению [90].

Большая группа неспецифических к сахару многофункциональных ферментов обладает высокой селективностью в отношении оц ДНК и РНК [91]. Некоторые из этих ферментов также проявляют 3'- или 5'-фосфомоноэстеразную

активность. Скорость гидролиза различных субстратов варьирует в зависимости от источника фермента. Так, субстратную специфичность нуклеазы P1 можно представить в виде цепочки предпочтений $3'AMФ > РНК > оц ДНК \gg дц ДНК$ [92], а нуклеазы табака – $оц ДНК > 3'AMФ > РНК > дц ДНК$ [93]. Следует отметить, что оц-специфические нуклеазы при высоких концентрациях фермента всё же гидролизуют дц ДНК [91].

Структурно-селективные эндонуклеазы (резольвазы), специфичные к повреждению ДНК, опознают и расщепляют разветвленные структуры ДНК, препятствующие прогрессии клеточного цикла, такие как остановившиеся вилки репликации или промежуточные продукты четырехсторонней рекомбинации, называемые соединениями Холлидея. Важно отметить, что соединения Холлидея физически соединяют две отдельные спирали ДНК и, таким образом, представляют собой потенциально вредные структуры, которые необходимо удалить, чтобы обеспечить эффективную репарацию ДНК, точную сегрегацию хромосом и поддержание стабильности генома [94-98]. Структуры, распознаваемые этой группой ферментов, схематически показаны на рисунке 1.3.

Белки семейства РНКазы H способны расщеплять РНК только в дц ДНК/РНК гибридах [99, 100]. Эти ферменты присутствуют во всех биологических царствах про- и эукариот, а РНКазы H-подобные домены обнаружены также в обратных транскриптазах ретровирусов и ретроэлементов [101, 102]. РНКазы H делятся на два основных класса: РНКазы H1 и H2, или ферменты типа 1 и типа 2, — с общими структурными особенностями каталитического домена, но разным набором субстратов для ферментативного расщепления. Кроме того, у некоторых архей и бактерий встречается третий тип — РНКазы H3.

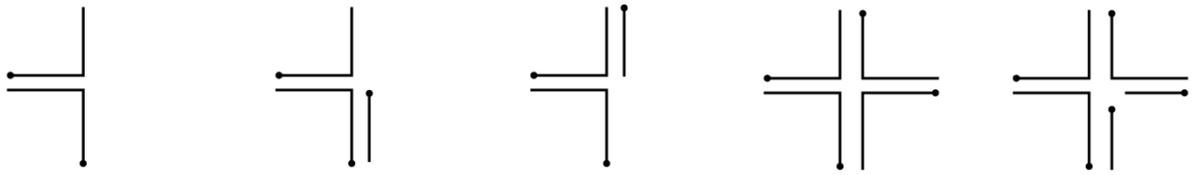


Рисунок 1.3. Структуры, распознаваемые нуклеазами, специфичными к повреждению (адаптировано из Nowotny & Gaur, 2016) [95]

РНКаза Н1 расщепляет цепь РНК в гибридах РНК/ДНК, но не обладает активностью по отношению к дц РНК или дц ДНК [99, 100, 103]. Главной ее функцией считается удаление РНК-затравок при репликации ДНК, и деградация РНК в случайных РНК-ДНК гибридах (R-петлях) [99, 104]. Субстраты, распознаваемые РНКазой Н1, не имеют специфических последовательностей, однако, ей необходим участок для связываний, включающий не менее четырех рибонуклеотидов [105].

В отличие от РНКазы Н1, РНКазы Н2 может отщеплять единичные рибонуклеотиды, включенные в ДНК-дуплекс [106, 107]. Недавние исследования показали, что работа РНКазы Н2, в отличие от РНКазы Н1, строго связана с фазой клеточного цикла (G2/M) [108]. Мутации РНКазы Н2 человека могут быть ассоциированы с развитием наследственного аутоиммунного заболевания синдрома Айкарди-Гутьерес, обусловленного накоплением остаточных РНК и ДНК [109, 110], а делеция субъединицы В РНКазы Н2 (RNASEH2B) регистрируется у больных с кастрационно-резистентным раком предстательной железы и хроническим лимфолейкозом [111].

РНКаза Н3 идентифицирована у архея и бактерий и тесно связана с РНКазой Н2 по последовательности и структуре, но обладает биохимическими свойствами, подобными РНКазе Н1 [100, 112-114].

Эндонуклеазы рестрикции (ЭР) — большая группа эндоДНКаз, которые узнают строго специфические последовательности нуклеотидов на дц ДНК. ЭР встречаются повсеместно среди прокариотических организмов, являясь частью системы рестрикции-модификации, вовлеченной в защиту генома хозяина от

чужеродной ДНК, такой как ДНК бактериофагов [115-117]. В настоящее время выделяют четыре основных типа ЭР.

ЭР типа I представляют собой многофункциональные ферменты, которые могут катализировать как рестрикцию, так и модификацию ДНК. S-аденозилметионин (AdoMet) является кофактором и донором метила для активности метилтрансферазы; эндонуклеолитическая активность требует АТФ, AdoMet и Mg^{2+} . Фермент рестрикции типа I состоит из трех субъединиц (S, M и R). Субъединица S обеспечивает распознавание мишени, M — включает сайт связывания AdoMet и активный сайт, ответственный за метилирование ДНК; субъединица R содержит активный сайт, ответственный за гидролиз АТФ, N-концевой эндонуклеазный домен, слитый с так называемым «моторным» доменом, характерным для ферментов процессинга ДНК и РНК и обладающим транслокационной или геликазной активностью [118]. Нуклеотидные последовательности, распознаваемые ферментами типа I, асимметричны и состоят из двух компонентов, разделенных неспецифическим спейсером длиной 6-8 п.о.: один длиной 3-4 п.о., а другой — 4-5 п.н. Все известные ферменты типа I метилируют остатки аденина, по одному в каждом компоненте последовательности-мишени, но на противоположных цепях. Активность ЭР типа I в качестве эндонуклеазы или метилтрансферазы определяется состоянием метилирования последовательности-мишени [119]. Если ни одна часть участка распознавания не метилирована, R-субъединицы перемещают фланкирующую ДНК и расщепляют ее примерно посередине между соседними сайтами узнавания [120, 121]. Мутации в эндонуклеазном домене могут предотвращать расщепление ДНК, не влияя на транслокационную функцию ДНК [122].

Основным критерием для классификации ЭР как фермента типа II является то, что она расщепляет ДНК специфически внутри или вблизи своего сайта узнавания и не требует гидролиза АТФ для своей активности. Часть ЭР типа II катализирует в дополнение к расщеплению и метилирование ДНК. Этим ферментам для функционирования требуется AdoMet. Классические ЭР II типа представляют собой гомодимеры, которые распознают палиндромную

последовательность длиной 4–8 п.н. и в присутствии Mg^{2+} расщепляют две нити ДНК внутри или непосредственно рядом с участком распознавания с образованием продуктов реакции с 5'-фосфатом и 3'-гидроксильной группой. Типичными представителями ЭР типа II являются EcoRI (в процессе расщепления ДНК образуются липкие концы с 5'-выступами), EcoRV (в процессе расщепления ДНК образуются тупые концы) и BglI (в процессе расщепления ДНК образуются липкие концы с 3'-выступами) [123-126]. На основании различий в сайтах узнавания и особенностей рестрикции ДНК была предложена классификация ЭР II типа с выделением одиннадцати подгрупп [127].

ЭР типа IIS распознают асимметричные последовательности и режут ДНК на определенном расстоянии от участка распознавания [128]. Эти ферменты используют разные подходы для связывания и расщепления ДНК, многие содержат отдельные ДНК-связывающий и ДНК-расщепляющий домены. Устройство каталитического центра ЭР IIS весьма разнообразно: некоторые содержат мотивы $\beta\alpha Me$, подобные тем, которые обнаруживаются в неспецифических нуклеазах; для других характерно наличие канонического каталитического мотива PD-(D/E)XK; третьи имеют «мотив фосфолипазы D» и расщепляют ДНК в отсутствие ионов двухвалентного металла [126, 128-130].

К подгруппе IIA принадлежат ЭР, также распознающие асимметричные участки, но позиция разреза находится вблизи или внутри распознаваемого участка [126].

ЭР типа IIR всегда распознают палиндромные последовательности и представляют собой гомодимеры, в которых каждая из двух субъединиц распознает один «полусайт» ДНК-мишени. Большинство ЭР типа IIR обладают каноническим каталитическим мотивом PD-(D/E)XK [126, 131].

ЭР подтипов IIE и IIF взаимодействуют с двумя копиями участков распознавания, при этом в первом случае гидролизу подвергается только один из участков, а во втором случае — оба. Ферменты типа IIF представляют собой тетрамеры, которые демонстрируют сходное с IIR связывание отдельных «полусайтов» ДНК с отдельными белковыми субъединицами и расщепление

палиндромов. В отличие от ферментов типа ПР, которые связываются с одним сайтом за раз, эти тетрамерные ЭР обычно связываются с двумя сайтами-мишенями, расщепляя один или оба аллостерическим или кооперативным способом [131].

К типу ПТ относят ферменты, которые распознают асимметричные последовательности и расщепляют ДНК внутри участка распознавания или очень близко к одной из его сторон. Эти ферменты имеют два разных каталитических сайта, которые могут быть расположены на двух различных субъединицах или в двух разных каталитических доменах на одной полипептидной цепи. В некоторых случаях две субъединицы/домена имеют одинаковый размер и оба участвуют в распознавании последовательности, а также в катализе; в других случаях субъединицы/домены имеют разный размер, и большая субъединица распознает последовательность целиком и расщепляет одну цепь, тогда как меньшая субъединица не обладает специфичностью к последовательности и только расщепляет вторую цепь. ЭР ПТ ассоциированы с двумя метилтрансферазами, которые модифицируют цепи своей ДНК в участке распознавания. Поскольку ЭР типа ПТ имеют два различных каталитических сайта, они могут быть преобразованы в никизы — эндонуклеазы, вносящие разрывы в одну из цепей ДНК [126, 131].

ЭР типа ПВ расщепляют ДНК с двух сторон от участка распознавания. Они могут распознавать как асимметричные (например, *VcsI*), так и симметричные последовательности (*VprII*) ДНК. Как правило, эти ферменты состоят из разных субъединиц и обладают как рестрикционной, так и модифицирующей активностью, которая проявляется в присутствии *AdoMet* [132, 133]. Для *VcsI* было показано, что каталитические центры рестрикции и модификации расположены в α -субъединице, тогда как β -субъединица содержит целевой домен распознавания [134].

ЭР типа ПГ, так же, как и ЭР ПВ, проявляют рестрикционную и модификационную активности и нуждаются в присутствии *AdoMet*. Однако в

отличие от ЭР II все активные сайты этих ферментов расположены на одной полипептидной цепи [135].

Эндонуклеазы рестрикции типа IIM распознают метилированную ДНК. Наиболее известным примером является эндонуклеаза DpnI [136]. DpnI — мономерный белок, состоящий из каталитического домена PD-D/EХК и С-концевого домена-активатора в виде крылатой спирали (winged helix). Оба домена связываются с ДНК в зависимости от последовательности и статуса метилирования. DpnI расщепляет ДНК внутри участка распознавания по механизму «двойного удара», то есть разрезая по одной цепи за раз. Интересно, что аллельный вариант DpnI, названный DpnII, расщепляет ту же последовательность ДНК — GATC, но только если она не метилирована [126].

ЭР III типа включают ограниченное число биохимически охарактеризованных ферментов. Это многосубъединичные и многофункциональные белки, обладающие эндонуклеазной, метилтрансферазной и АТФазной активностью. Сайты узнавания ферментов рестрикции III типа не являются палиндромными и имеют длину 5–6 п.о. [137, 138]. Наиболее хорошо охарактеризованы EcoP15I, EcoP1I, HinfIII и PstII, которые распознают асимметричные последовательности, 5'-CAGCAG-3', 5'-AGACC-3', 5'-GGAAT-3' и 5'-CTGATG-3' соответственно [137]. EcoP15I, EcoP1I, HinfIII и PstII требуют гидролиза АТФ и Mg^{2+} для расщепления ДНК и обладают АТФазной активностью. В то же время было показано, что PstII расщепляет ДНК в присутствии GTP и CTP [138]. Всем ЭР типа III для связывания и гидролиза ДНК необходимо наличие нескольких расположенных в противоположной ориентации сайтов узнавания. Таким образом, можно считать, что сайты распознавания ЭР типа III являются «симметричными» и прерываются в центре симметрии неспецифическим спейсером переменной длины [139].

Относительно новым классом эндонуклеаз рестрикции являются ЭР типа IV. Они распознают ассимметричные сайты, состоящие из двух разделенных случайной последовательностью частей. Эти ферменты способны расщеплять только модифицированные ДНК-субстраты, основания в которых были метилированы,

гидроксиметилированы и глюкозил-гидроксиметилированы. Для функционирования ЭР типа IV нуждаются в ионах металлов и AdoMet. Разрезание ДНК происходит около одного из участков распознавания [56, 140, 141].

Другая группа нуклеаз, которые распознают специфические последовательности ДНК, получила название эндонуклеазы конверсии генов или хоуминг-эндонуклеазы (ХЭ). ХЭ распознают протяженные специфические (или вырожденные) последовательности ДНК длиной от 15 до 38 п.о. и могут вносить разрывы в обе или в одну из цепей ДНК. Кодированные последовательности ХЭ были обнаружены в виде открытых рамок считывания в интронах и интраэонах во всех биологических надцарствах; каждая связана с уникальным геномом хозяина. ХЭ вовлечены в процессы гомологичной рекомбинации и вносят разрывы в определенных сайтах-мишенях внутри родственных аллелей, в которых отсутствует последовательность, кодирующая эндонуклеазу. Репарация приводит к невзаимному переносу кодирующей последовательности ХЭ в «ХЭ-аллель», и это событие может быть связано с ко-конверсией маркеров, фланкирующих сайт встраивания последовательности ХЭ [131, 142-146]. В отличие от ЭР, которые устанавливают контакты с каждым основанием в пределах сайта распознавания, ХЭ используют стратегию, в которой различное количество контактов устанавливается с отдельными парами оснований в длинном целевом сайте. Это позволяет достичь общей высокой специфичности, но различной точности распознавания в разных местах участка связывания и нивелировать влияние полиморфизма на работу фермента [147]. На основании консервативных мотивов активных сайтов ХЭ были сгруппированы в пять семейств (LHE или LAGLIDADG, HNH, GIY-YIG, His-Cys box и PD-(D/E)XK) [144-146].

Некоторые эндонуклеазы, получившие название «никазы», вносят разрывы только одну из цепей двухцепочечной ДНК. Сайтами распознавания для никаз, так же, как и для ЭР, служат короткие специфические последовательности ДНК. Никазы расщепляют ДНК в фиксированном положении относительно участка распознавания. К этой группе ферментов относят прежде всего субъединицы гетеродимерных ЭР [148], однако ряд никаз, по-видимому, представляет собой

природные мутанты ЭР, которые потеряли способность к димеризации и, как следствие, утратили активность расщеплять вторую цепь. К таким ферментам относятся никаза бактериофага f1, необходимая для его репликации [149], высокочастотные никующие эндонуклеазы вируса хлореллы [150-152] и белки, ассоциированные с ДНК-метилтрансферазами, которые атакуют ДНК рядом с неканонической парой G/T [127]. ДНК-трансферазы релаксосом, обнаруженные у ряда бактерий и принимающие участие в инициации переноса ДНК в ходе конъюгации, также могут выступать в роли никаз [153]. Также были идентифицированы и охарактеризованы редкочастотные никующие хоуминг-эндонуклеазы I-HmuI и I-BasI с сайтом узнавания более 20 п.о. [154-156]. Дополнительно на основе ряда ЭР и ХЭ были созданы генно-инженерные продукты с активностью никаз [131, 157].

1.1.4 Механизмы катализа

В зависимости от организации активного центра и механизма катализа с участием ионов двухвалентных металлов выделяют безкатионные, однокатионные и двухкатионные нуклеазы. В активных сайтах некоторых нуклеаз обнаружены три иона металлов, однако этот вариант можно рассматривать как разновидность двухкатионного катализа [67, 68]. Ферментативная реакция нуклеаз достигается либо за счет координации боковой цепи, либо за счет координации металлов. Механизмы расщепления сильно варьируют. Даже ферменты, выполняющие функционально идентичные реакции, могут иметь разные механизмы расщепления субстрата [68, 158]. Например, среди ЭР выделяют, по крайней мере, пять структурных семейств эндонуклеаз, которые реализуют, как минимум, четыре различных каталитических механизма [68, 123]. С другой стороны, ферменты со сходными третичными структурами могут иметь разные каталитические системы. Например, Cas6 и Cas2 нуклеазы имеют одинаковую ферредоксиновую укладку, тогда как их активные центры различаются [159-161]. Одновременно, сходство

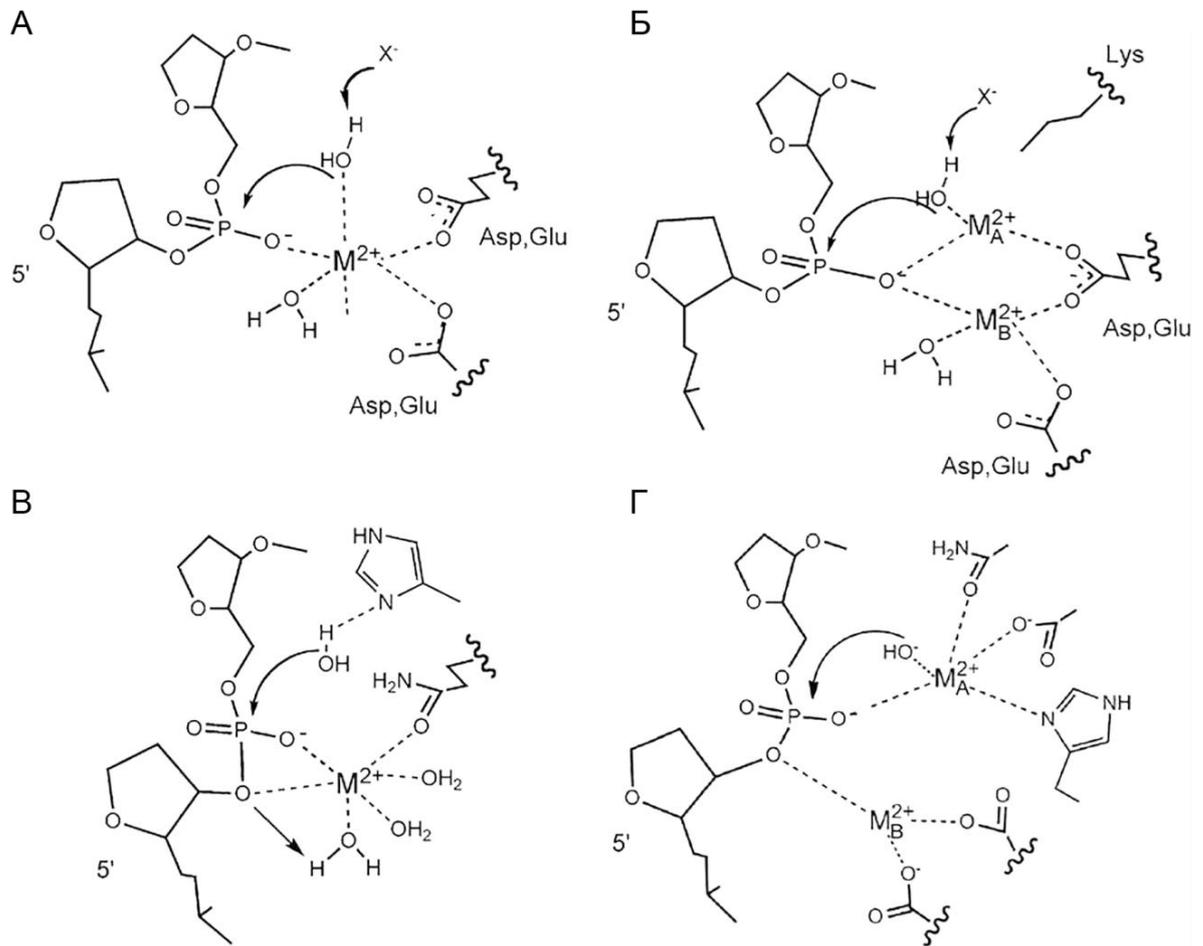
третичной структуры и механизм работы активного центра необязательно связаны с одинаковыми функциями.

Большинство охарактеризованных нуклеаз являются металлоферментами. Однако небольшое число нуклеаз не требуют для катализа ионов двухвалентных металлов. Эффект истощения ионов металлов используется для остановки реакции нуклеаз, зависимых от ионов металлов, например, для защиты от РНКаз, в то время как независимые от ионов металлов ферменты могут использоваться в буферных растворах, содержащих хелаторы [162]. Интересным исключением является CEAN (Chelated EGTA Activated Nuclease), которая была обнаружена в люминальных жидкостях репродуктивного тракта самцов мышей. Эта нуклеаза проявляет каталитическую активность исключительно в отношении плазмидной ДНК при добавлении как ионов Ca^{2+} , так и ЭГТА (этиленгликоль-бис-(β -аминоэтиловый эфир)- N,N,N',N' -тетрауксусная кислота). Удаление ЭГТА- Ca^{2+} блокирует каталитическую активность, которую можно повторно активировать добавлением новой порции ЭГТА- Ca^{2+} . Добавление ЭДТА (этилендиаминтетрауксусной кислоты) в сочетании с ионами металлов также активирует фермент, в то время как другие хелаторы, включая цитрат натрия, не стимулировали деградацию НК [163].

Кристаллография долгое время была предпочтительным методом для выяснения различных аспектов механизма металл-зависимого катализа расщепления НК и предложения моделей для выдвижения аргументов о количестве вовлеченных ионов металлов и их роли в реакции [68, 164, 165]. На рисунке 1.4 представлены модели катализа расщепления фосфорноэфирных связей нуклеазами с участием одного или двух ионов металла. Во всех случаях продукты металл-зависимой реакции расщепления НК имеют на концах 5'-фосфатную и 3'-гидроксильную группы.

Большинство металлонуклеаз в качестве кофактора используют Mg^{2+} . Многие могут использовать Mn^{2+} вместо Mg^{2+} , а некоторые могут использовать различные катионы, включая Co^{2+} , Zn^{2+} , Ni^{2+} , Fe^{2+} и Cu^{2+} [67, 68, 166, 167]. Известны Ca^{2+} -зависимые нуклеазы, однако в ряде случаев, например, для некоторых ЭР типа II, ионы Ca^{2+} ингибируют катализ [125]. Ионы металлов могут участвовать в

повышении сродства к субстрату [168] и в модулировании специфичности для конкретной последовательности или структурного субстрата [169].



(А) Механизм с участием иона одного металла в кислом активном центре эндонуклеазы EcoRI; (Б) Механизм с участием двух ионов металлов в кислом активном центре эндонуклеазы EcoRV4; (B) Механизм с участием иона одного металла в активном центре I-ProI; (Г) Механизм с участием двух ионов металлов с активным центром эндонуклеазы APE-1

Рисунок 1.4. Примеры общих механизмов катализа с участием ионов металлов в белковых металлонуклеазах (заимствовано из Dupireur, 2010) [165]

У нуклеаз обнаруживаются несколько каталитических модулей консервативной структуры. Широкое распространение определенного каталитического модуля, по-видимому, отражает его каталитическую важность. Например, DEDD мотив встречается у большой группы 3'-экзонуклеаз различной субстратной специфичности (ДНК или РНК). Широкой распространенностью обладают два каталитических мотива, которые связаны с наиболее разнообразными

функциями: это PD-(D/E)XK мотив и $\beta\beta\alpha$ Me мотив. Оба мотива состоят из модулей и имеют небольшой размер, они часто расположены в аминокислотной последовательности длиной 20–30 аминокислотных остатков. Мотив DEK расположен в структуре β -шпильки, а $\beta\beta\alpha$ Me мотив состоит из β -шпильки, за которой следует α -спираль. Таким образом, PD-(D/E)XK и $\beta\beta\alpha$ -Me мотивы легко встраиваются в разные четвертичные структуры и адаптируются к разным нуклеазным активностям и путям метаболизма. Также, и PD-(D/E)XK, и $\beta\beta\alpha$ Me мотивы очень устойчивы к аминокислотным заменам. За исключением первого D в мотиве PD-(D/E)XK и His, выступающего в качестве общего основания в $\beta\beta\alpha$ Me мотиве, другие каталитические остатки могут сильно варьировать [68].

Ниже будут рассмотрены некоторые семейства нуклеаз, демонстрирующие различные каталитические механизмы.

1.1.5 Нуклеазы, катализирующие расщепление субстрата с участием двух или более ионов металла

Предположение о механизме катализа НК с участием двух ионов металлов было сделано на основе данных кристаллографии 3'-5'-экзонуклеазы DnaQ, фрагмента Кленова и щелочной фосфатазы. Изучение рентгеновских структур с высоким разрешением, представляющих каталитический центр с участием двух металлов, вместе с исследованиями мутагенеза и кинетики продемонстрировали функциональную роль ионов для катализа во многих металлонуклеазах, включая самосплайсирующиеся рибозимы. Два иона двухвалентных металлов (A и B, рис. 1.4) располагаются в активном центре фермента и координируются кислыми остатками активного центра и кислородом фосфата нуклеиновой кислоты, или фосфатного стержня в случае рибозимов. Ион металла A, как правило, располагается на нуклеофильной стороне, а B – на стороне 3'-O остаточной группы. Нуклеазы осуществляют реакцию переноса фосфорила через октаэдрическую геометрию Mg^{2+} , в которой ионы одного или двух металлов обычно координируются шестью лигандами в целом, но обычно тремя различными типами

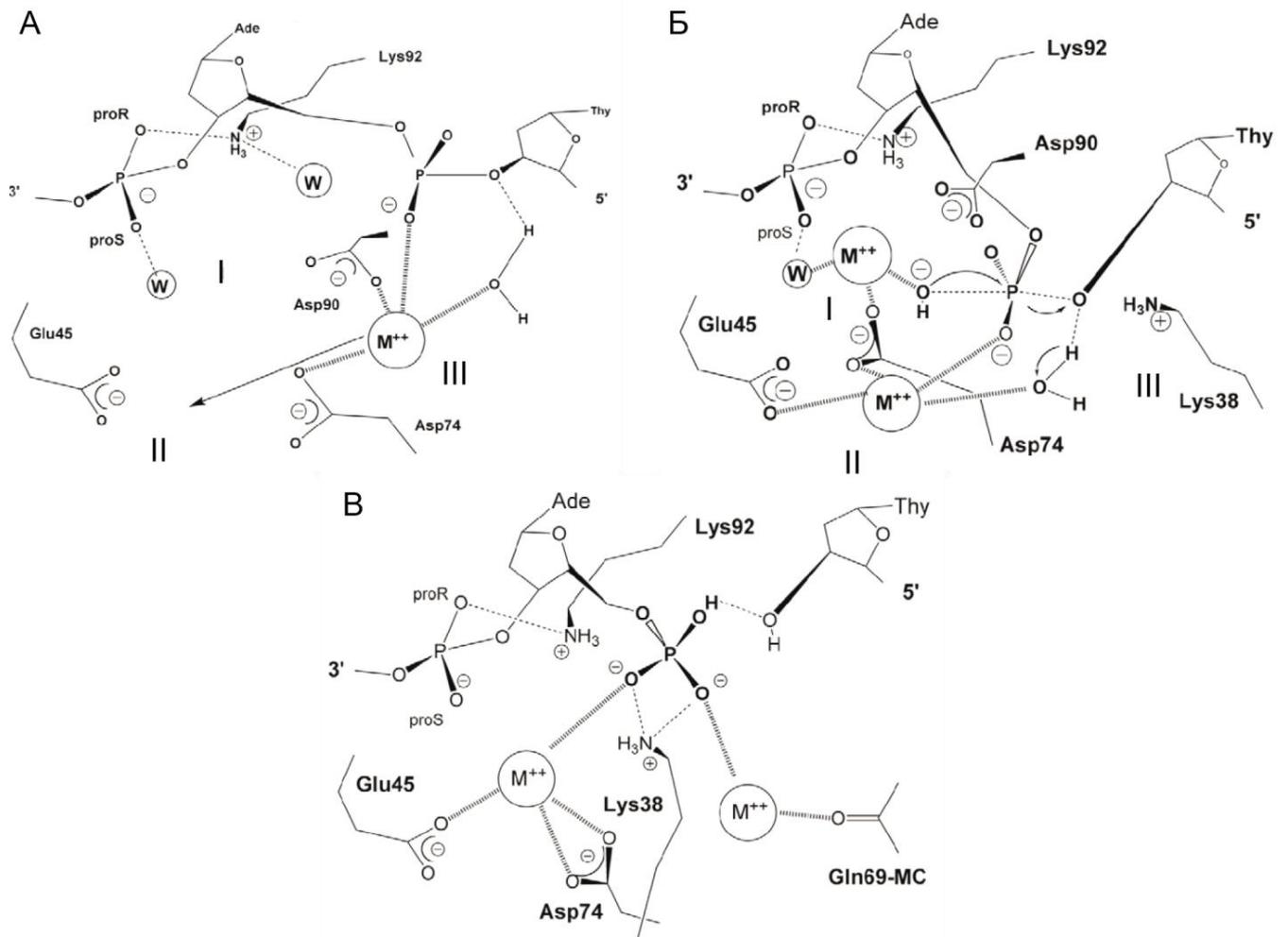
групп: (1) консервативными карбоксилатными остатками (Asp и Glu) в активных центрах нуклеаз, (2) расщепляющимися фосфатами НК и (3) молекулами воды. В катализе с двумя ионами металлов ион металла А депротонирует молекулу воды, чтобы инициировать атаку нуклеофила через ион гидроксида, тогда как ион металла В стабилизирует промежуточный пятивалентный фосфат, временно связывая нуклеофил и фосфат, что в дальнейшем приводит к реакции расщепления. После этого ион металла В диссоциирует и дестабилизирует комплекс субстрат-фермент. Таким образом, роль иона металла В уравнивает потребности химического переходного состояния и позволяет своевременно высвободить продукт. Дополнительные карбоксилаты или полярные группы помогают координировать ионы металла и удалять молекулы воды от иона В [68, 167, 170].

В большинстве случаев, Mg^{2+} представляет собой лучший вариант для двухкатионного катализа, в то время как Ca^{2+} может ингибировать перенос фосфорила [171]. Ввиду строгой координационной геометрии иона Mg^{2+} , связывание двух таких ионов чаще всего высоко избирательно и зависит от субстрата. В общем случае, Mn^{2+} может ослаблять субстратную специфичность нуклеаз и помогать дефектным ферментам с измененным вследствие мутации положением субстрата в активном центре и смещением разделительного фосфата [68, 171]. Такое смягчение и расширение субстратной специфичности широко распространено в металлоферментах, по-видимому, из-за того, что Mn^{2+} , являясь переходным элементом, менее требователен к строгой координации, чем Mg^{2+} . Было показано, что активный сайт А может принимать разные ионы металлов, тогда как сайт В более нуждается в участии Mg^{2+} [172].

Была предложена модель, описывающая каталитическую активность, которая предполагает, что один ион металла способствует образованию нуклеофила, в то время как второй ион металла способствует отщеплению уходящей группы во время гидролиза РНК. Оба металла были ответственны за стабилизацию ферментативного переходного состояния [173]. Надо отметить, что в каталитическом центре некоторых нуклеаз было обнаружено три иона металла. Однако для них был постулирован механизм катализа с участием двух ионов. В

системе EcoRV три различных положения ионов металлов были выведены из нескольких кристаллических структур активных центров фермента дикого типа и генно-инженерных мутантов [174]. В рибозиме Tetrahymena доказательства наличия этих сайтов появились в результате измерений зависимости параметров расщепления от ионов металлов на атом-специфически модифицированных рибозимах [175]. В обоих случаях дальнейшие исследования независимо привели к модели катализа, при которой одни и те же ионы металлов могли перемещаться в молекуле белка, занимая в ходе реакции несколько различных мест связывания (рисунок 1.5). Сходные данные, указывающие на участие в катализе только одного иона Mn^{2+} при наличии двух сайтов связывания металла в каталитическом центре, были получены для РНКазы H1 из *Escherichia coli* [176].

Принимая во внимание, что одни только статические рентгеновские структуры не могут исчерпывающим образом раскрыть, как два иона выполняют свою функциональную роль в ходе ферментативной реакции при процессинге цепей ДНК или РНК при переходе от субстрата к продуктам, по-прежнему остаются дискуссионными вопросы о функциональной роли каждого из двухвалентных катионов в осуществлении и координации катализа. Показано, что ферменты, имеющие два иона в каталитическом центре, сохраняют активность при внесении замен аминокислотных остатков, которые координируют ион металла А, но не В. Это предполагает альтернативный механизм образования нуклеофила без участия иона металла А, и указывает на то, что одного иона металла В достаточно для проведения реакции переноса фосфила. Можно предполагать, что механизм каталитической реакции с участием ионов двухвалентных металлов может быть еще пересмотрен [164, 165, 167].



(А) В исходном комплексе занят только участок III. (Б) В переходном состоянии ион металла в этом месте перемещается в положение II, и место I также становится занятым. (В) После расщепления один ион металла переходит к Gln69.

Рисунок 1.5. Предлагаемый механизм участия двух ионов металлов в катализе субстрата эндонуклеазой EcoRV, содержащей три сайта связывания ионов металлов (заимствовано из Horton & Perona, 2004) [177]

Нуклеазы, в каталитическом центре которых обнаружено два иона металла, представляют большую группу ферментов. В нее входят белковые ферменты и рибозимы, сиквенс-специфичные и неспецифичные нуклеазы, эндо- и экзонуклеазы с 5'-3' или 3'-5' полярностью, РНКазы и ДНКазы.

1.1.5.1 DEDD-мотив содержащие (или ДнкQ-подобные) нуклеазы

DEDD-мотив содержащие (или ДнкQ-подобные) нуклеазы являются ярким примером экзонуклеаз, каталитический центр которых содержит два сайта связывания ионов металлов. Эти ферменты характеризуются наличием консервативного мотива DEDD и обладают 3'-5' экзонуклеазной ДНКазной или РНКазной активностью. Структура каталитического ядра представляет собой смешанную структуру из альфа и бета цепей ($\beta 1$ - $\beta 2$ - $\beta 3$ - αA - $\beta 4$ - αB - $\beta 5$ - αC), формирующих центральную смешанную β -складчатую оболочку. Первые два консервативных карбоксилата лежат снаружи $\beta 1$ цепи, другие два расположены на αB и αC каждой стороны β -оболочки. Два иона металла координированы между первым аспарагином и DEDD мотивом, и pro-Sp атомом кислорода разделительного фосфата. Каждый из оставшихся трех карбоксилатов координирует один из двух ионов металла. Нуклеофильная вода может быть ориентирована и активирована с помощью высоко консервативных остатков Thr или His, поступающих к последнему аспарагину DEDD мотива, что объясняет названия экзонуклеаз DEDDy или DEDDh, соответственно. Интересно отметить, что отличия в последовательности мотива мало влияют на функцию. Например, и DEDDy, и DEDDh мотивы обнаружены у высокоточных (proofreading) экзонуклеаз [68, 178]. ДНК-полимеразы I и II прокариот и РНКазы D принадлежат к подсемейству DEDDy, тогда как корректирующая ϵ -субъединица ДНК-полимеразы III (также известная как MutD или ДнкQ), экзоны РНК-зависимых РНК-полимераз SARS и других HCoV относятся к подсемейству DEDDh [179].

Бактериальные РНКазы T и олигорибонуклеазы представляют собой гомодимерные 3'-5' экзонуклеазы с DEDDh мотивом [178]. Субстратом для этих ферментов служит оц РНК. РНКазы T необходима для созревания 5S и 23S рибосомальных РНК (рРНК), кроме того, она участвует в процессинге предшественников транспортных РНК (тРНК) [180]. Олигорибонуклеазы проявляют высокую степень консервативности последовательностей и обнаружены во всех изученных геномах эукариот, а также у протеобактерий и

актиномицетов. Они специфичны для небольших олигонуклеотидов. У бактерий олигорибонуклеаза необходима для роста, по-видимому потому, что ни один из других ферментов не может эффективно действовать на небольшие субстраты. В отсутствие олигорибонуклеазы в клетках накапливаются олигорибонуклеотиды из мРНК длиной 2–5 н.о. У *Streptomyces* олигорибонуклеаза не является существенной, но ее отсутствие приводит к замедлению роста [178].

К подсемейству DEDDh относятся также 3'-5' экзонуклеазы млекопитающих TREX1 и TREX2, которые предпочтительно расщепляют оц ДНК и с меньшей эффективностью — дц ДНК. Эти ферменты представляют собой облигаторные димеры, которые вовлечены в процессы ДНК репарации, и могут оказывать корректирующую поддержку ДНК-полимеразам [181, 182]. TREX1 содержит N-концевой экзонуклеазный домен и C-концевой трансмембранный домен, который локализует фермент в эндоплазматическом ретикулуме. TREX1 разрушает эндогенную ДНК (например, ДНК ретроэлементов) в цитоплазме, что предотвращает развитие аутоиммунной патологии [109, 183, 184]. Мутации TREX1 связаны с аутоиммунными и аутовоспалительными заболеваниями, такими как синдром Айкарди-Гутьера, системная красная волчанка. Также сообщалось, что TREX1 помогает вирусу ВИЧ-1 ускользнуть от иммунного ответа в клетках-хозяевах [185, 186]. TREX1 участвует в апоптозе, опосредованном гранзимом А, для деградации оц ДНК [187].

TREX2 в отличие от TREX1 не имеет C-концевого домена. Роль TREX2 менее изучена. Исследования эмбриональных стволовых клеток мышей показали, что делеция TREX2 приводит к спонтанным хромосомным перестройкам, а дефицит этой экзонуклеазы вызывает двухцепочечные разрывы ДНК [188]. Уровень экспрессии TREX2 варьируется в зависимости от клеточного цикла, с самым низким уровнем экспрессии в G2/M, что указывает на то, что TREX2 участвует в определенных клеточных фазах, таких как G1/S и S [189].

Также представителем подсемейства DEDDh является поли А-специфичная рибонуклеаза PARN, которая взаимодействует с оц поли(А) мРНК и участвует в деградации мРНК в качестве средства удаления поли(А) хвоста. PARN

экспрессируется практически во всех тканях эукариотических организмов [190-192]. PARN уникальна среди рибонуклеаз благодаря своей необычной способности взаимодействовать с 5'-кэп-концом и поли(А)-хвостом мРНК. Считается, что взаимодействие PARN с 5'-кэпом мРНК увеличивает скорость деградации мРНК [193]. Фрагмент массой 54 кДа от общей 74 кДа поли(А)-специфичной рибонуклеазы представляет собой олигомерную процессивную и кэп-взаимодействующую поли(А)-специфическую 3'-экзонуклеазу, в то время как функция другого фрагмента 20 кДа неизвестна [194]. У *Saccharomyces cerevisiae* PARN не обнаружена, но найден фермент, участвующий в деаденилировании мРНК POP2. Консенсусный мотив DEDD Pop2p заменен на SEDQ в *S. cerevisiae*, так что ферменту почкующихся дрожжей не хватает нескольких высококонсервативных остатков, необходимых для катализа. В то же время кристаллическая структура Pop2p *S. cerevisiae* подтвердила его общее структурное сходство с нуклеазами DEDD, и, кроме того, было показано, что фермент действует как активная экзонуклеаза на некоторых последовательностях РНК *in vitro*. Pop2p вместо Mg^{2+} предпочитает Mn^{2+} и Zn^{2+} при физиологических концентрациях ионов. Эксперименты по замещению ионов показывают, что деаденилирование мРНК может тонко регулироваться локальными уровнями Zn^{2+} в клетке [195-197].

К подсемейству DEDDu относятся экзонуклеазы WRN, обладающие 3'-5' экзонуклеазной активностью по отношению к дц ДНК. В состав белка также входит хеликазный домен. Изолированный экзонуклеазный WRN домен представляет собой мономер и может деградировать 3' концы дц ДНК с 5' выступающим концом [198]. Дефицит WRN приводит к развитию синдрома Вернера, или преждевременному старению [199].

Другое семейство, относящееся к DEDDu белкам — семейство РНКазы D, проявляющее 3'→5' экзорибонуклеазную активность. РНКазы этого семейства обнаружены у эукариота и ограниченной группы бактерий. Высказываются предположения, что РНКазы D могла появиться у бактерий путем горизонтального переноса генов от эукариотического организма [178]. РНКазы D кроме экзонуклеазного домена содержит два С-концевых хеликазных домена (HRDC, от

англ. helicase RNase D C-terminal), связывающих нуклеиновые кислоты. Фермент *E.coli* участвует в созревании малых стабильных РНК [200]. Активный центр РНКазы D содержит два иона Zn^{2+} , один из которых прочно связан с ферментом, а второй - обладает слабой связью [201]. Гомологи РНКазы D эукариот носят общее название Rrp6 и являются частью экзосомы — многосубъединичного рибонуклеазного комплекса, который способствует деградации или процессингу почти всех классов РНК как в ядре, так и в цитоплазме. Почкующиеся штаммы дрожжей, лишённые Rrp6, проявляют чувствительный к температуре фенотип роста и накапливают предшественники ядерной РНК и криптические нестабильные транскрипты [202, 203].

Ферменты семейства РНКазы H относятся к эндонуклеазам, содержащим два иона металла в каталитическом центре. Они имеют ту же топологию нуклеазного домена в виде β -листа, окруженного α -спиралями. Многие ферменты содержат консервативный DEDD-мотив, формирующий каталитическое ядро. Однако консервативность этого мотива менее выражена. Только первый Asp в $\beta 1$ сохраняется у всех членов семейства и, вероятно, координирует оба иона металлов. Asp, расположенный на конце соседнего $\beta 4$, является вторым наиболее консервативным, но он заменен Glu в HJ-резольвазе RuvC и Ydc2 [68, 204]. Механизм катализа с участием двух ионов металла для РНКазы H1, содержащий классический DEDD-мотив, был установлен в сериях экспериментов с помощью рентгеноструктурного анализа [103, 205]. Первый и третий Asp координируют два иона металла. Glu, расположенный на α -спирали, связывает 2'-гидроксил РНК-субстрата и ион металла В, тем самым усиливая специфичность катализа. Четвёртый Asp координирует только ион металла А. Было показано, что замена четвертого Asp на Asn значительно снижает каталитическую активность фермента, она может быть восстановлена заменой Mn^{2+} на Mg^{2+} [103, 205]. Хотя для РНКазы H был предложен классический каталитический механизм с двумя ионами металлов, вопрос о том, один или два иона используются в катализе остается спорным [171, 206].

Сходной с РНКазой Н пространственной структурой и DDE мотивом, который приблизительно эквивалентен 1-, 3- и 4-му карбоксилатам РНКазы Н1 обладают многие клеточные и ретровирусные ферменты, катализирующие сиквенс-специфичную транспозицию/рекомбинацию. В эту группу входят нуклеаза RAG1, участвующая в перестройке генов Т- и В-клеточных рецепторов; многочисленные транспозазы (например, Tn3, Tn5 и Mu, а также транспозазы семейств Mariner/Tc3, Sleeping Beauty и hAT), опосредующие вырезание и вставку транспозонов ДНК и интегразы, которые катализируют вставку ретровирусной ДНК в геном хозяина [207-209]. Реакция, катализируемая интегразами, обычно включает две стадии: процессинг 3'-конца и перенос цепи. Например, в случае интегразы ВИЧ-1 процессинг 3'-конца состоит из удаления терминального динуклеотида GT с образованием 5'-липкого конца и 3'-конца со свободной ОН-группой. В реакции переноса цепи ДНК-мишень геномная ДНК клетки-хозяина расщепляется, и 3'-концы ДНК соединяются с мишенью [210].

Бактериальная резольваза RuvC, расщепляющая четырехсторонние структуры ДНК, называемые соединениями Холлидея, на двух цепях одинаковой полярности, также демонстрирует сходство нуклеазного домена с РНКазой Н [209]. Связывание субстрата с RuvC не зависит от последовательности ДНК, но расщепление специфично и происходит только в родственной последовательности (A/T)TT↓(G/C) [211].

Белок UvrC является одним из важнейших элементов эксцизионной репарации ДНК бактерий. UvrC состоит из двух нуклеазных доменов, каждый из которых делает надрез ДНК рядом с поврежденным участком, таким образом, чтобы фрагмент, содержащий повреждение, мог быть удален с помощью хеликазы. Один из нуклеазных доменов UvrC сходен с ХЭ семейства GIY-YIG, а другой имеет РНКазу Н-подобную укладку [212, 213]. В этом домене выявляются два остатка каталитического центра, гомологичные таковым РНКазы Н1 (1-й и 3-й). В тоже время, третий аминокислотный остаток UvrC – гистидин участвует в координации единственного иона металла в отсутствие ДНК субстрата. Следует отметить, что

замена этого консервативного гистидина на аланин незначительно уменьшает активность фермента [212].

РНКаза Н-подобный механизм катализа характерен и для эндонуклеаз, относящихся к семейству «аргонавты» (Ago), которые используют для связывания с субстратом коротких «гидовых» РНК и ДНК в качестве ориентира для нацеливания на комплементарную НК для последующего расщепления [214]. Белки семейства Ago экспрессируются почти во всех организмах и играют критическую роль в метаболизме малых РНК, а также защищают организм от чужеродных НК через интерференцию генов, опосредованную малыми РНК или ДНК [215-217]. Эукариотические белки eAgo и прокариотические белки pAgo имеют очень низкую гомологию последовательностей, но их общая архитектура и функции чрезвычайно консервативны [215, 218]. eAgo содержат варибельный N-концевой домен и PAZ-домен, которые через внутренние линкерные последовательности L1 и L2 связаны с C-концевой частью, состоящей из среднего MID и C-концевого PIWI доменов. pAgo имеют большее структурное и функциональное разнообразие, чем eAgo, и могут использовать в качестве ориентира для ДНК-мишени как РНК, так и ДНК. Большинство длинных pAgo имеют структурную гомологию с eAgo, сохраняя все четыре глобулярных домена и два линкера, тогда как короткие pAgo состоят только из доменов MID и PIWI [215, 219, 220].

Домен PAZ представляет собой олигонуклеотид-связывающую складку, которая закрепляет двухнуклеотидный 3'-выступ гидовой РНК [221-223]. Домен MID образует специальный «карман» для встраивания для 5'-концевого фосфата гидовой РНК. Гидовые РНК находятся также в контакте с оставшимися двумя доменами, однако центральная их последовательность в основном остается свободной для взаимодействия с субстратом-мишенью. Связывание с мишенью также косвенно зависит от доменов MID и PAZ [224]. Домен PIWI реализует эндонуклеазную активность и проявляет сходство укладки с нуклеазным доменом РНКазы Н [225]. Показано, что каталитический активный центр РНКазы Н содержит мотив DEDD. Большинство охарактеризованных длинных pAgo также

содержат каталитическую тетраду DEDX (X = N, D или H), которая определяет эндонуклеазную активность этих белков [215, 220]. На примере белка CrAgo из *Clostridium perfringens*, имеющего мотив DEDD в домене PIWI, было продемонстрировано, что замена D614 на Ala полностью подавляет нуклеазную активность [226].

Короткие pAgo, в которых отсутствует домен PAZ, тем не менее сохраняют способность связывать и расщеплять мишени. Например, AfAgo, типичный короткий pAgo, состоит из 427 аминокислот с коротким N-концевым элементом, доменом PIWI и доменом MID. Без домена PAZ AfAgo связывает дц РНК через канал, образованный доменом MID и доменом PIWI, а 5'-фосфат дц РНК погружается в консервативный карман, расположенный в домене MID и домене PIWI [227].

Открытая группа pAgo — белки PIWI-RE — обладает сходной архитектурой с короткими pAgo, но имеет другие важные для связывания НК консервативные остатки в доменах PIWI и MID. Полный набор каталитических остатков РНКаз обнаружен путем множественного выравнивания только у некоторых членов подсемейства PIWI-RE последовательностей, что указывает на то, что не все белки PIWI-RE активны как нуклеазы [228].

У некоторых вышеперечисленных ферментов складка РНКазы H может быть нарушена вставкой различных структур, которые чаще всего расположены перед последним каталитическим остатком. Так, у UvrC, Tn5-транспозазы и Ago эта вставка преимущественно состоит из β -тяжей, а крупная вставка у транспозазы Hermes содержит только α -спирали [207, 209].

1.1.5.2 LAGLIDADG-содержащие хоуминг-эндонуклеазы

К ферментам с двухвалентным механизмом катализа также относятся ХЭ семейства LHE. Это семейство характеризуется консервативным аминокислотным мотивом LAGLIDADG. Это самое крупное и наиболее изученное семейство ХЭ. Как правило, LHE генерируют ступенчатые разрезы в своем сайте расщепления,

оставляя четыре выступающих 3'-нуклеотида. LHE могут кодировать белки с одним или двумя мотивами LAGLIDADG. Предполагается, что LHE с одним мотивом является наследственной версией, которая после события дупликации гена, за которым следует слияние продуктов дупликации, дает версии LHE с двойным мотивом [144, 229]. LHE с одним мотивом активны в виде гомодимеров, тогда как LHE с двойным мотивом являются мономерами. Вторые, по-видимому, в большей степени способны опознавать вырожденную последовательность в хоуминг-сайте, поскольку две части белка могут приобретать различную специфичность [144, 229].

Многие белки LHE, кодируемые интронами группы I, также функционируют как матуразы и помогают в эффективном сплайсинге интронов. Эти белки не расщепляют РНК; считается, что они просто помогают сворачивать РНК в конформацию, способствующую самосплайсингу. Однако многие матуразы LHE также являются функциональными эндонуклеазами, включая I-AniI из *Aspergillus nidulans* и I-ScaI из *Saccharomyces cerevisiae*. Тесная связь между эндонуклеазами LHE и матуразами подтверждена мутагенезом: одиночная мутация E→K в эндонуклеазе I-SceII из *S. cerevisiae* активизирует латентную активность матуразы, матураза *S. cerevisiae* может быть трансформирована в ХЭ путем замены двух аминокислот. Эти результаты предполагают, что активность эндонуклеазы и матуразы тесно связана как с функцией, так и с эволюцией белков LHE, кодируемых в интронах группы I [145].

С помощью рентгеновской кристаллографии ряда белков этого семейства были предложены структурные модели, раскрывающие детали их каталитического механизма. Несмотря на небольшую гомологию первичной последовательности за пределами мотивов LAGLIDADG, топология I-CreI, I-DmoI и эндонуклеазных доменов PI-SceI и PI-PfuI имеют заметное сходство. Коровая $\alpha\beta\beta\alpha\beta\beta\alpha$ -складка субъединицы I-CreI дважды повторяется в I-DmoI, PI-SceI и PI-PfuI и придает всем трем мономерам псевдодимерную структуру. Мотив LAGLIDADG содержится в первой α -спирали каждого домена или субъединицы. Две спирали LAGLIDADG каждого белка образуют плотно упакованную структуру с прямыми ван-дер-

ваальсовыми контактами между белковыми остовами. Эта упаковка спирального остова и крутой поворот, образованный высококонсервативными остатками глицина в нижней части спиралей, обеспечивают сохранение кислотных остатков (Asp или Glu) вблизи поверхности раздела белков. Эти остатки координируют катионы двухвалентных металлов, используемые двумя соседними активными центрами. Из-за плотной упаковки спиралей LAGLIDADG оба активных центра размещены в пределах ширины малой бороздки и расположены таким образом, чтобы окружить четыре 3'-нуклеотидных основания. ДНК-связывающий интерфейс образован четырьмя β -тяжами каждого домена, которые сворачиваются в антипараллельный β -лист в форме седла. Длина ДНК-связывающих седел варьирует в зависимости от длины распознаваемых сайтов [230].

Молекулы I-CreI и I-SceI содержат три иона металла в двух близко расположенных каталитических активных центрах, при этом два внешних металла используются каждым активным центром индивидуально; третий расположен на границе раздела белковых субъединиц. По сути, этот фермент использует механизм, аналогичный показанному выше на рис. 1.5. В то же время фермент не вступает в прямой контакт ни с одним из компонентов химической реакции, включая нуклеофильную воду, расщепляемый фосфат и уходящую 3'-группу, контактируя только с ионами металлов и большим количеством хорошо упорядоченных молекул воды. Нуклеофил позиционируется внешним металлом и, вероятно, активируется для атаки одной из двух периферийных молекул воды, которые действуют как основное основание [230-233]. Однако для I-CeuI и I-DmoI было продемонстрировано наличие только двух катионов Mn^{2+} . Аналогично для белков AniI и PI-SceI в центральном положении не было обнаружено ни катионов, ни молекул воды. Предполагается, что третий центр связывания ионов металла оказывается временно занят во время катализа. Возможно, что нуклеазная реакция в этих случаях реализуется ассиметрично с предпочтительным расщеплением одной цепи субстрата [145, 234].

1.1.5.3 Суперсемейство PD-(D/E)XK

К нуклеазам с двухкатионным механизмом катализа относится большинство нуклеаз их суперсемейства PD-(D/E)XK. В это семейство входит около 70% ЭР типа II, ХЭ, резольвазы, лямбда-эксонуклеаза, нуклеазные домены кэп-захватывающих белков [123, 235-241]. Мотив PD-(D/E)XK характерен и для нуклеазы FAN1, специфически распознающей соединение оц/дц ДНК на 5'-конце молекулы НК или в области разрыва. FAN1 делает надрез в дц ДНК с отступом на 2–4 нуклеотида от места оц/дц ДНК соединения. Кроме вышеописанной эндонуклеазной активности, FAN1 обладает 5'-3' экзонуклеазной активностью. Длина дц ДНК субстрата около оц/дц соединения должна быть не менее 10 п.о. [242, 243].

Все члены суперсемейства PD-(D/E)XK имеют общее структурное ядро, состоящее из смешанных β -слоев из 4 или 5 тяжей, фланкированных α -спиралями [244, 245]. Эти вторичные структуры часто встроены в очень разные периферические элементы, которые иногда составляют большую часть белка. Фосфодиэстеразы PD-(D/E)XK проявляют очень небольшое сходство последовательностей, несмотря на сохранение общей укладки ядра и нескольких остатков, ответственных за расщепление. Общий β -лист служит каркасом для слабоконсервативного активного центра, обычно состоящего из двух или трех кислотных остатков (Asp или Glu) и одного остатка Lys, которые вместе образуют характерный каталитический мотив (P)D...Xn...(D/E)XK (где X — любая аминокислота). Остаток Lys служит для позиционирования молекулы воды для прямой атаки на расщепляемую фосфорноэфирную связь, в то время как карбоксилатные остатки координируют ион Mg^{2+} , который действует как кофактор [68, 124, 125, 240].

У членов суперсемейства PD-(D/E)XK возникли разные варианты активного сайта. Так у никующей эндонуклеазы Vsr (от англ. very short patch repair) и ее гомологов мотив (D/E)XK был заменен на «FхН», а в другой части общей трехмерной складки появился дополнительный уникальный каталитический

остаток His [246, 247]. Сходный мотив EDxHD был обнаружен у группы хоуминг-эндонуклеаз. ХЭ с такой организацией каталитического центра получили название EDxHD-эндонуклеазы [68, 235, 246]. Эндонуклеазы Vsr и «EDxHD», по-видимому, используют активированный гистидин в качестве общего основания вместо лизина [246]. Было также продемонстрировано, что в некоторых ЭР кислотный остаток из мотива (D/E)ХК «мигрировал» в другую область полипептида таким образом, что сохранилось положение карбоксилатной группы в активном центре, несмотря на то, что боковая цепь оказалась прикреплена к другому месту в остове белка [237, 248-250]. В некоторых ферментах было обнаружено, что консервативный Lys заменен остатком Glu, Gln или Asn [125, 237].

Нуклеазы PD-(D/E)хК обнаруживают значительные вариации как по количеству, так и по точному положению связанных ионов металлов во время катализа, а также по точному структурному положению внутри коровой белковой складки каждого каталитического остатка. У различных ферментов было обнаружено разное число связанных каталитическим центром ионов металла: один, два или три [237, 251, 252]. Механизм катализа с участием двух катионов постулирован для большинства ЭР и ХЭ [68, 253], а также для кЭП-захватывающих белков вирусов гриппа [254-257], в то же время хоуминг-эндонуклеаза I-Ssp6803I, по-видимому, связывает один двухвалентный ион металла [246].

1.1.5.4 FEN-подобные нуклеазы

По всей видимости, зависимый от двух металлов катализ характерен также для суперсемейства FEN-подобных нуклеаз. Нуклеазы этого семейства имеют общую консервативную $\beta\alpha$ -складку, состоящую из центрального параллельного пятицепочечного β -слоя, окруженного α -спиралями. Каталитические аминокислотные остатки (как правило, Asp и Glu) в количестве 5-8 штук располагаются между концами нитей ($\beta 1$, $\beta 3$ и $\beta 4$) и началом следующих за ними спиралей. Как правило, непосредственно в расщеплении фосфорноэфирной связи

участвует четыре консервативных аминокислоты. Многие члены этого суперсемейства обладают двойной активностью 5'-экзо- и 5'-флип-эндонуклеазы.

Члены суперсемейства FEN1 (flap-endonuclease I) и EXO1 эукариот, как и упоминавшаяся ранее структуро-специфичная нуклеаза FAN1, распознают соединение оц/дц ДНК на 5'-конце молекулы НК или в области разрыва, однако в отличие от FAN1, они расщепляют дц ДНК на границе оц/дц-контакта [126]. Также к суперсемейству FEN1-подобных белков относятся белок комплекса факторов транскрипции ПН XPG, который участвует в репарации выпетливаний ДНК в геноме, резольваза GEN1 (gap-endonuclease I), которая предположительно является главным ферментом человека, опознающим соединения Холлидея, а также T4RNaseH бактериофагов [258]. На основании данных кристаллографии было сделано заключение, что эти нуклеазы используют каталитический механизм с участием двух ионов металлов, как и другие нуклеазы, содержащие абсолютно консервативный Asp [68, 258-260].

1.1.6 Нуклеазы, катализирующие расщепление субстрата с участием одного иона металла

В ходе катализа с участием одного иона металла используется только атом металла В, а роль металла А выполняют аминокислоты активного центра. Для обеспечения нуклеофильной атаки обычно используется гистидиновый остаток, который берет на себя роль иона металла А, тогда как ион одного металла В существует и играет ту же роль в стабилизации пентаковалентного промежуточного соединения [68, 171, 261, 262].

Основными классами ферментов, для которых свойственен такой катализ, являются His-Me нуклеазы и нуклеазы HUH. Оба класса ферментов содержат, по крайней мере, один остаток гистидина в каталитическом центре и имеют сходные пространственные характеристики каталитической реакции (расположение иона металла, разделительного фосфата и нуклеофила). Как и в случае двухкатионного катализа, продукты реакции содержат 5'-фосфатную и 3'-гидроксильную группы.

1.1.6.1 Нуклеазы His-Me

В большинстве случаев нуклеазы His-Me содержат в качестве общего основания гистидин, который участвует в депротонировании и активации нуклеофильной воды, а нуклеазы HUH всегда в качестве нуклеофила для расщепления ДНК-субстрата используют тирозин [262]. Множество нуклеаз $\beta\beta\alpha$ -Me в равной степени способны гидролизовать РНК и ДНК (сахаронеспецифичные) с небольшой специфичностью к самой последовательности. Например, хоуминг-эндонуклеаза I-ProI и эндонуклеаза рестрикции II типа KpnI, зависят от дополнительных ДНК-связывающих доменов для проявления своей специфичности [146, 263].

Нуклеазы His-Me обладают разными функциями и опознают различные субстраты. К His-Me нуклеазам относятся сиквенс- и сахар-неспецифичные нуклеазы, нуклеазы семейства колицина E, ДНКазы участвующие в апоптозе, некоторые эндонуклеазы рестрикции, HNH и His-Cys хоуминг-эндонуклеазы и HJ резольваза. Некоторые нуклеазы имеют большие спиральные вставки между β элементами. Тем не менее, структура каталитического ядра и сайтов связывания иона металла исключительно консервативна [83, 84, 164, 235, 261, 264].

Для этой группы ферментов была продемонстрирована общая кристаллическая структура нуклеазного домена, включающая два антипараллельных β -тяжа, связанных с C-концевой α -спиралью, которая содержит центральный ион двухвалентного металла. Этот мотив был назван « $\beta\beta\alpha$ Me». На основании обширной структурной информации было отмечено, что мотив $\beta\beta\alpha$ Me включает строго консервативный остаток His и имеет топологическое сходство с «цинковым пальцем» — ДНК-связывающим модулем, который также состоит из двух β -цепей и одной α -спирали, связанной через структурный ион цинка. На основании этого сходства эта группа нуклеаз также была названа нуклеазами His-Me или «металлическими пальцами». Все His-Me нуклеазы демонстрируют топологию $\beta\beta\alpha$, подобную цинковым пальцам, имеют строго консервативный остаток His и содержат центрально расположенный каталитический ион металла

(Me). При этом как His, так и Me необходимы для гидролиза НК [261, 265-267]. Например, неспецифическая *Serratia marcescens* нуклеаза (SmNucA), хоуминг-эндонуклеаза I-PpoI и неспецифическая нуклеаза из *Anabaena* sp. (AnNucA), представляющие ферменты His-Me, имеют в каталитическом центре важный остаток His, определяемый как общее основание, ион Mg^{2+} , который связывает расщепляющийся фосфат, консервативный Asp и три молекулы воды [84, 264, 268] (рисунок 1.6). В активном центре I-PpoI имеются два дополнительных остатка His, но их мутагенез существенно не снижает активность фермента [269]. Эндонуклеаза I-PpoI и SmNucA имеют в активном центре остаток Arg, который предположительно взаимодействует только с продуктом расщепления. В AnNucA в этом же месте находится остаток Asp, а взаимодействие с продуктом реакции осуществляет Arg93 [84, 164, 264].

Для ряда His-Me нуклеаз, включая хоуминг-эндонуклеазы семейства HNH, ДНК-расщепляющие колицины, транспозазы и некоторые эндонуклеазы рестрикции II типа (такие как KpnI), общим является мотив HNH [123, 146, 266]. Ферменты семейства бактериальных колицинов обычно содержат канонический мотив HNH вблизи своего С-конца, хоуминг-эндонуклеазы семейства HNH — ближе к N-концу.

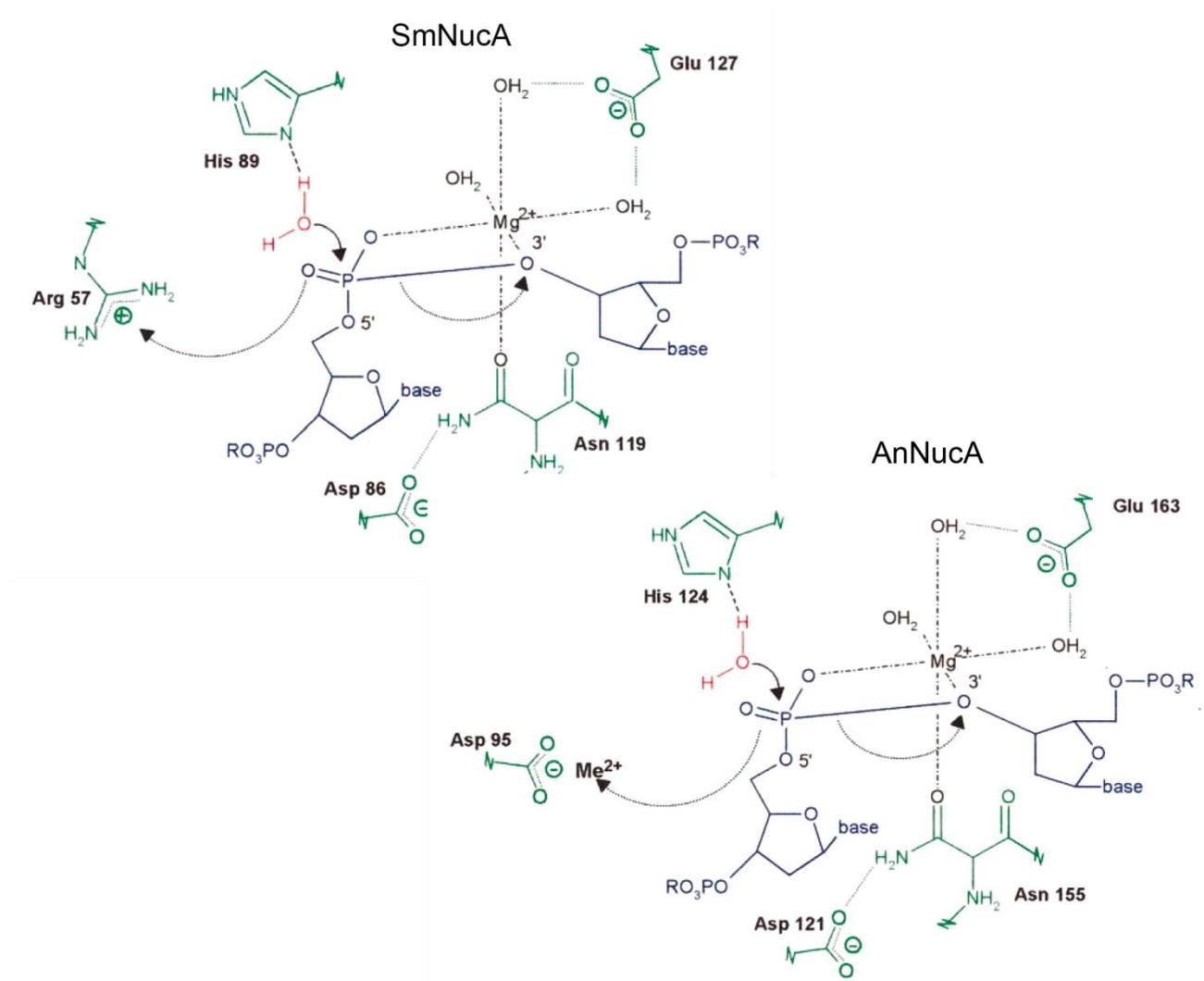


Рисунок 1.6. Каталитический центр сахар-неспецифических нуклеаз SmNucA и AnNucA (заимствовано из Meiss et al., 2000) [264]

Активный центр HNH-содержащих нуклеаз обычно не содержит консервативных каталитически незаменимых карбоксилатов, а His боковой цепи белка заменяет ион металла А при активации нуклеофила [171]. Различная расщепляющая активность этих ферментов отражает их каталитическое разнообразие: колицины неспецифически переваривают дц и оц ДНК, а ХЭ индуцируют двухцепочечные либо одноцепочечные разрывы в своих дц ДНК-мишенях [83, 146, 231, 270].

Высокую степень гомологии и HNH мотив имеют ДНК-расщепляющие колицины. Колицины представляют собой сильнодействующие антимикробные соединения, продуцируемые некоторыми членами семейства Enterobacteriaceae, которые сдерживают рост, а также выживание других бактерий, занимающих

общую нишу. Система номенклатуры колицинов определяется рецептором, с которым связывается фермент. Для белков, имеющих общий рецептор, к алфавитному символу добавляется цифровой индекс, как в случае E-колицинов (E1-E9) [270]. Все колицины E связываются с одним и тем же рецептором VtuB, который используется для усвоения витамина B12 [271]. На основании цитотоксической активности колицины подразделяются на нуклеазы и порообразующие колицины. Колицины-нуклеазы далее подразделяются на два типа: один, который расщепляет ДНК (колицины E2, E7, E8 и E9), и другой, который специфически расщепляет РНК (колицины D, E3, E4, E5 и E6) [270]. Структура каталитического центра колицинов-ДНКаз хорошо изучена на примере колицинов E7 и E9. ColE7 содержит ион Zn^{2+} в мотиве HNH. Нуклеазный домен ColE9 не активен с Zn^{2+} , но активен с Ni^{2+} или Mg^{2+} . Хотя ColE7 содержит ион переходного металла, а не ион щелочноземельного металла, в его активном центре используется гидролитический механизм, аналогичный механизму ранее рассмотренных NucA нуклеаз [83, 267, 272, 273]. His545 функционирует как основное основание, а также удерживается и поляризуется водородной связью с карбонильной группой основной цепи. Ион Zn^{2+} связан с расщепляющимся фосфатом и стабилизирует переходное состояние фосфоаниона. Основным остатком, расположенным рядом с расщепляемой фосфатной группой, вероятно, связывается с фосфатом после расщепления ДНК (рисунок 1.7).

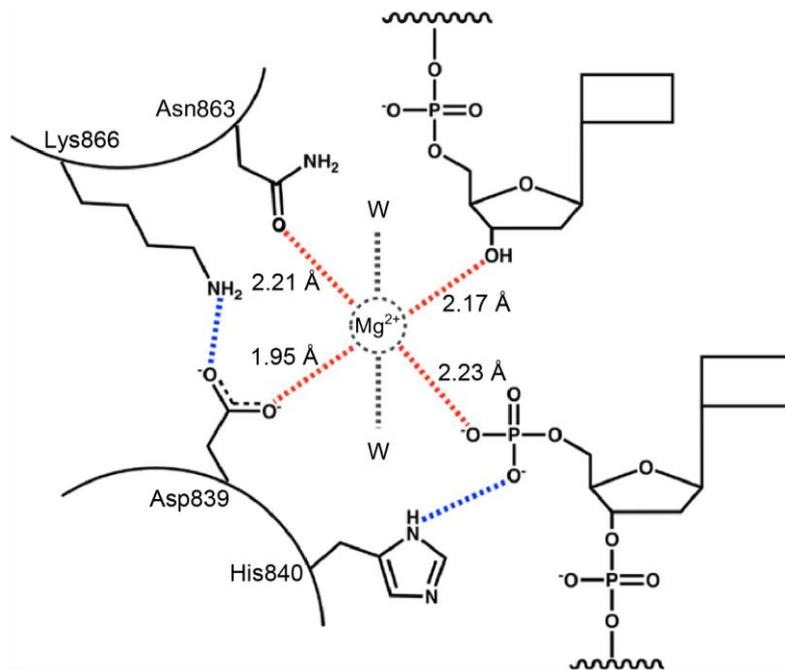


Рисунок 1.8. Каталитический центр HNH нуклеазного домена Cas9 (заимствовано из Zhu et al., 2019) [274]

1.1.6.2 HUN эндонуклеазы (релаксазы)

HUN эндонуклеазы — многочисленная группа повсеместно распространенных ферментов, вовлеченных в процессы репликации плазмид и бактериофагов по типу катящегося кольца, переноса плазмид, репликации некоторых эукариотических вирусов и в различные типы транспозиций [275, 276].

HUN эндонуклеазы реализуют расщепление и воссоединение оц ДНК с использованием остатка Tyr в активном центре для образования временной 5'-фосфотирозиновой связи с ДНК-субстратом. HUN эндонуклеазы содержат несколько консервативных белковых мотивов, включая мотив HUN, состоящий из двух остатков His, разделенных объемным гидрофобным остатком (U), и мотив Y, содержащий один или два остатка Tyr, разделенных несколькими аминокислотами. Эти ферменты катализируют разрыв и соединение оц ДНК с помощью уникального механизма с использованием Tyr для создания промежуточного соединения 5'-фосфотирозина и свободного 3'-ОН в месте расщепления. Эндонуклеазный домен

НУН вместе с лигандом двухвалентного иона металла (M^{2+}) разрывает оц ДНК с образованием ковалентного промежуточного соединения и высвобождает отщепленную уходящую группу (рисунок 1.9) [275]. Впоследствии 3'-ОН используется для праймирования репликации, как это наблюдается для доменов НУН в белках Rep одноцепочечных фагов и RCR-плазмид, или действует как нуклеофил для переноса цепи на стадии терминации репликации, в ходе конъюгативного переноса и транспозиции (там же).

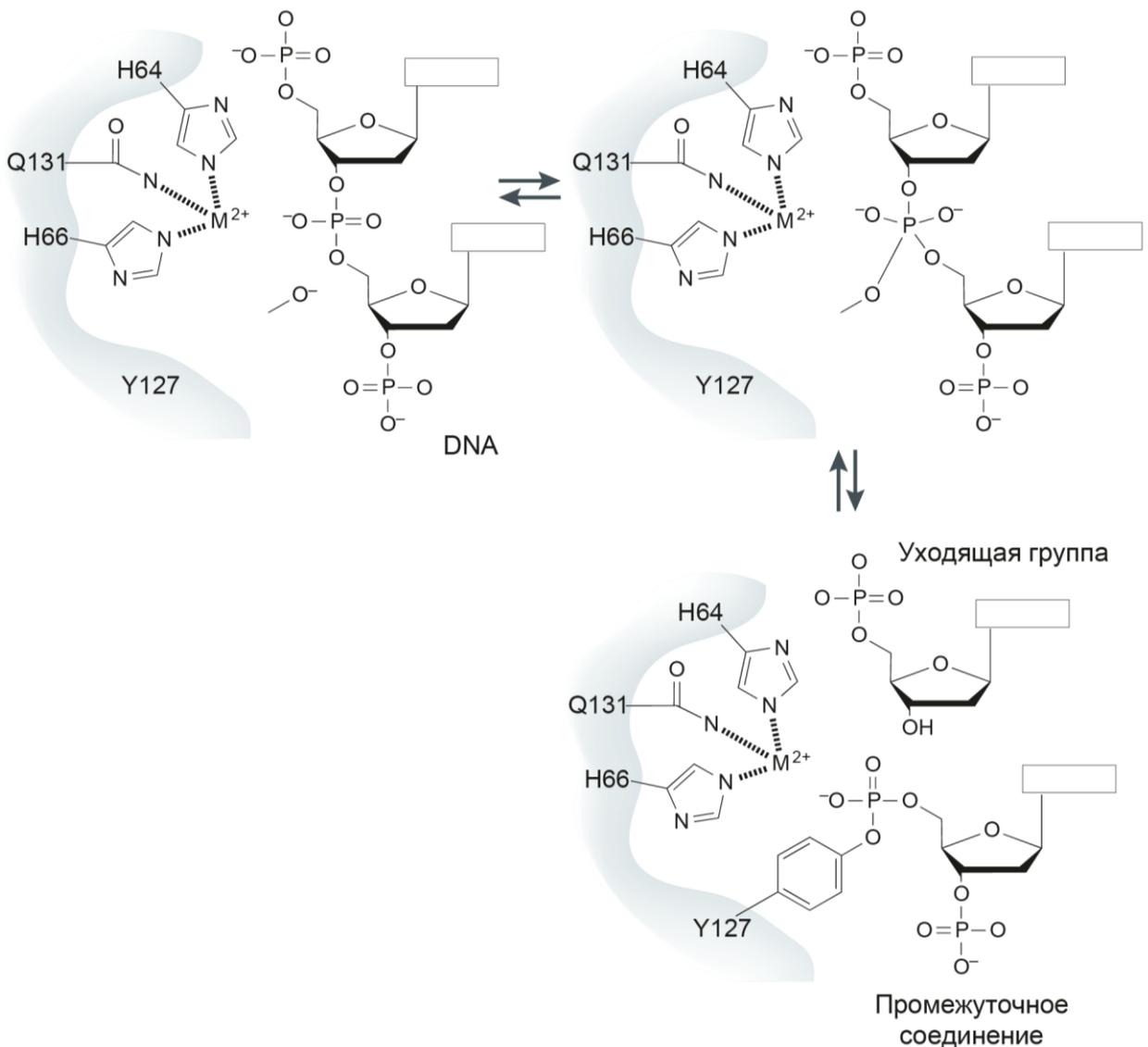


Рисунок 1.9. Катализ с использованием одного иона металла, реализуемый НУН эндонуклеазами (заимствовано из Chandler et al., 2013 с изменениями) [275]

1.1.6.3 Хоуминг-эндонуклеазы, содержащие His-Cys мотив

Хоуминг-эндонуклеазы, содержащие His-Cys мотив, представляют еще одну группу His-Me нуклеаз, которая, по-видимому, имеет те же эволюционные корни что и HNH хоуминг-эндонуклеазы. His-Cys хоуминг-эндонуклеазы найдены в интронах группы I, прерывающих гены ядерной рРНК низших эукариот, и характеризуются высоким содержанием гистидиновых и цистеиновых аминокислотных остатков [146, 235]. Наиболее изученный фермент этой подгруппы — I-ProI из слизевика *Physarum polycephalum* [277]. Кристаллическая структура I-ProI представляет собой димер, где каждый мономер содержит две последовательности, богатые гистидином и цистеином, которые координируют отдельные ионы цинка и участвуют в стабилизации структуры белка [278].

T4 эндонуклеаза VII (Endo VII) – единственная HJ резольваза, которая содержит $\beta\beta$ Me мотив. Она представляет собой гомодимер, узнает и расщепляет крестообразные структуры Холлидея. Каталитический центр фермента проявляет сходство с каталитическими центрами I-HmuI и I-ProI. В отличие от эндонуклеазы I фага T7, которая обладает типичным для РНКаз активным сайтом и использует двухкатионный механизм катализа [279], эндонуклеаза VII фага T4 не является специфичной для четырехконцевого соединения и может расщеплять трехконцевые соединения и ошибочно спаренные ДНК-гетеродуплексы, которые характеризуются искаженными двуспиральными структурами [280, 281].

1.1.6.4 ДНКазы I подобные нуклеазы

Семейство ДНКазы I подобных нуклеаз включает эндоДНКазы и некоторые РНКазы, с различными биологическими функциями от расщепления ДНК в биологических жидкостях до репарации ДНК и процессинга РНК [44, 68, 282, 283]. ДНКазы I и ее гомологи (DNASE1L1, DNASE1L2 и DNASE1L3) относятся к $\text{Ca}^{2+}/\text{Mg}^{2+}$ -зависимым эндоДНКазам, причем совместно присутствующие в

растворе ионы Ca^{2+} и Mg^{2+} оказывают синергетический эффект на ферментативную активность [282]. Эндонуклеаза APE1, кроме эндонуклеазной активности против ДНК, имеет 3'-5'-экзонуклеазную, 3'-фосфодиэстеразную активности и способна расщеплять РНК [81, 82].

ДНКазы I, EcoIII и APE1 имеют общую четырехслойную структуру $\alpha\beta\beta\alpha$, содержащую α/β -сэндвич, как у металлозависимых фосфатаз, хотя и иную топологию β -листа, и окружение вокруг каталитических центров. Каталитический центр расположен на одной стороне β -листа, который объединяет несколько консервативных кислотных остатков Asp, Glu и His. Кристаллические структуры ДНКазы I [284] и EcoIII [285] выявили единственный ион металла, связанный с активным центром. Также один ион металла, координированный с аминокислотными остатками каталитического центра и расщепленным фосфатом, был выявлен для APE1 [286-288].

РНКазы E представляет собой эндорибонуклеазу, участвующую в деградации матричной и регуляторной РНК и в процессинге предшественников разных видов РНК [289]. РНКазы E образует гомотетрамер с двумя каталитическими центрами. Каждый каталитический центр является составным и состоит из двух полипептидных цепей, каждая из которых вносит вклад в двухслойную структуру $\alpha\beta$. Каталитический домен одной субъединицы РНКазы E (одна структура $\alpha\beta$) аналогичен половине ДНКазы I, но имеет другую топологию в молекуле белка [290].

1.1.6.5 GIY-YIG эндонуклеазы

GIY-YIG эндонуклеазы — суперсемейство, включающее GIY-YIG хоуминг-эндонуклеазы, N-концевой домен UvrC, фаговую эндонуклеазу T4 EndoII, резольвазу Slx1 [68, 95, 97, 123, 146, 235, 291, 292]. Эти ферменты имеют модульную структуру и каталитическое ядро, представляющее собой трехскладчатый антипараллельный β -лист, окруженный α спиралями с каждой

стороны. На β -шпилечной структуре, за которой идет α -спираль расположен характерный аминокислотный мотив GIY-X10/11-YIG. Таким образом, структура GIY-YIG мотива подобна $\beta\beta\alpha$ модулю. Активный сайт содержит два остатка тирозина из GIY-YIG мотива. В то же время активный сайт включает остатки (Arg, His или Tyr, Glu и Asn) вне GIY-YIG мотива [235, 293, 294]. Предполагаемый каталитический механизм эндонуклеаз GIY-YIG включает активацию поступающей нуклеофильной воды тирозином в активном центре, который сам активируется через взаимодействие с окружающими основными боковыми цепями. Одиночный двухвалентный ион металла координируется расщепляющимся фосфатом и соседними боковыми цепями активного центра. Этот механизм и каталитический центр имеют заметное сходство с тем, что наблюдается у нуклеаз HNH, но на фоне совершенно другой топологии окружающих белковых складок [246].

1.1.7 Суперсемейство TOPRIM — нуклеазы, катализирующие расщепление нуклеиновых кислот с участием одного или двух ионов металла

TOPRIM представляет собой модуль связывания металла, обнаруженный среди топоизомераз типа IA и типа II, DnaG-подобных примаз, бактериальных АТФ-зависимых нуклеаз OLD-семейства, фермента процессинга 5S рРНК (матуразы), РНКазы M5 и др. [68, 295, 296].

Модуль TOPRIM представляет собой $\beta\alpha$ -складку и состоит из центрального параллельного четырехцепочечного β -листа, окруженного α -спиралями. TOPRIM характеризуется двумя консервативными аминокислотными мотивами E и DXD на концах 1-й и 3-й β -цепи соответственно, которые координируют двухвалентные катионы. Не все домены TOPRIM участвуют в реакциях переноса фосфорила. RecR, который является важным компонентом фактора рекомбинации RecFOR, имеет структуру TOPRIM, но лишен мотивов связывания металлов и нуклеазной активности [68].

Классическими представителями суперсемейства TOPRIM являются топоизомеразы типа IA и II. Топоизомеразы — ферменты, которые расщепляют ДНК, используя боковую цепь тирозина в качестве нуклеофила, чтобы получить ковалентный промежуточный продукт «фосфотирозил ДНК-фермент» для передачи цепи [46, 297]. После расщепления и смены топологии ДНК снова лигируется с помощью своей гидроксильной группы, полученной при расщеплении для разрыва фосфотирозильной связи. Неотъемлемой частью активных сайтов всех топоизомераз является тирозин, который выступает в качестве нуклеофила. Топоизомеразы обычно состоят из множества доменов с различными функциями [298].

Для релаксации сверхскрученной ДНК топоизомеразам типа IA и II требуются катионы Mg^{2+} . Топоизомеразы IB не нуждаются в бивалентном катионе для осуществления реакции, однако, показано, что ионы Mg^{2+} положительно влияют на её эффективность [299]. Функционально топоизомеразы типа I не зависят от гидролиза АТФ и используют для катализа энергию деформации сверхскрученной ДНК. Топоизомеразы типа II обычно АТФ-зависимы и обладают отдельным АТФ-связывающим доменом [46, 298]. Топоизомеразы типа II и IA формируют продукты расщепления, содержащие 5'-фосфотирозин и 3'-ОН [68]. В отличие от металл-независимых топоизомераз типа IB (см. ниже), топоизомеразы типа IA и типа II отдают мало предпочтения определенным последовательностям ДНК (там же).

Топоизомеразы типа IA представлены эукариотическими топоизомеразами III α и III β , состоят из четырех доменов, которые координируют связывание ДНК, расщепление и лигирование цепи. Домен I содержит так называемую укладку TOPRIM, которая может связывать ионы магния. Домен II состоит в основном из β -тяжей, образующих центральное ядро и связывающих домен III, содержащий каталитический остаток тирозина, с доменом IV [300]. Эти ферменты осуществляют реакцию по механизму «перехода цепи»: первая цепь ДНК связывается с доменами I и III топоизомеразы и расщепляется каталитическим остатком тирозина, расположенным в домене III, создавая 5'-фосфодиэфирную

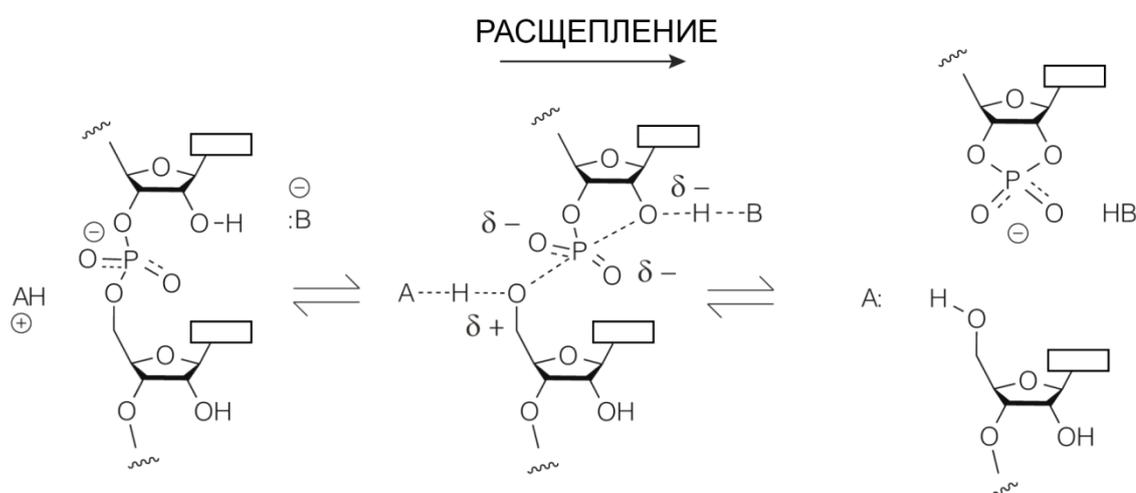
связь между ферментом и ДНК. Затем домен II, действуя как шарнир, разделяет расщепленную цепь ДНК и позволяет пройти через неё второй цепи ДНК, после чего домены I и III снова сходятся, и расщепленная цепь ДНК ре-лигируется.

Топоизомеразы типа II подразделяются на отдельные семейства IIA и IIB, расщепляющие дц ДНК [297]. Это димерные или гетеротетрамерные ферменты, имеющие составные каталитические центры, каждый из которых включает Tug в домене WHD одной субъединицы и карбоксилаты в домене TOPRIM другой субъединицы [46]. В кристаллических структурах топоизомераз типа II обнаружен единственный Mg^{2+} , координируемый консервативными мотивами E и DxD в складке TOPRIM [301, 302]. Как и топоизомеразы типа IA, топоизомеразы типа IIA проявляют свое действие через механизм прохождения нити, в данном случае с использованием, так называемого механизма «двух ворот»: дц ДНК, называемая «G-сегментом», связывается с топоизомеразой в центральной ДНК-связывающей области, называемой «ДНК-воротами». Вторая дц ДНК, «Т-сегмент», связывается с N-концевым доменом АТФазы. Затем происходит образование дц разрыва с выступающими четырьмя основаниями в G-сегменте, которое следует за образованием 5'-фосфодиэфирных связей между двумя остатками тирозина и отдельными цепями ДНК G-сегмента. Этот процесс сопровождается высвобождением неорганического фосфата и требует гидролиза АТФ. ДНК-связывающие ворота отделяются, и Т-сегмент проходит через G-сегмент в C-концевую область, или «С-ворота». С-ворота открываются и высвобождают ДНК Т-сегмента, а продукт АДФ высвобождается, перезапуская фермент для следующего каталитического цикла [298].

В отличие от топоизомераз бактериальная РНКаза M5, родственные нуклеазы OLD и DnaG-подобные примазы содержат в каталитическом центре дополнительный консервативный Asp, расположенный через три остатка после первого E и реализуют катализ, зависящий от двух ионов металла, который, видимо, сходен с механизмом катализа FEN1 [68].

1.1.8 Нуклеазы, осуществляющие катализ, не зависящий от ионов металла

Известно ограниченное число нуклеаз, не использующих ионы металлов при катализе расщепления НК. Зависимые от ионов металлов РНКазы в реакциях расщепления обычно создают продукты, содержащие 3'-ОН и 5'-фосфатные группы, в то время как независимые от ионов металлов ферменты часто генерируют продукты, имеющие 2',3'-циклические фосфаты (рисунок 1.10).



B^- — общее основание, которое используется для депротонирования нуклеофила $O2'$;

AH^+ — кислота, которая протонирует уходящую группу 5'-оксианиона.

