Федеральное государственное бюджетное учреждение науки Институт общей генетики им. Н.И. Вавилова Российской академии наук

На правах рукописи

Михайлова Алина Геннадьевна

### Мутационные спектры мтДНК животных

ДИССЕРТАЦИЯ

на соискание ученой степени кандидата биологических наук

1.5.7 – Генетика

Научный руководитель: к.б.н. Попадьин К. Ю.

Москва, 2022

### СОДЕРЖАНИЕ

СОДЕРЖАНИЕ
СПИСОК СОКРАЩЕНИЙ
ВВЕДЕНИЕ
ГЛАВА 1. ОБЗОР ЛИТЕРАТУРЫ 11
ГЛАВА 2. МАТЕРИАЛЫ И МЕТОДЫ 17
2.1 Обозначение тяжелой цепи 12-компонентного мутационного спектра мтДНК. 
2.2 Анализ дуплексного секвенирования мтДНК 17
2.3 Реконструкция видоспецифичного мутационного спектра всех доступных видов млекопитающих
2.4 Подбор экологических параметров 19
2.5 Анализ полных митохондриальных геномов 19
2.6 Анализ времени, проведенного ДНК в одноцепочечном состоянии
2.7 Анализ дополнительных метрик 20
2.8 Нормализация мутационных спектров 21
ГЛАВА 3. РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ
3.1 Мутационные спектры млекопитающих: замена А <sub>т</sub> >Г <sub>т</sub> ассоциирована с продолжительностью жизни
3.1.1 Частота de novo мутаций А <sub>т</sub> >Г <sub>т</sub> увеличивается с возрастом в соматических тканях и тканях зародышевой линии
3.1.2 А <sub>Т</sub> >Г <sub>Т</sub> наиболее часты у млекопитающих с большим временем генерации поколения
3.1.3 МтДНК млекопитающих с большим временем генерации поколения более бедна А <sub>т</sub> и богата Г <sub>т</sub> вследствие интенсивного мутагенеза в направлении А <sub>т</sub> >Г <sub>т</sub>
3.1.4 Г <sub>т</sub> А <sub>т</sub> -сдвиг является функцией как ВПДОС, так и времени генерации поколения
3.2 Мутационные спектры экзотермическими животных: замена A <sub>T</sub> >Г <sub>T</sub> ассоциирована с температурой тела69
3.2.1 Видоспецифичный мутационный спектр мтДНК рыб связан с температурой окружающей среды посредством повышенной асимметрии A <sub>T</sub> >Г <sub>т</sub>
3.2.2 Мутационный спектр мтДНК рыб влияет на их нейтральный нуклеотидный состав: тепловодные и долгоживущие виды, как правило, бедны АтЦт и богаты ГтТт

3.2.3 Мутационный спектр мтДНК чувствителен к температуре у всех позвоночных	классов 88
3.3 Мутационный спектр мтДНК в опухолевых тканях человека изменяет время канцерогенеза и связан со скоростью пролиферации клеток в т	гся во канях
предшественниках	
ГЛАВА 4. ОБСУЖДЕНИЕ	101
ВЫВОДЫ	110
СПИСОК ЛИТЕРАТУРЫ	111
СПИСОК РАБОТ, ОПУБЛИКОВАННЫХ ПО ТЕМЕ ДИССЕРТАЦИИ	122
БЛАГОДАРНОСТИ	128

#### СПИСОК СОКРАЩЕНИЙ

- мтДНК митохондриальная ДНК
- тРНК транспортная РНК
- рРНК рибосомальная РНК
- АФК активные формы кислорода
- яДНК ядерная ДНК
- 8-ОхоG 8-оксогуанин
- АП-сайты апуриновые/апиримидиновые сайты
- 5-OHU 5-Гидроксиурацил
- 5-ОНС 5-Гидроксицитозин
- 8-ОхоА 8-оксоаденин
- ВПДОС время, проведенное ДНК в одноцепочечном состоянии
- ЦОГ1 ген субъединицы I цитохром с-оксидазы
- ЦИТb ген цитохром b
- ЧВА частота вариантного аллеля

#### **ВВЕДЕНИЕ**

Актуальность и степень разработанности проблемы. Ожидаемая продолжительность жизни человека неустанно растет, и с каждым днем человечество все чаще сталкивается с болезнями старения. Наиболее революционные открытия в области геронтологии происходят не только при исследовании одного лишь человека как объекта, а также при сравнительных геномных анализах короткоживущих и долгоживущих видов животных. Отличным примером такого подхода являются недавние открытия, сделанные в лаборатории профессора Горбуновой [1,2], которые подчеркивают важность эволюционного подхода при изучении процесса старения человека.

Старение, обусловленное накоплением поломок (мутаций) ДНК, наиболее сильно выражается в митохондриальной ДНК (мтДНК), которая мутирует в десятки раз быстрее ядерного генома, а также способна к быстрому внутриклеточному отбору в сторону эгоистичных форм мтДНК. Эти качества мтДНК делают ее очень интересным и важным объектом изучения старения [3]. Следовательно, главной целью данной работы является понимание роли митохондриальной компоненты в процессе старения различных тканей у позвоночных животных. Ранние исследования на опухолевых тканях [4] показали, что мутационный спектр (вероятности перехода одного нуклеотида в любой другой) митохондрий уникален и обусловлен действием специфичного для митохондрий мутагена, который воздействует схожим образом на разные клетки. Предположительно, небольшие различия в мутационных спектрах митохондриальной ДНК разных животных и даже разных тканей могут свидетельствовать о особенностях внутри митохондриальной среды и специфичных для митохондрий мутагенах.

Восстановление мутационных спектров мтДНК разных видов животных позволит провести сравнительно-видовые анализы, которые прольют свет на видоспецифичные для митохондрий мутагены. Проведение биоинформатического анализа всех доступных данных с последовательностями

мтДНК позвоночных животных и человека и, как следствие, выявление митохондриальных мутагенов, которые могут быть ответственными за наблюдаемые ассоциации (например, такие, как уровень окислительного метаболизма), представляют собой важный этап В расшифровке митохондриальных мутационных сигнатур. Этот подход позволит открыть новую главу в интерпретации изменений в митохондриальных геномах. Мутационные спектры, восстановленные по нейтральным синонимичным полиморфным заменам мтДНК, смогут раскрыть физиологические И метаболические особенности изучаемых видов животных и тканей.

Соматические митохондриальные делеции, которые накапливаются у стареющих людей, также связаны с нейродегенерацией и саркопенией. Однако, несмотря на существующие исследования, нет общепринятых мнений о механизмах образования делеций в мтДНК [5]. Понимание этих механизмов поможет развивать митохондриальную медицину для снижения риска нейродегенеративных заболеваний.

В итоге, многомасштабные эмпирические данные по мутагенезу мтДНК, накопленные в различных базах данных, позволяют изучить и внедрить анализ мутационных спектров мтДНК, что приведет не только к открытию новых фундаментальных законов, касающихся особенностей мутагенеза мтДНК, накопления соматических мутаций и образования делеций, и, как следствие, описанию молекулярных сигнатур ключевых митохондриальных мутагенов, но и к прикладным результатам.

Цель исследования. Целью настоящей работы является изучение особенностей мутационных спектров мтДНК различных видов животных и человека с последующим выделением факторов мутагенеза и установлением мутационных сигнатур митохондриального генома позвоночных.

#### Задачи исследования:

1. Анализ мутационных спектров мтДНК, полученных из данных соматических и герменативных мутаций здоровых тканей человека и мышей;

- 2. Реконструкция и анализ мутационных спектров мтДНК, полученных из полиморфных геномных данных млекопитающих;
- 3. Реконструкция и анализ мутационных спектров мтДНК, полученных из полиморфных геномных данных рыб;
- 4. Анализ мутационных спектров мтДНК соматических мутаций, полученных из данных опухолевых тканей человека;
- Анализ эффектов обусловленных различиями базового метаболизма позвоночных животных (продолжительности жизни и температуры) на митохондриальный мутационный спектр;
- Анализ эффектов обусловленных различиями клеточных параметров и метаболизма (гипоксии и скорости пролиферации клеток) на митохондриальный мутационный спектр опухолевых тканей человека.

Научная новизна работы. В данной работе впервые показаны особенности мутационных процессов в митохондриальной ДНК, выявленные посредством анализа мутационных спектров позвоночных животных, полученных с использованием новейшего алгоритма реконструкции.

Высокая скорость мутации митохондриальной ДНК по сравнению с ядерной ДНК и увеличенное число копий мтДНК создают высокий уровень митохондриальных мутаций как в герминальных, так и в соматических тканях. Эти мутации мтДНК могут быть связаны с наследственными генетическими болезнями, фенотипами старения (например, нейродегенерация и саркопения) и даже развитием злокачественных опухолей [6–8]. Несмотря на высокую частоту возникновения мутаций мтДНК и их значительное влияние на здоровье человека, процесс митохондриального мутагенеза изучен недостаточно. Вопросы, связанные с пониманием природы такого мутагена или мутагенов, его связью с возрастом человека и животных, межвидовой изменчивостью мутационных спектров и их причинами, по-прежнему остаются открытыми.