Рисунок 1.10. Химическая реакция кислотно-щелочного катализа, приводящая к образованию продуктов с 2',3'-циклическим фосфатом (заимствовано Lilley, 2019) [303]

Отличительной чертой металлнезависимого катализа расщепления ДНК является ковалентное промежуточное соединение между ферментом и субстратом, например, фосфорил-серин или фосфорил-тирозин [68, 74, 171]. В качестве примера можно привести неспецифичные к последовательности топоизомеразы и специфичные к последовательности Cre-подобные рекомбиназы, которые используют тирозин в качестве нуклеофила и формируют 3'-фосфотиозиновый ковалентный промежуточный продукт [46, 48], а также ДНКазы суперсемейства фосфолипазы D (PLD), реакции расщепления ДНК у которых проходят с помощью

двухэтапного механизма через формирование фосфогистидинового промежуточного продукта [304, 305].

1.1.8.1 Нуклеазы суперсемейства фосфолипазы D

К нуклеазам, обладающим независимой от ионов металлов активностью, суперсемейства PLD относятся ДНКазы из *Trichinella spiralis* [306], несколько гомологичных сахар-неспецифичных нуклеаз Nuc бактерий [162, 307, 308], а также гомологичные им ДНКазы II лизосом макрофагов [309]. Эти ферменты были отнесены к суперсемейству PLD на основе консервативного НКD-мотива: $\text{HXK(X)}_4\text{D(X)}_6\text{GSXN}$ (где X – любая аминокислота). Каталитический механизм PLD-семейства высоко консервативен, в качестве субстрата могут выступать фосфолипиды, ДНК или фосфорилированные белки [310]. Консервативный мотив представлен дважды у каждого белка из-за гомодимеризации, либо внутренней дубликации [311]. His и Lys НКD-мотива вовлечены в катализ, тогда как Asp стабилизирует структуру с помощью водородной связи с амидами главной цепи. Один из двух консервативных His остатков ведет себя как нуклеофил, а другой служит донором протона для уходящей группы. Еще одним членом PDL суперсемейства является эндонуклеаза рестрикции VfiI. Белок состоит из двух структурных доменов. N-концевой каталитический домен аналогичен Nuc нуклеазе, устойчивой к ЭДТА, из суперсемейства фосфолипаз D. С-концевой ДНК-связывающий домен VfiI имеет β -бочкообразную структуру, очень похожую на эффекторный ДНК-связывающий домен Mg^{2+} -зависимой рестриктазы EcoRII и на ВЗ-подобный ДНК-связывающий домен растительных факторов транскрипции. VfiI предположительно возник в результате слияния доменов ДНК-узнающего элемента с неспецифической нуклеазой, родственной Nuc, и использует механизм, ограничивающий расщепление ДНК одним дц разрывом рядом со специфической последовательностью распознавания. Кристаллическая структура предполагает, что междоменный линкер может действовать как аутоингибитор, контролирующий

каталитическую активность VfiI в отсутствие специфической последовательности ДНК [312-314].

1.1.8.2 РНКазы-колицины

Все РНКазы-колицины (ColE3, ColE4, ColE5, ColE6 и ColD) вызывают гибель клеток за счет ингибирования синтеза белка путем расщепления специфических фосфорноэфирных связей в РНК, и как большинство рибонуклеаз не нуждаются в кофакторах ионов металлов. РНКазы-колицины делятся на две отдельные группы: те, которые расщепляют 16S рРНК (рРНКазы) (колицины E3, E4 и E6), и те, которые расщепляют антикодоновые петли определенных тРНК (тРНКазы) (колицины E5 и D). Эти белки не проявляют заметного структурного сходства друг с другом [270, 315].

ColE5 и ColD дают идентичные продукты расщепления с 5'-ОН и 2',3'-циклическим фосфатом [316, 317]. В отличие от классического двухгистидинового кислотно-основного механизма РНКазы А ни колицин E5, ни колицин D не имеют двух незаменимых гистидинов в своем активном центре. Домен колицина E5 вообще не содержит остатков гистидина, в то время как колицин D имеет один незаменимый гистидин (His611). В организации каталитического центра колицина E5 участвуют расположенные на противоположных сторонах каталитической щели Asp46, который действует как предполагаемое общее основание в начальном отрыве протона от 2'-ОН, и Arg48, который участвует в стабилизации переходного состояния [318]. У колицина D в дополнение к His611 были обнаружены две критические аминокислоты: Lys608 и Lys610, вовлеченные, по-видимому, в стабилизацию переходного состояния [319].

Механизмы каталитического расщепления РНК, реализуемые колицинами-рРНКазами изучены недостаточно. Для ColE3 показано, что он состоит из скрученного центрального β -слоя с периферийными α -спиралями, а его каталитический центр содержит три консервативных остатка, необходимых для каталитической активности: Asp510, His513 и Glu517 (эквивалент Asp55) [315].

1.1.8.3 Суперсемейство РНКазы А

Рибонуклеаза поджелудочной железы крупного рогатого скота (РНКазы А) является классическим представителем суперсемейства РНКаз А. Кроме того членами семейства являются ангиогенин (РНКазы 5), эозинофильный катионный белок (РНКазы 3), эозинофильный нейротоксин (РНКазы 2), антипатогенная рибонуклеаза (РНКазы 7) и несколько цитотоксины ооцитов лягушки-быка [320-322].

Большая часть членов суперсемейства РНКазы А идентифицирована у позвоночных. Члены этого суперсемейства обладают большим разнообразием последовательностей, но сходной структурной складкой с характерными тремя или четырьмя дисульфидными связями и консервативной каталитической триадой His-Lys-His [320, 323]. Белки суперсемейства действуют на РНК, катализируя расщепление фосфорноэфирной связи без участия ионов металла с образованием продуктов с циклическим 2'-3'-фосфорноэфирным фрагментом и фрагментом с 5'-концевым гидроксилом (там же).

В 2018 году был описан новый бактериальный токсин CdiA-STYkris, принадлежащий к системе межбактериальной конкуренции *Yersinia kristensenii*. Было показано, что он демонстрирует ту же структурную укладку, что и члены семейства РНКаз А, и проявляет рибонуклеазную активность *in vitro* и *in vivo*. В тоже время этот токсин не имеет сходства последовательностей с РНКазой А, консервативных дисульфидных связей и каталитической триады РНКазы А. Предполагается, что этот белок может быть членом отдельной ветки ферментов суперсемейства [320].

1.1.8.4 Топоизомеразы IV

Топоизомеразы IV катализируют релаксацию сверхскрученной ДНК по механизму «затрудненного вращения»: фермент связывается с ДНК и вносит разрыв в одну из цепей ДНК, 5'-конец разрезанной цепи вращается вокруг второй

цепи ДНК, расслабляя ДНК. Каталитически активный тирозин (Tyr723) запускает каталитический процесс посредством нуклеофильной атаки на остов ДНК, что приводит к разрыву одной из цепей с ферментом, ковалентно присоединенным к 3'-фосфату. Вторая нуклеофильная атака происходит после изменения числа связей, которое управляется 5'-гидрокси концом ДНК, что приводит к восстановлению интактной двухцепочечной ДНК. Все топоизомеразы IV обладают консервативным каталитическим ядром и похожим механизмом катализа [47, 298]. Кроме нуклеазного домена эти ферменты содержат один или более ДНК-связывающих доменов. В каталитический центр нуклеазного домена входит абсолютно консервативный остаток тирозина, который играет роль нуклеофила и формирует фосфотирозильную связь между ДНК и белком. Каталитический центр также содержит два аргинина, один лизин и один гистидин. Основные аминокислотные остатки ориентируют и нейтрализуют ДНК. Один консервативный аргинин, связанный с помощью водорода как с разделительным фосфатом, так и нуклеофильным тирозином, выступает в качестве основания для депротонирования тирозина, второй аргинин, по-видимому, вместе с лизином, служит донором протонов для уходящей 5'-группы [324].

1.1.8.5 Сиквенс-специфические рекомбиназы

При сайт-специфической рекомбинации каждая из двух коротких последовательностей ДНК (сайтов) разрезается в определенных точках обеих цепей, а разрезанные концы присоединяются к новым партнерам. Ферменты, которые опосредуют распознавание сайтов и последующие этапы разрезания и воссоединения, называются сайт-специфическими рекомбиназами (ССР). Большинство ССР попадают в одно из двух семейств в соответствии со сходством их белковых последовательностей и механизмов катализа. Эти семейства известны как тирозин-(Tyr)-рекомбиназы и серин-(Ser)-рекомбиназы, названия которых относятся к консервативному аминокислотному остатку, который атакует фосфодиэфир ДНК и ковалентно связывается с концом цепи ДНК во время

катализа [325]. Типичными представителями суперсемейства тирозиновых рекомбиназ являются фаговые интегразы (λ и HP1), рекомбиназы Cre, XerC/D и FLP, тогда как резольвазы/инвертазы, такие как ϕ C31, $\gamma\delta$ и Tn3, являются членами семейства сериновых рекомбиназ [326]. Два сайта рекомбинации сериновой рекомбиназы различны, тогда как два сайта рекомбинации тирозина идентичны. Это указывает на то, что рекомбинация на основе тирозина в принципе может быть обратимой, тогда как рекомбинация на основе серина является однонаправленной (там же).

Семейство сериновых рекомбиназ состоит из большого числа бактериальных и архейных резольваз/инвертаз, транспозаз и ферментов, разделяющих ДНК, которые используют консервативный Ser в качестве нуклеофила для расщепления и лигирования ДНК с образованием ковалентного промежуточного соединения 5'-фосфосерина. Для рекомбинации ДНК требуется пара длинных сайтов узнавания ДНК, которые расщепляются, и между ними происходит перестройка ДНК. Затем фермент осуществляет обратное лигирование по тем же сайтам [48, 325].

Ser-рекомбиназы содержат связанные длинной линкерной последовательностью N-концевой каталитический домен и C-концевой ДНК-связывающий домен, специфичный к последовательности, и функционируют в димерной форме. Каталитический домен состоит из ~120 N-концевых остатков, которые образуют смешанный пятицепочечный центральный β -лист, окруженный α -спиралями. Нуклеофил Ser находится на конце 1-й β -цепи. Каталитический центр необычен среди фосфотрансфераз тем, что не требует ионов двухвалентных металлов и содержит кластер боковых цепей аргинина, в том числе консервативный аргинин, который действует как общая кислота для протонирования 3'-кислорода во время начальной реакции расщепления. В то же время ион металла необходим для обратной реакции — лигирования [327].

В отличие от тирозиновых рекомбиназ Ser-рекомбиназы не образуют промежуточных структур Холлидея в ДНК-субстрате, осуществляя реаранжировку сходную с изменениями, катализируемыми топоизомеразами типа II, за исключением того, что происходит два расщепления дц ДНК, а не одно. Интересно,

что рекомбиназы Ser имеют структурное и функциональное сходство с топоизомеразами TOPRIM, а прохождение одного дуплекса ДНК через другой при сайт-специфической рекомбинации, осуществляемой Ser-рекомбиназой, не отличается от прохождения двухцепочечной ДНК в топологических изменениях, осуществляемых топоизомеразами типа II [68, 328].

Туг-рекомбиназы обладают консервативным каталитическим мотивом RHR в дополнение тирозину, выступающему в качестве нуклеофила. На основании анализа кристаллических структур этих ферментов было выявлено сходство организации каталитического центра этих ферментов и топоизомераз IV. Каталитический активный центр может состоять из аминокислотных остатков двух смежных субъединиц. Тирозин, выступающий в роли нуклеофила, по-видимому, депротонируется одним из двух консервативных остатков аргинина так же, как и в случае человеческой топоизомеразы и Торо IV вируса коровьей оспы. Другие консервативные остатки аргинина и лизина, располагающиеся в активном сайте, ориентируют разделительный фосфат, стабилизируют переходное состояние, а также могут протонировать уходящую 5'-группу [68]. Реакция, катализируемая Туг-рекомбиназами, включает образование и разрешение промежуточного соединения Холлидея. Дуплексы ДНК рекомбинируют в ходе двух последовательных этапов разрыва цепи и воссоединения. Перенос цепи и лигирование происходят не синхронно, а в ходе двух последовательных стадий отжига. Реакция протекает через ковалентный промежуточный продукт с образованием 3'-фосфотирозиновой связи [68, 329].

1.1.8.6 Другие металлонеинзависимые нуклеазы

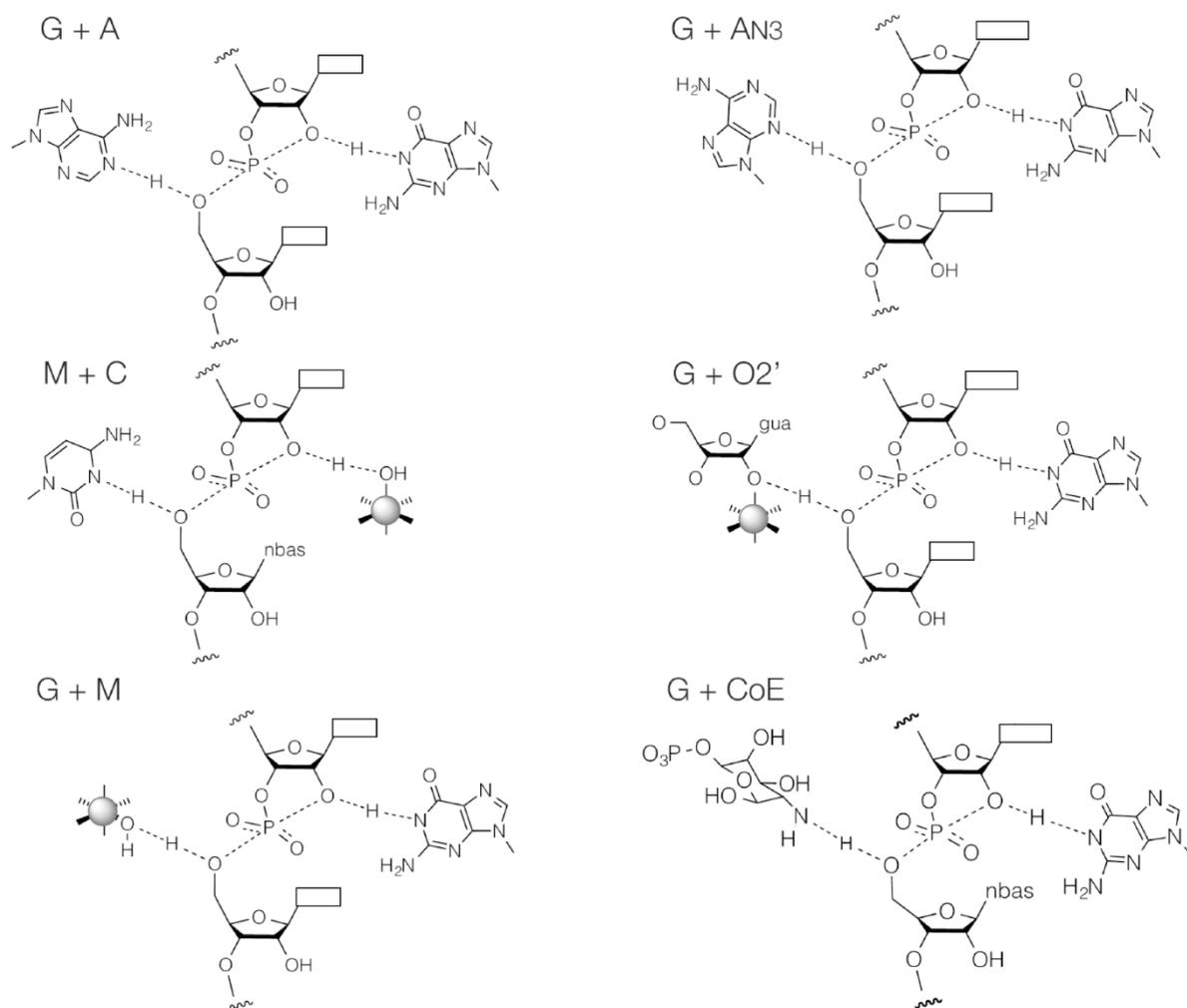
Известно очень мало ферментов-кандидатов, не содержащих металлов, за пределами вышеперечисленных семейств.

Сайт-специфической независимой от ионов металлов нуклеазой является ЭР типа II LlaKI из *Lactococcus lactis*. Она распознает и расщепляет тот же сайт (5'-GGATG-3'), что и FokI, но использует иной механизм катализа [330].

Неспецифическая нуклеаза WSSV-NSN на основании структурных особенностей родственна NucA из *S. marcescens*, которая зависит от ионов металла. Более того, в ее структуре выявляется потенциальный карман для связывания ионов металла. Было показано, что фермент действует не специфически, толерантно к ЭДТА только до концентрации 6 мМ. Поэтому было высказано предположение, что присутствие ионов металлов все же влияет на его каталитическую активность [331, 332]. Аналогичные предположения можно сделать касательно неспецифической нуклеазы из бактериофага, инфицирующего *Geobacillus*. Хотя фермент сохраняет активность при высоких концентрациях ЭДТА, добавление ионов двухвалентных металлов (10 мМ Mn^{2+} или Zn^{2+}) увеличивает его активность до 30% [333].

1.1.9 Каталитические механизмы рибозимов и дезоксирибозимов

Рибозимы, обладающие нуклеазной активностью, осуществляют сайт-специфические реакции расщепления РНК путем нуклеофильной атаки 2'-атома кислорода на соседний атом фосфора используя согласованный кислотно-щелочной катализ. В настоящее время известно девять природных рибозимов. Все они являются эндонуклеазами и образуют на концах продуктов расщепления 2'3'-циклический фосфат и 5'-гидроксильную группу [85, 303, 334-336]. Некоторые рибозимы способны осуществлять и обратную реакцию лигирования [303, 335]. Химическая реакция кислотно-щелочного катализа, реализуемая рибозимами показана на рисунке 1.11 [303, 337]. Рибозим вируса гепатита D (HDV) и TS разрывают экзон-интронные связи двухцепочечной РНК путем катализа с участием ионов металла [88, 338], остальные разрушают избыточную одноцепочечную РНК [85].



Около каждого варианта показана пара «основание + кислота».

Рисунок 1.11. Механизмы катализа нуклеаз-рибозимов (заимствовано из Lilley, 2019)

[303]

Рибозимы используют различные функциональные комбинации для осуществления катализа. На рисунке 1.11 суммированы известные варианты каталитических механизмов нуклеаз-рибозимов [303]. Наиболее распространенным элементом в механизмах рибозимов-нуклеаз является использование гуанина в качестве общего основания. В роли кислоты может использоваться аденин, кофактор GlcN6P, 2'-гидроксил, двухвалентный катион и т.д. Таким образом, общий кислотно-щелочной механизм этих рибозимов можно описать как «G/X» [303, 335]. В качестве разновидности общего кислотно-щелочного механизма у рибозимов HDV и TS цитозин действует как общая кислота, а молекула воды, связанная с ионами металла — как основное основание.

В то же время Теплова и соавт. предлагают другой вариант катализа для рибозимов HDV и TS с участием двух ионов металла, как показано на рисунке 1.12 [339].

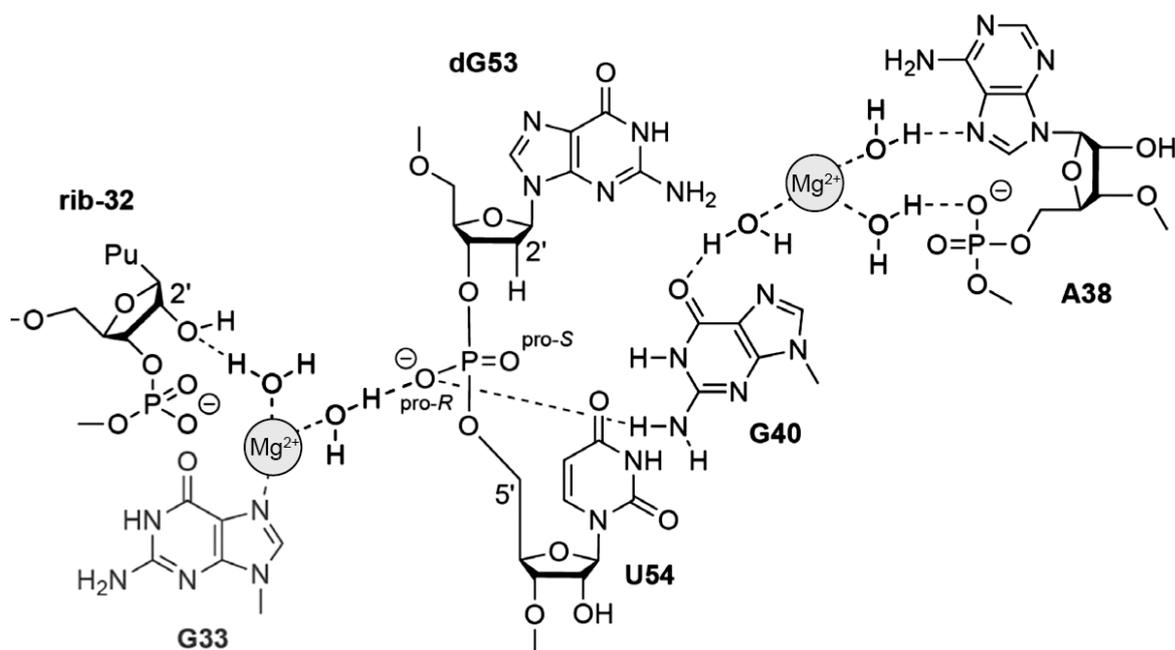


Рисунок 1.12. Предполагаемый механизмы катализа нуклеаз-рибозимов с участием двух ионов металла (заимствовано из Teplova et al., 2020) [339]

Дезоксирибозимы (ДНКзимы) представляют собой молекулы ДНК, обладающие каталитической активностью. В отличие от рибозимов, дезоксирибозимы в природе не обнаружены, однако такие молекулы были получены с помощью эволюции *in vitro* [340]. Первые полученные ДНКзимы расщепляли РНК. Были разработаны дезоксирибозимы с практической способностью к расщеплению почти всех целевых последовательностей динуклеотидов РНК. Большинству, но не всем этим ДНК-ферментам, требовался в качестве кофактора двухвалентный ион металла, такой как Mg^{2+} . Продукты реакции содержали на концах 2'3'-циклический фосфат и 5'-гидроксильную группу [341].

В дальнейшем были отобраны последовательности, способные расщеплять ДНК. Они использовали несколько различных механизмов расщепления, в том числе окислительное расщепление и разрывы цепей в апуриновых участках после

ДНК-катализируемого дегликозилирования. В частности, была получена структурно сложная оц ДНК-последовательность, которая катализировала гидролиз N-гликозидной связи определенного пуринового нуклеозида в ДНК, что приводило к расщеплению цепи. Скорость каталитической реакции депуринизации увеличивалась в $8,8 \times 10^5$ раз по сравнению со скоростью некаталитической реакции в тех же условиях. Было показано, что ДНКзим расщепляет ДНК по последовательному механизму, включающему зависимость от двухвалентного катиона депуринизацию ДНК с последующим ее расщеплением [342]. Были также получены ДНКзимы, которым для активности требовались одновременно как Mn^{2+} , так и Zn^{2+} , а также Zn^{2+} -зависимые расщепляющие ДНК ДНКзимы. Следует отметить, что эти ДНКзимы расщепляли только фосфорамидатную связь, но не амидную или фосфорноэфирную связь субстрата [340, 343-346].

1.2. Использование нуклеаз в биотехнологии и медицине

Сиквенс-специфичные эндонуклеазы, преимущественно ЭР типа II, открывшие эру технологий рекомбинантных ДНК, к настоящему времени стали общепринятым инструментом биотехнологии, что демонстрируют более 600 коммерчески доступных ЭР [314]. Рутинной практикой использования ЭР стало дробление ДНК при приготовлении библиотек ДНК с перекрывающимися фрагментами для геномного секвенирования, направленная делеция генов или областей промоторов, создание экспрессионных конструкторов и др. [5, 116].

Важное место в биотехнологии занимают никазы, которые способны превращать дц ДНК в никованную молекулу, доступную для удлинения ДНК-полимеразами, замещающими цепи этой молекулы. Место разреза можно многократно регенерировать для следующего раунда расширения надрезов без этапа денатурации. Амплифицированный ДНК-сигнал может быть детектирован различными методами: капиллярным электрофорезом, обращенно-фазовой жидкостной хроматографией, использованием ДНК-связывающих флуоресцентных красителей [131, 347]. Кроме природных ферментов был

разработан и создан ряд генно-инженерных нуклеотидных матриц. Такие нуклеотидные матрицы были получены несколькими путями: (1) через нарушение интерфейса димеризации в ЭР, (2) через инактивацию одного из сайтов связывания металла и (3) через ненаправленный мутагенез с отбором вариантов с нужными параметрами [347].

Востребованной альтернативой ПЦР стала изотермическая амплификация ДНК, позволяющая амплифицировать протяженные ДНК-матрицы. Амплификация проводится при одной и той же температуре. Для освобождения участка связывания затравочных праймеров используется нуклеотидаза, которая вносит разрывы в дц-ДНК. Для амплификации ДНК используется полимераза, способная вести синтез ДНК с вытеснением цепи (например, фрагмент Кленова) [348]. Одним из популярных вариантов изотермической амплификации стала амплификация оц ДНК по методу «катящегося конца». В этой схеме амплификации используется кольцевая оц ДНК-матрица, которая содержит последовательность, комплементарную ДНК-мишени нуклеотидазы [349].

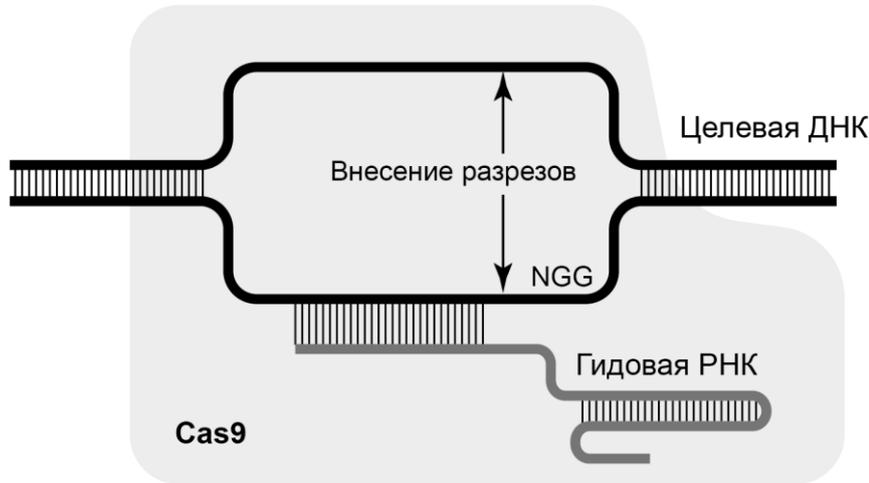
Было также предложено использование нуклеотидаз для ряда технологий обнаружения чужеродной (например, вирусной) ДНК и РНК в биологических образцах и методов детекции однонуклеотидных полиморфизмов [131, 347, 350-352]. В частности, в некоторых из этих методов нуклеотидазы используются для создания последовательность-специфических разрывов в ДНК-мишенях, после чего образовавшиеся оц ДНК-фрагменты становятся доступными для отжига зонда, содержащего 5'-флуорофор и 3'-гаситель флуоресценции [131, 350]. Методы флуоресцентной детекции могут сочетаться с вариантами изотермической амплификации, что приводит к заметному усилению сигнала при выявлении ДНК [351, 353]. В методе хромосомного нанокодирования созданные генно-инженерным путем нуклеотидазы были использованы для внесения специфических для последовательности разрывов; затем область разрыва заполнялась за счет активности ДНК-полимеразы флуоресцентно-мечеными нуклеотидами. Меченую ДНК протягивали через нанощель на подложку и использовали для оптического картирования [352].

За последние два десятилетия возникли новые возможности для целенаправленного создания более специализированных сайт-специфических нуклеаз, таких как ферменты рестрикции и нуклеазы с индивидуально подобранной субстратной специфичностью [352, 354-356]. Полученные генно-инженерным путем нуклеазы цинковых пальцев (ZFN), эффекторные нуклеазы, подобные активаторам транскрипции (TALEN), и CRISPR-Cas9 РНК-управляемые нуклеазы нашли применение не только в научных исследованиях, но и в медицинской практике. ZFN, TALEN и CRISPR/Cas представляют собой три типа сконструированных ДНК нуклеаз, которые способны создавать сайт-специфические двухцепочечные разрывы в точках-мишенях интересующей последовательности ДНК. Это белки, которые состоят из двух частей, включая программируемый и специфичный к последовательности ДНК-связывающий домен, и неспецифический домен расщепления ДНК. ДНК-связывающий домен ZFN, TALEN и CRISPR/Cas является настраиваемым и может быть запрограммирован на распознавание и соединение с любой интересующей последовательностью [354, 357, 358].

Использование нуклеаз *in vivo* и *ex vivo* для редактирования генов прокладывает путь к новым персонализированным методам лечения онкологических заболеваний и хронических инфекций [116, 355, 359, 360]. Сконструированные ZFN и TALEN успешно применяются для генной коррекции в зародышевых линиях, а также в клеточных линиях и в соматических тканях модельных животных [361-366]. После оптимизации конструкций удалось достичь генной коррекции с эффективностью до 18% и 5% в клетках K562 и CD4+ Т-клетках человека, соответственно, при сниженной токсичности [367]. В дальнейшем была продемонстрирована возможность замены домена нуклеазы FokI каталитическим доменом хоминг-эндонуклеазы I-TevI с целью создания сайт-специфической нуклеазы с повышенной специфичностью, а также специфического нуклеазного домена RvuII за счет слияния с эффекторными доменами TAL или ZF [368-370]. Как в ZFN, так и в TALEN домен расщепления ДНК происходит от

фермента рестрикции FokI. У ZFN за связывание с ДНК отвечает набор белков ZF, а у TALEN — ДНК-связывающие белки TALE *Xanthomonas* sp. [371].

Системе CRISPR/Cas уделяется сегодня огромное внимание, благодаря ее способности осуществлять направленную модификацию генома в живых эукариотических клетках [12-16]. Система CRISPR/Cas состоит из одной гидовой РНК и эндонуклеазы Cas (такой как Cas9 или Cas12). Гидовая РНК, состоящая из 20 нуклеотидов действует как ДНК-связывающий домен. Для нацеливания CRISPR/Cas на нужный регион ДНК требуется, чтобы комплементарная гидовой РНК последовательность, известная как мотив «protospacer-adjacent» (PAM), находилась рядом с сайтом ДНК-мишени (рисунок 1.13). Несколько вариантов PAM было найдено в природе, и еще больше разработано с помощью генно-инженерных подходов [372, 373]. Система CRISPR/Cas9 была неоднократно использована для изменения геномов лекарственных растений: для подавления гена синтазы розмариновой кислоты (SmRAS) у *Salvia miltiorrhiza* [374]; подавления гена фермента фарнезилпирофосфатсинтазы и, как следствие, снижения содержания сквалена у *Dioscorea zingiberensis* [375]; для нокаута гена SmCPS1, который является ключевым геном в биосинтезе таншинона у *Salvia miltiorrhiza* [376].



Белок Cas9 из *S. pyogenes* специфически связывается с мишенями ДНК за счет гидовой РНК.

Расщепление дц ДНК происходит с отступом на три нуклеотида в 5'-зону от PAM (NGG, где N — любой нуклеотид).

Рисунок 1.13. Схематическое изображение расщепления ДНК с помощью Cas9 (адаптировано из Gu et al., 2016) [377]

Базовые редакторы CRISPR — новая технология, которая облегчает введение целевых точечных мутаций в ДНК. В этой технологии в качестве нуклеазы используется белок слияния нуклеазы nCas9 (мутант нуклеазы Cas9) и дезаминаза, которая инициирует редактирование оснований, катализируя направленную реакцию дезаминирования. Были разработаны: редакторы оснований цитозина (CBE), которые могут катализировать дезаминирование целевого цитозина для облегчения его превращения в тимин; редакторы оснований аденозина (ABE), которые могут дезаминировать целевой аденозин, чтобы облегчить его переход в гуанозин; редакторы оснований гуанозина (GBE), которые в определенных условиях могут создавать трансверсии C > G [378-380]. Среди первых успешных примеров, демонстрирующих потенциал новой технологии, было использование ABE для восстановления экспрессии дистрофина в мышечной модели мышечной дистрофии Дюшана [381]. В настоящее время ведутся многочисленные разработки применения технологии для лечения других заболеваний [382].

Система CRISPR/Cas9 открыла новые возможности для разработки стратегий выявления различных типов нуклеиновых кислот, включая обнаружение геномной ДНК, негеномной ДНК, РНК и геномов патогенных микробов [358, 383]. Кроме редактирования генома система CRISPR/Cas9 была использована в технологиях удаления нежелательных последовательностей из библиотек кДНК перед секвенированием [377].

Еще одним инструментом генной инженерии стали сайт-специфические системы рекомбинации. В дополнение к природным ферментам был разработан целый спектр синтетических рекомбиназ для осуществления различных манипуляций с геномом [326]. Многие ССР могут функционировать в гетерологичных организмах, где их можно использовать для манипулирования или конструирования генома гетерологичных хозяев. Так система Cre/loxP остается наиболее широко используемой системой для нокаута генов в животных, растениях и микроорганизмах [384]. В зависимости от относительной ориентации спейсеров возможно проведение различных манипуляций, включая опосредованное рекомбиназой удаление, интеграцию, инверсию и даже обмен кассетами экспрессии. Система на основе Cre-рекомбиназы использовалась в международном проекте Synthetic Yeast Genome Project, Sc2.0 [385] для синтетических рекомбинаций хромосом и изменения порядка генетической информации на хромосомах. Система LoxPsym-mediated Evolution (SCRaMbLE) обеспечивает быструю крупномасштабную рекомбинацию генома и способна осуществлять направленную эволюцию генома в селективных условиях [326, 386].

Нуклеазы, не обладающие специфичностью к последовательностям, так же активно используются в различных приложениях. Широкий спектр ферментов с различной специфичностью позволяет проводить очистку ДНК от РНК, или оц от дц НК.

Нуклеазы, специфичные к оц ДНК, нашли применение в качестве аналитических инструментов для исследования гибридизации ДНК-ДНК и ДНК-РНК. Так, нуклеаза S1 активно использовалась для выявления и разрушения дц-участков в оц ДНК, определения сайтов инициации транскрипции и двумерного

анализа сложных транскриптов [91]. Также на основе способности оц-специфичных нуклеаз расщеплять несовпадающие пары оснований в гетеродуплексах ДНК/ДНК и распознавать структурные изменения, индуцированные различными мутагенными агентами, разработаны методики использования их для гетеродуплексного анализа продуктов ПЦР, скрининга генетических мутаций и локально измененных структур в ДНК [387-390]. Кроме того, нуклеазы, специфичные к оц ДНК, используются для отделения дц ДНК фракции НК, например, при очистке плазмид [391, 392], удалении одноцепочечных выступающих концов перед лигированием ДНК [393], анализе структур РНК [91].

Удаление НК из среды при производстве вирусов является хорошей иллюстрацией использования последовательность-неспецифических ДНКаз и РНКаз. Например, при производстве аденовирусов, лентивирусов, вирусов гриппа и вирусоподобных частиц обработка ДНКазами и РНКазами выполняется до или после осветления среды для снижения содержания НК в клетках или на последних стадиях после хроматографической очистки [392, 394-397]. Этот шаг повышает чистоту и безопасность вирусного продукта, а также снижает высокую вязкость, связанную с ДНК, и агрегацию вирусных частиц. Нуклеазы также можно использовать для расщепления векторов плазмидной ДНК, если для получения вирусных частиц используется стратегия временной трансфекции [396, 398].

Неспецифические нуклеазы, ДНКазы и РНКазы также используются для удаления фракций НК из различных препаратов. Так субстратная специфичность по отношению к ДНК делает ДНКазу I эффективным инструментом очистки РНК, тогда как РНКазы A используются для очистки плазмидной ДНК [84, 314, 392]. Основным промышленным применением сахар-неспецифических нуклеаз также является удаление НК при производстве рекомбинантных белков, вирусных вакцин, вирусоподобных частиц или вирусных векторов, которые предназначены для использования в клеточной и генной терапии [392, 399].

Некоторые нуклеазы нашли применение как биомаркеры различных состояний человека и в судебно-медицинской практике. ДНКазы I человека имеет шесть аллелей, типирование которых может использоваться при установлении

отцовства и идентификации личности. Фенотипы ДНКазы I надежно выявлялись в экстрактах из пятен спермы (равных 0,3 мкл жидкой спермы), хранившихся при комнатной температуре до года [400]. Кроме того, ДНКазы I может быть использована в терапии муковисцидоза и как фактор прогноза рака. Рекомбинантная человеческая ДНКазы I (рчДНКазы; дорназа альфа) является коммерчески доступной. У пациентов с муковисцидозом введение рчДНКазы в виде аэрозоля уменьшает обострения респираторных симптомов, что приводит к значительному улучшению легочной функции и хорошо переносится [401]. У больных раком снижение активности щелочной дезоксирибонуклеазы в сыворотке во время лечения обычно ассоциируется с первичным клиническим ответом, тогда как отсутствие снижения указывает на отсутствие ответа на терапию. У больных с полной регрессией опухоли за начальным снижением щелочной дезоксирибонуклеазы в сыворотке обычно следовал рост, превышающий исходный уровень. Подобного повышения не наблюдалось у больных с опухолевой прогрессией. Тесная статистическая связь была обнаружена между пациентами с карциномой желудка и высокой частотой ДНКазы I фенотипа 2. Полученные данные позволяют предположить, что ДНКазы I фенотипа 2 может быть маркером рака желудка [402, 403]. Кроме ДНКазы I ряд нуклеаз, таких как: FEN1, APE1, TREX2, РНКазы L и I, демонстрируют изменение активности и экспрессии при различных онкологических заболеваниях и рассматриваются как биомаркеры рака [404], а топоизомеразы I, II α и II β выступают мишенями имеющихся в настоящее время химиотерапевтических средств против рака [298]. Например, в качестве противоопухолевых препаратов в течение последних 20 лет активно применяются ингибиторы топоизомеразы I – иринотекан и топотекан [405].

Таким образом, нуклеазы относятся к необходимым для жизни ферментам, вовлеченным во все биологические процессы, в которых принимают участие НК. В ряде случаев они могут служить биомаркерами патологических состояний или объектами направленной терапии. Замечательная каталитическая эффективность и широкий спектр специфических характеристик делают нуклеазы очень ценным инструментом для различных биотехнологических приложений.

1.3. Некоторые методы анализа нуклеиновых кислот, требующие выделения одноцепочечной ДНК

1.3.1 Нормализация кДНК и геномной ДНК

Как уже упоминалось выше исследование новых геномов и транскриптомов эукариотических организмов осложняется разницей в представленности различных последовательностей в исследуемых образцах. В типичной эукариотической клетке на основе разницы концентраций различных кодирующих РНК может быть выделено три класса: высокопредставленные, промежуточные и редкие транскрипты [406-409]. Количество видов мРНК в каждом классе и их относительная численность показаны в таблице 1-1.

Таблица 1-1. Представленность различных видов кодирующей РНК в клетке (адаптировано из Carninci et al., 2000 и Gadgil et al., 2002) [406, 407]

Класс мРНК	Количество различных последовательностей	Усредненное количество копий на клетку	Относительное содержание класса мРНК
Высоко-представленные транскрипты	5–20	~12 500	20%
Транскрипты с промежуточной представленностью	~750	~300	40%–60%
Редкие и очень редкие транскрипты	~12 000	10–15	20–40%

Широкий диапазон вариаций в концентрациях клеточных мРНК разных классов осложняет полномасштабные исследования транскриптома. Обычно при

получении РНК для секвенирования используют образец, состоящий из клеток, выделенных из изучаемой ткани/органа. В ходе секвенирования преобладающие транскрипты имеют более высокую вероятность секвенирования несколько раз по сравнению с транскриптами, которые экспрессируются на низких уровнях или экспрессируются в небольшой части клеточных типов ткани и/или на определенных фазах развития и могут составлять небольшую долю от общего количества транскриптов образца. Такие транскрипты могут вообще не попасть в число исследованных молекул при заданной (приемлемой для исследователя по времени и стоимости) глубине секвенирования [409, 410]. В то же время многие из транскриптов, экспрессируемых на низких уровнях и тканеспецифическим образом (включая кодирующие и некодирующие транскрипты), представляют собой важные регуляторы клеточных процессов. Идентификация и понимание роли таких транскриптов является принципиальным для описания функционирования того или иного типа клеток в норме и при патологии [411].

Достижения в области технологий секвенирования последних лет позволили получать большое количество транскриптомных данных за относительно короткое время и по доступной цене. Технологии высокопроизводительного секвенирования включают сегодня как технологии короткого считывания, позволяющие получать короткие (до 600 п.о.) прочтения, так и методы длинного чтения со средней длиной прочтений более 10 т.п.о. [1, 4, 412-414]. Однако качество и скорость транскриптомного анализа, включая профилирование транскриптов и анализ дифференциальной экспрессии, в значительной степени остается зависимым от качества доступного эталонного транскриптома [415]. Изучение динамических изменений в транскриптоме и экспрессии генов, которые являются ключом к пониманию происходящих в клетках биологических процессов, в ответ на внутренние сигналы и сигналы окружающей среды, существенно упрощается при наличии высококачественного эталонного транскриптома. Создание такого эталонного транскриптома, который включает все многообразие изоформ транскриптов, до сих пор представляет сложную задачу. Так у высших растений эта работа осложняется тем, что до 70% интронсодержащих генов продуцируют

различные транскриптные изоформы, образующиеся в результате альтернативного сплайсинга [416-418]. Увеличение охвата редких транскриптов и уменьшение обилия транскриптов с высокой экспрессией может быть достигнуто с помощью выравнивания (нормализации) концентраций разнопредставленных последовательностей РНК в образце на этапе его подготовки к секвенированию.

Еще одной проблемой, в решении которой могла бы помочь процедура нормализации, является подготовка к секвенированию геномной ДНК эукариот. Повторяющиеся элементы могут составлять до 90% в геномах высших растений и до 69% в геномах млекопитающих [18-22]. При этом процент генов и известных функциональных элементов, например, у позвоночных представлен всего ~ 5-10% от генома [19]. Учитывая множество и сложность повторяющихся элементов внутри генома, было предложено разделить их на два больших семейства: IR (insert repeats) и TR (tandem repeats). При этом IR представлены повторяющимися последовательностями, разбросанными по всему геному, тогда как TR располагаются непосредственно рядом друг с другом в определенной локальной области генома. IR представлены в основном мобильными элементами, которых очень много в геноме человека. Однако, в то время как активные TE (Transposable Element) обладают способностью «прыгать» и встраиваться в новое место в геноме, более 99% TE представляют собой элементы древнего генома, не способные к перемещению. IR подразделяются на транспозоны и ретротранспозоны. Ретротранспозоны, в свою очередь, – на длинные концевые повторы (LTR), такие как, например, эндогенные ретровирусы человека (HERV), и прочие элементы, например, длинные повторы LINE и короткие SINE, которые в основном обнаруживаются в перичентромерах хромосом человека. Вместе LINE и SINE составляют до 30% генома человека [419]. TR характеризуются как непрерывные повторяющиеся единицы и имеют очень высокий процент идентичности друг с другом. Считается, что специфическая генетическая структура TR возникает в результате мутагенных процессов, включая рекомбинацию ближнего и дальнего действия и генную конверсию, что делает эти повторяющиеся последовательности одними из наиболее быстро развивающихся в геномах высших эукариот.

Основываясь на длине повторяющихся единиц, TR разделяют на (1) микросателлиты, состоящие из 2-6 п.о., также известные как повторы простой последовательности (SSR) или короткие тандемные повторы (STR); (2) минисателлиты – размером 10-100 п.о., представляющие наиболее вариабельную часть генома человека; (3) сателлитную ДНК длиной до 8 Мб, как в случае некоторых центромер человека; (4) макросателлиты, состоящие из нескольких т.п.о. повторяющихся макросателлитных повторов (MSR); и (5) мегасателлиты — большие области TR, которые, вероятно, возникли в результате неравных событий кроссинговера. Дополнительные классы тандемных повторов существуют в отдельных кластерах, таких как сегментарные дупликации (SD) [18, 420].

Появление технологий NGS в определенной степени решило проблемы, связанные полногеномным секвенированием, в том числе сделав доступным как ре-секвенирование уже известных геномов, так и секвенирование *de novo*. Однако проблема обогащения образцов ДНК целевыми последовательностями перед секвенированием сохранила свою актуальность. В настоящее время аннотации транскриптома немодельных организмов часто оказываются неполными или неточными. А при отсутствии референсного генома обработка данных секвенирования, отягощенных огромным количеством повторяющихся последовательностей, и большое число ошибок, возникающих в ходе пробоподготовки и собственно секвенирования, оказывается нетривиальной задачей, требующей специализированных инструментов биоинформатики и сложных конвейеров постобработки данных для сборки полногеномного проекта [412, 421-423]. Кроме того, недостаточная глубина прочтения приводит к потере значимой информации о редких уникальных последовательностях.

Одним из способов упрощения образцов геномной ДНК до секвенирования является их обогащение уникальными последовательностями с помощью технологии, получившей название «обогащение через захват» (capture-enrichment). Технология предполагает гибридизацию ДНК с биотинилированными зондами с последующим осаждением гибридов на магнитных шариках, покрытых стрептавидином, и секвенированием связавшейся фракции [424]. Технология

позволяет обогащать исследуемый образец митохондриальной ДНК или ДНК, содержащей известные однонуклеотидные полиморфизмы, или иной целевой ДНК и широко используется в исследованиях «древней» ДНК [425-427]. Очевидно, что данная технология применима только в случае обогащения известными молекулами.

Другой подход состоит в использовании, так называемого, таргетного секвенирования. В рамках этого подхода были созданы инструменты для выделения и секвенирования отдельных, интересующих исследователя, частей генома, например, последовательностей экзона или целевых панелей генов [428-430]. Такое таргетное секвенирование позволяет сократить количество необходимых прочтений и увеличить количество образцов в одном эксперименте. Однако эта методика применима в случаях ресеквенирования и анализа родственных геномов, когда биоинформатическая обработка прочтений идет путем сравнения с «референсным» геномом, но не решает проблему выявления редких последовательностей при секвенировании *de novo*.

Были предложены способы упрощения перед секвенированием эукариотической геномной ДНК путем удаления повторяющихся последовательностей, которые часто являются гиперметилированными. Подходы были основаны на использовании систем ЭР, специфичных или чувствительных к метилированию. Было предложено несколько вариантов, различающихся по количеству используемых ЭР: частичная рестрикция гиперметилированных участков [431-432], методы секвенирования фрагментов ДНК, ассоциированных с сайтами рестрикции [433, 434], метод уменьшения сложности полиморфных последовательностей, основанный на рестрикции ДНК в сочетании с эмульсионной ПЦР [435]; метод генотипирования путем секвенирования [436]. Однако вследствие очевидных ограничений этой группы методов, большинство исследователей ориентировались на технологии упрощения образцов ДНК через выравнивание концентраций разнопредставленных последовательностей с помощью нормализации, основанной на кинетике гибридизации НК после денатурации [17].

Гибридизация НК, при которой оц молекулы образуют дуплексы или дц гибриды за счет нековалентных комплементарных взаимодействий, является одним из наиболее фундаментальных процессов в биологии. Гибридизация и плавление комплементарных НК играют ключевую роль в процессах репликации, транскрипции и трансляции в клетках всех живых организмов [437]. В растворе при ренатурации денатурированной ДНК сначала образуются водородные связи в одной паре оснований между двумя комплементарными нуклеотидами реассоциирующих цепей ДНК. Вслед за тем, как произошло формирование «совершенной» пары оснований, происходит быстрый процесс «схлопывания» и образования дуплекса, состоящего из двух одноцепочечных молекулами [438]. Кинетика гибридизации денатурированной гомогенной НК в растворе пропорциональна ее концентрации и представляет собой реакцию второго порядка. Так, скорость гибридизации ДНК в ситуации, где два вида одноцепочечной НК «А» и «В» в растворе образуют дц ДНК «АВ», можно описать формулой:

$$dC_{AB}/dt = k_d \times C_A \times C_B - k_r C_{AB},$$

где C_A и C_B — концентрации комплементарных цепей А и В в нулевой момент времени, C_{AB} — концентрация продукта АВ, k_d и k_r — константы скоростей прямой и обратной реакции, частично определяемые последовательностью и длиной нуклеотидов, t — время реакции [439, 440]. Принимая во внимание, что $C_A = C_B = C_0$ (где C_0 — исходная концентрация дуплекса АВ до денатурации образца), не трудно заметить, что скорость гибридизации НК пропорциональна квадрату её исходной концентрации.

Интегрируя это уравнение, можно получить формулу, описывающую изменение концентрации C_{AB} как функцию времени:

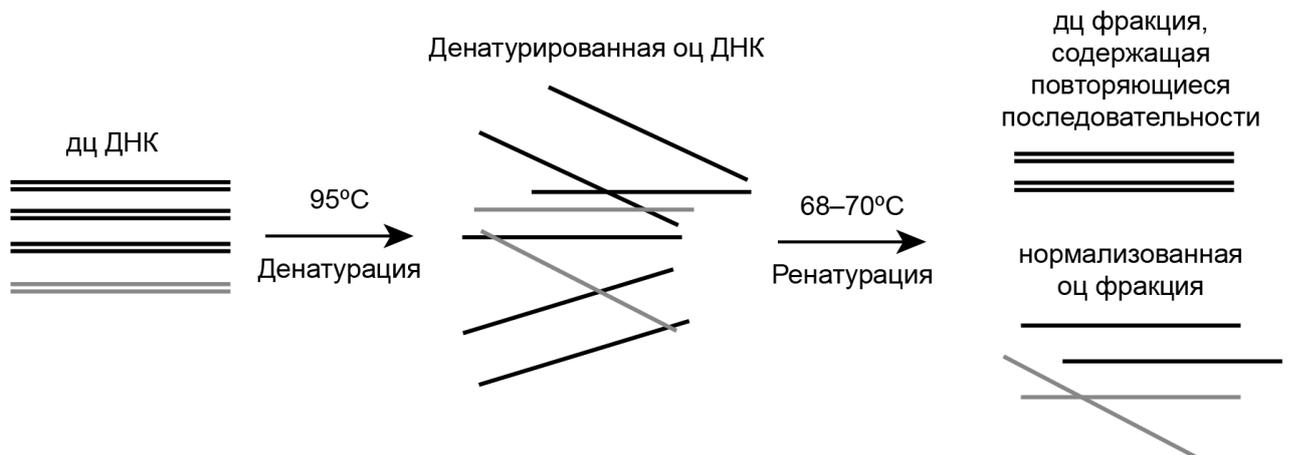
$$C(t) = C_0/[1+kC_0t],$$

где C_0 — исходная концентрация ДНК

Различия в скорости гибридизации последовательностей с различной представленностью в образцах геномной ДНК в растворе принято характеризовать параметром C_0t , равном произведению концентрации нуклеотидов (C_0) в молях на литр на время реассоциации в секундах (t) [441]. В типичном исследовании кинетики ренатурации образцы расщепленной геномной ДНК подвергают тепловой денатурации и оставляют реассоциировать до различных значений C_0t . Для каждого образца ренатурированную ДНК отделяют от одноцепочечной ДНК с помощью хроматографии на гидроксипатите и определяют процент оц ДНК. График зависимости логарифма значения C_0t образца от процента оц ДНК называется кривой C_0t . Математический анализ кривой C_0t позволяет оценить размер генома, долю генома, содержащуюся в однокопийных и повторяющихся компонентах ДНК, и кинетическую сложность каждого компонента [442]. Например, для ДНК теленка было определено значение C_0t фракции медленно реассоциирующей ДНК равное 0,03 и C_0t фракции быстро реассоциирующей ДНК равное 3000, что указывало на 100000-кратную разницу в концентрациях ДНК в этих фракциях (там же). Сходные данные были получены для других геномов эукариот, что позволило сделать вывод о наличии в геномах фракций, с тремя значениями параметров C_0t , которые соответствуют высокопредставленным (сателлитным) последовательностям, последовательностям со средней представленностью и уникальным последовательностям. На основе фракционирования ДНК с помощью C_0t были разработаны методы получения нормализованных библиотек геномной ДНК, которые специфичны для высоко повторяющихся, умеренно повторяющихся и однокопийных последовательностей [420, 443]. Примечательно, что, хотя метод C_0t немного занижал реальное количество повторяющихся последовательностей, вероятно, из-за их медленной ренатурации, он дал глобально точную картину состава генома [420].

Общий принцип нормализации ДНК путем денатурации и реассоциации в растворе показан на рисунке 1.14. В идеальных условиях в ходе гибридизации должно быть достигнуто полное выравнивание концентраций различных

последовательностей в оц фракции. На практике, естественно, такого результат достичь невозможно. Это определяется тем, что на эффективность гибридизации помимо легко адаптируемых параметров, таких как ионная сила и состав буфера, температура гибридизации, существенное влияние оказывает природа и сложность гибридизуемых молекул ДНК. В частности, скорость гибридизации обратно пропорциональна длине реассоциирующих фрагментов и сложности образца [438]. Наблюдается снижение эффективности гибридизации при использовании длинных молекул, склонных к образованию внутримолекулярных дуплексов, это особенно сильно заметно в случае гибридизации оц ДНК с РНК, которая имеет, как правило, множество вторичных структур [439]. Нуклеотидный состав конкретных последовательностей ДНК также сказывается на эффективности их гибридизации. Так высокое содержание GC пар увеличивает скорость реассоциации, а для библиотек с введенными на концы ДНК адаптерами влияние на эффективность гибридизации оказывает последовательность адаптера [17].



Черные прямоугольники — повторяющиеся последовательности, серые прямоугольники — уникальные последовательности.

Рисунок 1.14. Упрощенная схема нормализации фракции оц ДНК в ходе гибридизации

Тем не менее, поскольку высокопредставленные последовательности ДНК после денатурации реассоциируют быстрее, чем редкие, любой метод, который может надежно выделить фракцию оц ДНК, может обогащать малочисленные

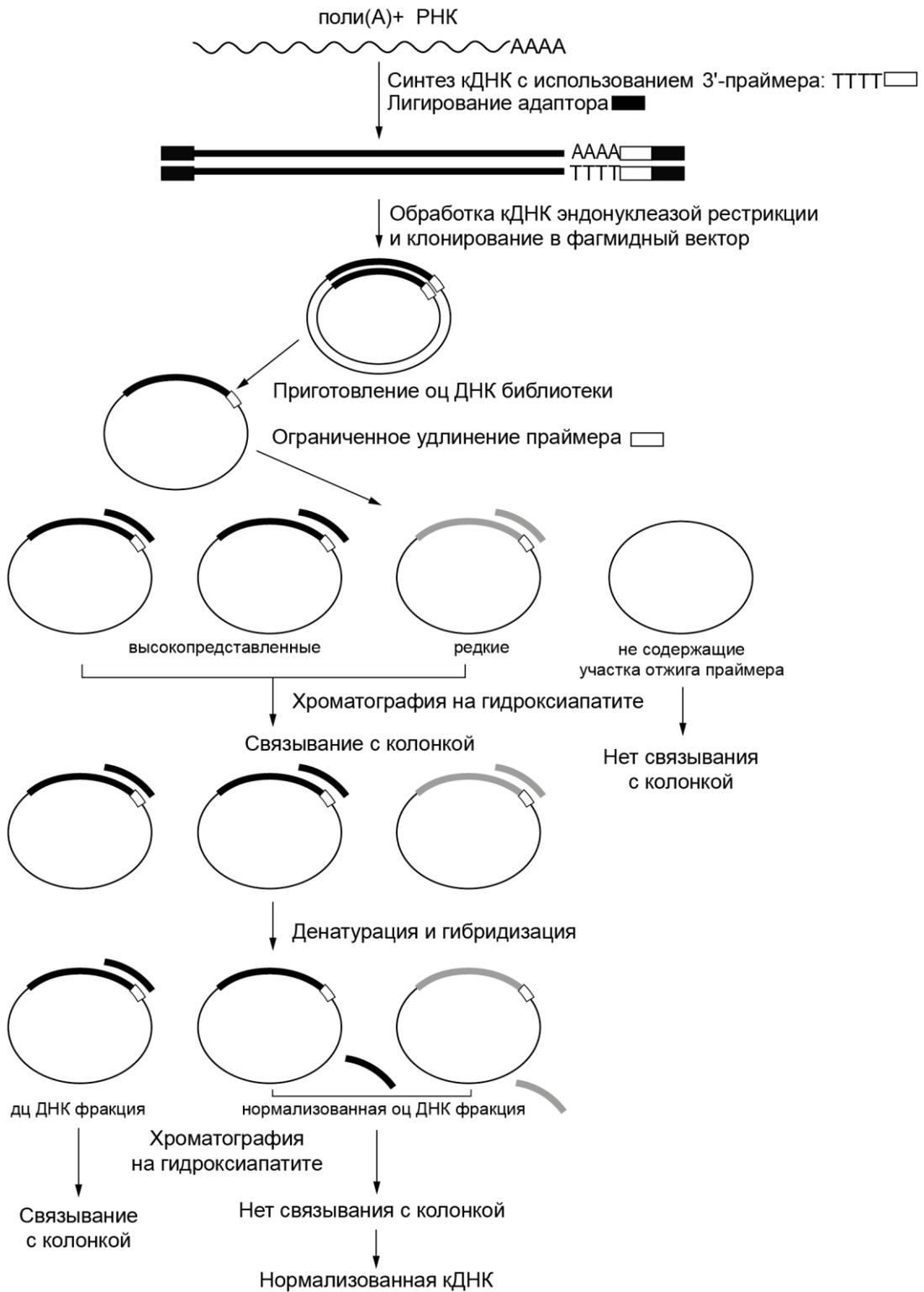
последовательности [443, 444]. В отсутствие ферментов, специфически расщепляющих дц ДНК при температуре гибридизации, для выделения оц ДНК из гибридизационной смеси предлагались различные процедуры, основная часть которых предполагала физическое разделение фракций.

В ряде протоколов нормализации было предложено проводить разделение оц и дц ДНК фракций с помощью хроматографии на гидроксиапатитовых колонках, связывающих дц ДНК [410, 445-447]. В этом случае не связавшаяся с колонкой фракция преимущественно представляет собой оц ДНК, с выровненными концентрациями разных последовательностей. В основном для нормализации использовалась дробленая ДНК или кДНК, последняя была представлена в основном последовательностями, соответствующими 3'-некодирующим областям мРНК. Для достижения адекватного выравнивания образцов часто использовалось несколько последовательных раундов денатурации – гибридизации – хроматографии. Один из предложенных методов ориентировался на обогащение полноразмерными последовательностями кДНК, которые представляют особый интерес при исследовании транскриптома [410]. Схематически метод показан на рисунке 1.15.

Для получения направленной библиотеки кДНК, обогащенной полноразмерными последовательностями, использовался фаговый вектор, который трансформировали в бактериальные клетки вместе с фагом-помощником, который превращал дц плазмиды, содержащие вставки ДНК, в кольцевые одноцепочечные ДНК. Кольцевые оц ДНК выделяли, обрабатывали эндонуклеазой рестрикции для устранения контаминации двухцепочечной репликативной формой ДНК, очищали от дц ДНК хроматографией на гидроксиапатите и подвергали электрофорезу в легкоплавком агарозном геле для удаления ДНК хелперного фага и остаточных примесей тРНК или олигорибонуклеотидов. Оц ДНК вырезали и очищали из геля и использовали для двух последовательных раундов нормализации. На ДНК-вставке как на матрице синтезировали ограниченную по длине (примерно 200 п.о.) вторую цепь ДНК, отбирали дц ДНК на гидроксиапатитной колонке, денатурировали и оставляли

реассоциировать. Негибридизованные одноцепочечные кольца (нормализованная библиотека) отделяли от реассоциированных частичных дуплексов с помощью хроматографии. Хотя предложенный метод был применен в нескольких проектах по секвенированию транскриптомов [448], его трудоемкость, особый температурный режим при хроматографии и необходимость в использовании больших количеств стартового материала требовала поиска других подходов.

Другой подход физического выделения нормализованной оц ДНК фракции состоял в использовании биотинилированной или иммобилизованной на носителе драйверной ДНК [449].



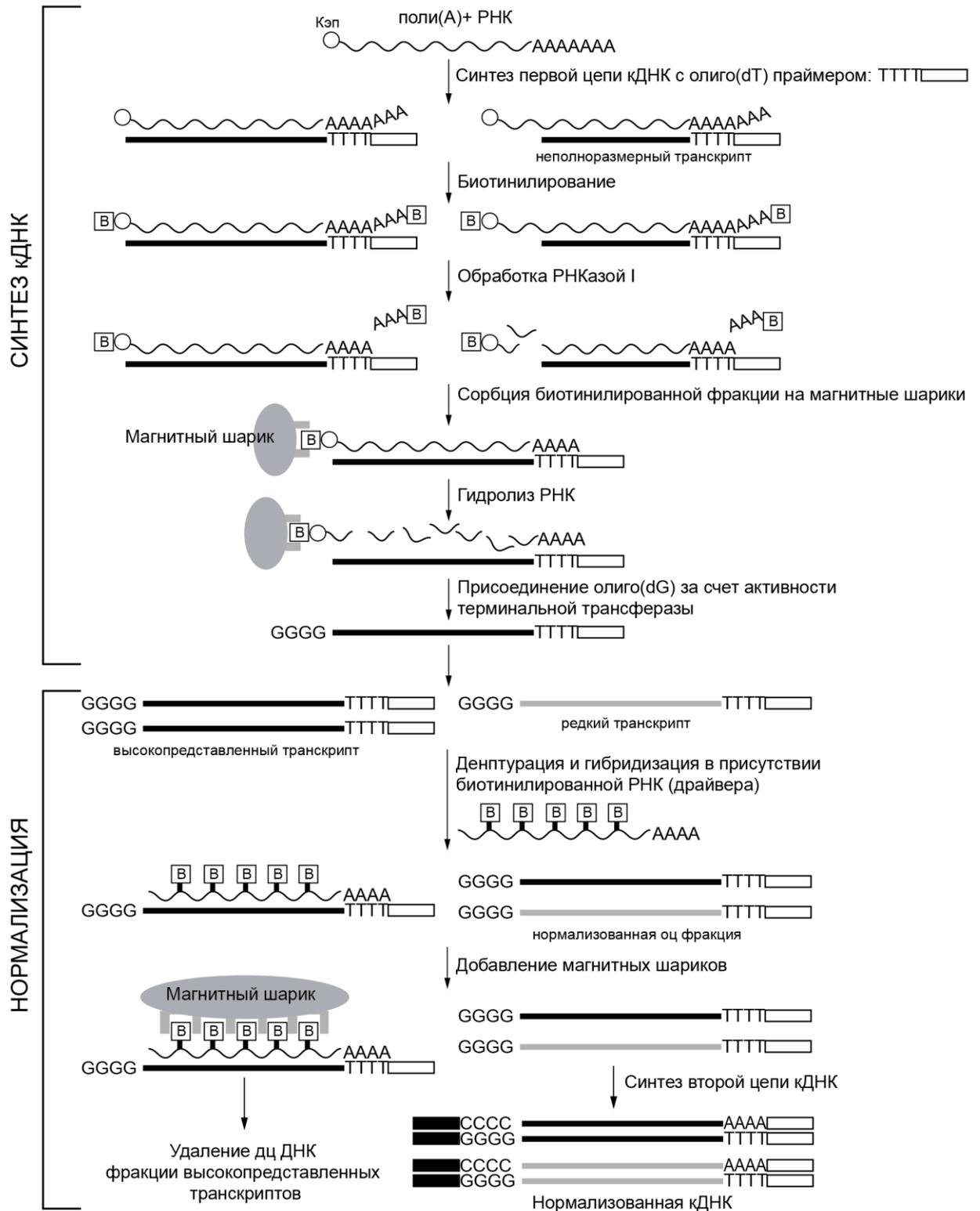
Толстые черные линии — высокопредставленные транскрипты, толстая серая линия — редкий транскрипт. Прямоугольник — адаптер и комплементарная ему последовательность.

Рисунок 1.15. Схема нормализации полноразмерных библиотек кДНК с использованием хроматографии на колонке с гидроксиапатитом (адаптировано из Soares et al., 1994) [410]

В варианте, предложенном Sasaki и соавторами, в качестве драйвера использовалась первая цепь кДНК, синтезированная с олиго(dT) затравки, иммобилизованной на латексных шариках. В ходе нескольких раундов гибридизации мРНК с иммобилизованной ДНК и удаления связавшихся фракций, оставшаяся мРНК использовалась для синтеза кДНК библиотеки. По данным авторов было удалено около 97% мРНК и получен образец кДНК с разностью в представленности транскриптов не более 30 раз [449]. Метод не получил широкого распространения из-за высокого риска потери мРНК в результате деградации в ходе гибридизации, необходимости использования больших количеств мРНК и более низкой скорости гибридизации НК, иммобилизованных на твердой фазе, по сравнению с НК в растворе [406]. Альтернативный подход заключался в использовании биотинилированной РНК в качестве драйвера, который гибридизовали с первой цепью ДНК. Оставшаяся нормализованная кДНК после удаления повторяющихся последовательностей с помощью магнитных шариков, связанных со стрептавидином, использовалась для амплификации и секвенирования (там же). Этот подход позволяет получать библиотеки кДНК со схожими параметрами выравнивания концентраций, что и метод с использованием иммобилизованного драйвера, и так же включает ряд трудоемких процедур: многостадийную процедуру биотинилирования мРНК и несколько раундов гибридизации (рисунок 1.16).

Все вышеперечисленные методы оказались довольно трудоемкими (некоторые требовали значительных навыков) и поэтому плохо подходили для рутинной практики.

Были предприняты попытки обойти стадию физического разделения фракций ДНК после гибридизации, используя энзиматическое расщепление ДНК с помощью ЭР.



Черные линии — высокопредставленные транскрипты, серая линия — редкий транскрипт.

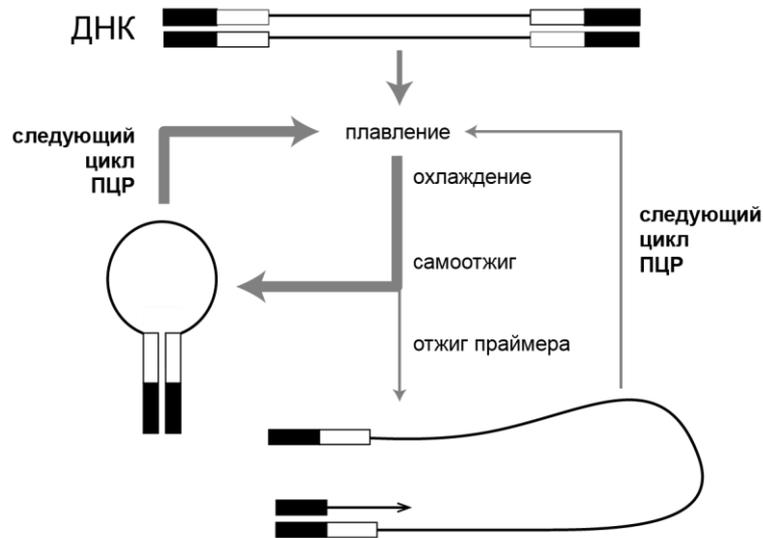
Прямоугольник — адаптер и комплементарная ему последовательность. В — биотин.

Рисунок 1.16. Схема нормализации полноразмерной обогащенной кДНК с использованием биотинилированной РНК (адаптировано из Carninci et al., 2000) [406]

В варианте, предложенном Coche и Dewez, для нормализации использовали библиотеку тотальной кДНК, иммобилизованную на твердой фазе. К свободным концам ДНК лигировали линкерные последовательности, содержащие сайт рестрикции. После цикла денатурации/самогибридизации образец обрабатывали эндонуклеазами рестрикции, которые удаляли линкерную последовательность в дц ДНК. Далее оставшуюся ДНК амплифицировали. При этом молекулы, утратившие линкерные последовательности, не вовлекались в амплификацию, так как не имели сайтов отжига ПЦР-праймеров. Дальнейшие циклы денатурации/самогибридизации проводили в растворе [450, 451]. Методика оказалась низкоэффективной вследствие заметной потери разнообразия нормализованной библиотеки, возможно, из-за того, что обработка ЭР велась при температуре 37°C, когда ДНК легко образует дополнительные дуплексы и шпильки.

Еще один подход по изолированию нормализованной оц ДНК фракции был разработан на основе эффекта селективной супрессии ПЦР [452, 453]. Суть этого эффекта заключается в том, что молекулы, содержащие на концах длинные инвертированные повторы, не подвергаются амплификации в ПЦР с использованием праймера, который комплементарен внешней части этого повтора (рисунок 1.17). Причиной этого является то, что при охлаждении реакционной смеси после денатурации самоотжиг инвертированных концевых повторов происходит при более высокой температуре, чем отжиг праймера. Таким образом большая часть молекул с такими повторами не амплифицируются в ходе ПЦР.

Для нормализации с использованием эффекта селективной супрессии ПЦР используется дробленая с помощью ЭР дц кДНК. Предварительно образец кДНК разделяют на две части, к каждой на 5'-конец лигируют свой адаптер, обеспечивающий формирование на концах ДНК инвертированного повтора. Образцы денатурируют и оставляют реассоциировать.

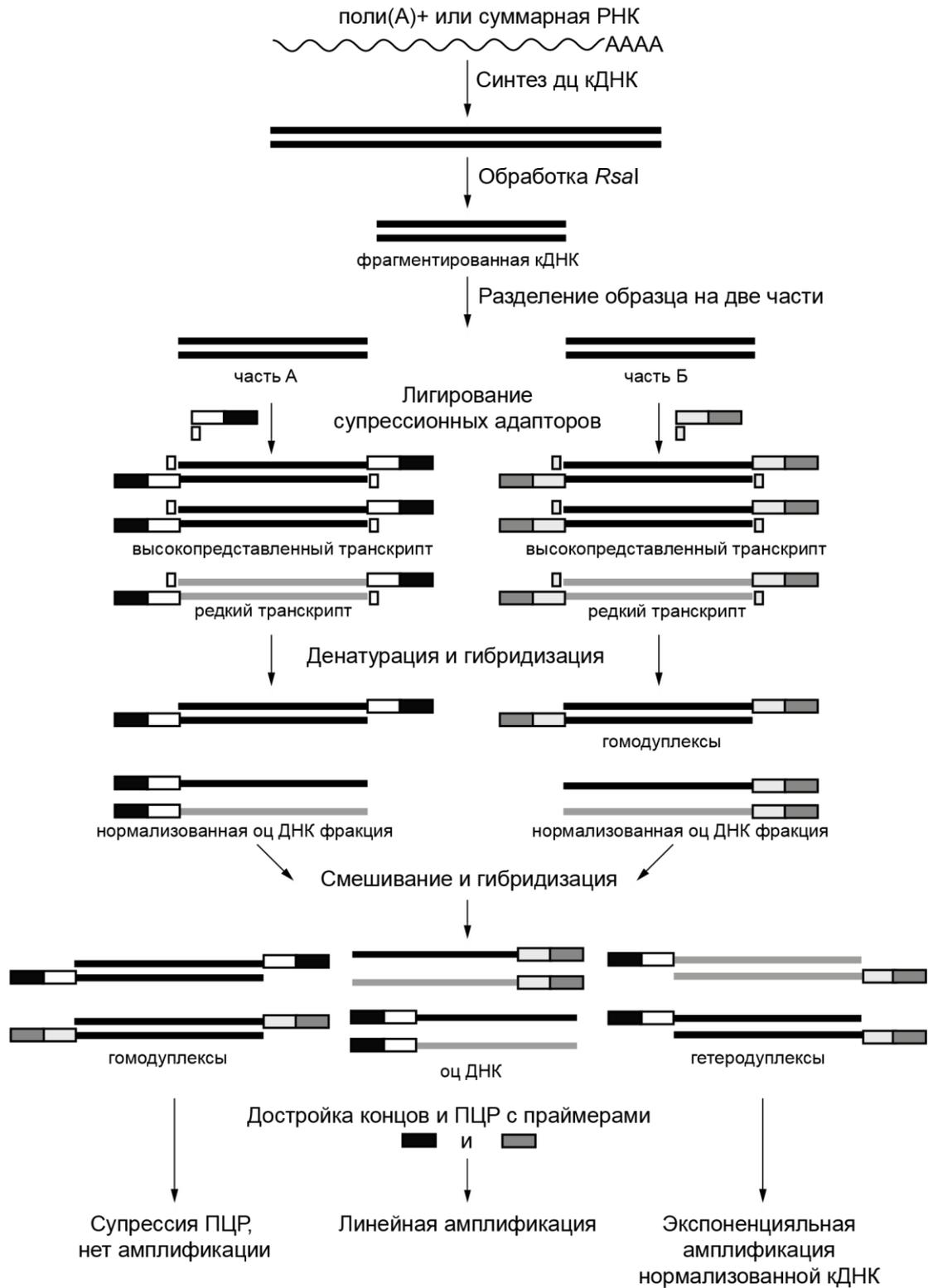


Инвертированные концевые повторы изображены в виде прямоугольников. Черным отмечена последовательность праймера, используемого в ПЦР и комплементарная ей последовательность.

Рисунок 1.17. Эффект селективной супрессии ПЦР (адаптировано из Siebert et al., 1995) [452]

В каждой гибридизационной смеси формируется дц ДНК с одинаковыми адаптерами на обоих 5'-концах (гомодуплексы) и нормализованная оц ДНК. Далее образцы смешивают и продолжают гибридизацию. Теперь в растворе оставшаяся оц фракция формирует дц ДНК, которая содержит разные адаптеры на 5'-концах (гетеродуплексы).

После гибридизации аликвоту образца используют для ПЦР. До начала ПЦР происходит достройка цепей и на концах гомодуплексной дц ДНК образуются инвертированные концевые повторы. Таким образом, в ПЦР с праймерами, которые комплементарны внешним частям адаптеров, амплифицируется только нормализованная фракция гетеродуплексов. Схема метода нормализации показана на рисунке 1.18. Метод нормализации кДНК на основе эффекта селективной супрессии ПЦР является довольно простым и воспроизводимым. К его недостаткам можно отнести невозможность нормализовать полноразмерные библиотеки кДНК.



Черные линии — высокопредставленные транскрипты, серая линия — редкий транскрипт.

Прямоугольник — адаптер и комплементарная ему последовательность.

Рисунок 1.18. Схема нормализации кДНК с использованием эффекта супрессии ПЦР (адаптировано из Лукьянов и др., 1999) [453]

1.3.2 Удаление нецелевых транскриптов из библиотек кДНК

рРНК составляет около 80–90% от тотальной РНК клетки [454-456]. В большинстве случаев эта популяция молекул не представляет интереса для исследователей. Информация о последовательностях рРНК может быть отфильтрована из данных секвенирования на стадии биоинформатического анализа, однако оптимальным является удаление этих последовательностей из популяции на стадии подготовки образцов к секвенированию [17, 456, 457].

Кроме необходимости удалять рРНК время от времени возникает потребность в удалении иных нецелевых транскриптов. Высокопредставленные транскрипты конкурируют с редкими РНК, препятствуют их обнаружению и скрывают их биологическую функцию, например, при функциональном скрининге, особенно в случаях сходных механизмов действия [458]. В частности, изоформы транскриптов часто уникальны для разных типов клеток, тканей, стадий развития и организмов. Они возникают в ходе процессинга РНК за счет использования альтернативных сайтов начала транскрипции, сплайсинга и полиаденилирования. Появление таких изоформ способствует разнообразию клеток и организмов у высших эукариот, адаптации и видообразованию, а также может отражать определенные патологические состояния [459-461]. Возможность проведения комбинированного анализа дифференциальной экспрессии транскриптома на немодельных организмах осложнена отсутствием их аннотированных полных транскриптомов и низкой доступностью биоматериалов. Например, это относится к охраняемым и находящимся под угрозой исчезновения видам, для которых возможно лишь получить образцы крови [462-464]. Однако поли(А)⁺ РНК, выделенная из цельной крови, может содержать до 76% транскриптов гемоглобина [465]. Проблема удаления высокопредставленных нежелательных транскриптов, наблюдается и в области молекулярной клинической диагностики. В случае диагностики опухолей по анализу крови или при анализе внеклеточных НК эта проблема стоит наиболее остро. Дополнительно, при профилировании рака доля мутантных последовательностей, происходящих от опухоли, может быть

исчезающе малой на фоне последовательностей дикого типа из-за обилия иммунных клеток или распределения опухолевых клеток внутри нормальной ткани [466, 467].

Для удаления нецелевых последовательностей используются несколько методик. Часть из них разработана прицельно для удаления рРНК, другие могут быть адаптированы под различные исследовательские задачи.

Селекция поли(А) РНК путем гибридизационного захвата с использованием олиго(dT)-связанной твердой фазы или амплификация первой цепи кДНК с помощью олиго(dT)-праймера давно стали рутинными процедурами, позволяющими избавиться образец от рРНК и незрелых или не полностью процессированных гетерогенных ядерных РНК [17, 468]. Однако очевидным недостатком этого подхода является удаление из рассмотрения всех остальных поли(А)-негативных РНК, которые потенциально могут представлять интерес, например, некодирующих РНК, поли(А)-негативных транскриптов биморфных генов, недавно транскрибированных транскриптов [469-471]. Также известно, что поли(А)-последовательность отсутствует у гистоновых мРНК, мРНК HEG1 и DUX [472]. Кроме того, как правило, поли(А) последовательности в большинстве случаев отсутствуют в бактериальных транскриптах, что делает указанный подход неприменимым к бактериальным транскриптомам [473]. Дополнительно, следует отметить, для деградированной РНК (особенно в образцах, фиксированных формалином и залитых парафином, отбор поли(А)⁺ фракции позволяет изолировать только 3'-часть транскрипта [470, 474].

Очевидным решением проблемы может быть использование нормализации кДНК перед секвенированием, которая обсуждалась выше. Однако, выравнивание концентраций всех транскриптов не применимо, например, в случаях, когда исследователя интересует дифференциальный анализ генной экспрессии.

Альтернативная тактика, которая использовалась с целью деплеции рРНК, — неслучайное праймирование гексамерным праймером. Так, с помощью компьютерного вычитания гексамеров, комплементарных рРНК, из библиотеки случайных гексамерных праймеров, был отобран сет из 749 «не совсем случайных

гексамеров», который в условиях высокого содержания солей может избирательно затравлять синтез первой цепи кДНК не являющихся рРНК транскриптов [475]. Как было показано позднее, нечувствительность к одному или двум несовпадениям в сайте связывания праймера обратной транскриптазы позволяет использовать около 50 праймеров, которые дают приемлемое покрытие транскриптома при секвенировании. Однако около 10% прочтений все равно относились к цитоплазматической рРНК, что продемонстрировала невысокую специфичность данного подхода [476]. К тому же, требуется тщательный дизайн для создания набора праймеров.

Популярным способом является удаление (деплеция) последовательностей рРНК с использованием смеси специфичных к их последовательностям зондов, иммобилизованных на твердом носителе. Однако недостаточная эффективность такого отбора приводит, как правило, к худшему результату, нежели описанная выше поли(А)-селекция [17, 477, 478]. В качестве альтернативы было предложено использовать более эффективную гибридизацию НК (РНК или одноцепочечной первой цепи кДНК) в растворе с мечеными биотином зондами, комплементарными нежелательным последовательностям, с последующим удалением фракции связавшихся с зондами молекул с помощью магнитных шариков, связанных со стрептовидином [467, 470, 479, 480]. Недостатки такого подхода уже обсуждались выше, в главе 1.3.1. В другом варианте, для селективного удаления рРНК фракции связавшейся с зондами использовали коллекцию антител против человеческой и митохондриальной рРНК [457]. Этот подход дает лучшие результаты, чем описанные выше способы. Однако применим только при наличии специфических антител против нецелевых последовательностей интересующего исследователя организма.

Другая группа методов базируется на энзиматическом расщеплении нецелевых последовательностей.

Был предложен метод с использованием РНКазы H и ДНКазы I, приводящий к высокому уровню деплеции нецелевых последовательностей. Тотальную РНК смешивают с набором олигонуклеотидных праймеров к последовательностям,

которые должны быть удалены, и проводят синтез первой цепи в присутствии этих праймеров. Затем реакционную смесь обрабатывают РНКазой H, что приводит к селективной деградации РНК в РНК/ДНК дуплексах. Обработка ДНКазой I используется для удаления праймеров и первой цепи кДНК, полученной на первом этапе [454, 481-483]. В модификации этого метода вместо синтеза первой цепи кДНК до обработки РНКазой H используется гибридизация тотальной РНК с набором олигодезоксирибонуклеотидных зондов к нежелательным последовательностям [484].

Другим многообещающим методом является технология DASH (истощение обильных последовательностей путем гибридизации), в котором деплеция достигается путем ферментативного расщепления последовательностей рекомбинантной нуклеазой Cas9 при добавлении в образцы гидовых РНК [377]. Вариант этой технологии был применен для удаления полноразмерных транскриптов гемоглобина из образцов цельной крови, что позволило выявить более 6000 изоформ и ряд новых генов у белых медведей *Ursus maritimus* [458].

Важно отметить, что технологии с использованием РНКазы H и DASH не применимы в случаях работы с малыми количествами РНК, извлеченными из одной или очень ограниченного числа клеток [17].

1.3.3 Анализ генетических полиморфизмов и выявление целевых последовательностей НК в биологических образцах

Однонуклеотидный полиморфизм (ОНП) — изменение (замена, вставка, делеция) одного нуклеотида в определенном положении в геномной ДНК — является одной из наиболее распространенных наследственных вариаций в геномах различных организмов. У человека описано более 9 миллионов ОНП [485, 486].

Большинство ОНП расположены в некодирующих областях генома и, хотя они и не изменяют кодируемые белки, эти ОНП служат важными генетическими или физическими маркерами для сравнительных или эволюционных исследований. В кодирующих областях ОНП могут вызывать изменения в структуре белка и,

следовательно, в его функционировании, что может быть причиной развития заболевания или приводить к изменению ответа на лекарство или токсин окружающей среды. Такие ОНП используются в качестве молекулярных маркеров во многих исследованиях генетики заболеваний и фармакогеномных исследованиях [487].

С развитием методов высокопроизводительного секвенирования выявление мутаций с их помощью стало одним из популярных подходов. NGS активно используется для обнаружения генетических вариантов, таких как ОНП, небольшие вставки/делеции, структурные вариации и вариации числа копий, которые могут выявить индивидуальные различия по сравнению с эталонным геномом. Следует отметить, что методы NGS в применении к ОНП нуждаются в аккуратном биоинформатическом анализе данных с фильтрацией ошибок прочтения и пробоподготовки, и требуют программного обеспечения (ПО) высочайшего качества, что в ряде случаев затрудняет их применение, особенно для клинических приложений [488]. Поэтому все еще остаются востребованными иные методы анализа, способные осуществлять высокоспецифическую детекцию ОНП.

Методы генотипирования можно разделить на пять основных типов: (1) методы, основанные на кинетике гибридизации ДНК; (2) методы, основанные на использовании технологий удлинения праймера (primer extension); (3) методы, основанные на использовании технологий лигирования; (4) методы, использующие нуклеазы; (5) методы разделения молекул на основании их конформационных особенностей.

Многие методы, используемые для анализа ОНП, так же нашли применение при детекции целевых последовательностей (например, микроорганизмов или биомаркеров заболеваний) в различных биологических образцах. И наоборот, многие методы, предложенные для детекции целевых последовательностей в биообразцах, затем были применены для анализа ОНП. Так как в обоих случаях задачей является выявление целевой ДНК (или РНК)-мишени в образце с высокой разрешающей способностью, можно объединить эти методы в одну группу.

Ниже приведены некоторые примеры, иллюстрирующие разнообразие используемых подходов к анализу ОНП и детекции целевых последовательностей в образцах НК.

1.3.3.1 Технологии на основе кинетики гибридизации НК

Температура плавления (T_m) представляет собой элементарную термодинамическую характеристику ДНК, которая может изменяться за счет различий в нуклеотидной последовательности, содержании GC и длине ампликона. Анализ кривой плавления, особенно анализ кривой плавления с высоким разрешением (High Resolution Melting, HRM), широко используется для обнаружения как известных, так и неизвестных ОНП. Метод основан на ПЦР в реальном времени (РВ-ПЦР) с использованием одного или нескольких зондов и интеркалирующего флуоресцентного красителя. После ПЦР температуру реакции снижают для отжига одноцепочечной ДНК, а затем повышают. Диссоциация насыщающего интеркалирующего флуоресцентного красителя из двухцепочечной ДНК по мере ее плавления приводит к снижению флуоресцентного сигнала, при этом уровень флуоресценции оказывается пропорционален количеству оставшейся двухцепочечной ДНК. Изменение флуоресценции указывает на сдвиг T_m , который виден при анализе кривой плавления [489]. Основываясь на вариациях флуоресценции ДНК-связывающего красителя, HRM-анализ позволяет обнаружить наличие ОНП и подтвердить тип нуклеотидной замены [488, 490]. С помощью HRM-анализа можно выделить четыре класса ОНП с разными изменениями T_m . Базовые замены оснований G/A и C/T составляют первый класс ОНП, тогда как G/T и C/A образуют второй класс. Третий и четвертый классы состоят из ОНП C/G и A/T соответственно. ОНП первого и второго классов могут быть четко генотипированы с помощью HRM из-за их высоких различий T_m , равных $0,5^\circ\text{C}$. ОНП третьего класса вызывает разницу T_m в $0,4^\circ\text{C}$, в то время как разница в $0,3^\circ\text{C}$ возникает для ОНП четвертого класса (там же). Метод, сочетающий HRM с баркодированием ДНК, известный как Var-HRM, оказался

успешным для анализа видовой разнообразия, например, количественного определения примесей близкородственных видов в агропищевых продуктах. В этой методологии в качестве ПЦР праймеров и зондов используются короткие ортологичные последовательности ДНК, характерные для определенных видов, которые авторы назвали «баркодами» [491]. В отличие от обычного РТ-ПЦР, Var-HRM позволяет проводить количественный анализ и увеличивает разрешающую способность обычного анализа кривой плавления [492].

Молекулярные маяки (molecular beacons, MB) представляют собой короткие олигонуклеотидные зонды, имеющие структуру шпильки и содержащие флуоресцентный репортерный краситель (R), например, 5-FAM (5-карбоксихлорофлуоресцеин), VIC или TET (тетрахорофлуоресцеин) на 5'-конце и соединение, гасящее флуоресценцию (Q), такое как TAMRA (тетраметилродамин) на 3'-конце. В отсутствие ДНК-мишени зонд имеет низкий (фоновый) уровень флуоресценции из-за FRET (Фёрстеровский перенос энергии флуоресценции). При связывании зонда с комплементарной ДНК-мишенью он линейаризуется, и R и Q оказываются пространственно разделены, что приводит к возрастанию эмиссии флуоресценции. ДНК-мишени. Не полностью комплементарные зонду последовательности не образуют с зондом гибридов, поэтому при наличии в ДНК-мишени ОНП не происходит роста эмиссии флуоресценции (рисунок 1.19).

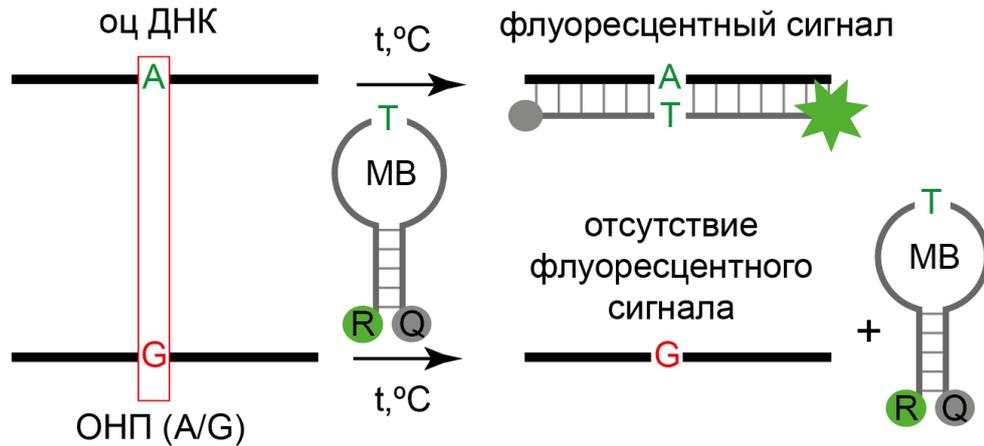


Рисунок 1.19. Схематическое изображение технологии детекции ОНП с использованием молекулярных маяков (адаптировано из Tuagi et al., 1998) [493]

Для одновременной детекции аллельных вариантов возможно использование в одной реакционной смеси нескольких МВ с различными R [493, 494]. Специфическая структура МВ и пара флуорофор/гаситель позволяют идентифицировать целевые НК в режиме реального времени с высокой чувствительностью и селективностью. МВ не активны в отсутствие конкретной мишени, просты в использовании и пригодны для различных модификаций. Модифицированные МВ, такие как 2'-О-метилованные МВ, МВ на основе пептидных нуклеиновых кислот (PNA), МВ на основе заблокированных нуклеиновых кислот (LNA) и МВ-аптамеры обладают высокой стабильностью и используются для детекции различных мишеней, таких как ОНП, мРНК, катионы, тяжелые металлы, ДНК, микроРНК (миРНК) и так далее [495, 496].