В мировой литературе существует огромное количество информации о мутационных спектрах ядерных геномов, однако митохондриальные геномы фактически не изучены. Например, ранее было предположено, что избыток Г- нуклеотидов в митохондриальном геноме у долгоживущих млекопитающих может быть следствием благоприятного отбора в пользу более стабильных геномов у таких видов [9], однако для достоверных выводов необходимо сравнить мутационные процессы у короткоживущих и долгоживущих млекопитающих. В других исследованиях мутационные спектры мтДНК оказались различными между разными видами [10,11], но до сих пор не установлены ключевые факторы, объясняющие эти различия.

Теоретическая и практическая значимость исследования. Результаты рабочую гипотезу: данного исследования позволили сформулировать мутационный спектр митохондриальной ДНК, особенно частота мутаций из аденина в гуанин на тяжелой цепи мтДНК, может быть чувствителен к уровню аэробного метаболизма на уровне клеток и тканей. Ранее, в работе Людвиг и др. 2019 (Ludwig et al. 2019) было показано, что митохондриальные соматические мутации можно утилизировать для отслеживания траектории развития отдельных клеток в сложных организмах [12]. Однако, основываясь на результатах проделанного исследования, мутационный спектр мтДНК может информативным маркером для оценки интенсивности аэробного стать метаболизма в различных клетках, тканях, организмах и видов. Исследование и использование этого маркера могут открыть новые возможности для решения как фундаментальных, так и прикладных задач, и расширить применение митохондриальных мутаций.

#### Основные положения, выносимые на защиту:

- Мутационные спектры позвоночных ассоциированы с такими параметрами жизненного цикла позвоночных животных, как температура тела и продолжительность жизни.
- Вероятность транзиций А<sub>т</sub>>Г<sub>т</sub> увеличивается у долгоживущих животных и у животных с высокой температурой тела.
- 3. Вероятность транзиций A<sub>T</sub>>Г<sub>T</sub> уменьшается в опухолевых образцах на более поздней стадии канцерогенеза (предположительно более

гипоксичных) и в тканях, с высокой скоростью пролиферации клеток (также предположительно более гипоксичных).

 Мутация А<sub>т</sub>>Г<sub>т</sub> может быть чувствительна к уровню молекулярного кислорода и являться уникальным маркером окислительного метаболизма клеток, тканей и организмов.

Личный вклад автора. Автор настоящей диссертации принимал непосредственное участие в подготовке баз данных, анализе мутационных спектров, интерпретации результатов и написании текста, за исключением: разработка программного комплекса, реконструирующего мутационные спектры митохондрий из разных типов геномных данных (см. Материалы и методы) – с. н. с. БФУ им. Канта Гунбин К. В., анализ вторичной структуры мтДНК (см. Литературный обзор) – м. н. с. БФУ им. Канта Шаманский В. А., моделирование ожидаемого нуклеотидного состава на основе системы дифференциальных уравнений - с. н. с. БФУ им. Канта Юров В. А.

Степень достоверности и апробация работы. Надежность полученных подтверждается выводами, основанными результатов на статистически значимых данных. Результаты диссертационной работы были представлены на следующих конференциях: международная конференция Society for Molecular Biology and Evolution (SMBE) Satellite Meeting, Эйн-Геди, Израиль, 3-6 сентября 2017; международная конференция SMBE, Йокогама, Япония, 8-12 июля 2018; конференция Bioinformatics of Genome Regulation and международная Structure/Systems Biology (BGRS/SB), 20-25 августа 2018; международная конференция Moscow Conference on Computational Molecular Biology (MCCMB), Москва, Россия, 27-30 июля 2019; международная конференция SMBE Fitch Symposium, 30 июня 2020; BGRS/SB, 6-10 июля 2020; международная конференция European Society of Human Genetics (ESHG), 28-31 августа 2021; международная конференция МССМВ, Москва, Россия, 3-6 августа 2021; международная конференция ESHG, 11-14 июня 2022; международная SMBE Everywhere Global Symposia 3, 2 конференция августа 2022; международная конференция Mitochondria 2022 Workshop, Эйн-Геди, Израиль,

13 – 16 ноября 2022; международная конференция ESHG, Глазго, Великобритания, 10-13 июня 2023; международная конференция SMBE, Феррара, Италия, 23-27 июля 2023; всероссийская конференция МССМВ, Москва, Россия, 3-6 августа 2023.

Публикации. Результаты исследования представлены в 6 научных публикациях, в том числе в 3 статьях в ведущих научных журналах, индексируемых в базах данных Scopus и Web of Science и рекомендованных ВАК для защиты диссертаций.

Объем и структура диссертации. Диссертация изложена на 128 страницах и включает следующие разделы: «Введение», «Обзор литературы», «Материалы и методы», «Результаты исследования», «Обсуждение», «Выводы», «Благодарности» «Список литературы». Работа содержит 40 рисунков и 29 таблиц. Список литературы включает 123 литературных источников.

Работа выполнена при поддержке гранта РНФ 21-75-20143 - (2021-2024 гг) и Российской Федеральной программой академического лидерства "Приоритет 2030" в Балтийском федеральном университете им. Иммануила Канта.

#### ГЛАВА 1. ОБЗОР ЛИТЕРАТУРЫ

Митохондриальная ДНК человека (мтДНК) представляет собой замкнутую кольцевую молекулу, состоящую примерно из 16 600 пар оснований. Этот компактный геном содержит 37 генов, которые кодируют 13 полипептидов, 22 транспортных РНК (тРНК) и 2 рибосомальных РНК (рРНК). 13 полипептидов являются важными компонентами митохондриальных дыхательных комплексов, расположенных во внутренней митохондриальной мембране. Хотя мтДНК кодирует лишь небольшую часть из примерно 90 белков, участвующих в дыхательной цепи, её роль жизненно важна для клеточного дыхания. Клетки, лишенные мтДНК, известные как р0-клетки, демонстрируют нарушенную функцию дыхания, что подчеркивает важность мтДНК для поддержания нормальной клеточной активности [13].

Значение поддержания целостности мтДНК дополнительно подчеркивается исследованиями на генетически модифицированных мышахмутаторах. Эти мыши, накапливающие большое количество мутаций в своей мтДНК, проявляют фенотипы преждевременного старения и сокращенную продолжительность жизни, что подчеркивает критическую роль поддержания мтДНК для здоровья и долголетия [14].

Исследование митохондриальных мутаций И митохондриальных мутационных сигнатур имеет первостепенное значение из-за их глубокого влияния на здоровье человека и болезни. Соматические и наследственные митохондриальной ЛНК ассоциированы мутации В с различными митохондриальными заболеваниями, которые часто поражают ткани с высоким энергетическим спросом, такие как мышцы и нервная система, включая потерю слуха [15] и нейродегенеративные расстройства [6]. Понимание этих мутаций может помочь в диагностике и разработке целенаправленных терапий для этих состояний [16]. Кроме того, митохондриальные мутационные сигнатуры специфические модели мутаций, возникающие в результате различных эндогенных и экзогенных факторов — могут дать представление о механизмах

митохондриальной дисфункции. Анализируя эти сигнатуры, исследователи могут проследить происхождение митохондриальных мутаций, понять их влияние на клеточный метаболизм и разработать стратегии для смягчения их последствий. Более того, поскольку митохондриальная дисфункция связана со старением и рядом возрастных заболеваний, изучение этих мутаций и их сигнатур может способствовать разработке вмешательств для поддержания здорового старения и увеличения продолжительности жизни. Следовательно, продолжение исследований в области митохондриального мутагенеза имеет важное значение для продвижения понимания митохондриальной биологии и улучшения здоровья человека. Высокая частота митохондриальных мутаций обеспечивает богатый источник вариантов, которые уже широко используются для отслеживания истории видов, популяций, организмов внутри популяций [17,18], а в последнее время и клеток в тканях [12,19]. Суммируя вышесказанное, изучение эволюции мтДНК не только существенно для нахождения вариаций мтДНК в различных популяциях, понимания генетического разнообразия и экологических динамик, но и обладает значительным потенциалом для продвижения медицинских исследований.

Молекулярная эволюция является функцией как мутационных процессов, так и естественного отбора. Чтобы грамотно интерпретировать действие силы естественного отбора, крайне важно изучить процесс возникновения мутаций. предположение, что избыток Г-нуклеотидов в Ранее было высказано митохондриальном геноме долгоживущих млекопитающих является результатом отбора, благоприятствующего более стабильным геномам у долгоживущих видов за счет избытка ГЦ пар нуклеотидов [9]. Однако данное заключение может быть недостаточно обоснованным, если не провести сравнительный анализ мутационных процессов у короткоживущих И долгоживущих млекопитающих. Действительно, были показаны значительные отличия в мутационных спектрах мтДНК между разными видами [10,11], однако до сих пор не было предложено никаких движущих факторов, объясняющих эти отличия.

Интересно, что аналогичный пробел в знаниях о мутационных спектрах мтДНК существует на уровне сравнения тканей. Исследования, проводимые на популяции пациентов с различными типами опухолей показали, что мтДНК имеет уникальную мутационную сигнатуру, которая отличается от всех известных ядерных сигнатур [20]. Более того, хорошо известные сильные экзогенные мутагены, такие как табачный дым при раке легких у курящих или ультрафиолетовый свет при меланоме, не оказывают ожидаемого воздействия на митохондриальный мутационный спектр. Таким образом, основной мутаген мтДНК, а также причины изменчивости мутационных спектров мтДНК неизвестны как на сравнительно-тканевом, так и на сравнительно-видовом уровне.