Принцип COLD-PCR основан на использовании критических температур денатурации (T_c) последовательностей ДНК, которые ниже, чем их T_m . Эффективность амплификации ПЦР для последовательности ДНК резко падает, если температура денатурации устанавливается ниже ее T_c . T_c сильно зависит от последовательности ДНК и определяется эмпирически для каждой конкретной последовательности. В технологии COLD-PCR для выявления ОНП используется промежуточная температура отжига во время цикла ПЦР, которая обеспечивает

перекрестную гибридизацию мутантных аллелей и аллелей дикого типа. Образуются гетеродуплексы, которые плавятся при более низких температурах, чем гомодуплексы. Далее осуществляется селективная денатурация гетеродуплексов и амплификация при T_c , во время которой гомодуплексы остаются двухцепочечными и эффективно не амплифицируются. Этот подход позволяет селективно обогащать образцы ДНК редкими аллелями, отличающимися одним или несколькими нуклеотидами в любом положении последовательности [497]. Для идентификации аллелей обычно используют прямое секвенирование или гибридизацию с флуоресцентно-мечеными зондами, такими, например, как MB [488].

Технология использования ДНК-чипов, содержащих коллекции зондов к различным аллельным вариантам или целевым последовательностям, является высокопроизводительной и, как правило автоматизированной, процедурой. Зонды гибридизуются с образцом ДНК, кДНК или РНК для подтверждения специфического аллеля ОНП, выявления целевых последовательностей или для анализа генной экспрессии. Специфичность анализа достигается за счет разницы в температурах гибридизации совершенных и несовершенных НК дуплексов [498]. ДНК чипы можно использовать в генетических исследованиях, включая геномную селекцию, картирование генов, анализ структуры популяции и построение карт сцепления [488]. Обработка данных требует знаний в области биоинформатики, однако, она существенно проще, чем анализ данных NGS. ДНК чипы плохо применимы при поиске ОНП в полиплоидных и сложных геномах (там же).

1.3.3.2 Технологии на основе реакции удлинения праймера

Анализ ОНП на основе технологий удлинения праймеров подразумевают аллельное распознавание за счет специфичности фермента, осуществляющего аллель-специфическое включение нуклеотидов в ходе удлинения праймера на ДНК матрице. Используется либо общий праймер для обнаружения обеих аллелей, либо специфические праймеры для обнаружения каждой аллели.

В первом случае используется праймер с 3'-концом, прилегающим к сайту ОНП, для удлинения которого используется ДНК полимеразы [499]. Вариант включенного в ходе реакции нуклеотида, определяющего вариант ОНП, выявляется либо по флуоресценции, либо по массе. Выбор праймеров под определенную нуклеотидную замену и дизайн анализа достаточно просты и позволяют одновременно обнаруживать несколько ОНП, благодаря чему данный подход реализован во многих коммерческих наборах. Анализ PinPoint использует дидезоксирибонуклеотидтрифосфаты (ddNTP) для однонуклеотидного удлинения праймера и является самым простым среди методов, основанных на этом подходе [500]. Он не требует какой-либо химической модификации нуклеотидов или праймеров. Детекция продукта основана на матрично-активированной лазерной десорбции/ионизации масс-спектрометрии в полете (matrix assisted laser desorption/ionization time-of-flight mass spectrometry, MALDI-TOF MS). Схематически метод показан на рисунке 1.20.

Низкое разрешение спектральных пиков было основным ограничением такого подхода. Это было преодолено благодаря использованию ddNTP с массовой меткой для увеличения разницы в массе между аллелями [501]. В методе MassEXTEND™ (Sequenom, CA) была использована смесь обычных нуклеотидтрифосфатов и ddNTP для удлинения праймера, что позволило увеличить разницу в массе между продуктами удлинения, соответствующим обоим аллелям [502]. Был предложен ряд вариантов этого подхода, которые отличаются модификациями нуклеотидов, используемых в реакции удлинения праймера. Также были предложены ddNTP, которые присоединены к фрагменту биотина, что позволяет экстрагировать продукт реакции из смеси НК и упрощает дальнейший анализ [503-506].

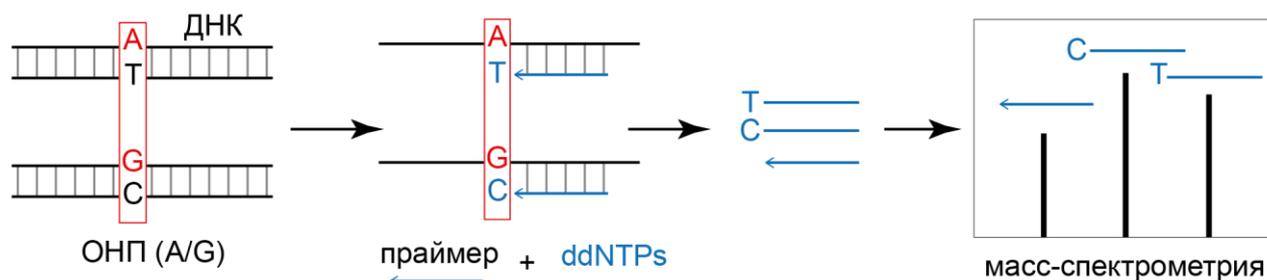


Рисунок 1.20. Схематическое изображение метода анализа ОНП PinPoint (адаптировано из Kim & Misra, 2007) [487]

В некоторых модификациях этого метода используется детекция продуктов реакции с помощью флуоресценции. В этих случаях в реакцию удлинения праймера добавляют флуоресцентно-меченные ddNTP. Варианты анализа были реализованы в различных форматах, включая реакцию в растворе и детекцию флуоресценции; реакцию в растворе с последующим захватом продуктов удлинения на твердой фазе и детекцией флуоресценции; или реакцию, опосредованную твердой фазой [487, 507-509].

Группа методов, основанная на использовании аллель-специфических праймеров для амплификации фрагментов ДНК, примыкающих к ОНП, получила название аллель-специфической ПЦР (АС-ПЦР). В реакции используется термостабильная ДНК-полимераза, у которой отсутствует корректирующая экзонуклеазная активность, два (или более) по-разному меченных аллель-специфических праймера, структуры которых отличаются одной 3'-концевой буквой, и общий встречный праймер. В идеальных условиях удлинение аллель-специфического праймера происходит только тогда, когда он связывается с сайтом ОНП с корректной комплементарностью (рисунок 1.21). АС-ПЦР можно использовать как в классическом формате, так и в режиме РТ-ПЦР. Продукт ПЦР детектируют с помощью различных методов, например, электрофоретическим разделением продуктов реакции или по флуоресценции, специфической для определенного ОНП [487, 510, 511].

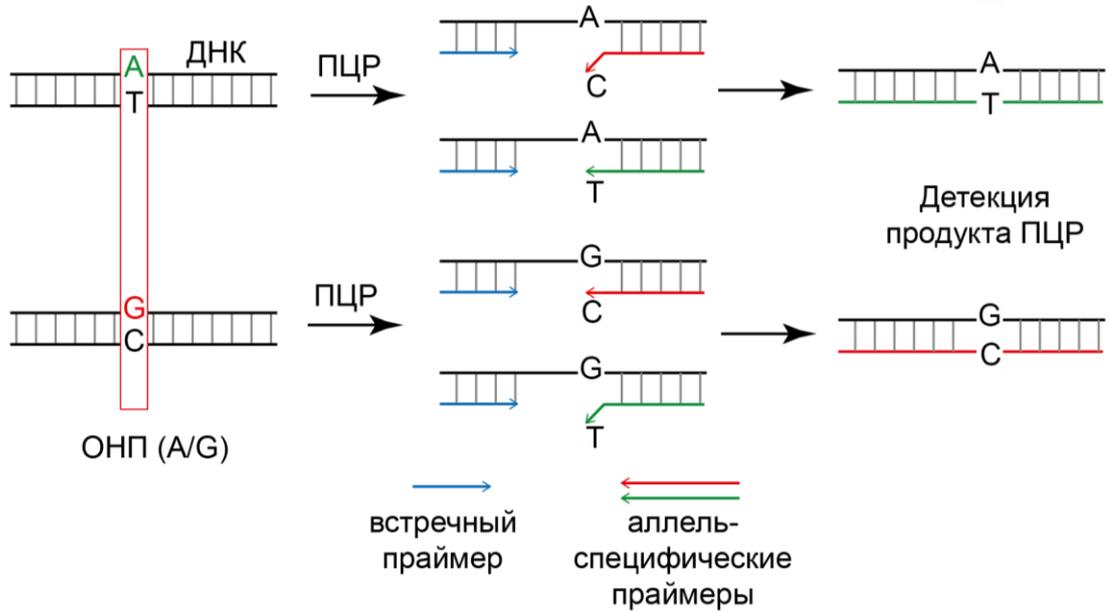


Рисунок 1.21. Классическая схема аллель-специфической ПЦР (адаптировано из Kim & Misra, 2007) [487]

Другим широко применяемым методом ПЦР-детекции ОНП является генотипирование с использованием зондов TaqMan. Принцип TaqMan-анализа основан на способности Taq-полимеразы с 5'→3' экзонуклеазной активностью расщеплять флуоресцентный зонд, комплементарно связанный с дц ДНК-матрицей. Процедура TaqMan начинается с ПЦР в присутствии прямого и обратного праймеров для амплификации интересующей полиморфной последовательности и двух аллель-специфических зондов, меченых флуоресцентным красителем (R) на 5'-конце и содержащих гаситель флуоресценции (Q) на 3'-конце. В процессе ПЦР ДНК денатурируют, а при температуре отжига праймеры и зонды одновременно связываются со своей мишенью. Каждый уникально помеченный зонд предпочтительно связывается с одним из двух аллелей интересующего ОНП с различной аффинностью. Флуоресцентный сигнал генерируется из-за расщепления Taq-полимеразой связавшегося с ДНК зонда. Гомозиготность по одной аллели определяется по флуоресцентному сигналу первого красителя, по второй аллели — по флуоресцентному сигналу второго красителя, в то время как наличие обоих

флуоресцентных сигналов указывает на гетерозиготность по исследуемой аллели ОНП [512]. Схематически метод показан на рисунке 1.22.

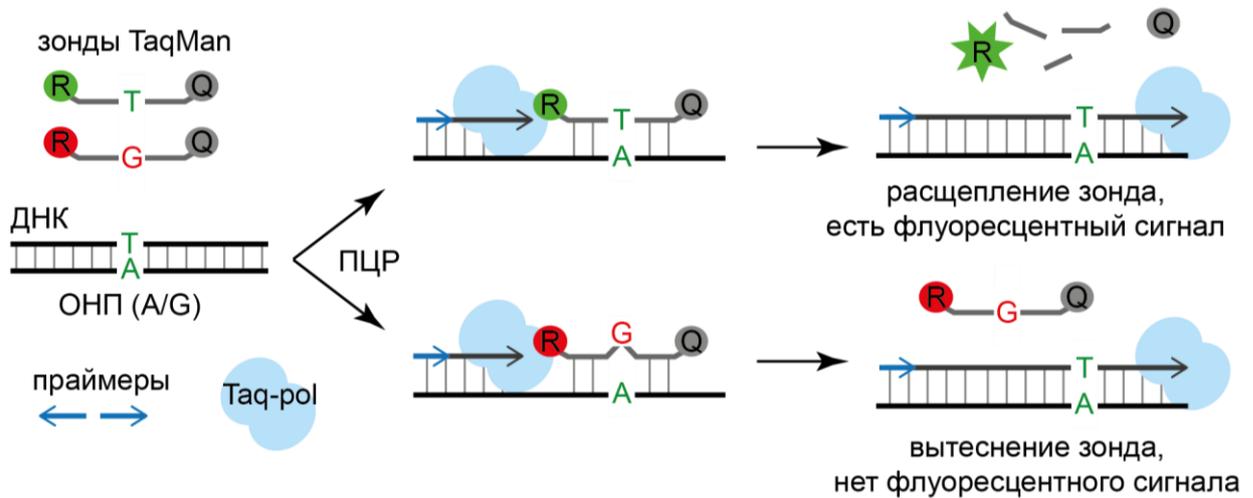


Рисунок 1.22. Анализ ОНП с помощью проб TaqMan и Taq-полимеразы (Taq-pol) с 5'→3' экзонуклеазной активностью (адаптировано из Holden & Wang, 2008) [512]

Методы, основанные на изотермической амплификации, применяются как для обнаружения патогенов, так и для генотипирования, и используются в качестве альтернативы ПЦР. Метод петлевой изотермической амплификации (LAMP, loop-mediated isothermal amplification), представляет собой вариант изотермической амплификации, осуществляющей реакцию в присутствии четырех различных праймеров. Комбинируя праймеры, распознающие несколько последовательностей целевой области, и добавив в реакцию дополнительную пару «петлевых праймеров», можно наработать большое количество ДНК за короткое время [513].

1.3.3.3 Методы, основанные на использовании нуклеазной активности

Рестрикционный анализ продуктов ПЦР — технология, основанная на использовании ЭР, которыерезают ДНК по уникальному сайту рестрикции, создавая несколько фрагментов ДНК разного размера. В случае, если в сайте узнавания/рестрикции находится ОНП, размер возникших фрагментов позволяет выявлять аллельные варианты [514]. Метод применяется, например, при анализе

мутаций серповидноклеточной анемии, где замена основания А на Т в гене β -глобина, приводящая к замене глутаминовой кислоты на валин, разрушает сайт распознавания рестриктазы MstII. В этом случае нуклеотидная замена легко выявляется после амплификации участка ДНК, включающего мутацию, и инкубации с соответствующей ЭР [515]. Другие применения рестрикционного анализа продуктов ПЦР включают выявление эпидемиологической паразитарной инфекции, диагностику *Helicobacter pylori*, *Giardia lamblia*, анализ некоторых характерных мутаций при канцерогенезе [516]. Однако использование рестрикционного анализа сильно ограничено, так как не все ОНП находятся в сайтах рестрикции ЭР [517].

Обойти это ограничение удалось с помощью применения системы РНК-гидовых инженерных нуклеаз (RNA-guided engineered nucleases, RGEN) на основе Cas9 и TALEN, с помощью которых были разработаны новые методики выявления различных типов нуклеиновых кислот, включая обнаружение геномной ДНК, негеномной ДНК, РНК и геномов патогенных микробов [358]. По сути, сиквенс-специфическое взаимодействие комплекса «Cas9-гидовая РНК» с ее двухцепочечной ДНК-мишенью представляет собой гибридизацию РНК-ДНК с участием белка. В отличие всех методов на основе гибридизации и ПЦР, этот процесс происходит при низкой температуре, что позволяет избежать разрывов ДНК-матрицы и проблем неспецифичной температурно-зависимой гибридизации НК. Высвобожденные фрагменты ДНК-мишени могут быть лигированы с универсальным адаптером и амплифицированы в ПЦР [518, 519].

Также была разработана группа методов обнаружения мутаций с использованием нуклеаз Surveyor и T7E1, специфически разрезающих одну цепь ДНК в месте нарушения комплементарности. В ходе анализа осуществляется ПЦР гетерозиготных аллелей, за которой следует тепловая денатурация и медленный отжиг, создающий ДНК популяцию, содержащую гетеродуплексы. Ферментативное расщепление этих гетеродуплексов с помощью нуклеазы приводит к появлению фрагментов, свидетельствующих о нуклеотидных несоответствиях в образце [390, 520]. Методы на основе нуклеаз, чувствительных

к нарушениям комплементарности в ДНК, активно используются для генотипирования индуцированных мутаций при геномном редактировании [519].

Был предложен анализ ОНП по технологии Invader®, основанной на расщеплении ДНК структурно-специфической эндонуклеазой. В данном варианте используется три зонда для генотипирования ОНП: два аллель-специфических зонда, меченных FRET парой «R и Q», и общий зонд «инвадер», который комплементарен 3'-области полиморфного сайта и имеет несовпадение с матрицей в ОНП-сайте. При этом, аллель-специфические зонды комплементарны 5'-области полиморфного сайта и содержат ОНП-специфический нуклеотид и выступающий нуклеотид на 5'-концах. При гибридизации с матрицей ДНК «инвадер» и полностью комплементарный аллель-специфический зонд образуют трехмерную структуру, которая распознается структурно-специфической эндонуклеазой. Фермент расщепляет выступ аллель-специфического зонда в сайте ОНП, что приводит к возникновению флуоресцентного сигнала. Второй аллель-специфический зонд остается целым и не дает флуоресцентного сигнала [487, 521].

1.3.3.4 Технологии на основе лигирования

Методы выявления геномных вариаций на основе специфичности ДНК-лигазы изначально были предложены для выявления делеций и инсерций. Однако они оказались применимы и в случае ОНП. В классическом исполнении метода детекции ОНП с помощью лигирования используются два аллель-специфических зонда, которые связываются с матрицей в ОНП-сайте и общий встречный олигонуклеотид, примыкающий к ОНП непосредственно рядом с аллель-специфическим зондом. В случае, если два соседствующие друг с другом олигонуклеотида гибридизуются с одноцепочечной матричной ДНК с полной комплементарностью, лигаза соединяет их, образуя один олигонуклеотид. В большинстве методов лигирования используются аллель-специфические зонды с 3'-концами в сайтах ОНП, так как лигазы более чувствительны к несовпадениям на 3'-конце. Чтобы определить нуклеотид в ОНП-сайте, анализируют продукты

лигирования. Для визуализации результатов лигирования было предложено использовать метки с комбинаторным переносом энергии флуоресценции (CFET), состоящие из флуоресцентных красителей, передающих энергию, когда они находятся в непосредственной близости друг от друга. Различные профили эмиссии флуоресценции создаются с помощью ограниченного набора красителей за счет варьирования их количества и расстояния между ними. Для генотипирования ОНП используются два аллель-специфических зонда, меченные разными метками CFET, и общий биотинилированный зонд. После реакции лигирования продукт отделяют с помощью стрептавидина и анализируют электрофорезом на капиллярной матрице или иным способом. Показано, что метод хорошо работает на образцах из свежей или замороженной ткани, но плохо применим к образцам из тканей, фиксированных формалином [522-524].

Интересной разработкой для анализа ОНП являются многоцветные флуоресцентно-меченые самолигирующиеся зонды, содержащие гасители флуоресценции. Когда такие зонды связываются в соседних сайтах на ДНК-мишени, они образуют между собой ковалентную связь вследствие химической реакции аутолигирования. При этом происходит удаление гасителя флуоресценции и появление флуоресцентного сигнала. Такие зонды находят применение как в экспериментах *in vitro*, так и в живых клетках микроорганизмов [525].

1.3.3.5 Методы на основе анализа конформационных изменений

Методы, основанные на анализе конформации, включают в себя анализ одноцепочечного конформационного полиморфизма и анализ гетеродуплексной подвижности.

Технология анализа гетеродуплексной подвижности основаны на том, что гетеродуплексы и гомодуплексы ДНК имеют различную подвижность. При смешивании ДНК дикого типа и мутантной ДНК, денатурации и последующего медленного схлопывания образуются гомо- и гетеродуплексы, которые могут быть разделены по подвижности в полиакриламидном геле. Гомодуплексы мигрируют

через гель со значительно более высокой скоростью, чем гетеродуплексы. Скорость движения гетеродуплесов зависит от их структурных характеристик. Большие вставки или делеции (более 3 п.н.) создают очень стабильные формы. Эта стабильность нечувствительна к условиям электрофореза и позволяет легко обнаруживать такого рода мутации. Однако гетеродуплексы с заменами одного основания нестабильны и сверхчувствительны к изменениям окружающей среды. Действительно, для некоторых однонуклеотидных замен различия в подвижности между гомо- и гетеродуплексами минимальны или даже отсутствуют. Это снижает чувствительность метода к данному типу мутаций [526].

Основная концепция анализа одноцепочечного конформационного полиморфизма основана на начальном разделении дц НК на одноцепочечные формы при нагревании. В разделенном состоянии молекулы имеют тенденцию складываться в трехмерную конформацию в соответствии с последовательностями нуклеотидов. Таким образом, при электрофорезе в полиакриламидном геле отдельные молекулы ДНК могут иметь одинаковый размер, но образовывать разные трехмерные структуры с разной подвижностью. Замена нуклеотида (мутация) в первичной последовательности может изменять физическую конформацию денатурированных одноцепочечных молекул и, следовательно, их подвижность, что приводит к изменению картины расположения полос в нейтральных полиакриламидных гелях [526, 527]. Недостаток этого метода заключается в том, что в настоящее время отсутствует теоретическая модель для предсказания точной конформации ДНК при различных параметрах, таких как: мутация, размер фрагмента ДНК, содержание G и C, плотность геля, концентрация ДНК, ионная сила и pH. Кроме того, отмечается низкая воспроизводимость метода при анализе серий образцов [528].

Денатурирующая высокоэффективная жидкостная хроматография (ДВЖХ) позволяет выявлять мутации в смеси денатурированных и повторно отожденных ампликонов ПЦР путем дифференциального удержания гомо- и гетеродуплексной ДНК на носителях для обращенно-фазовой хроматографии при частичной

денатурации. Как и в случае HRM-анализа чувствительность метода зависит от температуры. Исследование проводится с помощью УФ- или флуоресцентного мониторинга неочищенных ампликонов размером до 1500 п.о. в течение 2–3 минут. Кроме мутационного анализа, метод ДВЖХ оказался полезным в исследованиях эволюции человека и картирования генов. В полностью денатурирующих условиях, ДВЖХ применяется для анализа известных полиморфизмов за счет использования способности поли(стирол-дивинилбензола) разделять одноцепочечные молекулы ДНК идентичного размера, отличающиеся одним основанием. В условиях полной денатурации удается разрешить все возможные замены оснований, за исключением трансверсий C→G [529]. Однако ДВЖХ имеет ограничения, поскольку не может определять конкретные типы мутаций или непосредственно обнаруживать гомозиготные мутации. Как следствие, ДВЖХ был в значительной степени заменен другими методами обнаружения [488].

Приведенные примеры иллюстрируют разнообразие подходов, применяемых для детекции ОНП, анализа целевых последовательностей в образцах ДНК и РНК и других изменений генома. В некоторых методах используются комбинации двух или более подходов. Например, в технологии BeadArray™ (Illumina, CA) для выявления ОНП сочетаются гибридизация, удлинение праймера и лигирование с последующей амплификацией аллель-специфического продукта [530]. Выбор способа детекции ОНП, целевых последовательностей или анализа их разнообразия определяется, как правило, для каждой исследовательской задачи с учетом оборудования и компетенций конкретного коллектива.

ГЛАВА 2. МАТЕРИАЛЫ И МЕТОДЫ

2.1. Олигонуклеотиды

В данной работе использовали олигонуклеотиды, входящие в состав наборов реактивов, или синтезированные на заказ в ЗАО Евrogen (Россия). Олигонуклеотиды проходили обязательную стадию очистки с помощью высокоэффективной жидкостной хроматографии.

2.2. Выделение и очистка НК

Здесь и далее в данной работе всегда использовали стерильную, свободную от нуклеаз деионизованную воду, полученную с помощью системы очистки milliQ (Германия).

Выделение суммарной РНК осуществляли с использованием реактива TRIzol (Invitrogen, США). Для выделения суммарной РНК использовали свежие ткани гепатопанкреаса ракообразных, свежие ткани коралловых полипов Zoantus, клетки линии MCF-7 человека, бактериальные клетки *E. coli*, а также мононуклеарные клетки человека, выделенные из свежей периферической крови здорового добровольца с помощью седиментации в градиенте плотности фиколл-урографина по стандартному протоколу [531].

Для выделения плазмидной ДНК использовали набор реактивов QiaPrep Spin (Qiagen, США). Для очистки ДНК из реакционных смесей (продукты ПЦР, лигирования, обработки ЭР) — наборы реактивов Qia-quick PCR purification kit (Qiagen) или QJ Aquich (Qiagen). Образцы деплецированной кДНК для приготовления библиотек очищали с использованием колонок CHROMA SPINTM-400 (Clontech). Процедуры осуществляли в соответствии с приложенными инструкциями производителей. В некоторых случаях для очистки НК использовали

фенольно-хлороформную экстракцию с осаждением этанолом, как описано Matz M.V. (2003) [532].

2.3. Секвенирование ДНК

Для автоматизированного секвенирования использовали автоматический секвенатор серии 454 GS FLX Standard (Roche, Швейцария). В каждом случае использовали по ~3 мкг каждого образца ДНК. Предварительно проводили фрагментацию ДНК и лигирование адаптера, как описано в Cheung et al. (2006) [533]. Секвенирование осуществляли в центре «Биоинженерия» РАН.

Подготовку образцов и секвенирование по методу Сэнгера с использованием удлиненного праймера, меченного полинуклеотидкиназой T4, проводили в соответствии со стандартным протоколом [534].

2.4. Синтез первой цепи кДНК

Для получения первой цепи кДНК использовали суммарную РНК, выделенную, как описано выше, а также суммарную РНК и поли(А)⁺ РНК из различных тканей человека (Clontech) и суммарную РНК из *N. furzeri*, предоставленную Институтом исследования возраста им. Лейбница (Германия).

Синтез первой цепи кДНК осуществляли на основе 1 мкг суммарной РНК или 0,5 мкг поли(А)⁺ РНК. Перед забором аликвот размороженной РНК образцы прогревали при 65°C в течение 1 мин для предотвращения агрегации РНК и перемешивали содержимое, осторожно встряхивая пробирку.

В работе использовали наборы реактивов, адаптированные для получения образцов кДНК, обогащенных полноразмерными последовательностями: SMART RACE (Clontech, США), SMARTScribe (Clontech) или Mint-Universal (Евроген). В некоторых случаях вместо адаптеров, указанных производителем набора реактивов, использовали адаптеры, указанные на рис. 3.37. В случае кДНК прокариот вместо олиго(dT) содержащего CDS-3 праймера использовали праймер

со случайной гексамерной последовательностью: 5'-CAGTGGTATCAACGCAGAGTACG(N)10.

2.5. Амплификация и клонирование ДНК

Для ПЦР (если не указано иное) использовали наборы реактивов Advantage 2 PCR Kit (Clontech) или Encyclo PCR Kit (Евроген, Россия). Для приготовления амплифицированной кДНК, обогащенной полноразмерными последовательностями, использовали наборы реактивов SMARTScribe (Clontech), Advantage 2 PCR Kit (Clontech) или Encyclo PCR Kit (Евроген, Россия) с наборами праймеров, обозначенных при описании соответствующих процедур в главе «Результаты и обсуждение».

Реакционная смесь содержала 1 × реакционный буфер для ПЦР, 200 мкМ dNTP, 0,3 мкМ каждого ген-специфического праймера, ДНК матрицу в конечной концентрации 20-40 пг/мкл, и рекомендованное производителем количество смеси ДНК полимераз. При амплификации фрагментов ДНК для анализа генов человека в образцах кДНК использовали праймеры из набора Human Amplimer Set (Clontech). Для оценки размеров вставок клонированных библиотек кДНК использовали плазмидные M13 праймеры. Структуры остальных ген-специфических и выродженных праймеров, использованных в амплификации, а также температура отжига (T_0) приведены в таблице 2-1.

Таблица 2-1. Структуры использованных праймеров

Праймер	Последовательность (5'→3')	Назначение	T_0
3Pr1	ССТ CAG TGG CA(G/A) GCT TT(C/T) AAT	3'-RACE нуклеазы из камчатского краба. Первая ПЦР	64°C
3Pr2	CAG GCC TTT AAT AAT GG(T/A/G/C) AA(T/C) TGG	3'-RACE нуклеазы из камчатского краба. Заглубленная ПЦР	64°C

5Pr1- Par_DSN	GGC CAG GTC TCG GGT CGC	5'-RACE нуклеазы из камчатского краба. Первая ПЦР	64°C
5Pr2- Par_DSN	GGG TCG CGT ATT CTA GGT A	5'-RACE нуклеазы из камчатского краба. Заглубленная ПЦР 1	64°C
5Pr3- Par_DSN	CCA TTA TTG AAG GCC TGC CA	5'-RACE нуклеазы из камчатского краба. Заглубленная ПЦР 2	64°C
AFN2	CCTCAGTGGCA(G/A)GCTTT(C/T) AAC	3'-RACE гомологов нуклеазы из камчатского краба. Первый ПЦР	60°C
GNW1	CAGGCCTTTAATAATGGTAATT GG	3'-RACE нуклеаз Pal_Nuc и Uca_Nuc. Заглубленная ПЦР	64°C
5Pr1- Uca_Nuc	GGATTGCCATTAATGTCGTC	5'-RACE нуклеазы Uca_Nuc. Первая ПЦР	64°C
5Pr2- Uca_Nuc	CCACTGTACACCCGAAGGTC	5'-RACE нуклеазы Uca_Nuc. Заглубленная ПЦР 1	64°C
5Pr3- Uca_Nuc	AACCAAGGCTCGCCAAGTCC	5'-RACE нуклеазы Uca_Nuc. Заглубленная ПЦР 2	64°C
5Pr1- Gam_Nuc	CAATGGTCCGAATTCTGTTCTC	5'-RACE нуклеазы Gam_Nuc. Первая ПЦР	64°C
5Pr2- Gam_Nuc	GTGACTACGCGCAGAGTGGC	5'-RACE нуклеазы Gam_Nuc. Заглубленная ПЦР	64°C
5Pr1- Pal_Nuc	CCAGCACTCCCAACCTCC	5'-RACE нуклеазы Pal_Nuc. Первая ПЦР	64°C

5Pr2-Pal_Nuc	GTCAGGTCAGTGCCGTGGGC	5'-RACE нуклеазы Pal_Nuc. Заглубленная ПЦР	64°C
FC1-Uca_Nuc	GGATTGCCATTAATGTCGTC	Амплификация полной кодирующей последовательности Uca_Nuc	60°C
FC2-Uca_Nuc	CCACTGTACACCCGAAGGTC		
FC1-Pal_Nuc	CCAGCACTCCCCAACCTCC	Амплификация полной кодирующей последовательности Pal_Nuc	60°C
FC2-Pal_Nuc	GTCAGGTCAGTGCCGTGGGC		
FC1-Gam_Nuc	CAATGGTCCGAATTCTGTTCTC	Амплификация полной кодирующей последовательности Gam_Nuc	60°C
FC2-Gam_Nuc	GTGACTACGCGCAGAGTGGC		
FC1-Par_DSN	TTGGATCCTGTCAATGGCCAGG ACTGTGTGTGG	Амплификация кодирующей последовательности Par_DSN без лидерного пептида	58°C
FC2-Par_DSN	AAAAGCTTAGTGAGGAGTCCGA CATTGCCAG		
p53-dir2	AAGGGGAGCCTCACCACG	Амплификация фрагмента гена p53, содержащего ОНП С309Т	65°C
p53-rev2	CCACGGATCTGAAGGGT		
F2-dir	ATGGTTCCCAATAAAAAGTGAC	Амплификация фрагмента гена F2, содержащего ОНП G20210A	63°C
F2-rev	AATAGCACTGGGAGCATTGA		
MTHFR-dir	CTTGAAGGAGAAGGTGTCTG	Амплификация фрагмента гена MTHFR, содержащего ОНП С677Т	63°C
MTHFR-rev	AAGAAAAGCTGCGTGATGATG		

BRCA1-dir	GCCCCCTGAAGATCTTTCTG	Амплификация фрагмента гена BRCA1, содержащего ОНП 5382insC	65°C
BRCA1-rev	CCAAGGTCCAAAGCGAGCA		
F5-dir	TACTAATCTGTAAGAGCAGATC	Амплификация фрагмента гена F5, содержащего ОНП G1698A	65°C
F5-rev	GTTACTTCAAGGACAAAATACC		
ApoE-dir	GGCCTGCACCTCGCCGCGGTA	Амплификация фрагмента гена ApoE, содержащего ОНП C388T	65°C
ApoE-rev	GCGGACATGGAGGACGT		
hRas-dir	CGGAATATAAGCTGGTGGTG	Амплификация фрагмента гена hRas, содержащего ОНП G35T	65°C
hRas-rev	GATGGTCAGCGCACTCTTG		
nRas-dir	CTGAGTACAAACTGGTGGTG	Амплификация фрагмента гена nRas, содержащего ОНП G34A, G35C, G35A	65°C
nRas-rev	GCTGGATTGTCAGTGCGCTT		
kRas-dir	TATAAACTTGTGGTAGTTGGAG	Амплификация фрагмента гена kRas, содержащего ОНП G34A, G35T, G35A, G38A	65°C
kRas-rev	CTGTATCGTCAAGGCACTCT		
COMT-dir	CAGCCACAGTGGTGCAGAG	Амплификация фрагмента гена COMT, содержащего ОНП rs165599	65°C
COMT-rev	GTCCACCTGTCCCCAGCG		
DKC1-dir	AAGGCATACATTTCCATGG	Амплификация фрагмента гена DKC1, содержащего инверсию 166_167invCT	63°C
DKC1-rev	CAAGGATGCCAGCAGTAAG		

COX-dir1	CCGCAACCTCAACACCACCT	Амплификация фрагмента гена COX1, содержащего ОНП С7028Т	63°C
COXrev1	ACAGCTCCTATTGATAGGAC		
COX-dir2	AATAGACGTAGACACACGAG	Амплификация фрагмента гена COX1, содержащего ОНП С7028Т	63°C
COX-rev2	TTCGCTTCGAAGCGAAGGCT		
COX-dir3	CTAGACATCGTACTACACGA	Амплификация фрагмента гена COX1, содержащего ОНП С7028Т	63°C
COX-rev3	CAGCTCCTATTGATAGGACA		
18D17-1	AAAGCAGCATGTCCTGGAGT	Амплификация фрагмента слитого транскрипта 18D17	58°C
18D17-2	ACCCACCATGAAGGGATTG		
13D07-1	GGGATTGGGGTAGAGGTTTT	Амплификация фрагмента слитого транскрипта 13D07	62°C
13D07-2	TTCAATGCTGCGATTATCCTC		
8F09-1	GCTGGACCTATAGCAACTGAAT G	Амплификация фрагмента слитого транскрипта 8F09	62°C
8F09-2	GAAGAAGAGGAACCGGATGG		
23L23-1	AGCAAATTGTGCAATGGAAT	Амплификация фрагмента слитого транскрипта 23L23	62°C
23L23-2	GGTTTGTGCTGACCATCTCC		
T1P1	ACGTCGGATCCTGGGTTCTTCAG CGTCTCCATG	Амплификация Par_DSN-T1	64°C
T1P2	ACGTCAAGCTTCCCGTGAGCCA GGTCGTCC		64°C
T2P	ACGTCGGATCCTTATGAACTGA TCAGCGTGTGTTT	Амплификация Par_DSN-T2	64°C
T3P	ACGTCGGATCCTACAACCTTGC GCACCGAGCA	Амплификация Par_DSN-T3	64°C
T4P	ACGTCGGATCCTAACCACGTGA AGGAGACC	Амплификация Par_DSN-T4	64°C

TCR-C-Uni	AGGCGGCTGCTCAGGCAGT	ТКР восемнадцатого семейства	63°C
BV18	AGACACCTGGTCAGGAGGAG		
DKC1 ex3F-1	AAAGGCATACATTTCCATGG	Амплификация фрагмента гена DKC1	63°C
DKC1 ex3F-2	CAAGGATGCCAGCAGTAAG		
RPS19-1	CCCAGCTCGTTAGAATGCACC	Амплификация фрагмента гена RPS19	63°C
RPS19-2	GAAAAGTCAGACAGGACTGTCC		
TERC-1	AGAGTGACTCTCACGAGAGC	Амплификация фрагмента гена TERC	63°C
TERC-2	ATGTGTGAGCCGAGTCCTG		
GJB2-1	ACGCTGCAGACGATCCTGGG	Амплификация фрагмента гена GJB2	63°C
GJB2-2	GGCGTGGACACGAAGATCAG		
ELA2-1	GCCAATCCAGCGTCTTGTCC	Амплификация фрагмента гена ELA2	63°C
ELA2-2	CACCGCTCAGAACCTCAGTC		
MYL2-1	GGCGAGTGAACGTGAAAAAT	Амплификация фрагмента гена MYL2	63°C
MYL2-2	CCCCCGAAGAAACATAGACA		
MYL2-3	TAAGCCCTGTGTTCCGTTTC	Амплификация фрагмента гена MYL2	63°C
MYL2-4	TAAGCCCTGTGTTCCGTTTC		
MYBPC3-1	CGGGAGGAGACCTTCAAATA	Амплификация фрагмента гена MYBPC3	63°C
MYBPC3-2	CAAGCCCTAAAGCCTCATGT		
TNNC1-1	GGGGCAGGGCTATTTAAGTC	Амплификация фрагмента гена TNNC1	63°C
TNNC1-2	ACGTAGGCCTTCTCCTTGGT		
MYLK2-1	AAAAGGAAGATGTCCCAGCA	Амплификация фрагмента гена MYLK2	63°C
MYLK2-2	CTCCAGCAACACCATTTCT		
LAMP2-1	GCTTTGCTGGCCACTTTTAT		63°C

LAMP2-2	TCACCACAGCAAAGACAAAGTC	Амплификация фрагмента гена LAMP2	
PDE1A-dir	CAAACAACCTTTCACATTCAG	Амплификация фрагмента гена PDE1A	59°C
PDE1A-rev	CCTGACATGTCTGTAGATAAAA CCA		
PDE1B-dir	ATCCATGATTATGAGCACAC	Амплификация фрагмента гена PDE1B	61°C
PDE1B-rev	GGAGTTCTACAAACTCATCCTTG GT		
PDE1C-dir	CCAACAATTTCCACATTCAG	Амплификация фрагмента гена PDE1C	61°C
PDE1C-rev	TCACCATTTCAATTACCAAGGTT CG		
PDE2A-dir	CTTTGCCTTGTTTATTTCT	Амплификация фрагмента гена PDE2A	59°C
PDE2A-rev	GGAGAAATGATCAAAGATGTTG CAG		
PDE3A-dir	TGATCATCCAGGAAGGACTA	Амплификация фрагмента гена PDE3A	63°C
PDE3A-rev	TAAATTCCACATGGTCAAGGTT AAT		
PDE3B-dir	AGAATTGATGGCTCTATACG	Амплификация фрагмента гена PDE3B	61°C
PDE3B-rev	TGTATTCTGGGCGAGAAAGATA TAG		
PDE4A-dir	GTCTCCAACCAGTTCCTCAT	Амплификация фрагмента гена PDE4A	65°C
PDE4A-rev	CGATGACCATCTTGCGTAGGCT CTG		
PDE4B2-dir	CCAATCAGTTTCTCATCAAC	Амплификация фрагмента гена PDE4B	61°C
PDE4B2-rev	ATATCAGTTGCTAACACCATGTC AA		
PDE4C-dir	TCTGATTAACACCAACTCAG	Амплификация фрагмента гена PDE4C	61°C
PDE4C-rev	GTTCATGTGTTTGGACATGTCTG TG		
PDE5A-dir	GGTGTGAATAACTCTTACAT	Амплификация фрагмента гена PDE5A	59°C
PDE5A-rev	AACGTGGTCTTATATTCTTCAAT GG		
PDE8A-dir	CCAACTCCTTCCTGTGTAAT		61°C

PDE8A-rev	GATAATCATTCCTCTCCATGTTT TT	Амплификация фрагмента гена PDE8A	
PDE10A-dir	TAACAGCTACCTGCAGAAGT	Амплификация фрагмента гена PDE10A	63°C
PDE10A- rev	AGTATAAAGCAAGGTCTGTGGC AAT		
PDE6G-dir	TTTAAGCAGCGACAGACCAG	Амплификация фрагмента гена PDE6G	65°C
PDE6G-rev	GCCTCGTGCTAGATGATGCCAT ATT		
23SPHK-1	GAACCTCGCTGTGAAGATGCAGT	Амплификация фрагмента 23S РНК E.coli	63°C
23SPHK-2	GTCACGCTCGCAGTCAAGCT		
16SPHK-1	СТААСАСАТGCAAGTCGAAC	Амплификация фрагмента 16S РНК E.coli	63°C
16SPHK-2	TGACTТААСАААССGССТGC		
Adk1-1	ATGCGTATCATTTCTGCTTGG	Амплификация фрагмента гена Adk1 E.coli	63°C
Adk1-2	ACCATTACGGCAGTCTTCCT		
Ksg-1	TGGTGATGCGGCTCAACACG	Амплификация фрагмента гена Ksg E.coli	65°C
Ksg-2	ATCAGCAGGATGCGATGAC		
Din-1	AAATTCCTGGCGTCGGCAAAGT	Амплификация фрагмента гена Din E.coli	65°C
Din-2	TGAGGGTCAAGCAACGTCACAT		
TalB-1	ATCGCGGCAATGAAGCTGTA	Амплификация фрагмента гена TalB E.coli	65°C
TalB-2	CGCAATTGACGCTTCGGTGT		
NarP-1	CGCGGTGTTTCGTCAGTTACT	Амплификация фрагмента гена NarP E.coli	65°C
NarP-2	TCAACACCGAGGCAATCTGT		

Амплификацию проводили в амплификаторе MJ Research PTC-200 DNA Thermal Cycler (BioRad, США), используя следующий режим ПЦР: 94°C — 20 с, T₀ — 30 с, 72°C — 20-30 с для коротких фрагментов ДНК, 3 мин для кДНК,

обогащенной полноразмерными последовательностями. Во всех случаях подбор оптимального числа циклов для амплификации кДНК и выход реакции на плато определяли, контролируя накопление продуктов ПЦР с помощью электрофореза в 1-2% агарозе с окрашиванием бромистым этидием. Обычно для амплификации целевых фрагментов требовалось 26-30 цикла ПЦР, однако в редких случаях это число могло быть увеличено.

Для амплификации фрагментов кодирующих последовательностей нуклеаз ракообразных в процедуре RACE использовали набор реактивов SMART RACE (Clontech) и протокол, разработанный нами ранее [535]. В качестве стартового материала использовали амплифицированную кДНК, приготовленную на основе суммарной РНК из ракообразных. Продукты ПЦР очищали от реакционной смеси и клонировали в плазмидный вектор pGEM-T Easy Vector (Promega, США) по протоколу производителя и использовали для трансформации *E.coli*. Вставки проверяли с помощью секвенирования по методу Сэнгера.

Кодирующую последовательность Par_DSN, амплифицированную на матрице первой цепи кДНК из камчатского краба с помощью праймеров FC1-Par_DSN и FC2-Par_DSN, клонировали в вектор pQE-30 (Qiagen) при получении неактивного рекомбинантного белка и в вектор pET-22b (Novagen) для сайт-направленного мутагенеза и получения активного рекомбинантного белка.

Для получения укороченных вариантов Par_DSN, в качестве ДНК матрицы использовали фрагменты кодирующую последовательность нуклеазы, клонированную в вектор pET-22b. Для варианта Par_DSN-T1 использовали праймеры T1P1 и T1P2, для остальных вариантов — один ген-специфический праймер и праймер к T7-терминатору (Novagen, США). Амплификацию проводили в 10 циклах ПЦР с 1 нг плазмидной ДНК в режиме 95°C — 10 с, 64°C — 10 с, 72 °C — 1 мин. Продукты ПЦР очищали от компонентов реакционной смеси и клонировали в вектор pET22b по сайтам рестрикции BamHI и EcoRI. Отсутствие дополнительных мутаций подтверждали секвенированием по методу Сэнгера.

Для амплификации фрагментов гена DKC1, содержащих инверсию 166_167invCT, использовали ранее полученную плазмидную ДНК, несущую фрагменты мутантного гена DKC1 и гена дикого типа [536]. Для амплификации фрагментов гена COMT, содержащих ОНП rs16559, использовали геномную ДНК из периферической крови здоровых женщин-добровольцев, предоставленную Центром воспалительных заболеваний Университета Северной Каролины (США). Амплификацию остальных фрагментов ДНК человека, содержащих ОНП, проводили из образцов геномной ДНК, любезно предоставленных проф. А.Г. Татосьяном, проф. А.В. Терских, проф. Н.К. Янковским и проф. А.В. Поляковым.

Для создания библиотеки фрагментов ТКР восемнадцатого семейства первую цепь кДНК из мононуклеарных клеток крови человека использовали в качестве матрицы для амплификации в ПЦР с праймерами TCR-C-Uni и BV18. Продукт ПЦР клонировали в вектор pGEM-T-Easy (Promega, США) и использовали для трансформации *E. coli*. Амплификацию вставок для секвенирования по методу Сэнгера осуществляли со стандартными M13-праймерами с использованием в качестве матрицы 3 нг плазмидной ДНК. Проводили 15 циклов ПЦР (95°C — 7 сек, 63°C — 20 сек, 72°C — 40 сек).

Для создания библиотек нормализованной и ненормализованной геномной ДНК свежую амплифицированную ДНК (~200 нг) клонировали в вектор pGEM-T-Easy (Promega).

Для амплификации фрагментов генов фосфодиэстераз проводили от 18 до 25 циклов ПЦР в следующем режиме: 95°C — 7 с, 55-65°C — 20 с и 72°C — 15 с. Продукты ПЦР клонировали в вектор pGemT-easy (Promega). Очищенные плазмиды с целевыми вставками использовали в качестве матриц для амплификации (5 нг каждой плазмидной ДНК на 25 мл реакционной смеси) с ген-специфичными праймерами. В каждом случае проводили 12 циклов ПЦР. Очищенные амплифицированные продукты смешивали вместе в эквимолрных количествах для получения драйвера для ДСН-деплеции.

Для приготовления библиотек из *Zoanthus* sp. использовали очищенные образцы амплифицированной и деплецированной кДНК, которые клонировали в вектор pTriplEx2 (Clontech) по сайтам рестрикции Sfi I.

В работе, если не указано иное, использовали ЭР производства New England Biolabs (Великобритания). Обработка ЭР осуществлялась по инструкции производителя. Для лигирования использовали ДНК лигазу T4 (Promega), согласно инструкции производителя.

В работе, если не указано иное, использовали компетентные клетки *E.coli*, штамм XL1-blue (Евроген, Россия), подготовленные для трансформации посредством электропорации. Трансформацию и выращивание колоний осуществляли по инструкции производителя. Для электропорации использовали микропульсатор Biorad Micropulser (Biorad, США).

2.6. ПЦР в реальном времени

РВ-ПЦР проводили на амплификаторе MX3005P QPCR Systems (Stratagene, США). Производительность машины ПЦР по умолчанию составляла 45 циклов.

Образцы амплифицированного контроля и деплецированной кДНК очищали с использованием набора для очистки ПЦР QJ Aquich (Qiagen), осаждали этанолом, растворяли в 50 мкл стерильной воды и определяли количество кДНК на спектрофотометре. Из-за невозможности исключить изменения концентраций эталонных генов во время процедур нормализации и вычитания измерение не нормализовали по эталонным генам, а просто представляли относительные концентрации каждого гена с использованием эквивалентных количеств кДНК для каждой библиотеки.

РВ-ПЦР на образцах кДНК из плаценты проводили с использованием коммерчески доступных праймеров и TaqMan зондов, соответствующих каталитическим доменам различных генов фосфодиэстераз (Applied Biosystems, США). РВ-ПЦР на образцах кДНК *E.coli* осуществляли с использованием готовой

смеси qPCRmix-HS SYBR (Евроген, Россия) с использованием праймеров, указанных в таблице 2-1.

2.7. Сайт-направленный мутагенез

Кодирующую последовательность Par_DSN амплифицировали с матрицы первой цепи кДНК, полученной как описано выше, с помощью праймеров FC1-Par_DSN и FC2-Par_DSN (таблица 2-1) и клонировали в вектор pET-22b (Novagen, США). Сайт-направленный мутагенез осуществляли с использованием ПЦР с праймерами к T7-промотору или T7-терминатору (Novagen, США) и ген-специфическими праймерами, содержащими целевые замены (таблица 2-2). Фрагменты очищали от реакционной смеси с помощью электрофореза в 1% агарозном геле с экстракцией с помощью набора реактивов QIAquick PCR purification kit (Qiagen) по инструкции производителя. Далее очищенные фрагменты использовали в реакции элонгации с помощью Taq-полимеразы в следующем режиме: 94 °C — 20 с, 58 °C — 2 мин, 72 °C — 1 мин, 5 циклов. Полноразмерные последовательности амплифицировали в ПЦР с праймерами к T7-промотору и T7-терминатору и клонировали в pET-22b по сайтам рестрикции NdeI и HindIII по стандартному протоколу. Последовательности вставок подтверждали с помощью секвенирования по методу Сэнгера.

Таблица 2-2. Структуры праймеров, использованных для сайт-направленного мутагенеза

Вносимая аминокислотная замена	Последовательности прямого и встречного праймеров (5'→3')
His237Ala	ACTTCGCCAAAGGTGCCATGTCTCCTGACGC и GCGTCAGGAGACATGGCACCTTTGGCGAAGT
Arg235FAla	GTACTTCGCCGCAGGTCACATG и GACATGTGACCTGCGGCGAAG

Arg235Lys	GTACTTCGCCAGAGGTCACATG и GACATGTGACCTCTGGCGAAG
Arg184Ala	GAGACTAGTGCTCCCTCCTTCAAGA и TGAAGGAGGGAGCACTAGTCTCGAT
Arg184Lys	GAGACTAGTAAACCCTCCTTCAAGA и TGAAGGAGGGTTTACTAGTCTCGAT
His171Ala	CGAGCATGTCCTCGCCGGCGCCAACATT и GCAATGTTGGCGCCGGCGAGGACATGCT
His168Ala	AACCTTGCGCACCGAGGCTGTCCTCCAC и CGCCGTGGAGGACAGCCTCGGTGCGCAA
Gly139Val	ATAGGCACAGTATTGTCTTCGAGTACTACG и CGTAGTACTCGAAGACAATACTGTGCCTATC
Gly133Val	GGAAGCGGACCAGGTGGATAGGCACA и TACTGTGCCTATCCACCTGGTCCGCT
Arg121Ala + Arg122Ala	AAGGAGACCATCGCAGCTAACCTTGGAACA и GTTCCAAGGTTAGCTGCGATGGTCTCCTTC
Arg122Ala	AGGAGACCATCCGCGCTAACCTTGGAACAT и CATGTTCCAAGGTTAGCGCGGATGGTCTCC
Arg122Lys	AGGAGACCATCCGCAAGAACCTTGGAACAT и CATGTTCCAAGGTTCTTGCGGATGGTCTCC
Arg121Lys	GTGAAGGAGACCATCAAGCGCAACCTTGGAACA и TGTTCCAAGGTTGCGCTTGATGGTCTCCTTC
Arg121Ala	GTGAAGGAGACCATCGCCCGCAACCTTGGAА и TGTTCCAAGGTTGCGGGCGATGGTCTCCTTCAC

2.8. Рестрикционный анализ

Рестрикционный анализ использовали для подтверждения аллельного статуса образцов геномной ДНК по ОНП rs165599 гена COMT. В реакции рестрикции 500 нг очищенного продукта ПЦР обрабатывали 10 ед. Nae I. Продукты реакции анализировали с помощью электрофореза на 2,5% агарозе с окраской бромистым этидием. Один фрагмент ДНК (150 п.о.) наблюдался у гомозигот по

аллели А, два фрагмента ДНК (76 и 74 п.о.) — для гомозигот по аллели G, все три фрагмента — для гетерозиготы.

2.9. Гибридизация НК с радиоактивно мечеными пробами

³²P-меченые ген-специфические пробы получали с использованием праймеров из набора Human Amplimer Set (Clontech) с помощью набора реактивов Prime-a-Gene (Promega). Радиоактивное мечение образцов кДНК осуществляли с использованием того же набора реактивов. Гибридизацию проводили, как предложили Engler-Blum et al. (1993) [537].

Для виртуального Нозерн-блот анализа использовали метод Franz et al. (1999) [538]. Фракционирование образцов кДНК и перенос ДНК мембраны Hybond-N (Amersham, США) осуществляли по стандартному протоколу Sambrook et al. (1989) [534]. Для гибридизации с меченой кДНК использовали нейлоновые мембраны, содержащие 183 ПЦР пробы и 2300 олигонуклеотидных проб, комплементарных генам человека (Clontech). Анализ данных проводили с помощью AtlasImage™ Software (Clontech).

2.10. Получение рекомбинантных белков

В первых экспериментах для трансформации *Escherichia coli* штамм Ad 494 (Novagen) использовали кодирующую последовательность Par_DSN, клонированную в вектор pQE-30 (Qiagen, США). Культуру *E. coli* после ночи инкубации использовали для выделения рекомбинантного белка в денатурирующих условиях с помощью TALON® Metal Affinity Resins (Clontech, США). Рекомбинантный белок использовали для иммунизации кролика и получения поликлональных антител. Для получения рекомбинантных белков, обладающих нуклеазной активностью, использовали кодирующие последовательность Par_DSN дикого типа, укороченные и мутантные варианты,

полученные, как описано выше, и клонированные в вектор pET-22b(+). Выделение и обработку активных белков проводили, как описано в работе Анисимовой и соавт. [539].

2.11. Получение поликлональных антител

Поликлональные антитела получали как описано Cooper & Paterson (1999) [540]. Для инъекций использовали эмульсию рекомбинантного Par_DSN в полном адьюванте Фрейнда. Было сделано четыре инъекции (1 мг Par_DSN на инъекцию) с 30-дневными интервалами. Сыворотку отбирали через 10 дней после инъекции. Титр антител оценивали с помощью ELISA и вестерн-иммуоблотинга по стандартным протоколам [541].

2.12. Приготовление ацетонового порошка из гепатопанкреаса камчатского краба

Свежий гепатопанкреас краба гомогенизировали в двух объемах 100 мМ Трис-НСl с 100 мМ ЭДТА и центрифугировали при 10000–12000 g в течение 30 мин. Для приготовления ацетонового порошка супернатант разбавляли 1,5 объемами ацетона, инкубировали 12 ч и центрифугировали при 10 000 g в течение 1 ч. Осадок высушивали для хранения. Все процедуры осуществляли при 4°C. В первых экспериментах использовали буферный раствор Трис-НСl с рН 8,0. В окончательном протоколе использовали буферный раствор Трис-НСl с рН 7,1.

2.13. Выделение Par_DSN из камчатского краба

Свежий супернатант, полученный как описано выше, или раствор из ацетонового порошка использовали для анионообменной хроматографии на сорбенте DEAE MacroPrep, уравновешенным 0,05 М Трис-НСl. Для промывки использовали четыре объема того же буфера. Белок элюировали ступенчатым

градиентом 0–1 М NaCl в 0,05 М Трис-HCl. Все операции проводили при скорости потока 3 мл/мин. Фракции анализировали с помощью гель-электрофореза в полиакриламидном геле и иммуноблоттинга с ПА-нук по стандартному протоколу [541], а также путем определения ДНКазной активности фракций. Для определения концентрации белка использовали метод Бредфорд [542], как стандарт использовали разведения бычьего сывороточного альбумина. Обогащенные Par_DSN фракции с высокой удельной активностью собирали и добавляли в раствор кристаллический NaCl до конечной концентрации 5 М. Далее препарат подвергали хроматографии на фенил-сефарозе (Amersham Pharmacia Biotech), уравновешенной 0,05 М Трис-HCl с 5 М NaCl. После промывки колонки четырьмя объемами того же буфера элюировали фракции в ступенчатом градиенте 5–3 М NaCl в том же буфере. Обогащенные целевым белком фракции с высокой удельной активностью собирали, доводили до 5 М твердым раствором NaCl и повторно хроматографировали на фенил-сефарозе в тех же условиях.

Выделенный фермент лиофилизировали для длительного хранения. Для работы фермент растворяли в 50 мМ Трис-HCl, pH 8.0 и добавляли равный объем глицерина.

2.14. Определение активности Par_DSN

Для определения активности нуклеазы по отношению к ДНК субстрату использовали предложенную Liao et al. (1974) [543] модификацию методики Куница [544]. Анализ проводили в реакционном буфере, содержащем 7 мМ MgCl₂, 50 мМ Трис-HCl (pH 8,0). Также использовали метод, основанный на выявлении накопления растворимых продуктов расщепления ДНК как описано Мензоровой и др. (1993) [27] с модификацией: инкубацию с субстратом проводили при 60°. Для остановки реакции в реакционную смесь добавляли 10 мкл 3 М калий-ацетатного буфера, (pH 4,8) и 300 мкл 96% этанола. В качестве субстрата использовали 40 мкг ДНК тимуса теленка.

Определение РНКазной активности осуществляли, как описано Но et al. (1998) [545]. Реакционную смесь объемом 20 мкл, содержащую 1 × буфер DSN (100 мМ Трис-НСl, рН 7,5; 10 мМ MgCl₂, 2 мМ дитиотриэтол) и 0,6 ед. Куница Par_DSN, прогревали при 55°C в течение 5 мин и добавляли в нее свежеприготовленную РНК из *Saccharomyces cerevisiae* (Sigma, США) в количестве 150 нг. Инкубацию продолжали 1 ч и останавливали реакцию добавлением 0,1 объема ацетата аммония (7,5 М) и 3 объема 96% этанола. Продукты реакции перемешивали, центрифугировали, пятикратно разбавляли водой и определяли оптическую плотность при 260 нм.

2.15. Исследование параметров, влияющих на активность Par_DSN

Влияние на активность Par_DSN ионов двухвалентных металлов определяли в инкубационных смесях, содержащих различные концентрации (0–100 мМ) индивидуальных двухвалентных катионов (Mg²⁺, Ca²⁺, Mn²⁺, Co²⁺, Cd²⁺) или смесей катионов, объединенных в равных количествах. Реакционные смеси инкубировали при температуре от 25 до 85°C. Активность измеряли при значениях рН в диапазоне от 3 до 11, скорректированных с использованием 50 мМ глицин-НСl, ацетата натрия, MES-NaOH, Трис-НСl и боратно-натриевых буферов. ДНКазную активность определяли, как описано выше. Зависимость активности ДСН от рН анализировали с помощью реакции Куница с использованием буферных растворов 50 мМ формиат натрия (рН 3,0–3,5), 50 мМ ацетат натрия (рН 3,5–6,0), 50 мМ Mes-NaOH (рН 6,0–7,0), 50 мМ Трис-НСl (рН 7,0–9,5) и 50 мМ глицин-NaOH (рН 9,0–10,0) в присутствии 7 мМ MgCl₂. Зависимость активности Par_DSN от температуры анализировали при в диапазоне от 20 до 100°C (с интервалом 5°C). Исследование температурной стабильности Par_DSN осуществляли в диапазоне температур от 25 до 100°C в 1 × буфере DSN. Исследование рН стабильности осуществляли при температуре 25°C при значениях рН от 0 до 14 с использованием буферных растворов 50 мМ формиат натрия (рН 3,0–3,5), 50 мМ ацетат натрия (рН 3,5–6,0), 50 мМ Mes-NaOH (рН 6,0–7,0), 50 мМ Трис-НСl (рН 7,0–9,5) и 50 мМ

глицин-NaOH (pH 9,0–10,0) в присутствии 7 мМ MgCl₂. Тестирование влияния потенциальных ингибиторов и стимуляторов, добавляемых в инкубационные смеси осуществляли в 1 × буфере DSN в присутствии 7 мМ MgCl₂ и при различных температурах.

2.16. Определение субстратной специфичности

Определение субстратной специфичности Par_DSN проверяли в реакционных смесях в общем объеме 20 мкл, содержащем 1 × буфер DSN, 0,06 ед. Куница Par_DSN, субстрат. Использовали следующие комбинации субстратов: 150 нг ДНК фага λ и 50 нг ДНК M13; 200 нг геномной ДНК человека и 50 нг ДНК M13; 200 нг плазмидной ДНК pQE-30 (Qiagen), 100 нг суммарная РНК человека + 100 нг ДНК фага λ. Для предотвращения образования дц структур в ДНК фага M13 инкубацию реакционных смесей осуществляли при 70°C в течение различных промежутков времени. Инкубацию с РНК субстратом осуществляли при 37°C. После 30 мин инкубации реакции останавливали добавлением ЭДТА до конечной концентрации 5 мМ. Продукты реакции визуализировали на 0,9% агарозном геле после окрашивания бромистым этидием.

Для приготовления дц РНК субстрата фрагмент кДНК гена p53 амплифицировали в ПЦР с ген-специфическими праймерами: p53dir1 (5'-AGGGGAGCCTCACACG-3') и p53rev1 (5'-CCACGGATCTGAAGGGTGAA-3'). После очистки продукта амплификации, к нему лигировали последовательность T7 промотора и использовали продукт лигирования в качестве матрицы для двух ПРЦ: с праймерами p53dir1 и T7, и с праймерами p53rev1 и T7. Полученные после амплификации фрагменты очищали от реакционной смеси и использовали для синтеза РНК с промотора T7. Одна из полученных синтетических РНК соответствует смысловой, а другая – антисмысловой РНК гена p53. Для приготовления дц РНК субстрата эти РНК смешивали в эквимольных количествах.

Для создания синтетических дц субстратов меченые олигонуклеотиды смешивали с эквимольными количествами комплементарных или частично комплементарных немеченых олигонуклеотидов. Структуры субстратов показаны в таблице 3-4 и на рисунке 3.23. Реакцию расщепления проводили в общем объеме 20 мкл, содержащем 1 × буфер DSN, 0,6 единицы Куница Par_DSN и 0,3 мкМ субстрата при 30°C. Время инкубации варьировали в зависимости от требований эксперимента, отбирая аликвоты реакционной смеси через определенные промежутки времени.

Эффективность расщепления синтетических субстратов, содержащих на 5'-конце флуорофор и на 3'-конце гаситель флуоресценции, оценивали по изменению интенсивности флуоресценции реакционной смеси, которую измеряли на спектрофлуориметре Cary Eclipse (Varian, Австралия). Здесь и далее, фоновой флуоресценцией считали значение интенсивности эмиссии флуоресценции нерасщепленного субстрата. Максимальной флуоресценцией считали значение интенсивности эмиссии флуоресценции реакционной смеси, содержащей соответствующее количество флуорофора.

Субстратную специфичность по отношению к радиоактивно меченым синтетическим ДНК-ДНК-дуплексами проводили при 30°C; реакции с ДНК-РНК дуплексами — при 45°C в течение 1 ч. Реакционная смесь объемом 30 мкл содержала 1 × буфер DSN, 1 ед. Куница Par_DSN, по 10 пМ каждого синтетического олигонуклеотида. Реакции останавливали добавлением ЭДТА до конечной концентрации 5 мМ. Анализ продуктов реакции осуществляли с помощью электрофореза в полиакриламидном геле в присутствии мочевины с последующей автордиографией на пленке BioMax (Kodak, США). Информация о структуре синтетических субстратов суммирована в таблице 2-3. Олигонуклеотиды p7 — p9 и p25 и p30 были помечены с помощью γ -³²P рАТФ с использованием полинуклеотидкиназы (СибЭнзим, Россия).

Таблица 2-3. Структуры олигонуклеотидов, использованных для анализа субстратной специфичности Par_DSN

Название	Последовательность (5'→3')	Описание
Na21T7	TGTAGCGTGAAGACGAC AGAAGTAATACGACTCA CTATAGGGC	Олигодезоксирибонуклеотид, использованный для формирования дц ДНК дуплексов с мечеными γ - ^{32}P рАТФ олигонуклеотидами p7 – p10
рибо30	СТААТАСГACTCACGGA TGGGCGGGAATAA	^{32}P -меченый олигорибонуклеотид, формирующий дуплекс с p25
p7	GCCCTAT	^{32}P -меченый олигодезоксирибонуклеотид
p8	GCCCTATA	^{32}P -меченый олигодезоксирибонуклеотид
p9	GCCCTATAG	^{32}P -меченый олигодезоксирибонуклеотид
p25	CCCGCCCATCCGTGAGT CGTATTAG	^{32}P -меченый олигодезоксирибонуклеотид

2.17. Масс-спектрометрический анализ

Выделенный природный белок подвергали электрофорезу и расщеплению в полиакриламидном геле, как описано Shevchenko et al. (2006) [546]. Масс-спектрометрический анализ проводили с использованием MALDI-TOF/TOF с возможностью LIFT (Ultraflex TOF/TOF, Bruker Daltonik GmbH, Бремен, Германия) на базе ЦКП ИБХ РАН при участии Р.Х. Заганшина. Ионизацию проводили импульсным азотным лазером с длиной волны 337 нм. Эксперименты MS/MS на четырех наиболее интенсивных пептидных пиках переваренного белка проводили с помощью лазерно-индуцированной диссоциации (LID). Все данные были обработаны с использованием программных пакетов FlexAnalysis версии 2.4 и BioTools версии 3.0 (Bruker Daltonics GmbH). Информация о фрагментации из

нескольких спектров LID, объединенных со спектрами TOF, была обработана с помощью программного обеспечения Mascot (<http://www.matrixscience.com/>) для определения возможной идентичности с использованием допуска точности 0,1 Да для исходного пептида и 0,5 Да для фрагментов MS/MS, и следующие параметры поиска: база данных геномов – NCBI; таксономия – все; фермент – трипсин.

2.18. Анализ последовательностей НК и аминокислот

Для анализа последовательностей НК использовали пакет программ BLAST (<http://www.ncbi.nlm.nih.gov/BLAST>). Доменную организацию предсказанных белков анализировали с помощью программного обеспечения SMART (<http://smart.embl-heidelberg.de>), поиск сигнальных последовательностей осуществляли с помощью программного обеспечения SignalP (версии 3.0, 5.0 и 6.0; <https://services.healthtech.dtu.dk/service.php/SignalP>). Дополнительный анализ трансмембранных структур проводили с помощью программного обеспечения TMHMM-2.0 (<https://services.healthtech.dtu.dk/service.php/TMHMM-2.0>) и PredictProtein (<https://predictprotein.org/>) со встроенными модулями TMSEG и RePROF. Информационную структуру белка анализировали как описано в работе А.Н. Некрасова (2004) [547]. Предполагаемые N- и/или C-концевые концы генов выявляли в последовательностях аннотированных геномов с использованием программного обеспечения FGENESH (<http://sun1.softberry.com>). Повторяющиеся элементы генома в данных секвенирования анализировали с помощью программного обеспечения Repeat-Masker (www.repeatmasker.org/cgi-bin/WEBRepeatMasker). Множественное выравнивание последовательностей осуществляли с помощью программного обеспечения ClustalX [548] с визуализацией в программе GeneDoc [549].

Для построения филогенетического дерева использовали программное обеспечение TREE-PUZZLE, версия 5.0 [550]. Деревья максимального правдоподобия были построены с использованием алгоритма «quartet puzzling» с 10000 шагов. Мы использовали дискретную модель G-распределения (с восемью

категориями) для неоднородных участков [551] и модель замещения Уилана и Голдмана [552].

2.19. ДСНП-анализ

В одном исполнении ДСНП-анализа (от англ. duplex-specific nuclease preference) использовали очищенные продукты ПЦР, растворенные в 50 мМ Tris-HCl, pH 8,0. Аликвоту продукта ПЦР, содержащую 75 нг ДНК, или раствор плазмидной ДНК (500 нг) в том же буфере инкубировали 1 ч при 35°C в реакционной смеси с конечным объемом 15 мкл, содержащей 0,6 ед. Куница Par_DSN, 0,2 ед. фрагмента Кленова (Fermentas, США), 1 × буфер для фрагмента Кленова (Fermentas) и ДСНП-зонды в конечной концентрации 0,3 мкМ.

Для подготовки субстрата длиной 4 т.п.о. фрагмент гена COX1 длиной 69 п.о. клонировали в вектор pT-Adv (Clontech), очищенные плазмиды, содержащие целевые ОНП, использовали в качестве субстрата для ДСНП-анализа.

В другом исполнении ДСНП-анализа использовали неочищенные продукты ПЦР. Аликвоту реакционной смеси ПЦР объемом 5 мкл, содержащую продукт амплификации (~ 75-100 нг ДНК), смешивали с буфером DSN (до 1-кратного конечного разведения), 0,6-0,75 ед. Куница Par_DSN и ДСНП-зондами (до конечной концентрации пробы 0,3 мкМ). Объем реакционной смеси довели до 15 мкл. Полученную смесь инкубировали при 65°C 10 мин, затем — при 35°C 40 мин.

В обоих случаях реакцию останавливали добавлением ЭДТА до конечной концентрации 5 мМ. Эффективность расщепления субстратов оценивали по изменению интенсивности эмиссии флуоресценции, которую измеряли на спектрофлуориметре Cary Eclipse (Varian) или на фотометре-флуориметре ФФМ-01 (Кортэк, Россия). Визуализацию реакционных смесей осуществляли на флуоресцентном стереомикроскопе Olympus SZX12 с зеленым для FAM (эмиссия 520 нм, возбуждение при 480 нм) и красным для TAMRA (эмиссия 590 нм,

возбуждение при 560 нм) фильтрами. Структуры ДСНП-зондов показаны в таблице 2-4.

Таблица 2-4. Структуры ДСНП-зондов для анализа ОНП в генах человека

Ген (ОНП)	Праймер	Последовательность (5'→3')
p53(C309T)	p53-wt	FAM-TGGGCAGTGC-DABCYL
	p53-309T	TAMRA-TGGACAGTGC-DABCYL
F2 (G20210A)	F2-wt	FAM-GCTCGCTGAG-DABCYL
	F2-20210A	TAMRA-GCTTGCTGAG-DABCYL
MTHFR (C677T)	MTHFR-wt	FAM-TCGGCTCCCG-DABCYL
	MTHFR-677T	TAMRA-TCGACTCCCG-DABCYL
BRCA1 (5382insC)	BRCA-wt	FAM-AATCCCAGGA-DABCYL
	BRCA- 5382insC	TAMRA-AATCCCCAGG-DABCYL
F5 (G1698A)	FVLeiden-wt	TAMRA-TCC TCG CCT G-DABCYL
	FVLeiden- 1698A	FAM-TCC TTG CCT G-DABCYL
ApoE (C388T)	ApoE-388-wt	FAM-CGTGCGCGGC-DABCYL
	ApoE-388T	TAMRA-CGTGTGCGGC-DABCYL
hRas (G35T)	hRAS-35wt	FAM-ACC GCC GGC G-DABCYL
	hRAS-35T	TAMRA-ACC GAC GGC G-DABCYL
nRas (G34A, G35C, G35A)	nRAS-34/35wt	FAM-CCA CCT GCT C-DABCYL
	nRAS-35A	TAMRA-CCA TCT GCT C-DABCYL
	nRAS-35C	TAMRA-CCA GCT GCT C-DABCYL
	nRAS-34A	TAMRA-CCA CTT GCT C-DABCYL
kRas (G34A, G35T, G35A, G38A)	KRAS-34/35wt	FAM-GCT GGT GGC G-DABCYL
	kRas-35A	TAMRA-GCT GAT GGC G-DABCYL
	kRas-35T	TAMRA-GCT GTT GGC G-DABCYL

	kRas-34A	TAMRA-GCT AGT GGC G-DABCYL
	kRas-38A	TAMRA-GGT GAC GTA G-DABCYL
COMT (rs16559)	COMT-wt	FAM- TGCCAGCCTG-DABCYL
	COMT-16559G	TAMRA- TGCCGGCCTG-DABCYL
COX1 (C7028T)	COX-7028T	FAM-GTGAGCTACA-DABCYL
Wt — зонд, комплементарный последовательности гена дикого типа, для остальных зондов указан комплементарный зонду вариант ОНП.		

2.20. Статистическая обработка данных

Мера болевой чувствительности была рассчитана для каждого индивидуума на основе шестнадцати экспериментальных болевых процедур: четыре измерения порога болевой чувствительности при надавливании, девять показателей теплового порога/толерантности и два показателя ишемического порога/толерантности, один показатель термической боли, опосредованной С-волоконками. Все шестнадцать показателей были стандартизированы для единиц нормальных отклонений (z-показателей) путем вычитания среднего значения выборки, а затем деления на стандартное отклонение выборки. У трех из 202 генотипированных индивидуумов отсутствовал один показатель боли, у одного отсутствовали два показателя боли и у одного отсутствовали 10 показателей боли, и мы вменили среднее значение, равное нулю, в каждом случае. Затем был рассчитан суммарный z-показатель каждого индивидуума путем сложения всех шестнадцати единиц нормальных отклонений. Это дало приблизительно нормальное распределение, которое удовлетворяло предположениям об однородности дисперсии во всех последующих t-тестах и моделях ANOVA.

Статистическая оценка ассоциаций между ОНП, гаплотипами и чувствительностью к боли проводилась с использованием моделей дисперсионного

анализа ANOVA [553] и t-критерия Стьюдента для сравнения гомозигот. Проверялась нулевая гипотеза о равенстве средних среди аллелей.

Затем была проведена оценка независимых эффектов ОНП в обобщенной линейной модели с несколькими переменными. Кроме данных о ОНП rs165599 гена COMT использовались данные, полученные на тех же образцах членами консорциума для аллелей rs2097903, rs6269, rs4633; rs4818; rs4680 (Val158Met). Каждый ОНП был введен как отдельная пара фиктивных переменных (с двумя степенями свободы). ОНП вводили в следующей последовательности шагов: сначала rs4680 (met158val), затем rs4818, rs4633 и rs6269, и, наконец, rs2097903 и rs165599. Эта последовательность была основана на следующих теоретических основаниях: сначала мы контролировали один ОНП в кодирующей области, который вызвал синонимичное изменение в аминокислотной последовательности COMT; затем исследовали эффекты ОНП, связанные с несинонимичным изменением аминокислотной последовательности; и наконец, оценили любые дополнительные эффекты ОНП в промоторной и нетранслируемой областях. ОНП, введенные на третьем этапе, внесли менее одного процента в общую вариацию (R-квадрат) модели каждый.

Анализ QTL, коррелирующих с продолжительностью жизни, проводился силами коллег из Института исследования возраста им. Лейбница (Германия). Для анализа использовали результаты скрещивания короткоживущей и долгоживущей популяций *N. furzeri*. Перед анализом QTL были исключены 25% потомства F2 ($n = 71$), которые рано умерли. Данные о продолжительности жизни остальных особей F2 были преобразованы в целые числа (ранговый порядок) от 1 (первая рыба, погибшая после удаления 25% ранних смертей) до 213 для самой долгоживущей рыбы. Затем ранги были преобразованы в распределение Гаусса путем распределения их по девяти классам. Данные о продолжительности жизни были проверены на нормальное распределение с использованием критерия Колмогорова-Смирнова. Классы показали нормальное распределение (среднее значение = 5,0, стандартное отклонение = 1,40). Пол и максимальная масса тела не влияли на ранжированные данные о продолжительности жизни (пол: $P = 0,733$,

максимальная масса тела: $P = 0,097$). Расчеты проводились с помощью SigmaPlot 11 (Systat Software, Inc., Германия). Были определены минимальные, максимальные, средние и медианные значения признаков, а также стандартное отклонение и дисперсия. QTL картировали с использованием метода GridQTL [554, 555]. Полногеномное сканирование QTL выполняли с использованием подхода картирования интервалов прямой селекции [556]. Полученные пороги значимости для тестов с одним QTL и без QTL оценивались с использованием теста перестановки [557]; были проанализированы 1000 перестановок данных. Были получены следующие пороговые значения: высокосignificant для всего генома ($P = 0,01$) и значимый ($P = 0,05$). Известно, порог $P = 0,05$ для всего генома соответствует предполагаемому порогу $P = 0,63$ для всего генома у мыши [558], оцененному на основании количества хромосом. Поскольку *N. furzeri* имеет примерно такое же количество хромосом, что и мышь, разумно применить в качестве «предполагаемого» порога для всего генома. Для оценки 95% доверительного интервала одного QTL был выполнен параметрический бутстреп с 1000 итераций [559].

2.21. Выявление целевых молекул РНК

Реакционную смесь объемом 20 мкл, содержащую 50 нг смысловой или антисмысловой синтетической РНК гена p53, полученной как описано выше, 1 × буфер DSN, 0,6 ед. Куница Par_DSN и ДСНП-зонд, комплементарный смысловой РНК (5'-TAMRA-TGG GCA GTG CTC GCT-DABCYL-3') в конечной концентрации 0,3 мкМ, инкубировали 3 часа при 35°C. Реакцию останавливали добавлением ЭДТА до конечной концентрации 5 мМ. Эффективность расщепления субстратов оценивали по изменению интенсивности эмиссии флуоресценции, как описано выше.

2.22. ДСН-детекция

Для ДСН-детекции объединяли 15 мкл свежего ПЦР-продукта, 2 мкл 10 × буфера ДСН, 1 мкл 5 мкМ раствора ГСО (конечная концентрация 0,25 мкМ), 1 мкл 5 мкМ раствора универсального зонда (конечная концентрация 0,25 мкМ) и 1 мкл раствора Par_DSN (0,5 ед. Куница/мкл). В течение 20 мин реакционную смесь инкубировали при 60°C, а затем 1 ч при 37°C. Реакцию останавливали добавлением ЭДТА до конечной концентрации 5 мМ. Степень гидролиза универсального зонда оценивали по изменению интенсивности эмиссии флуоресценции, которую измеряли с помощью спектрофлуориметра Cary Eclipse (Varian).

При определении оптимальной длины ГСО (ген-специфического олигонуклеотида) в качестве ДНК субстрата использовали продукт амплификации фрагмента ТРК, полученный на матрице плазмидной ДНК с подтвержденной вставкой в ПЦР с M13-праймерами.

Для тестирования ДСН-детекции применяли продукты амплификации фрагмента гена DKC1, полученные на матрице плазмидной ДНК с использованием праймеров DKC1-dir и DKC1-rev, и фрагмента гена СОМТ, полученные на матрице геномной ДНК с известным аллельным статусом с использованием праймеров СОМТ-dir и СОМТ-rev. Структуры ГСО и универсальных зондов приведены в таблице 2-5.

Таблица 2-5. Структуры олигонуклеотидов, использованных для ДСН-детекции

Олиго- нуклеотид	Последовательность (5'→3')	Назначение
ТКР19	GCTGGTGGCGCTTTTGGGTGTGGGAGA ТС	ГСО разной длины для выявления целевой
ТКР17	GCTGGTGGCGCTTTTGGGTGTGGGAGA	
ТКР15	GCTGGTGGCGCTTTTGGGTGTGGGA	

TKP14	GCTGGTGGCGCTTTTGGGTGTGGG	последовательность и ТКР18β
TKP13	GCTGGTGGCGCTTTTGGGTGTGG	
TKP12	GCTGGTGGCGCTTTTGGGTGTG	
TKP11	GCTGGTGGCGCTTTTGGGTGT	
DKC1- Norm	GCTGGTGGCGATACCTTTAGCAAAAGG	ГСО для выявления последовательность и DKC1 дикого типа
DKC1-Mut	GCTGGTGGCGATACCTTTGACAAAAGG	ГСО для выявления последовательность и DKC1, содержащей инверсию 166_167invCT
COMT-A1	GCTGGTGGCGCGACTGCCAGCCTGGG	ГСО для выявления последовательность и «аллели А» гена COMT
COMT-C	GCTGGTGGCGCCCAGGCCGGCAGTC	ГСО для выявления последовательность и «аллели G» гена COMT
UZ-Fam	FAM-GCTGGTGGCG-DABCYL	Универсальный зонд
UZ-Tamra	TAMRA-CGTGTGCGGC-DABCYL	Универсальный зонд

2.23. Аллель-специфическая ПЦР, совмещенная с ДСН-детекцией

Для аллель-специфической ПЦР использовали набор реактивов Encyclo PCR Kit (Евроген, Россия). В качестве матрицы использовали образцы ранее охарактеризованной геномной ДНК человека, содержащей различные аллельные варианты гена COMT. Амплификацию проводили в ПЦР с 5 нг ДНК матрицы в 50 мкл реакционной смеси в присутствии праймеров COMT-out-dir, COMT-A2 (5'-GCT GGT GGC GCC TCT TCG TTT CCC AGT CT-3') и COMT-G (5'-CGT-GTG CGG CTC CTC TTC GTT TCC CAG ACC-3'). В первых трех циклах ПЦР использовали режим: 95°C — 7 с; 62°C — 20 с; 72°C — 30 с; в следующих 25 циклах: 95°C — 7 с; 67°C — 20 с; 72°C — 30 с. Далее в реакционные смеси добавляли 10 ед. акт. экзонуклеазы 1(USB, США) и инкубировали их при 37°C в течение 20 мин. После инкубации для инактивации экзонуклеазы 1, смеси прогревали при 95°C 5 мин, а затем сразу использовали для ДСН-детекции. Для этого отбирали алиquotы объемом 15 мкл реакционных смесей и смешивали их с 2 мкл 10 × буфера DSN, 1 мкл 5 мкМ зонда UZ-Fam, 1 мкл 5 мкМ зонда UZ-Tamra (см. таблицу 3-6) и 1 мкл раствора Par_DSN (0,5 ед. Куница/мкл). Полученные смеси инкубировали при 60°C в течение 20 мин и далее при 37°C в течение 30 мин. Степень гидролиза универсальных зондов оценивали, как описано выше.

2.24. ДСН-нормализация кДНК

Для нормализации амплифицированной кДНК, очищенную кДНК (600–1200 нг) помещали в 1 × гибридизационный буфер (50 мМ Hepes, pH 7.5, 0,4 М NaCl) в конечном объеме реакционной смеси 16 мкл. Разделяли гибридизационную смесь на четыре части (S1, S1/2, S1/4 и K), в каждую пробирку добавляли каплю минерального масла, чтобы избежать испарения реакционной смеси, и после центрифугирования в течение 2 мин на максимальной скорости настольной центрифуге инкубировали реакционные смеси при 98°C в течение 2 мин в

амплификаторе. Затем понижали температуру до 68°C и продолжали инкубацию при этой температуре 5–6 ч для гибридизации денатурированной кДНК. Незадолго до окончания процедуры гибридизации подготавливали разведения Par_DSN до конечных концентраций 1 ед. Куница/мкл, 5 ед. Куница /мкл и 0,25 ед. Куница /мкл в 50 мМ Трис-НСl, рН 8,0. Растворы Par_DSN сохраняли на льду.

По окончании гибридизации в реакционные смеси, содержащие реассоциированную кДНК, добавляли 5 мкл предварительно нагретый 2 × DSN буфер. Растворы быстро перемешивали, возвращали реакционные смеси в амплификатор и продолжали инкубацию при 68°C в течение 10 мин. Затем добавляли в реакционные смеси по 1 мкл раствора Par_DSN: в образец S1 — 1 ед. Куница/мкл, в образец S1/2 — 0,5 ед. Куница/мкл, в образец S1/4 — 0,25 ед. Куница/мкл. В образец К добавляли 50 мМ Трис-НСl, рН 8,0. Немедленно после добавления нуклеазы возвращали реакционные смеси в нагретый амплификатор и продолжали инкубацию при 68°C в течение 25 мин. По окончании инкубации ингибировали нуклеазную активность Par_DSN добавлением в реакционные смеси 10 мкл 5 мМ ЭДТА. Содержимое пробирок перемешивали и продолжали инкубацию при 68°C в течение 5 мин. Затем переносили пробирки с реакционными смесями на лед, добавляли в каждую реакционную смесь 20 мкл воды и использовали 1 мкл аликвоты для амплификации нормализованной кДНК.

Амплификацию нормализованной кДНК осуществляли в объеме реакционной смеси 50 мкл. В ПЦР использовали праймер M1 (см. рис. 2.36). Реакционные смеси подвергали предварительной денатурации при 95°C в течение 1 мин; затем осуществляли ПЦР в следующем режиме: 95°C — 15 с, 66°C — 20 с, 72°C — 3 мин. Эффективность нормализации и накопление продуктов ПЦР оценивали с помощью электрофореза в агарозном геле с окрашиванием бромистым этидием. Для увеличения средней длины продукта ПЦР 2 мкл аликвоту нормализованной амплифицированной кДНК реамплифицировали в ПЦР в 50 мкл реакционной смеси с праймером M2 (5'-AAG CAG TGG TAT CAA CGC AG-3'). Вначале смеси подвергали предварительной денатурации при 95°C в течение 1

мин; затем осуществляли 12 циклов ПЦР в следующем режиме: 95°C — 15 с, 64°C — 20 с, 72°C — 3 мин, и 1 цикл в режиме: 64°C — 15 с, 72°C — 3 мин.

Для нормализации первой цепи кДНК свежеприготовленную первую цепь кДНК осаждали этанолом и растворяли в воде до конечной концентрации кДНК ~100 нг/мкл. Для гибридизации и обработки нуклеазой использовали аликвоту 1,5 мкл полученного раствора первой цепи, как описано выше в течение 4 ч. Инкубацию в присутствии Paq_DSN продолжали 20 мин. Амплификацию нормализованной фракции оц ДНК осуществляли, как описано выше.

2.25. ДСН-нормализация геномной ДНК

Геномную ДНК человека (Clontech) фрагментировали с помощью ультразвукового процессора Cole-Parmer CP750 (Vernon Hills, IL, США) в течение 20 циклов по 3 с с интервалами 10 с между циклами до средней длины 500 т.п.о. (диапазон длин ДНК 200–750 т.п.н.). Фрагментированную ДНК очищали фенольно-хлороформной экстракции, осаждали этанолом и растворяли 20 mM Трис-НСl (рН 8,0) до конечной концентрации 1 мкг/мкл. Аликвоту объемом 1 мкл обрабатывали ДНК-полимеразой Т4 (Fermentas, Канада) в соответствии с инструкциями производителя для создания тупых концов, а затем очищали в ходе фенольно-хлороформной экстракции, осаждали этанолом и повторно растворяли в 20 mM Tris-НСl (рН 8,0) до конечной концентрации ДНК 100 нг/мкл. 5'-концы фрагментов ДНК фосфорилировали путем инкубации 8 мкл образца ДНК с 1 мкл 10-кратного лигазного буфера (Fermentas), riboАТР (конечная концентрация 3 mM) и 0,7 мкл полинуклеотидкиназы Т4 (5 ед./мкл; Fermentas) при 37°C в течение 15 мин с последующей тепловой инактивацией фермента при 60°C в течение 20 мин. Фосфорилированную ДНК лигировали с адаптером Not1S (5'-GGT CGC GGC CGA GGT-3') и амплифицировали с праймером Not1S (4 нг матрицы на 100 мкл реакционной смеси для ПЦР). После предварительного нагревания при 72°C в течение 2 мин проводили 14 циклов ПЦР в режиме: 95°C — 7 с, 62°C — 20 с и 72°C

—1 мин. После очистки продукты ПЦР осаждали этанолом и разводили водой до конечной концентрации ДНК 75 нг/мкл.

Аликвоту, содержащую ~ 150 нг ДНК объединяли с 1 мкг ДНК человека C0t-1 (Invitrogen) в 1× гибридизационном буфере. Реакционные смеси покрывали минеральным маслом, чтобы избежать испарения воды, и денатурировали при 98°C в течение 3 мин. Затем смеси инкубировали для ренатурации ДНК при 68°C в течение 5 ч (C0t ~50). После инкубации к реакционной смеси добавляли 5 мкл 2 × DSN-буфера, предварительно нагретого до 70°C. Далее в реакцию добавляли 1 мкл раствора Par_DSN (1 ед. Куница/мкл) и продолжали инкубацию в течение 20 мин при 65°C. Затем останавливали реакцию добавлением 10 мкл 5 мМ ЭДТА и разводили в 20 мкл воды. Аликвоты нормализованной оц ДНК объемом 5 мкл амплифицировали в пяти реакционных смесях с праймером Not1S в общем реакционном объеме 50 мкл. Было выполнено 18 циклов ПЦР в следующем режиме: 95°C — 7 с, 62°C — 20 с и 72°C — 1 мин. Продукты ПЦР очищали от реакционной смеси, осаждали этанолом, разбавляли водой до конечной концентрации ДНК 50 нг/мкл и использовали для клонирования библиотек ДНК и секвенирования с использованием автоматического секвенатора серии 454 GS FLX Standard (Roche, Швейцария).