Электронный утечки в митохондриальной дыхательной цепи приводить приводить к образованию активных форм кислорода (АФК) [21]. АФК могут нанести ущерб митохондриальной ДНК, повреждая основания ДНК и индуцируя разрывы в цепях ДНК, которые, если не репарируются должным образом, могут привести к мутациям и нестабильности генома [22,23]. Скорость мутаций в некоторых участках человеческой митохондриальной ДНК, включая последовательности рРНК и тРНК, может быть в сотни раз выше, чем в ядерной ДНК [24,25].

Активные формы кислорода являются нормальным побочным продуктом аэробного метаболизма, следовательно, более высокий уровень метаболической активности может приводить к увеличению окислительных повреждений [26–28]. Некоторые исследования показывали, что мтДНК более чувствительна к окислительным повреждениям, чем ядерная ДНК (яДНК) [29,30]. Хорошо документированным мутационным признаком, индуцированным АФК, является модификация основания ДНК гуанина в 8-оксогуанин (8-ОхоG), которая после неправильного спаривания с аденином приводит к Г>Т/Ц>А трансверсиям. Хотя замены Г>Т считаются признаком окислительного повреждения ядерной ДНК (сигнатура COSMIC SBS18) [31–33], они довольно редки в мтДНК и незначительно увеличиваются с возрастом в мтДНК [4,34]. Важно отметить, что

параллельно с повышением уровня метаболизма развиваются многочисленные антиоксидантные защитные системы и системы репарации, эффективно компенсирующие повышенное окислительное повреждение [35–37]. Кроме того, редкость данной замены, несмотря на довольно высокий уровень образование окислительного повреждения мтДНК нуклеотидных И модификаций, может быть объяснена наличием высокоэффективных систем репарации.

Эксцизионная репарация оснований (англ. - BER, base excision repair) является основным механизмом восстановления повреждений оснований ДНК, вызванных АФК [38]. Эксцизионная репарация окислительных повреждений оснований и урацила обнаружена в митохондриях [39,40]. Например, большинство окислительных нуклеотидных модификаций мтДНК восстанавливаются главным образом при помощи OGG1 гликозилазы [40,41]), одной из двух главных бифункциональных гликозилаз, репарирующих окисленные нуклеотидные модификации [42,43]. OGG1 распознает и удаляет 8-ОхоG из ДНК, создавая апуриновые/апиримидиновые сайты (АП-сайты), которые затем обрабатывается АП-эндонуклеазой [41,44].

Таким образом, до сих пор не существует четко установленных мутационных признаков окислительного повреждения мтДНК [34,45]. Однако нельзя исключать возможность появления других окисленных оснований в митохондриях, например, таких как 5-Гидроксиурацил (5-OHU), чей уровень в ДНК тканей млекопитающих и клеток человека сравним с таковым у 8-ОхоG [43], и 5-Гидроксицитозин (5-OHC) [46–48]. Окисление цитозина приводит к образованию цитозин гликоля, которое может либо дегидрироваться до 5-OHC, либо деаминироваться до урацил гликоля. 5-OHU возникает либо в результате деаминации 5-OHC, либо дегидрирования урацил гликоля. Все три модификации (за исключением нестабильного цитозин гликоля) обладают мутагенным эффектом и могут вызывать Ц>Т/Г>А транзиции [49], хотя наиболее мутагенным эффектом обладает 5-OHU (83%) [50]. Редко встречается окисленная модификация аденина, аналогичная модификации гуанина - 8-оксоаденин (8ОхоА), которая показала слабую способность вызывать мутации А>Г и А>Ц в клетках млекопитающих [51].

Кроме того, широко известно, что в митохондриях спонтанное дезаминирование цитозина и аденина, преимущественно происходящее на одноцепочечной мтДНК, приводит к заменам Ц>Т/Г>А и А>Г/Т>Ц соответственно [52,53]. В недавней работе Ильющенко и др. 2024 (Iliushchenko et al. 2024) авторы обсуждают потенциальные механизмы образования мутаций в нейтральных позициях митохондриального генома, выделяя два основных фактора репликационные повреждения, связанные с ошибками митохондриальной ДНК-полимеразы гамма, и повреждения, связанные с присутствием в митохондрии эндогенных мутагенов. Последние представляют собой разнообразные агенты химического мутагенеза, индуцирующие как окисление, метилирование и дезаминирование [54]. Согласно их результатам, химическое повреждение мтДНК усиливается с повышением метаболизма и уровня гипоксии у различных групп хордовых по сравнению с опухолевыми образцами человека. Анализируя 192-компонентный мутационный спектр авторы показали, что замены А>Г на тяжелой цепи митохондриальной ДНК заметно преобладают у птиц - группы с самой высокой скоростью метаболизма по сравнению с другими группами хордовых. И наоборот, злокачественные опухоли человека, вероятно, в связи с их относительно гипоксическим состоянием, демонстрируют пониженную частоту А>Г на тяжелой цепи.