2.26. ДСН-деплегция

В качестве драйвера использовали продукты амплификации нескольких генов фосфодиэстераз, приготовленные, как описано выше. В другом эксперименте в качестве дц драйвера использовали синтетические олигонуклеотиды. Было приготовлено четыре олигонуклеотида, формирующих два полностью комплементарных дуплекса: один (zGFP-Dr) с последовательностью идентичной последовательности гена флуоресцентного белка zGFP, второй (zRFP-Dr) — с последовательностью гена флуоресцентного белка zRFP. Парные олигонуклеотиды

смешивали в конечной концентрации 1 пмоль каждый. Структуры олигонуклеотидов показаны в таблице 2-6.

Очищенную кДНК трейсера осаждали этанолом и растворяли в стерильной воде до конечной концентрации 75 нг/мкл. Для каждого эксперимента аликвоту раствора объемом 2 мкл смешивали с 1 мкл 4 × гибридизационного и 1 мкл ДНК драйвера, содержащим примерно 100-кратный избыток ДНК. Образцы покрывали минеральным маслом, денатурировали при 98°C в течение 5 мин и оставляли ренатурировать. Для драйверов, состоящих из ПЦР-фрагментов, ренатурацию проводили при 70°C в течение 1 ч. При использовании в качестве драйвера синтетических олигонуклеотидов образцы инкубировали при 65°C в течение 40 мин. В конце гибридизации к каждому образцу добавляли 5 мл предварительно нагретого раствора Par_DSN в 1× DSN буфере и инкубировали при 68°C в течение 30 мин. Для остановки реакции использовали прогревание при 95°C в течение 7 мин. Образцы немедленно разводили стерильной водой до конечного объема 40 мкл и использовали аликвоты объемом 1 мкл для ПЦР с праймером M1. Всего было проведено 20 циклов ПЦР (95°C — 7 с; 65°C — 20 с; 72°C — 3 мин) для кДНК плаценты человека и 17 для кДНК кораллового полипа.

Таблица 2-6. Структуры олигонуклеотидов, использованных для DSN-деплеции последовательностей флуоресцентных белков

Дуплекс	Последовательности (5'→3') олигонуклеотидов
zGFP-Dr	ACATGTGCATACCATTTACGCTGATGACAATGTAGTTCAATTCAA ACCA _и GGTTTGAATTGAACTACATTTGTCATCAGCGTAATGGTATG CACATGT
zRFP-Dr	CGTTTATGACCATTAACTGATAAGATTGTAGTTCTAACATGCT ATTGCACGTTTATGA _и TCATAAACGTGCAATAGCATGTTAGAAC TACAATCTTATCAGGTTAATGGTCATAAACG

Для DSN-деплеции рРНК 80 нг амплифицированной кДНК прокариот, приготовленной на основе суммарной РНК с использованием случайной затравки,

денатурировали в 1 × гибридационном буфере в реакционной смеси объемом 20 мкл при 98°C в течение 3 мин и ренатурировали при 68°C в течение 4,5 ч. Затем к реакционной смеси добавляли равный объем 2 × буфера DSN и 1 ед. Куница Par_DSN. Смесь инкубировали при 67°C в течение 25 мин. Реакцию останавливали, добавляя в реакционную смесь 40 мкл 5мМ ЭДТА. Аликвоту объемом 1 мкл амплифицировали в 20 циклах ПЦР в 30 мкл реакционной смеси с праймером M1.

2.27. Вычитающая гибридизация

Вычитающую гибридизацию осуществляли с помощью набора реактивов PCR-Select DNA Subtraction Kit (Clontech) согласно протоколу.

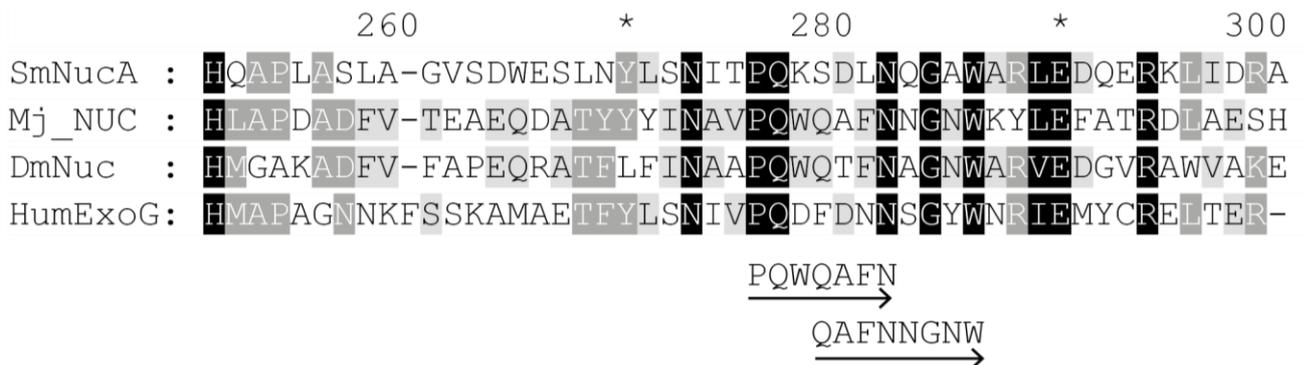
ГЛАВА 3. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

3.1. Идентификация термостабильной нуклеазы, селективно расщепляющей двухцепочную ДНК, и её характеристика

Отсутствие термостабильных ферментов специфических к дц ДНК в палитре нуклеаз, доступных для разработки методов выделения оц фракции ДНК из смесей НК, ограничивало возможности создания технологий анализа сложных образцов ДНК и кДНК. В то же время было описано, что ДНКазы некоторых организмов обладают предпочтительностью к дц ДНК по сравнению с оц ДНК. Было показано, что ДНКазы I теленка в оптимальных условиях расщепляет дц ДНК примерно в 100 раз более эффективно, чем денатурированную ДНК [89]. Для ряда беспозвоночных также было описано наличие нуклеазной активности избирательной для дц ДНК [27-30]. Из плодовой мушки *Drosophila melanogaster* была выделена митохондриальная Mg^{2+} -зависимая нуклеаза, преимущественно расщепляющая дц ДНК [29]. Температурная зависимость авторами не была исследована, данных о структуре фермента также получено не было. Другой группой исследователей был выделен термолабильный белок из креветки *Pandalus borealis*, проявляющий нуклеазную активность со специфичностью к дц ДНК [30]. Однако фермент инактивировался при умеренном нагревании с периодом полураспада 1 мин при 65°C (там же). Кодированная последовательность этого белка не была определена, и не было получено данных о его структуре. В 1990 г. была очищена ДНКазы креветки *Marspeaneus japonicus*, обладающая функциональным сходством с ДНКазой I быка [560]. Кодированная ее последовательность была идентифицирована в 2000 г. [561]. Было показано, что фермент обладает низким уровнем гидролитической активности по отношению к РНК в присутствии Mg^{2+} и Ca^{2+} , однако не было получено данных о различиях в его субстратной предпочтительности к оц и дц ДНК. Наконец, было продемонстрировано наличие Ca^{2+} , Mg^{2+} -зависимой ДНКазы К, устойчивой к прогреванию при 60°C, у камчатского краба *Paralithodes camtschaticus* [27, 28]. На образцах дц ДНК был

показан синергетический эффект ионов Ca^{2+} и Mg^{2+} на нуклеазную активность ДНКазы K, а также продемонстрировано, что фермент значительно эффективнее расщепляет дц ДНК, чем оц ДНК [27]. Нуклеаза была выделена в гомогенном состоянии, определен ее аминокислотный состав. Однако получить данные о структуре фермента авторам не удалось.

Нами было выдвинуто предположение, что термостабильный фермент камчатского краба может быть гомологичен ДНКазе креветки *M. japonicus*. Было принято решение клонировать нуклеазу из камчатского краба. Для осуществления этой задачи было проведено сравнение аминокислотных последовательностей известных нуклеаз с нуклеазой *M. japonicus* и выявлены регионы для конструирования вырожденных праймеров (Рисунок 3.1).



Под выравниванием указаны регионы, выбранные для конструирования вырожденных праймеров для двух раундов амплификации 3'-концевого фрагмента кодирующей последовательности.

Рисунок 3.1. Множественное выравнивание аминокислотных последовательностей наиболее консервативного региона нуклеазы из *M. japonicus* (Mj_NUC), нуклеазы из *S. marcescens* (SmNucA), эндонуклеазы G человека (HumEndoG) и предсказанной на основании сравнения нуклеазы из *D. melanogaster* (DmNuc).

Последовательности нуклеаз из разных организмов отличаются низким процентом идентичности. В связи с этим не удалось подобрать пары праймеров для амплификации внутреннего фрагмента кодирующей последовательности целевого белка, поэтому для изолирования нуклеотидной последовательности нуклеазы

камчатского краба была применена ранее разработанная стратегия, названная нами Step-Out RACE [535], с использованием в качестве матрицы полноразмерной амплифицированной кДНК, содержащей на обоих концах одинаковые последовательности адаптеров, а в качестве праймеров – вырожденные олигонуклеотиды к высококонсервативному участку нуклеазы креветки *M. japonicus* и специальную смесь олигонуклеотидов (Step-out mix).

Схема изолирования 3'-концевого фрагмента изображена на рисунке 3.2.

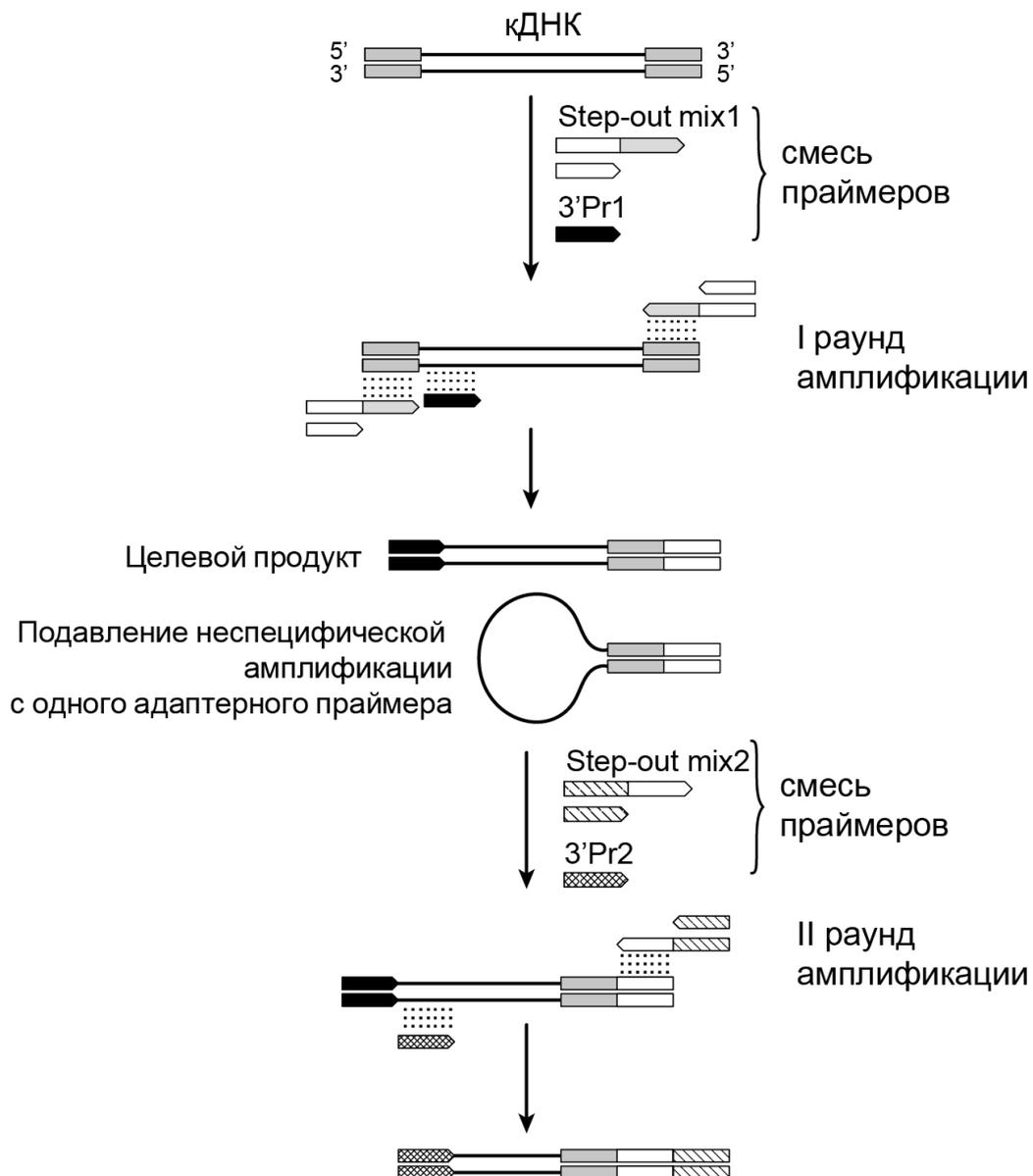


Рисунок 3.2. Схема изолирования неизвестной последовательности по технологии Step-Out 3'-RACE с использованием двух смесей для ПЦР «с выходом наружу» и двух вырожденных праймеров для заглубленной ПЦР (3'Pr1 и 3'Pr2)

Для амплификации целевой последовательности ДНК в каждом раунде ПЦР используется один вырожденный ген-специфический праймер и специальная смесь олигонуклеотидов (Step-out mix), в которой в низкой концентрации содержится длинный олигонуклеотидный праймер, комплементарный своей внутренней частью фланкирующему ДНК-матрицу адаптеру, а в высокой концентрации – короткий олигонуклеотидный праймер, идентичный внешней части длинного олигонуклеотида. Длинный олигонуклеотидный праймер быстро расходуется, и амплификация осуществляется со второго, короткого праймера. В результате происходит подавление фоновой амплификации за счет эффекта супрессии ПЦР [452, 453]. В результате эффективно нарабатывается только целевой продукт.

Данные о нуклеотидной последовательности полученного 3'-концевого фрагмента были использованы для изолирования и клонирования полной кодирующей последовательности нуклеазы. Последовательность нуклеотидов, кодирующая полноразмерную нуклеазу, была депонирована в базу данных GenBank под номером AF520591 (номер последовательности аминокислот — AAN86143). Анализ нуклеотидной последовательности нуклеазы выявил открытую рамку считывания длиной в 1221 п.о. (начиная с иницирующего кодона), кодирующую белок, состоящий из 407 аминокислот. Предсказанная аминокислотная последовательность белка демонстрировала 64% идентичности с аминокислотной последовательностью нуклеазы из *M. japonicus*, включала в себя сигнальный пептид и нуклеазный домен. Доменная организация предсказанного белка представлена на рисунке 3.3.

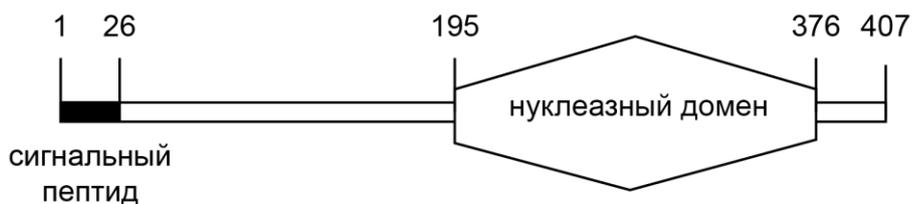


Рисунок 3.3. Доменная организация нуклеазы камчатского краба, предсказанная на основе анализа ее аминокислотной последовательности с помощью программного пакета SMART (Schultz et al. 2000) [562]

Получить активный рекомбинантный белок в гетерологической бактериальной системе экспрессии не удалось – весь продукт обнаруживался в нерастворимой фракции в тельцах включения. Поэтому было принято решение осуществить очистку природного фермента. Для контроля выделения и для подтверждения связи выделяемого природного фермента с клонированным нами белком было решено использовать поликлональные антитела (ПА-нук) против неактивного рекомбинантного продукта, выделенного в денатурирующих условиях. Для выделения природного фермента на первом этапе работы использовали свежий гепатопанкреас, в дальнейшем перешли на коммерчески доступный ацетоновый порошок из камчатского краба. Стадии хроматографического разделения фракций контролировали по анализу продуктов в денатурирующем полиакриламидном геле, наличию нуклеазной активности к ДНК и связыванию фракции в денатурирующих условиях с ПА-нук. Используя комбинацию из нескольких последовательных стадий хроматографии: анионообменная хроматография, два цикла очистки на фенил-сефарозе, хроматография на гидроксиапатитной колонке, хроматография на гепарин-сефарозе и гель-фильтрация, нами был получен гомогенный белок с высокой удельной нуклеазной активностью, и степенью очистки примерно в 618 раз от исходного. Молекулярная масса выделенного белка, определенная с помощью эксклюзионной хроматографии (~40 кДа), хорошо согласовывалась с молекулярной массой, определенной с помощью SDSGE (~44 кДа), что указывало на то, что выделенный белок присутствовал в растворе в мономерной форме.

Для подтверждения идентичности первичной структуры выделенного белка и аминокислотной последовательности, предсказанной по клонированной кодирующей нуклеотидной последовательности, дополнительно к окрашиванию антителами был использован масс-спектрометрический анализ. Параллельная масс-спектрометрия продуктов обработки очищенного препарата природной нуклеазы трипсином, выявила с высокой степенью достоверности четыре пептида: DIETSRPSFK, YLEYATR, DLAESHGSDLR и EVVPVPSLTWK. Каждая из этих аминокислотных последовательностей была представлена в аминокислотной

последовательности, предсказанной на основе анализа клонированной НК. Таким образом, идентичность денатурированного рекомбинантного белка (кодируемого клонированной нуклеотидной последовательностью гена нуклеазы) и активного фермента, очищенного из гепатопанкреаса краба, была подтверждена масс-спектрометрическим анализом пептидов и иммуноблоттингом.

Анализ физико-химических свойств выделенного природного фермента осуществляли с использованием в качестве субстрата дц ДНК. Было показано, что нуклеаза проявляет ферментативную активность только в присутствии двухвалентных катионов металлов и ингибируется ЭДТА, то есть является металл-зависимым ферментом. Данные зависимости нуклеазы от различных катионов двухвалентных металлов суммированы в Таблице 3-1 и на рисунке 3.4.

Таблица 3-1. Зависимость каталитической активности нуклеазы камчатского краба от катионов двухвалентных металлов

[C], мМ	Каталитическая активность, ед. Куница							
	Mg ²⁺	Ca ²⁺	Mn ²⁺	Co ²⁺	Cd ²⁺	Mg ²⁺ + Ca ²⁺	Mn ²⁺ + Ca ²⁺	Co ²⁺ + Ca ²⁺
0	0	0	0	0	0	0	0	0
0,5			4,6	3	2	5	30	25
1	0,5	0,1	23,7	15	4	17	70,1	60
2	1	0,2	69	40	4,2	49,6	150	130
3	3,1	0,7	115	60	4,1	89,9	220	170
4	5,6	0,8	149,5	80	4,3	119,4	270	195
5	8	0,9	189,8	95	4,5	136,4	290	210
6	9,5	1	191,2	110	4,7	141,1	300	230
7	16	1,5	201	135,1	4,9	150,4	303,6	243
8	17	1,7	212	165	5,1	152	305	250
9	18	1,9	220	178	5,4	156,6	302	255
10	19	2	231,4	185,4	5,3	152	301	258,7
15	21,5	2,3	248,7	183,2	5,2			
17,5	23	2,4	251	182	5,4			

20	23	2,5	261,6	181,1	5,3			
22,5	23	2,5	251	170	5,2			
25	21	2,5	244,4	155	5,1			
30	21	2,5	241	140	4,8			
40	16,5	2	230	130	4			
50	10,5	1,5	221,4	119,3	3			
100	3,8	0,5	112	61,8	1			
[С] — концентрация. Жирным шрифтом выделены оптимальные концентрации для разных ионов								

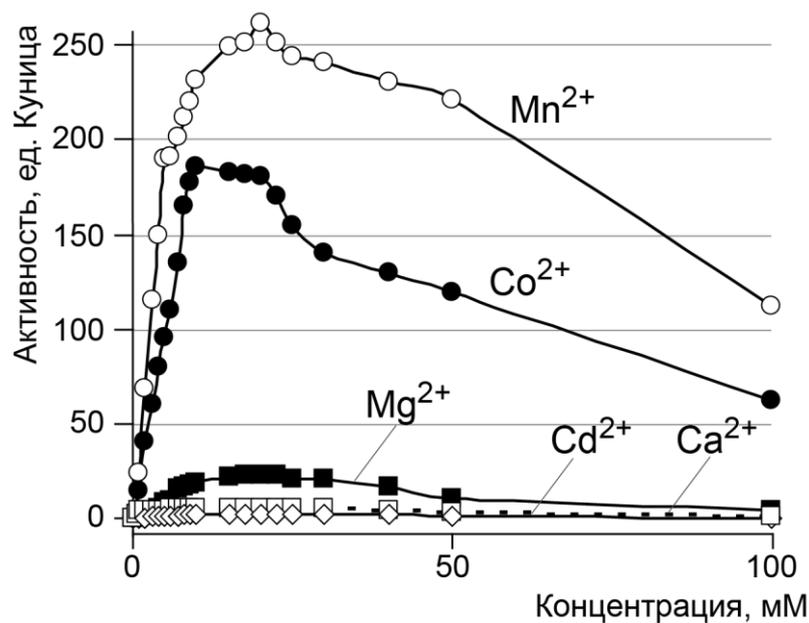


Рисунок 3.4. Зависимость активности нуклеазы от концентрации двухвалентных катионов

Максимальный уровень активности наблюдался в присутствии ионов марганца в концентрации 20 мМ. В присутствии оптимальной концентрации ионов кобальта (10 мМ) фермент демонстрировал 70% от максимальной активности, а в присутствии оптимальных концентраций ионов магния (7–30 мМ) только 10%. В присутствии только ионов Ca^{2+} и Cd^{2+} фермент не проявлял нуклеазной активности. Добавление к реакционной смеси, содержащей ионы магния, ионов Ca^{2+} производило синергетический эффект, увеличивая активность фермента примерно в пять раз при оптимальной концентрации обоих ионов. Этот эффект при

комбинациях ионов Ca^{2+} и Co^{2+} или Ca^{2+} и Mn^{2+} был выражен в значительно меньшей степени (рисунок 3.5). Для остальных комбинаций исследуемых ионов синергетического эффекта обнаружено не было.

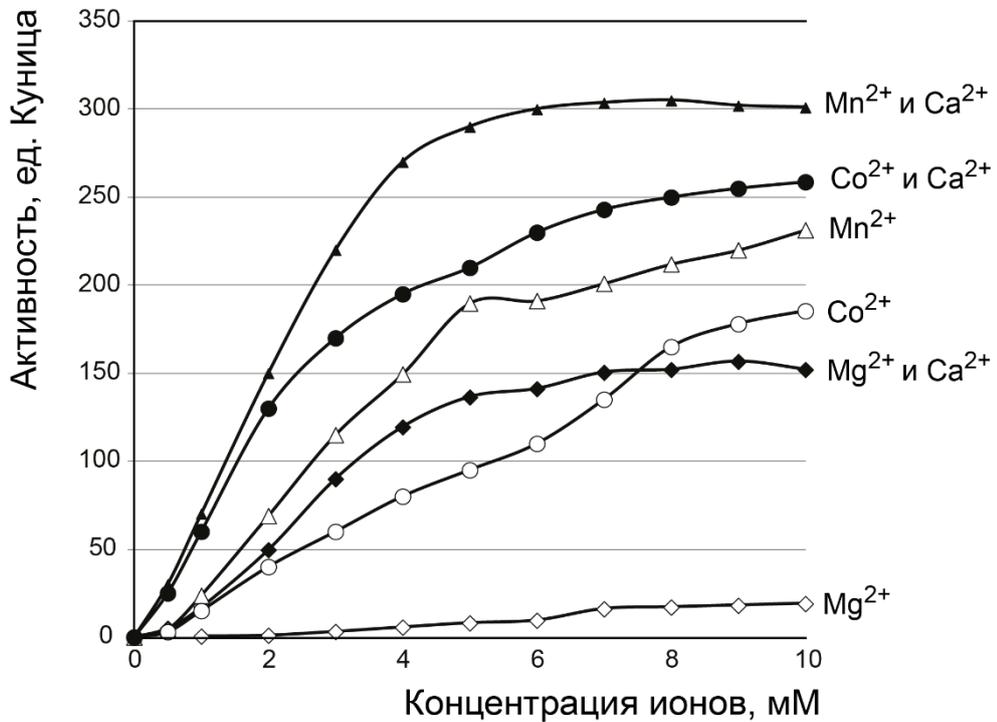


Рисунок 3.5. Зависимость активности нуклеазы от концентрации двухвалентных катионов

Фермент проявлял активность в широком диапазоне температур. При использовании в качестве субстрата дц ДНК температурный оптимум наблюдался при $\sim 60^\circ\text{C}$. Дальнейшее повышение температуры сопровождалось резким снижением каталитической активности (рисунок 3.6). Это могло быть связано как с денатурацией фермента, так и (с большей вероятностью) с плавлением дц ДНК-субстрата. Кинетика тепловой денатурации нуклеазы показана на рисунке 3.7. Инкубация в течение 30 мин при температуре 70°C сопровождалась 30% падением активности, при 80°C – 60% падением активности. Тем не менее, 5% активности фермента сохранялось даже при прогревании при 100°C . На рисунке 3.8. приведен график температурной стабильности нуклеазы.

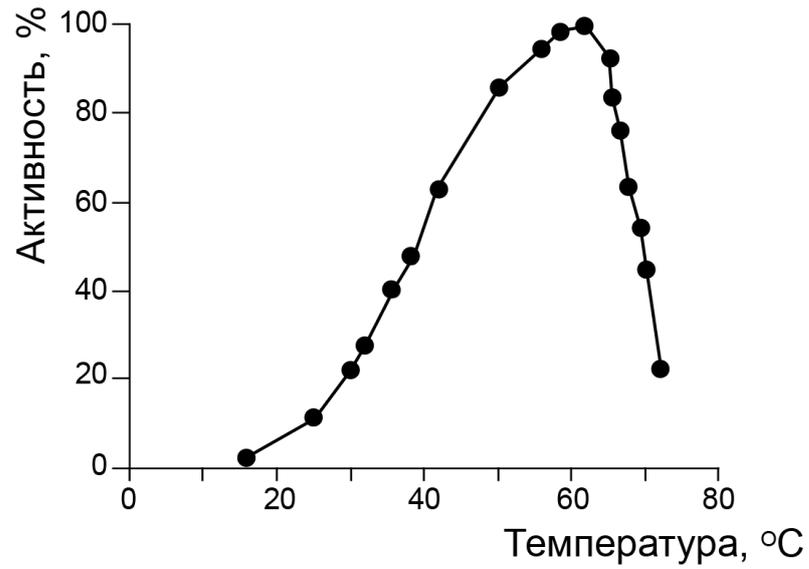


Рисунок 3.6. Зависимость активности нуклеазы от температуры

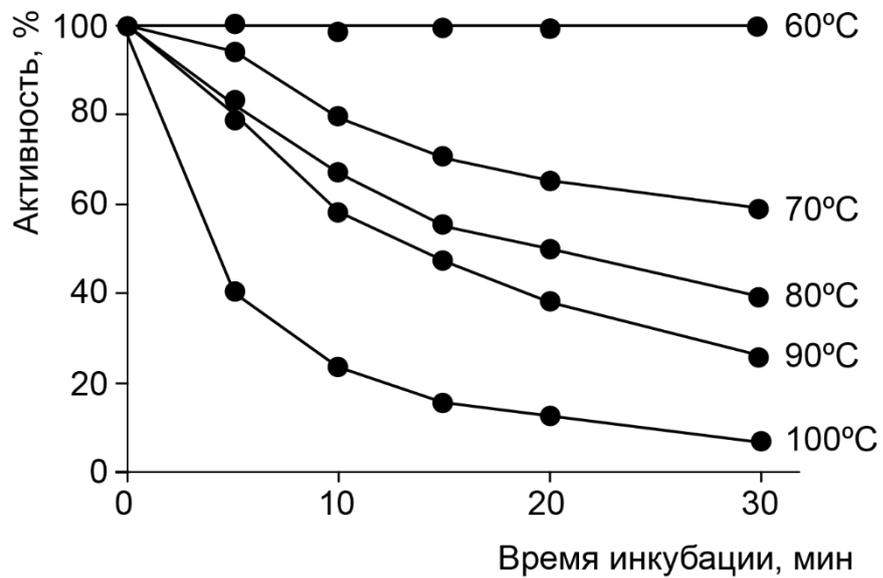


Рисунок 3.7. Кинетика тепловой денатурации нуклеазы камчатского краба при 30 мин инкубации при разных температурах

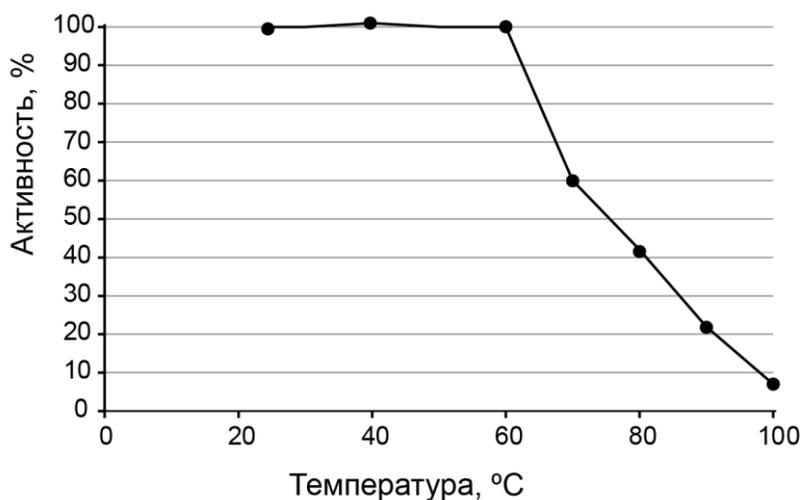


Рисунок 3.8. Температурная стабильность нуклеазы, определенная по остаточной активности после тепловой инкубации в течение 30 мин

Анализ рН стабильности фермента показал, что он сохраняет более 90% активности после 30 мин инкубации в диапазоне рН от 4 до 12. Фермент имеет широкий оптимум рН с максимумом при рН 6,8. При значениях рН менее 3,0 и более 9,0 детектируемой каталитической активности фермента не обнаруживалось (таблица 3-2).

Таблица 3-2. Зависимость каталитической активности нуклеазы камчатского краба от рН реакционной смеси

рН	Активность, ед. Куница	рН	Активность, ед. Куница
3,2	4,5	6,8	83
3,6	7,6	7	76,5
4	8,9	7,2	73,5
4,6	12,1	7,4	71,5
5	15,7	7,8	41,4
5,2	18,2	8	40,1
5,6	40,4	8,2	32,7
5,8	48,5	8,6	17,4

6	62	8,8	7,4
6,6	81,5	9	1,3

Повышение ионной силы раствора сопровождалось снижением каталитической активности нуклеазы. Так, добавление 0,2 М NaCl снижало ферментативную активность нуклеазы в 10 раз, а при концентрации 0,5 М NaCl сохранялось всего около 1,5% ее активности. Мочевина была менее эффективна в ингибировании ферментативной активности, чем NaCl, и для 10-кратного ингибирования активности нуклеазы камчатского краба требовалось 3 М мочевины. При 4 М мочевины фермент сохранял 3,5% активности (Таблица 3-3).

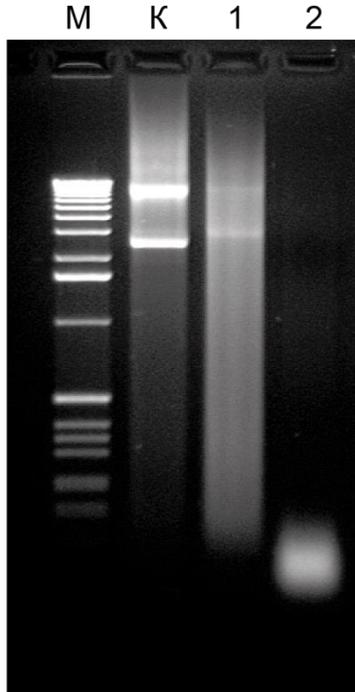
Таблица 3-3. Зависимость каталитической активности нуклеазы камчатского краба от концентрации мочевины и ионной силы раствора

Концентрация, моль	Активность, ед. Куница	
	NaCl	мочевина
0	110	110
0,05	66	
0,1	30	97
0,15	19	
0,2	11,5	
0,25	8	
0,5	1,6	60
1	0,3	47
2	0	19
4	0	3,0909091

Инкубация нуклеазы в течение 1 ч при 37°C в присутствии 1% SDS, 10 мМ меркаптоэтанола или 0.3% перекиси водорода снижала активность фермента только на 8-10%. Однако инкубация в присутствии 1% SDS при 60°C полностью

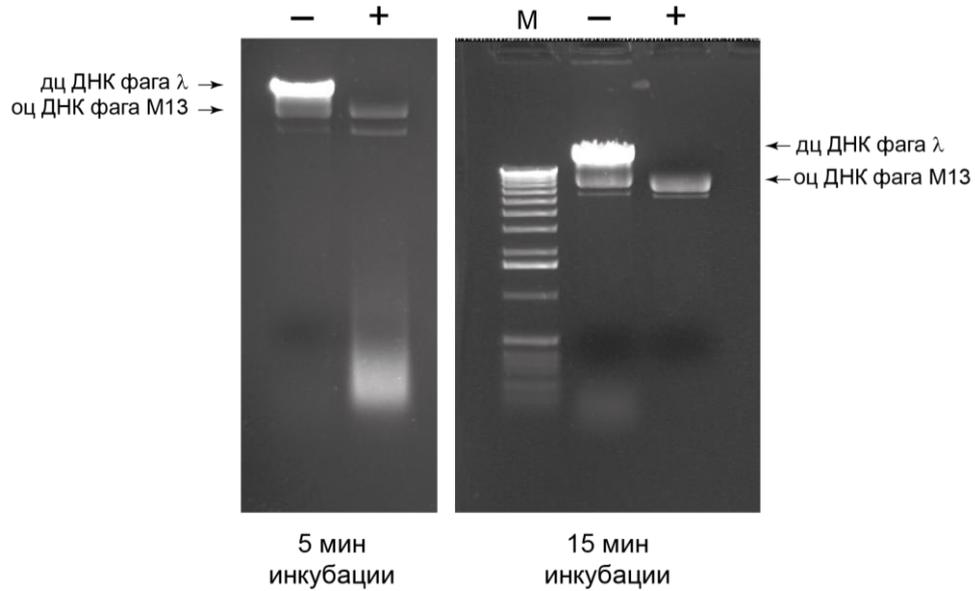
инактивировала фермент. А в присутствии меркаптоэтанола и перекиси водорода в тех же условиях приводила к потере 70% и 80% активности, соответственно. Инкубация со спермидином в концентрациях менее 5 мМ не оказывала ингибирующего действия на нуклеазу камчатского краба, однако при 10 мМ спермидина наблюдалось падение ее активности до 50%, а при концентрации 8 мМ – до 7-8%. Кроме того, было выявлено, что фермент обладает устойчивостью к обработке протеиназой К и папаину. Инкубация в течение 30 мин в присутствии этих протеиназ при 37°C не приводила к изменению активности фермента.

На образцах природной оц и дц ДНК было выявлено, что фермент обладает селективной избирательностью по отношению к дц ДНК, оставляя оц ДНК интактной после 10-15 мин инкубации при 70°C (рисунки 3.9 – 3.11). Нуклеаза расщепляла любые дц ДНК субстраты (геномную ДНК, фаговую ДНК, суперскрученную плазмидную ДНК), но не проявляла заметной активности против оц ДНК.



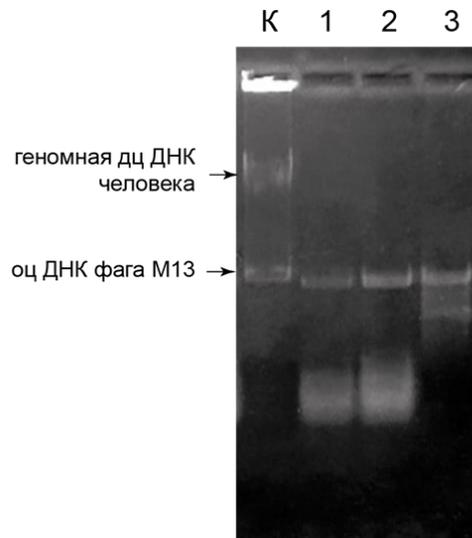
Контроль (К) — инкубация 200 нг плазмидной ДНК без нуклеазы; дорожки 1, 2 — инкубация 200 нг плазмидной ДНК при 65°C в течение 1,5 мин (дорожка 1) и 7 мин (дорожка 2) в присутствии 0,2 ед. Куница/мкл нуклеазы. М — маркер длин ДНК, 1 kb ladder (Invitrogen, США).

Рисунок 3.9 — Анализ расщепления плазмидной ДНК нуклеазой камчатского краба



Электрофорез в агарозном геле продуктов реакции расщепления фаговой дц и оц ДНК в присутствии 0,2 ед. Куница/мкл (+) или в отсутствии (-) нуклеазы камчатского краба в ходе инкубация при 65°C. Время инкубации указано под электрофореграммами. М — маркер длин ДНК, 1 kb ladder (Invitrogen, США).

Рисунок 3.10 — Анализ расщепления фаговой дц и оц ДНК нуклеазой камчатского краба



Контроль (К) — инкубация смеси геномной дц ДНК человека и оц ДНК фага M13 при 65°C в течение 9 мин без нуклеазы; дорожка 1 — инкубация в течение 3 мин в присутствии нуклеазы (0,2 ед. Куница/мкл); дорожка 2 — инкубация в течение 9 мин в присутствии нуклеазы (0,2 ед. Куница/мкл); дорожка 3 — инкубация в течение 9 мин в присутствии нуклеазы (2 ед.

Куница/мкл).

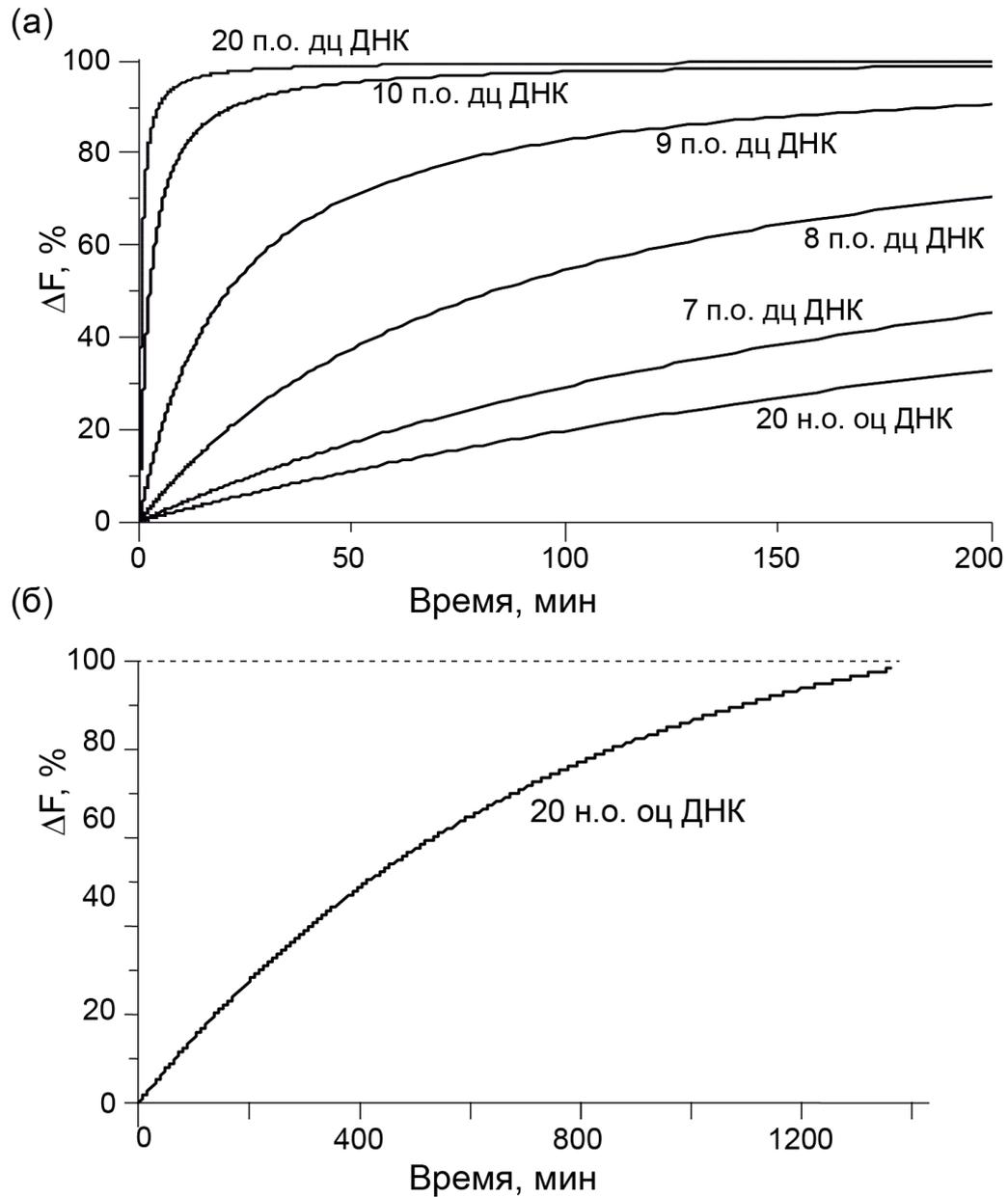
Рисунок 3.11 — Электрофорез в агарозном геле продуктов реакции расщепления геномной дц ДНК и фаговой оц ДНК в присутствии нуклеазы камчатского краба

Сравнение эффективности расщепления оц и дц ДНК далее проводили с использованием синтетических оц и дц олигонуклеотидов. Для формирования дц ДНК субстратов меченный донором и гасителем флуоресценции 20-мерный олигонуклеотид гибридизовали в эквимольных количествах с комплементарными ему немечеными олигонуклеотидами разной длины. Структуры олигонуклеотидных субстратов и время их полурасщепления ($T_{1/2}$) представлены в таблице 3-4. Результаты анализа показаны на рисунке 3.12.

Таблица 3-4. Каталитическая активность нуклеазы по отношению к синтетическим ДНК субстратам

Субстрат	Длина дц ДНК дуплекса	$T_{1/2}$, мин
5' -TAMRA-tgggcagtgctcgcttagtg-DABCYL-3'	0	406,8 ± 37,0
5' -TAMRA-tgggcagtgctcgcttagtg-DABCYL-3' 3' -cacgagc-5'	7	241,9 ± 39,2
5' -TAMRA-tgggcagtgctcgcttagtg-DABCYL-3' 3' -cacgagcg-5'	8	83,2 ± 15,5
5' -TAMRA-tgggcagtgctcgcttagtg-DABCYL-3' 3' -tcacgagcg-5'	9	20,8 ± 6,15
5' -TAMRA-tgggcagtgctcgcttagtg-DABCYL-3' 3' -tcacgagcga-5'	10	2,4 ± 0,4
5' -TAMRA-tgggcagtgctcgcttagtg-DABCYL-3' 3' -accggtcacgagcgaatcac-5'	20	0,6 ± 0,2

В серии подобных экспериментов с использованием различных концентраций нуклеазы было показано, что эффективный гидролиз синтетических дц ДНК субстратов наблюдался уже в присутствии 0,01 ед. Куница и при низких концентрациях субстрата, тогда как детектируемая дегградация оц ДНК субстратов происходила только в присутствии 10 ед. Куница и высокой концентрации субстрата.



Относительное изменение флуоресценции реакционной смеси (ΔF) определяли по формуле:

$$\Delta F = (F_R - F_0 / F_M - F_0) \times 100\%,$$

где F_R — уровень флуоресценции реакционной смеси после инкубации в присутствии нуклеазы,

F_M — уровень флуоресценции реакционной смеси, содержащей полностью гидролизованный субстрат, F_0 — фоновая флуоресценция реакционной смеси в отсутствие фермента.

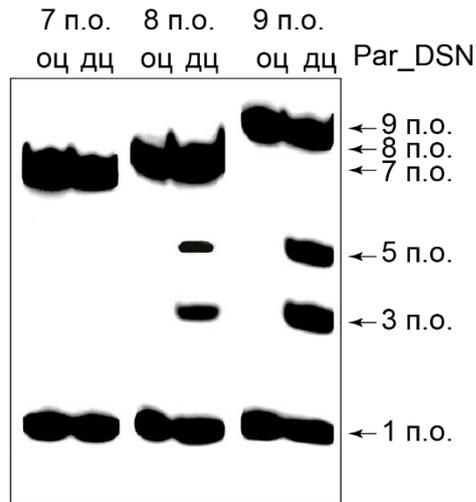
Использовали 0,2 ед. Куница / мкл нуклеазы, инкубация проводилась при 30°C.

Рисунок 3.12. Анализ расщепления синтетических оц и дц ДНК субстратов нуклеазой камчатского краба

Подсчитанное соотношение гидролитической активности фермента к дц и оц ДНК субстрату составило не менее 1000:1, что значительно выше, чем для всех известных ранее нуклеаз. При изменении условия гидролиза (добавление в реакционную среду различных комбинаций активирующих ионов двухвалентных металлов, детергентов, хаотропных агентов, полиаминов) это соотношение не менялось. Благодаря столь высокой избирательности нуклеазы по отношению к дц ДНК, она была нами названа дуплекс-специфической нуклеазой (DSN или Par_DSN, от. англ. *Paralithodes camchaticus duplex-specific nuclease*).

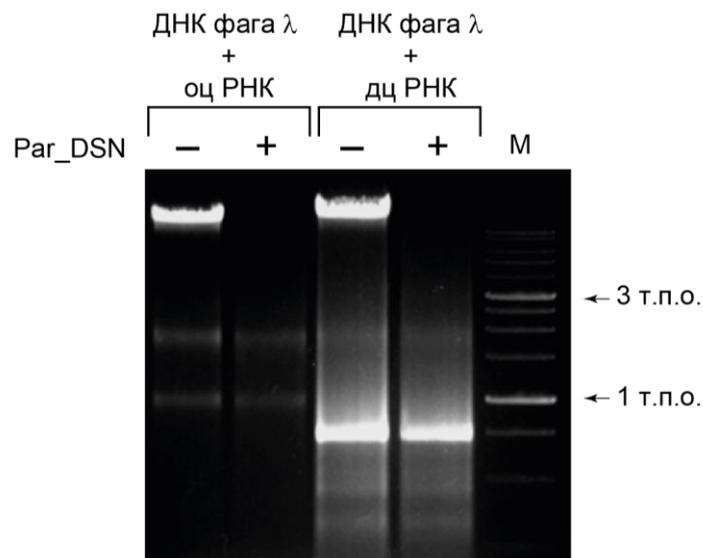
Для анализа особенностей расщепления субстрата Par_DSN использовали радиоактивно меченые олигонуклеотиды длиной (длиной 7–9 н.о.). Дц ДНК субстраты формировали гибридизацией этих олигонуклеотидов с комплементарной областью немеченого олигонуклеотида длиной 43 н.о. В ходе инкубации радиоактивно меченых субстратов в течение 1 ч при 30°C детектируемого расщепления 7-звеньевых дц ДНК-дуплексов не наблюдалось, дуплексы длиной 8 п.н. разрушались частично, а дуплексы длиной 9 п.о. — расщеплялись полностью (рисунок 3.13). ДНК-ДНК субстраты длиной 10 п.о. фермент расщеплял в основном между четвертым и пятым или пятым и шестым нуклеотидами считая от 5'-конца, образуя продукты длиной от пяти и меньше п.о.

Нуклеаза камчатского краба не проявляла детектируемой активности по отношению к оц и дц РНК-субстратам (рисунок 3.14), в тоже время она эффективно расщепляла ДНК цепь в ДНК-РНК гибридах (рисунок 3.15).



Длины и структуры, использованных ДНК-субстратов указаны над дорожками. Справа показан размер продукта, которому соответствует данная полоса.

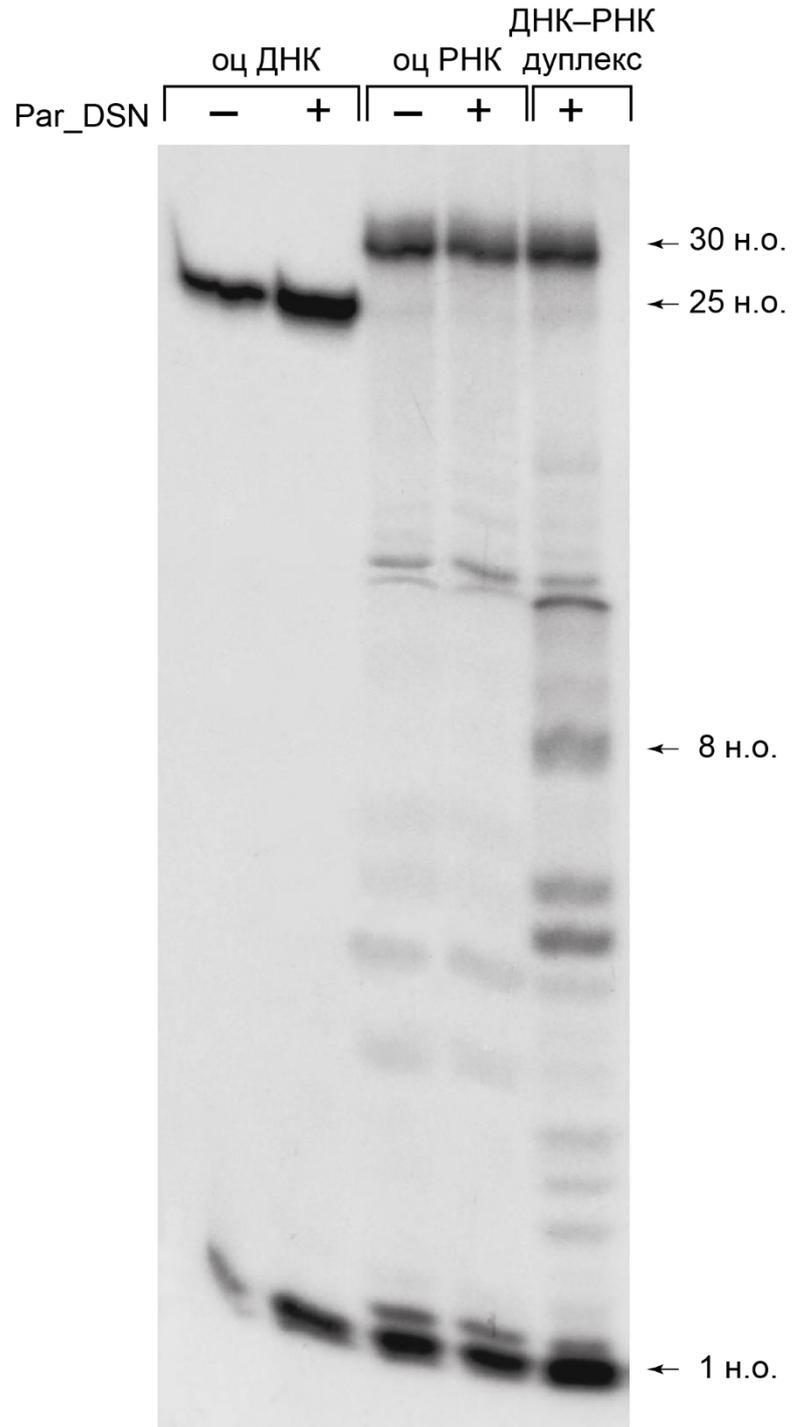
Рисунок 3.13. Анализ расщепления синтетических радиоактивно меченых дц ДНК субстратов нуклеазой камчатского краба



Суммарную клеточную РНК (оц РНК) и синтетическую дц РНК инкубировали в присутствии нуклеазы камчатского краба течение 30 мин при 37°C. В качестве положительного контроля к РНК добавляли дц ДНК фага λ. Продукты реакции визуализировали с помощью электрофореза в агарозном геле. Используемая комбинация субстратов указан над дорожками.

М — маркер длин ДНК, 1 kb ladder (Сибэнзим, Россия). Т.п.о. — тысяч пар оснований.

Рисунок 3.14. Анализ активности нуклеазы камчатского краба по отношению к оц и дц РНК



Олигодезоксирибонуклеотид длиной 25 н.о. и комплементарный ему олигорибонуклеотид длиной 30 н.о. инкубировали 1 ч по отдельности и в смеси при 45°C в присутствии (+) или в отсутствии (-) дуплекс-специфической нуклеазы. Продукты гидролиза анализировали электрофорезом в ПААГ и визуализировали с помощью автордиографии.

Рисунок 3.15. Расщепление синтетических ДНК-РНК гибридов нуклеазой камчатского краба

Было проведено сравнение свойств препарата выделенной нами нуклеазы и препарата ДНКазы К, охарактеризованной коллегами из Дальневосточного института биоорганической химии [27, 28]. Несмотря на совпадение ряда полученных в независимых экспериментах характеристик, таких как термостабильность и широкий оптимум рН, по ряду параметров ферменты существенно различались. Особенно значительными были отличия в предпочтении различных ионов двухвалентных металлов для оптимальной активности (таблица 3-5).

Таблица 3-5. Сравнение свойств Par_DSN и ДНКазы К

Фермент	ДНКазы К (Мензорова и др., 1993; Мензорова и др., 1994)	Par_DSN
Влияние ионов двухвалентных металлов	$Mg^{2+} + Ca^{2+} > Mg^{2+} > Mn^{2+} > Co^{2+}$	$Mn^{2+} > Co^{2+} > Mg^{2+} + Ca^{2+} > Mg^{2+}$
Температурный оптимум	50°C	59–60°C
рН оптимум	6,0–8,5	6,6–7,4
Влияние 2-меркаптоэтанола	100% активность при 20 мин инкубации при 60°C	30% активность при 30 мин инкубации при 60°C
рН стабильность	4–12	4–12
NaCl, 50% инактивации	0,1 М	< 0,05
NaCl, 95% инактивации	0,3 М	~ 0,3 М
Мочевина, 50% остаточная активность	1 М	1 М
Мочевина, инактивация	2,5 М	> 4М
Активность после 30 мин прогревания при 100	10%	7%
Молекулярная масса по данным гель-электрофореза в ПААГ	42 кДа	44 кДа
Связывание с ПА-нук	+?	+++
Аминокислотный состав	357 аа (предполагаемое)	407 аа
Дисульфидных связей	18	5

Заявленная степень очистки	1100 раз	617 раз
Процедура очистки	анионообменная хроматография (ДЭАЭ-Toyopearl), хроматография на оксиапатите, два цикла хроматографии на моноQ, гель-фильтрация	анионообменная хроматография (ДЭАЭ-MacroPrep), два цикла очистки на фенил-сефарозе, хроматография на гидроксиапатитной колонке, хроматография на гепарин-сефарозе, гель-фильтрация

Par_DSN продемонстрировала также меньшую устойчивость к ионной силе раствора и присутствию 2-меркаптоэтанола, в то же время была несколько более устойчива к высоким концентрациям мочевины. Оба фермента не проявляли активности по отношению к РНК и ДНК фага M13, и эффективно расщепляли дц ДНК. Однако, в отношении образцов гибридизованной после ренатурации кДНК ферменты вели себя по-разному: после обработки Par_DSN в течение 5 мин при 68°C оц кДНК-фракция сохранялась без заметной потери длины, в то время как обработка ДНКазой К в тех же условиях приводила к заметной деградации оц кДНК. Различия наблюдались также в длине и составе аминокислотной последовательности, количестве потенциальных дисульфидных связей (таблица 3-5). Тем не менее, препарат ДНКазы К давал слабый сигнал с антителами ПА-нук. Анализ обработанного трипсином препарата ДНКазы К с помощью масс-спектрометрии показал наличие загрязнения по крайней мере коллагенолитическими протеазами, что не позволило выявить входящие в ее состав пептидные последовательности. Загрязнение другими белками препаратов ДНКазы К может объяснить наблюдаемые различия в аминокислотном составе, но не разницу в существенных ферментативных и физико-химических свойствах, имеющих отношение к катализу. Таким образом, вопрос об идентичности выделенного нами фермента и ДНКазы К остался не до конца закрытым. Возможно предположить наличие у камчатского краба еще одного родственного фермента-

нуклеазы, имеющего несколько иные свойства, нежели были описаны для Par_DSN.

В связи с тем, что препарат ДНКазы К оказался непригоден для удаления ДНК-дуплексов при гибридизации денатурированной ДНК, в дальнейшей работе мы сосредоточились на исследовании Par_DSN и ее применении к анализу сложных смесей НК.

Таким образом, нами был клонирован и охарактеризован термостабильный фермент, обладающий непревзойденной селективностью по отношению к дц ДНК–дуплекс-специфическая нуклеаза Par_DSN. Было показано, что Par_DSN практически неактивна по отношению к РНК и оц-ДНК субстрату и проявляет свойства металл-зависимой ДНКазы по отношению к ДНК в ДНК-ДНК и ДНК-РНК гибридах. Минимальная длина дуплекса, разрушаемого Par_DSN, составляет 8 п.о. Важно отметить, что Par_DSN сочетает в себе высокую специфичность и высокую стабильность, что и обусловило большую привлекательность этого фермента как инструмента для биотехнологии.

К сожалению, Par_DSN, экспрессированная в *E. coli*, была локализована в тельцах включения, и наши попытки экспрессировать функциональный рекомбинантный белок сначала не увенчались успехом. Однако мы смогли использовать очищенный, денатурированный рекомбинантный белок для получения препарата антител и разработать успешную стратегию очистки нуклеазы из гепатопанкреаса камчатского краба. Вскоре Анисимовой с соавт. удалось разработать схему реактивации полученного рекомбинантного белка Par_DSN, основанную на устойчивости этой нуклеазы к обработке протеиназой К [563]. Свойства реактивированного рекомбинантного белка оказались идентичны свойствам выделенной из камчатского краба нуклеазы. Это стало окончательным подтверждением удачно проведенного исследования.

3.2. Идентификация семейства нуклеаз, специфически расщепляющих дц ДНК, и его структурно-функциональная характеристика

Наше исследование структуры и свойств ДСН гепатопанкреаса камчатского краба было в значительной степени индуцировано идентификацией нуклеазы из гепатопанкреаса креветки *M. japonicus*. Анализ ее аминокислотной последовательности выявил сходство ее нуклеазного домена с нуклеазными доменами Serratia-подобных ДНК-РНК неспецифических эндонуклеаз [561]. Serratia-подобные ДНК-РНК неспецифические эндонуклеазы (SNF от англ. Serratia family of nucleases), архетипом которых является нуклеаза SmNucA, относятся к суперсемейству His-Me [261]. Большинство SNF-нуклеаз расщепляют ДНК и РНК в дц и оц форме с одинаковой эффективностью или, как эндонуклеаза из *Syncephalastrum racemosum*, с небольшим предпочтением к ДНК [84, 268, 564]. Для всех представителей семейства был постулирован сходный механизм работы и наличие единственного каталитического центра, связывающего как ДНК, так и РНК. Характерной особенностью структурной организации SNF является наличие так называемого нуклеазного (NUC) домена с консервативным мотивом RGH, в состав которого входит гистидиновый остаток каталитического центра [84, 264].

Нуклеаза креветки эффективно расщепляла ДНК и проявляла слабую активность в отношении РНК субстратов. Par_DSN обнаруживала 64% идентичности аминокислотной последовательности с нуклеазой креветки и сходную комбинацию аминокислотных остатков каталитического центра, предсказанного на основе сравнения ее аминокислотной последовательности с последовательностями SNF-эндонуклеаз. В то же время Par_DSN проявляла уникальную субстратную специфичность, расщепляя только ДНК в составе дц молекул ДНК-ДНК и ДНК-РНК. Кроме того, Par_DSN обладала выраженной термостабильностью, не характерной для SNF. В литературных источниках не было описания подобной активности. Механизм, лежащий в основе особых свойств Par_DSN, требовал дальнейшего изучения.

В 2006 г. появились данные об аминокислотной последовательности еще одной нуклеазы, обладающей сходной с Par_DSN субстратной специфичностью – CiquEndo из семиполосного комара [565]. Интересно, что CiquEndo также была отнесена авторами к SNF нуклеазам (там же).

Мы предположили, что сравнительный анализ Par_DSN, нуклеазы креветки *M. japonicus* и CiquEndo с «классическими» членами SNF позволит выявить особенности этих ферментов, обуславливающие их уникальную субстратную специфичность. Так как точность множественного выравнивания последовательностей тем выше, чем больше используется родственных последовательностей [566], было принято решение о клонировании гомологов Par_DSN из других ракообразных и проведении дополнительного поиска родственных последовательностей в доступных базах данных.

С помощью технологии, которую мы использовали для клонирования нуклеотидной последовательности Par_DSN, нами были получены полные кодирующие последовательности потенциальных нуклеаз трех видов ракообразных: Gam_Nuc из рачка *Gammarus* sp., Pal_Nuc из креветки *Paleomonidae* sp. и Uca_Nuc из краба *Uca crassipes*. Эти последовательности были депонированы в базу данных GenBank: последовательность нуклеотидов Gam_Nuc DQ862539, Pal_Nuc — DQ862538, Uca_Nuc — DQ862540; последовательности аминокислот — ABI18973, ABI18972 и ABI18974 соответственно. Предсказанные аминокислотные последовательности Pal_Nuc и Uca_Nuc имели 50-65% идентичности с последовательностью Par_DSN, в то время как последовательность Gam_Nuc — только 32% идентичности. С помощью ПО SignalP-3.0/5.0 [567, 568] во всех предсказанных белках ракообразных (включая идентифицированные нами нуклеазы и нуклеазу креветки Mj_Nuc) был обнаружен сигнальный пептид, указывающий на возможный секреторный характер их экспрессии (таблица 3-б). Однако точность выявления сигнального пептида у Gam_Nuc оказалась существенно ниже, чем у других исследуемых нуклеаз. Согласно анализу, выполненному с помощью ПО SMART, Pal_Nuc, Uca_Nuc и Gam_Nuc имеют доменную организацию сходную с организацией доменов Par_DSN и Mj_Nuc,

содержат сигнальный пептид и отделенный от него линкерной последовательностью нуклеазный домен, гомологичный домену эндонуклеаз SNF (таблица 3-7). В структуре Gam_Nuc в линкерной последовательности был также выявлен трансмембранный helix регион (таблица 3-7).

Таблица 3-6. Результат анализа аминокислотных последовательностей нуклеаз ракообразных с помощью программного обеспечения SignalP-5.0

Белок	Вероятность наличия сигнальной последовательности	Вероятность, что последовательность не является сигнальным пептидом	Позиция отрезания сигнального пептида и вероятность сделанного предсказания (Pr)
Par_DSN	0.975693	0.024307	26-27 (Pr = 0.9553)
Mj_Nuc	0.986276	0.013724	21-22 (Pr = 0.9460)
Pal_Nuc	0.946196	0.053804	23-24 (Pr = 0.8938)
Uca_Nuc	0.975649	0.024351	26-27 (Pr = 0.7060)
Gam_Nuc	0.505635	0.494365	28-29 (Pr = 0.1721)

Таблица 3-7. Доменная организация нуклеаз ракообразных, предсказанная с помощью программного обеспечения SMART

Белок	Сигнальный пептид	Трансмембранный регион	Нуклеазный домен	Длина полипептида, aa
Par_DSN	1–27	не выявлено	195–376	407
Mj_Nuc	1–21	не выявлено	190–372	402
Pal_Nuc	1–24	не выявлено	146–374	404
Uca_Nuc	1–27	не выявлено	196–375	405

Gam_Nuc	1–22	44–66	137–368	392
Приведены позиции аминокислот начала и конца выявленного региона/домена и длина предсказанной аминокислотной последовательности (aa – аминокислоты). Номера аминокислотных остатков указаны с учетом первого метионина				

К сожалению, сделать окончательный вывод о структуре N-конца Gam_Nuc оказалось затруднительно. Анализ с помощью ПО TMSEG [569], предназначенного для выявления трансмембранных регионов, с высокой достоверностью выявил в последовательности Gam_Nuc сигнальный пептид длиной 20 аминокислотных остатков. Одновременно анализ с помощью альтернативного ПО ТМНММ-2.0 [570] и модуля RePROF для предсказания спиральных трансмембранных белков пакета программ PredictProtein [571] указывал скорее на наличие двух трансмембранных регионов (рисунок 3.16). По данным ТМНММ-2.0 вероятность нахождения С-концевой области белка в цитоплазме оценивается в 98%.

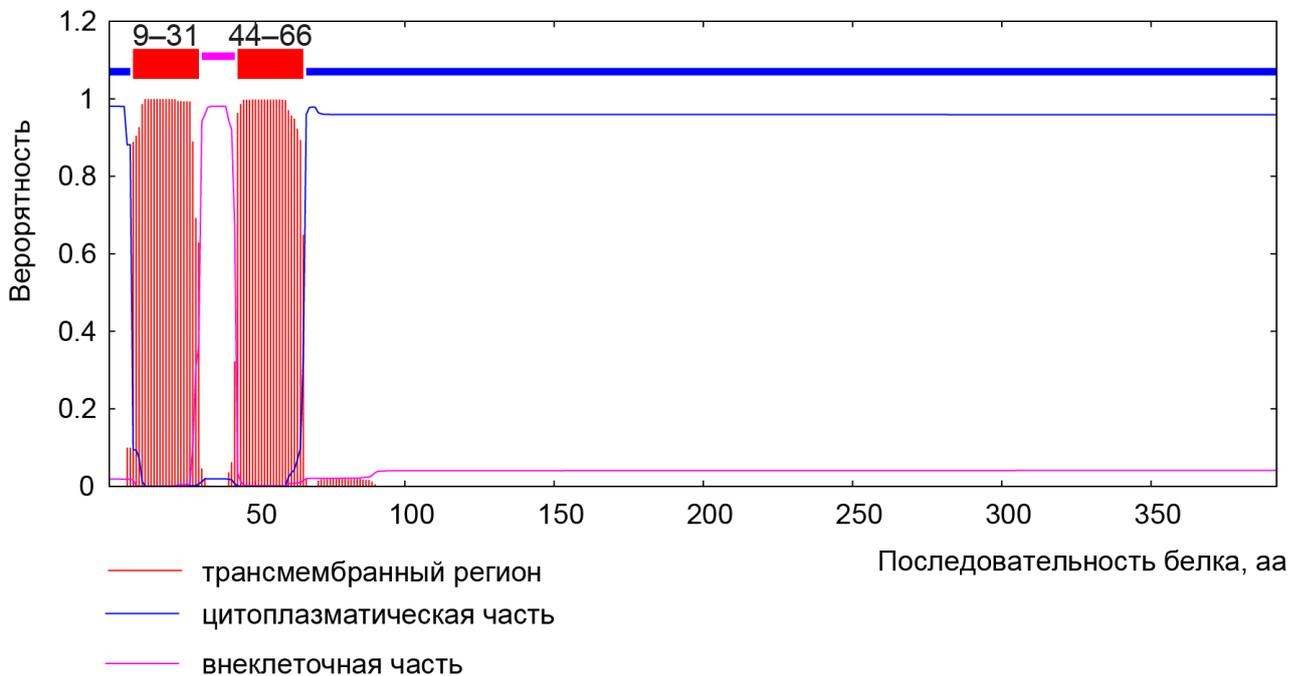


Рисунок 3.16. Анализ аминокислотной последовательности Gam_Nuc с помощью программного обеспечения ТМНММ-2.0.

Дополнительно в доступных базах данных с помощью пакета программ NCBI BLAST был выявлен ряд последовательностей гомологов нуклеаз ракообразных. Большая часть найденных аминокислотных последовательностей относилась к неохарактеризованным белкам, предсказанным на основе данных геномного секвенирования насекомых (таблица 3-8). Также в таблицу добавлены сведения об использованных в выравнивании нуклеазах SNF и о нуклеазах ракообразных.

Таблица 3-8. Аминокислотные последовательности, отобранные для сравнительного анализа

Название	Организм	Номер в GenBank ¹	Краткая аннотация
Par_DSN	камчатский краб <i>P. camtschaticus</i> (ракообразные)	AAN86143	дц ДНК специфическая эндонуклеаза
Gam_Nuc	гаммарус <i>Gammarus sp.</i> (ракообразные)	ABI18973	Последовательность аминокислот, предсказанная на основе последовательности мРНК
Mj_Nuc	креветка кумура <i>M. japonicus</i> (ракообразные)	CAB55635	ДНКаза со слабой РНКазной активностью [560, 561]
Pal_Nuc	креветка <i>Palaemonidae sp.</i> (ракообразные)	ABI18972	Последовательность аминокислот, предсказанная на основе мРНК
Uca_Nuc	краб-скрипач <i>Uca crassipes</i> (ракообразные)	ABI18974	Последовательность аминокислот, предсказанная на основе последовательности мРНК
Gm_Tsal1	муха цеце <i>Glossina morsitans morsitans</i> (насекомые)	AAF82097	Последовательность аминокислот, предсказанная на основе последовательности мРНК из слюнных желез [572]

Gm_Tsal2	<i>G. morsitans morsitans</i> (насекомые)	AAF82098	Последовательность аминокислот, предсказанная на основе последовательности мРНК из слюнных желез [572]
Cmae_Nuc	зеленый краб <i>Carcinus maenas</i> (ракообразные)	нет	Частичная последовательность аминокислот, предсказанная на основе мРНК, депонированных в GenBank DN738756; DV944764 в рамках проекта по исследованию EST [573]. Содержит неполный С-конец
Нaме_Nuc	лобстер <i>Homarus americanus</i> (ракообразные)	нет	Частичная последовательность аминокислот, предсказанная на основе мРНК, депонированных в GenBank CN853059 в рамках проекта по исследованию EST [573]. Содержит неполный С-конец
Dm_CG14120	плодовая муха <i>D. melanogaster</i> (насекомые)	NP_648610	Последовательность аминокислот, предсказанная при аннотировании генома [574]
Dm_CG3819	<i>D. melanogaster</i> (насекомые)	CG3819	Последовательность аминокислот, предсказанная при аннотировании генома [574]
LOC4812298	<i>D. pseudoobscura</i> (насекомые)	XP_0013529 14.1	Последовательность аминокислот, предсказанная при аннотировании генома [575]

Dm_CG6839	<i>D. melanogaster</i> (насекомые)	NP_649076	Последовательность аминокислот, предсказанная при аннотировании генома [574]
AGAP011047	комар <i>Anopheles gambiae</i> (насекомые)	XP_308089	Последовательность аминокислот, предсказанная при аннотировании генома [576]
AGAP011048	<i>A. gambiae</i> (насекомые)	XP_308088	Последовательность аминокислот, предсказанная при аннотировании генома [576]. В рамках этой работы выявлена его полная кодирующая последовательность
AGAP007845	<i>A. gambiae</i> (насекомые)	XP_317654	Последовательность аминокислот, предсказанная при аннотировании генома [576].
Dm_CG14118	<i>D. melanogaster</i> (насекомые)	NP_648612	Последовательность аминокислот, предсказанная при аннотировании генома [574]
Dm_IP10440	<i>D. melanogaster</i> (насекомые)	AAU55811	Последовательность аминокислот, предсказанная при аннотировании генома (Berkeley Drosophila Genome Project)
Dm_LP23408p	<i>D. melanogaster</i> (насекомые)	AAS77426	Последовательность аминокислот, предсказанная при аннотировании генома [577]
DmEndoG	<i>D. melanogaster</i> (насекомые)	NP_610737	Последовательность аминокислот, предсказанная при аннотировании генома [574]

CuquEndo	москит <i>Culex quinquefasciatus</i> (насекомые)	AAR18449	дц ДНК специфическая эндонуклеаза из слюнных железах [565]
Dm_CG9989	<i>D. melanogaster</i> (насекомые)	AAF56806	ДНКаза, индуцированная стрессом, последовательность аминокислот предсказана при аннотировании генома [574]
Lundep	<i>Lutzomyia longipalpis</i> (насекомые)	AAS16916	Последовательность аминокислот, предсказанная на основе последовательности мРНК из слюнных желез [578]
SmNucA	<i>Serratia marcescens</i> (прокариоты)	P13717	Сахар-неспецифическая ДНК/РНК эндонуклеаза [268]
AnNucA	<i>Anabaena variabilis</i> (прокариоты)	P38446	Сахар-неспецифическая ДНК/РНК эндонуклеаза [579]
Sr_Nuc1	<i>Syncephalastrum racemosum</i> (прокариоты)	P81204	Сахар-неспецифическая ДНК/РНК эндонуклеаза [564, 580]
Calb_Nuc	<i>Candida albicans</i> (грибы)	XP_716805	Митохондриальная нуклеаза, последовательность аминокислот предсказана при аннотировании генома [581]
Cun_NucC1A	<i>Cunninghamella echinulata</i> (грибы)	AAC78769	Сахар-неспецифическая ДНК/РНК эндонуклеаза [545]
Sac_Nuc	<i>Saccharomyces cerevisiae</i> (грибы)	NP_012327	Митохондриальная нуклеаза, последовательность аминокислот предсказана при аннотировании генома [582]

HumEndoG	<i>Homo sapiens</i> (млекопитающие)	NP_005098	Митохондриальная экзо/эндонуклеаза [583]
CaenEndoG	<i>Caenorhabditis elegans</i> (нематоды)	NP_491371	Митохондриальная нуклеаза, последовательность аминокислот предсказана при аннотировании генома (<i>C. elegans</i> Sequencing Consortium, 1999)

Для всех экстрагированных из GenBank последовательностей была исследована структурная организация полипептидной цепи. По данным ПО SMART все они содержали нуклеазный домен, характерный для SNF. При анализе с помощью ПО SignalIP-5.0 сигнальный пептид был обнаружен у большинства белков, кроме группы митохондриальных эндонуклеаз (CaenEndoG, Hum_EndoG, DmEndoG, Calb_Nuc), а также белков AGAP007845, Cun_NucC1A, Dm_CG3819, Sr_Nuc1 и AGAP011048. Последовательность AGAP011048 была предсказана в ходе аннотирования генома *A. gambitae* и содержала неполные N- и C-концы. Дополнительный поиск «потерянных» экзонов для AGAP011048 был осуществлён с помощью ПО FGGENESH [584] и позволил обнаружить некорректно выявленные при аннотировании места вырезания интронов и недостающие 3' и 5'-концевые экзоны (таблица 3-9). В дальнейшем сравнении мы использовали уточненный вариант последовательности. Его сходство с последовательностями из родственных нуклеаз позволяет утверждать, что этот вариант является более верным. Уточненный вариант последовательности аминокислот AGAP011048 по данным SignalIP-5.0 содержит сигнальный пептид. Результат множественного выравнивания последовательностей приведен на рисунке 3.17. Высококонсервативный участок нуклеазных доменов использовали для построения древа филогенетического сродства с помощью метода максимального правдоподобия.

Таблица 3-9. Уточнение последовательности предсказанного белка AGAP011048

Последовательность, депонированная в GenBank (XP_308088)	Уточненная последовательность
<p>FFLKSSQNKIAALPAGGCTVPFADLPYPE QPLILIPGTEKYWYPLDETREILVPTGAPIEL ACQQGFRLFPTQRSITIQCVTNSTFSFGGSP YPMKSLACTSYWLSSAKTTQARCHNESVI VKVGFELADARWVNVFDVCYDEQLYHTH FVRHYMNRRANGGYQSGNPRPGWYQGAY YTGVNINTLYTVNRQRETIATILNSQARAD VLVQNTTNGIYMARGHIAARADRVYGGTEQ NATFWFLNAAPQWQNFNGVNWERVESSV RDFVGKRDLELTVYSGTYGVQKLADGNG DYREIWLDFDPVEGRRRAPAPMLYYKILH DEASNAGIALVGVNNVHVPVELILREYVLC KDIGDEVEWIDWERKNLTIGYCYACEVNA FNDAIGRPHPQLNVAKLLTS</p>	<p>MLGRVLAVSCIFLVAYLRGAAARDVPTS EEIPRAGGCTVPFADLPYPEQPLILIPGTEK YWYPLDETREILVPTGAPIELACQQGFRLFP TQRSITIQCVTNSTFSFGGSPYPMKSLACTS YWLSSAKTTQARCHNESVIVKVGFELADA RWVNVFDVCYDEQLYHTHFVRHYMNRRAN GGYQSGNPRPGWYQGAYYTGVNINTLYT VNRQRETIATILNSQARADVLVQNTTNGIY MARGHIAARADRVYGGTEQNATFWFLNAAP QWQNFNGVNWERVESSVRDFVGKRDLEL TVYSGTYGVQKLADGNGDYREIWLDFDPV EGRRRAPAPMLYYKILHDEASNAGIALVG VNNVHVPVELILREYVLC KDIGDEVEWIDWERKNLTIGYCYACEVNA FNDAIGRPHPQLNVAKLLTSSGGRGAVFALSMLAGLGL HALLFVVRSSALRA</p>
<p>Цветом выделены добавленные и удаленные (зачеркнуты) фрагменты последовательности</p>	


```

*          140          *          160          *          180
SmNucA   : -----MRFNNKMLLAAALFFAQASADTLESI-DN : 29
AnNucA   : -----MGICGKLGVAALVALVICSPVQSQVPLTELSPSISVHLLLG : 42
Hum_EndoG : KSIASRLRGSRRFLSGFVAGAVVGAAGAGLAAALQFFRSQGAEGALTGKQPDGSAEKA-VL : 62
CaenEndoG : IAGISFLAGKYSNDDLPIFRNVQSATNVPMNQIQVSEPMTVKPAASNADAMGPRSA-EI : 69
AGAP007845 : LERKRLTASDYINTGKAAEFCSKPLPIFGTVSAATPIPAD-SPPSAMVNVSRVG-QI : 84
DmEndoG   : GAFYLGTHVERERQHNGSTSGLPRLPGLPTFGTVSAASLI PAQENNVSLTATPSRIG-QI : 76
Calb_Nuc  : MSKLLVNTLSLGTIGGTALYFWGKSSIPPTPVQSPNLPAPITNPSGGVFDPSAINPQ-GF : 59
Sac_Nuc   : MCSRILLSGLVGLGAGTGLTYLLLNKHSPTQI IETPYPTQKPNQSHSFNVDPG-GF : 59
Sr_Nuc1   : -----MPIIRTSRATASAEPSLHRLMNPIRRNTLSARASS-DI : 39
Cun_NucC1A : -----SPNSE-SI : 7
Gam_Nuc   : ----LYRQLGNTSLERDSEYLSLQKPVENFE BSSIQYNVLPPLGGGNEEDEEDSIG--GV : 138
CuquEndo  : -----LTCSPSS-----INAKRFDCAKPPDVTERDL-NRE CNGKSS-----KT : 113
Lundep    : -----VETITCNTN-----FDLASYS CNKSTSTDTIET-EEV CGGSG-----KV : 122
Gm_Tsall  : -KNGEESVTLMCHNNEFD--DGQGNGLDLFTGVKPTAELRKT-KERC SLGDL-----YV : 131
Gm_Tsall2 : KSNQGETTTLSCNGNQFE--SAAKEKLDMSCKSMKAVVEQT-TKRCMGNDF-----NL : 127
Dm_CG9989 : -----TTFLCRYDSVF-----SPALSSAACAPPDPVVVKVP-DTSCSIPS-----AT : 110
Dm_IP10440 : ----SVFKETVQDNGQF-----SVPLPRLCLSPMQPVTKHIRDGD CAG-----NL : 116
Dm_LP23408 : ----RDIVATTCQAGRVPAF-----QPPI P-MT RAAPAAITTPVQDRRC PA-----TM : 106
Dm_CG14118 : LDAKLR SIRVKVQDTTFEWMGAKIHFSDFVNSHSMPTVERL-DRS CGSDT PSPSTSYL : 170
Dm_CG14120 : P FVN RWQV TATCLQNKYFLVDDLIYPFANFSCTAWPIFTALRS-GKDCNGGT-----DL : 151
Dm_CG3819 : -----SVTAKVGGTTFKIDDKHDL SAIKCTSWPVFVGKKS-GSS CNGGT-----TL : 132
AGAP011048 : --PTQRSITITQCVTNSTFSFGSPYPMKSLACTSYWLSSAKTT-QARCHNES-----VI : 140
AGAP011047 : --PSETSIVVTCLVDDQFNYESKMYAFTEFGCTANWRSVARRT-ANRCYNDA-----TI : 151
LOC4812298 : --LSGKSVTAKVEGTTFRIGDKDHDVSDIKC SSWPAFVGKKS-GSS CNGGT-----TL : 151
Dm_CG6839 : ---SENLLTASCVSGTTFVSGSNFEFKDLYCKSWPGFKAVKS-GAT CNGG-----IV : 157
Par_DSN   : --LGMEAVEAKCAGGVMLAIEGTEWEIWSLGC SNHVKETIRRN-LGTCGEADQG-----DR : 134
Uca_NUC   : --LGEAEVQARCLSSGLLTIGDAEWDLASLGCSSDVKETIFRD-LGTCGAGGVG-----IL : 133
Mj_Nuc    : --LGTAVDVQCGGNLLVVDGTEWRMDELGC SKKDKESIHRN-LGSCGDGGVG-----VF : 130
Pal_Nuc   : SRSRGCWREPVLAACLITGDGTDNALNEIGVTPHLRACRRTWVPVRCR--SW-----YL : 132
Cmae_Nuc  : AVFGEMVVEATCLGGIMLDVNGKNKSLKEIGCSQKIRSTIKKN-LGSCGAENLG-----QL : 129
Hame_Nuc  : --LDVAVVEAVCAGGDL LAVAGTEWQLFELGCTEKLSESIHRD-MGSCGSQDMG-----VI : 131

```

```

*          200          *          220          *          240
SmNucA   : CAVGCGPTG--GSSNVSIVRHATITNNNSTTKFANWVAYHITKDTP---ASGKTR---NWK : 81
AnNucA   : NPSGATPTKLPDNYLMVKNOVALSYNNSKGTANWVAVQLNSSW----GNAERQ-DNFR : 97
Hum_EndoG : EQFGFPI T----GTEARCYTNHALSYDAQRVPRWVLEHISKSKL---MGDADRKHKCFK : 115
CaenEndoG : MKHGYPGF-----TNVRTYEDVLSYDYKTRTAHWVCEHLTPERL-KHAEGVDRKLCEFK : 123
AGAP007845 : MKYGFPGF----DNVRSFDDVLSYDRRTRVAHWVFEHLTPATV-KHNDAVDRAKCDFK : 138
DmEndoG   : MKYGFPGF----DHVRSFSDVLSYDRRNRVPHWVFEHLTAESV-AKNDAVDRSKCDFK : 130
Calb_Nuc  : FKFGFPGP----IHDLANRTEVSCYDRSTRNPYVIEHITKASLAKAESNGDRKRSVFK : 115
Sac_Nuc   : FKYGFPGP----IHDLQNRREEISCYNRQTONPYVLEHITPESL--AARNADRKNSFFK : 113
Sr_Nuc1   : LKLGNGPG----VSDLLERSGMI LSYNRRDRLAHWVGEHLTSASL-QAGQGVDRDKSNFQ : 94
Cun_NucC1A : LQFGDPGT----ARDFLERESVLSYNRRDRVASWTGEHLTADSL-KTGDGVD RDHSKFK : 62
Gam_Nuc   : SAFGYD LK-----ETFYTYLKFEEFKTKMMSSRVDHILHGREL--LKAEPHRDY-KFR : 188
CuquEndo  : FQMGFRAS-----GGLFVPEYELCYLMDTCSVSHVKHRLLGASF--GGVSGTREK-SFQ : 164
Lundep    : YKVGFP L P-----SGNFHSIXQTCFDKKNLTPLYSIHILNGQAV--GYHLKHTRG--SFR : 173
Gm_Tsall  : YDVVFRIS-----EEELVGPVYDLCYNEYSQKPAYSRNIIINGAAV--NYRVPESETDLTV : 184
Gm_Tsall2 : YEAGYKVN-----GKFYGSVYDLCYNGKGQSSGYTHNFIYGRTW--KYKLPKPYEHYS : 179
Dm_CG9989 : FAVGFSFN-----GRMELVNRNCFDGYSLA--FQ-HSIYKAYR--YVNTVPRPNPTWQ : 158
Dm_IP10440 : YAVGYTID-----GKDLELVRTCFDCGQGRLVYSQSDVYYKTF-----FPKRFVFEV : 164
Dm_LP23408 : YRVGYNVG-----NNQFLELYRACFDTRAVRAIFVEHRVYGKPF-----YITRCPVQFS : 155
Dm_CG14118 : YRVGYDTG-----DGRFVATMELCHDPNQLRTHYAHHQLT PANV-HFQKCLKRP--RFS : 221
Dm_CG14120 : VQVGFVEV-----DGGFLQSYELCHDAEAETRYVHHVLYPSSY-DYQHGVARP--NFL : 202
Dm_CG3819 : IKVGFELSGSR-----FATQVEVCFNEDEEVTRYVYHRLEPGNN-YYATGVDR I--TFG : 183
AGAP011048 : VKVGFELA-----DARVNVVDVYDEQLYHTHFVRHYMNRANG-GYQSGNPRP--GWY : 191
AGAP011047 : VELGFEMG-----AREPKIMDVCHDEVTYDNHYLVHEFTPANA-GFQTGVBRP--GWI : 201
LOC4812298 : IRVGFELD-----GGRFATEVCFNEDEEVTRYVYHTLAPGNN-YYATGVDR I--TFG : 202
Dm_CG6839 : IRVGF EIT-----SSRFAEQMOCFNEEEVTRYTRHKLEPGSN-YYETGVARI--TFQ : 208
Par_DSN   : HSI GF EY Y-----GGSIIYELISVCFNPGVSETTLRTEHVHLHGANA AAKDIETSRP--SFK : 187
Uca_NUC   : NNGFQIF-----SLNYDKVLIINVCFEAASETTLFTHHILHGADA AAKDVEASRP--SFK : 186
Mj_Nuc    : EGI GF E I F-----GSDSFYELIRVCFEPKAE T TLYSEHVHLHGANA AAKDIDSSRP--SFK : 183
Pal_Nuc   : HAVGFNIA-----TTGSFHEFESTCFDHAAE T TLYTKHTLHGANA IAKDVDP SRP--PFK : 185
Cmae_Nuc  : HIIGFQFA-----NKFEYQVLVCFDHDKQATLYTRHLHGANA AAKDIDSSRP--SFK : 180
Hame_Nuc  : HSI GF E I F-----GPEDFYELIRVCFDPKAE T TLYSENVHLHGANA AAKDKDNSRP--GFK : 184

```

Рис. 3.17 Продолжение


```

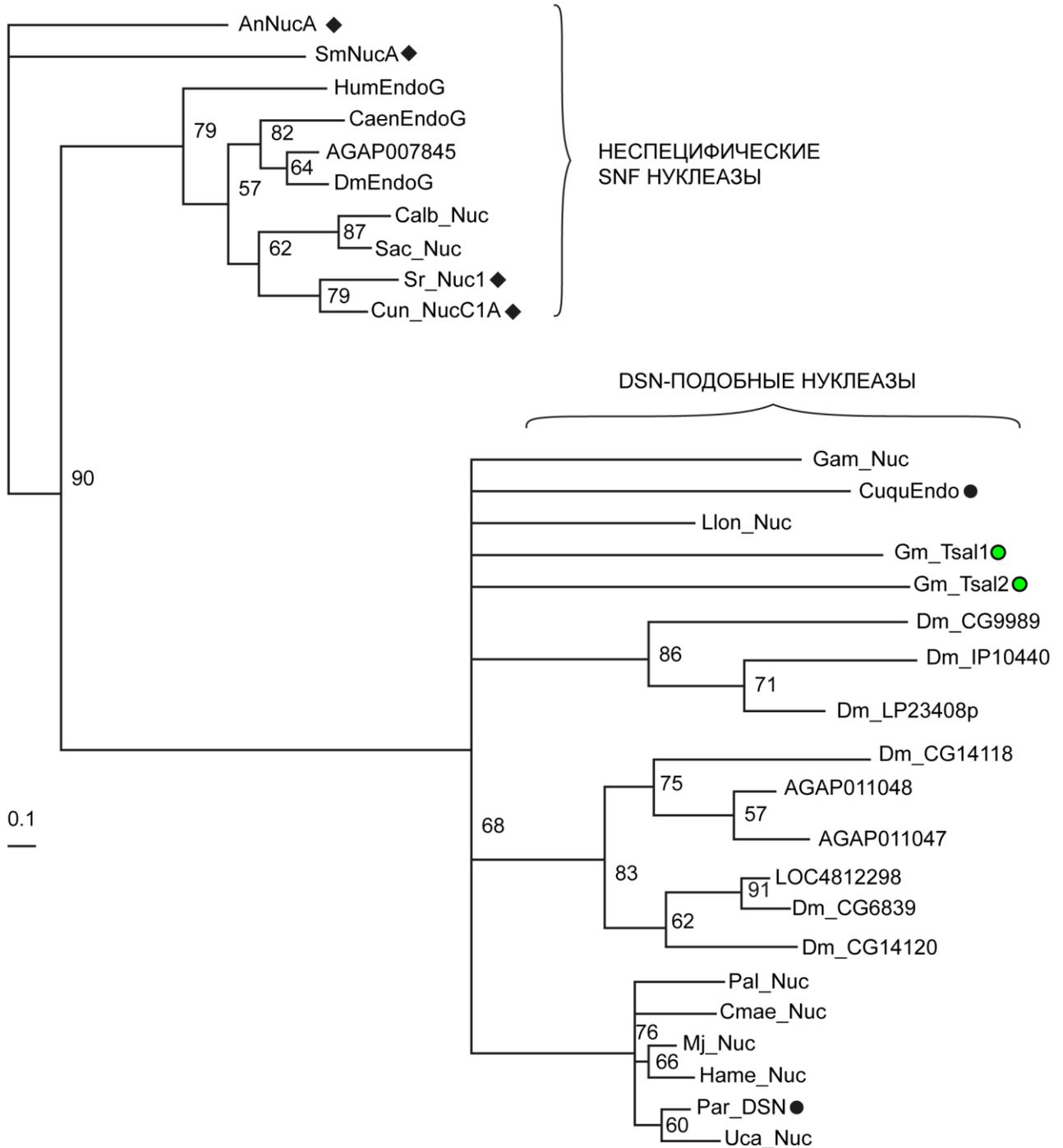
*           500           *
SmNucA      : EKRTGLIIWAGIPDDVQASLKSKPGVLPPELMGCKN-- : 266
AnNucA      : ESLTGYDPLSNVSPNIQTSIESKVDN----- : 273
Hum_EndoG   : EKLSVLVFPFHDRTSDIRNICSVDTCLLDFQEF-- : 307
CaenEndoG   : ERSAGLEIFARIDPKSIVKENGAKK-----GGLLW-- : 308
AGAP007845  : ERAAGLLFDKIA---RNKLAKINGKKV----- : 323
DmEndoG     : ERSAGLLFDQIN---RKQLTTINGKKV----- : 310
Calb_Nuc    : ERSSGLDLLQKVPYYNKQDLCKQVKCEIIVREFPKQV : 309
Sac_Nuc     : ERSTGLELLQKVPPSKKKALCKEVNQCIVVRDFSNA : 312
Sr_Nuc1     : EKAAGLVFDKIDRSKFADLCSKTTCQVR----- : 284
Cun_NucC1A  : EKASGLTFDKIDVSTLGDLC AATTC AVSSSGGD-- : 251
Gam_Nuc     : --TRGALVCLTPGDL-CGVEPKACGYKDIM----- : 393
CuquEndo    : EASRGITHCYWKDLARKNPLPIEVQREINNKKN--- : 364
Lundep      : NYSIGLTICCKIEEFVKRNKIILPKEVNKNYTKKLL : 387
Gm_Tsall    : DWTQGYTCCITWNNVKDYVRALPQDIE-INNLLKFN- : 399
Gm_Tsal2    : TQSAGSYGDFGNRIRTLPKDIYVK-----GLLKY-- : 388
Dm_CG9989   : TKDVGHTVCCDPYRFITINAPHLTGVC----- : 370
Dm_IP10440  : -PNDGYTCCDPKRFPY----- : 356
Dm_LP23408  : -AQAGFSFCNPNATFRP----- : 346
Dm_CG14118  : NLKNGYLYACSWANLARAVGHLPRPLLEVDELLT--- : 439
Dm_CG14120  : DLHEGYSYACSWEDFIEVVKDLPLEDLHTNGILGL-- : 426
Dm_CG3819   : DITAGYSYACEVPEFR--KKVTHLPEFSVSGLLV--- : 404
AGAP011048  : NLTIGYSYACEVNAFNDAIGRPHQPQNV-AKLLTSSG : 420
AGAP011047  : NLSRGYSYACEVNEFNRTGHLQPQNV--SLLI--- : 425
LOC4812298  : DISAGYSYACDVAEFRKKVDHLPQFAVS--GLLV--- : 423
Dm_CG6839   : DIKAGWSYACEVADFLKTVKHLPALTAGG-GLLV--- : 430
Par_DSN     : DLAHGTYCCAVDDL-RQAIPIYIPDLGNV-GLLTN-- : 406
Uca_NUC     : SLASGYTRCSVAEL-RASVPHVPDLGNV-CLLTD-- : 405
Mj_Nuc      : DLGHGTYCCAVDDL-RAAI PHVPDLG SV-GLLDK-- : 402
Pal_Nuc     : DLFRGFTYCCAVDDL-RAAI PHVPDLG NV-GLLDS-- : 404
Cmae_Nuc    : ----- : -
Hame_Nuc    : ----- : -

```

Множественные выравнивания последовательностей были подготовлены с помощью программного обеспечения Clustal X [548] с визуализацией в программе GeneDoc [549]. Консервативные остатки выделены черным цветом — процент консервативности выше 80%, темно серым — 50%, светло серым — 35%. Подчеркнут участок, который использовали для построения филогенетического дерева. Сокращения и номера GenBank для последовательностей показаны в таблице 3-8.