Предположительно, мутационный спектр мтДНК всё-таки может быть чувствителен к химическому повреждению из-за видоспецифичной скорости аэробного метаболизма. Например, уровень метаболизма сильно зависит от температуры тела животного. Влияние температуры на скорость мутаций [55-58],интенсивно изучалось протяжении десятилетий на однако температурно-специфическая мутационная сигнатура, то есть влияние температуры на изменения в мутационном спектре совсем недавно стало предметом первых исследований, где мтДНК не была в центре внимания. Недавние эксперименты по накоплению мутаций в Chironomus riparius и E. coli показали, что доля транзиций увеличивается в высокотемпературных условиях [59,60].

Дополнительно, помимо точечных мутаций для лучшего понимания мутационных процессов в мтДНК, необходимо проанализировать делеционные спектры. Известно, что соматические делеции мтДНК могут накапливаться с возрастом в постмитотических тканях и ассоциированы с различными фенотипами старения, такими как нейродегенерация и саркопения. Авторами работы Шаманский и др. 2023 (Shamanski et al. 2023) были проведены первые анализы по подсчету делеционного груза и его связи с гаплогруппами [61]. Глубокое понимание причин формирования соматических мтДНК делеций человека сможет более эффективно направить развитие митохондриальной медицины в сторону уменьшения риска делеций. Научившись кластеризовать наблюдаемые соматические делеции В разных гаплогруппах, можно предсказывать гаплогрупп-специфические спектры делеций и, потенциально, риски болезней.

Принимая во внимание недавний прогресс в расшифровке вариаций мутационных спектров ядерного генома в зависимости от различных типов опухолей [62], факторов окружающей среды [33], нокаутов генов [63], человеческих популяций [64] и видов приматов [65] в данной работе фокус сосредоточен на мутационных спектрах мтДНК. Учитывая тесную связь уровня метаболизма и параметрами функционирования митохондриального генома (и, следовательно, потенциальных мутагенов мтДНК) у позвоночных животных предполагается существование ассоциации между мутационным спектром мтДНК и видоспецифичными особенностями жизненного цикла.

#### ГЛАВА 2. МАТЕРИАЛЫ И МЕТОДЫ

### 2.1 Обозначение тяжелой цепи 12-компонентного мутационного спектра мтДНК.

Хотя традиционно принято оперировать заменами мтДНК в нотации легкой цепи, которая соответствует референсным последовательностям мтДНК, в данной работе используется комплементарной тяжелой цепи, как это было сделано ранее рядом других авторов [66,67]. Тяжелая цепь была выбрана для нотации, поскольку она более чувствительна к приобретению мутаций и, следовательно, такая нотация будет более значимо отражать природу митохондриального мутагенеза [4,34,52]. В дальнейшем, чтобы упростить биологическую интерпретацию мутационного спектра мтДНК, используется 12компонентный спектр, основанный в нотации тяжелой цепи, обозначаемый подстрочным индексом т.

#### 2.2 Анализ дуплексного секвенирования мтДНК.

Все данные о соматических мутациях мтДНК, полученные с помощью метода дуплексного секвенирования, были получены от проф. Санчес-Контрерас и др. 2021 (Sanchez-Contreras et al. 2021) [52]. Дополнительно, были использованы данные дуплексного секвенирования контрольной мтДНК человека, полученные из двух источников [68,69] с указанным возрастным интервалом 10-30 и 80-90 лет соответственно.

# 2.3 Реконструкция видоспецифичного мутационного спектра всех доступных видов млекопитающих.

Используя все доступные внутривидовые последовательности (апрель 2016 г.) генов, кодирующих митохондриальные белки, был получен мутационный спектр для каждого вида. Все доступные последовательности мтДНК любых доступных белок-кодирующих генов для всех видов позвоночных были выгружены из баз данных, была реконструирована внутривидовая филогения, затем используя последовательность аутгруппы (ближайший вид к анализируемому), были реконструированы спектры предковых состояний во всех позициях во всех узлах внутреннего дерева и, наконец, получен список однонуклеотидных замен для каждого гена каждого вида. Для выполнения данной задачи был разработан программный комплекс [70]. Наиболее важными этапами программного комплекса были следующие. После получения всех последовательностей доступных нуклеотидных генов, кодирующих митохондриальные белки, создается локальная нуклеотидная база данных BLAST. Далее извлекаются внутривидовые последовательности из этой базы данных с использованием программного обеспечения tblastn (пакет ncbi-blast 2.6.0+) и поданной на вход RefSeq последовательности запрашиваемого белка данного вида и данного митохондриального гена. На следующем этапе все внутривидовые последовательности для данного вида и данного гена несколько выравниваются с использованием macse v1.01b. Это программное раз обеспечение выполняет множественные выравнивания на основе используемых кодонов (с использованием стандартного митохондриального генетического кода) и позволяет проверить наличие внутренних стоп-кодонов в анализируемых последовательностях (такие позиции удаляются из выравнивания). Затем реконструируются наследственные последовательности в каждом внутреннем узле дерева. Реконструкция внутривидовых предковых последовательностей основана на топологии дерева максимального правдоподобия и была создана с использованием двух альтернативных подходов: максимальной парсимонии и максимального правдоподобия. Неукорененная топология дерева внутривидовых последовательностей была получена с помощью RaxML v.8.2.9 и опции «-т GTRGAMMAIX». Для выбора лучшей топологии дерева использовалось лучшее дерево из 50 альтернативных прогонов, выполненных на основе различных начальных деревьях (опция «-N 50»). После этого дерево было аутгруппы. укоренено на последовательности Эта последовательность ближайших видов была найдена как первая последовательность выравнивания blast, следующая за последовательностями данных видов в подготовленной базе данных последовательностей. Предки с использованием подхода максимальной

парсимонии реконструированы с использованием программы dnapars из пакета philip v.3.697. Реконструкция максимального правдоподобия (реконструкция маргинальных предковых состояний) выполнялась с использованием опций RaxML v.8.2.9 и «-m GTRGAMMAIX -f A». Для реконструкции предков использовались модели замены на основе нуклеотидов (а не кодонов), поскольку анализировались только четырехкратно вырожденные позиции в кодонах.

#### 2.4 Подбор экологических параметров

Время генерации поколения в днях (как средний возраст родителей текущей когорты, отражающий скорость оборота размножающихся особей в популяции) млекопитающих было загружено из базы данных «Dryad»: https://doi.org/10.5061/dryad.gd0m3 [71].

Среднегодовая температура окружающей среды (воды) в градусах Цельсия и время полового созревания в годах (средний или медианный возраст наступления первой зрелости, при котором 50% популяции нерестится впервые) для рыб были загружены с https://www.fishbase.se/ (по состоянию на сентябрь 2019 г.)[72].

Температура тела была получена из базы данных AnAge (https://genomics.senescence.info/species/index.html) [73]. Температура птиц была получена из данного исследования [74].

#### 2.5 Анализ полных митохондриальных геномов

Полные митохондриальные геномы были загружены из базы данных GenBank, используя следующий поисковый запрос: 'Chordata [Organism] AND (complete genome [All Fields] AND mitochondrion [All Fields] AND mitochondrion [filter]'. На первом этапе были экстрагированы неперекрывающиеся области генов, кодирующих белки, определили использование кодонов и извлекли фракции нуклеотидов  $A_T$ ,  $\Gamma_T$ ,  $T_T$  и  $U_T$  в синонимичных четырехкратно вырожденных позициях.  $\Gamma_T A_T$ -сдвиг (англ. skew) нуклеотидов рассчитывался по формуле  $\Gamma_T A_T$ -сдвиг= $(\Gamma_T - A_T)/(\Gamma_T + A_T)$ , используя только синонимичные четырехкратно вырожденные позиции для каждого белок-кодирующего гена каждого RefSeq митохондриального генома всех доступных видов млекопитающих.

#### 2.6 Анализ времени, проведенного ДНК в одноцепочечном состоянии

Основываясь на асинхронном режиме репликации мтДНК и предполагая постоянную скорость репликации ДНК-полимеразой в пределах большой и малой дуг репликации мтДНК, относительное время проведенное ДНК в одноцепочечном состоянии (ВПДОС) для белок-кодирующих генов, кодируемых на тяжелой цепи мтДНК человека, было рассчитано по формуле ВПДОС<sub>для большой дуги</sub>=(Расположение гена – Точка начала репликации легкой цепи)\*2. Для всех видов млекопитающих ранг геноспецифического ВПДОС был одинаковым.

#### 2.7 Анализ дополнительных метрик

Асимметрия кодонов XXЦ<sub>L</sub> определялась как медианное значение XXЦ<sub>L</sub>/(XXЦ<sub>L</sub>+XXT<sub>L</sub>) для каждой аминокислоты. Аналогично, асимметрия XXA<sub>L</sub> определялась как медианное значение XXA<sub>L</sub>/(XXA<sub>L</sub>+XXГ<sub>L</sub>).

 $S_{TT}$ - $S_{AII}$  рассчитывалась как разность сумм пар относительных частот нуклеотидов по формуле  $S_{TT}$ - $S_{AII}$ =(сумма( $T_T$ + $\Gamma_T$ )сумма( $A_T$ + $II_T$ )/(сумма( $A_T$ + $II_T$ )+сумма( $T_T$ + $\Gamma_T$ )), где сумма( $A_T$ + $II_T$ )+сумма( $T_T$ + $\Gamma_T$ ) = 1.