Рисунок 3.17. Множественное выравнивание аминокислотных последовательностей Par_DSN и ее гомологов из других членистоногих с нуклеазами SNF

На филогенетическом древе (рисунок 3.18) изучаемые белки разделились на две большие клады. Первая включала в себя «классические» SNF нуклеазы. Вторая клада, названная нами подгруппой DSN-подобных нуклеаз, состояла из нуклеаз насекомых и ракообразных. В свою очередь, эта подгруппа делилась на несколько подклад, при этом нуклеазы ракообразных, за исключением нуклеазы рачка бокоплава Gam_Nuc, группировались на дендрограмме вместе.

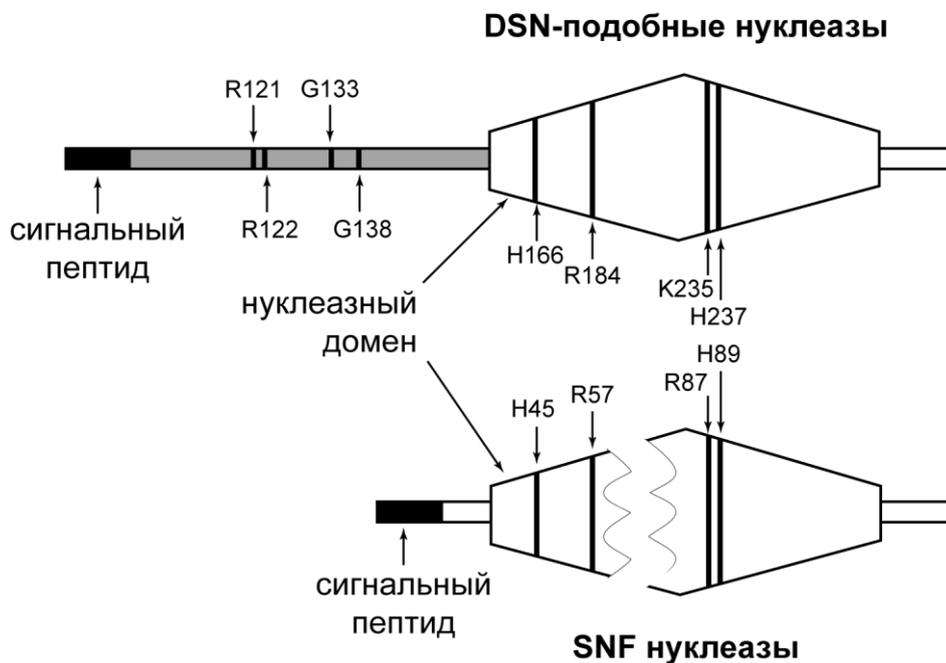


Черными кружками отмечены нуклеазы, имеющие подтвержденную субстратную избирательность к дц ДНК, зелеными кружками — нуклеазы, для которых эта специфичность была подтверждена позднее. Ромбами отмечены SNF нуклеазы, расцепляющие дц и оц ДНК и РНК. Цифрами показана вероятностная поддержка каждой ветви.

Рисунок 3.18. Филогенетическое древо максимального правдоподобия, построенное на основе множественных выравниваний нуклеаз (рисунок 3.17) с использованием алгоритма «quartet puzzling» с 10000 шагов.

Хотя большая часть нуклеаз насекомых оказались в одной кладе с Par_DSN, предсказанные белки DmEndoG и AGAP007845 вошли в кладу «классических» SNF нуклеаз, что характерно для митохондриальных эндонуклеаз G. Интересно, что нуклеаза CуquEndo, аминокислотная последовательность которой имеет чуть большее сходство с SmNucA (26% идентичности и 42% сходства), чем с Par_DSN (25% идентичности и 40% сходства), оказалась в одной подгруппе с DSN-подобными нуклеазами. На основании анализа филогенетического древа и данных об известной субстратной специфичности некоторых нуклеаз нами было выдвинуто предположение, что DSN-подобные нуклеазы и SNF нуклеазы эволюционировали по дивергентному пути, о чем свидетельствуют их различные биологические функции и субстратные специфичности.

Сравнение доменной организации DSN-подобных и SNF нуклеаз по данным множественного выравнивания показано на рис. 3.19.



Серым отмечен КР регион. Указаны консервативные аминокислоты каталитического центра SmNucA и соответствующие ей консервативные аминокислоты Par_DSN, а также консервативные аминокислоты КР.

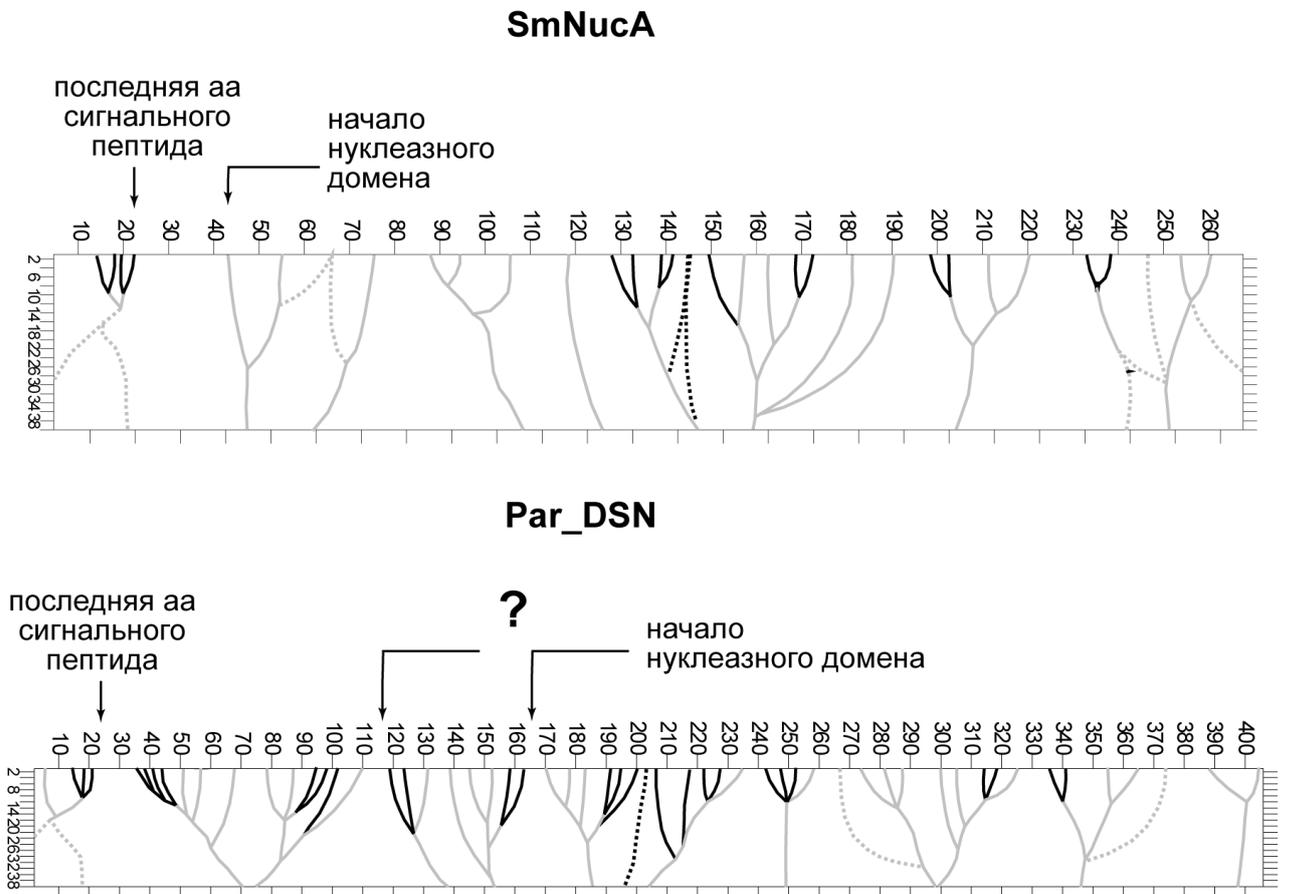
Рисунок 3.19. Сравнение доменной организации SNF и DSN-подобных нуклеаз

Было обнаружено, что у всех DSN-подобных нуклеаз нуклеазный домен более чем на 50 аминокислотных остатков длиннее нуклеазного домена, характерного для SNF и содержит «вставку» из, по крайней мере, шестнадцати аминокислотных остатков перед консервативным RGH-мотивом. Также у всех DSN-подобных нуклеаз между сигнальным пептидом и нуклеазным доменом находится последовательность длиной 120-180 аминокислотных остатков (КР) с неизвестной функцией. Родственных ей последовательностей в доступных базах данных обнаружить не удалось. Высококонсервативные участки у DSN-подобных нуклеаз оказались расположены не только внутри нуклеазного домена, определенного множественным выравниванием, но и по всей длине КР. Интересно, что характерной особенностью этих белков оказались консервативные остатки цистеина, распределенные по всей длине выравнивания, которые в том числе были найдены и внутри КР. Мы предположили, что нуклеазный домен DSN-подобных нуклеаз может иметь иную точку начала, нежели нуклеазный домен у SNF.

Для определения возможных границ нуклеазного домена у DSN-подобных нуклеаз был проведен анализ информационной структуры белка по методу А.Н. Некрасова [547, 585-586]. А.Н. Некрасов показал, что в белковых последовательностях существуют иерархически организованные элементы, образованные перекрывающимися «информационными единицами» – фрагментами аминокислотной последовательности длиной не более шести аминокислотных остатков. Совокупность иерархических элементов (информационная ассоциация) описывает организацию структурной информации в исследуемом белке, а ее анализ позволяет выявить области богатые иерархически организованными элементами, соответствующие функциональным доменам белка, и разделяющие их области с низким содержанием таких элементов.

Результат анализа информационной структуры Par_DSN и SmNucA показан на рисунке 3.20. Анализ информационной структуры Par_DSN выявил наличие шести информационных ассоциаций, первая из которых соответствовала области сигнального пептида. Вторая ассоциация была расположена между 35 и 114 аминокислотными остатками, третья – между 114 и 238, четвертая – между 238 и

260, пятая – между 265 и 325, а шестая начиналась с 336 остатка и заканчивалась в районе 376 аминокислотного остатка, в месте окончания нуклеазного домена.



Показаны пронумерованные аминокислотные остатки полипептидных цепей и образуемые ими информационные ассоциации. Аа — аминокислота.

Рисунок 3.20. Анализ информационной структуры Par_DSN и SmNucA

На основании информационной структуры Par_DNS было выдвинуто предположение, что началу нуклеазного домена соответствует 114 аминокислотная позиция. Менее вероятной представлялась позиция 162, но ее нельзя было теоретически исключить. Для проверки этого предположения были клонированы и экспрессированы несколько укороченных с 5'-конца вариантов фермента. Первый вариант (Par_DSN-T1) представлял собой полипептид, включающий в себя аминокислотные остатки 192–379 Par_DSN, связанные с бактериальным сигнальным пептидом и 6xHis последовательностью. Эта последовательность

соответствовала нуклеазному домену, определенному с помощью программы SMART. Полипептид был очищен и ренатурирован. К сожалению, он оказался нерастворимым и неактивным.

Второй вариант, Par_DSN-T2, содержащий аминокислотные остатки с 149 по 406, соответствовал NUC домену, определенному с помощью множественного выравнивания. После ренатурации и очистки был получен растворимый белок, не проявляющий ДНКазную активность. Инкубация Par_DSN-T2 с протеиназой К приводила к его полному расщеплению.

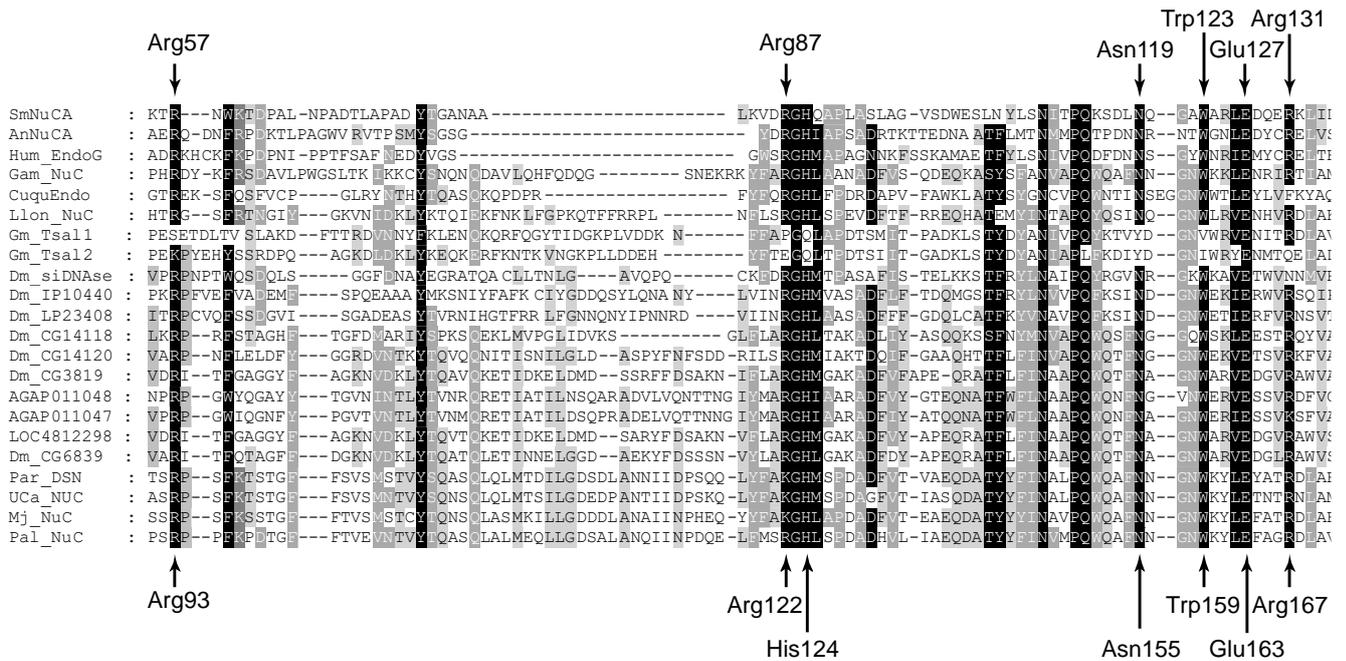
Третий вариант, Par_DSN-T3, содержал остатки 162–406. Из его структуры был исключен неспаренный остаток цистеина. Как и предыдущий вариант Par_DSN-T3 был растворим, чувствителен к протеиназе К и не обладал ДНКазной активностью.

Наконец, вариант Par_DSN-T4, содержащий аминокислотные остатки с 114 по 406, оказался растворимым белком, способным гидролизовать ДНК субстрат. Однако Par_DSN-T4 обладал свойствами, сильно отличающимися его от Par_DSN дикого типа. Например, он проявлял значительно меньшую каталитическую активность (6 единиц Куница против 6000 единиц активности интактной Par_DSN), имел температурным оптимумом при 37°C (против 60°C полноразмерной Par_DSN-T4) и пониженную температурную стабильность. Инкубация в течение 20 мин при 70°C приводила к полной инактивации этого белка, тогда как полноразмерная Par_DSN сохраняла при этих условиях около 60% активности. Кроме того, Par_DSN-T4 оказался чувствителен к обработке протеиназой К. Наконец, соотношение гидролитической активности дц к оц ДНК субстратам для этого варианта не превышало 10.

На основании полученных данных нами был сделан вывод, что нуклеазный домен Par_DSN, необходимый для проявления энзиматической активности, включает аминокислотные остатки с 114 по 406. Фрагмент с 28-го по 113-й аминокислотный остаток, по-видимому, вовлечен в модуляцию субстратной специфичности Par_DSN и важен для селективного связывания дц ДНК.

Структура каталитического центра была описана ранее для ряда SNF нуклеаз, включая SmNucA, AnNucA и эндонуклеазу G быка [84, 164, 264, 268, 587]. Было показано, что у Sm-нуклеазы в формировании каталитического центра принимают участие аминокислотные остатки His89, Arg57, Arg87, Asn119, Glu127, Trp123 и Arg131, где His89 является ключевой аминокислотой, участвующей в катализе, а Asn119 обеспечивает связывание двухвалентного катиона металла [264]. Кроме того, было продемонстрировано, что Asp86 опосредованно взаимодействует с металлосвязывающим сайтом, образуя водородные связи с амидными азотами Asn119 и Gln120 [588]. Аналогичный каталитический центр, включающий аминокислотные остатки His124, Arg93, Arg122, Asn155, Glu 163, Trp159 и Arg167, был описан также для AnNucA нуклеазы [164, 264]. Основным отличием в функционировании каталитических центров SNF нуклеаз является стабилизация переходного состояния: у SmNucA в нейтрализации негативного заряда в переходном состоянии участвует Arg57, у AnNucA — неконсервативный Asp95 [264], а у эндонуклеазы G — Cys115, расположенный в позиции, соответствующей Arg57 SmNucA [587].

На основании множественного выравнивания DSN-подобных нуклеаз и представителей SNF было показано, что большинство DSN-подобных нуклеаз содержат консервативные Glu, Trp, Asn, His и Arg остатки, соответствующие Glu127, Trp123, Asn119, His89 и Arg57 SmNucA (рисунок 3.21). «Замены» в этих позициях были обнаружены только у Tsal белков слюнных желез *G. morsitans* [589]. Остатки аргинина, соответствующие Arg167 AnNucA, были обнаружены во всех DSN-подобных нуклеазах за исключением белков Dm_CG9989, CuquEndo и Tsal2. У Tsal белков не выявлялось аминокислотного остатка, соответствующего Arg122 нуклеазы AnNucA, а у четырех из шести нуклеаз ракообразных в этой позиции находился остаток лизина. Аминокислотный остаток в позиции, соответствующий Asp86 SmNucA, у DSN-подобных нуклеаз оказался неконсервативным. У всех нуклеаз ракообразных в этом месте находился остаток аланина.



Сверху стрелками указаны аминокислотные остатки каталитического центра SmNucA, снизу — AnNucA.

Рисунок 3.21. Расположение аминокислотных остатков активных центров SmNucA и AnNucA на фрагменте множественного выравнивания аминокислотных последовательностей нуклеаз (полное выравнивание см. на рис. 3.17)

Наиболее заметной особенностью, характерной для ракообразных, входящих в группу DSN-подобных нуклеаз, является замена в нуклеазном домене высококонсервативного мотива DRGH на AKGH. Предполагается, что в SmNucA остаток D86 косвенно участвует в связывании Mg^{2+} , координируя N119 [588]. Было показано, что мутация D86A в SmNucA приводит к увеличению потребности нуклеазы в магнии, при этом без подавления ее ферментной активности [590]. Возможно, замена аспарагиновой кислоты на аланин у нуклеаз ракообразных определяет особенности металлозависимости, сходной с мутантным вариантом SmNucA.

Для идентификации аминокислотных остатков, важных для проявления энзиматической активности Par_DSN, а также участвующих в формировании ее трехмерной структуры, нами был осуществлен сайт-направленный мутагенез с последующим анализом активности рекомбинантных мутантных белков. Список полученных мутантов представлен в таблице 3-10. Так как устойчивость к

протеолизу свидетельствует о правильном формировании дисульфидных связей в белке, нами была изучена устойчивость мутантных белков к протеиназе К. Энзиматическая активность полученных рекомбинантных белков оценивалась по способности гидролизовать геномную дц ДНК и оц ДНК-субстраты (синтетические олигонуклеотиды).

Таблица 3-10 – Анализ активности и чувствительности к протеиназе К мутантов Par_DSN

Внесенная замена	Соответствующая позиция аминокислотного остатка у SmNucA	Активность по отношению к дц ДНК, ед. Куница /мг	Выход белка после обработки протеиназой К (%)
нет	–	6070	20%
His 237 Ala	His 89	Не детектируется	19%
Lys 235 Ala	Arg 87	110	21%
Lys 235 Arg	Arg 87	5980	20%
Arg 184 Lys	Arg 57	6100	18%
Arg 184 Ala	Arg 57	1810	22%
Arg 122 Lys	Arg25	6050	20%
Arg 122 Ala	Arg25	6030	20%
Arg 121 Lys	Arg24	5990	19%
Arg 121 Ala	Arg24	5960	19%
Arg 121 Ala + Arg 122 Ala	Arg24	Не детектируется	4%
His 168 Ala	His45	Не детектируется	0
His 168 Lys	His45	Не детектируется	0
His 171 Asp	-	Не детектируется	0
Gly 133 Val	-	Не детектируется	0
Gly 139 Val	-	Не детектируется	0

Как и ожидалось, замена His237 (соответствующего His89 SmNucA) на Ala привела к потере энзиматической активности Par_DSN, что говорит о ключевой роли этого аминокислотного остатка в катализе. Замена Lys235, соответствующего вовлеченному в связывание субстрата Arg87 SmNucA, на Ala привела к резкому снижению нуклеазной активности. Такой же эффект был продемонстрирован ранее при замене Arg87Ala для SmNucA [590]. Следует отметить, что мутант Lys235Arg обладал нуклеазной активностью соизмеримой с активностью Par_DSN дикого типа. Все эти замены не влияли на устойчивость фермента к протеолизу протеиназой К. На основе полученных результатов, нами был сделан вывод, что His237 и Lys235 Par_DSN и родственных ей нуклеаз играют ту же роль в работе каталитического центра, что и His89 и Arg87 у SmNucA.

По результатам множественного выравнивания позиция Arg184 Par_DSN соответствует Arg57 SmNucA. Ранее было показано, что у SmNucA замена этого аминокислотного остатка на остатки лизина или аланина приводит к падению активности фермента на 94 и 99 %, соответственно [590]. В отличие от мутантов SmNucA, мутант Par_DSN, содержащий замену Arg184Lys, обладал активностью, сравнимой с активностью Par_DSN дикого типа, а замена Arg184Ala приводила только к 3-кратному падению активности. Обе замены не влияли на устойчивость фермента к протеолизу протеиназой К. Иными словами, в отличие от Arg57 SmNucA, Arg184 Par_DSN, по-видимому, не является существенным для функционирования каталитического центра. Мутанты Par_DSN, содержащие замены His168, соответствующего высококонсервативному у SNF нуклеаз остатку His45 (нумерация по SmNucA), на остатки аланина и лизина, оказались чувствительными к протеолизу протеиназой К и не обладали гидролитической нуклеазной активностью. К такому же эффекту приводили замены His171Asp, Gly139Val и Gly133Val. Был сделан вывод, что эти остатки могут играть важную роль в формировании правильной трехмерной структуры белка. Интересно, что His171 консервативен только у Par_DSN-подобных нуклеаз ракообразных, Gly139 – у всех изученных нуклеаз членистоногих, за исключением Tsall1, а Gly133 не консервативен. Дополнительно, была продемонстрирована важная для

функционирования фермента роль консервативных аминокислотных остатков в КР области. Мутантный белок 121Ala-122Ala оказался полностью неактивен, тогда как мутанты Arg121Ala и Arg122Ala обладали высокой гидролитической активностью, соизмеримой с Par_DSN дикого типа. Мы предположили, что, по крайней мере, один из остатков аргинина каким-то образом участвует либо в формировании, либо регуляции каталитической активности активного центра Par_DSN. При замене одного из остатков аргинина второй берет на себя эту функцию. Следует заметить, что большинство Par_DSN-подобных нуклеаз содержат остаток аргинина или лизина по крайней мере в одной из этих двух позиций.

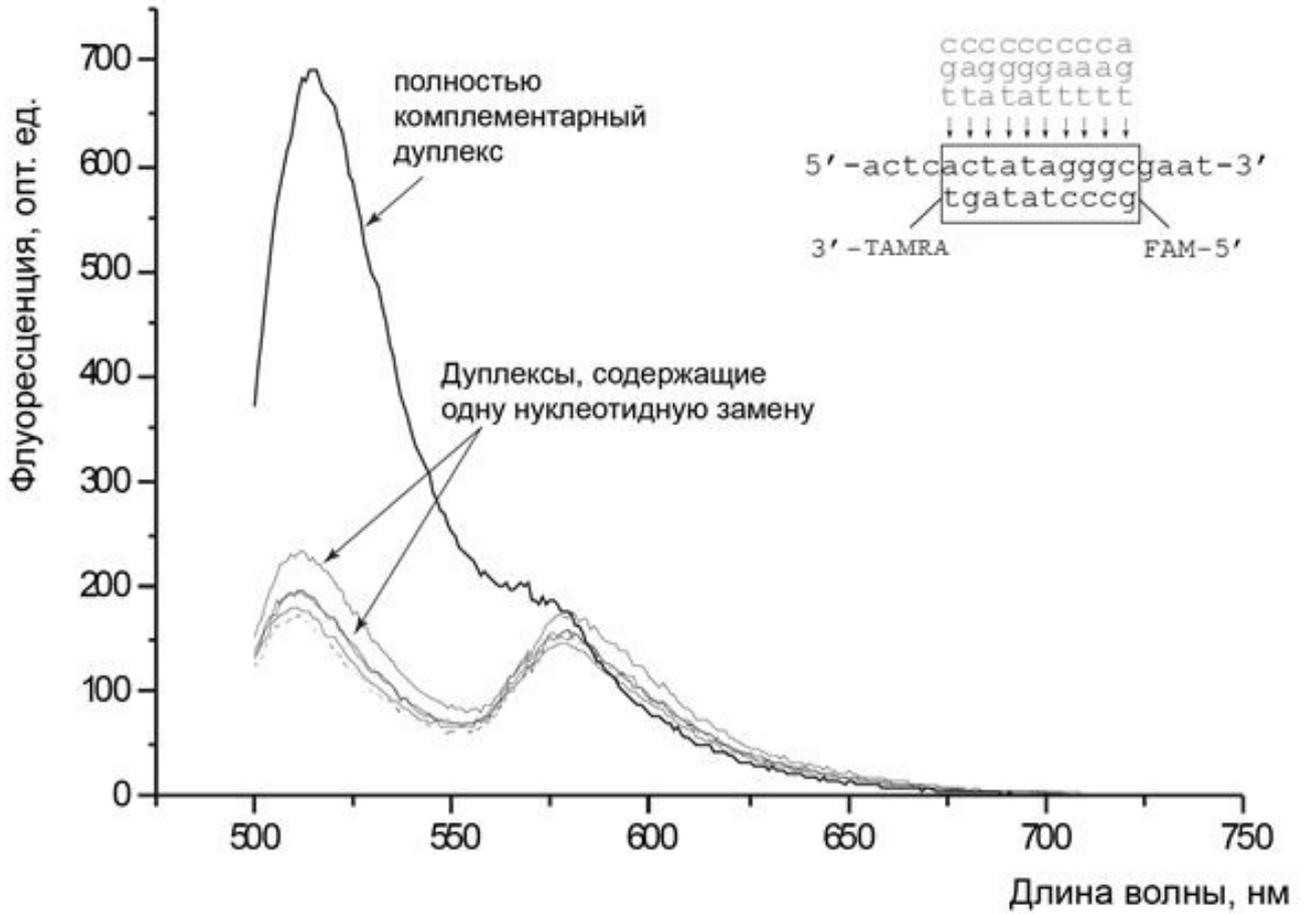
Таким образом, Par_DSN и ее гомологи содержат в своем составе нуклеазные домены, которые длиннее соответствующего домена нуклеаз SNF, а их каталитический центр, хотя и сходен с каталитическим центром SNF нуклеаз, имеет ряд принципиальных отличий. Для нуклеаз ракообразных, в частности, характерна замена консервативного мотива DRGH на AKGH. Между нуклеазным доменом и сигнальным пептидом Par_DSN-подобные нуклеазы содержат регион длиной около 90 аминокислотных остатков, который, по-видимому, важен для определения субстратной специфичности и стабильности фермента. На основании этих отличий, нами было предложено выделить Par_DSN-подобные нуклеазы в новое семейство внутри суперсемейства His-Me нуклеаз. Исходя из определенного сходства членов этого семейства с нуклеазами SNF, мы предположили их дивергентное расхождение в эволюции и постулировали, что нуклеазы с дц-ДНКазной активностью могут быть найдены у других членистоногих. Нами было показано, что нуклеазный домен Par_DSN отделенный от КР области сохраняет каталитические свойства, но в значительной степени утрачивает избирательность по отношению к дц ДНК. Таким образом, среди Par_DSN-подобных нуклеаз можно ожидать как неспецифической активности по отношению к дц и оц ДНК и РНК, так и дуплекс-специфической активности.

Наше предположение было в дальнейшем подтверждено рядом работ независимых лабораторий, которым удалось продемонстрировать активность некоторых ферментов, отнесенных нами к семейству Par_DSN-подобных нуклеаз.

Так, были исследованы белки мухи цеце Tsal1 и Tsal2 [591]. Они оказались ДНК-связывающими белками с остаточной дц ДНКазной активностью. Активности по отношению к РНК и оц ДНК обнаружено не было (там же). Кроме того, была клонирована кодирующая последовательность описанной ранее дц ДНК-специфичной магний-зависимой термолабильной нуклеазы из креветки *Pandalus borealis* (далее Pbor_DSN) [592]. Анализ ее аминокислотной последовательности показал наличие характерных для дуплекс-специфических нуклеаз особенностей нуклеазного домена. Авторы провели сравнение субстратной специфичности этой нуклеазы с Par_DSN и ДНКазой I быка. Сравняя относительную нуклеазную активность к денатурированной и неденатурированной ДНК, авторы показали, что разница в эффективности расщепления этих субстратов для ДНКазы I в пять раз ниже, чем для Pbor_DSN. Также было показано, что нуклеазная активность Pbor_DSN в отношении денатурированной ДНК составляет около 1–5% активности от неденатурированной ДНК. На синтетических субстратах относительный коэффициент расщепления оц ДНК по сравнению с дц ДНК для нативной Pbor_DSN составил 0,4%, для рекомбинантной — 0,1%. Для Par_DSN этот параметр оказался ниже чувствительности использованного метода (там же). Дуплекс-специфическая нуклеаза, активная при температуре до 63°C и устойчивая к обработке протеиназой K, была выделена из снежного краба *Chionoecetes opilio*. Нуклеаза демонстрировала энзиматическую активность в присутствии ионов Mn^{2+} и Mg^{2+} , но не в присутствии ионов Ca^{2+} [593]. Появились данные о нескольких неспецифических эндонуклеазах, попавших в семейство Par_DSN-подобных нуклеаз. Было показано, что нуклеаза Lunder из *L. longipalpis* с одинаковой эффективностью расщепляет дц и оц ДНК и РНК [594]. Для стресс-индуцируемой нуклеазы *D. melanogaster* CG9989 была продемонстрирована примерно одинаковая эффективность расщепления дц и оц ДНК и РНК в присутствии ионов Cu^{2+} . В присутствии Zn^{2+} нуклеаза расщепляла оц ДНК несколько хуже, чем дц ДНК, а в присутствии Mn^{2+} расщепляла дц ДНК и связывала оц ДНК, сохраняя ее в лунке во время электрофоретического разделения НК [595].

3.3. Методы анализа мутаций и выявления ДНК-мишеней в сложных смесях нуклеиновых кислот

Исследование активности Par_DSN против полностью комплементарных ДНК-ДНК дуплексов и дуплексов, содержащих одну нуклеотидную замену, показало, что Par_DSN расщепляет полностью комплементарные молекулы ДНК с заметно более высокой эффективностью. Различия в активности в отношении таких субстратов наблюдались при размерах дуплекса от 8 до 12 п.о., с наибольшей разницей при длине дуплекса 10 п.о. (данные не показаны). С помощью коллекции 18-звеньевых олигонуклеотидов и 10-звеньевых флуоресцентно-меченого зонда было показано, что тип и положение точечной замены не влияет на эффективность гидролиза субстрата нуклеазой (рисунок 3.22).

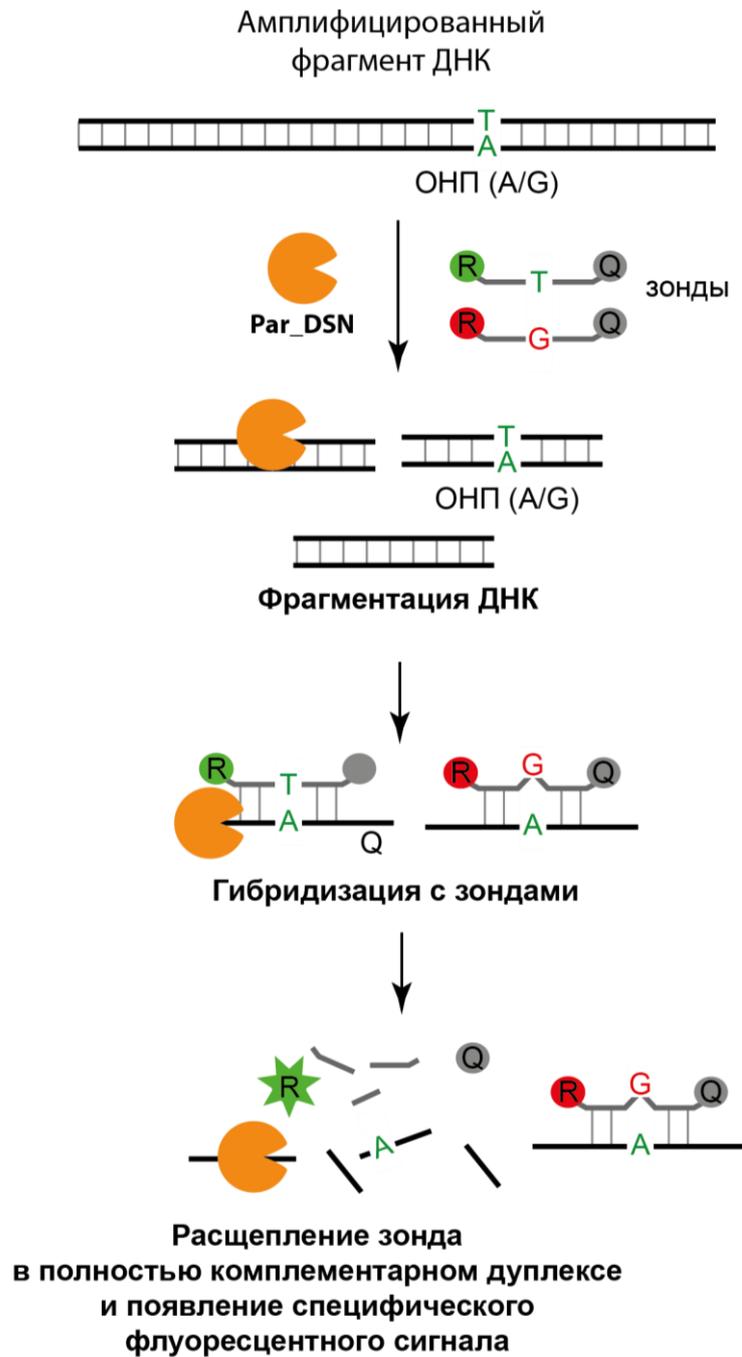


Для создания ДНК-ДНК дуплексов объединяли зонд длиной 10 н.о., меченный парой «донор и гаситель флуоресценции» и 18-звеньевой немеченый олигонуклеотид, содержащий полностью или частично комплементарную зонду последовательность.

Для создания дуплексов, содержащих некомплементарную пару, в последовательность 18-звеньевого олигонуклеотида в каждой позиции комплементарного зонду участка вносились все возможные замены нуклеотида (показано серым цветом над стрелками).

Рисунок 3.22. Анализ нуклеазной активности Pac_DSN по отношению к полностью комплементарным ДНК-ДНК субстратам и субстратам, содержащим неспаренную нуклеотидную пару

В то же время влияние на эффективность расщепления дуплексов с одной неспаренной парой оказывал тип добавленной на 5' или 3'-конец зонда флуоресцентной краски. Так при замене гасителя флуоресценции TAMRA на DABCYL эффективность гидролиза содержащих некомплементарную пару дуплексов возрастала у субстратов, где эта пара находилась на 5'-конце 18-звеньевого нуклеотида (рисунок 3.23).



R — донор флуоресценции, Q — гаситель флуоресценции.

Рисунок 3.24. Схема выявления мутантной последовательности в биологических образцах с помощью Par_DSN

Для выявления мутантной последовательности используется два 10-звеньевых зонда (далее ДСНП-зонды), каждый из которых содержит флуорофор и гаситель флуоресценции. После расщепления зонда происходит высвобождение

флуорофора и появление специфического флуоресцентного сигнала. Один зонд содержит последовательность нуклеотидов, которая полностью комплементарна ДНК последовательности дикого типа, второй- последовательность нуклеотидов, комплементарную мутантной последовательности.

При добавлении ДСНП-зондов и Par_DSN в продукт ПЦР, содержащий амплифицированный целевой фрагмент с полиморфным сайтом, на первом этапе происходит расщепления ПЦР-продукта с появлением коротких фрагментов ДНК, которые способны образовывать дуплексы с зондами. Полностью комплементарные дуплексы также подвергаются атаке Par_DSN, что приводит к появлению флуоресценции определенной длины волны от соответствующего флуоресцентного донора, освобожденного от гасителя флуоресценции. Несовершенные дуплексы остаются интактными.

В первых вариантах предложенного метода использовали очищенный продукт ПЦР, который гибридизовали с ДСНП-зондами при 30–35°C. Для амплификации флуоресцентного сигнала в реакционную смесь добавляли фрагмент Кленова. За счет его полимеразной активности с затравок, образованных при расщеплении ДНК-субстрата нуклеазой, в образце продолжали нарабатываться новые молекулы ДНК, в том числе содержащие комплементарные зондам участки.

В другой модификации метода мы использовали неочищенный продукт ПЦР, содержащий ДНК-полимеразу. В этой модификации метода продукт ПЦР инкубировали с Par_DSN и ДСНП-зондами при температуре 60°C. В этих условиях Par_DSN расщепляет дц ДНК, создавая короткие ДНК-фрагменты. Эти фрагменты денатурируют и выступают как затравки для дополнительного синтеза ДНК за счет активности ДНК-полимеразы. Далее температуру инкубации снижали до 30–35°C, обеспечивая гибридизацию DSNP-зондов с короткими ДНК-мишенями. Par_DSN расщепляла образовавшиеся ДНК-ДНК-дуплексы, высвобождая флуорофор.

Было проведено исследование длины фрагментов ДНК, которые могут быть использованы для анализа ОНП. На примере ДНК, содержащей ОНП С7028Т гена COX1, мы показали, что метод применим при использовании фрагментов ДНК в диапазоне до 4000 п.о. (таблица 3-11, рисунок 3.25).

Таблица 3-11. Результаты измерений флуоресценции образцов, содержащих фрагменты ДНК различной длины с «Т» или «С» нуклеотидом в тестируемой позиции, в присутствии Par_DSN и ДСНП-зонда, полностью комплементарного «Т»-варианту.

Фрагмент ДНК	F₀	F_R	F_R – F₀	(F_R – F₀)/ F₀	cF, %
69 п.о. «Т»	233,05	824,03	590,98	2,54	100
69 п.о. «С»	215,18	239,04	23,86	0,11	4,4
534 п.о. «Т»	222,65	656,32	433,67	1,95	76,7
534 п.о. «С»	221,65	233,85	12,19	0,06	2,2
952 п.о. «Т»	222,45	518,26	295,81	1,33	52,4
952 п.о. «С»	212,13	244,19	32,06	0,15	5,9
4000 п.о. «Т»	214,12	461,12	247,00	1,15	45,4
4000 п.о. «С»	222,13	244,19	22,06	0,10	3,9

cF — изменение флуоресценции в тестируемых образцах, рассчитанное как % от максимального значения $cF = (F_R - F_0) / F_0$, где F_R — уровень флуоресценции реакционной смеси после инкубации в присутствии нуклеазы, F_0 — фоновая флуоресценция реакционной смеси в отсутствие фермента. В связи с тем, что в эксперименте использовали равные количества ДНК-субстратов, максимальная концентрация высвобожденного донора флуоресценции, принятая за 100%, достигалась при расщеплении самого короткого субстрата. Продукты ПЦР и плазмиду pT-Adv, содержащие в тестируемой позиции нуклеотид «Т» или «С», использовали для анализа с ДСНП-зондом, полностью комплементарным «Т»-варианту.

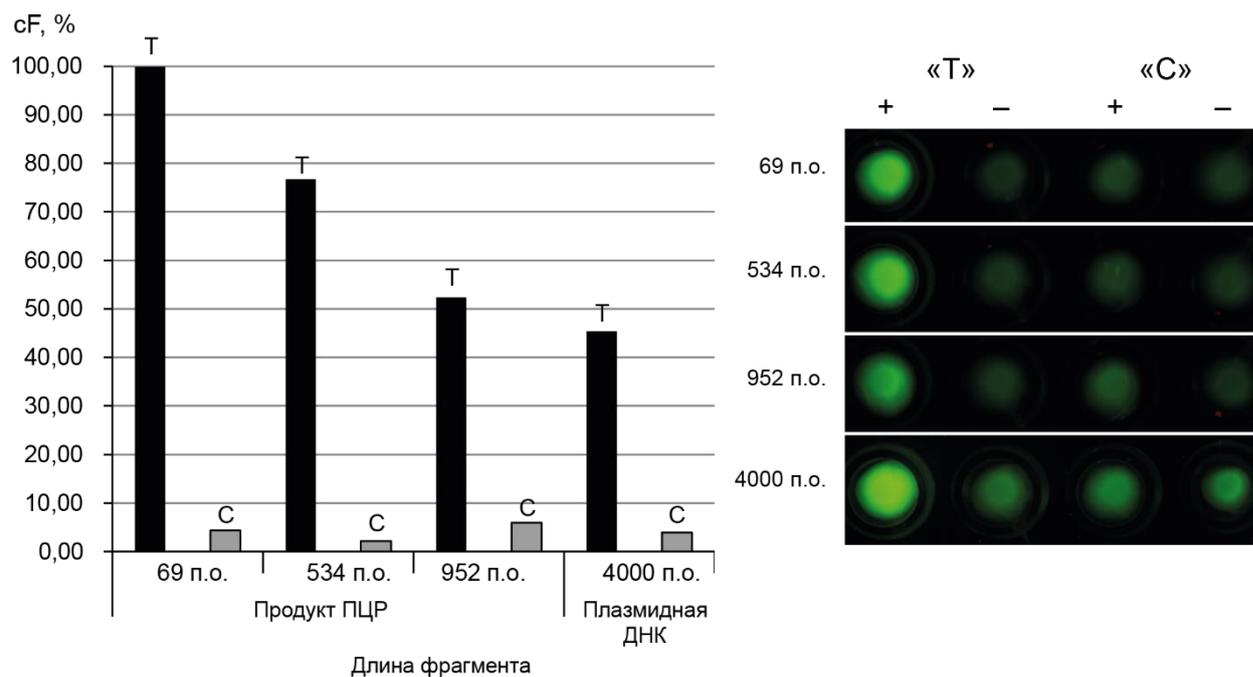


График построен на основе данных, показанных в таблице 3-11. Черные столбцы — образцы ДНК, содержащие в тестируемой позиции нуклеотид «Т», серые столбцы — образцы ДНК, содержащие в тестируемой позиции нуклеотид «С». Рядом показана фотография плашки с тестируемыми образцами (флуоресцентная спектроскопия, эмиссия при 520 нм).

«+» — инкубация в присутствии Par_DSN, «-» — инкубация без Par_DSN.

Рисунок 3.25. Выявление мутантных последовательностей при использовании ДНК-субстратов различной длины

Метод ДСНП-анализа был опробован в модельных экспериментах на образцах геномной ДНК человека с различным аллельным статусом по исследуемому ОНП, ассоциированным с рядом заболеваний человека (таблица 3-12). В некоторых случаях для ДСНП-анализа использовались синтетические 60-звеньевые олигонуклеотидные субстраты. Аллельный статус исследуемых образцов подтверждали с помощью секвенирования.

Таблица 3-12. Выбранные для ДСНП-анализа ОНП

Ген, ОНП	Заболевание, с которым ассоциирован мутантный аллель	Ссылки
<i>P53</i> , C309T	Развитие онкологических заболеваний	Mesaeli and Phillipson, 2004 [596]
<i>F2</i> , G20210A	Нарушение свертываемости крови	de Moerloose et al., 1998 [597]
<i>F5</i> , G1698A	Нарушения свертываемости крови	
<i>MTHFR</i> , C677T	Нарушения в развитии нервной системы	Kirke et al., 2004 [598]
<i>KRAS</i> , G34A, G35T, G35A, G38A	Развитие онкологических заболеваний	Chan et al., 2003; Mu et al., 2004 [599, 600]
<i>NRAS</i> , G34A, G35C, G35A	Развитие онкологических заболеваний	Shih et al., 2004 [601]
<i>HRAS</i> , G35T	Развитие онкологических заболеваний	Galiana et al., 1993 [602]
<i>APOE</i> , C388T	Изменение антикоагулянтной активности	Bryk et al., 2018 [603]
<i>BRCA1</i> , 5382insC	Развитие онкологических заболеваний	Sokolenko et al., 2006; Wang et al., 2012 [604, 605]

В каждом случае с помощью пары ген-специфических праймеров осуществляли амплификацию фрагмента ДНК, содержащего ОНП, смешивали неочищенный продукт с двумя ДСНП-зондами, обеспечивающими однозначное выявление аллели дикого типа и/или мутантной аллели. Для зонда, комплементарного аллели дикого типа, использовали пару «флуорофор и гаситель флуоресценции» FAM-DABCYL, для зонда, комплементарного мутантной аллели, - пару TAMRA-DABCYL. Таким образом, расщепление зонда, специфического для аллели дикого типа, приводило к возникновению флуоресценции с максимумом

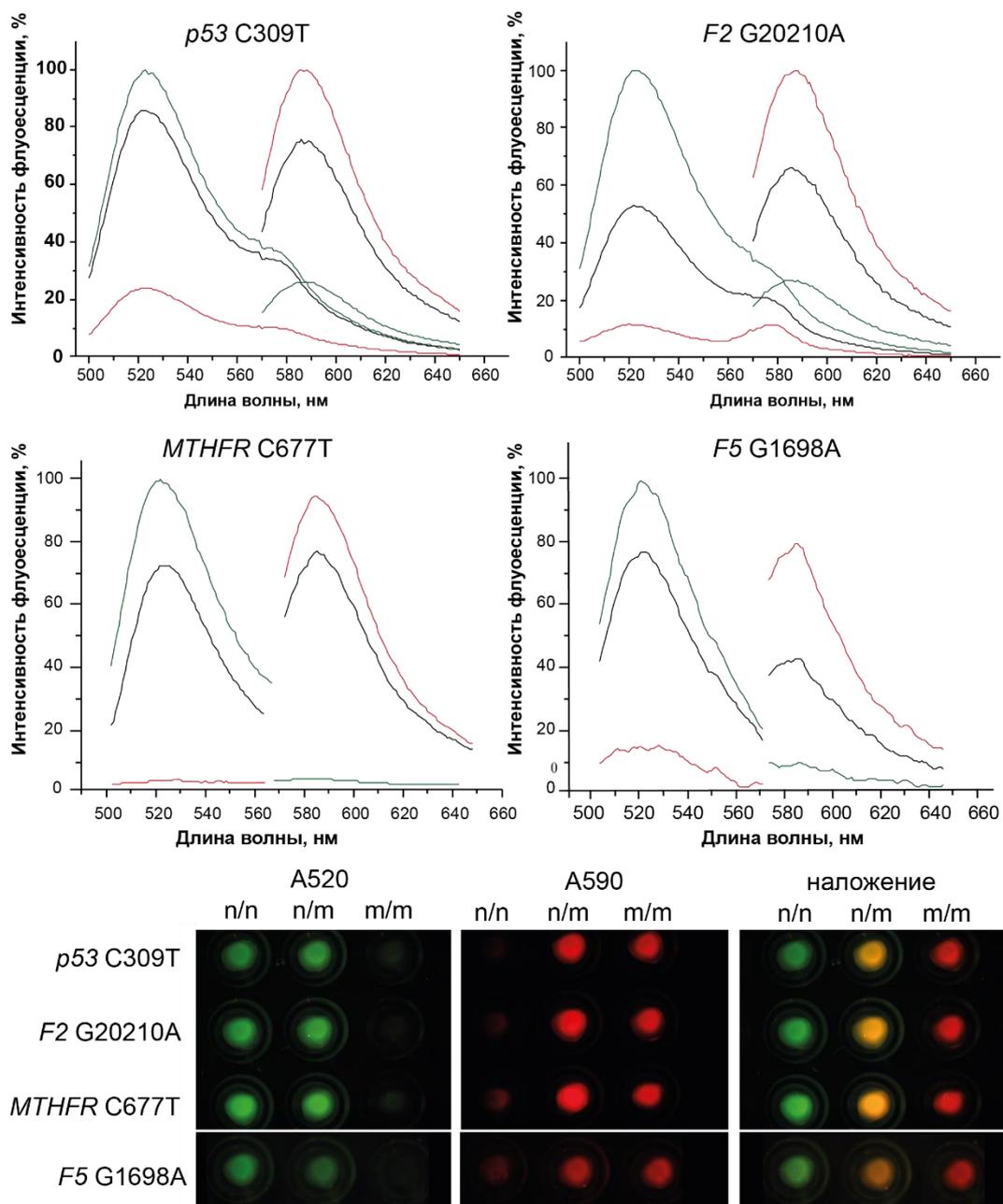
эмиссии при 520 нм (зеленый цвет), а зонда, специфического для мутантной аллели, - с максимумом эмиссии при 590 нм (красный цвет).

Результаты экспериментов представлены в таблице 3-13 и рисунке 3.26. Во всех случаях с помощью ДСНП-анализа удалось достоверно установить аллельный статус исследуемого образца. На примере полиморфизма 5382insC в гене *BRCA1* также была продемонстрирована применимость анализа для обнаружения однонуклеотидных делеций и инсерций. Следует отметить, что для выявления различных мутаций применяли стандартные условия реакции, что позволяет использовать метод ДСНП для широкого спектра мутаций.

Таблица 3-13. Результаты ДСНП-анализа аллельных вариантов в некоторых генах человека

Аллель	n/n			n/m			m/m		
	F ₅₂₀	F ₅₉₀	ΔF	F ₅₂₀	F ₅₉₀	ΔF	F ₅₂₀	F ₅₉₀	ΔF
<i>HRAS</i> G35T	594	68	8,7	469	536	0,9	49	555	0,1
<i>KRAS</i> G34A	750	56	13,4	507	548	0,9	56	658	0,1
<i>KRAS</i> G35A	744	54	13,8	567	611	0,9	52	798	0,1
<i>KRAS</i> G35T	756	63	12,0	617	591	1,0	61	753	0,1
<i>KRAS</i> G38A	472	48	9,8	354	360	1,0	54	421	0,1
<i>NRAS</i> G34A	724	67	10,8	633	704	0,9	42	873	0,0
<i>NRAS</i> G35C	790	52	15,2	629	689	0,9	71	853	0,1
<i>NRAS</i> G35A	773	66	11,7	600	675	0,9	54	799	0,1
<i>BRCA1</i> 5382insC	390	63	6,2	317	480	0,7	57	591	0,1
<i>P53</i> C309T	415	51	8,1	355	325	1,1	44	580	0,1
<i>F2</i> G20210A	407	45	9,0	310	272	1,1	38	340	0,1
<i>F5</i> G1698A (Leiden)	431	42	10,3	346	383	0,9	51	436	0,1
<i>MTHFR</i> C677T	421	56	7,5	257	318	0,8	50	396	0,1
<i>APOE</i> C388T	258	21	12,3	189	201	0,9	27	266	0,1

F₅₂₀ — интенсивность эмиссии флуоресценции при 520 нм, F₅₉₀ — интенсивность эмиссии флуоресценции при 590 нм (данные приведены в оптических единицах).
ΔF — отношение F₅₂₀ / F₅₉₀. n — аллель дикого типа, m — мутантный аллель.



Нормализованные графики флуоресценции образцов ДНК: зеленый цвет — образцы ДНК, гомозиготные по аллели дикого типа, красный — образцы ДНК, гомозиготные по мутантной аллели, черный — гетерозиготные образцы ДНК. Ниже показаны фотографии плашек с тестируемыми образцами, сделанные при помощи флуоресцентного стереомикроскопа с зеленым (эмиссия 520 нм) и красным (эмиссия 590 нм) фильтрами, и компьютерное наложение этих изображений. n — аллель дикого типа, m — мутантный аллель.

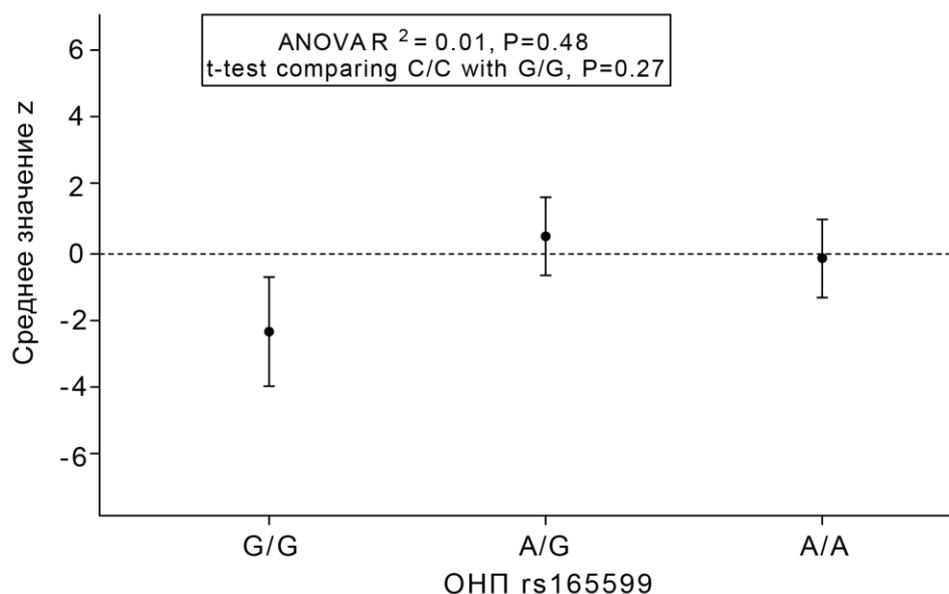
Рисунок 3.26. Примеры использования ДСНП-анализа для детекции ОНП в генах человека

ДСНП-анализ был использован в экспериментальной работе для анализа ОНП rs165599 гена *COMT*. Этот ген кодирует катехоламин-О-метилтрансферазу-фермент, вовлеченный в определение болевой чувствительности у млекопитающих. В рамках этой работы были идентифицированы три генетических гаплотипа гена *COMT*, которые определяли низкую, среднюю и высокую болевую чувствительность. Эти гаплотипы охватывают 96% человеческой популяции, и пять комбинаций этих гаплотипов тесно связаны ($P = 0,0004$) с изменением чувствительности к экспериментальной боли. Нами был проведен анализ ОНП rs165599 гена *COMT* в образцах геномной ДНК из периферической крови здоровых женщин-добровольцев, принявших участие в 3-летнем проспективном когортном исследовании для выявления факторов риска развития миогенного поражения височно-нижнечелюстного сустава. ОНП rs165599 расположен в самом конце 3'-некодирующей транскрибируемой области (UTR) гена *COMT*; по литературным данным его аллель G связан с шизофренией [606]. Результаты анализа ОНП rs165599 гена *COMT* представлены в таблице 3-14. Все определенные с помощью ДСНП-анализа аллельные статусы были подтверждены с помощью рестрикционного анализа. На рисунке 3.27 показаны результаты статистического анализа степени ассоциированности ОНП rs165599 гена *COMT* с болевой чувствительностью.

Таблица 3-14. Результаты ДСНП-анализа аллельных вариантов ОНП rs165599 гена COMT

Генотип	Число носителей	Частота встречаемости аллеля	Среднее значение (z-оценка)	ANOVA R^2	P	t
G/G	27	0.138	-2.4 (8.5)	0.008	0.48	0.27
A/G	87	0.446	0.4 (11.4)			
A/A	81	0.415	-0.2 (10.3)			

ANOVA — дисперсионный анализ, проверяющий нулевую гипотезу о равенстве средних среди трех аллелей; t — критерий Стьюдента, проверяющий нулевую гипотезу о равенстве средних между гомозиготами; p — значение вероятности для данной статистической модели о равенстве средних между гомозиготами



Первый представленный аллель связан с наименьшей чувствительностью к боли. Различия между комбинациями аллелей оценивали с помощью дисперсионного анализа (ANOVA) с использованием z-показателя в качестве зависимой переменной. Значимость различий в средних z-показателях, связанных с каждым гомозиготным генотипом, определяли с помощью t-критерия Стьюдента. Каждое значение представляет собой средний z-показатель.

Рисунок 3.27. Анализ ассоциации фенотипов с различной болевой чувствительностью с ОНП rs165599 гена COMT

Было показано, что ОНП rs165599 не оказывает существенного влияния на болевую чувствительность, что вместе с аналогичным результатом, полученным для ОНП, расположенным в 5'-UTR этого гена, позволило сделать вывод о низком вкладе UTR в модуляцию активности гена *COMT*, связанной с болевой чувствительностью.

Также нами была продемонстрирована возможность применения ДСНП-анализа для тестирования нескольких точечных мутаций на ограниченном участке ДНК. Такие мутации описаны, например, для генов *BRCA1* и *KRAS* при колоректальном раке [599], гена *NRAS* при острой миелоидной лейкемии [601] и гена *PVT1* при диабетической терминальной почечной недостаточности [607]. Высокая плотность ОНП также была описана для генов МНС [608]. Мутационный анализ таких регионов представляет большие сложности, большинство методов анализа мутаций не позволяет различать близкорасположенные ОНП.

В модельном эксперименте на примере мутаций в положениях 34 и 35 гена *KRAS*: G34A, G35A или G35T [599], мы продемонстрировали высокую разрешающую способность ДСНП-анализа даже в такой непростой ситуации (таблица 3-15). Было показано, что и при одновременном добавлении в реакционную смесь четырех различных ДСНП-зондов, специфический сигнал возникал только при наличии полностью комплементарной последовательности в исследуемом образце.

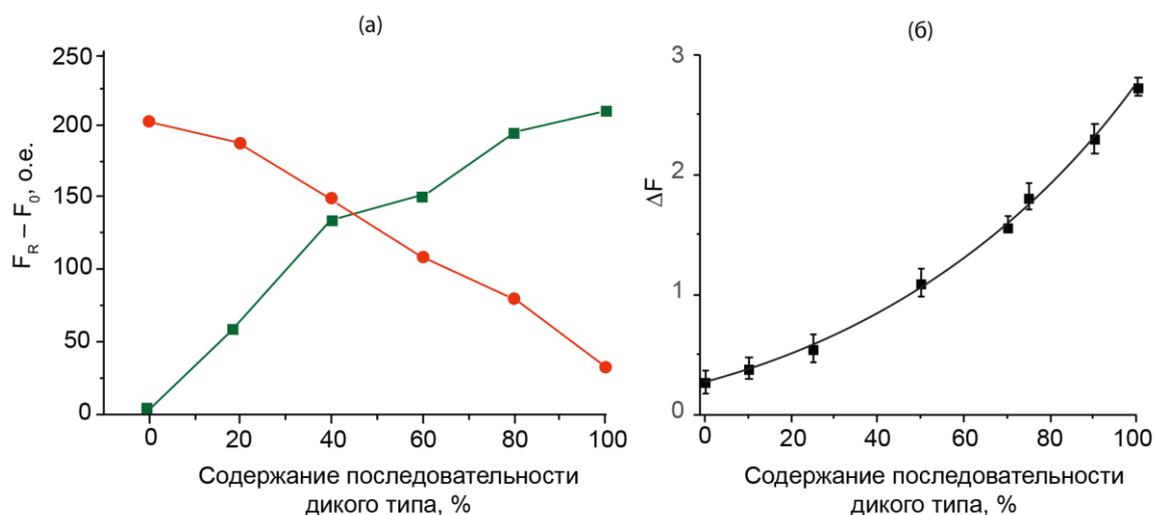
Таблица 3-15. ДСНП-анализ близкорасположенных мутаций в гене *KRAS*

Аллель	DSNP-зонды	F520	F590
Дикий тип*	<i>KRAS</i> -34/35wt+ <i>KRAS</i> -G35A+ <i>KRAS</i> -G35T+ <i>KRAS</i> -G34A	754	98
G35A*	<i>KRAS</i> -34/35wt+ <i>KRAS</i> -G35A	63	776
G35T	<i>KRAS</i> -34/35wt+ <i>KRAS</i> -G35A	45	42
G34A	<i>KRAS</i> -34/35wt+ <i>KRAS</i> -G35A	70	58
G35A	<i>KRAS</i> -34/35wt+ <i>KRAS</i> -G35T	70	63
G35T*	<i>KRAS</i> -34/35wt+ <i>KRAS</i> -G35T	85	853
G34A	<i>KRAS</i> -34/35wt+ <i>KRAS</i> -G35T	74	50
G35A	<i>KRAS</i> -34/35wt+ <i>KRAS</i> -G34A	78	67
G35T	<i>KRAS</i> -34/35wt+ <i>KRAS</i> -G34A	62	62
G34A*	<i>KRAS</i> -34/35wt+ <i>KRAS</i> -G34A	88	698

F520 — интенсивность эмиссии флуоресценции при 520 нм, возбуждение при 480 нм, F590 — интенсивность эмиссии флуоресценции при 590 нм, возбуждение при 560 нм. Звездочками отмечены образцы, содержащие зонд, полностью комплементарный анализируемому ДНК-фрагменту.

На примере анализа полиморфизма G35A в гене *KRAS* мы продемонстрировали применимость ДСНП-анализа для полуколичественной оценки соотношения различных аллелей. Такое исследование востребовано, например, при анализе суммарных образцов ДНК, полученных от разных доноров или при изучении гетерогенной опухолевой ткани [609, 610]. В модельном эксперименте в качестве ДНК-матрицы для ПЦР использовали синтетические 60-звеньевые олигонуклеотиды, один содержал последовательность гена *KRAS* дикого типа, а второй — последовательность с заменой G35A. Олигонуклеотиды

смешивали в различных пропорциях перед амплификацией. Продукты ПЦР были затем подвергнуты ДСНП-анализу в присутствии специфических зондов. Зонд, последовательность которого была полностью комплементарна последовательности гена дикого типа, содержал донор флуоресценции FAM, зонд, последовательность которого была комплементарна последовательности мутантного гена – TAMRA. Оба зонда содержали гаситель флуоресценции DABCYL. На первом этапе в присутствии полностью комплементарных зондов и субстратов было определено оптимальное время инкубации, составившее 30 мин. На следующем этапе было проведено исследование зависимости интенсивности флуоресценции и ее относительного изменения от соотношения концентраций ДНК-фрагментов, содержащих мутантную аллель и аллель дикого типа (рисунок 3.28).

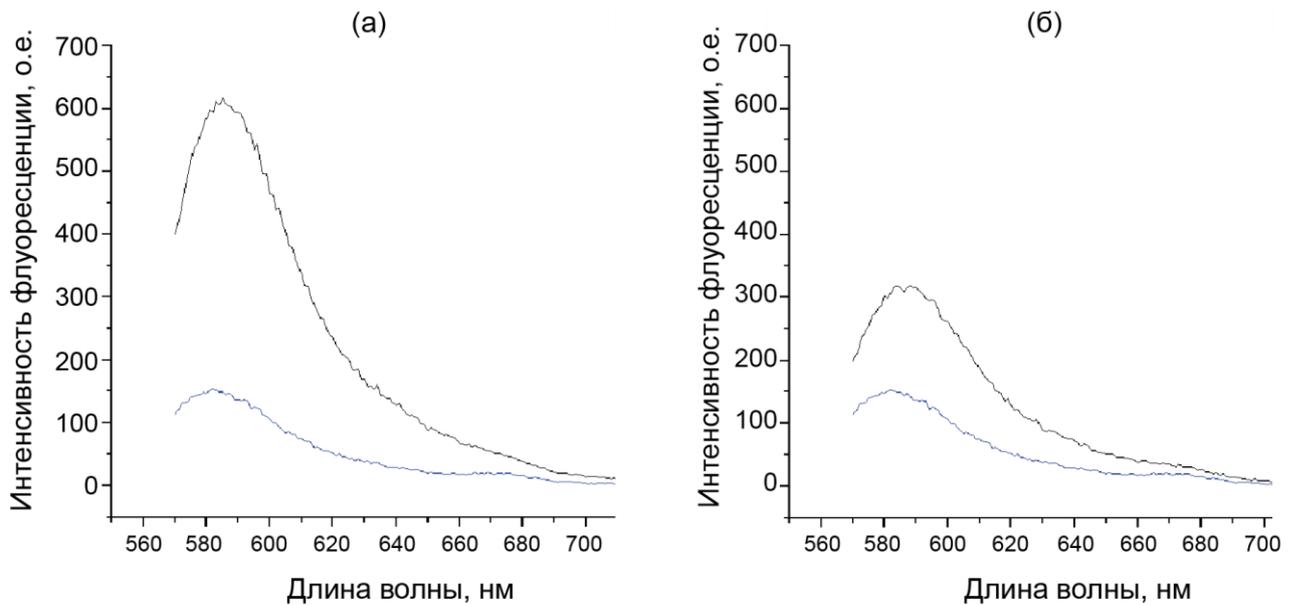


(а) График зависимости интенсивности флуоресценции образцов ДНК после инкубации с Paq_DSN от относительного содержания ДНК дикого типа: зеленый — флуоресценция FAM, красный — флуоресценция TAMRA; (б) график зависимости относительного изменения флуоресценции ΔF от относительного содержания ДНК дикого типа, где $\Delta F = F_{590} \setminus F_{520}$. F_R — уровень флуоресценции реакционной смеси после инкубации в присутствии нуклеазы, F_0 — фоновая флуоресценция реакционной смеси в отсутствие фермента, $F_{590} = F_R - F_0$ для TAMRA, $F_{520} = F_R - F_0$ для FAM, о.е. — оптические единицы.

Рисунок 3.28. Использование ДСНП-анализа для оценки относительной концентрации аллеля в образцах ДНК

Полученные калибровочные кривые позволяют оценивать относительное (процентное) содержание мутантной аллели и аллели дикого типа в образцах ДНК с точностью до 10%. Нелинейный характер полученной зависимости объясняется различиями в интенсивности флуоресценции использованных флуорофоров.

Очевидным применением ДСНП-анализа, кроме тестирования мутационного статуса образцов ДНК, является выявление любых целевых последовательностей в образцах ДНК. А в связи с тем, что Par_DSN так же эффективно расщепляет ДНК как в ДНК-ДНК гибридах, так и в ДНК-РНК гибридах, протокол ДСНП-анализа легко адаптируется для поиска целевых последовательностей и ОНП в образцах РНК. Благодаря тому, что РНК не подвергается атаке со стороны Par_DSN, метод анализа РНК не нуждается в дополнительных ферментативных реакциях, направленных на амплификацию сигнала. На рисунке 3.29 показаны результаты эксперимента по выявлению целевой последовательности в образцах синтетической РНК. В эксперименте использовали образец синтетической РНК и ДСНП-зонд, меченный парой «флуорофор-гаситель флуоресценции» TAMRA-DABCYL.



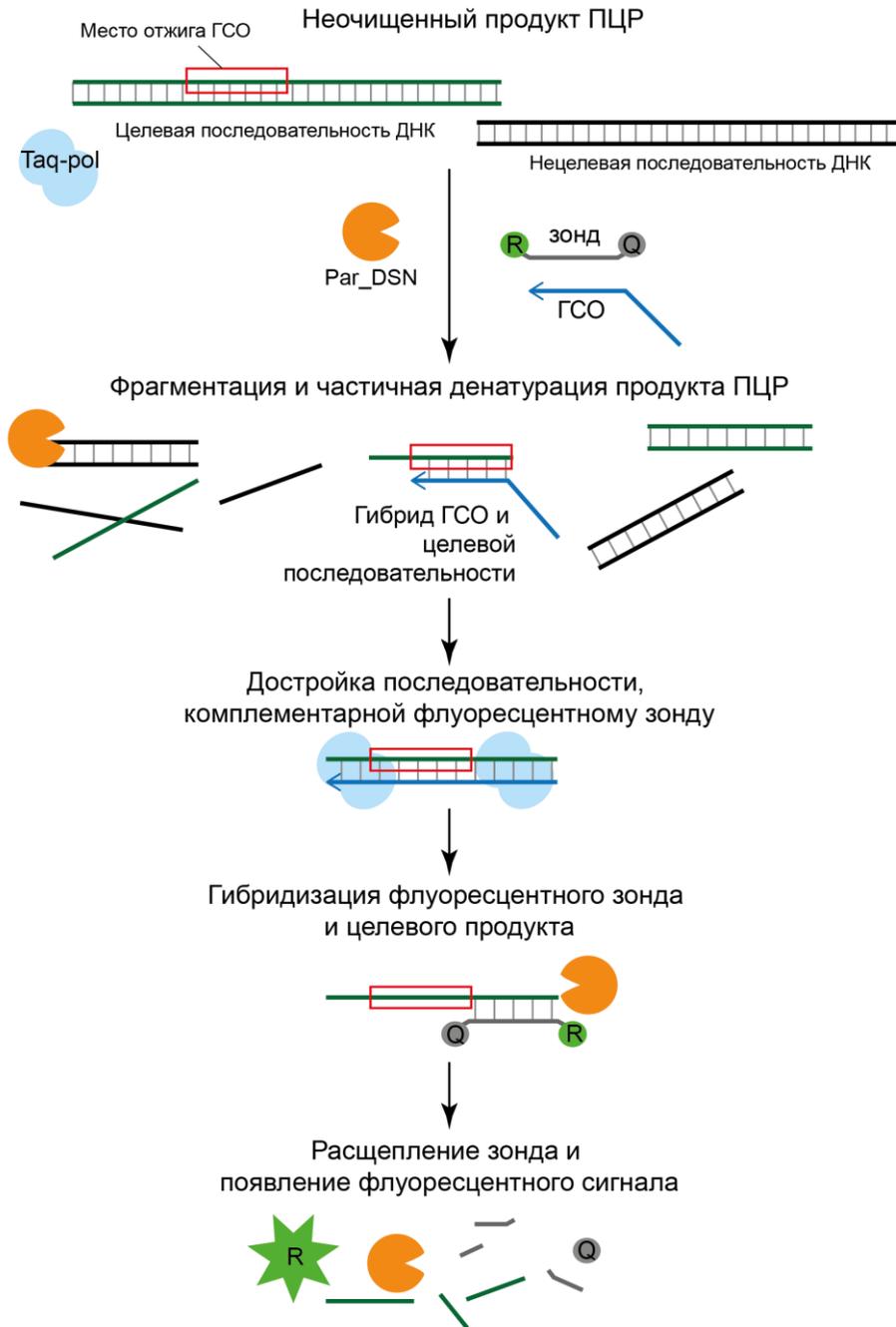
(а) реакция со смысловой РНК и комплементарным ей зондом; (б) реакция с антисмысловой РНК (отрицательный контроль). Синяя линия — флуоресценция реакционной смеси после инкубации в отсутствие Par_DSN; черная линия — флуоресценция реакционной смеси после инкубации в присутствии Par_DSN.

Рисунок 3.29. Использование ДСНП-анализа для выявления целевой последовательности в образце РНК

В описанном выше варианте ДСНП-анализа для каждой анализируемой последовательности необходим синтез своего ДСНП-зонда, что в определенной степени удорожает проведение и усложняет дизайн эксперимента. Эта же проблема присуща ряду других популярных методов, таких как метод с использованием молекулярных маяков [493, 494] и метод TaqMan анализ [611]. В ряде исследований была постулирована необходимость разработки методов ОНП-анализа с универсальными зондами, позволяющими избежать подобного дорогостоящего синтеза [494, 612]. Другая проблема связана с выявлением целевых ДНК мишеней в комплексных продуктах ПЦР. Одна из часто возникающих проблем ПЦР — образование нецелевых продуктов в ходе реакции при использовании в качестве матрицы ДНК высокой сложности (например, образцов геномной ДНК или суммарной кДНК). Содержание целевой последовательности в таких образцах часто довольно низкое, и поэтому высока вероятность получения в ходе ПЦР

нецелевых продуктов, в том числе нецелевых продуктов той же длины, что и целевая амплифицируемая последовательность. Для изоляции целевых последовательностей из таких смешанных продуктов ПЦР используют молекулярное клонирование и производят отбор клонов с помощью рестрикционного анализа, дот-блот гибридизации или их прямого секвенирования. Таким образом, выявление целевого продукта в образцах ДНК без дополнительных трудоемких стадий является актуальной задачей.

Для решения этой задачи нами была предложена модификация метода ДСНП-анализа с использованием ген-специфического олигонуклеотида (ГСО) и 10-звеньевое универсального зонда, содержащего пару «флуорофор и гаситель флуоресценции». ГСО включает последовательность, комплементарную последовательности целевого продукта, и ковалентно связанную с ней последовательность, идентичную последовательности универсального зонда. Схематически метод, получивший название ДСН-детекция, представлен на рисунке 3.30.



R — флуорофор, Q — гаситель флуоресценции, TaqPol — ДНК-полимераза, ГСО — ген-специфический олигонуклеотид.

Рисунок 3.30. Способ выявления целевой последовательности с использованием Par_DSN

Свежий продукт ПЦР инкубируют короткое время в присутствии ГСО, зонда и Par_DSN при 60°C. Во время инкубации за счет нуклеазной активности Par_DSN происходит расщепление ДНК и гибридизация ГСО с этой фрагментированной и частично денатурированной ДНК. ДНК-полимераза, оставшаяся в реакционной смеси, используя в качестве матрицы ГСО, образует продукт амплификации,

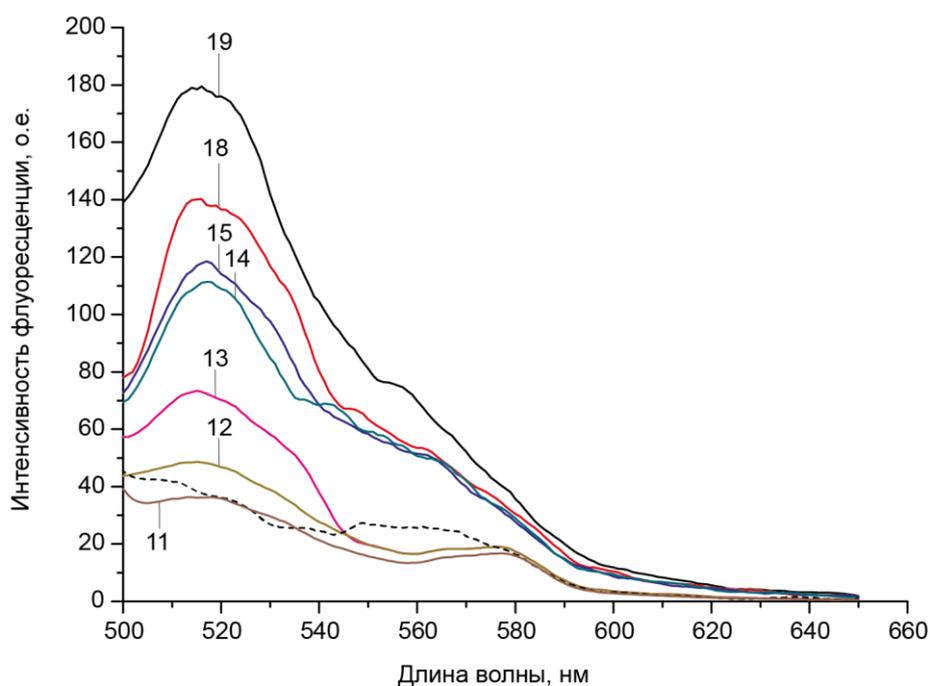
содержащий последовательность, комплементарную универсальному зонду. На следующем этапе температуру гибридизации понижают, обеспечивая условия для гибридизации зонда с комплементарной ему последовательностью. В результате расщепления зонда в образовавшихся ДНК-ДНК дуплексах нуклеазой Par_DSN происходит высвобождение флуорофора и появляется флуоресценция специфической длины волны, которая может быть детектирована любым доступным способом. В отсутствии целевой ДНК для ГСО не находится комплементарной последовательности, в результате не появляется места для отжига зонда и флуоресцентный сигнал не формируется.

Метод был протестирован в модельной системе для выявления гипервариабельных мотивов вариабельных доменов β -субъединицы Т-клеточного рецептора (ТКР). Анализ этих последовательностей востребован при исследованиях аутоиммунных, инфекционных и онкологических заболеваний [613-617]. На основании анализа мотивов, взаимодействующих с молекулами главного комплекса гистосовместимости, для β -субъединицы ТКР выделяют 24 основных семейства. Последовательности гипервариабельных мотивов этих семейств можно амплифицировать в ПЦР с использованием известных пар геноспецифических праймеров [618].

Продукт амплификации фрагмента вариабельного домена ТКР восемнадцатого семейства (ТКР18 β), полученный с использованием универсального праймера (TCR-C-Uni) и встречного праймера, специфического для последовательностей этого семейства (BV18), выглядит при электрофоретическом разделении как одна полоса. Однако при анализе библиотеки, полученной путем клонирования этого ПЦР-продукта, было обнаружено, что более 60% вставок представлены нецелевыми последовательностями. Эта библиотека была выбрана в качестве модельного объекта для ДСН-детекции.

Предварительно была подобрана оптимальная длина специфической последовательности ГСО. В эксперименте использовали универсальный зонд (UZ-

Fam), меченый парой «донор-гаситель флуоресценции» FAM-DABCYL, и ГСО различной длины. Плазмиду, содержащую целевой фрагмент гена ТКР18 β , и контрольную плазмиду, содержащую нецелевую вставку, амплифицировали с универсальных плазмидных праймеров M13 и продукт ПЦР использовали в качестве ДНК-субстрата. Было показано, что уменьшение длины ген-специфической части ГСО сопровождается уменьшением эффективности ДСН-детекции. При длине ген-специфической части ГСО 10-14 н.о. интенсивность специфической флуоресценции, возникающей при расщеплении зонда, мало отличима от фона (рисунок 3.31).



Цифрами показана длина ген-специфической последовательности ГСО. Прерывистая линия — интенсивность флуоресценции контрольного образца (без целевого продукта ПЦР). О.е. — оптические единицы.

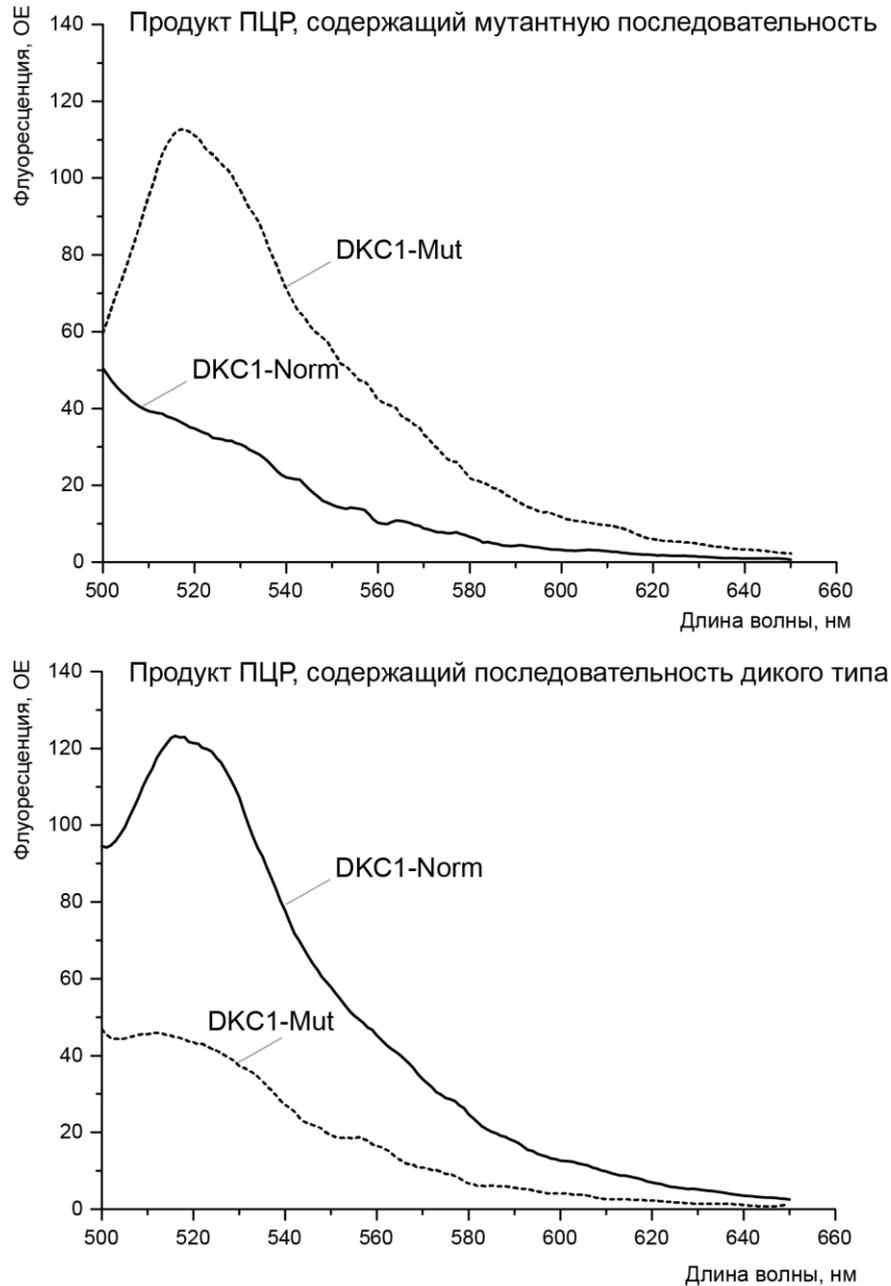
Рисунок 3.31. Зависимость эффективности ДСН-детекции от длины геноспецифической последовательности ген-специфического праймера ГСО

Следующим этапом работы стало тестирование клонированного продукта ПЦР, полученного при амплификации кДНК человека с использованием праймеров TCR-C-Uni и BV18. Для детекции использовали ГСО с геноспецифической

последовательностью длиной 17 н.о. Результаты, полученные с помощью ДСН детекции верифицировали секвенированием. Во всех случаях секвенирование подтвердило результаты ДСН-детекции.

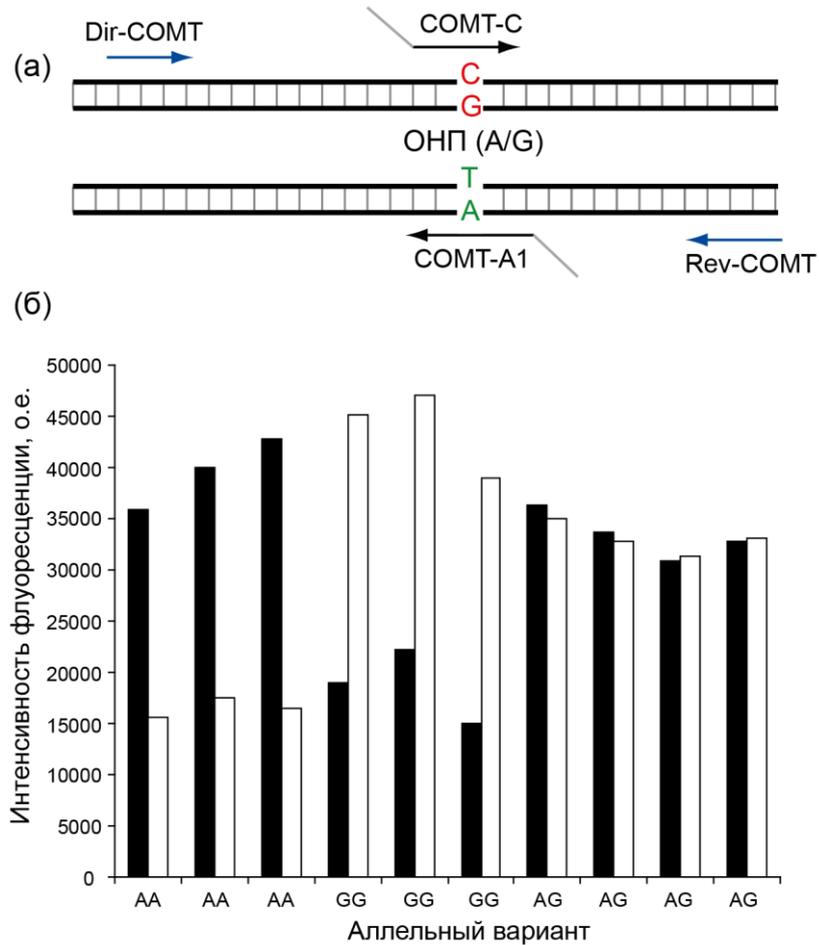
Метод ДСН-детекции был также использован для тестирования в модельной системе новой мутации – двухнуклеотидной инверсии 166_167invCT в третьем экзоне гена *DKC1* (GenBank NM_001363). Эта мутация была выявлена нами при исследовании случая врожденного дискератоза [536]. Для анализа использовали два ГСО, один включал ген-специфическую последовательность длиной 17 н.о., полностью комплементарную варианту *DKC1* дикого типа (*DKC1-Norm*), второй – мутантному варианту (*DKC1-Mut*). Оба ГСО содержали на 5'-конце последовательность UZ-Fam. ДНК для анализа получали путем ПЦР с плазмидной ДНК, содержащей целевые вставки. Результаты эксперимента представлены на рисунке 3.32. Было продемонстрировано, что ДСН-детекция позволяет уверенно выявлять последовательности ДНК, отличающиеся двухнуклеотидной инверсией.

Для исследования возможности различать последовательности, отличающиеся однонуклеотидной заменой, был проведен анализ девяти ранее охарактеризованных образцов ДНК, содержащих ОНП rs165599 гена *SOMT*. В работе использовали два ГСО: *SOMT-A1*, содержащий геноспецифическую последовательность длиной 16 н.о., совпадающую с последовательностью «аллеля А», и *SOMT-C*, содержащий геноспецифическую последовательность длиной 15 н.о., комплементарную последовательности «аллеля G» (рисунок 3.33а). Фрагменты, содержащие целевую последовательность ДНК, амплифицировали в ПЦР с использованием двух встречных праймеров и использовали в ДСН-детекции образцов гомо- и гетерозиготной ДНК. Результаты анализа продемонстрированы на рисунке 3.33б. Показано, что метод позволяет уверенно различать аллельные варианты в образцах геномной ДНК.



Каждый из продуктов ПЦР инкубировали в присутствии ГСО DKC1-Norm (сплошная линия) или DKC1-Mut (пунктирная линия), зонда UZ-Fam и Par_DSN; затем измеряли флуоресценцию, появившуюся в результате высвобождения флуорофора FAM.

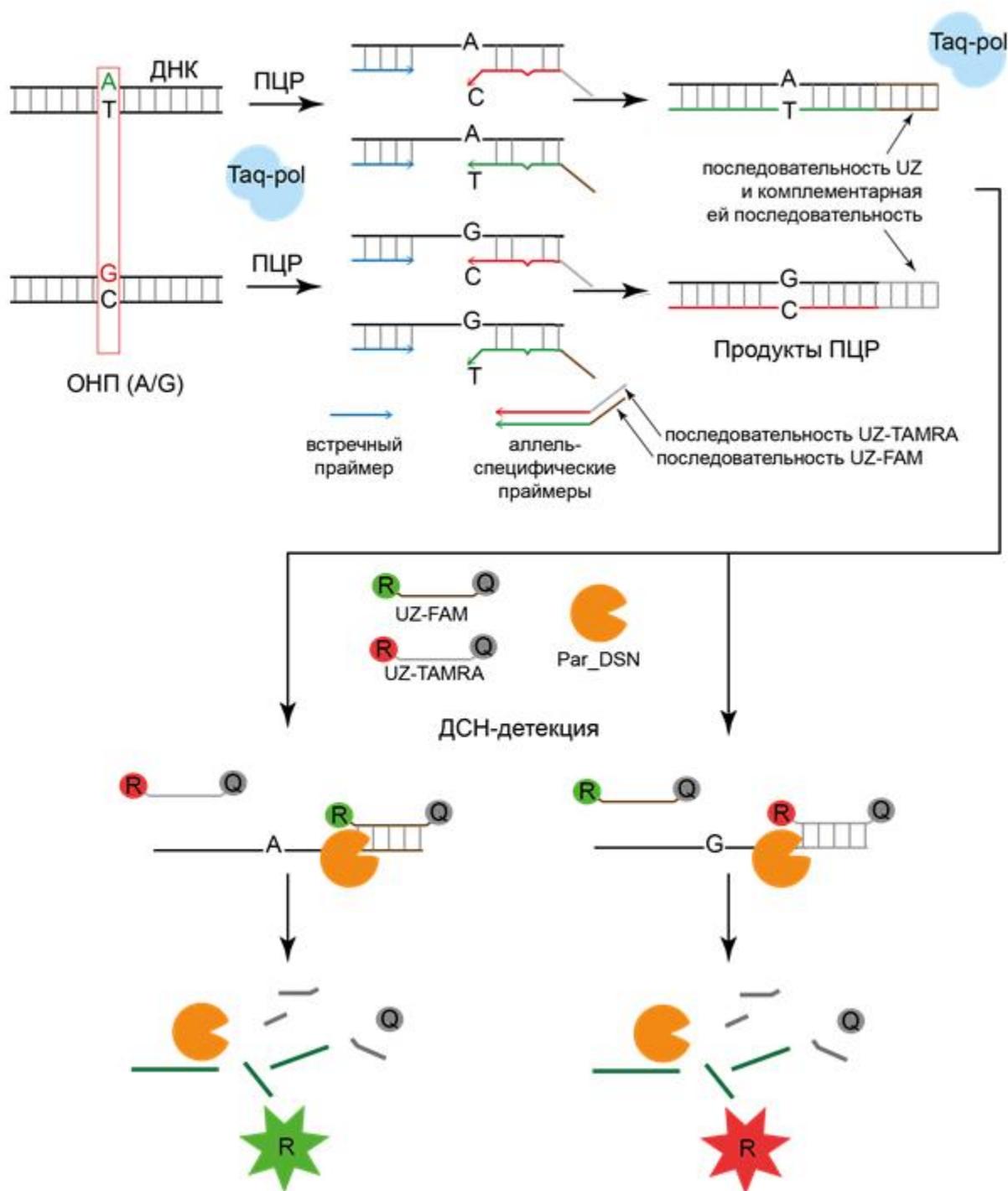
Рисунок 3.32. ДСН-детекция двухнуклеотидной инверсии 166_167invCT гена DKC1 в модельном эксперименте



(а) Структура использованных ГСО. Серым показана последовательность, идентичная последовательности универсального зонда. (б) Результат тестирования ОНП rs165599 гена *COMT* в гомо- и гетерозиготных образцах геномной ДНК. Аллельный статус образца указан по оси X. О.е. — оптические единицы.

Рисунок 3.33. ДСН-детекция ОНП rs165599 гена *COMT* в образцах геномной ДНК

Наконец, мы совместили технологию ДСН-детекции с аллель-специфической ПЦР. Схематически метод представлен на рисунке 3.34. Для амплификации используются аллель-специфические праймеры: первый праймер содержит на 5'-конце последовательность UZ-FAM, второй – последовательность другого универсального зонда UZ-TAMRA. На 3'-концы праймеров в третьем положении была добавлена дополнительная замена нуклеотидного основания. Как было показано Kwok et al. (1994) [619], этот прием повышает эффективность выявления ОНП в ходе аллель-специфической ПЦР.



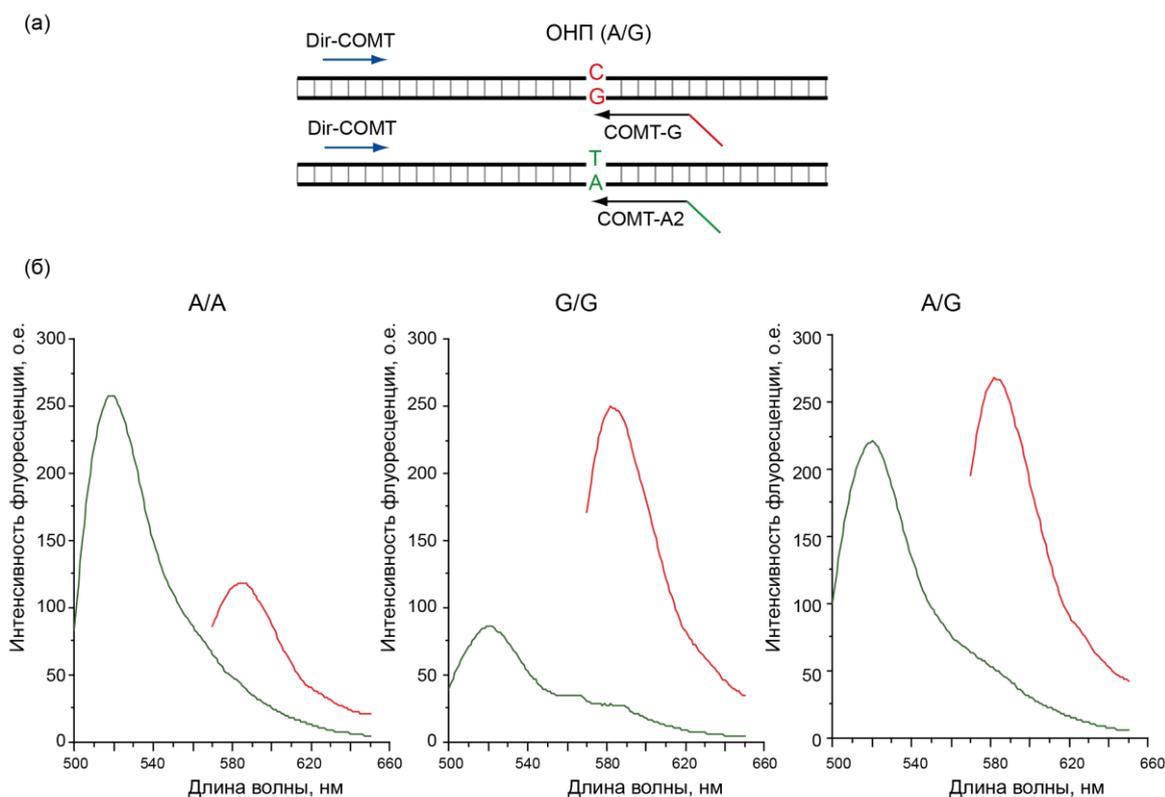
R — флуорофор; Q — гаситель флуоресценции. Подробное описание стадий ДСН-детекции показано на рис. 3.30.

Рисунок 3.34. Схема ДСН-детекции, совмещенной с аллель-специфической ПЦР

Аmplification фрагментов ДНК на тестируемой ДНК матрице проводят в ПЦР с конкурирующими за место отжига аллель-специфическими праймерами и общим встречным праймером. После амплификации реакцию смесь

инкубируют с эндонуклеазой I для удаления избытков праймеров, а затем используют ее для ДСН-детекции.

Метод был апробирован на модели ОНП rs165599 гена *COMT* с использованием 30 ранее охарактеризованных образцов. Во всех случаях результаты ДСН-детекции совпали с полученными ранее данными. Результаты эксперимента показаны на рис. 3.35.



(а) Схема расположения праймеров, использованных для амплификации фрагмента гена *COMT*. Зеленым и красным цветом показаны последовательности, идентичные универсальным зондам UZ-FAM и UZ-TAMRA соответственно. (б) Графики флуоресценции гомо- и гетерозиготных по аллели rs165599 образцов ДНК после ДСН-детекции. Зеленая линия — интенсивность эмиссии флуоресценции зонда UZ-FAM; красная линия — интенсивность эмиссии флуоресценции зонда UZ-TAMRA; о.е. — оптические единицы. AA — образец, гомозиготный по аллели А, GG — образец, гомозиготный по аллели G, AG — гетерозиготный образец.

Рисунок 3.35. Результат выявления ОНП rs165599 гена *COMT* в образцах геномной ДНК с помощью ДСН-детекцией, совмещенной с аллель-специфической ПЦР.

Таким образом, на основе Par_DSN нами был разработан комплекс методов для выявления целевых последовательностей в образцах ДНК и РНК, обладающий

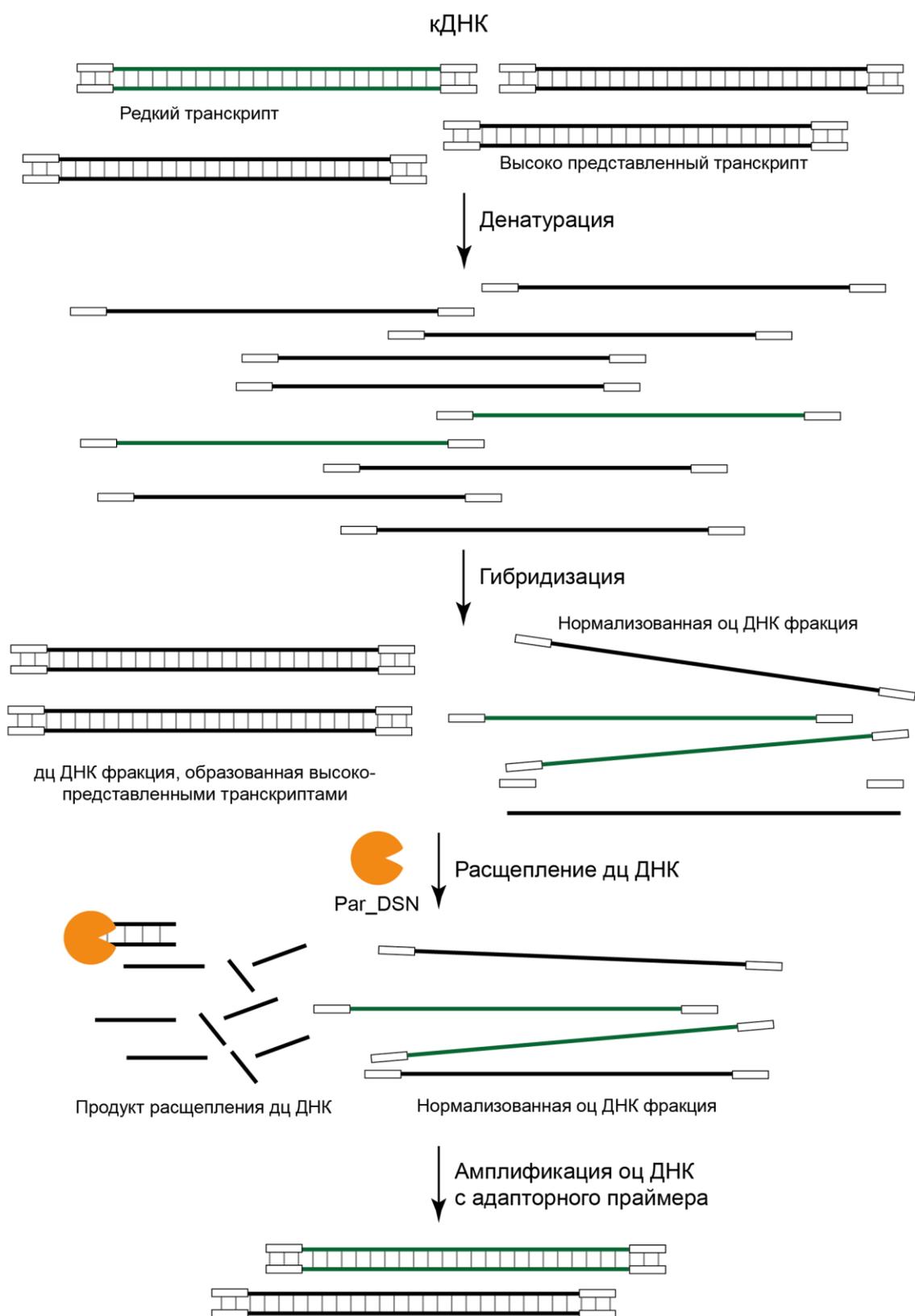
высокой специфичностью и позволяющий различать нуклеиновые кислоты, отличающиеся всего одной нуклеотидной заменой. Высокая адаптивность предложенного технологий и отсутствие трудозатратных процедур позволяет легко адаптировать их под самые различные нужды исследователей и различный формат экспериментов. Система легко масштабируется как на увеличение числа анализируемых проб, так и на расширение спектра последовательностей, анализируемых в одном образце. В качестве стартового материала для анализа может использоваться геномная ДНК, плазмидная ДНК, кДНК и РНК. Разработанные нами методы были апробированы в модельных экспериментах и на экспериментальных биологических образцах.

Уникальная возможность использовать в качестве стартового материала РНК и наличие коммерчески доступного препарата Par_DSN (коммерческое название «Duplex-specific nuclease» или DSN, производитель ЗАО Евроген, Россия) сделала нашу технологию высоко востребованной в мировой лабораторной практике. Основным фокусом исследователей стал анализ микроРНК в различных биологических системах. МикроРНК – популяция некодирующих молекул оц РНК длиной 20–24 н.о., которые играют важную роль в дифференцировке клеток и развитии заболеваний [620]. Абберантно экспрессируемые специфические популяции микроРНК ассоциированы с рядом опасных заболеваний человека, включая сердечно-сосудистые и онкологические заболевания [621, 622]. Анализ таких популяций в последние годы приобретает все большее значение как в лабораторной практике, так и для клинической диагностики. Потребность в способах быстрой детекции микроРНК привела к появлению целого веера модификаций, созданных в различных лабораториях на основе разработанных нами технологий ДСН-детекции и ДСНП-анализа. В основе всех этих методов лежит использование рециркуляции РНК-мишеней в присутствии сигнальных оц ДНК-зондов, которые за счет комплементарности образуют с РНК ДНК-РНК дуплексы и подвергаются Par_DSN-опосредованному расщеплению. Благодаря рециркуляции РНК-мишеней, следовые микроРНК вызывают высвобождение значительного количества «сигнальных молекул», исходно связанных с оц ДНК-

зондом [34, 623]. Во множестве работ были предложены разнообразные «сигнальные молекулы» и различные варианты их анализа, получившие общее название «биосенсоры микроРНК на основе DSN». В настоящее время различных модификация таких биосенсоров насчитывается более 30 видов, включая флуоресцентные биосенсоры [623, 624], колориметрические биосенсоры [625-627], люминесцентные биосенсоры [628, 629], меченные биотином биосенсоры [630], электрохимические биосенсоры [631, 632], проточные цитометрические биосенсоры [633], масс-спектрометрические биосенсоры [634] и так далее.

3.4. Нормализация кДНК

Как уже упоминалось выше, удаление дц ДНК фракции является основной проблемой методов нормализации образцов НК, предназначенных для исчерпывающего секвенирования или для приготовления библиотек. Par_DSN, обладающая крайне высокой специфичностью к дц ДНК и термостабильностью, позволила решить эту проблему. Общий принцип нормализации с использованием Par_DSN (далее ДСН-нормализация) амплифицированной кДНК схематически представлен на рисунке 3.36. Метод основан на использовании кинетики гибридизации НК и отличается от альтернативных технологий способом удаления дц ДНК фракции.



Прямоугольниками обозначены последовательности адаптеров. Редкий транскрипт обозначен зеленой линией, высокопредставленный — черной линией.

Рисунок 3.36. Схема нормализации кДНК с использованием Par_DSN

В ходе ДСН-нормализации дц ДНК денатурируют и гибридизуют в условиях, обеспечивающих реассоциацию комплементарных молекул ДНК. Во время реассоциации высокопредставленные в образце молекулы образуют дц фракцию, подлежащую удалению, а в оц фракции происходит выравнивание концентраций различных транскриптов. По окончании гибридизации для удаления нецелевой дц кДНК фракции реакционную смесь обрабатывают Par_DSN. Оставшуюся нормализованную оц фракцию кДНК далее амплифицируют с помощью ПЦР. Так как Par_DSN сохраняет высокую активность при температуре 70°C, удаление дц ДНК из смеси НК может осуществляться при температуре гибридизации. Благодаря этому во время нуклеазной обработки не происходит потери транскриптов, склонных к образованию вторичных структур.

Для нормализации используется амплифицированная кДНК, содержащая на 3'- и 5'-концах последовательности специфических адаптеров. Такие адаптеры могут быть введены на концы кДНК с помощью различных методов, включая методы лигирования [635-637], методы с использованием терминальной трансферазы [638] и методы введения адаптеров в ходе синтеза кДНК [532, 639].

Наибольший интерес представляет нормализация образцов кДНК, предназначенных для различных исследований транскриптома, включая приготовление клонированных библиотек полноразмерных кДНК и высокопроизводительное секвенирование. Иными словами, особый интерес представляет нормализация репрезентативных образцов кДНК, обогащенных полноразмерными транскриптами. Среди многочисленных методик, направленных на получение таких образцов, технология синтеза кДНК SMART [639] является одной из наиболее простых в применении, воспроизводимых и широко используемых [640, 641]. Учитывая преимущества метода SMART по сравнению с альтернативными подходами, мы в своей работе использовали образцы кДНК, полученные этим способом. Кратко, метод SMART основан на способности обратной транскриптазы MMLV при достижении конца РНК-матрицы добавлять на 3'-конец новосинтезированной первой цепи кДНК несколько нематричных остатков цитозина [642]. Адаптер Smart Oligo, содержащий на 5'-конце участок

поли(рибоG), образует комплементарные связи с поли(С)-участком конца первой цепи кДНК и воспринимается обратной транскриптазой как продолжение считываемой РНК. Таким образом, обратная транскриптаза «перескакивает» на новую матрицу и вводит адаптерную последовательность, комплементарную Smart Oligo, на 3'-конец первой цепи кДНК. В результате первая цепь содержит на 5'-конце адаптерную последовательность праймера (обычно олиго(Т)-праймер), который служил в качестве затравки синтеза первой цепи, а на 3'-конце – последовательность, введенную с помощью Smart Oligo [532, 639]. С учетом различных нужд исследователей нами были проверены различные варианты адаптеров для синтеза кДНК, позволяющие получать образцы нормализованной кДНК, фланкированные как одинаковыми, так и различными адаптерными последовательностями, что позволяло в дальнейшем проводить как ненаправленное, так и направленное клонирование кДНК библиотек. Мы использовали коммерчески доступные адаптеры Smart Oligo II, Smart Oligo IV и CDS-3 (Clontech, США), а также адаптеры, разработанные в рамках этой работы. Структуры апробированных адаптеров показана на рисунке 3.37.

Был разработан протокол нормализации, включающий этапы контроля двух принципиальных стадий: амплификации кДНК и обработки нормализованной кДНК нуклеазой. Так как сохранение репрезентативности исходной библиотеки прямо и непосредственно влияет на качество нормализованного образца, первым этапом является подбор оптимальных условий амплификации кДНК. Оптимальным при амплификации считали число циклов ПЦР на 2-3 цикла меньше, чем необходимо для выхода реакции на плато. Также необходимо строго контролировать стадию ферментативного удаления дц ДНК фракции, которая потенциально может привести к утрате целевых последовательностей и обеднению образца. Поэтому в ходе амплификации нормализованной фракции ДНК после обработки образцов нуклеазой также необходим анализ качества нормализации. В оптимально нормализованном образце при разрешении на электрофореze в агарозе с окраской интеркалирующим красителем наблюдается исчезновение характерных

полос, соответствующих высокопредставленным транскриптам. При этом средний размер длины ДНК фрагментов должен сохраняться.

Последовательности адаптеров для направленного клонирования

SMART Oligo IV: 5'-AAGCAGTGGTATCAACGCAGAGTGGCCATTACGGCCrGrGrG-3'
SfiI

CDS-3M адаптор: 5'-AAGCAGTGGTATCAACGCAGAGTGGCCGAGGGCGGCC (T)₂₀-3'
SfiI

SMART Oligo II: 5'-AAGCAGTGGTATCAACGCAGAGTACGCrGrGrG-3'
RsaI

CapNot1 адаптор: 5'-AAGCAGTGGTATCAACGCAGAGTGC GGCCGC (T)₂₀-3'
NotI

Последовательности адаптеров для ненаправленного клонирования

SMART Oligo II: 5'-AAGCAGTGGTATCAACGCAGAGTACGCrGrGrG-3'
RsaI

CDS-3 адаптор: 5'-AAGCAGTGGTATCAACGCAGAGTAC (T)₃₀-3'
RsaI

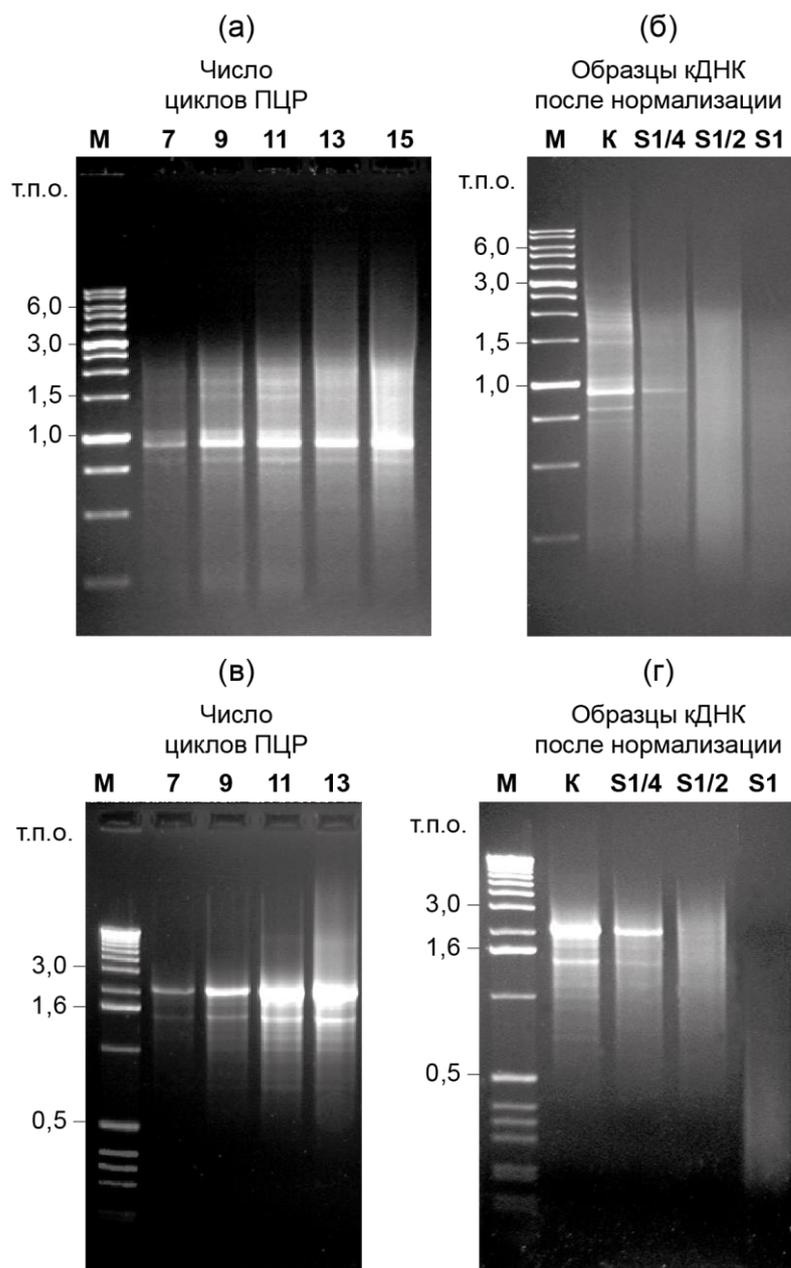
Праймер для амплификации нормализованной библиотеки

ПЦР праймер M1: 5'-AAGCAGTGGTATCAACGCAGAGT-3'

Серой линией отмечены общие части адаптеров, подчеркнуты и обозначены сайты рестрикции.

Рисунок 3.37. Последовательности адаптеров, апробированных для ДСН-нормализации

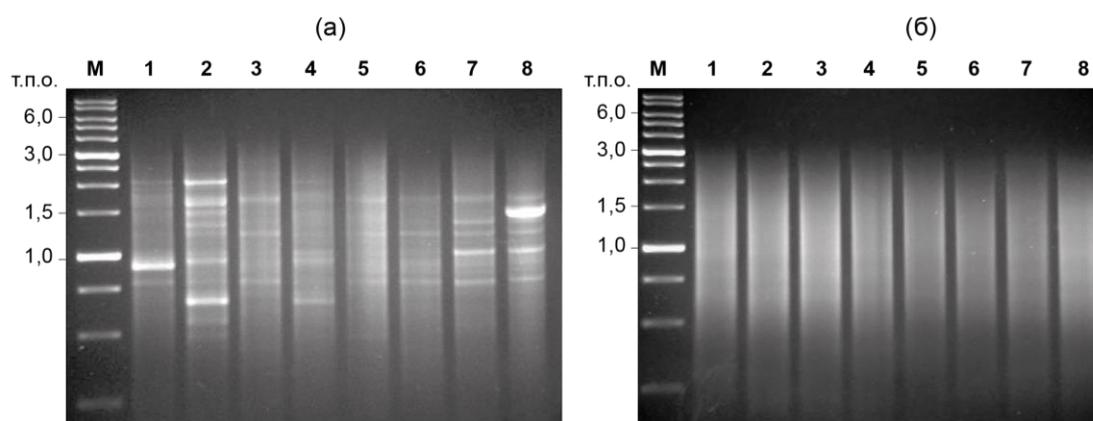
На рисунке 3.38 показаны примеры ДСН-нормализации с контролем критических стадий. В качестве стартового материала использовалась суммарная РНК из плаценты и из скелетной мышцы человека. В первом случае для приготовления первой цепи и амплификации кДНК была использована комбинация адаптеров SMART Oligo II + CDS-3, во втором – SMART Oligo IV + CDS-3M.



(а, в) Результат электрофореза в агарозном геле продуктов амплификации кДНК. Число циклов ПЦР указано над дорожками. Оптимальная амплификация была достигнута после 11 циклов ПЦР в случае (а) и 10 циклов ПЦР в случае (в). (б, г) Результат электрофореза в агарозном геле этих кДНК после нормализации. S1/4 — нормализация в присутствии 0,25 ед Куница Par_DSN, S1/2 — в присутствии 0,5 ед Куница Par_DSN, S1 — в присутствии 1 ед Куница Par_DSN, К — контрольный образец, инкубация без нуклеазы. М (а, б) — маркер длин ДНК, 1 kb ladder (Сибэнзим, Россия); М (в, г) — маркер длин ДНК, 1 kb ladder (Invitrogen, США).

Рисунок 3.38. Нормализация кДНК, полученной на основе суммарной РНК из плаценты (а, б) и скелетной мышцы (в, г) человека с использованием комбинаций адаптеров SMART Oligo II + CDS-3 и SMART Oligo IV + CDS-3M соответственно

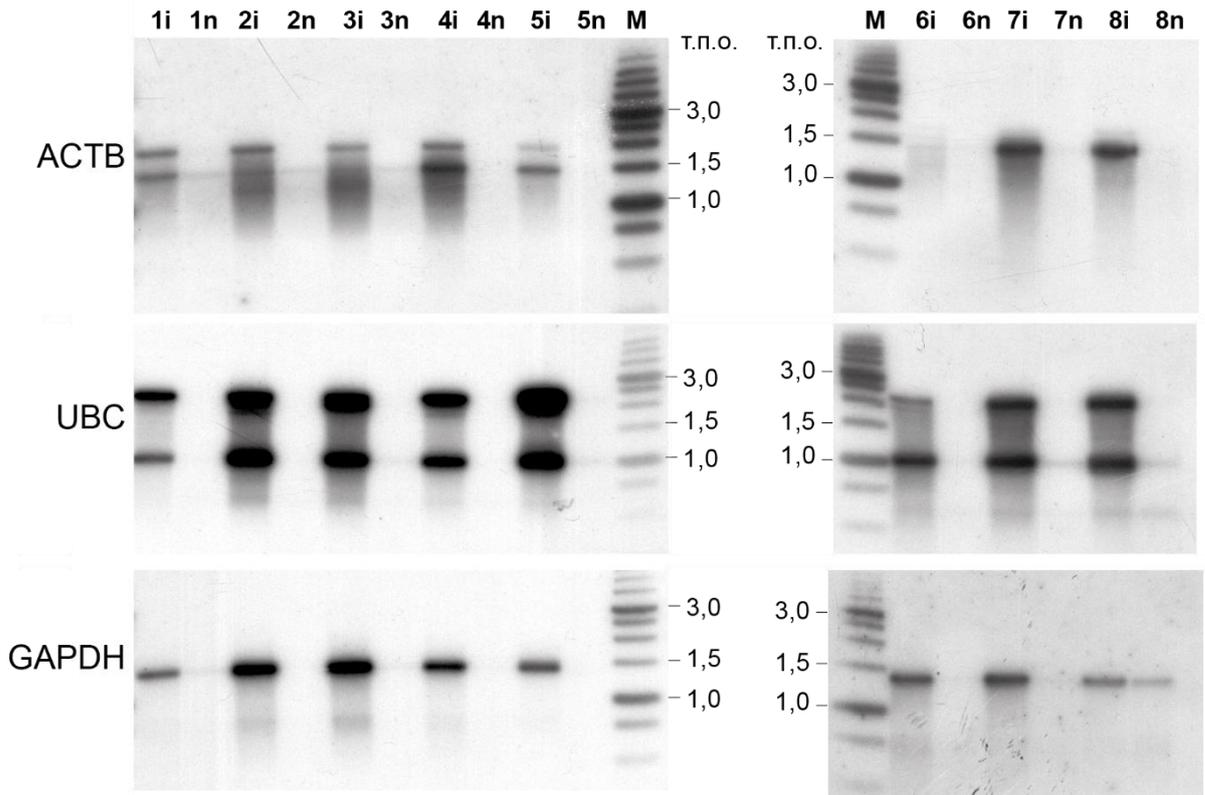
Результаты разделения на гель-электрофорезе продуктов нормализации амплифицированной кДНК, полученной на основе суммарной РНК из некоторых тканей человека, с использованием комбинаций адаптеров SMART Oligo II + CapNotI показаны на рисунке 3.39. Рядом показаны образцы ненормализованной кДНК. Во всех случаях после нормализации наблюдалось исчезновение полос, соответствующих высокопредставленным, характерным для данной ткани, транскриптами.



Результат электрофореза в агарозном геле образцов ненормализованной (а) и нормализованной (б) кДНК, приготовленной на основе суммарной РНК из плаценты (дорожка 1), эмбриональной печени (дорожка 2), мозга (дорожка 3), эмбриональных легких (дорожка 4), яичек (дорожка 5), почек (дорожка 6), тонкой кишки (дорожка 7) и желудка (дорожка 8). М — маркер длин ДНК, 1 kb ladder (Сибэнзим, Россия).

Рисунок 3.39. Результат нормализации образцов кДНК из тканей человека, полученный с использованием комбинации адаптеров SMART Oligo II + CapNotI

Уменьшение относительной концентрации высокопредставленных транскриптов было подтверждено с помощью виртуального Нозерн блот анализа (рисунок 3.40). Изменение концентраций различных транскриптов и сохранение репрезентативности библиотек кДНК было проверено на примере нормализованного и ненормализованного образцов кДНК из эмбриональной печени человека путем их гибридизации на макрочипе с 183 ПЦР-пробами. Результаты гибридизации показаны на рисунке 3.41. В обоих случаях было продемонстрировано выравнивание концентраций разнопредставленных транскриптов в образце нормализованной кДНК.



Образцы, использованные для гибридизации, описаны на рис. 2.39. ACTB (GenBank NM_001101), UBC (GenBank NM_021009), GAPDH (GenBank NM_002046). М — маркер длин ДНК, 1 kb ladder (Сибэнзим, Россия)

Рисунок 3.40. Виртуальный Нозерн блот анализ нормализованной (n) и ненормализованной (i) кДНК с P32-мечеными пробами, комплементарными высокопредставленным генам

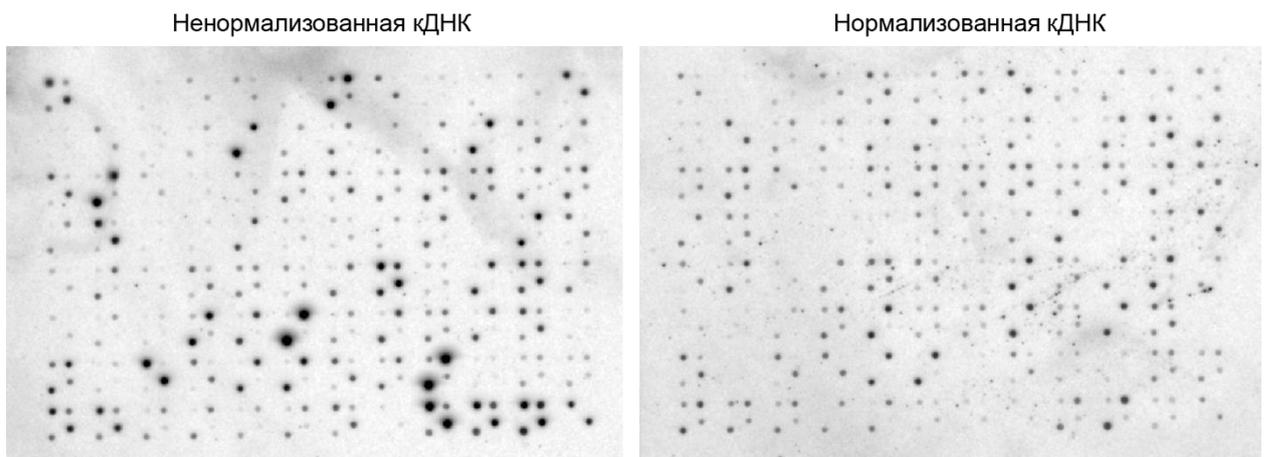
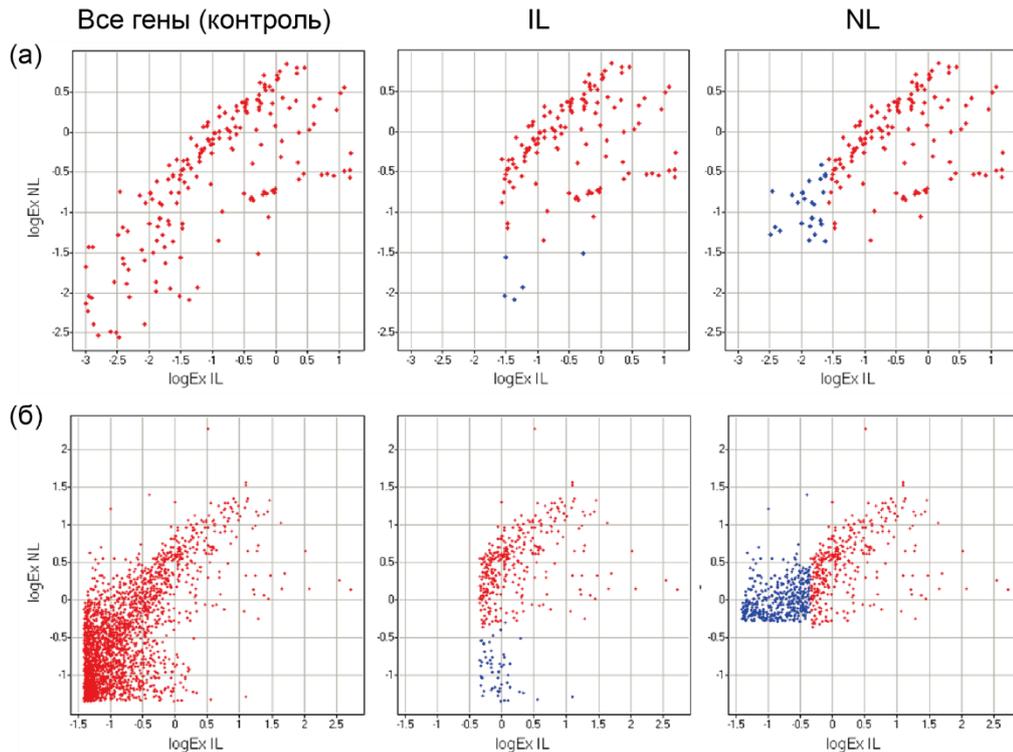


Рисунок 3.41. Результат гибридизации нормализованного и ненормализованного образцов кДНК из эмбриональной печени с макрочипом, содержащим 183 пробы

При гибридизации на микрочипе с 183 пробами в ненормализованной библиотеке кДНК было выявлено 130 последовательностей, комплементарных нанесенных на микрочип зондам. В нормализованной библиотеке – 151 последовательность. Таким образом концентрация для 26 транскриптов была повышена в ходе нормализации до детектируемого с помощью гибридизации уровня. При этом было обнаружено падение в ходе нормализации концентрации до недетектируемого с помощью гибридизации уровня для пяти транскриптов (рисунок 3.42а). Аналогичный эксперимент был также проведен с использованием макрочипа, содержащего 2300 олигонуклеотидных проб. В ходе этого эксперимента, в ненормализованной библиотеке кДНК было выявлено 388 последовательности, в нормализованной библиотеке – 771 последовательность (т.е. добавилось 382 последовательности, с низким уровнем экспрессии в ненормализованной библиотеке). Падение в ходе нормализации концентрации до недетектируемого уровня было обнаружено для 51 последовательности (рисунок 3.42б). Пять последовательностей, уровень которых в ходе нормализации упал ниже ожидаемого уровня, были проанализированы в образцах NL и PL с помощью ПЦР. Четыре из них были обнаружены в нормализованном образце после 30-32 циклов ПЦР, одна последовательность так и не была выявлена (данные не приведены).

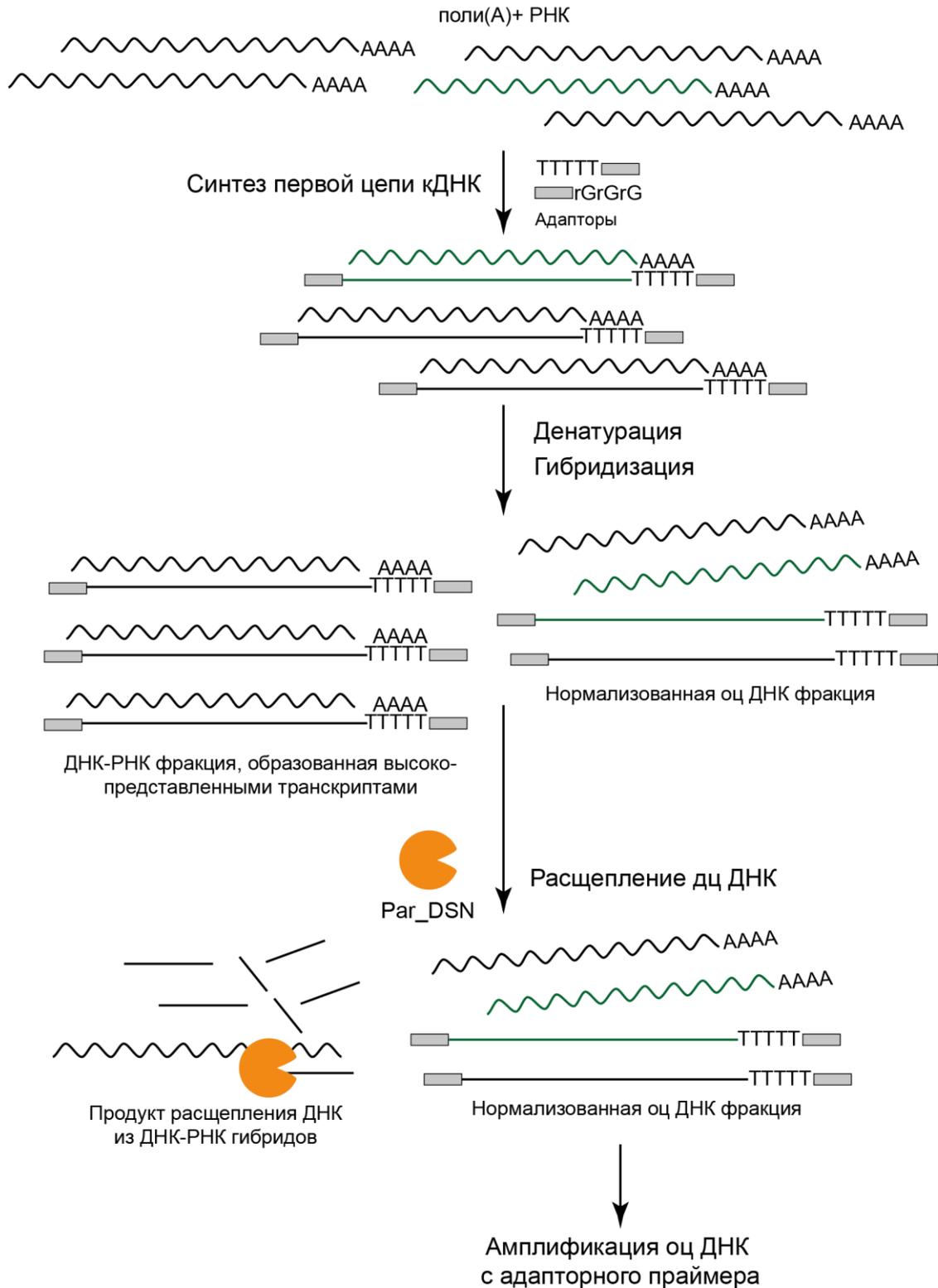
Чтобы избежать потери длинных транскриптов при амплификации нормализованной кДНК в протокол нормализации была добавлена дополнительная стадия второй амплификации, направленная на регуляцию средней длины продукта ПЦР. Для этого коллекция праймеров была дополнена укороченным на несколько нуклеотидов с 3'-конца вариантом праймера для амплификации (далее – праймер M2). Ранее нами было показано, что при наличии инвертированных концевых повторов на концах ДНК-матрицы в ходе ПЦР с праймером, комплементарным только части этого повтора, преимущество в амплификации получают более длинные молекулы [643].



(а) Гибридизация с микрочипом, содержащим 183 ПЦР-пробы; (б) с макрочипом, содержащим 2300 олигонуклеотидных проб. Контроль — гибридизация с образцом, содержащим все последовательности макрочипа. Данные об уровне экспрессии были преобразованы в \log_2 для обеих библиотек и нанесены на плот-диаграмму.

Рисунок 3.42. Дифференциальный анализ экспрессии с использованием данных массива гибридизации на макрочипах нормализованной (NL) и ненормализованной (IL) кДНК из эмбриональной печени

Также нами была разработана модификация метода, позволяющая нормализовать первую цепь кДНК против мРНК, выступающей в качестве драйвера. Схематически эта модификация ДСН-нормализации показана на рисунке 3.43. Поли(А)⁺ РНК используется для синтеза первой цепи кДНК, затем продукты реакции денатурируют и подвергают реассоциации. В ходе реассоциации первая цепь кДНК высокопредставленных транскриптов образует гибриды с комплементарной ей РНК, тогда как первая цепь кДНК редких транскриптов остается в оц форме. Далее в ходе инкубации реакционной смеси с Par_DSN происходит расщепление ДНК в ДНК-РНК гибридах с сохранением нормализованной оц-фракции кДНК.



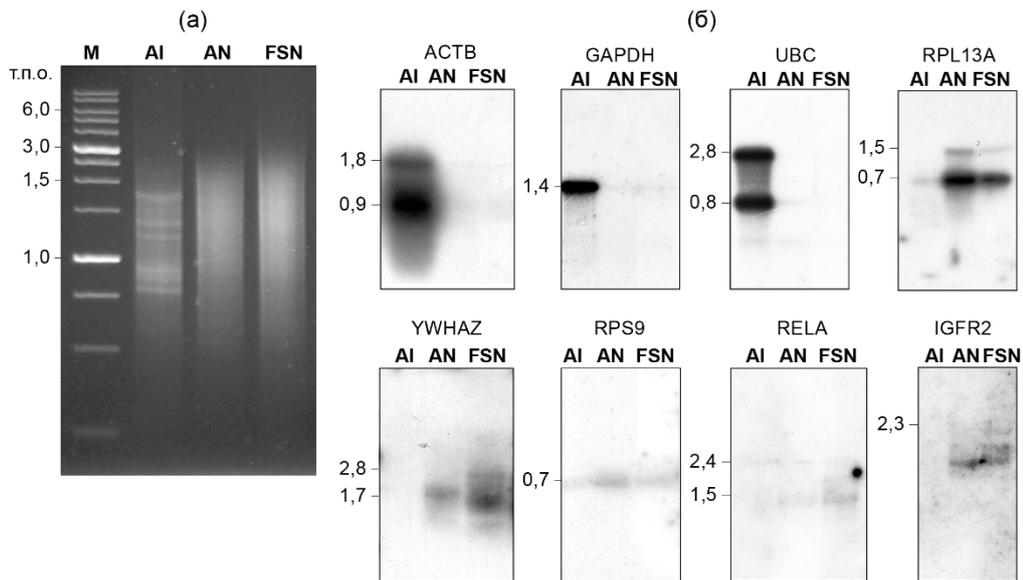
Прямоугольниками обозначены последовательности адаптеров.

Редкий транскрипт обозначен зеленой линией, высокопредставленный — черной линией.

Рисунок 3.43. Схема ДСН-нормализации первой цепи кДНК

Для апробации модификации метода ДСН-нормализации первой цепи кДНК была использована поли(А)+ РНК из скелетной мышцы человека. На ее основе были получены образцы амплифицированной нормализованной кДНК с использованием в качестве стартового материала для нормализации как первой цепи кДНК (образец FSN), так и амплифицированной кДНК (образец AN). Эти образцы сравнивали с образцом ненормализованной контрольной кДНК (образец AI).

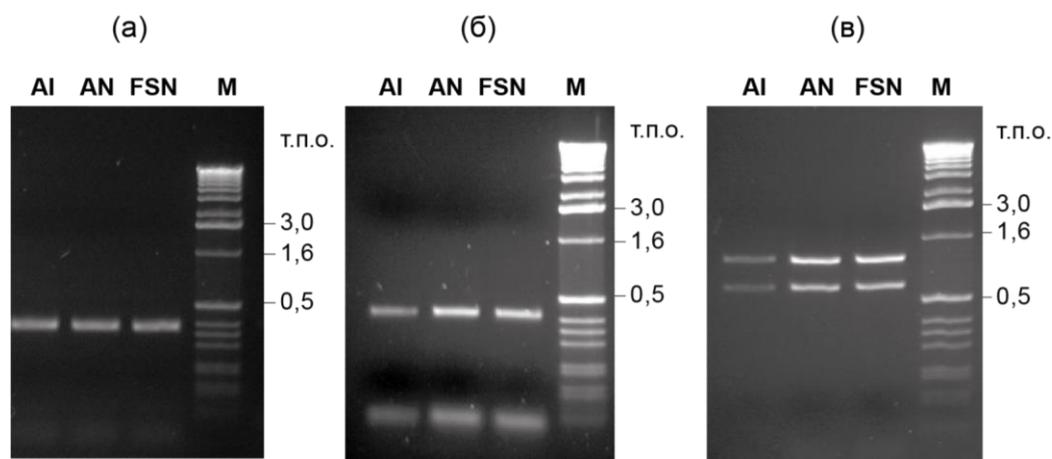
На рисунке 3.44 показан результат электрофореза этих образцов в агарозном геле и результаты виртуального Нозерн блот анализа указанных образцов с P^{32} -мечеными пробами, комплементарными транскриптам с различной представленностью. Согласно данным виртуального Нозерн-блот анализа в ходе нормализации наблюдается увеличение концентрации редких и уменьшения концентрации высокопредставленных транскриптов.



(а) Результат электрофореза в агарозном геле образцов ненормализованной (AI), нормализованной амплифицированной (AN) и нормализованной первой цепи (FSN) кДНК. М — маркер длин ДНК, 1 kb ladder (Сибэнзим, Россия). (б) Результат виртуального Нозерн-блот анализа этих образцов с P^{32} -мечеными пробами, комплементарными генам ACTB (GenBank NM_001101), GAPDH (GenBank NM_002046), UBC (GenBank NM_021009), YWHAZ (GenBank NM_003406), RPS9 (NM_001013), RPL13A (GenBank NM_012423), IGF2R (GenBank NM_001329128).

Рисунок 3.44. Результат нормализации образцов кДНК из скелетной мышцы человека

Дополнительно концентрации нескольких очень редких транскриптов были проанализированы в образцах AI, AN, FSN с помощью ПЦР. Типичный результат амплификации представлен на рисунке 3.45. Во всех проанализированных случаях было выявлено сохранение или небольшое увеличение концентрации исследованных фрагментов в нормализованных библиотеках.

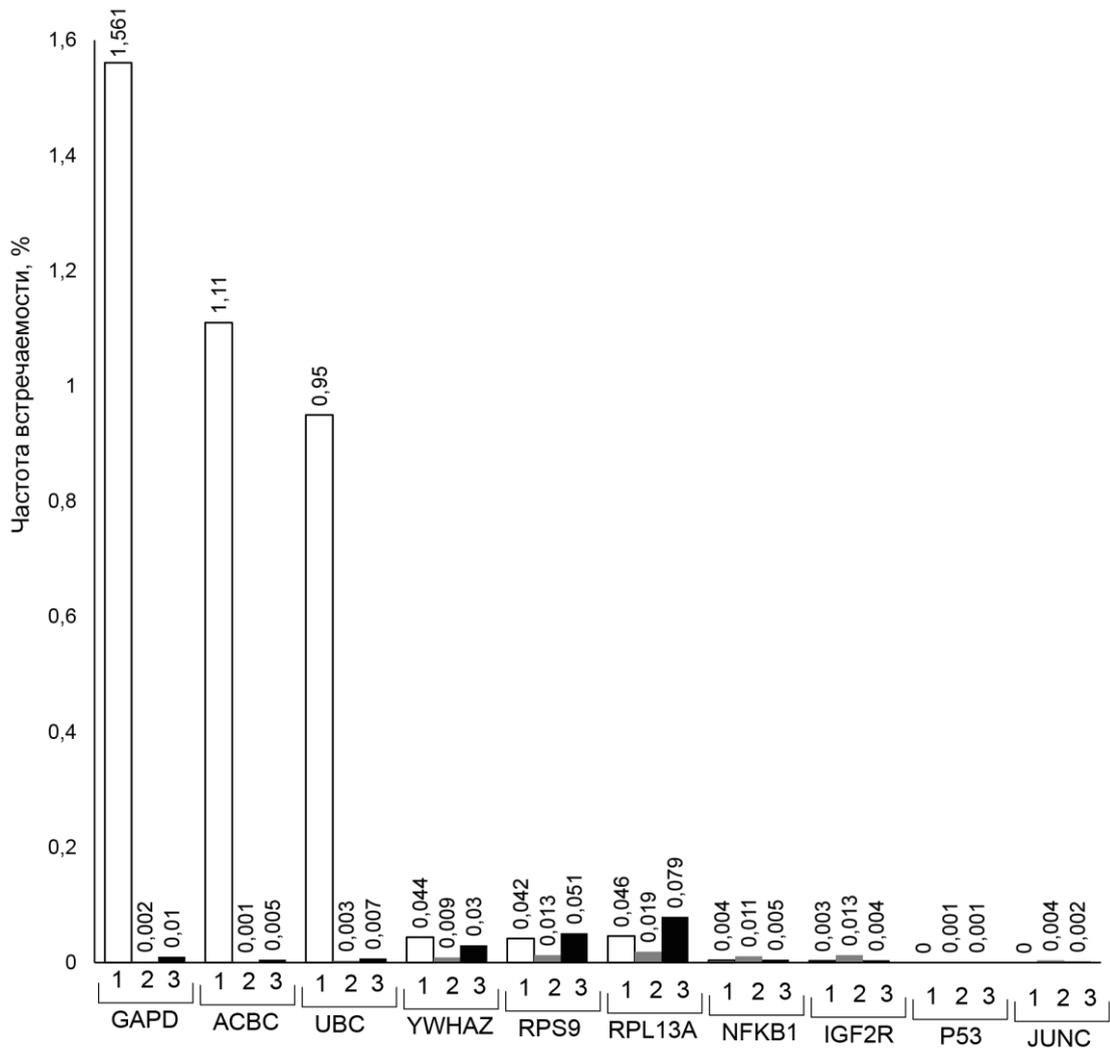


(а) IGF2R (GenBank NM_000876); (б) P53 (GenBank NM_000546), (в) JUNC (GenBank NM_002228). 27 циклов ПЦР. Описание образцов кДНК представлено на рис. 2.43. М — маркер длин ДНК, 1 kb ladder (Invitrogen, США).

Рисунок 3.45. Электрофорез в агарозном геле продуктов амплификации фрагментов редких генов в нормализованных и ненормализованном образцах кДНК скелетной мышцы человека

Образцы AI, AN и FSN были использованы для конструирования библиотек кДНК, с дальнейшей гибридизацией этих библиотек с набором радиоактивно-меченых проб к ряду генов человека. Было обнаружено заметное уменьшение числа клонов, несущих вставки с последовательностями высокопредставленных транскриптов (ACTB, GAPD, UBC) в библиотеках AN и FSN. Число клонов, содержащих вставки последовательностей генов средней представленности (RPL13A, RPS9, YWHAZ) осталось практически неизменным в AN библиотеке по сравнению с AI библиотекой и несколько снизилось в FSN библиотеке. В FSN библиотеке также наблюдалось 3-4-кратное увеличение, по сравнению с AI библиотекой, числа клонов, содержащих вставки последовательностей редких

транскриптов (IGF2R, JUNC, NFKB1). В AN библиотеке концентрация этих транскриптов не изменилась (рисунок 3.46).



Встречаемость транскриптов, определяемая как отношение положительных сигналов гибридизации колоний с радиоактивно мечеными пробами к общему числу колоний, указана над столбцами. 1 — библиотека AI, 2 — библиотека FSN, 3 — библиотека AN. В качестве проб использовали фрагменты генов, указанных под столбцами.

Рисунок 3.46. Частота встречаемости различных транскриптов в AI, FSN и AN библиотеках

В таблице 3-16 представлены данные анализа средней длины вставок в полученной библиотеке, подсчитанные на основании 106 случайно отобранных содержащих вставку клонов, показавшие сохранение среднего размера длины кДНК в ходе нормализации.

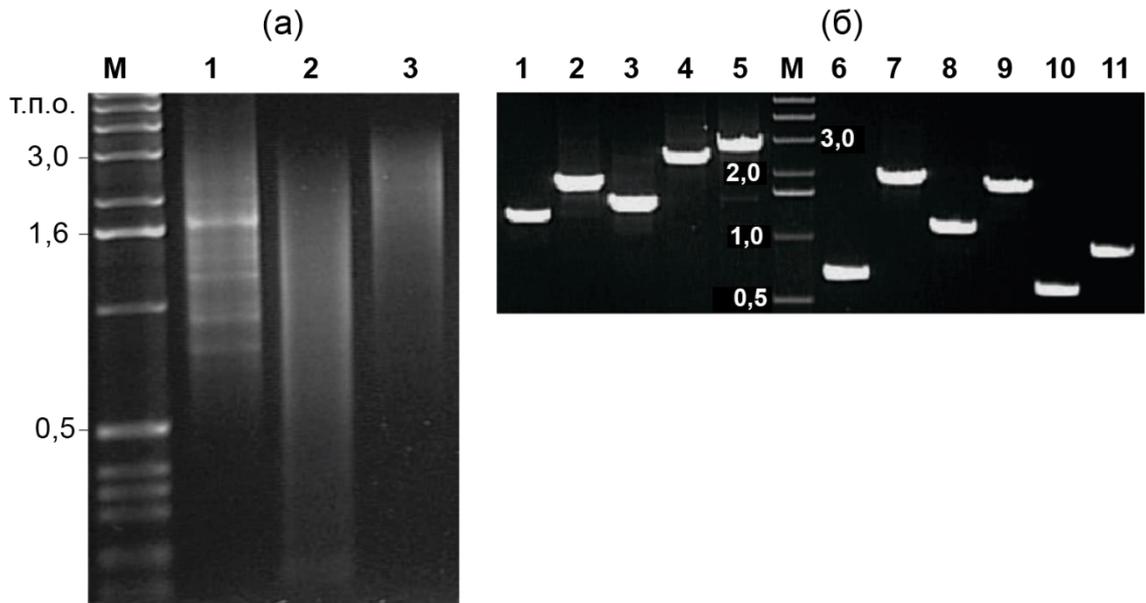
Таблица 3-16. Размер вставок в библиотеках кДНК AI, AN и FSN

Размер вставки, п.о.	Библиотека		
	AI	FSN	AN
> 2700	1	1	2
2400-2700	4	3	4
2100-2400	1	3	2
1800-2100	20	19	21
1500-1800	22	28	25
1200-1500	42	37	33
900-1200	6	8	9
600-900	4	5	7
300-600	5	2	3
< 300	1	0	0

Ранее было продемонстрировано, что если в образце ненормализованной кДНК присутствует ~10000 разных типов молекул, то при его нормализации концентрация каждого типа молекул кДНК будет стремиться к значению 0.01% [644]. В этом случае, при исчерпывающей нормализации концентрация транскриптов со средней представленностью должна снизиться в несколько раз или остаться неизменной. Концентрация высокопредставленных транскриптов снизится на один-два порядка. Концентрация же редких транскриптов должна вырасти, но не более чем в три раза. Результаты анализа библиотек кДНК, полученных с помощью ДСН-нормализации, хорошо согласуются с расчетными данными, подтверждая эффективность разработанного нами метода. Примечательно, что в библиотеке FSN количество клонов, соответствующих высокопредставленным транскриптам, оказалось меньше, чем количество клонов, соответствующих редким транскриптам. Подобная “сверхнормализация” может объясняться тем, что при гидролизе кДНК-РНК гибридов с помощью Par_DSN происходит расщепление только цепи кДНК, а высвобождаемые молекулы РНК

образуют новые гибриды с молекулами кДНК оц-фракции. Фракция молекул кДНК во вновь образованных ДНК/РНК гибридах также подвергается расщеплению. Исходя из полученных данных можно сделать вывод, что нормализация первой цепи кДНК является более эффективной по сравнению с нормализацией амплифицированной кДНК. Тем не менее, ДСН-нормализация амплифицированной кДНК также ведет к значительному снижению в образце высокопредставленных транскриптов, что позволяет значимо повысить эффективность поиска редких последовательностей в ситуации, когда доступна только суммарная РНК.

ДСН-нормализация была использована при исследовании aberrantly-spliced транскриптов в клеточной линии клеток рака молочной железы MCF-7. Хорошо известно, что геномы злокачественных опухолей накапливают большое количество геномных перестроек во время онкогенеза. Важный вопрос заключается в том, являются ли обнаруживаемые при исследовании опухолей геномные перестройки функционально значимыми или это просто результат нестабильности генома. Мы предположили, что обширные перестройки генома, присутствующие в клетках MCF-7, создают слитые транскрипты. Чтобы проверить эту гипотезу, нами была сконструирована полноразмерная нормализованная библиотека кДНК из тотальной РНК клеток MCF-7 (рисунок 3.47). 5089 клонов из этой библиотеки были секвенированы. Для 2700 клонов удалось картировать оба конца молекулы кДНК. При этом у 44 из них концы были локализованы на разных хромосомах, а 56 имели нарушения в ориентации при картировании на хромосоме.



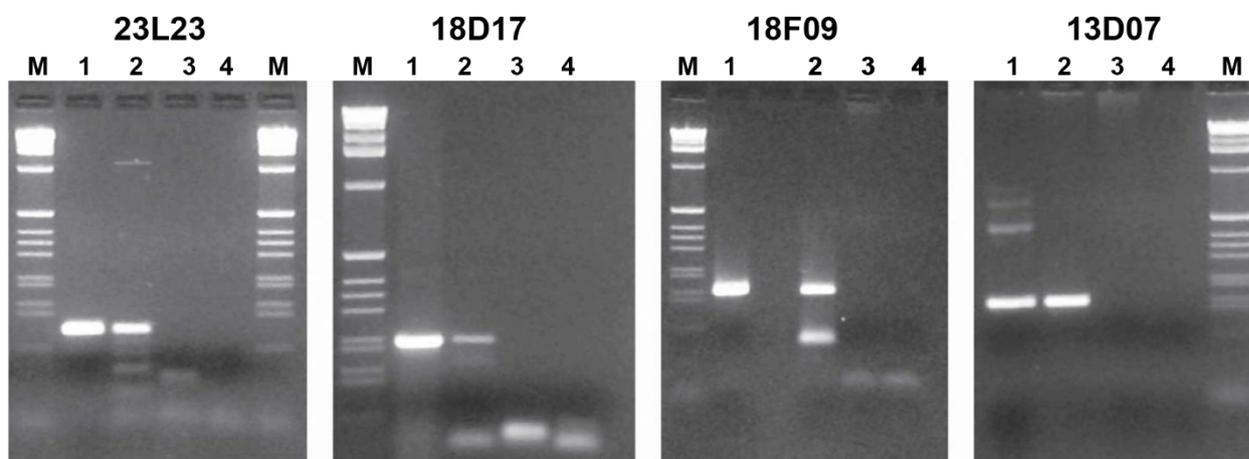
(а) Результат электрофореза в агарозном геле образцов ненормализованной (дорожка 1) и нормализованной кДНК после первой амплификации в ПЦР с праймером М1 (дорожка 2) и второй амплификации с праймером М2 (дорожка 3). М — маркер длин ДНК, 1 kb ladder (Invitrogen, США). (б) Результат амплификации случайных клонов, сделанный для оценки размера вставок нормализованной библиотеки.

Рисунок 3.47. Результат нормализации образцов кДНК, приготовленных на основе суммарной РНК из клеток MCF-7

Также было идентифицировано 24 новых транскрипта, которые содержали фрагменты из разных известных генов. Для четырех из этих транскриптов был подтвержден их аберрантный статус и экспрессия в независимых образцах клеток MCF-7 (рисунок 3.48, таблица 3-17). Выявленные транскрипты состояли из фрагментов, слитых из различных генов, причем в слиянии участвовали как экзоны, так и интроны и UTR. Для проверки трех из этих четырех транскриптов потребовалось использование «гнездовой» ПЦР с парами заглубленных праймеров, что указывает на то, что многие из обнаруженных слияний являются очень редкими.

Полученные нами данные показывают, что транскриптомы злокачественных опухолей человека могут содержать большое количество транскриптов слияния. Применение ДСН-нормализации сделало возможным при разумном объеме секвенирования выявление крайне редких транскриптов слияния.

Их идентификация подтвердила функциональное значение геномных перестроек в онкогенезе клеток МСF-7. Кроме того, такие опухолеспецифические транскрипты могут иметь определенное клиническое значение как возможные мишени терапевтического воздействия и маркеры опухолевого развития [645].



Дорожка 1 — клон ДНК из нормализованной библиотеки (положительный контроль); дорожка 2 — независимый препарат кДНК МСF-7; дорожка 3 — нормальная кДНК молочной железы взрослой женщины (Stratagene, США); дорожка 4 — отрицательный контроль. М — маркер длин ДНК, 1 kb ladder (Invitrogen, США).

Рисунок 3.48. Результаты ПЦР-валидации слитых транскриптов на независимых препаратах кДНК из МСF-7 с использованием клон-специфических праймеров, охватывающих точки слияния

Таблица 3-17. Характеристика транскриптов слияния, выявленных с помощью ДСН-нормализации

Название клона	Длины слитых фрагментов, п.о.	Координаты фрагмента на хромосоме	Ориентация фрагмента на хромосоме	Ген, к которому относится фрагмент
18D17	310	chr2:55448021	плюс	<i>MTIF2</i>
	553	chr2:43913586	плюс	<i>PLEKHH2</i>
13D07	526	chr6:100,979,123	минус	<i>HELIC1</i>
		chr6:101,033,719	минус	

	556	chr6:101,002,220	плюс	
18F09	630	chr7:128,003,608	минус	<i>CALU</i>
		chr7:128,003,709	плюс	
	547	chr7:128,005,309	минус	<i>CALU</i>
23L23	403	chr16:85,490,770	плюс	<i>KIAA0182</i>
	405	chr19:16,074,534	плюс	<i>AK023385</i>

Метод ДСН-нормализации кДНК был также использован для создания коллекции генетических маркеров, определяющих продолжительность жизни *Nothobranchius furzeri*. *N. furzeri* демонстрируют исключительно короткую продолжительность жизни, возрастное снижение поведенческих функций, экспрессию возрастных биомаркеров и восприимчивость к манипуляциям с продолжительностью жизни. Для поиска локусов количественного признака (QTL), которые коррелируют с продолжительностью жизни, были получены гибриды короткоживущей и долгоживущей популяций этой рыбы. Объединив данные о маркерах со значениями продолжительности жизни, был идентифицирован один высокозначимый QTL, который объясняет 11,3% дисперсии продолжительности жизни у популяций *N. furzeri*, и три локуса, которые могут быть так же ассоциированы с продолжительностью жизни. Для идентификации генов-кандидатов осуществлялось секвенирование геномов популяций *N. furzeri*, данные которого были дополнены секвенированием нормализованной библиотеки кДНК. В результате работы было выделено 342 транскрипта, относящихся к высокозначимому QTL. Было доказано, что определение продолжительности жизни у *N. furzeri* является полигенным, и что обнаружение генов-кандидатов может быть осуществлено с помощью межвидового анализа. Полученные первые результаты по выявлению локусов, контролирующих продолжительность жизни у *N. furzeri* иллюстрируют потенциал этого вида позвоночных в качестве генетической модели для исследования старения.

Таким образом, нами была предложена технология нормализации кДНК, основанная на кинетике гибридизации НК и использовании уникальных свойств Par_DSN. ДСН-нормализация отличается от аналогичных методов простотой исполнения и решает задачу быстрого и эффективного выделения нормализованной оц кДНК фракции из гибридизационной смеси. ДСН-нормализация применима как образцам амплифицированной кДНК, приготовленным из суммарной РНК, так и к первой цепи кДНК, приготовленной на основе поли(А)+ РНК, и адаптирована для работы с образцами, обогащенными полноразмерными последовательностями. Благодаря простоте исполнения и доступности технологии ДСН-нормализация быстро стала самым популярным методом нормализации кДНК перед высокопроизводительным секвенированием. Разработанный нами метод был использован более чем в двух сотнях работ. В рамках этих работ было подтверждено, что технология в целом сохраняет исходное разнообразие библиотеки кДНК и позволяет эффективно идентифицировать редкие и крайне редкие транскрипты, а также улучшает эффективность секвенирования и качество данных из образцов кДНК, полученной из деградированной РНК или из малых количеств РНК [646-650]. В то же время в работе Matvienko et al. (2013) было показано, что в связи с тем, что последовательности с высоким GC-содержанием более эффективно реассоциируют, чем АТ-богатые последовательности, они могут теряться в ходе ДСН-нормализации [651]. Так же Boon et al. (2018) показали, что некоторые адаптеры могут влиять на эффективность нормализации [17]. Следует заметить, что эти недостатки являются общими для всех методов нормализации, основанных на кинетике гибридизации НК.

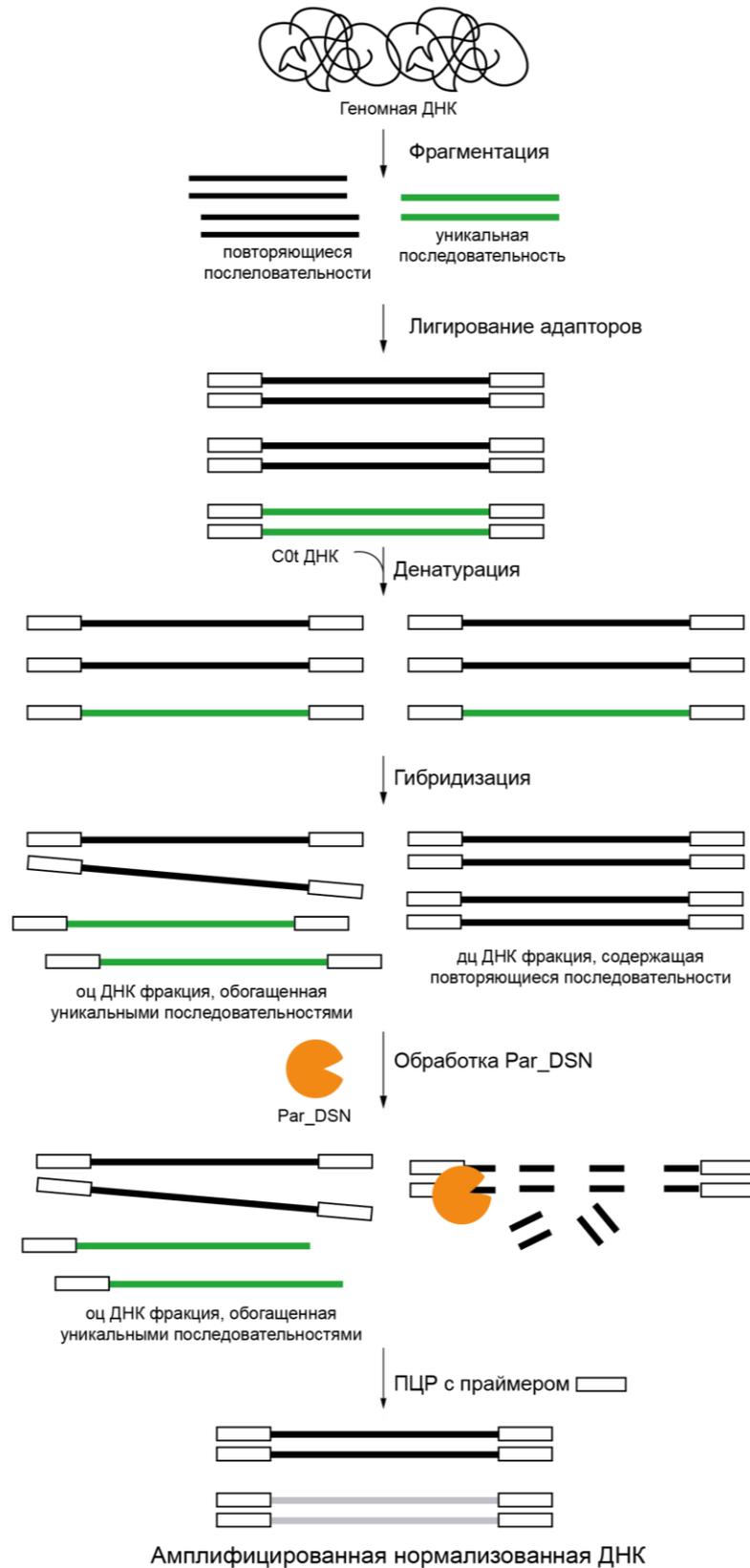
3.5. Нормализация геномной ДНК

В связи с бурным развитием технологий высокопроизводительного секвенирования полногеномное секвенирование стало популярным подходом для получения генетической информации об организмах. Однако значительную часть

геномной ДНК составляют высокогомологичные повторяющиеся элементы. Их присутствие не только увеличивает число прочтений, необходимое для секвенирования генома, но затрудняет, а иногда делает невозможным, сборку генома в ходе биоинформатической обработки.

Мы исследовали возможность применения технологии ДСН-нормализации для удаления высокогомологичных повторяющихся элементов из эукариотической геномной ДНК. Схема метода геномной нормализации показана на рисунке 3.49. Геномную ДНК фрагментируют и лигируют на концы адаптерные последовательности. После денатурации и последующей ренатурации фрагментированной геномной ДНК дц ДНК фракцию повторяющихся последовательностей подвергают гидролизу под действием *Par_DSN*. Оц фракцию геномной ДНК, обогащенную низкокопийными последовательностями, амплифицируют в ПЦР с праймером, комплементарным последовательности адаптера. Степень желаемой нормализации может зависеть от потребностей исследовательского коллектива. В частности, модулировать ее можно, добавляя в реакционную смесь для гибридизации фракции *C0t* (см. выше).

В модельном эксперименте мы использовали геномную ДНК человека. Хотя содержание повторяющихся элементов в геноме человека значительно ниже, чем в геномах высших растений, геномная ДНК человека представляется более удобной для апробации метода, так как для нее известен точный состав и частота встречаемости повторяющихся элементов [652].



Прямоугольниками обозначены последовательности.

Редкий транскрипт обозначен зеленой линией, высокопредставленный — черной линией.

Рисунок 3.49. Схема ДСН-нормализации геномной ДНК

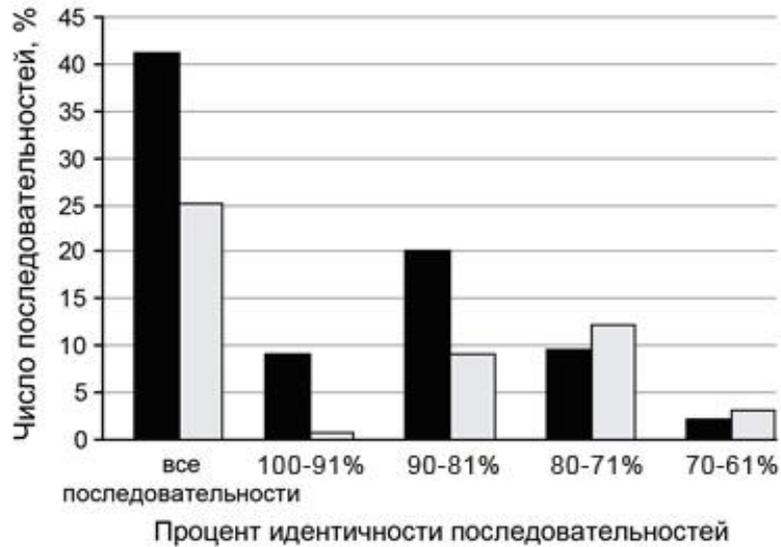
В модельном эксперименте нормализацию осуществляли с добавлением фракции C0t-1 ДНК человека, которая представляет собой последовательности, содержащие высокую концентрацию повторов LINE и SINE, и традиционно используется для блокирования неспецифического связывания богатой повторами ДНК с зондами гибридизации. При проведении гибридизации важно учитывать, что оптимальные результаты могут быть достигнуты и с другими повторяющимися фракциями ДНК от C0t 2 до C0t 10 [653].

Нормализованный и контрольный (ненормализованный) образцы ДНК секвенировали с использованием автоматизированной системы секвенирования 454 GS FLX (Roche). Было получено 29240 прочтений для нормализованного и 31789 для ненормализованного образцов, состоящих из 6269460 и 6643277 нуклеотидов соответственно. В таблице 3-18 приведены данные о представленности последовательностей различных семейств повторяющихся элементов в ненормализованном и нормализованном образцах геномной ДНК человека. Анализ уровня встречаемости различных повторяющихся элементов по результатам секвенирования показал, что в результате нормализации происходит снижение процентного содержания повторяющихся элементов примерно с 40% до 25%. При этом в образце нормализованной геномной ДНК можно отметить существенное снижение уровней Alu, LINE L1P, ERV-K и ERV1 повторов, а также сателлитных последовательностей. В то же время концентрация других повторов не претерпевает существенных изменений. Анализ репрезентативности повторов с различной степенью дивергенции выявил, что в ненормализованном образце ДНК примерно 10% последовательностей имеют 100-91% идентичности; 20% – 90-71%, а остальные последовательности – менее 70% идентичности. В результате нормализации произошло 15-кратное снижение концентрации повторяющихся элементов с высокой степенью нуклеотидной идентичности (100-91%) и 2-кратное – элементов, имеющих 81-90% идентичности. Концентрация последовательностей, идентичность которых ниже 80% практически не изменилась (рис. 3.50). Иными словами, наблюдается корреляция эффективности нормализации со степенью

дивергенции последовательностей семейства: чем ниже дивергенция последовательностей, тем эффективнее они нормализуются.

Таблица 3-18. Выявление последовательностей повторяющихся элементов различных семейств в геномной ДНК человека до и после нормализации

Семейства повторяющихся элементов	Ненормализованная ДНК			Нормализованная ДНК		
	п.н.	%	Ср. див., %	п.н.	%	Ср. див., %
Общее число нуклеотидов	6643277	100	-	6269460	100	-
LINE L1P	371253	5.6	9	69670	1.1	15
ERV-K	20597	0.3	8	6575	0.1	12
SSR (simple sequence repeats)	68367	1.0	10	18499	0.3	11
Alu	631112	9.5	12	6572	0.1	16
Последовательности-сателлиты	293827	4.4	15	7915	0.1	20
ERV1	165461	2.5	15	105155	1.7	17
ERV-L	323437	4.9	19	245039	3.9	22
LINE L1M/HAL	496398	7.5	19	605965	9.7	21
ДНК транспозоны	161304	2.4	19	170331	2.7	21
MIR	118927	1.8	27	142060	2.3	27
LINE L2	104535	1.6	28	147241	2.3	28
LINE CR1/L3	10226	0.1	29	13318	0.2	29
LINE L4-L5	3508	0.05	33	3701	0.05	30
Ср. див. — средняя дивергенция						



Черные столбцы — ненормализованная ДНК, серые — нормализованная ДНК

Рисунок 3.50. Представленность повторяющихся элементов в нормализованном и ненормализованном образцах геномной ДНК человека

Для оценки вклада $C0t-1$ -фракции в эффективность нормализации было проведено сравнение эффективности нормализации с добавлением и без добавления $C0t-1$. Нормализованную и ненормализованную ДНК без добавления $C0t-1$ использовали для конструирования библиотек ДНК и секвенирования методом Сэнгера 300 случайно выбранных клонов из каждой библиотеки. Полученные данные сравнивали с данными нормализации геномной ДНК человека в присутствии $C0t-1$.

Было показано, что в отсутствие $C0t-1$ наблюдается только 7-кратное снижение концентрации высокопредставленных (100-91% идентичности) повторов. Для остальных типов повторов значимой разницы в эффективности нормализации с или без $C0t-1$ ДНК выявлено не было (таблица 3-19). Таким образом, добавление $C0t-1$ фракции повышает эффективность нормализации, по крайней мере, для высокомолекулярных повторов.

Таблица 3-19. Эффективность нормализации повторяющихся элементов генома в присутствии и отсутствии C0t-1 фракции

Образец	Секвенирование методом 454				Секвенирование методом Сэнгера			
	До нормализации (контроль)		Нормализация в присутствии C0t-1ДНК		До нормализации		Нормализация в отсутствии C0t-1 ДНК	
	п.о.	%	п.о.	%	п.о.	%	п.о.	%
Всего последовательностей	6643277	100	6269460	100	118930	100	123890	100
Повторяющихся элементов:	2768952	41.68	1573235	25.09	49395	41.5	34048	27.5
С идентичностью 91%–100%	601509	9.05	39846	0.63	8658	7.2	1254	1.0
С идентичностью 81%–90%	1333536	20.07	563969	8.99	19189	16.1	8759	7.1
С идентичностью 71%–80%	634379	9.55	758635	12.1	15527	13.1	17782	14.3
С идентичностью 61%–70%	147972	2.23	194503	3.1	6021	5.1	6253	5.0
С идентичностью менее 60%	15446	0.23	16282	0.26	-	-	-	-

Для проверки сохранности в ходе нормализации уникальных последовательностей ДНК фрагменты 11 различных генов были проанализированы с помощью ПЦР в режиме реального времени в образцах геномной ДНК до и после нормализации. Для всех 11 фрагментов не было выявлено значимых изменений в уровнях представленности (таблица 3-20).

Таблица 3-20. Анализ изменений в концентрации уникальных последовательностей ДНК в ходе геномной нормализации

Ген	Идентификационный номер GenBank	Число циклов ПЦР для амплификации фрагмента	
		ДНК до нормализации	ДНК после нормализации
<i>DKC1</i>	NM_001363	29	28
<i>RPS19</i>	NM_001022	28	28
<i>TERC</i>	NR_001566	32	33
<i>GJB2</i>	NM_004004	29	29
<i>ELA2</i>	EU617980	33	34
<i>MYL2</i>	NM_000432	32	32
<i>MYBPC3</i>	NM_000256	32	32
<i>TNNC1</i>	NM_003280	32	33
<i>MYLK2</i>	NM_033118	30	34
<i>LAMP2</i>	NM_002294	30	30

На основании полученных данных был сделан вывод, что ДСН-нормализация может быть полезной для снижения в образце геномной ДНК эволюционно молодых геномных повторяющихся элементов с низкой степенью дивергенции. Известно, что большое количество высококонсервативных повторяющихся элементов содержат геномы многих высших растений. Так,

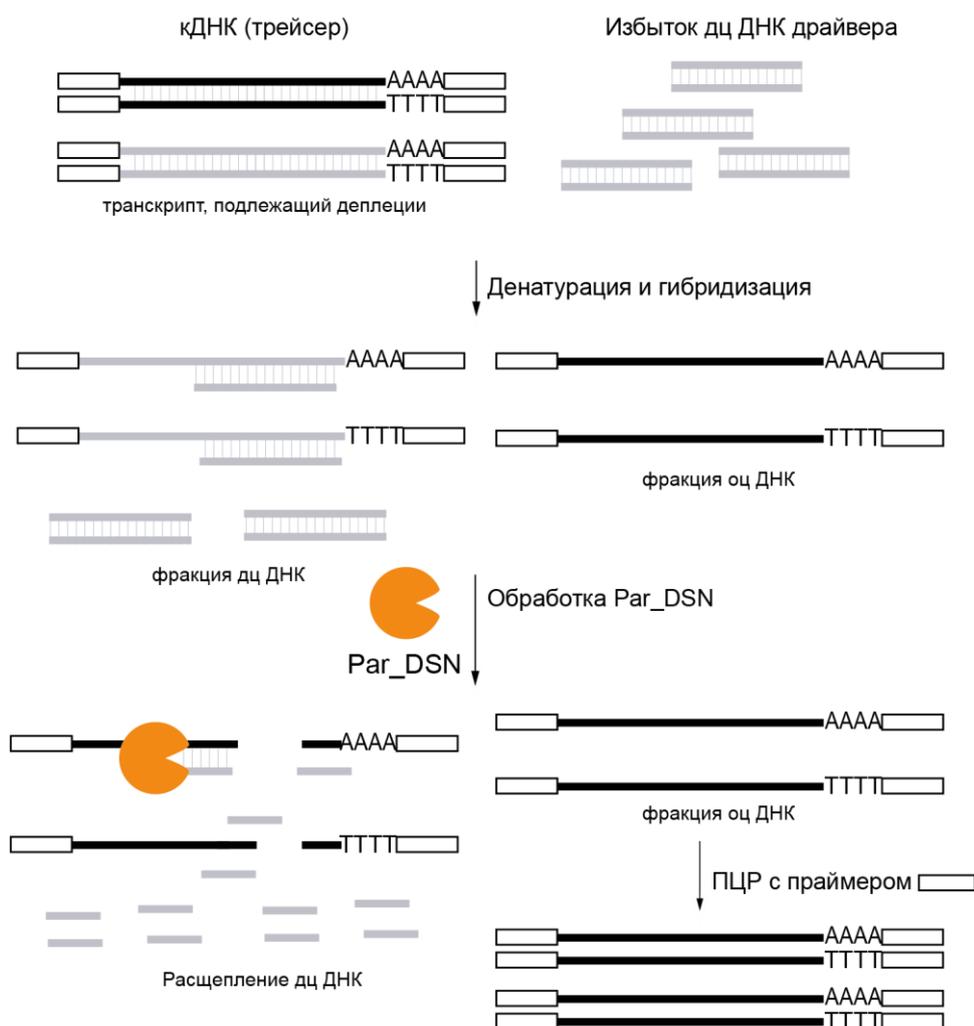
например, более 80% геномной ДНК кукурузы и пшеницы состоят из повторяющихся последовательностей ДНК, в основном – ретротранспозонов, имеющих более 95% идентичности [20-22, 654]. 91-95% идентичности имеют небольшие неавтономные САСТА (SНАС) элементы пшеницы [655]. С другой стороны, ДСН-нормализация геномной ДНК не оказывает значимого влияния на повторы с высокой степенью дивергенции. Возможно, это связано с жесткими условиями гибридизации. Смягчение условий гибридизации (например, за счет снижения температуры реассоциации и/или повышения катионной концентрации в буфере) может изменить ситуацию, хотя и с возможной, из-за неспецифической гибридизации, частичной потерей уникальных последовательностей.

Метод нормализации геномной ДНК с помощью Par_DSN был апробирован в ряде лабораторий. Так, Conner et al. (2016) [656] на примере картирования апоспория-специфической геномной области *Pennisetum squamulatum* (8X) и *Cenchrus ciliaris* (4X) продемонстрировали эффективность ДСН-нормализации геномной ДНК для целевого картирования физически больших геномных локусов, богатых повторяющимися последовательностями. Matvienko et al. (2013) [651] использовали геномную ДСН-нормализацию перед высокопроизводительным секвенированием для ДНК из арабидопсиса и из салата. В этой работе было показано, что при 22-часовой гибридизации происходит 12-кратное падение концентраций наиболее распространенных повторяющихся последовательностей, и 10-15-кратное падение концентраций умеренно повторяющиеся последовательностей, таких как *Athila*, при сохранении редких повторов и уникальных последовательностей. Повторная гибридизация в течение 46 и 70 часов приводила к дальнейшему падению представленности повторяющихся элементов, имеющих высокий процент идентичности (там же). На основании данных секвенирования нормализованной геномной ДНК авторы рекомендовали ДСН-нормализацию, как протокол, обеспечивающий эффективный доступ к низкокопийной части геномов, снижающий количество прочтений и, следовательно, стоимость, необходимую для секвенирования, а также упрощающий сборку наиболее информативных компонентов генома (там же). В

работе Ichida & Abe (2019) [657] показано, что ДСН-нормализация геномной ДНК позволила снизить относительное содержание 18S и 25S рибосомальной ДНК в образцах геномной ДНК до 1,15% и 3,54% в рисе и 1,70–1,95% и 14,71–20,01% в трех сортах пшеницы. Анализ полногеномного секвенирования диплоидного сорта пшеницы KU104-1 показал, что обработка Par_DSN резко снижает количество повторяющихся элементов генома, таких как ретротранспозоны Tu1-Copia и Tu3-Gypsy, а также ДНК-транспозоны. В то же время количество прочтений, полученных для малокопийных последовательностей и генов, кодирующих белки, увеличилось более чем на 50% [635]. Gagic et al. (2015) [658] рекомендовали ДСН-нормализацию для подготовки метагеномов бактерий для секвенирования. Авторы использовали метагеном из пяти бактериальных штаммов и сравнили предложенный нами метод с методом нормализации, основанном на выделении нормализованной оц ДНК фракции на гидроксипатитовых колонках. В заключении к работе авторы указали, что ДСН-нормализация привела к значительному увеличению представленности последовательностей самых редких членов сообщества. Кроме того, результатом нормализации стала возможность идентифицировать редкие гены, которые не могли быть обнаружены до нормализации с использованием той же глубины секвенирования (там же).

3.6. Селективное удаление нецелевых последовательностей из образцов кДНК. Удаление рибосомальной РНК перед полномасштабным секвенированием транскриптомов прокариот

Уникальные свойства Par_DSN легли в основу технологии удаления выбранных транскриптов из библиотек кДНК. Метод получил название ДСН-деплеция кДНК. Предложенная технология проста в исполнении и может быть применена как к фрагментированной кДНК, так и к кДНК, обогащенной полноразмерными последовательностями. Схема технологии показана на рис. 3.51.



Прямоугольниками обозначены последовательности адаптеров.

Рисунок 3.51. Схема ДСН-деплеции кДНК

Как и большинство известных на сегодняшний день методов вычитающей гибридизации и нормализации кДНК, ДСН-деплеция основана на кинетике гибридизации НК [439]. Глубину нормализации образца можно варьировать, меняя время гибридизации. Гидролиз дц ДНК фракции, происходящий за счет нуклеазной активности Par_DSN, является ключевым этапом ДСН-деплеции. Так же, как и в случае с ДСН-нормализацией, обработка Par_DSN происходит при высоких температурах, что позволяет избежать потери транскриптов из-за формирования вторичных структур и неспецифической гибридизации. В качестве трейсера используют образец дц кДНК, фланкированный последовательностями адаптера. Трейсер смешивают с избытком ДНК драйвера, представляющего собой

фрагменты генов, последовательности которых нужно удалить из образца трейсера, денатурируют и оставляют реассоциировать. Во время гибридизации, молекулы драйвера формируют гибриды с комплементарными им последовательностями. Полученную фракцию дц кДНК гидролизуют, инкубируя с дуплекс-специфичной нуклеазой краба. Оставшуюся в смеси оц кДНК амплифицируют при помощи ПЦР и используют для клонирования. Процесс ДСН-деплеции сопровождается частичной нормализацией кДНК, поскольку в ходе гибридизации наиболее часто встречающиеся молекулы кДНК успевают перейти в двухцепочечную форму и затем подвергаются гидролизу.

Для ДСН-деплеции может быть использована кДНК, полученная практически любым из известных методов. Единственное предъявляемое требование – наличие на концах молекул трейсера известных адаптерных последовательностей для последующей амплификации.

Как было ранее продемонстрировано в экспериментах по анализу ОНП, минимальным субстратом для Par_DSN является ДНК-ДНК дуплекс длиной 10 п.о. Таким образом, в качестве драйвера для ДСН-деплеции должны быть использованы фрагменты генов, образующие с подлежащими удалению молекулами трейсера, дуплексы длиной более 10 п.о. Учитывая, что гибридизация ведется при температуре в 68-70°C, длина таких дуплексов должна превышать 30 п.о. В модельных экспериментах в качестве драйвера нами были апробированы как амплифицированные фрагменты ДНК длиной примерно 100 п.о., так и синтетические 47- и 60-звеньевые олигонуклеотиды.

В первом модельном эксперименте в качестве трейсера использовали кДНК плаценты человека, из которой удаляли последовательности генов семейства фосфодиэстераз. Фосфодиэстеразы млекопитающих формируют суперсемейство, состоящее из 10 семейств, отличающихся по аминокислотным последовательностям, субстратной специфичности, чувствительности к ингибиторам, путям регуляции и тканевому распределению [659]. С-концевой фрагмент кодирующей последовательности генов этого суперсемейства содержит высококонсервативный регион длиной ~250 аминокислотных остатков,

формирующий каталитический домен [660]. Для приготовления драйвера фрагменты, соответствующие последовательности каталитических доменов девяти генов фосфодиэстераз человека, были амплифицированы, смешаны в равных пропорциях, подвергнуты очистке на колонке, и использованы в качестве драйвера.

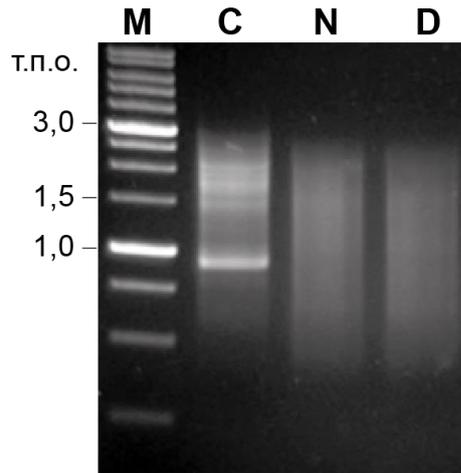
Драйвер (примерно 100-кратный избыток для каждого удаляемого гена) был добавлен к амплифицированной кДНК плаценты (трейсер). После этого образец денатурировали и гибридизовали при 68°C в течение 1 ч. Далее, гибридизационную смесь инкубировали с Par_DSN для расщепления дц ДНК фракции. Оставшуюся оц кДНК амплифицировали с использованием праймеров, комплементарных адаптерным последовательностям, фланкирующим кДНК. Для оценки вклада нормализации в процесс ДСН-деплеции использовали кДНК трейсера, подвергнутую аналогичным процедурам в отсутствие драйвера. В результате была получена коллекция образцов, характеристика которых представлена в таблице 3-21.

Таблица 3-21. Характеристика образцов, полученных в модельном эксперименте при тестировании метода ДСН-деплеции кДНК

Образец	Описание	Драйвер	Обработка Par_DSN после гибридизации
C	Контрольный образец	нет	нет
N	Контрольный частично нормализованный образец	нет	да
D	Экспериментальный образец, подвергшийся ДСН-деплеции	Амплифицированные фрагменты генов фосфодиэстераз	да

Результат анализа этих образцов с помощью электрофореза в агарозном геле показан на рис. 3.52. Сравнение электрофреграмм образцов выявило исчезновение

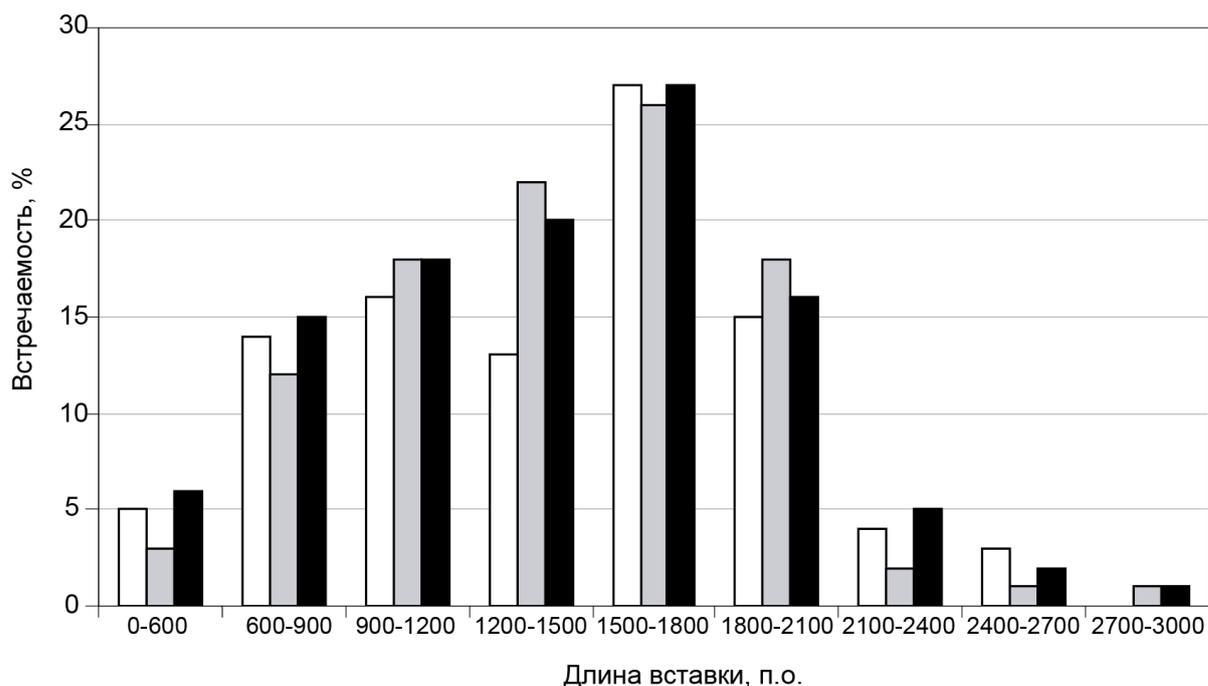
полос, соответствующих наиболее высокопредставленным транскриптам, в образцах N и D при сохранении средней длины кДНК на уровне контрольного образца. Сохранность средней длины вставок была подтверждена при анализе случайно выбранных клонов из библиотек кДНК, сконструированных из полученных образцов (рисунок 3.53).



Характеристика образцов, названия которых указаны над дорожками, приведена в таблице 3-21.

M — маркер длин ДНК, 1 kb ladder (Сибэнзим, Россия)

Рисунок 3.52. Результат электрофореза в агарозном геле контрольных (C, N) и экспериментального (D) образцов, полученных в модельном эксперименте при тестировании метода ДСН-деплеции кДНК



Белые столбцы — образец С, серые столбцы — образец N, черные — образец D.

Характеристика образцов приведена в таблице 3-21.

Рисунок 3.53. Частота встречаемости вставок различной длины в библиотеках кДНК, полученных в модельном эксперименте при тестировании метода ДСН-деплеции

Эффективность ДСН-деплеции была оценена с помощью количественной ПЦР с использованием TaqMan-проб, комплементарных 5'-последовательностям кодирующих областей генов фосфодиэстераз, лежащих выше консервативного каталитического домена. В качестве контроля использовали пробы к некоторым высокопредставленным транскриптам и к гену фосфодиэстеразы *PDE4D*, который не подвергся деплеции. Результаты анализа суммированы в таблице 3-22 и на рисунке 3.54.

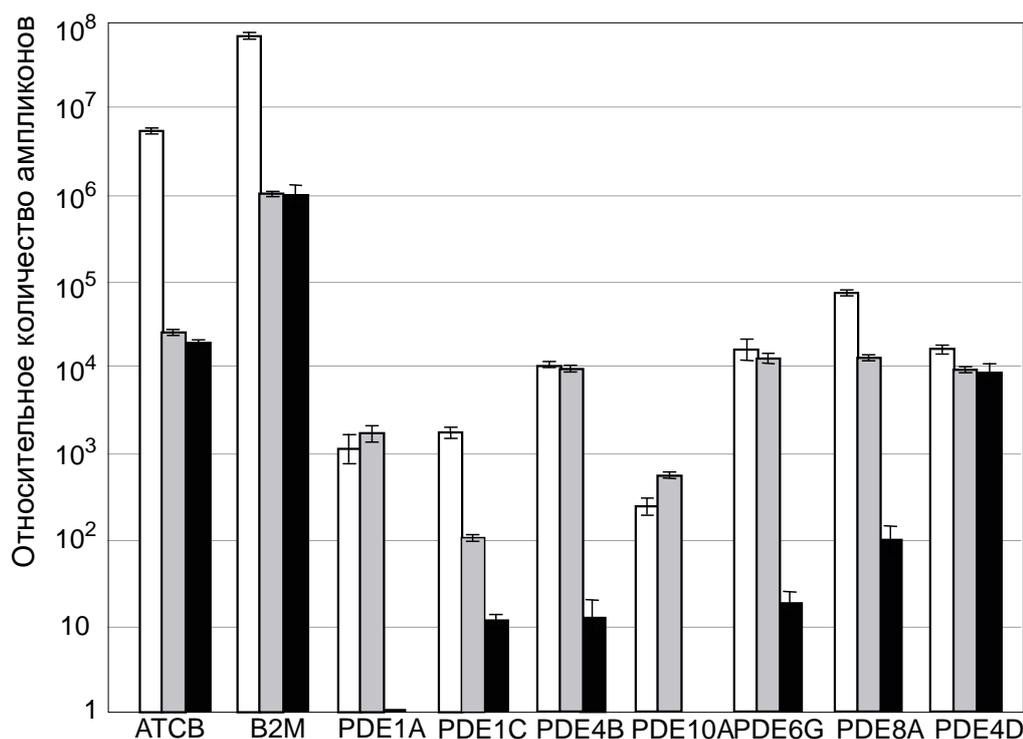
Таблица 3-22. Результаты амплификации в ходе РВ-ПЦР ряда транскриптов из образцов кДНК, полученных при тестировании метода ДСН-деплеции кДНК

Образец кДНК	Анализируемый транскрипт	Среднее C_t	σ	Относительное число ампликонов
С1	АСТВ	17,519	0,06879	5854035
С2	АСТВ	17,738	0,096542	5031301

N1	ACTB	25,377	0,505075	25241
N2	ACTB	25,390	0,486819	25006
D1	ACTB	25,910	0,057079	17439
D2	ACTB	25,696	0,036664	20227
C1	B2M	14,133	0,130697	61220012
C2	B2M	13,828	0,056115	75593036
N1	B2M	20,159	0,041701	939478
N2	B2M	19,966	0,046593	1073395
D1	B2M	19,760	0,150152	1238790
D2	B2M	20,742	0,048856	626845
C1	PDE1A	30,542	0,082582	704
C2	PDE1A	29,381	0,058232	1573
N1	PDE1A	28,979	0,02385	2078
N2	PDE1A	29,616	0,150614	1337
D1	PDE1A	39,494	0,383491	1
D2	PDE1A	39,453	0,523913	1
C1	PDE1C	29,445	0,048781	1504
C2	PDE1C	29,131	0,142787	1871
N1	PDE1C	33,424	0,361118	95
N2	PDE1C	33,202	0,700479	111
D1	PDE1C	36,808	1,367433	9
D2	PDE1C	36,374	1,023931	12
C1	PDE4B	26,614	0,10014	10705
C2	PDE4B	26,708	0,072652	10028
N1	PDE4B	26,605	0,027398	10772
N2	PDE4B	26,869	0,106351	8968
D1	PDE4B	35,816	0,089052	18
D2	PDE4B	37,487	0,490238	6
C1	PDE10A	31,754	0,229687	304
C2	PDE10A	32,556	0,097116	174
N1	PDE10A	30,802	0,026285	587
N2	PDE10A	31,058	0,05684	492
D1	PDE10A	Не выявлен	-	0
D2	PDE10A	Не выявлен	-	0
C1	PDE6G	26,521	0,989272	11420
C2	PDE6G	25,659	0,186557	20760
N1	PDE6G	26,244	0,147972	13837

N2	PDE6G	26,652	0,092819	10427
D1	PDE6G	34,949	0,616288	33
D2	PDE6G	36,417	0,671238	12
C1	PDE8A	24,010	0,105934	65083
C2	PDE8A	23,743	0,109353	78315
N1	PDE8A	26,421	0,177918	12239
N2	PDE8A	26,436	0,125855	12113
D1	PDE8A	33,568	0,100943	86
D2	PDE8A	32,647	0,424293	163
C1	PDE4D	26,474	0,982164	11794
C2	PDE4D	25,386	0,109436	25084
N1	PDE4D	26,649	0,494068	10448
N2	PDE4D	26,561	0,635244	11108
D1	PDE4D	26,639	0,50079	10523
D2	PDE4D	26,539	0,8074	11276
Приведены данные двух параллельных экспериментов. σ — стандартное отклонение для C_t , определенное по трем повторным измерениям.				

В образце, истощенном по последовательностям фосфодиэстераз наблюдалось, по крайней мере, 100-кратное снижение концентрации всех девяти транскриптов по сравнению как с контрольным, так и с частично нормализованным образцами. При этом концентрация молекул кДНК *PDE4D* оставалась неизменной во всех образцах. Частичная нормализация снижала концентрацию высокопредставленных транскриптов в опытных образцах, однако не влияла на концентрацию членов семейства фосфодиэстераз, относящихся к редким и среднепредставленным транскриптам.



Белые столбцы — образец С, серые столбцы — образец N, черные — образец D.

Характеристика образцов приведена в таблице 3-22.

Рисунок 3.54. Влияние ДСН-деплеции кДНК на представленность различных транскриптов по результатам количественной ПЦР

Влияние ДСН-деплеции на концентрацию высокомолекулярных транскриптов, подлежащих удалению, было исследовано в следующем эксперименте. Ранее нами было показано, что коралловый полип *Zoanthus* экспрессирует два гомологичных флуоресцентных белка различных цветов. Эти белки имеют 87%-гомологию на нуклеотидном уровне. В модельном эксперименте, кДНК *Zoanthus* была разделена на 2 образца: первый использовали для удаления транскриптов зеленого флуоресцентного белка (zGFP, GenBank ID 1080018), другой – красного флуоресцентного белка (zRFP, GenBank ID 1080019). В качестве драйвера использовали синтетические олигонуклеотиды, комплементарные негомологичному участку кДНК генов флуоресцентных белков. После процедуры деплеции, были сконструированы две экспрессионные библиотеки кДНК и проведен их скрининг по выявлению клонов, продуцирующих флуоресцентные белки. Был проведен анализ 30000 колоний из каждой

библиотеки. Результаты подсчета клонов представлены в таблице 3-23. В обоих случаях, число колоний, экспрессирующих флуоресцентный белок, кДНК которого не подвергалась деплеции, было близко к числу таких колоний в контрольной библиотеке, в то время как число колоний, экспрессирующих флуоресцентный белок, кДНК которого подвергалась целевой деплеции, было снижено.

Таблица 3-23. Результат скрининга контрольной и деплецированных библиотек кДНК из *Zoantus*

Библиотека кДНК	Контроль	С деплецией zGFP	С деплецией zRFP
Зеленые колонии	33	1	28
Красные колонии	17	14	0

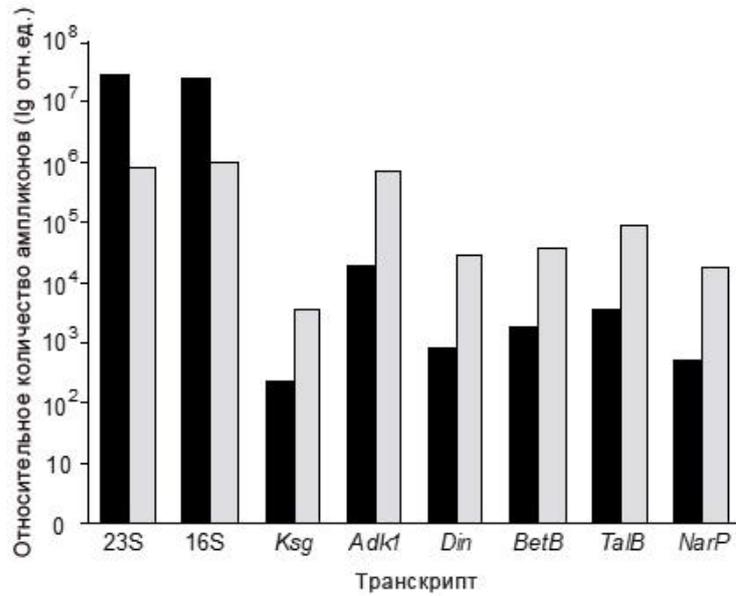
Таким образом, ДСН-деплегция позволяет эффективно удалить из образца уже известные транскрипты, одновременно снижая концентрацию высокоповторяющихся кДНК. Это обеспечивает возможность более глубокого анализа экспрессионных библиотек кДНК методом функционального скрининга.

Одним из важнейших применений Par-DSN стало удаление рРНК и тРНК из образцов кДНК перед анализом бактериального транскриптома. При работе с прокариотами использование стандартных технологий анализа транскриптома сильно затруднено из-за невозможности применения традиционных подходов селекции мРНК. Использование олиго(dT)-праймера в качестве затравки при синтезе кДНК прокариот малоэффективно, так как мРНК бактерий по большей части неполиаденирована [473]. При этом кДНК, если ее синтезировать с использованием случайной затравки, более чем на 95% будет состоять из копий рРНК и тРНК [661]. Таким образом, актуальной задачей при разработке технологий анализа транскриптома прокариот является удаление из образцов суммарной кДНК бактерий избытка копий рРНК и тРНК при сохранении относительной

представленности кодирующих последовательностей. Сходная задача стоит и при работе с образцами деградированной РНК из фиксированных тканей человека, для которых также не всегда применима селекция поли(А)+ РНК с помощью олиго(dT)-затравки [470, 474].

Мы разработали модификацию ДСН-деплеции для подготовки транскриптомов прокариот к полномасштабному секвенированию (далее – ДСН-деплеция рРНК). Для синтеза первой цепи кДНК на основе суммарной РНК прокариот мы адаптировали технологию SMART [639], заменив 3'-адаптер, содержащий олиго(dT)-последовательность на адаптер, содержащий декамерную случайную последовательность на 3'-конце и общую с SMART Oligo 5'-часть. Первая цепь кДНК, получаемая в ходе синтеза, оказывается фланкирована последовательностью адаптера, которая используется для амплификации кДНК. После амплификации дц кДНК денатурируют и оставляют ренатурировать в гибридизационных условиях. Образовавшиеся дц ДНК молекулы удаляются из реакционной смеси с помощью обработки Par_DSN. В отличие от ДСН-нормализации, условия гибридизации в данном случае подобраны таким образом, что дц ДНК фракцию образуют многокопийные молекулы рРНК и тРНК, а копии мРНК остаются одноцепочечными и сохраняют исходное распределение в образце. Эффективность предложенной процедуры истощения кДНК тРНК и рРНК была подтверждена в модельном эксперименте на примере *E. coli* штамм XL1-blue.

В таблице 3-24 и на рисунке 3.55 представлены результаты анализа профилей экспрессии различных генов в образцах ДСН-деплецированных и контрольных образцов кДНК. Из приведенных данных видно, что в ходе деплеции происходит обеднение образцов кДНК рибосомальной РНК примерно в 50 раз, а концентрация молекул мРНК возрастает примерно в 30 раз. При этом профиль распределения мРНК в образце меняется незначительно. Очевидно, что увеличить глубину истощения рРНК можно дополнительно добавив в реакционную смесь для гибридизации драйвер, содержащий фрагменты транскриптов, подлежащих деплеции, как это было описано выше.



Относительное число транскриптов (логарифмическая шкала) в контрольном (черные столбцы) и истощенном по рРНК (серые столбцы) образцах кДНК.

Рисунок 3.55. Влияние ДСН-деплеции рРНК на представленность различных транскриптов по результатам количественной ПЦР

Таблица 3-24. Результаты ПЦР в реальном времени ряда транскриптов из образцов бактериальных кДНК, полученных при ДСН-деплеции рРНК

Ген	Длина амплифицируемого фрагмента	Относительное число ампликонов	
		Образец С (контрольный)	Образец D (после DSN-деплеции рРНК)
<i>23S pPHK</i>	342	29920083	857637,585
<i>16S pPHK</i>	553	26099353,9	983188,62
<i>Ksg</i>	380	236,222642	3630,63243
<i>Adk1</i>	240	20028,354	722998,685
<i>Din</i>	487	835,895763	28182,3922
<i>BetB</i>	249	1833,65274	39656,1409
<i>TalB</i>	270	3630,63243	90013,7132
<i>NarP</i>	477	518,186973	18077,796

Идентификационный номер GenBank (аннотированный геном): CP081007.
Структуры праймеров приведены в главе «Материалы и методы»

Для дальнейшей проверки применимости технологии ДСН-деплеции рРНК к дифференциальному анализу экспрессии, было проведено сравнение с помощью супрессионной вычитающей гибридизации [662, 663] деплецированных транскриптомов двух близких штаммов *E. coli* штамм XL1-blue и BL 21.DE3. В качестве контроля использовали образцы кДНК, не подвергавшиеся ДСН-деплеции рРНК.

Полученные после вычитающей гибридизации образцы клонировали и анализировали полученные библиотеки методом дифференциального скрининга. Выявленные клоны секвенировали и подтверждали дифференциальный характер их экспрессии с помощью ПЦР в реальном времени (таблица 3-25, рисунок 3.56). Все клоны из экспериментальных библиотек подтвердили дифференциальный характер экспрессии, тогда как клоны из контрольных библиотек оказались ложнопозитивными.

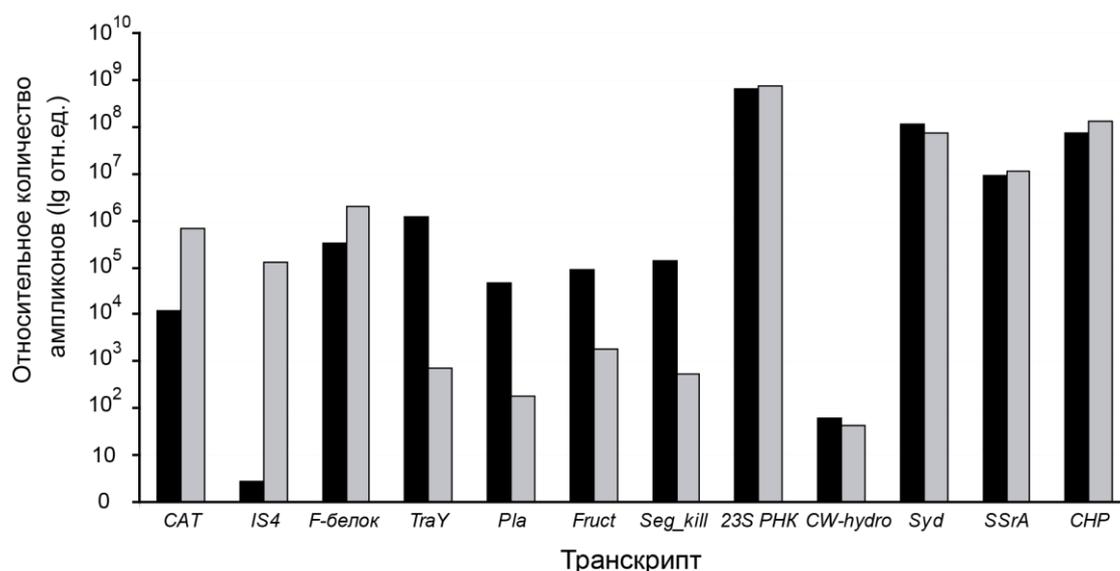


Рисунок 3.56. Результат количественной ПЦР на матрице контрольной первой цепи кДНК из штаммов *E. coli* XL1 blue (черные столбцы) и BL 21.DE3 (серые столбцы) с использованием праймеров, специфических к дифференциально экспрессирующимся генам, выявленным в ходе дифференциального скрининга вычитенных библиотек кДНК

Таблица 3-25. Результаты дифференциального скрининга контрольных и экспериментальных образцов после вычитающей гибридизации штаммов *E. coli*

Библиотека кДНК	Клонов в контиге (штук)	Транскрипт	Дифференциальный характер экспрессии
Библиотеки после ДСН-деплеции рРНК			
Обогащенная последовательностями, характерными для штамма <i>E. coli</i> BL 21.DE3 (BL 21.DE3-D)	13	<i>CAT</i>	подтвержден
	7	<i>IS4</i>	подтвержден
	1	<i>F-белок</i>	подтвержден
Обогащенная последовательностями, характерными для штамма <i>E. coli</i> XL1-blue (XL1-D)	12	<i>TraY</i>	подтвержден
	5	<i>Pla</i>	подтвержден
	4	<i>Fruct</i>	подтвержден
	2	<i>Seg_kill</i>	подтвержден
Контрольные библиотеки			
Обогащенная последовательностями, характерными для штамма <i>E. coli</i> BL 21.DE3 (BL 21.DE3-C)	3	23S РНК	не подтвержден
	1	<i>CW-hydro</i>	не подтвержден
Обогащенная последовательностями, характерными для штамма <i>E. coli</i> XL1-blue (XL1-C)	2	<i>Syd</i>	не подтвержден
	4	<i>SSrA</i>	не подтвержден
	2	<i>CHP</i>	не подтвержден

Таким образом, метод ДСН-деплеции рРНК позволяет эффективно уменьшать в суммарной кДНК прокариот содержание сверхпредставленных транскриптов (копии рРНК и тРНК), сохраняя относительную представленность копий мРНК. При этом деплецированные образцы кДНК могут быть использованы для количественного сравнения уровней экспрессии генов, а также для секвенирования транскриптома.

Разработанные нами метод ДСН-деплеции был апробирован во множестве лабораторий, в основном для удаления рРНК из различных образцов кДНК

прокариот. Например, Archer et al. (2014) [664] использовали 50 проб длиной 20 п.о., соответствующих фрагментам рРНК для ее деплеции в образцах кДНК *Saccharomyces cerevisiae* и подтвердили эффективность этого подхода. Также ДСН-деплегция со специфическими пробами для удаления рРНК была включена в протокол исследования 5'-3'котрансляционного распада мРНК *Saccharomyces cerevisiae* [665]. Chung et al. (2015) рекомендовали ДСН-деплецию рРНК как удобный инструмент удаления избытков рРНК из образцов для количественного анализа транскриптомов микроорганизмов [666]. Эффективность ДСН-деплеции рРНК была также подтверждена в работах Garoutte et al. (2016) [667] на примере зондирования метатранскриптома пастбищной почвы, и Hoeijmakers et al. (2012) [668] при анализе дифференциальной экспрессии *Plasmodium*. Дополнительно Fang & Akinci-Tolun (2016) [669] продемонстрировали возможность использования ДСН-деплеции рРНК для истощения рРНК при приготовлении библиотек из единичных клеток.

ЗАКЛЮЧЕНИЕ

Значение нуклеаз для процессов, происходящих в живых организмах, сложно переоценить. Не менее важным является практическое использование нуклеаз в биотехнологии и биомедицине. Поэтому поиск и изучение нуклеаз с новыми свойствами представляет собой важную научную и практическую задачи. Несмотря на широкий спектр нуклеаз, используемых в молекулярно-биологических исследованиях на момент начала данной работы, среди них отсутствовали клонированные термостабильные ферменты, способные избирательно гидролизовать двухцепочечную ДНК, оставляя одноцепочечную ДНК интактной. В то же время ферменты с такой активностью были необходимы для создания технологий анализа сложных смесей НК.

В рамках настоящей работы была клонирована полноразмерная кодирующая последовательность нового фермента из гепатопанкреаса камчатского краба – дуплекс-специфическая нуклеаза (Par_DSN). Анализ нуклеотидной последовательности Par_DSN выявил открытую рамку считывания длиной в 1221 п.о., кодирующую белок, состоящий из 407 аминокислот. Предсказанная аминокислотная последовательность нуклеазы краба включала в себя сигнальный пептид и нуклеазный домен. Получить активный рекомбинантный белок в гетерологической бактериальной системе экспрессии не удалось – весь продукт обнаруживался в нерастворимой фракции в тельцах включения. Поэтому была предложена и реализована методика очистки природного фермента, состоящая из комбинации нескольких последовательных стадий хроматографии. Идентичность очищенного природного белка нуклеазы из камчатского краба клонированному нами ферменту Par_DSN была подтверждена с помощью поликлональных антител (полученных к рекомбинантному белку), а также при помощи масс-спектрометрического анализа. Был проведен детальный анализ физико-химических свойств выделенного природного фермента, включая зависимость нуклеазы от катионов двухвалентных металлов, температуры, pH, мочевины,

ионной силы раствора и т.д. Было показано, что нуклеаза проявляет ферментативную активность только в присутствии двухвалентных катионов металлов и ингибируется ЭДТА, то есть является металл-зависимым ферментом. Было обнаружено, что Par_DSN проявляет активность по разрушению дц ДНК в широком диапазоне температур. При использовании в качестве субстрата дц ДНК температурный оптимум наблюдался при 60°C, а 5% активности фермента сохранялось даже после его инкубации в течение 30 мин при температуре 100°C. Было продемонстрировано, что фермент обладает селективной избирательностью по отношению к дц ДНК, оставляя оц ДНК интактной. Нуклеаза расщепляла любые дц ДНК субстраты (геномную ДНК, фаговую ДНК, суперскрученную плазмидную ДНК), но не проявляла заметной активности против оц ДНК. Подсчитанное соотношение гидролитической активности фермента к дц и оц ДНК субстрату составило не менее 1000:1, что значительно выше, чем для всех известных нуклеаз. Было показано, что Par_DSN практически неактивна по отношению к РНК субстрату, но проявляет свойства металл-зависимой ДНКазы по отношению к ДНК в ДНК-РНК гибридах. Была определена минимальная длина дуплекса ДНК, разрушаемого нуклеазой краба, составившая 8 п.о. Также продемонстрирована преимущественная активность фермента к полностью комплементарным дуплексам ДНК, по сравнению с несущими в своем составе хотя бы одну нуклеотидную замену в одной из цепей ДНК.

Был проведен сравнительный анализ последовательностей Par_DSN и ее гомологов из других членистоногих (в том числе клонированных нами полноразмерных кодирующих последовательности нуклеаз еще из трех видов ракообразных) с нуклеазами SNF. Анализ филогенетического древа выявил две большие клады. Первая включала в себя «классические» SNF нуклеазы. Вторая клада, названная нами подгруппой DSN-подобных нуклеаз, состояла из нуклеаз насекомых и ракообразных. На основании анализа филогенетического древа и данных об известной субстратной специфичности некоторых нуклеаз нами было выдвинуто предположение, что DSN-подобные нуклеазы и SNF нуклеазы эволюционировали по дивергентному пути, о чем свидетельствуют их различные

биологические функции и субстратные специфичности. Анализ нуклеазной активности сконструированных укороченных форм Par_DSN, а также рекомбинантных мутантных белков, полученных сайт-направленным мутагенезом, позволил сделать вывод, что Par_DSN и ее гомологи содержат в своем составе нуклеазные домены, которые длиннее соответствующего домена нуклеаз SNF, а их каталитический центр, хотя и сходен с каталитическим центром SNF нуклеаз, имеет ряд принципиальных отличий. Кроме того, между нуклеазным доменом и сигнальным пептидом DSN-подобные нуклеазы содержат регион длиной около 90 аминокислотных остатков, важный для определения субстратной специфичности и стабильности фермента. На основании этих отличий, нами было предложено выделить DSN-подобные нуклеазы в новое семейство – Дуплекс специфических нуклеаз (ДСН) внутри суперсемейства His-Me нуклеаз. Исходя из определенного сходства членов этого семейства с нуклеазами SNF, мы предположили их дивергентное расхождение в эволюции и постулировали, что нуклеазы с дц ДНКазной активностью могут быть найдены у других членистоногих, что и было в дальнейшем подтверждено рядом работ независимых лабораторий.

Таким образом, нами впервые в мире был клонирован и охарактеризован термостабильный фермент, обладающий непревзойденной селективностью по отношению к дц ДНК – дуплекс-специфическая нуклеаза Par_DSN.

На основе уникальных свойств Par_DSN нами были разработаны новые высокоэффективные технологии: анализа ОНП в генах эукариот, выявления целевых ДНК-мишеней в комплексной смеси нуклеиновых кислот, нормализации кДНК и геномной ДНК, селективного удаления нецелевых транскриптов из популяций кДНК эукариотических организмов и удаления последовательностей рРНК при создании библиотек кДНК прокариотических организмов. Разработанные нами технологии нашли свое применение в совместных работах в рамках международного научного консорциума по исследованию хронической боли, международной программы по исследованию рака молочной железы, в рамках проекта по изучению механизмов старения, и были высоко оценены зарубежными коллегами.

Данные настоящей работы легли в основу создания ряда коммерчески доступных продуктов, что привело к широкому внедрению полученных в исследовании результатов в мировую лабораторную практику. Методы нормализации и деплеции с использованием Par_DSN стали методами выбора при подготовке биологических образцов для высокопроизводительного секвенирования. В настоящее время число статей, упоминающих дуплекс-специфическую нуклеазу камчатского краба, в базе данных Google Scholar превышает 2000.

ВЫВОДЫ

1. Открыто новое семейство нуклеаз, специфически разрушающих двухцепочечную ДНК, и получившее название дуплекс специфические нуклеазы (ДСН).
2. Проведен филогенетический анализ белков данного семейства, а также исследованы функциональные и структурные особенности ферментов данного семейства на примере дуплекс специфической нуклеазы Par_DSN из камчатского краба (*Paralithodes camtschaticus*).
3. Определены границы нуклеазного домена Par_DSN, а также проведен анализ сайтов, ответственных за каталитическую активность данного фермента.
4. На основе уникальной способности Par_DSN расщеплять полностью комплементарные короткие ДНК-ДНК дуплексы со значительно более высокой эффективностью, чем дуплексы, содержащие одну и более нуклеотидных замен, разработана группа технологий анализа мутаций в генах эукариот, а также технология выявления целевых ДНК-мишеней в комплексной смеси нуклеиновых кислот.
5. На основе уникальных особенностей Par_DSN – специфичности по отношению к дц ДНК, а также термостабильности фермента – разработана группа высокоэффективных технологий выделения целевой фракции из сложных смесей нуклеиновых кислот, включая технологию создания нормализованных библиотек кДНК из клеток и тканей эукариотических организмов, технологию создания нормализованных библиотек геномной ДНК из клеток и тканей эукариотических организмов, технологию селективного удаления нецелевых транскриптов из популяций кДНК эукариотических организмов, технологию удаления последовательностей рРНК при создании библиотек кДНК прокариотических организмов.

Список сокращений и условных обозначений

АМФ	Аденозинмонофосфат
АТФ	Аденозинтрифосфат
ГСО	Ген-специфический олигонуклеотид
ДВЖХ	Денатурирующая высокоэффективная жидкостная хроматография
ДНК	Дезоксирибонуклеиновая кислота
ДНКзимы	Дезоксирибозимы
дц	Двухцепочечный
ДСН	Дуплекс-специфическая нуклеаза
ДСНП-анализ	от англ. duplex-specific nuclease preference
кДНК	Комплементарная дезоксирибонуклеиновая кислота
мРНК	Матричная рибонуклеиновая кислота
НК	Нуклеиновая кислота
н.о.	Нуклеотидное основание
ОНП	Однонуклеотидный полиморфизм
оц	Одноцепочечный
ПА-нук	Поликлональные антитела
п.о.	Пара оснований
ПЦР	Полимеразная цепная реакция
РНК	Рибонуклеиновая кислота
РВ-ПЦР	Полимеразная цепная реакция в реальном времени
рРНК	Рибосомальная рибонуклеиновая кислота
ССР	Сайт-специфическая рекомбиназа

т.п.о.	Тысяча пар оснований
тРНК	Транспортная рибонуклеиновая кислота
ХЭ	Хоуминг-эндонуклеазы
ЭР	Эндонуклеазы рестрикции
ЭДТА	Этилендиаминтетрауксусная кислота
C0t	Произведение концентрации нуклеотидов (C0) в молях на литр на время реассоциации в секундах (t)
dNTP	Дезоксинуклеозидтрифосфат
ddNTP	Дидезоксинуклеозидтрифосфат
HRM	High Resolution Melting
IR	Insert repeats
MB	Molecular beacons
NGS	Next generation sequencing (секвенирование нового поколения)
Par_DSN	Дуплекс-специфической нуклеаза из камчатского краба <i>Paralithodes camtschaticus</i>
TALEN	Transcription activator-like effector nuclease
Tc	Critical temperature
Tm	Temperature melting
TR	Tandem repeats
ZFN	Zinc-finger nuclease

СПИСОК ЛИТЕРАТУРЫ

1. Behzadi P., Ranjbar R. DNA microarray technology and bioinformatic web services // *Acta Microbiol. Immunol. Hung.* *Acta Microbiol Immunol Hung*, 2019. Vol. 66, № 1. P. 19–30.
2. Bahassi E.M., Stambrook P.J. Next-generation sequencing technologies: breaking the sound barrier of human genetics // *Mutagenesis*. 2014. Vol. 29, № 5. P. 303–310.
3. De Maria Marchiano R. et al. Translational Research in the Era of Precision Medicine: Where We Are and Where We Will Go // *J Pers Med*. 2021. Vol. 11, № 3. P. 216.
4. Giani A.M. et al. Long walk to genomics: History and current approaches to genome sequencing and assembly // *Comput. Struct. Biotechnol. J*. 2020. Vol. 18. P. 9–19.
5. Heather J.M., Chain B. The sequence of sequencers: The history of sequencing DNA // *Genomics*. 2016. Vol. 107, № 1. P. 1–8.
6. Hu C. et al. A Population-Based Study of Genes Previously Implicated in Breast Cancer // *N. Engl. J. Med*. 2021. Vol. 384, № 5. P. 440–451.
7. Han F. et al. Genome-wide characterization of Toll-like receptors in Japanese meagre *Argyrosomus japonicus* and their response to poly (I:C) injection // *Aquaculture*. 2021. Vol. 542. P. 736907.
8. Jackson D.A., Symons R.H., Berg P. Biochemical method for inserting new genetic information into DNA of Simian Virus 40: circular SV40 DNA molecules containing lambda phage genes and the galactose operon of *Escherichia coli* // *Proc. Natl. Acad. Sci. U. S. A*. 1972. Vol. 69, № 10. P. 2904–2909.

9. Lehman I.R. DNA ligase: structure, mechanism, and function // *Science*. 1974. Vol. 186, № 4166. P. 790–797.
10. Roberts R.J. How restriction enzymes became the workhorses of molecular biology // *Proceedings of the National Academy of Sciences of the United States of America*. 2005. Vol. 102, № 17. P. 5905–5908.
11. Saiki R.K. et al. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase // *Science*. 1988. Vol. 239, № 4839. P. 487–491.
12. Burmistrz M., Krakowski K., Krawczyk-Balska A. RNA-Targeting CRISPR-Cas Systems and Their Applications // *Int. J. Mol. Sci.* 2020. Vol. 21, № 3. P. 1122
13. Chylinski K. et al. Classification and evolution of type II CRISPR-Cas systems // *Nucleic Acids Res.* 2014. Vol. 42, № 10. P. 6091–6105.
14. Khosravi M.A. et al. Targeted deletion of BCL11A gene by CRISPR-Cas9 system for fetal hemoglobin reactivation: A promising approach for gene therapy of beta thalassemia disease // *Eur. J. Pharmacol.* 2019. Vol. 854. P. 398–405.
15. Sorek R., Kunin V., Hugenholtz P. CRISPR--a widespread system that provides acquired resistance against phages in bacteria and archaea // *Nat. Rev. Microbiol.* 2008. Vol. 6, № 3. P. 181–186.
16. Valenti M.T. et al. CRISPR/Cas system: An emerging technology in stem cell research // *World J. Stem Cells.* 2019. Vol. 11, № 11. P. 937–956.
17. Boone M., De Koker A., Callewaert N. Capturing the “ome”: the expanding molecular toolbox for RNA and DNA library construction // *Nucleic Acids Res.* 2018. Vol. 46, № 6. P. 2701–2721.
18. Balzano E., Pelliccia F., Giunta S. Genome (in)stability at tandem repeats // *Semin. Cell Dev. Biol.* 2021. Vol. 113. P. 97–112.

19. de Koning A.P.J. et al. Repetitive elements may comprise over two-thirds of the human genome // *PLoS Genet.* 2011. Vol. 7, № 12. P. e1002384.
20. Janicki M., Rooke R., Yang G. Bioinformatics and genomic analysis of transposable elements in eukaryotic genomes // *Chromosome Res.* 2011. Vol. 19, № 6. P. 787–808.
21. Lapitanz N.L.V. Organization and evolution of higher plant nuclear genomes // *Genome.* 1992. Vol. 35, № 2. P. 171–181.
22. Wanjugi H. et al. Rapid development of PCR-based genome-specific repetitive DNA junction markers in wheat // *Genome.* 2009. Vol. 52, № 6. P. 576–587.
23. González-Porta M. et al. Transcriptome analysis of human tissues and cell lines reveals one dominant transcript per gene // *Genome Biol.* 2013. Vol. 14, № 7. P. R70.
24. Wang Q. et al. Comparative transcriptomics in three Passerida species provides insights into the evolution of avian mitochondrial complex I // *Comp. Biochem. Physiol. Part D Genomics Proteomics.* 2018. Vol. 28. P. 27–36.
25. Ngara T.R., Zhang H. Recent Advances in Function-based Metagenomic Screening // *Genomics Proteomics Bioinformatics.* 2018. Vol. 16, № 6. P. 405–415.
26. Simon C., Daniel R. Metagenomic Analyses: Past and Future Trends // *Applied and Environmental Microbiology.* 2011. Vol. 77, № 4. P. 1153–1161.
27. Мензорова Н.И., Зинатулин Р.Ф., Фаворов В.В., Рассказов В.А. Выделение и исследование свойств Са, Mg-зависимой эндонуклеазы из гепатопанкреаса краба *Paralithodes camtschaticall* Биохимия. 1993. Т. 58. С. 681-691.
28. Menzorova N.I., Markova A.V., Rasskazov V.A. Highly stable Ca²⁺, Mg²⁺ - dependent Dnase from crab hepatopancreas // *Biochemistry.* 1994. Vol. 59, No 3. P. 449-456.

29. Harosh I. et al. Purification and characterization of a mitochondrial endonuclease from *Drosophila melanogaster* embryos // *Eur. J. Biochem.* 1992. Vol. 210, № 2. P. 455–460.
30. Nilsen I.W. et al. Protein purification and gene isolation of chlamysin, a cold-active lysozyme-like enzyme with antibacterial activity // *FEBS Lett.* 1999. Vol. 464, № 3. P. 153–158.
31. Ma Y. et al. Short-probe-based duplex-specific nuclease signal amplification strategy enables imaging of endogenous microRNAs in living cells with ultrahigh specificity // *Talanta.* 2018. Vol. 186. P. 256–264.
32. Pang J.K.S., Phua Q.H., Soh B.-S. Applications of miRNAs in cardiac development, disease progression and regeneration // *Stem Cell Res. Ther.* 2019. Vol. 10, № 1. P. 336.
33. Huang P. et al. A novel DSN-based fluorescence assay for MicroRNA-133a detection and its application for LVH diagnosis in maintenance hemodialysis patients // *J. Clin. Lab. Anal.* 2020. Vol. 34, № 10. P. e23438.
34. Wu Y. et al. Recent advances in duplex-specific nuclease-based signal amplification strategies for microRNA detection // *Biosens. Bioelectron.* 2020. Vol. 165. P. 112449.
35. Xu W. et al. Au@Ag core-shell nanoparticles for microRNA-21 determination based on duplex-specific nuclease signal amplification and surface-enhanced Raman scattering // *Mikrochim. Acta.* 2020. Vol. 187, № 7. P. 384.
36. Araki T. Über die Nucleinsäure aus der Schleimhaut des Dünndarms. Vorläufige Mitteilung // *Hoppe-Seyler's Zeitschrift für physiologische Chemie.* 1903. Vol. 38, № 1-2. P. 98–100.

37. Iwanoff L. Über die fermentative Zersetzung der Thymonucleinsäure durch Schimmelpilze // Hoppe-Seyler's Zeitschrift für physiologische Chemie. 1903. Vol. 39, № 1. P. 31–43.
38. Shen B. et al. Multiple but dissectible functions of FEN-1 nucleases in nucleic acid processing, genome stability and diseases // Bioessays. 2005. Vol. 27, № 7. P. 717–729.
39. Kao H.-I., Bambara R.A. The protein components and mechanism of eukaryotic Okazaki fragment maturation // Crit. Rev. Biochem. Mol. Biol. 2003. Vol. 38, № 5. P. 433–452.
40. Shevelev I.V., Hübscher U. The 3' 5' exonucleases // Nat. Rev. Mol. Cell Biol. 2002. Vol. 3, № 5. P. 364–376.
41. Reha-Krantz L.J. DNA polymerase proofreading: Multiple roles maintain genome stability // Biochim. Biophys. Acta. 2010. Vol. 1804, № 5. P. 1049–1063.
42. Marti T.M., Fleck O. DNA repair nucleases // Cell. Mol. Life Sci. 2004. Vol. 61, № 3. P. 336–354.
43. Mimitou E.P., Symington L.S. DNA end resection: many nucleases make light work // DNA Repair. 2009. Vol. 8, № 9. P. 983–995.
44. Nishino T., Morikawa K. Structure and function of nucleases in DNA repair: shape, grip and blade of the DNA scissors // Oncogene. 2002. Vol. 21, № 58. P. 9022–9032.
45. Di Felice F., Camilloni G. Why Should DNA Topoisomerase I Have a Scaffold Activity? // Biology. 2021. Vol. 10, № 3. P. 190.
46. Schoeffler A.J., Berger J.M. DNA topoisomerases: harnessing and constraining energy to govern chromosome topology // Q. Rev. Biophys. 2008. Vol. 41, № 1. P. 41–101.

47. Soren B.C. et al. Topoisomerase IB: a relaxing enzyme for stressed DNA // *Cancer Drug Resistance*. 2020. Vol. 3, № 1. P. 18–25
48. Grindley N.D.F., Whiteson K.L., Rice P.A. Mechanisms of Site-Specific Recombination // *Annual Review of Biochemistry*. 2006. Vol. 75, № 1. P. 567–605.
49. Butterfield R.J. et al. Congenital lethal motor neuron disease with a novel defect in ribosome biogenesis // *Neurology*. 2014. Vol. 82, № 15. P. 1322–1330.
50. Gasse L., Flemming D., Hurt E. Coordinated Ribosomal ITS2 RNA Processing by the Las1 Complex Integrating Endonuclease, Polynucleotide Kinase, and Exonuclease Activities // *Mol. Cell*. 2015. Vol. 60, № 5. P. 808–815.
51. Mattijssen S., Welting T.J.M., Pruijn G.J.M. RNase MRP and disease // *Wiley Interdiscip. Rev. RNA*. 2010. Vol. 1, № 1. P. 102–116.
52. Patel A.A., Steitz J.A. Splicing double: insights from the second spliceosome // *Nat. Rev. Mol. Cell Biol*. 2003. Vol. 4, № 12. P. 960–970.
53. Tafforeau L. et al. The complexity of human ribosome biogenesis revealed by systematic nucleolar screening of Pre-rRNA processing factors // *Mol. Cell*. 2013. Vol. 51, № 4. P. 539–551.
54. Cooper A.M. et al. Molecular mechanisms influencing efficiency of RNA interference in insects // *Pest Manag. Sci*. 2019. Vol. 75, № 1. P. 18–28.
55. James P., Halladay J., Craig E.A. Genomic Libraries and a Host Strain Designed for Highly Efficient Two-Hybrid Selection in Yeast // *Genetics*. 1996. Vol. 144, № 4. P. 1425–1436.
56. Tock M.R., Dryden D.T.F. The biology of restriction and anti-restriction // *Curr. Opin. Microbiol*. 2005. Vol. 8, № 4. P. 466–472.

57. Millman A. et al. Bacterial Retrons Function In Anti-Phage Defense // *Cell*. 2020. Vol. 183, № 6. P. 1551–1561.e12.
58. Gupta S.K. et al. The mammalian secreted RNases: mechanisms of action in host defence // *Innate Immun*. 2013. Vol. 19, № 1. P. 86–97.
59. Counis M.-F., Torriglia A. Acid DNases and their interest among apoptotic endonucleases // *Biochimie*. 2006. Vol. 88, № 12. P. 1851–1858.
60. Parrish J.Z., Xue D. Cuts can kill: the roles of apoptotic nucleases in cell death and animal development // *Chromosoma*. 2006. Vol. 115, № 2. P. 89–97.
61. Samejima K., Earnshaw W.C. Trashing the genome: the role of nucleases during apoptosis // *Nat. Rev. Mol. Cell Biol*. 2005. Vol. 6, № 9. P. 677–688.
62. Zhang L., Reha-Krantz L.J. Nuclease // *Brenner's Encyclopedia of Genetics*. 2013. P. 118–123.
63. Santa P. et al. The Role of Nucleases and Nucleic Acid Editing Enzymes in the Regulation of Self-Nucleic Acid Sensing // *Front. Immunol*. 2021. Vol. 12. P. 629922.
64. Barnard E.A. Ribonucleases // *Annual Review of Biochemistry*. 1969. Vol. 38, № 1. P. 677–732.
65. Laskowski M. [34] Purification and properties of the mung bean nuclease // *Nucleic Acids Part I*. 1980. P. 263–276.
66. Linn L. DSM-III // *JAMA*. 1982. Vol. 247, № 23. P. 3207–9.
67. Dupureur C. An Integrated Look at Metallonuclease Mechanism // *Current Chemical Biology*. 2008. Vol. 2, № 2. P. 159–173.
68. Yang W. Nucleases: diversity of structure, function and mechanism // *Q. Rev. Biophys*. 2011. Vol. 44, № 1. P. 1–93.

69. Doudna J.A., Cech T.R. The chemical repertoire of natural ribozymes // *Nature*. 2002. Vol. 418, № 6894. P. 222–228.
70. van Gent D.C., Mizuuchi K., Gellert M. Similarities between initiation of V(D)J recombination and retroviral integration // *Science*. 1996. Vol. 271, № 5255. P. 1592–1594.
71. Kirsebom L.A., Trobro S. RNase P RNA-mediated cleavage // *IUBMB Life*. 2009. Vol. 61, № 3. P. 189–200.
72. Raines R.T. Ribonuclease A // *Chem. Rev.* 1998. Vol. 98, № 3. P. 1045–1066.
73. Deutscher M.P. et al. RNase PH: an *Escherichia coli* phosphate-dependent nuclease distinct from polynucleotide phosphorylase // *Proceedings of the National Academy of Sciences*. 1988. Vol. 85, № 13. P. 4710–4714.
74. Champoux J.J. DNA topoisomerases: structure, function, and mechanism // *Annu. Rev. Biochem.* 2001. Vol. 70. P. 369–413.
75. Abelson J., Trotta C.R., Li H. tRNA Splicing // *Journal of Biological Chemistry*. 1998. Vol. 273, № 21. P. 12685–12688.
76. Weir A.F. Deoxyribonuclease I (EC 3.1.21.1) and II (EC 3.1.22.1) // *Methods Mol. Biol.* 1993. Vol. 16. P. 7–16.
77. Melgar E., Goldthwait D.A., Ukstins I. Deoxyribonucleic Acid Nucleases // *Journal of Biological Chemistry*. 1968. Vol. 243, № 17. P. 4409–4416.
78. Weir A.F. Nucleases: an overview // *Methods Mol. Biol.* 1993. Vol. 16. P. 1–6.
79. Rangarajan S., Shankar V. Extracellular nuclease from *Rhizopus stolonifer*: purification and characteristics of - single strand preferential - deoxyribonuclease activity // *Biochim. Biophys. Acta*. 1999. Vol. 1473, № 2-3. P. 293–304.

80. Finger L.D. et al. The wonders of flap endonucleases: structure, function, mechanism and regulation // *Subcell. Biochem.* 2012. Vol. 62. P. 301–326.
81. Chou K.-M., Cheng Y.-C. The Exonuclease Activity of Human Apurinic/Apyrimidinic Endonuclease (APE1) // *Journal of Biological Chemistry.* 2003. Vol. 278, № 20. P. 18289–18296.
82. Li M., Wilson D.M. Human Apurinic/Apyrimidinic Endonuclease 1 // *Antioxidants & Redox Signaling.* 2014. Vol. 20, № 4. P. 678–707.
83. Hsia K.-C., Li C.-L., Yuan H.S. Structural and functional insight into sugar-nonspecific nucleases in host defense // *Curr. Opin. Struct. Biol.* 2005. Vol. 15, № 1. P. 126–134.
84. Rangarajan E.S., Shankar V. Sugar non-specific endonucleases // *FEMS Microbiol. Rev.* 2001. Vol. 25, № 5. P. 583–613.
85. Cochrane J.C., Strobel S.A. Catalytic strategies of self-cleaving ribozymes // *Acc. Chem. Res.* 2008. Vol. 41, № 8. P. 1027–1035.
86. Correll C.C. et al. RNA recognition and base flipping by the toxin sarcin // *J. Synchrotron Radiat.* 2004. Vol. 11, № Pt 1. P. 93–96.
87. Nowotny M., Yang W. Structural and functional modules in RNA interference // *Curr. Opin. Struct. Biol.* 2009. Vol. 19, № 3. P. 286–293.
88. Stahley M.R., Strobel S.A. RNA splicing: group I intron crystal structures reveal the basis of splice site selection and metal ion catalysis // *Curr. Opin. Struct. Biol.* 2006. Vol. 16, № 3. P. 319–326.
89. Drew H.R. Structural specificities of five commonly used DNA nucleases // *J. Mol. Biol.* 1984. Vol. 176, № 4. P. 535–557.

90. Baranovskii A.G., Buneva V.N., Nevinsky G.A. Human Deoxyribonucleases // *Biochemistry (Moscow)*. 2004. Vol. 69, № 6. P. 587–601.
91. Desai N.A., Shankar V. Single-strand-specific nucleases // *FEMS Microbiol. Rev.* 2003. Vol. 26, № 5. P. 457–491.
92. Fujimoto M., Kuninaka A., Yoshino H. Identity of Phosphodiesterase and Phosphomonoesterase Activities with Nuclease P1 (a Nuclease from *Penicillium citrinum*) // *Agricultural and Biological Chemistry*. 1974. Vol. 38, № 4. P. 785–790.
93. Oleson A.E., Janski A.M., Clark E.T. An extracellular nuclease from suspension cultures of tobacco // *Biochim. Biophys. Acta*. 1974. Vol. 366, № 1. P. 89–100.
94. Nishino T., Ishino Y., Morikawa K. Structure-specific DNA nucleases: structural basis for 3D-scissors // *Curr. Opin. Struct. Biol.* 2006. Vol. 16, № 1. P. 60–67.
95. Nowotny M., Gaur V. Structure and mechanism of nucleases regulated by SLX4 // *Curr. Opin. Struct. Biol.* 2016. Vol. 36. P. 97–105.
96. Rass U. Resolving branched DNA intermediates with structure-specific nucleases during replication in eukaryotes // *Chromosoma*. 2013. Vol. 122, № 6. P. 499–515.
97. Schwartz E.K., Heyer W.-D. Processing of joint molecule intermediates by structure-selective endonucleases during homologous recombination in eukaryotes // *Chromosoma*. 2011. Vol. 120, № 2. P. 109–127.
98. Suck D. DNA recognition by structure-selective nucleases // *Biopolymers*. 1997. Vol. 44, № 4. P. 405–421.
99. Cerritelli S.M., Crouch R.J. Ribonuclease H: the enzymes in eukaryotes // *FEBS J.* 2009. Vol. 276, № 6. P. 1494–1505.
100. Hyjek M., Figiel M., Nowotny M. RNases H: Structure and mechanism // *DNA Repair*. 2019. Vol. 84. P. 102672.

101. Champoux J.J., Schultz S.J. Ribonuclease H: properties, substrate specificity and roles in retroviral reverse transcription // FEBS J. 2009. Vol. 276, № 6. P. 1506–1516.
102. Malik H.S., Eickbush T.H. Phylogenetic analysis of ribonuclease H domains suggests a late, chimeric origin of LTR retrotransposable elements and retroviruses // Genome Res. 2001. Vol. 11, № 7. P. 1187–1197.
103. Nowotny M. et al. Crystal structure of *B. halodurans* RNase H catalytic domain mutant D132N. // Cell. 2005. Vol. 121. № 7. P. 1005-1016.
104. Nguyen T.-N. et al. Crystal structure of LC9-RNase H1, a type 1 RNase H with the type 2 active-site motif. // FEBS Lett. 2013. Vol. 587. № 9. P. 1418-1423.
105. Tadokoro T., Kanaya S. Ribonuclease H: molecular diversities, substrate binding domains, and catalytic mechanism of the prokaryotic enzymes // FEBS J. 2009. Vol. 276, № 6. P. 1482–1493.
106. Tsukiashi M. et al. Construction and characterization of ribonuclease H2 knockout NIH3T3 cells // J. Biochem. 2019. Vol. 165, № 3. P. 249–256.
107. Ohtani N. et al. Molecular diversities of RNases H // J. Biosci. Bioeng. Elsevier BV, 1999. Vol. 88, № 1. P. 12–19.
108. Lockhart A. et al. RNase H1 and H2 Are Differentially Regulated to Process RNA-DNA Hybrids // Cell Rep. 2019. Vol. 29, № 9. P. 2890–2900.e5.
109. Crow Y.J. et al. Mutations in the gene encoding the 3'-5' DNA exonuclease TREX1 cause Aicardi-Goutières syndrome at the AGS1 locus // Nat. Genet. 2006. Vol. 38, № 8. P. 917–920.
110. Crow Y.J., Rehwinkel J. Aicardi-Goutieres syndrome and related phenotypes: linking nucleic acid metabolism with autoimmunity // Hum. Mol. Genet. 2009. Vol. 18, № R2. P. R130–R136.

111. Zimmermann M. et al. CRISPR screens identify genomic ribonucleotides as a source of PARP-trapping lesions // *Nature*. 2018. Vol. 559, № 7713. P. 285–289.
112. Chon H. et al. Crystal Structure and Structure-based Mutational Analyses of RNase HIII from *Bacillus stearothermophilus*: A New Type 2 RNase H with TBP-like Substrate-binding Domain at the N Terminus // *Journal of Molecular Biology*. 2006. Vol. 356, № 1. P. 165–178.
113. Kochiwa H., Tomita M., Kanai A. Evolution of ribonuclease H genes in prokaryotes to avoid inheritance of redundant genes // *BMC Evol. Biol.* 2007. Vol. 7. P. 128.
114. Lang K.S. et al. Replication-Transcription Conflicts Generate R-Loops that Orchestrate Bacterial Stress Survival and Pathogenesis // *Cell*. 2017. Vol. 170, № 4. P. 787–799.e18.
115. Gupta R., Capalash N., Sharma P. Restriction endonucleases: natural and directed evolution // *Appl. Microbiol. Biotechnol.* 2012. Vol. 94, № 3. P. 583–599.
116. DI Felice F., Micheli G., Camilloni G. Restriction enzymes and their use in molecular biology: An overview // *J. Biosci.* 2019. Vol. 44, № 2. P. 38.
117. Murray N.E. 2001 Fred Griffith review lecture. Immigration control of DNA in bacteria: self versus non-self // *Microbiology*. 2002. Vol. 148, № Pt 1. P. 3–20.
118. Singleton M.R., Dillingham M.S., Wigley D.B. Structure and mechanism of helicases and nucleic acid translocases // *Annu. Rev. Biochem.* 2007. Vol. 76. P. 23–50.
119. Murray N.E. Type I restriction systems: sophisticated molecular machines (a legacy of Bertani and Weigle) // *Microbiol. Mol. Biol. Rev.* 2000. Vol. 64, № 2. P. 412–434.

120. Neaves K.J. et al. Direct visualization of G-quadruplexes in DNA using atomic force microscopy // *Nucleic Acids Res.* 2009. Vol. 37, № 18. P. 6269–6275.
121. Rosamond J. et al. Mechanisms of action of the type-I restriction endonuclease, *ecoB*, and the *recBC* DNase from *Escherichia coli* // *Cold Spring Harb. Symp. Quant. Biol.* 1979. Vol. 43 Pt 2. P. 1049–1057.
122. Davies D.R. et al. The Three-dimensional Structure of a Tn5Transposase-related Protein Determined to 2.9-Å Resolution // *Journal of Biological Chemistry.* 1999. Vol. 274, № 17. P. 11904–11913.
123. Orłowski J., Bujnicki J.M. Structural and evolutionary classification of Type II restriction enzymes based on theoretical and experimental analyses // *Nucleic Acids Res.* 2008. Vol. 36, № 11. P. 3552–3569.
124. Pingoud A., Jeltsch A. Structure and function of type II restriction endonucleases // *Nucleic Acids Res.* 2001. Vol. 29, № 18. P. 3705–3727.
125. Pingoud V. et al. Specificity changes in the evolution of type II restriction endonucleases: a biochemical and bioinformatic analysis of restriction enzymes that recognize unrelated sequences // *J. Biol. Chem.* 2005. Vol. 280, № 6. P. 4289–4298.
126. Pingoud A., Wilson G.G., Wende W. Type II restriction endonucleases--a historical perspective and more // *Nucleic Acids Res.* 2014. Vol. 42, № 12. P. 7489–7527.
127. Roberts R.J. et al. A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes // *Nucleic Acids Res.* 2003. Vol. 31, № 7. P. 1805–1812.
128. Szybalski W. et al. Class-IIIS restriction enzymes--a review // *Gene.* 1991. Vol. 100. P. 13–26.

129. Jakubauskas A. et al. Identification of a single HNH active site in type IIS restriction endonuclease Eco31I // *J. Mol. Biol.* 2007. Vol. 370, № 1. P. 157–169.
130. Sapranaukas R. et al. Novel subtype of type IIS restriction enzymes. BfiI endonuclease exhibits similarities to the EDTA-resistant nuclease Nuc of *Salmonella typhimurium* // *J. Biol. Chem.* 2000. Vol. 275, № 40. P. 30878–30885.
131. Chan S.-H., Stoddard B.L., Xu S.-Y. Natural and engineered nicking endonucleases--from cleavage mechanism to engineering of strand-specificity // *Nucleic Acids Res.* 2011. Vol. 39, № 1. P. 1–18.
132. Marshall J.J.T., Halford S.E. The type IIB restriction endonucleases // *Biochem. Soc. Trans.* 2010. Vol. 38, № 2. P. 410–416.
133. Vitkute J. et al. BpII, a new BcgI-like restriction endonuclease, which recognizes a symmetric sequence // *Nucleic Acids Res.* 1997. Vol. 25, № 22. P. 4444–4446.
134. Kong H. Analyzing the functional organization of a novel restriction modification system, the BcgI system // *J. Mol. Biol.* 1998. Vol. 279, № 4. P. 823–832.
135. Janulaitis A. et al. Cloning and sequence analysis of the genes coding for Eco57I type IV restriction-modification enzymes // *Nucleic Acids Research.* 1992. Vol. 20, № 22. P. 6051–6056.
136. Lacks S., Greenberg B. A deoxyribonuclease of *Diplococcus pneumoniae* specific for methylated DNA // *J. Biol. Chem.* 1975. Vol. 250, № 11. P. 4060–4066.
137. Raghavendra N.K., Bheemanaik S., Rao D.N. Mechanistic insights into type III restriction enzymes // *Front. Biosci.* . 2012. Vol. 17, № 3. P. 1094–1107.
138. Sears A. et al. Characterization of the Type III restriction endonuclease PstII from *Providencia stuartii* // *Nucleic Acids Res.* 2005. Vol. 33, № 15. P. 4775–4787.

139. Rao D.N., Dryden D.T.F., Bheemanaik S. Type III restriction-modification enzymes: a historical perspective // *Nucleic Acids Res.* 2014. Vol. 42, № 1. P. 45–55.
140. Fukuda E. et al. Cell death upon epigenetic genome methylation: a novel function of methyl-specific deoxyribonucleases // *Genome Biol.* 2008. Vol. 9, № 11. P. R163.
141. Loenen W.A.M., Raleigh E.A. The other face of restriction: modification-dependent enzymes // *Nucleic Acids Res.* 2014. Vol. 42, № 1. P. 56–69.
142. Barzel A. et al. Native homing endonucleases can target conserved genes in humans and in animal models // *Nucleic Acids Res.* 2011. Vol. 39, № 15. P. 6646–6659.
143. Silva G.H., Belfort M. Analysis of the LAGLIDADG interface of the monomeric homing endonuclease I-DmoI // *Nucleic Acids Res.* 2004. Vol. 32, № 10. P. 3156–3168.
144. Hafez M., Hausner G. Homing endonucleases: DNA scissors on a mission // *Genome.* 2012. Vol. 55, № 8. P. 553–569.
145. Marcaida M.J. et al. Homing endonucleases: from basics to therapeutic applications // *Cell. Mol. Life Sci.* 2010. Vol. 67, № 5. P. 727–748.
146. Stoddard B.L. Homing endonuclease structure and function // *Q. Rev. Biophys.* 2005. Vol. 38, № 1. P. 49–95.
147. Chevalier B. et al. Flexible DNA target site recognition by divergent homing endonuclease isoschizomers I-CreI and I-MsoI // *J. Mol. Biol. Elsevier BV*, 2003. Vol. 329, № 2. P. 253–269.
148. Zheleznaya L.A. et al. Nicking endonucleases // *Biochemistry.* 2009. Vol. 74, № 13. P. 1457–1466.
149. Geider K. et al. Proteins and nucleotide sequences involved in DNA replication of filamentous bacteriophage // *Adv. Exp. Med. Biol.* 1984. Vol. 179. P. 45–54.

150. Chan S.-H. et al. Cloning of CviPII nicking and modification system from chlorella virus NYs-1 and application of Nt.CviPII in random DNA amplification // *Nucleic Acids Res.* 2004. Vol. 32, № 21. P. 6187–6199.
151. Chan S.-H. et al. Cloning of Nt.CviQII nicking endonuclease and its cognate methyltransferase: M.CviQII methylates AG sequences // *Protein Expr. Purif.* 2006. Vol. 49, № 1. P. 138–150.
152. Xia Y.N. et al. A site-specific single strand endonuclease activity induced by NYs-1 virus infection of a Chlorella-like green alga // *Nucleic Acids Res.* 1988. Vol. 16, № 20. P. 9477–9487.
153. Francia M.V. et al. Catalytic domain of plasmid pAD1 relaxase TraX defines a group of relaxases related to restriction endonucleases // *Proc. Natl. Acad. Sci. U. S. A.* 2013. Vol. 110, № 33. P. 13606–13611.
154. Landthaler M., Shub D.A. The nicking homing endonuclease I-BasI is encoded by a group I intron in the DNA polymerase gene of the *Bacillus thuringiensis* phage Bastille // *Nucleic Acids Res.* 2003. Vol. 31, № 12. P. 3071–3077.
155. Landthaler M., Lau N.C., Shub D.A. Group I intron homing in *Bacillus* phages SPO1 and SP82: a gene conversion event initiated by a nicking homing endonuclease // *J. Bacteriol.* 2004. Vol. 186, № 13. P. 4307–4314.
156. Landthaler M. et al. I-BasI and I-HmuI: two phage intron-encoded endonucleases with homologous DNA recognition sequences but distinct DNA specificities // *J. Mol. Biol.* 2006. Vol. 358, № 4. P. 1137–1151.
157. Xu S.-Y. Sequence-specific DNA nicking endonucleases // *Biomol. Concepts.* 2015. Vol. 6, № 4. P. 253–267.

158. Syson K. et al. Structural insight into how *Streptomyces coelicolor* maltosyl transferase GlgE binds α -maltose 1-phosphate and forms a maltosyl-enzyme intermediate // *Biochemistry*. 2014. Vol. 53, № 15. P. 2494–2504.
159. Beloglazova N. et al. A novel family of sequence-specific endoribonucleases associated with the clustered regularly interspaced short palindromic repeats // *J. Biol. Chem.* 2008. Vol. 283, № 29. P. 20361–20371.
160. Carte J. et al. Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes // *Genes Dev.* 2008. Vol. 22, № 24. P. 3489–3496.
161. Wang R. et al. Interaction of the Cas6 riboendonuclease with CRISPR RNAs: recognition and cleavage // *Structure*. 2011. Vol. 19, № 2. P. 257–264.
162. Schmitz S., Nölle V., Elleuche S. A non-specific nucleolytic enzyme and its application potential in EDTA-containing buffer solutions // *Biotechnol. Lett.* 2019. Vol. 41, № 1. P. 129–136.
163. Dominguez K., Ward W.S. A novel nuclease activity that is activated by Ca(2+) chelated to EGTA // *Syst. Biol. Reprod. Med.* 2009. Vol. 55, № 5-6. P. 193–199.
164. Dupureur C.M. Roles of metal ions in nucleases // *Current Opinion in Chemical Biology*. 2008. Vol. 12, № 2. P. 250–255.
165. Dupureur C.M. One is enough: insights into the two-metal ion nuclease mechanism from global analysis and computational studies // *Metallomics*. 2010. Vol. 2, № 9. P. 609–620.
166. Maguire M.E., Cowan J.A. *BioMetals*. 2002. Vol. 15, № 3. P. 201–201.
167. Palermo G. et al. Catalytic metal ions and enzymatic processing of DNA and RNA // *Acc. Chem. Res.* 2015. Vol. 48, № 2. P. 220–228.

168. Conlan L.H., Dupureur C.M. Multiple metal ions drive DNA association by PvuII endonuclease // *Biochemistry*. 2002. Vol. 41, № 50. P. 14848–14855.
169. Bowen L.M., Dupureur C.M. Investigation of restriction enzyme cofactor requirements: a relationship between metal ion properties and sequence specificity // *Biochemistry*. 2003. Vol. 42, № 43. P. 12643–12653.
170. Kennedy A.K., Haniford D.B., Mizuuchi K. Single active site catalysis of the successive phosphoryl transfer steps by DNA transposases: insights from phosphorothioate stereoselectivity // *Cell*. 2000. Vol. 101, № 3. P. 295–305.
171. Yang W., Lee J.Y., Nowotny M. Making and breaking nucleic acids: two-Mg²⁺-ion catalysis and substrate specificity // *Mol. Cell*. 2006. Vol. 22, № 1. P. 5–13.
172. Brautigam C.A. et al. Structures of Normal Single-Stranded DNA and Deoxyribo-3'-S-phosphorothiolates Bound to the 3'-5' Exonucleolytic Active Site of DNA Polymerase I from *Escherichia coli*, // *Biochemistry*. 1999. Vol. 38, № 2. P. 696–704.
173. Steitz T.A., Steitz J.A. A general two-metal-ion mechanism for catalytic RNA // *Proc. Natl. Acad. Sci. U. S. A.* 1993. Vol. 90, № 14. P. 6498–6502.
174. Horton N.C., Newberry K.J., Perona J.J. Metal ion-mediated substrate-assisted catalysis in type II restriction endonucleases // *Proc. Natl. Acad. Sci. U. S. A.* 1998. Vol. 95, № 23. P. 13489–13494.
175. Shan S. et al. Defining the catalytic metal ion interactions in the *Tetrahymena* ribozyme reaction // *Biochemistry*. 2001. Vol. 40, № 17. P. 5161–5171.
176. Tsunaka Y. et al. Identification of single Mn(2+) binding sites required for activation of the mutant proteins of *E.coli* RNase HI at Glu48 and/or Asp134 by X-ray crystallography // *J. Mol. Biol.* 2005. Vol. 345, № 5. P. 1171–1183.

177. Horton N.C., Perona J.J. DNA cleavage by EcoRV endonuclease: two metal ions in three metal ion binding sites // *Biochemistry*. American Chemical Society (ACS), 2004. Vol. 43, № 22. P. 6841–6857.
178. Zuo Y., Deutscher M.P. Exoribonuclease superfamilies: structural analysis and phylogenetic distribution // *Nucleic Acids Res.* 2001. Vol. 29, № 5. P. 1017–1026.
179. Palanivelu P. An Overview of the Proofreading Functions in Bacteria and in Severe Acute Respiratory Syndrome-Coronaviruses // *International Journal of Biochemistry Research & Review*. 2021. P. 33–62.
180. Deutscher M.P., Li Z. Exoribonucleases and their multiple roles in RNA metabolism // *Prog. Nucleic Acid Res. Mol. Biol.* 2001. Vol. 66. P. 67–105.
181. Kim S.H. et al. TREX1 degrades the 3' end of the small DNA oligonucleotide products of nucleotide excision repair in human cells // *Nucleic Acids Res.* 2022. Vol. 50. № 7. P. 3974-3984.
182. Lauková L. et al. Deoxyribonucleases and Their Applications in Biomedicine // *Biomolecules*. 2020. Vol. 10, № 7. P.1036.
183. Crow Y.J., Manel N. Aicardi-Goutières syndrome and the type I interferonopathies // *Nat. Rev. Immunol.* 2015. Vol. 15, № 7. P. 429–440.
184. Stetson D.B. et al. Trex1 prevents cell-intrinsic initiation of autoimmunity // *Cell*. 2008. Vol. 134, № 4. P. 587–598.
185. Hasan M., Yan N. Safeguard against DNA sensing: the role of TREX1 in HIV-1 infection and autoimmune diseases // *Front. Microbiol.* 2014. Vol. 5. P. 193.
186. Yan N. et al. The cytosolic exonuclease TREX1 inhibits the innate immune response to human immunodeficiency virus type 1 // *Nat. Immunol.* 2010. Vol. 11, № 11. P. 1005–1013.

187. Chowdhury D. et al. The exonuclease TREX1 is in the SET complex and acts in concert with NM23-H1 to degrade DNA during granzyme A-mediated cell death // *Mol. Cell*. 2006. Vol. 23, № 1. P. 133–142.
188. Dumitrache L.C., Hu L., Hasty P. TREX2 exonuclease defective cells exhibit double-strand breaks and chromosomal fragments but not Robertsonian translocations // *Mutat. Res*. 2009. Vol. 662, № 1-2. P. 84–87.
189. Chen M.-J. et al. Biochemical and cellular characteristics of the 3' → 5' exonuclease TREX2 // *Nucleic Acids Res*. 2007. Vol. 35, № 8. P. 2682–2694.
190. Cevher M.A. et al. Nuclear deadenylation/polyadenylation factors regulate 3' processing in response to DNA damage // *EMBO J*. 2010. Vol. 29, № 10. P. 1674–1687.
191. Nanjappa D.P. et al. Poly (A)-specific ribonuclease (PARN): More than just “mRNA stock clearing” // *Life Sci*. 2021. Vol. 285. P. 119953.
192. Wu M. et al. Structural insight into poly(A) binding and catalytic mechanism of human PARN // *EMBO J*. 2005. Vol. 24, № 23. P. 4082–4093.
193. He G.-J., Yan Y.-B. Contributions of the C-terminal domain to poly(A)-specific ribonuclease (PARN) stability and self-association // *Biochemistry and Biophysics Reports*. 2019. Vol. 18. P. 100626.
194. Martinez J. et al. A 54-kDa fragment of the Poly(A)-specific ribonuclease is an oligomeric, processive, and cap-interacting Poly(A)-specific 3' exonuclease // *J. Biol. Chem*. 2000. Vol. 275, № 31. P. 24222–24230.
195. Andersen K.R. et al. The activity and selectivity of fission yeast Pop2p are affected by a high affinity for Zn²⁺ and Mn²⁺ in the active site // *RNA*. 2009. Vol. 15, № 5. P. 850–861.

196. Daugeron M.C., Mauxion F., Séraphin B. The yeast POP2 gene encodes a nuclease involved in mRNA deadenylation // *Nucleic Acids Res.* 2001. Vol. 29, № 12. P. 2448–2455.
197. Thore S. et al. X-ray structure and activity of the yeast Pop2 protein: a nuclease subunit of the mRNA deadenylase complex // *EMBO Rep.* 2003. Vol. 4, № 12. P. 1150–1155.
198. Perry J.J.P. et al. WRN exonuclease structure and molecular mechanism imply an editing role in DNA end processing // *Nat. Struct. Mol. Biol.* 2006. Vol. 13, № 5. P. 414–422.
199. Muftuoglu M. et al. The clinical characteristics of Werner syndrome: molecular and biochemical diagnosis // *Hum. Genet.* 2008. Vol. 124, № 4. P. 369–377.
200. Li Z., Pandit S., Deutscher M.P. 3' exoribonucleolytic trimming is a common feature of the maturation of small, stable RNAs in *Escherichia coli* // *Proc. Natl. Acad. Sci. U. S. A.* 1998. Vol. 95, № 6. P. 2856–2861.
201. Zuo Y., Wang Y., Malhotra A. Crystal structure of *Escherichia coli* RNase D, an exoribonuclease involved in structured RNA processing // *Structure.* 2005. Vol. 13, № 7. P. 973–984.
202. Houseley J., LaCava J., Tollervey D. RNA-quality control by the exosome // *Nat. Rev. Mol. Cell Biol.* 2006. Vol. 7, № 7. P. 529–539.
203. Januszyk K., Lima C.D. The eukaryotic RNA exosome // *Curr. Opin. Struct. Biol.* 2014. Vol. 24. P. 132–140.
204. Yang W., Steitz T.A. Recombining the structures of HIV integrase, RuvC and RNase H // *Structure.* 1995. Vol. 3, № 2. P. 131–134.

205. Nowotny M., Yang W. Stepwise analyses of metal ions in RNase H catalysis from substrate destabilization to product release // *EMBO J.* 2006. Vol. 25, № 9. P. 1924–1933.
206. Tramontano E., Corona A., Menéndez-Arias L. Ribonuclease H, an unexploited target for antiviral intervention against HIV and hepatitis B virus // *Antiviral Res.* 2019. Vol. 171. P. 104613.
207. Majorek K.A. et al. The RNase H-like superfamily: new members, comparative structural analysis and evolutionary classification // *Nucleic Acids Res.* 2014. Vol. 42, № 7. P. 4160–4179.
208. Moelling K. et al. RNase H As Gene Modifier, Driver of Evolution and Antiviral Defense // *Front. Microbiol.* 2017. Vol. 8. P. 1745.
209. Nowotny M. Retroviral integrase superfamily: the structural perspective // *EMBO Rep.* 2009. Vol. 10, № 2. P. 144–151.
210. Chiu T.K., Davies D.R. Structure and function of HIV-1 integrase // *Curr. Top. Med. Chem.* 2004. Vol. 4, № 9. P. 965–977.
211. Shida T. et al. Analysis of substrate specificity of the RuvC holliday junction resolvase with synthetic Holliday junctions // *J. Biol. Chem.* 1996. Vol. 271, № 42. P. 26105–26109.
212. Karakas E. et al. Structure of the C-terminal half of UvrC reveals an RNase H endonuclease domain with an Argonaute-like catalytic triad // *EMBO J.* 2007. Vol. 26, № 2. P. 613–622.
213. Verhoeven E.E. et al. Catalytic sites for 3' and 5' incision of Escherichia coli nucleotide excision repair are both located in UvrC // *J. Biol. Chem.* 2000. Vol. 275, № 7. P. 5120–5123.

214. Peters L., Meister G. Argonaute proteins: mediators of RNA silencing // *Mol. Cell.* 2007. Vol. 26, № 5. P. 611–623.
215. Jin S., Zhan J., Zhou Y. Argonaute proteins: structures and their endonuclease activity // *Mol. Biol. Rep.* 2021. Vol. 48, № 5. P. 4837–4849.
216. Willkomm S. et al. Structural and mechanistic insights into an archaeal DNA-guided Argonaute protein // *Nat Microbiol.* 2017. Vol. 2. P. 17035.
217. Wu J. 'en et al. Argonaute proteins: Structural features, functions and emerging roles // *J. Advert. Res.* 2020. Vol. 24. P. 317–324.
218. Olina A.V. et al. Argonaute Proteins and Mechanisms of RNA Interference in Eukaryotes and Prokaryotes // *Biochemistry.* 2018. Vol. 83, № 5. P. 483–497.
219. Hegge J.W., Swarts D.C., van der Oost J. Prokaryotic Argonaute proteins: novel genome-editing tools? // *Nat. Rev. Microbiol.* 2018. Vol. 16, № 1. P. 5–11.
220. Swarts D.C. et al. DNA-guided DNA interference by a prokaryotic Argonaute // *Nature.* 2014. Vol. 507, № 7491. P. 258–261.
221. Lingel A. et al. Structure and nucleic-acid binding of the *Drosophila* Argonaute 2 PAZ domain // *Nature.* 2003. Vol. 426, № 6965. P. 465–469.
222. Song J.-J. et al. The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes // *Nat. Struct. Biol.* 2003. Vol. 10, № 12. P. 1026–1032.
223. Yan K.S. et al. Structure and conserved RNA binding of the PAZ domain // *Nature.* 2003. Vol. 426, № 6965. P. 468–474.
224. Elkayam E. et al. The structure of human argonaute-2 in complex with miR-20a // *Cell.* 2012. Vol. 150, № 1. P. 100–110.

225. Darricarrère N. et al. Function of Piwi, a nuclear Piwi/Argonaute protein, is independent of its slicer activity // *Proc. Natl. Acad. Sci. U. S. A.* 2013. Vol. 110, № 4. P. 1297–1302.
226. Cao Y. et al. Argonaute proteins from human gastrointestinal bacteria catalyze DNA-guided cleavage of single- and double-stranded DNA at 37 °C // *Cell Discov.* 2019. Vol. 5. P. 38.
227. Yuan Y.-R. et al. Crystal structure of *A. aeolicus* argonaute, a site-specific DNA-guided endoribonuclease, provides insights into RISC-mediated mRNA cleavage // *Mol. Cell.* 2005. Vol. 19, № 3. P. 405–419.
228. Burroughs A.M., Iyer L.M., Aravind L. Two novel PIWI families: roles in inter-genomic conflicts in bacteria and Mediator-dependent modulation of transcription in eukaryotes // *Biol. Direct.* 2013. Vol. 8. P. 13.
229. Haugen P., Bhattacharya D. The spread of LAGLIDADG homing endonuclease genes in rDNA // *Nucleic Acids Res.* 2004. Vol. 32, № 6. P. 2049–2057.
230. Chevalier B.S., Stoddard B.L. Homing endonucleases: structural and functional insight into the catalysts of intron/intein mobility // *Nucleic Acids Res.* 2001. Vol. 29, № 18. P. 3757–3774.
231. Chevalier B.S., Monnat R.J. Jr, Stoddard B.L. The homing endonuclease I-CreI uses three metals, one of which is shared between the two active sites // *Nat. Struct. Biol.* 2001. Vol. 8, № 4. P. 312–316.
232. Moure C.M., Gimble F.S., Quioco F.A. The crystal structure of the gene targeting homing endonuclease I-SceI reveals the origins of its target site specificity // *J. Mol. Biol.* 2003. Vol. 334, № 4. P. 685–695.

233. Moure C.M., Gimble F.S., Quioco F.A. Crystal structures of I-SceI complexed to nicked DNA substrates: snapshots of intermediates along the DNA cleavage reaction pathway // *Nucleic Acids Res.* 2008. Vol. 36, № 10. P. 3287–3296.
234. Bolduc J.M. et al. Structural and biochemical analyses of DNA and RNA binding by a bifunctional homing endonuclease and group I intron splicing factor // *Genes Dev.* 2003. Vol. 17, № 23. P. 2875–2888.
235. Belfort M., Bonocora R.P. Homing endonucleases: from genetic anomalies to programmable genomic clippers // *Methods Mol. Biol.* 2014. Vol. 1123. P. 1–26.
236. Dias A. et al. The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit // *Nature.* 2009. Vol. 458, № 7240. P. 914–918.
237. Feder M., Bujnicki J.M. Identification of a new family of putative PD-(D/E)XK nucleases with unusual phylogenomic distribution and a new type of the active site // *BMC Genomics.* 2005. Vol. 6. P. 21.
238. Knizewski L. et al. Realm of PD-(D/E)XK nuclease superfamily revisited: detection of novel families with modified transitive meta profile searches // *BMC Struct. Biol.* 2007. Vol. 20. № 7. P. 40.
239. Kosinski J., Feder M., Bujnicki J.M. The PD-(D/E)XK superfamily revisited: identification of new members among proteins involved in DNA metabolism and functional predictions for domains of (hitherto) unknown function // *BMC Bioinformatics.* 2005. Vol. 12. № 6. P. 172.
240. Steczkiewicz K. et al. Sequence, structure and functional diversity of PD-(D/E)XK phosphodiesterase superfamily // *Nucleic Acids Res.* 2012. Vol. 40, № 15. P. 7016–7045.
241. Zhao L. et al. The restriction fold turns to the dark side: a bacterial homing endonuclease with a PD-(D/E)-XK motif // *EMBO J.* 2007. Vol. 26, № 9. P. 2432–2442.

242. Jin H., Cho Y. Structural and functional relationships of FAN1 // DNA Repair. 2017. Vol. 56. P. 135–143.
243. Pizzolato J. et al. FANCD2-associated nuclease 1, but not exonuclease 1 or flap endonuclease 1, is able to unhook DNA interstrand cross-links in vitro // J. Biol. Chem. 2015. Vol. 290, № 37. P. 22602–22611.
244. Aggarwal A.K. Structure and function of restriction endonucleases // Curr. Opin. Struct. Biol. 1995. Vol. 5, № 1. P. 11–19.
245. Venclovas C., Timinskas A., Siksnys V. Five-stranded beta-sheet sandwiched with two alpha-helices: a structural link between restriction endonucleases EcoRI and EcoRV // Proteins. 1994. Vol. 20, № 3. P. 279–282.
246. Taylor G.K., Stoddard B.L. Structural, functional and evolutionary relationships between homing endonucleases and proteins from their host organisms // Nucleic Acids Res. 2012. Vol. 40, № 12. P. 5189–5200.
247. Tsutakawa S.E., Jingami H., Morikawa K. Recognition of a TG mismatch: the crystal structure of very short patch repair endonuclease in complex with a DNA duplex // Cell. 1999. Vol. 99, № 6. P. 615–623.
248. Bujnicki J.M., Rychlewski L. Reassignment of specificities of two cap methyltransferase domains in the reovirus lambda 2 protein // Genome Biol. 2001. Vol. 2, № 9. P. RESEARCH0038.
249. Horton V. et al. Evolutionary relationship between different subgroups of restriction endonucleases // J. Biol. Chem. 2002. Vol. 277, № 16. P. 14306–14314.
250. Tamulaitis G., Solonin A.S., Siksnys V. Alternative arrangements of catalytic residues at the active sites of restriction enzymes // FEBS Lett. 2002. Vol. 518, № 1-3. P. 17–22.

251. Horton J.R. et al. Caught in the act: visualization of an intermediate in the DNA base-flipping pathway induced by HhaI methyltransferase // *Nucleic Acids Res.* 2004. Vol. 32, № 13. P. 3877–3886.
252. Kovall R.A., Matthews B.W. Type II restriction endonucleases: structural, functional and evolutionary relationships // *Curr. Opin. Chem. Biol.* 1999. Vol. 3, № 5. P. 578–583.
253. Groothuizen F.S., Sixma T.K. The conserved molecular machinery in DNA mismatch repair enzyme structures // *DNA Repair.* 2016. Vol. 38. P. 14–23.
254. Crépin T. et al. Mutational and metal binding analysis of the endonuclease domain of the influenza virus polymerase PA subunit // *J. Virol.* 2010. Vol. 84, № 18. P. 9096–9104.
255. Datta K. et al. Characterization of PA-N terminal domain of Influenza A polymerase reveals sequence specific RNA cleavage // *Nucleic Acids Res.* 2013. Vol. 41, № 17. P. 8289–8299.
256. Kowalinski E. et al. Structural analysis of specific metal chelating inhibitor binding to the endonuclease domain of influenza pH1N1 (2009) polymerase // *PLoS Pathog.* 2012. Vol. 8, № 8. P. e1002831.
257. Wandzik J.M., Kouba T., Cusack S. Structure and Function of Influenza Polymerase // *Cold Spring Harb. Perspect. Med.* 2021. Vol. 11, № 9. P. a038372
258. Grasby J.A. et al. Unpairing and gating: sequence-independent substrate recognition by FEN superfamily nucleases // *Trends Biochem. Sci.* 2012. Vol. 37, № 2. P. 74–84.
259. Tomlinson C.G. et al. Substrate recognition and catalysis by flap endonucleases and related enzymes // *Biochem. Soc. Trans.* 2010. Vol. 38, № 2. P. 433–437.

260. Tsutakawa S.E., Tainer J.A. Double strand binding-single strand incision mechanism for human flap endonuclease: implications for the superfamily // *Mech. Ageing Dev.* 2012. Vol. 133, № 4. P. 195–202.
261. Wu C.-C., Lin J.L.J., Yuan H.S. Structures, Mechanisms, and Functions of His-Me Finger Nucleases // *Trends Biochem. Sci.* 2020. Vol. 45, № 11. P. 935–946.
262. Yang W. An equivalent metal ion in one- and two-metal-ion catalysis // *Nat. Struct. Mol. Biol.* 2008. Vol. 15, № 11. P. 1228–1231.
263. Sokolowska M., Czapinska H., Bochtler M. Crystal structure of the beta beta alpha-Me type II restriction endonuclease Hpy99I with target DNA // *Nucleic Acids Res.* 2009. Vol. 37, № 11. P. 3799–3810.
264. Meiss G. et al. Mechanism of DNA cleavage by the DNA/RNA-non-specific *Anabaena* sp. PCC 7120 endonuclease NucA and its inhibition by NuiA // *J. Mol. Biol.* 2000. Vol. 297, № 2. P. 521–534.
265. Gimadutdinow O.A. et al. STRUCTURE, FUNCTION AND EVOLUTION OF *Serratia marcescens* ENDONUCLEASE // *Journal of Experimental Biology and Agricultural Sciences.* 2018. Vol. 6, № 1. P. 53–61.
266. Mehta P., Katta K., Krishnaswamy S. HNH family subclassification leads to identification of commonality in the His-Me endonuclease superfamily // *Protein Sci.* 2004. Vol. 13, № 1. P. 295–300.
267. Sui M.-J. et al. Metal ions and phosphate binding in the H-N-H motif: crystal structures of the nuclease domain of ColE7/Im7 in complex with a phosphate ion and different divalent metal ions // *Protein Sci.* 2002. Vol. 11, № 12. P. 2947–2957.
268. Benedik M.J., Strych U. *Serratia marcescens* and its extracellular nuclease // *FEMS Microbiol. Lett.* 1998. Vol. 165, № 1. P. 1–13.

269. Mannino S.J., Jenkins C.L., Raines R.T. Chemical mechanism of DNA cleavage by the homing endonuclease I-PpoI // *Biochemistry*. 1999. Vol. 38, № 49. P. 16178–16186.
270. Calcuttawala F. et al. Structural and functional insights into colicin: a new paradigm in drug discovery // *Archives of Microbiology*. 2022. Vol. 204, № 1. P. 37.
271. Mohanty A.K. et al. Enzymatic E-colicins bind to their target receptor BtuB by presentation of a small binding epitope on a coiled-coil scaffold // *J. Biol. Chem.* 2003. Vol. 278, № 42. P. 40953–40958.
272. Maté M.J., Kleanthous C. Structure-based analysis of the metal-dependent mechanism of H-N-H endonucleases // *J. Biol. Chem.* 2004. Vol. 279, № 33. P. 34763–34769.
273. Pommer A.J. et al. Mechanism and cleavage specificity of the H-N-H endonuclease colicin E9 // *J. Mol. Biol.* 2001. Vol. 314, № 4. P. 735–749.
274. Zhu X. et al. Cryo-EM structures reveal coordinated domain motions that govern DNA cleavage by Cas9 // *Nat. Struct. Mol. Biol.* 2019. Vol. 26, № 8. P. 679–685.
275. Chandler M. et al. Breaking and joining single-stranded DNA: the HUH endonuclease superfamily // *Nat. Rev. Microbiol.* 2013. Vol. 11, № 8. P. 525–538.
276. Guzmán-Herrador D.L., Llosa M. The secret life of conjugative relaxases // *Plasmid*. 2019. Vol. 104. P. 102415.
277. Flick K.E. et al. DNA binding and cleavage by the nuclear intron-encoded homing endonuclease I-PpoI // *Nature*. 1998. Vol. 394, № 6688. P. 96–101.
278. Kühlmann U.C. et al. Structural parsimony in endonuclease active sites: should the number of homing endonuclease families be redefined? // *FEBS Lett.* 1999. Vol. 463, № 1-2. P. 1–2.

279. Hadden J.M. et al. The structural basis of Holliday junction resolution by T7 endonuclease I // *Nature*. 2007. Vol. 449, № 7162. P. 621–624.
280. Biertümpfel C., Yang W., Suck D. Crystal structure of T4 endonuclease VII resolving a Holliday junction // *Nature*. 2007. Vol. 449, № 7162. P. 616–620.
281. Scholz S.R. et al. Experimental evidence for a beta beta alpha-Me-finger nuclease motif to represent the active site of the caspase-activated DNase // *Biochemistry*. 2003. Vol. 42, № 31. P. 9288–9294.
282. Han D.S.C., Lo Y.M.D. The Nexus of cfDNA and Nuclease Biology // *Trends Genet.* 2021. Vol. 37, № 8. P. 758–770.
283. Whitaker A.M., Freudenthal B.D. APE1: A skilled nucleic acid surgeon // *DNA Repair*. 2018. Vol. 71. P. 93–100.
284. Suck D., Lahm A., Oefner C. Structure refined to 2Å of a nicked DNA octanucleotide complex with DNase I // *Nature*. 1988. Vol. 332, № 6163. P. 464–468.
285. Mol C.D. et al. Structure and function of the multifunctional DNA-repair enzyme exonuclease III // *Nature*. 1995. Vol. 374, № 6520. P. 381–386.
286. Batebi H., Dragelj J., Imhof P. Role of AP-endonuclease (Ape1) active site residues in stabilization of the reactant enzyme-DNA complex // *Proteins*. 2018. Vol. 86, № 4. P. 439–453.
287. Mol C.D. et al. DNA-bound structures and mutants reveal abasic DNA binding by APE1 and DNA repair coordination [corrected] // *Nature*. 2000. Vol. 403, № 6768. P. 451–456.
288. Tsutakawa S.E. et al. Conserved structural chemistry for incision activity in structurally non-homologous apurinic/apyrimidinic endonuclease APE1 and

endonuclease IV DNA repair enzymes // *J. Biol. Chem.* 2013. Vol. 288, № 12. P. 8445–8455.

289. Worrall J.A.R., Luisi B.F. Information available at cut rates: structure and mechanism of ribonucleases // *Curr. Opin. Struct. Biol.* 2007. Vol. 17, № 1. P. 128–137.

290. Callaghan A.J. et al. Structure of *Escherichia coli* RNase E catalytic domain and implications for RNA turnover // *Nature*. 2005. Vol. 437, № 7062. P. 1187–1191.

291. Kaminska K.H. et al. Type II restriction endonuclease R.Hpy188I belongs to the GIY-YIG nuclease superfamily, but exhibits an unusual active site // *BMC Structural Biology*. 2008. Vol. 8, № 1. P.48.

292. Lagerbäck P. et al. Bacteriophage T4 endonuclease II, a promiscuous GIY-YIG nuclease, binds as a tetramer to two DNA substrates // *Nucleic Acids Res.* 2009. Vol. 37, № 18. P. 6174–6183.

293. Dunin-Horkawicz S., Feder M., Bujnicki J.M. Phylogenomic analysis of the GIY-YIG nuclease superfamily // *BMC Genomics*. 2006. Vol. 7. P. 98.

294. Van Roey P. et al. Catalytic domain structure and hypothesis for function of GIY-YIG intron endonuclease I-TevI // *Nat. Struct. Biol.* 2002. Vol. 9, № 11. P. 806–811.

295. Allemand F. et al. The 5S rRNA maturase, ribonuclease M5, is a Toprim domain family member // *Nucleic Acids Res.* 2005. Vol. 33, № 13. P. 4368–4376.

296. Oakley A.J. A structural view of bacterial DNA replication // *Protein Sci.* 2019. Vol. 28, № 6. P. 990–1004.

297. Wang J.C. Cellular roles of DNA topoisomerases: a molecular perspective // *Nat. Rev. Mol. Cell Biol.* 2002. Vol. 3, № 6. P. 430–440.

298. Hevener K. et al. Recent developments in topoisomerase-targeted cancer chemotherapy // *Acta Pharm Sin B*. 2018. Vol. 8, № 6. P. 844–861.

299. Sissi C., Palumbo M. Effects of magnesium and related divalent metal ions in topoisomerase structure and function // *Nucleic Acids Res.* 2009. Vol. 37, № 3. P. 702–711.
300. Garnier F., Debat H., Nadal M. Type IA DNA Topoisomerases: A Universal Core and Multiple Activities // *Methods Mol. Biol.* 2018. Vol. 1703. P. 1–20.
301. Nichols M.D. et al. Structure and function of an archaeal topoisomerase VI subunit with homology to the meiotic recombination factor Spo11 // *EMBO J.* 1999. Vol. 18, № 21. P. 6177–6188.
302. Dong K.C., Berger J.M. Structural basis for gate-DNA recognition and bending by type IIA topoisomerase IIA. // *Nature.* 2007. V. 450. P. 1201-1205.
303. Lilley D.M.J. Classification of the nucleolytic ribozymes based upon catalytic mechanism // *F1000Res.* 2019. Vol. 8. P. 1462.
304. Gottlin E.B. et al. Catalytic mechanism of the phospholipase D superfamily proceeds via a covalent phosphohistidine intermediate // *Proc. Natl. Acad. Sci. U. S. A.* 1998. Vol. 95, № 16. P. 9202–9207.
305. Sasnauskas G. et al. A novel mechanism for the scission of double-stranded DNA: BfiI cuts both 3'-5' and 5'-3' strands by rotating a single active site // *Nucleic Acids Res.* 2010. Vol. 38, № 7. P. 2399–2410.
306. Liao C. et al. Characterisation of a plancitoxin-1-like DNase II gene in *Trichinella spiralis* // *PLoS Negl. Trop. Dis.* 2014. Vol. 8, № 8. P. e3097.
307. Schmitz S. et al. Comparative analysis of two non-specific nucleases of the phospholipase D family from the plant pathogen competitor bacterium *Pantoea agglomerans* // *Appl. Microbiol. Biotechnol.* 2019. Vol. 103, № 6. P. 2635–2648.

308. Schmitz S. et al. Characterization of Single Amino Acid Variations in an EDTA-Tolerating Non-specific Nuclease from the Ice-Nucleating Bacterium *Pseudomonas syringae* // *Molecular Biotechnology*. 2020. Vol. 62, № 1. P. 67–78.
309. Nagata S. Autoimmune diseases caused by defects in clearing dead cells and nuclei expelled from erythroid precursors // *Immunol. Rev.* 2007. Vol. 220. P. 237–250.
310. Davies D.R. et al. Explorations of peptide and oligonucleotide binding sites of tyrosyl-DNA phosphodiesterase using vanadate complexes // *J. Med. Chem.* 2004. Vol. 47, № 4. P. 829–837.
311. Uesugi Y., Hatanaka T. Phospholipase D mechanism using *Streptomyces* PLD // *Biochim. Biophys. Acta.* 2009. Vol. 1791, № 9. P. 962–969.
312. Bao Y. et al. Expression and purification of BmrI restriction endonuclease and its N-terminal cleavage domain variants // *Protein Expr. Purif.* 2008. Vol. 58, № 1. P. 42–52.
313. Grazulis S. et al. Structure of the metal-independent restriction enzyme BfiI reveals fusion of a specific DNA-binding domain with a nonspecific nuclease // *Proc. Natl. Acad. Sci. U. S. A.* 2005. Vol. 102, № 44. P. 15797–15802.
314. Schwardmann L.S., Nölle V., Elleuche S. Bacterial non-specific nucleases of the phospholipase D superfamily and their biotechnological potential // *Appl. Microbiol. Biotechnol.* 2020. Vol. 104, № 8. P. 3293–3304.
315. Cascales E. et al. Colicin biology // *Microbiol. Mol. Biol. Rev.* 2007. Vol. 71, № 1. P. 158–229.
316. Ogawa T. et al. A cytotoxic ribonuclease targeting specific transfer RNA anticodons // *Science.* 1999. Vol. 283, № 5410. P. 2097–2100.

317. Tomita K. et al. A cytotoxic ribonuclease which specifically cleaves four isoaccepting arginine tRNAs at their anticodon loops // *Proc. Natl. Acad. Sci. U. S. A.* 2000. Vol. 97, № 15. P. 8278–8283.
318. Lin Y.-L., Elias Y., Huang R.H. Structural and mutational studies of the catalytic domain of colicin E5: a tRNA-specific ribonuclease // *Biochemistry.* 2005. Vol. 44, № 31. P. 10494–10500.
319. Graille M. et al. Structural inhibition of the colicin D tRNase by the tRNA-mimicking immunity protein // *EMBO J.* 2004. Vol. 23, № 7. P. 1474–1482.
320. Cuthbert B.J., Burley K.H., Goulding C.W. Introducing the new bacterial branch of the RNase A superfamily // *RNA Biol.* 2018. Vol. 15, № 1. P. 9–12.
321. Dyer K.D., Rosenberg H.F. The RNase a superfamily: generation of diversity and innate host defense // *Mol. Divers.* 2006. Vol. 10, № 4. P. 585–597.
322. Rosenberg H.F. RNase A ribonucleases and host defense: an evolving story // *J. Leukoc. Biol.* 2008. Vol. 83, № 5. P. 1079–1087.
323. Cuchillo C.M., Nogués M.V., Raines R.T. Bovine pancreatic ribonuclease: fifty years of the first enzymatic reaction mechanism // *Biochemistry.* 2011. Vol. 50, № 37. P. 7835–7841.
324. Yakovleva L. et al. Chemical and traditional mutagenesis of vaccinia DNA topoisomerase provides insights to cleavage site recognition and transesterification chemistry // *J. Biol. Chem.* 2008. Vol. 283, № 23. P. 16093–16103.
325. Stark W.M. The Serine Recombinases // *Microbiol Spectr.* 2014. Vol. 2, № 6. P. 46.
326. Wang J.-Y. et al. Artificial nondirectional site-specific recombination systems // *iScience.* 2022. Vol. 25, № 1. P. 103716.

327. Keenholtz R.A. et al. Arginine as a general acid catalyst in serine recombinase-mediated DNA cleavage // *J. Biol. Chem.* 2013. Vol. 288, № 40. P. 29206–29214.
328. Yang W. Topoisomerases and site-specific recombinases: similarities in structure and mechanism // *Crit. Rev. Biochem. Mol. Biol.* 2010. Vol. 45, № 6. P. 520–534.
329. Turan S., Bode J. Site-specific recombinases: from tag-and-target- to tag-and-exchange-based genomic modifications // *FASEB J.* 2011. Vol. 25, № 12. P. 4088–4107.
330. Belkebir A., Azeddoug H. Characterization of LlaKI, a New Metal Ion-Independent Restriction Endonuclease from *Lactococcus lactis* KLDS4 // *ISRN Biochem.* 2012. Vol. 2012. P. 287230.
331. Witteveldt J. et al. The influence of viral RNA secondary structure on interactions with innate host cell defences // *Nucleic Acids Research.* 2014. Vol. 42, № 5. P. 3314–3329.
332. Li L., Lin S., Yanga F. Functional identification of the non-specific nuclease from white spot syndrome virus // *Virology.* 2005. Vol. 337, № 2. P. 399–406.
333. Song Q., Zhang X. Characterization of a novel non-specific nuclease from thermophilic bacteriophage GBSV1 // *BMC Biotechnology.* 2008. Vol. 8, № 1. P. 43.
334. Lilley D.M.J. Structure, folding and mechanisms of ribozymes // *Current Opinion in Structural Biology.* 2005. Vol. 15, № 3. P. 313–323.
335. Peng H. et al. Self-cleaving ribozymes: substrate specificity and synthetic biology applications // *RSC Chem Biol.* 2021. Vol. 2, № 5. P. 1370–1383.
336. Weinberg C.E., Weinberg Z., Hammann C. Novel ribozymes: discovery, catalytic mechanisms, and the quest to understand biological function // *Nucleic Acids Res.* 2019. Vol. 47, № 18. P. 9480–9494.

337. Ke A. et al. A conformational switch controls hepatitis delta virus ribozyme catalysis // *Nature*. 2004. Vol. 429, № 6988. P. 201–205.
338. Toor N., Keating K.S., Pyle A.M. Structural insights into RNA splicing // *Curr. Opin. Struct. Biol.* 2009. Vol. 19, № 3. P. 260–266.
339. Teplova M. et al. Crucial Roles of Two Hydrated Mg Ions in Reaction Catalysis of the Pistol Ribozyme // *Angew. Chem. Int. Ed Engl.* 2020. Vol. 59, № 7. P. 2837–2843.
340. Micura R., Höbartner C. Fundamental studies of functional nucleic acids: aptamers, riboswitches, ribozymes and DNAzymes // *Chem. Soc. Rev.* 2020. Vol. 49, № 20. P. 7331–7353.
341. Silverman S.K. In vitro selection, characterization, and application of deoxyribozymes that cleave RNA // *Nucleic Acids Res.* 2005. Vol. 33, № 19. P. 6151–6163.
342. Sheppard T.L., Ordoukhanian P., Joyce G.F. A DNA enzyme with N-glycosylase activity // *Proc. Natl. Acad. Sci. U. S. A.* 2000. Vol. 97, № 14. P. 7802–7807.
343. Chandra M., Sachdeva A., Silverman S.K. DNA-catalyzed sequence-specific hydrolysis of DNA // *Nat. Chem. Biol.* 2009. Vol. 5, № 10. P. 718–720.
344. Dhamodharan V., Kobori S., Yokobayashi Y. Large Scale Mutational and Kinetic Analysis of a Self-Hydrolyzing Deoxyribozyme // *ACS Chem. Biol.* 2017. Vol. 12, № 12. P. 2940–2945.
345. Gu H. et al. Small, highly active DNAs that hydrolyze DNA // *J. Am. Chem. Soc.* 2013. Vol. 135, № 24. P. 9121–9129.
346. Schlosser K. et al. In vitro selection of small RNA-cleaving deoxyribozymes that cleave pyrimidine-pyrimidine junctions // *Nucleic Acids Res.* 2008. Vol. 36, № 14. P. 4768–4777.

347. Abrosimova L.A. et al. Kinetic Analysis of the Interaction of Nicking Endonuclease BspD6I with DNA // *Biomolecules*. 2021. Vol. 11, № 10. P. 1420.
348. Shi L. et al. Development of a DNA microarray-based multiplex assay of avian influenza virus subtypes H5, H7, H9, N1, and N2 // *Acta Virol*. 2014. Vol. 58, № 1. P. 14–19.
349. Murakami T., Sumaoka J., Komiyama M. Sensitive isothermal detection of nucleic-acid sequence by primer generation-rolling circle amplification // *Nucleic Acids Res*. 2009. Vol. 37, № 3. P. e19.
350. Hosoda K. et al. A novel sequence-specific RNA quantification method using nicking endonuclease, dual-labeled fluorescent DNA probe, and conformation-interchangeable oligo-DNA // *RNA*. 2008. Vol. 14, № 3. P. 584–592.
351. Wu D. et al. A label-free colorimetric isothermal cascade amplification for the detection of disease-related nucleic acids based on double-hairpin molecular beacon // *Anal. Chim. Acta*. 2017. Vol. 957. P. 55–62.
352. Zhang P. et al. Engineering BspQI nicking enzymes and application of N.BspQI in DNA labeling and production of single-strand DNA // *Protein Expression and Purification*. 2010. Vol. 69, № 2. P. 226–234.
353. Zou B. et al. Ultrasensitive DNA detection by cascade enzymatic signal amplification based on Afu flap endonuclease coupled with nicking endonuclease // *Angew. Chem. Int. Ed Engl*. 2011. Vol. 50, № 32. P. 7395–7398.
354. Gaj T., Gersbach C.A., Barbas C.F. 3rd. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering // *Trends Biotechnol*. 2013. Vol. 31, № 7. P. 397–405.
355. Pan Y. et al. Biological and Biomedical Applications of Engineered Nucleases // *Molecular Biotechnology*. 2013. Vol. 55, № 1. P. 54–62.

356. Niazian M. Application of genetics and biotechnology for improving medicinal plants // *Planta*. 2019. Vol. 249, № 4. P. 953–973.
357. Zhu X. et al. An efficient genotyping method for genome-modified animals and human cells generated with CRISPR/Cas9 system // *Sci. Rep.* 2014. Vol. 4. P. 6420.
358. Zhu H., Li C., Gao C. Applications of CRISPR–Cas in agriculture and plant biotechnology // *Nature Reviews Molecular Cell Biology*. 2020. Vol. 21, № 11. P. 661–677.
359. Knott G.J., Doudna J.A. CRISPR-Cas guides the future of genetic engineering // *Science*. 2018. Vol. 361, № 6405. P. 866–869.
360. Saha S.K. et al. Programmable Molecular Scissors: Applications of a New Tool for Genome Editing in Biotech // *Mol. Ther. Nucleic Acids*. 2019. Vol. 14. P. 212–238.
361. Alwin S. et al. Custom zinc-finger nucleases for use in human cells // *Mol. Ther.* 2005. Vol. 12, № 4. P. 610–617.
362. Bibikova M. et al. Targeted chromosomal cleavage and mutagenesis in *Drosophila* using zinc-finger nucleases // *Genetics*. 2002. Vol. 161, № 3. P. 1169–1175.
363. Durai S. et al. Zinc finger nucleases: custom-designed molecular scissors for genome engineering of plant and mammalian cells // *Nucleic Acids Res.* 2005. Vol. 33, № 18. P. 5978–5990.
364. Porteus M.H., Carroll D. Gene targeting using zinc finger nucleases // *Nat. Biotechnol.* 2005. Vol. 23, № 8. P. 967–973.
365. Porteus M.H. Mammalian gene targeting with designed zinc finger nucleases // *Mol. Ther.* 2006. Vol. 13, № 2. P. 438–446.

366. Pagant S. et al. ZFN-mediated in vivo gene editing in hepatocytes leads to supraphysiologic α -Gal A activity and effective substrate reduction in Fabry mice // *Mol. Ther.* 2021. Vol. 29, № 11. P. 3230–3242.
367. Urnov F.D. et al. Highly efficient endogenous human gene correction using designed zinc-finger nucleases // *Nature*. 2005. Vol. 435, № 7042. P. 646–651.
368. Kleinstiver B.P. et al. Monomeric site-specific nucleases for genome editing // *Proc. Natl. Acad. Sci. U. S. A.* 2012. Vol. 109, № 21. P. 8061–8066.
369. Schierling B. et al. A novel zinc-finger nuclease platform with a sequence-specific cleavage module // *Nucleic Acids Res.* 2012. Vol. 40, № 6. P. 2623–2638.
370. Yanik M. et al. TALE-PvuII fusion proteins--novel tools for gene targeting // *PLoS One*. 2013. Vol. 8, № 12. P. e82539.
371. Becker S., Boch J. TALE and TALEN genome editing technologies // *Gene and Genome Editing*. 2021. Vol. 2. P. 100007.
372. Anzalone A.V., Koblan L.W., Liu D.R. Genome editing with CRISPR-Cas nucleases, base editors, transposases and prime editors // *Nat. Biotechnol.* 2020. Vol. 38, № 7. P. 824–844.
373. Xu R. et al. Gene targeting using the *Agrobacterium tumefaciens*-mediated CRISPR-Cas system in rice // *Rice*. 2014. Vol. 7, № 1. P. 5.
374. Zhou Z. et al. CRISPR/Cas9-mediated efficient targeted mutagenesis of RAS in *Salvia miltiorrhiza* // *Phytochemistry*. 2018. Vol. 148. P. 63–70.
375. Feng S. et al. Application of the CRISPR/Cas9 system in *Dioscorea zingiberensis* // *Plant Cell, Tissue and Organ Culture (PCTOC)*. 2018. Vol. 135, № 1. P. 133–141.
376. Li X.-Y. et al. A selective knockout method for discovery of minor active components from plant extracts: Feasibility and challenges as illustrated by an application

- to *Salvia miltiorrhiza* // *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 2017. Vol. 1068-1069. P. 253–260.
377. Gu W. et al. Depletion of Abundant Sequences by Hybridization (DASH): using Cas9 to remove unwanted high-abundance species in sequencing libraries and molecular counting applications // *Genome Biol.* 2016. Vol. 17. P. 41.
378. Gaudelli N.M. et al. Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage // *Nature.* 2017. Vol. 551, № 7681. P. 464–471.
379. Komor A.C. et al. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage // *Nature.* 2016. Vol. 533, № 7603. P. 420–424.
380. Kurt I.C. et al. CRISPR C-to-G base editors for inducing targeted DNA transversions in human cells // *Nat. Biotechnol.* 2021. Vol. 39, № 1. P. 41–46.
381. Ryu S.-M. et al. Adenine base editing in mouse embryos and an adult mouse model of Duchenne muscular dystrophy // *Nat. Biotechnol.* 2018. Vol. 36, № 6. P. 536–539.
382. Zeballos C M.A., Gaj T. Next-Generation CRISPR Technologies and Their Applications in Gene and Cell Therapy // *Trends Biotechnol.* 2021. Vol. 39, № 7. P. 692–705.
383. Binnie A. et al. CRISPR-based strategies in infectious disease diagnosis and therapy // *Infection.* 2021. Vol. 49, № 3. P. 377–385.
384. Song A.J., Palmiter R.D. Detecting and Avoiding Problems When Using the Cre-lox System // *Trends in Genetics.* 2018. Vol. 34, № 5. P. 333–340.
385. van der Sloot A., Tyers M. Synthetic Genomics: Rewriting the Genome Chromosome by Chromosome // *Mol. Cell.* 2017. Vol. 66, № 4. P. 441–443.
386. Shen Y. et al. SCRaMbLE generates designed combinatorial stochastic diversity in synthetic chromosomes // *Genome Res.* 2016. Vol. 26, № 1. P. 36–49.

387. Howard J.T. et al. Heteroduplex cleavage analysis using S1 nuclease // *Biotechniques*. 1999. Vol. 27, № 1. P. 18–19.
388. Orita M. et al. Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction // *Genomics*. Elsevier BV, 1989. Vol. 5, № 4. P. 874–879.
389. Shishido K., Ando T. [S1 nuclease] // *Tanpakushitsu Kakusan Koso*. 1985. Vol. 30, № 8. P. 981–987.
390. Qiu P. et al. Mutation detection using Surveyor nuclease // *Biotechniques*. 2004. Vol. 36, № 4. P. 702–707.
391. Balagurumoorthy P., Adelstein S.J., Kassis A.I. Method to eliminate linear DNA from mixture containing nicked circular, supercoiled, and linear plasmid DNA // *Anal. Biochem*. 2008. Vol. 381, № 1. P. 172–174.
392. Prazeres D. M. F. Considerations on the Use of Enzymes in the Downstream Processing of Biopharmaceuticals // *Pharmaceutical Bioprocessing*. 2016. Vol. 4, №. 5. P. 95-99.
393. Shishido K., Ando T. Efficiency of T4 DNA ligase-catalyzed end joining after S1 endonuclease treatment on duplex DNA containing single-stranded portions // *Biochim. Biophys. Acta*. 1981. Vol. 656, № 1. P. 123–127.
394. Bandeira V. et al. Downstream processing of lentiviral vectors: releasing bottlenecks // *Hum. Gene Ther. Methods*. 2012. Vol. 23, № 4. P. 255–263.
395. Konz J.O. et al. Development of a purification process for adenovirus: controlling virus aggregation to improve the clearance of host cell DNA // *Biotechnol. Prog*. 2005. Vol. 21, № 2. P. 466–472.

396. Shaw A. et al. Using Pulmozyme DNase treatment in lentiviral vector production // *Hum. Gene Ther. Methods*. 2012. Vol. 23, № 1. P. 65–71.
397. Steppert P. et al. Purification of HIV-1 gag virus-like particles and separation of other extracellular particles // *J. Chromatogr. A*. 2016. Vol. 1455. P. 93–101.
398. Sastry L. et al. Evaluation of plasmid DNA removal from lentiviral vectors by benzonase treatment // *Hum. Gene Ther.* 2004. Vol. 15, № 2. P. 221–226.
399. Ausubel L.J. et al. Production of CGMP-Grade Lentiviral Vectors // *Bioprocess Int.* 2012. Vol. 10, № 2. P. 32–43.
400. Sawazaki K. et al. Deoxyribonuclease I (DNase I) typing from semen stains: low enzyme activity in vaginal fluids does not interfere with seminal DNase I typing from mixture stains // *J. Forensic Sci.* 1993. Vol. 38, № 5. P. 1051–1062.
401. Fuchs H.J. et al. Effect of aerosolized recombinant human DNase on exacerbations of respiratory symptoms and on pulmonary function in patients with cystic fibrosis. The Pulmozyme Study Group // *N. Engl. J. Med.* 1994. Vol. 331, № 10. P. 637–642.
402. Kishi K., Yasuda T., Takeshita H. DNase I: structure, function, and use in medicine and forensic science // *Leg. Med.* . Elsevier BV, 2001. Vol. 3, № 2. P. 69–83.
403. Tsutsumi S. et al. Phenotype 2 of deoxyribonuclease I may be used as a risk factor for gastric carcinoma // *Cancer*. 1998. Vol. 82, № 9. P. 1621–1625.
404. Balian A., Hernandez F.J. Nucleases as molecular targets for cancer diagnosis // *Biomark Res.* 2021. Vol. 9, № 1. P. 86.
405. Thomas A., Pommier Y. Targeting Topoisomerase I in the Era of Precision Medicine // *Clin. Cancer Res.* 2019. Vol. 25, № 22. P. 6581–6589.

406. Carninci P. et al. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes // *Genome Res.* 2000. Vol. 10, № 10. P. 1617–1630.
407. Gadgil C. et al. A mathematical model for suppression subtractive hybridization // *Comp. Funct. Genomics.* 2002. Vol. 3, № 5. P. 405–422.
408. Hastie N.D., Bishop J.O. The expression of three abundance classes of messenger RNA in mouse tissues // *Cell.* 1976. Vol. 9, № 4 PT 2. P. 761–774.
409. Hoang S.A. et al. Gene Expression Predicts Histological Severity and Reveals Distinct Molecular Profiles of Nonalcoholic Fatty Liver Disease // *Sci. Rep.* 2019. Vol. 9, № 1. P. 12541.
410. Soares M.B. et al. Construction and characterization of a normalized cDNA library // *Proc. Natl. Acad. Sci. U. S. A.* 1994. Vol. 91, № 20. P. 9228–9232.
411. Mita K. et al. The construction of an EST database for *Bombyx mori* and its application // *Proc. Natl. Acad. Sci. U. S. A.* 2003. Vol. 100, № 24. P. 14121–14126.
412. Athanasopoulou K. et al. Third-Generation Sequencing: The Spearhead towards the Radical Transformation of Modern Genomics // *Life.* 2021. Vol. 12, № 1. P. 30.
413. Goodwin S., McPherson J.D., McCombie W.R. Coming of age: ten years of next-generation sequencing technologies // *Nat. Rev. Genet.* 2016. Vol. 17, № 6. P. 333–351.
414. van Dijk E.L. et al. The Third Revolution in Sequencing Technology // *Trends Genet.* 2018. Vol. 34, № 9. P. 666–681.
415. Johnson L.K., Alexander H., Brown C.T. Re-assembly, quality evaluation, and annotation of 678 microbial eukaryotic reference transcriptomes // *Gigascience.* 2019. Vol. 8, № 4. P. giy158.

416. Chamala S. et al. Genome-wide identification of evolutionarily conserved alternative splicing events in flowering plants // *Front Bioeng Biotechnol.* 2015. Vol. 3. P. 33.
417. Huang C.-K., Lin W.-D., Wu S.-H. An improved repertoire of splicing variants and their potential roles in *Arabidopsis* photomorphogenic development // *Genome Biol.* 2022. Vol. 23, № 1. P. 50.
418. Lee Y., Rio D.C. Mechanisms and Regulation of Alternative Pre-mRNA Splicing // *Annual Review of Biochemistry.* 2015. Vol. 84, № 1. P. 291–323.
419. Makałowski W. et al. Transposable Elements: Classification, Identification, and Their Use As a Tool For Comparative Genomics // *Methods Mol. Biol.* 2019. Vol. 1910. P. 177–207.
420. Richard G.-F., Kerrest A., Dujon B. Comparative genomics and molecular dynamics of DNA repeats in eukaryotes // *Microbiol. Mol. Biol. Rev.* 2008. Vol. 72, № 4. P. 686–727.
421. Ekblom R., Galindo J. Applications of next generation sequencing in molecular ecology of non-model organisms // *Heredity.* 2011. Vol. 107, № 1. P. 1–15.
422. Liao Y., Smyth G.K., Shi W. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads // *Nucleic Acids Res.* 2019. Vol. 47, № 8. P. e47.
423. Ungaro A. et al. Challenges and advances for transcriptome assembly in non-model species // *PLoS One.* 2017. Vol. 12, № 9. P. e0185020.
424. Maricic T., Whitten M., Pääbo S. Multiplexed DNA sequence capture of mitochondrial genomes using PCR products // *PLoS One.* 2010. Vol. 5, № 11. P. e14004.

425. Cruz-Dávalos D.I. et al. In-solution Y-chromosome capture-enrichment on ancient DNA libraries // *BMC Genomics*. 2018. Vol. 19, № 1. P. 608.
426. Enk J.M. et al. Ancient whole genome enrichment using baits built from modern DNA // *Mol. Biol. Evol.* 2014. Vol. 31, № 5. P. 1292–1294.
427. Lazaridis I. et al. Genomic insights into the origin of farming in the ancient Near East // *Nature*. 2016. Vol. 536, № 7617. P. 419–424.
428. Di Resta C. et al. Next-generation sequencing approach for the diagnosis of human diseases: open challenges and new opportunities // *EJIFCC*. 2018. Vol. 29, № 1. P. 4–14.
429. Grada A., Weinbrecht K. Next-generation sequencing: methodology and application // *J. Invest. Dermatol.* 2013. Vol. 133, № 8. P. e11.
430. Williams E.S., Hegde M. Implementing genomic medicine in pathology // *Adv. Anat. Pathol.* 2013. Vol. 20, № 4. P. 238–244.
431. Emberton J. et al. Gene enrichment in maize with hypomethylated partial restriction (HMPR) libraries // *Genome Res.* 2005. Vol. 15, № 10. P. 1441–1446.
432. Larsson H. et al. The Hypomethylated Partial Restriction (HMPR) method reduces the repetitive content of genomic libraries in Norway spruce (*Picea abies*) // *Tree Genetics & Genomes*. 2013. Vol. 9, № 2. P. 601–612.
433. Baird N.A. et al. Rapid SNP discovery and genetic mapping using sequenced RAD markers // *PLoS One*. 2008. Vol. 3, № 10. P. e3376.
434. Peterson B.K. et al. Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-Model Species // *PLoS ONE*. 2012. Vol. 7, № 5. P. e37135.

435. van Orsouw N.J. et al. Complexity reduction of polymorphic sequences (CRoPS): a novel approach for large-scale polymorphism discovery in complex genomes // *PLoS One*. 2007. Vol. 2, № 11. P. e1172.
436. Elshire R.J. et al. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species // *PLoS One*. 2011. Vol. 6, № 5. P. e19379.
437. Michelini F. et al. From “Cellular” RNA to “Smart” RNA: Multiple Roles of RNA in Genome Stability and Beyond // *Chemical Reviews*. 2018. Vol. 118, № 8. P. 4365–4403.
438. Britten R. J., Davidson E. H. Hybridization strategy // *Nucleic acid hybridisation*. 1985. P. 3-15.
439. Young BD, Anderson M. Quantitative analysis of solution hybridisation. In *Nucleic Acids Hybridisation, a practical approach* (eds. Hames, B.D. and Higgins, S.J.) IRL Press, Oxford-Washington DC. 1985; 47-71
440. Zhang D.Y., Chen S.X., Yin P. Optimizing the specificity of nucleic acid hybridization // *Nat. Chem*. 2012. Vol. 4, № 3. P. 208–214.
441. Britten R.J., Kohne D.E. Repeated sequences in DNA. Hundreds of thousands of copies of DNA sequences have been incorporated into the genomes of higher organisms // *Science*. 1968. Vol. 161, № 3841. P. 529–540.
442. Peterson D.G., Pearson W.R., Stack S.M. Characterization of the tomato (*Lycopersicon esculentum*) genome using in vitro and in situ DNA reassociation // *Genome*. 1998. Vol. 41, № 3. P. 346–356.
443. Peterson D.G. et al. Integration of Cot analysis, DNA cloning, and high-throughput sequencing facilitates genome characterization and gene discovery // *Genome Res*. 2002. Vol. 12, № 5. P. 795–807.

444. Paterson A.H. Leafing through the genomes of our major crop plants: strategies for capturing unique information // *Nat. Rev. Genet.* 2006. Vol. 7, № 3. P. 174–184.
445. Ko M.S. An “equalized cDNA library” by the reassociation of short double-stranded cDNAs // *Nucleic Acids Res.* 1990. Vol. 18, № 19. P. 5705–5711.
446. Patanjali S.R., Parimoo S., Weissman S.M. Construction of a uniform-abundance (normalized) cDNA library // *Proc. Natl. Acad. Sci. U. S. A.* 1991. Vol. 88, № 5. P. 1943–1947.
447. Vandernoot V.A. et al. cDNA normalization by hydroxyapatite chromatography to enrich transcriptome diversity in RNA-seq applications // *Biotechniques.* 2012. Vol. 53, № 6. P. 373–380.
448. Bonaldo M.F., Lennon G., Soares M.B. Normalization and subtraction: two approaches to facilitate gene discovery // *Genome Res.* 1996. Vol. 6, № 9. P. 791–806.
449. Sasaki Y.F., Ayusawa D., Oishi M. Construction of a normalized cDNA library by introduction of a semi-solid mRNA-cDNA hybridization system // *Nucleic Acids Res.* 1994. Vol. 22, № 6. P. 987–992.
450. Coche T.G. Subtractive cDNA cloning using magnetic beads and PCR // *Methods Mol. Biol.* 1997. Vol. 67. P. 371–387.
451. Coche T., Dewez M. Reducing bias in cDNA sequence representation by molecular selection // *Nucleic Acids Res.* 1994. Vol. 22, № 21. P. 4545–4546.
452. Siebert P.D. et al. An improved PCR method for walking in uncloned genomic DNA // *Nucleic Acids Res.* 1995. Vol. 23, № 6. P. 1087–1088.
453. К. А. Лукьянов, Н. Г. Гурская, Е. А. Богданова, С. А. Лукьянов. Селективная супрессия полимеразной цепной реакции // *Биоорганическая Химия.* 1999. Т. 25, № 3. С. 163-170.

454. Adiconis X. et al. Comparative analysis of RNA sequencing methods for degraded or low-input samples // *Nat. Methods*. 2013. Vol. 10, № 7. P. 623–629.
455. von der Haar T. A quantitative estimation of the global translational activity in logarithmically growing yeast cells // *BMC Syst. Biol.* 2008. Vol. 2. P. 87.
456. Wilhelm B.T., Landry J.-R. RNA-Seq-quantitative measurement of expression through massively parallel RNA-sequencing // *Methods*. 2009. Vol. 48, № 3. P. 249–257.
457. O’Neil D., Glowatz H., Schlumpberger M. Ribosomal RNA depletion for efficient use of RNA-seq capacity // *Curr. Protoc. Mol. Biol.* 2013. Vol. 103. P. 4.19.1-8.
458. Byrne A. et al. Realizing the potential of full-length transcriptome sequencing // *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 2019. Vol. 374, № 1786. P. 20190097.
459. Graveley B.R. Alternative splicing: increasing diversity in the proteomic world // *Trends Genet. Elsevier BV*, 2001. Vol. 17, № 2. P. 100–107.
460. Ilagan J.O. et al. U2AF1 mutations alter splice site recognition in hematological malignancies // *Genome Res.* 2015. Vol. 25, № 1. P. 14–26.
461. Zhang W. et al. Isoform Switch of TET1 Regulates DNA Demethylation and Mouse Development // *Mol. Cell.* 2016. Vol. 64, № 6. P. 1062–1073.
462. Du H. et al. Comparative Transcriptome Analysis of Resistant and Susceptible Tomato Lines in Response to Infection by *Xanthomonas perforans* Race T3 // *Front. Plant Sci.* 2015. Vol. 6. P. 1173.
463. Hernández-Fernández J., Pinzón A., Mariño-Ramírez L. De novo transcriptome assembly of loggerhead sea turtle nesting of the Colombian Caribbean // *Genom Data.* 2017. Vol. 13. P. 18–20.
464. Huang R. et al. Transcriptome-wide discovery of coding and noncoding RNA-binding proteins // *Proc. Natl. Acad. Sci. U. S. A.* 2018. Vol. 115, № 17. P. E3879–E3887.

465. Mastrokolas A. et al. Increased sensitivity of next generation sequencing-based expression profiling after globin reduction in human blood RNA // *BMC Genomics*. 2012. Vol. 13. P. 28.
466. Gu J. et al. GFR α 2 prompts cell growth and chemoresistance through down-regulating tumor suppressor gene PTEN via Mir-17-5p in pancreatic cancer // *Cancer Lett*. 2016. Vol. 380, № 2. P. 434–441.
467. Shin K. et al. Cellular origin of bladder neoplasia and tissue dynamics of its progression to invasive carcinoma // *Nat. Cell Biol*. 2014. Vol. 16, № 5. P. 469–478.
468. Sultan M. et al. Influence of RNA extraction methods and library selection schemes on RNA-seq data // *BMC Genomics*. 2014. Vol. 15. P. 675.
469. Dieci G. et al. The expanding RNA polymerase III transcriptome // *Trends Genet*. 2007. Vol. 23, № 12. P. 614–622.
470. Kraus A.J., Brink B.G., Siegel T.N. Efficient and specific oligo-based depletion of rRNA // *Sci. Rep*. 2019. Vol. 9, № 1. P. 12281.
471. Yang L. et al. Genomewide characterization of non-polyadenylated RNAs // *Genome Biol*. 2011. Vol. 12, № 2. P. R16.
472. Dai X. et al. H2A.Z Represses Gene Expression by Modulating Promoter Nucleosome Structure and Enhancer Histone Modifications in Arabidopsis // *Molecular Plant*. 2018. Vol. 11, № 4. P. 635.
473. Lakey D.L. et al. Priming reverse transcription with oligo(dT) does not yield representative samples of *Mycobacterium tuberculosis* cDNA // *Microbiology*. 2002. Vol. 148, № 8. P. 2567–2572.

474. Mullins M. et al. Agreement in breast cancer classification between microarray and quantitative reverse transcription PCR from fresh-frozen and formalin-fixed, paraffin-embedded tissues // *Clin. Chem.* 2007. Vol. 53, № 7. P. 1273–1279.
475. Armour C.D. et al. Digital transcriptome profiling using selective hexamer priming for cDNA synthesis // *Nat. Methods.* 2009. Vol. 6, № 9. P. 647–649.
476. Arnaud O. et al. Targeted reduction of highly abundant transcripts using pseudo-random primers // *Biotechniques.* 2016. Vol. 60, № 4. P. 169–174.
477. Cui P. et al. A comparison between ribo-minus RNA-sequencing and polyA-selected RNA-sequencing // *Genomics.* 2010. Vol. 96, № 5. P. 259–265.
478. Huang R. et al. An RNA-Seq strategy to detect the complete coding and non-coding transcriptome including full-length imprinted macro ncRNAs // *PLoS One.* 2011. Vol. 6, № 11. P. e27288.
479. Ingolia N.T. et al. The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments // *Nat. Protoc.* 2012. Vol. 7, № 8. P. 1534–1550.
480. Field L.A. et al. Functional identity of genes detectable in expression profiling assays following globin mRNA reduction of peripheral blood samples // *Clin. Biochem.* 2007. Vol. 40, № 7. P. 499–502.
481. Matranga C.B. et al. Enhanced methods for unbiased deep sequencing of Lassa and Ebola RNA viruses from clinical and biological samples // *Genome Biol.* 2014. Vol. 15, № 11. P. 519.
482. Morlan J.D., Qu K., Sinicropi D.V. Selective depletion of rRNA enables whole transcriptome profiling of archival fixed tissue // *PLoS One.* 2012. Vol. 7, № 8. P. e42882.

483. Rosenow C. et al. Prokaryotic RNA preparation methods useful for high density array analysis: comparison of two approaches // *Nucleic Acids Res.* 2001. Vol. 29, № 22. P. E112.
484. Munafó D.B. et al. Selective Depletion of Abundant RNAs to Enable Transcriptome Analysis of Low-Input and Highly Degraded Human RNA // *Curr. Protoc. Mol. Biol.* 2016. Vol. 113, № 1. P. 7.22.1–7.22.9.
485. Altshuler D. et al. A haplotype map of the human genome // *Nature.* 2005. Vol. 437, №. 7063. P.1299–320.
486. Rocha D. et al. Seventh international meeting on single nucleotide polymorphism and complex genome analysis: “ever bigger scans and an increasingly variable genome” // *Human Genetics.* 2006. Vol. 119, № 4. P. 451–456.
487. Kim S., Misra A. SNP genotyping: technologies and biomedical applications // *Annu. Rev. Biomed. Eng.* 2007. Vol. 9. P. 289–320.
488. Jian Y., Li M. A narrative review of single-nucleotide polymorphism detection methods and their application in studies of *Staphylococcus aureus* // *Journal of Bio-X Research.* 2021. Vol. 4, № 1. P. 1–9.
489. Mehta B., Daniel R., McNevin D. HRM and SNaPshot as alternative forensic SNP genotyping methods // *Forensic Sci. Med. Pathol.* 2017. Vol. 13, № 3. P. 293–301.
490. Gomes S. et al. Alternative SNP detection platforms, HRM and biosensors, for varietal identification in *Vitis vinifera* L. using F3H and LDOX genes // *Sci. Rep.* 2018. Vol. 8, № 1. P. 5850.
491. Jaakola L., Suokas M., Häggman H. Novel approaches based on DNA barcoding and high-resolution melting of amplicons for authenticity analyses of berry species // *Food Chemistry.* 2010. Vol. 123, № 2. P. 494–500.

492. Sun W. et al. The Potential Power of Bar-HRM Technology in Herbal Medicine Identification // *Front. Plant Sci.* 2016. Vol. 7. P. 367.
493. Tyagi S., Bratu D.P., Kramer F.R. Multicolor molecular beacons for allele discrimination // *Nat. Biotechnol.* 1998. Vol. 16, № 1. P. 49–53.
494. Myakishev M.V. et al. High-throughput SNP genotyping by allele-specific PCR with universal energy-transfer-labeled primers // *Genome Res.* 2001. Vol. 11, № 1. P. 163–169.
495. Bidar N. et al. Monitoring of microRNA using molecular beacons approaches: Recent advances // *TrAC Trends in Analytical Chemistry.* 2020. Vol. 131. P. 116021.
496. Bidar N. et al. Molecular beacon strategies for sensing purpose // *TrAC Trends in Analytical Chemistry.* 2021. Vol. 134. P. 116143.
497. Zuo Z., Jabbar K.J. COLD-PCR: Applications and Advantages // *Methods Mol. Biol.* 2016. Vol. 1392. P. 17–25.
498. LaFramboise T. Single nucleotide polymorphism arrays: a decade of biological, computational and technological advances // *Nucleic Acids Res.* 2009. Vol. 37, № 13. P. 4181–4193.
499. Sokolov B. P. Primer extension technique for the detection of single nucleotide in genomic DNA // *Nucleic Acids Research.* 1990. Vol. 18, №. 12. P. 3671.
500. Ross P. et al. High level multiplex genotyping by MALDI-TOF mass spectrometry // *Nat. Biotechnol.* 1998. Vol. 16, № 13. P. 1347–1351.
501. Fei Z., Ono T., Smith L.M. MALDI-TOF mass spectrometric typing of single nucleotide polymorphisms with mass-tagged ddNTPs // *Nucleic Acids Res.* 1998. Vol. 26, № 11. P. 2827–2828.

502. Braun A., Little D.P., Köster H. Detecting CFTR gene mutations by using primer oligo base extension and mass spectrometry // *Clin. Chem.* 1997. Vol. 43, № 7. P. 1151–1158.
503. Kim S. et al. Solid phase capturable dideoxynucleotides for multiplex genotyping using mass spectrometry // *Nucleic Acids Res.* 2002. Vol. 30, № 16. P. e85.
504. Kim S. et al. Thirtyfold multiplex genotyping of the p53 gene using solid phase capturable dideoxynucleotides and mass spectrometry // *Genomics.* 2004. Vol. 83, № 5. P. 924–931.
505. Park J.H. et al. A mass spectrometry-based multiplex SNP genotyping by utilizing allele-specific ligation and strand displacement amplification // *Biosens. Bioelectron.* 2017. Vol. 91. P. 122–127.
506. Sauer S. et al. Facile method for automated genotyping of single nucleotide polymorphisms by mass spectrometry // *Nucleic Acids Res.* 2002. Vol. 30, № 5. P. e22.
507. Le Hellard S. et al. SNP genotyping on pooled DNAs: comparison of genotyping technologies and a semi automated method for data storage and analysis // *Nucleic Acids Res.* 2002. Vol. 30, № 15. P. e74.
508. Pastinen T. et al. Minisequencing: a specific tool for DNA analysis and diagnostics on oligonucleotide arrays // *Genome Res.* 1997. Vol. 7, № 6. P. 606–614.
509. Bell P.A. et al. SNPstream® UHT: Ultra-High Throughput SNP Genotyping for Pharmacogenomics and Drug Discovery // *BioTechniques.* 2002. Vol. 32, № 6S. P. S70–S77.
510. Budiarto B.R., Pohan P.U., Desriani. Nucleic acid amplification-based HER2 molecular detection for breast cancer // *Journal of Oncological Sciences.* 2019. Vol. 5, № 1. P. 31–41.

511. Kofiadi I.A., Rebrikov D.V. Methods for detecting single nucleotide polymorphisms: allele-specific PCR and hybridization with oligonucleotide probe // *Genetika*. 2006. Vol. 42, № 1. P. 22–32.
512. Holden M. J., Wang L. Quantitative real-time PCR: fluorescent probe options and issues // In *Standardization and quality assurance in fluorescence measurements II*. Springer, Berlin, Heidelberg. 2008. P. 489-508.
513. Gill P., Hadian Amree A. AS-LAMP: A New and Alternative Method for Genotyping // *Avicenna J. Med. Biotechnol.* 2020. Vol. 12, № 1. P. 2–8.
514. Narayanan S. Applications of restriction fragment length polymorphism // *Ann. Clin. Lab. Sci.* 1991. Vol. 21, № 4. P. 291–296.
515. Wilson J.T. et al. Use of restriction endonucleases for mapping the allele for beta s-globin // *Proc. Natl. Acad. Sci. U. S. A.* 1982. Vol. 79, № 11. P. 3628–3631.
516. Hashim H.O., Al-Shuhaib M.B. Exploring the potential and limitations of PCR-RFLP and PCR-SSCP for SNP detection: A review // *Appl. Biotechnol. Rep. Armenian Green Publishing Co.*, 2019. Vol. 6, № 4. P. 137–144.
517. Shan Q. et al. Genome editing in rice and wheat using the CRISPR/Cas system // *Nat. Protoc.* 2014. Vol. 9, № 10. P. 2395–2410.
518. Gao Y., Gao K., Yang H. CRISPR/Cas: a potential gene-editing tool in the nervous system // *Cell Regen.* 2020. Vol. 9, № 1. P. 12.
519. Kim S. et al. Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins // *Genome Res.* 2014. Vol. 24, № 6. P. 1012–1019.

520. Vouillot L., Thélie A., Pollet N. Comparison of T7E1 and surveyor mismatch cleavage assays to detect mutations triggered by engineered nucleases // *G3* . 2015. Vol. 5, № 3. P. 407–415.
521. Lyamichev V. et al. Polymorphism identification and quantitative detection of genomic DNA by invasive cleavage of oligonucleotide probes // *Nat. Biotechnol.* 1999. Vol. 17, № 3. P. 292–296.
522. Arboleda V.A., Xian R.R. An Overview of DNA Analytical Methods // *Methods Mol. Biol.* 2019. Vol. 1897. P. 385–402.
523. Hardenbol P. et al. Highly multiplexed molecular inversion probe genotyping: over 10,000 targeted SNPs genotyped in a single tube assay // *Genome Res.* 2005. Vol. 15, № 2. P. 269–275.
524. Tong A.K. et al. Combinatorial fluorescence energy transfer tags for multiplex biological assays // *Nat. Biotechnol.* 2001. Vol. 19, № 8. P. 756–759.
525. Sando S., Abe H., Kool E.T. Quenched Auto-Ligating DNAs: Multicolor Identification of Nucleic Acids at Single Nucleotide Resolution // *Journal of the American Chemical Society.* 2004. Vol. 126, № 4. P. 1081–1087.
526. Kakavas V.K. et al. PCR-SSCP: a method for the molecular analysis of genetic diseases // *Mol. Biotechnol.* 2008. Vol. 38, № 2. P. 155–163.
527. Gasser R.B. et al. Single-strand conformation polymorphism (SSCP) for the analysis of genetic variation // *Nat. Protoc.* 2006. Vol. 1, № 6. P. 3121–3128.
528. Tabit F.T. Advantages and limitations of potential methods for the analysis of bacteria in milk: a review // *J. Food Sci. Technol.* 2016. Vol. 53, № 1. P. 42–49.
529. Xiao W., Oefner P.J. Denaturing high-performance liquid chromatography: A review // *Hum. Mutat.* 2001. Vol. 17, № 6. P. 439–474.

530. Ferguson J.A., Steemers F.J., Walt D.R. High-density fiber-optic DNA random microsphere array // *Anal. Chem.* 2000. Vol. 72, № 22. P. 5618–5624.
531. Böyum A. Separation of leukocytes from blood and bone marrow. Introduction // *Scand. J. Clin. Lab. Invest. Suppl.* 1968. Vol. 97. P. 7.
532. Matz M.V. et al. Amplification of cDNA ends using PCR suppression effect and step-out PCR // *Methods Mol. Biol.* 2003. Vol. 221. P. 41–49.
533. Cheung F. et al. Sequencing Medicago truncatula expressed sequenced tags using 454 Life Sciences technology // *BMC Genomics.* 2006. Vol. 7. P. 272.
534. Sambrook J., Fritsch E.F., Maniatis T. *Molecular Cloning: A Laboratory Manual.* 1989. P. 1659.
535. Matz M. et al. Amplification of cDNA ends based on template-switching effect and step-out PCR // *Nucleic Acids Res.* 1999. Vol. 27, № 6. P. 1558–1560.
536. Kurnikova M. et al. Identification of a novel mutation in DKC1 in dyskeratosis congenita // *Pediatric Blood & Cancer.* 2009. Vol. 52, № 1. P. 135–137.
537. Engler-Blum G. et al. Reduction of background problems in nonradioactive northern and Southern blot analyses enables higher sensitivity than ³²P-based hybridizations // *Anal. Biochem.* 1993. Vol. 210, № 2. P. 235–244.
538. Franz O., Bruchhaus I., Roeder T. Verification of differential gene transcription using virtual northern blotting // *Nucleic Acids Res.* 1999. Vol. 27, № 11. P. e3.
539. Anisimova V.E. et al. Renaturation, activation, and practical use of recombinant duplex-specific nuclease from Kamchatka crab // *Biochemistry.* 2006. Vol. 71, № 5. P. 513–519.
540. Cooper H.M., Paterson Y. Production of Polyclonal Antisera // *Current Protocols in Cell Biology.* 1999. Vol. 3, № 1. P. 16.2.1-16.2.8.

541. Harlow E., Lane D., Lane E.H.D. *Antibodies: A Laboratory Manual*. CSHL Press, 1988. P. 726.
542. Bradford M.M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding // *Analytical Biochemistry*. 1976. Vol. 72, № 1-2. P. 248–254.
543. Liao T.-H. Bovine Pancreatic Deoxyribonuclease D // *Journal of Biological Chemistry*. 1974. Vol. 249, № 8. P. 2354–2356.
544. Kunitz M. Crystalline desoxyribonuclease; isolation and general properties; spectrophotometric method for the measurement of desoxyribonuclease activity // *J. Gen. Physiol.* 1950. Vol. 33, № 4. P. 349–362.
545. Ho H.-C. et al. Purification, characterization and complete amino acid sequence of nuclease C1 from *Cunninghamella echinulata* var. *echinulata* // *European Journal of Biochemistry*. 1998. Vol. 256, № 1. P. 112–118.
546. Shevchenko A. et al. In-gel digestion for mass spectrometric characterization of proteins and proteomes // *Nat. Protoc.* 2006. Vol. 1, № 6. P. 2856–2860.
547. Nekrasov A.N. Entropy of protein sequences: an integral approach // *J. Biomol. Struct. Dyn.* 2002. Vol. 20, № 1. P. 87–92.
548. Thompson J.D. et al. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools // *Nucleic Acids Res.* 1997. Vol. 25, № 24. P. 4876–4882.
549. Nicholas K. B., Nicholas Jr H. B. GeneDoc: a tool for editing and annotating multiple sequence alignments. Pittsburgh Supercomputing Center's National Resource for Biomedical Supercomputing, ver. 2.7. 000. – 1997.

550. Strimmer K., von Haeseler A. Likelihood-mapping: a simple method to visualize phylogenetic content of a sequence alignment // *Proc. Natl. Acad. Sci. U. S. A.* 1997. Vol. 94, № 13. P. 6815–6819.
551. Yang Z. Maximum-Likelihood Models for Combined Analyses of Multiple Sequence Data // *J. Mol. Evol.* 1996. Vol. 42, № 5. P. 587–596.
552. Whelan S., Goldman N. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach // *Mol. Biol. Evol.* 2001. Vol. 18, № 5. P. 691–699.
553. Cuevas A., Febrero M., Fraiman R. An anova test for functional data // *Computational Statistics & Data Analysis.* 2004. Vol. 47, № 1. P. 111–122.
554. Haley C.S., Knott S.A. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers // *Heredity.* 1992. Vol. 69, № 4. P. 315–324.
555. Hernández-Sánchez J., Grunchev J.-A., Knott S. A web application to perform linkage disequilibrium and linkage analyses on a computational grid // *Bioinformatics.* 2009. Vol. 25, № 11. P. 1377–1383.
556. Carlborg O., Brockmann G.A., Haley C.S. Simultaneous mapping of epistatic QTL in DU6i x DBA/2 mice // *Mamm. Genome.* 2005. Vol. 16, № 7. P. 481–494.
557. Churchill G.A., Doerge R.W. Empirical threshold values for quantitative trait mapping // *Genetics.* 1994. Vol. 138, № 3. P. 963–971.
558. Brockmann G.A. et al. Quantitative trait loci affecting body weight and fatness from a mouse line selected for extreme high growth // *Genetics.* 1998. Vol. 150, № 1. P. 369–381.
559. Visscher P.M., Thompson R., Haley C.S. Confidence intervals in QTL mapping by bootstrapping // *Genetics.* 1996. Vol. 143, № 2. P. 1013–1020.

560. Chou M.Y., Liao T.H. Shrimp hepatopancreatic deoxyribonuclease--purification and characterization as well as comparison with bovine pancreatic deoxyribonuclease // *Biochim. Biophys. Acta*. 1990. Vol. 1036, № 2. P. 95–100.
561. Wang W.Y., Liaw S.H., Liao T.H. Cloning and characterization of a novel nuclease from shrimp hepatopancreas, and prediction of its active site // *Biochem. J.* 2000. Vol. 346 Pt 3. P. 799–804.
562. Schultz J. et al. SMART: a web-based tool for the study of genetically mobile domains // *Nucleic Acids Res.* 2000. Vol. 28, № 1. P. 231–234.
563. Anisimova, V.E. et al. 2006. Renaturation, activation, and practical use of recombinant duplex-specific nuclease from Kamchatka crab// *Biochemistry*. 2006. Vol. 71. № 5. P. 513–519.
564. Ho H.C., Liao T.H. Protein structure and gene cloning of *Syncephalastrum racemosum* nuclease // *Biochem. J.* 1999. Vol. 339 (№ 5 Pt 2). P. 261–267.
565. Calvo E., Ribeiro J.M.C. A novel secreted endonuclease from *Culex quinquefasciatus* salivary glands // *J. Exp. Biol.* 2006. Vol. 209, № Pt 14. P. 2651–2659.
566. Katoh K. et al. MAFFT version 5: improvement in accuracy of multiple sequence alignment // *Nucleic Acids Res.* 2005. Vol. 33, № 2. P. 511–518.
567. Almagro Armenteros J.J. et al. SignalP 5.0 improves signal peptide predictions using deep neural networks // *Nat. Biotechnol.* 2019. Vol. 37, № 4. P. 420–423.
568. Bendtsen J.D. et al. Improved Prediction of Signal Peptides: SignalP 3.0 // *Journal of Molecular Biology*. 2004. Vol. 340, № 4. P. 783–795.
569. Bernhofer M. et al. TMSEG: Novel prediction of transmembrane helices // *Proteins*. 2016. Vol. 84, № 11. P. 1706–1716.

570. Möller S., Croning M.D., Apweiler R. Evaluation of methods for the prediction of membrane spanning regions // *Bioinformatics*. 2001. Vol. 17, № 7. P. 646–653.
571. Bernhofer M. et al. PredictProtein - Predicting Protein Structure and Function for 29 Years // *Nucleic Acids Res.* 2021. Vol. 49, № W1. P. W535–W540.
572. Li S., Kwon J., Aksoy S. Characterization of genes expressed in the salivary glands of the tsetse fly, *Glossina morsitans morsitans* // *Insect Mol. Biol.* 2001. Vol. 10, № 1. P. 69–76.
573. Towle D.W., Smith C.M. Gene discovery in *Carcinus maenas* and *Homarus americanus* via expressed sequence tags // *Integr. Comp. Biol.* 2006. Vol. 46, № 6. P. 912–918.
574. Celniker S.E. et al. Finishing a whole-genome shotgun: release 3 of the *Drosophila melanogaster* euchromatic genome sequence // *Genome Biol.* 2002. Vol. 3, № 12. P. RESEARCH0079.
575. Richards S. et al. Comparative genome sequencing of *Drosophila pseudoobscura*: chromosomal, gene, and cis-element evolution // *Genome Res.* 2005. Vol. 15, № 1. P. 1–18.
576. Holt R.A. et al. The genome sequence of the malaria mosquito *Anopheles gambiae* // *Science*. 2002. Vol. 298, № 5591. P. 129–149.
577. Stapleton M. et al. The *Drosophila* gene collection: identification of putative full-length cDNAs for 70% of *D. melanogaster* genes // *Genome Res.* 2002. Vol. 12, № 8. P. 1294–1300.
578. Valenzuela J.G. et al. Identification of the most abundant secreted proteins from the salivary glands of the sand fly *Lutzomyia longipalpis*, vector of *Leishmania chagasi* // *J. Exp. Biol.* 2004. Vol. 207, № Pt 21. P. 3717–3729.

579. Ghosh M. et al. Structural Insights into the Mechanism of Nuclease A, a $\beta\beta\alpha$ Metal Nuclease from *Anabaena* // *Journal of Biological Chemistry*. 2005. Vol. 280, № 30. P. 27990–27997.
580. Chen L.Y. et al. Deoxyribonuclease of *Syncephalastrum racemosum*--enzymatic properties and molecular structure // *Arch. Biochem. Biophys.* 1993. Vol. 303, № 1. P. 51–56.
581. Jones T. et al. The diploid genome sequence of *Candida albicans* // *Proc. Natl. Acad. Sci. U. S. A.* 2004. Vol. 101, № 19. P. 7329–7334.
582. Goffeau A. et al. Life with 6000 Genes // *Science*. 1996. Vol. 274, № 5287. P. 546–567.
583. Cymerman I.A. et al. EXOG, a novel paralog of Endonuclease G in higher eukaryotes // *Nucleic Acids Res.* 2008. Vol. 36, № 4. P. 1369–1379.
584. Salamov A.A., Solovyev V.V. Ab initio gene finding in *Drosophila* genomic DNA // *Genome Res.* 2000. Vol. 10, № 4. P. 516–522.
585. Rogov S.I., Nekrasov A.N. A numerical measure of amino acid residues similarity based on the analysis of their surroundings in natural protein sequences // *Protein Eng.* 2001. Vol. 14, № 7. P. 459–463.
586. Nekrasov A.N. Analysis of the information structure of protein sequences: a new method for analyzing the domain organization of proteins // *J. Biomol. Struct. Dyn.* 2004. Vol. 21, № 5. P. 615–624.
587. Schäfer P. et al. Structural and functional characterization of mitochondrial EndoG, a sugar non-specific nuclease which plays an important role during apoptosis // *J. Mol. Biol.* 2004. Vol. 338, № 2. P. 217–228.

588. Miller M.D., Cai J., Krause K.L. The active site of *Serratia* endonuclease contains a conserved magnesium-water cluster // *J. Mol. Biol.* 1999. Vol. 288, № 5. P. 975–987.
589. Haddow J.D. et al. Identification of major soluble salivary gland proteins in teneral *Glossina morsitans morsitans* // *Insect Biochem. Mol. Biol.* 2002. Vol. 32, № 9. P. 1045–1053.
590. Friedhoff P. et al. Kinetic analysis of the cleavage of natural and synthetic substrates by the *Serratia* nuclease // *Eur. J. Biochem.* 1996. Vol. 241, № 2. P. 572–580.
591. Caljon G. et al. Tsetse salivary gland proteins 1 and 2 are high affinity nucleic acid binding proteins with residual nuclease activity // *PLoS One.* 2012. Vol. 7, № 10. P. e47233.
592. Nilsen I.W. et al. The enzyme and the cDNA sequence of a thermolabile and double-strand specific DNase from Northern shrimps (*Pandalus borealis*) // *PLoS One.* 2010. Vol. 5, № 4. P. e10295.
593. Hirakawa & Ohiso, 2012 EUROPEAN PATENT SPECIFICATION EP 2161331B1 Date of publication and mention of the grant of the patent: 25.04.2012 Bulletin 2012/17
594. Chagas A.C. et al. Lundep, a sand fly salivary endonuclease increases *Leishmania* parasite survival in neutrophils and inhibits XIIa contact activation in human plasma // *PLoS Pathog.* 2014. Vol. 10, № 2. P. e1003923.
595. Seong C.-S. et al. Cloning and characterization of a novel *Drosophila* stress induced DNase // *PLoS One.* 2014. Vol. 9, № 8. P. e103564.
596. Mesaeli N., Phillipson C. Impaired p53 expression, function, and nuclear localization in calreticulin-deficient cells // *Mol. Biol. Cell.* 2004. Vol. 15, № 4. P. 1862–1870.

597. de Moerloose P. et al. Superficial vein thrombosis of lower limbs: influence of factor V Leiden, factor II G20210A and overweight // *Thromb. Haemost.* 1998. Vol. 80, № 2. P. 239–241.
598. Kirke P.N. et al. Impact of the MTHFR C677T polymorphism on risk of neural tube defects: case-control study // *BMJ.* 2004. Vol. 328, № 7455. P. 1535–1536.
599. Chan T.L. et al. BRAF and KRAS mutations in colorectal hyperplastic polyps and serrated adenomas // *Cancer Res.* 2003. Vol. 63, № 16. P. 4878–4881.
600. Mu D.-Q., Peng Y.-S., Xu Q.-J. Values of mutations of K-ras oncogene at codon 12 in detection of pancreatic cancer: 15-year experience // *World J. Gastroenterol.* 2004. Vol. 10, № 4. P. 471–475.
601. Shih F. et al. Oncogenic codon 13 NRAS mutation in a primary mesenchymal brain neoplasm and nevus of a child with neurocutaneous melanosis // *Acta Neuropathologica Communications.* 2014. Vol. 2, № 1. P. 140.
602. Galiana C. et al. Possible role of activated ras genes in human esophageal carcinogenesis // *Int. J. Cancer.* 1993. Vol. 54, № 6. P. 978–982.
603. Bryk A.H. et al. Bleeding predictors in patients following venous thromboembolism treated with vitamin K antagonists: Association with increased number of single nucleotide polymorphisms // *Vascul. Pharmacol.* 2018. Vol. 106. P. 22–27.
604. Sokolenko A.P. et al. High frequency of BRCA1 5382insC mutation in Russian breast cancer patients // *Eur. J. Cancer.* 2006. Vol. 42, № 10. P. 1380–1384.
605. Wang F. et al. Common BRCA1 and BRCA2 mutations in breast cancer families: a meta-analysis from systematic review // *Mol. Biol. Rep.* 2012. Vol. 39, № 3. P. 2109–2118.

606. Shifman S. et al. A highly significant association between a COMT haplotype and schizophrenia // *Am. J. Hum. Genet.* Elsevier BV, 2002. Vol. 71, № 6. P. 1296–1302.
607. Hanson R.L. et al. Identification of PVT1 as a candidate gene for end-stage renal disease in type 2 diabetes using a pooling-based genome-wide single nucleotide polymorphism association study // *Diabetes.* 2007. Vol. 56, № 4. P. 975–983.
608. Stewart C.A. et al. Complete MHC haplotype sequencing for common disease gene mapping // *Genome Res.* 2004. Vol. 14, № 6. P. 1176–1187.
609. Bidshahri R. et al. Quantitative Detection and Resolution of BRAF V600 Status in Colorectal Cancer Using Droplet Digital PCR and a Novel Wild-Type Negative Assay // *J. Mol. Diagn.* 2016. Vol. 18, № 2. P. 190–204.
610. Itabashi T. et al. Quantitative detection of mutant alleles of the K-ras gene with minor groove binder-conjugated fluorogenic DNA probes // *Int. J. Oncol.* 2004. Vol. 24, № 3. P. 687–696.
611. Holden M. J., Wang L. Quantitative real-time PCR: fluorescent probe options and issues // *Standardization and quality assurance in fluorescence measurements II.* Springer, Berlin, Heidelberg, 2008. P. 489-508.
612. Khripin Y. High-throughput genotyping with energy transfer-labeled primers // *Methods Mol. Biol.* 2006. Vol. 335. P. 215–240.
613. Alcover A., Alarcón B., Di Bartolo V. Cell Biology of T Cell Receptor Expression and Regulation // *Annual Review of Immunology.* 2018. Vol. 36, № 1. P. 103–125.
614. Davey M.P. et al. T-cell receptor variable β genes show differential expression in CD4 and CD8 T cells // *Human Immunology.* 1991. Vol. 32, № 3. P. 194–202.
615. Durgeau A. et al. Recent Advances in Targeting CD8 T-Cell Immunity for More Effective Cancer Immunotherapy // *Front. Immunol.* 2018. Vol. 9. P. 14.

616. Nolan S. et al. A large-scale database of T-cell receptor beta (TCR β) sequences and binding associations from natural and synthetic exposure to SARS-CoV-2 // *Res Sq.* 2020.
617. Zmievsckaya E. et al. Application of CAR-T Cell Therapy beyond Oncology: Autoimmune Diseases and Viral Infections // *Biomedicines.* 2021. Vol. 9, № 1. P. 59.
618. Arden B. et al. Human T-cell receptor variable gene segment families // *Immunogenetics.* 1995. Vol. 42, № 6. P. 455–500.
619. Kwok S. et al. A guide to the design and use of mismatched and degenerate primers // *PCR Methods Appl.* 1994. Vol. 3, № 4. P. S39–S47.
620. He L., Hannon G.J. MicroRNAs: small RNAs with a big role in gene regulation // *Nat. Rev. Genet.* 2004. Vol. 5, № 7. P. 522–531.
621. Kristensen H. et al. Novel diagnostic and prognostic classifiers for prostate cancer identified by genome-wide microRNA profiling // *Oncotarget.* 2016. Vol. 7, № 21. P. 30760–30771.
622. Zhou X. et al. Plasma miRNAs in diagnosis and prognosis of pancreatic cancer: A miRNA expression analysis // *Gene.* 2018. Vol. 673. P. 181–193.
623. Qiu X. et al. Duplex-specific nuclease-mediated bioanalysis // *Trends Biotechnol.* 2015. Vol. 33, № 3. P. 180–188.
624. Tian J. et al. Ultrathin graphitic carbon nitride nanosheets: a novel peroxidase mimetic, Fe doping-mediated catalytic performance enhancement and application to rapid, highly sensitive optical detection of glucose // *Nanoscale.* 2013. Vol. 5, № 23. P. 11604–11609.

625. Deng J. et al. Real-time colorimetric assay of inorganic pyrophosphatase activity based on reversibly competitive coordination of Cu^{2+} between cysteine and pyrophosphate ion // *Anal. Chem.* 2013. Vol. 85, № 19. P. 9409–9415.
626. Guo Y. et al. A plasmonic colorimetric strategy for biosensing through enzyme guided growth of silver nanoparticles on gold nanostars // *Biosens. Bioelectron.* 2016. Vol. 78. P. 267–273.
627. Shen Q. et al. A simple “clickable” biosensor for colorimetric detection of copper (II) ions based on unmodified gold nanoparticles // *Biosens. Bioelectron.* 2013. Vol. 41. P. 663–668.
628. Lu X. et al. Target-driven switch-on fluorescence aptasensor for trace aflatoxin B1 determination based on highly fluorescent ternary CdZnTe quantum dots // *Anal. Chim. Acta.* 2019. Vol. 1047. P. 163–171.
629. Wang G.-H. et al. Utility of *Ochrobactrum anthropi* YC152 in a Microbial Fuel Cell as an Early Warning Device for Hexavalent Chromium Determination // *Sensors.* 2016. Vol. 16, № 8. P. 1272.
630. Yue X. et al. Rapid detection of Salmonella in milk by a nuclear magnetic resonance biosensor based on the streptavidin-biotin system and O-carboxymethyl chitosan target gadolinium probe // *J. Dairy Sci.* 2021. Vol. 104, № 11. P. 11486–11498.
631. Kim J.H. et al. Technological advances in electrochemical biosensors for the detection of disease biomarkers // *Biomed Eng Lett.* 2021. Vol. 11, № 4. P. 309–334.
632. Zhang S. et al. Novel electrochemical sensors from poly[N-(ferrocenyl formacyl) pyrrole]@multi-walled carbon nanotubes nanocomposites for simultaneous determination of ascorbic acid, dopamine and uric acid // *Nanotechnology.* 2019. Vol. 31, № 8. P. 085503.

633. Qiu W. et al. Spin-Valve based magnetoresistive nanoparticle detector for applications in biosensing // *Sensors and Actuators A: Physical*. 2017. Vol. 265. P. 174–180.
634. Li N. et al. Eliminating sweet spot in MALDI-MS with hydrophobic ordered structure as target for quantifying biomolecules // *Talanta*. 2020. Vol. 218. P. 121172.
635. Chenchik A. et al. Full-length cDNA cloning and determination of mRNA 5' and 3' ends by amplification of adaptor-ligated cDNA // *Biotechniques*. 1996. Vol. 21, № 3. P. 526–534.
636. Maruyama K., Sugano S. Oligo-capping: a simple method to replace the cap structure of eukaryotic mRNAs with oligoribonucleotides // *Gene*. 1994. Vol. 138, № 1-2. P. 171–174.
637. Nagalakshmi U. et al. The transcriptional landscape of the yeast genome defined by RNA sequencing // *Science*. 2008. Vol. 320, № 5881. P. 1344–1349.
638. Carninci P. et al. High efficiency selection of full-length cDNA by improved biotinylated cap trapper // *DNA Res.* 1997. Vol. 4, № 1. P. 61–66.
639. Zhu Y.Y. et al. Reverse transcriptase template switching: a SMART approach for full-length cDNA library construction // *Biotechniques*. 2001. Vol. 30, № 4. P. 892–897.
640. Oikonomopoulos S. et al. Methodologies for Transcript Profiling Using Long-Read Technologies // *Front. Genet.* 2020. Vol. 11. P. 606.
641. Claro I.M. et al. Rapid viral metagenomics using SMART-9N amplification and nanopore sequencing // *Wellcome Open Research*. 2021. Vol. 6. P. 241.
642. Chenchik A. et al. Gene Cloning and Analysis by RT-PCR // *BioTechniques Books*. Natick, MA In Siebert P. and Larrick J. (eds). 1998. P. 305–319.

643. Shagin D.A. et al. Regulation of average length of complex PCR product // *Nucleic Acids Res.* 1999. Vol. 27, № 18. P. e23.
644. Moreno-Palanques R.F., Fuldner R.A. Construction of cDNA Libraries // *Automated DNA Sequencing and Analysis.* 1994. P. 102–109.
645. Rabbitts T.H., Stocks M.R. Chromosomal translocation products engender new intracellular therapeutic technologies // *Nat. Med.* 2003. Vol. 9, № 4. P. 383–386.
646. Ekblom R. et al. Comparison between Normalised and Unnormalised 454-Sequencing Libraries for Small-Scale RNA-Seq Studies // *Comp. Funct. Genomics.* 2012. Vol. 2012. P. 281693.
647. Hu Y. et al. Improving the diversity of captured full-length isoforms using a normalized single-molecule RNA-sequencing method // *Commun Biol.* 2020. Vol. 3, № 1. P. 403.
648. Sambol N., Creecy J. Duplex-specific nuclease (DSN): A method for the rehabilitation of low-copy number DNA profiles // *Forensic Science International: Genetics Supplement Series.* 2013. Vol. 4, № 1. P. e59–e60.
649. Tanguy M. et al. Sequence analysis of a normalized cDNA library of *Mytilus edulis* hemocytes exposed to *Vibrio splendidus* LGP32 strain // *Results Immunol.* 2013. Vol. 3. P. 40–50.
650. Ke T. et al. Construction of a Normalized Full-Length cDNA Library of Sesame Developing Seed by DSN and SMARTTM // *Agricultural Sciences in China.* 2011. Vol. 10, № 7. P. 1004–1009.
651. Matvienko M. et al. Consequences of normalizing transcriptomic and genomic libraries of plant genomes using a duplex-specific nuclease and tetramethylammonium chloride // *PLoS One.* 2013. Vol. 8, № 2. P. e55913.

652. Lander E.S. et al. Initial sequencing and analysis of the human genome // *Nature*. 2001. Vol. 409, № 6822. P. 860–921.
653. Trifonov V.A., Vorobieva N.N., Rens W. FISH With and Without COT1 DNA // *Fluorescence In Situ Hybridization (FISH) — Application Guide*. 2009. P. 99–109.
654. Ramakrishna W. et al. Different types and rates of genome evolution detected by comparative sequence analysis of orthologous segments from four cereal genomes // *Genetics*. 2002. Vol. 162, № 3. P. 1389–1400.
655. Wicker T. et al. CACTA transposons in Triticeae. A diverse family of high-copy repetitive elements // *Plant Physiol*. 2003. Vol. 132, № 1. P. 52–63.
656. Conner J. et al. Targeted sequencing of a complex locus within a polyploid genome using reduced representation libraries // *Molecular Breeding*. 2016. Vol. 36, № 5. P. 60.
657. Ichida H., Abe T. An improved and robust method to efficiently deplete repetitive elements from complex plant genomes // *Plant Sci*. 2019. Vol. 280. P. 455–460.
658. Gagic D. et al. Improving the genetic representation of rare taxa within complex microbial communities using DNAnormalization methods // *Molecular Ecology Resources*. 2015. Vol. 15, № 3. P. 464–476.
659. Beavo J.A. Cyclic nucleotide phosphodiesterases: functional implications of multiple isoforms // *Physiol. Rev*. 1995. Vol. 75, № 4. P. 725–748.
660. Manganiello V.C. et al. Type III cGMP-inhibited cyclic nucleotide phosphodiesterases (PDE3 gene family) // *Cell. Signal*. 1995. Vol. 7, № 5. P. 445–455.
661. Brown, T. *Genomes 4*. - Garland Science, 2018
662. Diatchenko L. et al. Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries // *Proc. Natl. Acad. Sci. U. S. A*. 1996. Vol. 93, № 12. P. 6025–6030.

663. Gurskaya N.G. et al. Equalizing cDNA subtraction based on selective suppression of polymerase chain reaction: cloning of Jurkat cell transcripts induced by phytohemagglutinin and phorbol 12-myristate 13-acetate // *Anal. Biochem.* 1996. Vol. 240, № 1. P. 90–97.
664. Archer S.K., Shirokikh N.E., Preiss T. Selective and flexible depletion of problematic sequences from RNA-seq libraries at the cDNA stage // *BMC Genomics.* 2014. Vol. 15. P. 401.
665. Zhang Y., Pelechano V. Application of high-throughput 5'P sequencing for the study of co-translational mRNA decay // *STAR Protoc.* 2021. Vol. 2, № 2. P. 100447.
666. Chung B.Y. et al. The use of duplex-specific nuclease in ribosome profiling and a user-friendly software package for Ribo-seq data analysis RNA // 2015. Vol. 21, № 10. P. 1731–1745.
667. Garoutte A. et al. Methodologies for probing the metatranscriptome of grassland soil // *J. Microbiol. Methods.* 2016. Vol. 131. P. 122–129.
668. Hoeijmakers W.A.M., Bártfai R., Stunnenberg H.G. Transcriptome analysis using RNA-Seq // *Methods Mol. Biol.* 2013. Vol. 923. P. 221–239.
669. Fang N., Akinci-Tolun R. Depletion of Ribosomal RNA Sequences from Single-Cell RNA-Sequencing Library // *Curr. Protoc. Mol. Biol.* 2016. Vol. 115. P. 7.27.1–7.27.20.