Весь статистический анализ проводился на языке R. Для анализа филогенетической инерции использовался филогенетический обобщенный метод наименьших квадратов (PGLS - англ. Phylogenetic Generalised Least Squares), пакет «Caper», версия 1.0.1. Все сырые данные и скрипты депонированы в репозитории GitHub по ссылкам https://github.com/polarsong/mtDNA\_mutspectrum/, https://github.com/mitoclub/MutSpecOfActinopterygii.

#### 2.8 Нормализация мутационных спектров

Доли наблюдаемых замен зависят от частот предковых нуклеотидов в синонимичных четырехкратно вырожденные позициях, которые крайне неоднородны. Нормализация наблюдаемых нуклеотидных замен проводилась по частоте предковых нуклеотидов, после чего был получен мутационный спектр как вероятность мутации данного нуклеотида в любой другой нуклеотид независимо от его частоты в геноме. Каждый видоспецифичный мутационный спектр, представленный в виде вектора двенадцати скоростей замен, дополнительно трансформирован в частоты так, чтобы общая сумма всех скоростей замен была равна единице.

#### ГЛАВА 3. РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

## 3.1 Мутационные спектры млекопитающих: замена A<sub>T</sub>>Г<sub>T</sub> ассоциирована с продолжительностью жизни.

## 3.1.1 Частота de novo мутаций A<sub>T</sub>>Г<sub>т</sub> увеличивается с возрастом в соматических тканях и тканях зародышевой линии.

Митохондриальный геном характеризуется сильной асимметрией цепей по нуклеотидному составу: тяжелая цепь богата гуанином ( $\Gamma_T$ ) и бедна цитозином ( $\Pi_T$ ), а легкая цепь наоборот - богата цитозином ( $\Pi_T$ ) и с низким содержанием гуанина ( $\Gamma_T$ ). Мутационное объяснение этой асимметрии основано на предположении, что тяжелая цепь мтДНК, будучи в одноцепочечном состоянии во время асинхронной репликации, более восприимчива к двум наиболее распространенным мутациям в мтДНК:  $\Pi_T > T_T$  и  $\Lambda_T > \Gamma_T$ , приводящим к дефициту  $\Pi_T$  и избытку  $\Gamma_T$ . Ранее анализы полных митохондриальных геномов млекопитающих дополнительно выявили, что эта нуклеотидная асимметрия формирует нуклеотидный градиент вдоль мтДНК [25,75,76]: глобальный дефицит  $\Pi_T$  по сравнению с  $T_T$  и  $\Lambda_T$  по сравнению с  $\Gamma_T$  в третьей позиции кодонов становится более выраженным вдоль большой дуги репликации мтДНК от гена субъединицы I цитохром с-оксидазы ( $\Pi_T$ ) к гену цитохрома b ( $\Pi_T$ ) на фоне асинхронного режима репликации мтДНК: две наиболее частых транзиции  $U_T > T_T$  и  $A_T > \Gamma_T$  чаще встречаются в области гена ЦИТЬ, которая провела значительно больше времени в одноцепочечном состоянии по сравнению с областью гена ЦОГ1 [66] (Рисунок 3.1). Недавно большая коллекция соматических мутаций, полученная с помощью высокочувствительного подхода дуплексного секвенирования, позволила точно реконструировать градиенты  $U_T > T_T$  и  $A_T > \Gamma_T$  и однозначно подтвердила мутагенный эффект ВПДОС во время асинхронной репликации [52] (Рисунок 3.1).



Рисунок 3.1 — Схема репликации мтДНК

Примечание: на рисунке изображены: дочерняя тяжелая цепь - пунктирная черная линия; родительская тяжелая цепь - сплошная утолщающаяся черная линия, отражающая ВПДОС; дочерняя легкая цепь - пунктирная серая линия; родительская легкая цепь - сплошная серая линия; О<sub>н</sub> - точка начала репликации дочерней тяжелой цепи; О<sub>L</sub> - точка начала репликации дочерней легкой цепи.

Недавнее подтверждение такой мутационной природы градиентов замен в мтДНК [52] обеспечивает надежную основу для дальнейших исследований мутационных спектров мтДНК. Было отмечено, например, что положительный градиент Г<sub>Т</sub>/А<sub>т</sub> значительно различается между видами приматов и выше у видов с более длительным периодом беременности, в то время как градиент Т<sub>т</sub>/Ц<sub>т</sub> не

показывает сильных видоспецифичных вариаций [66]. Это говорит о том, что мутации A<sub>T</sub>>Г<sub>T</sub>, формирующие градиент Г<sub>T</sub>/A<sub>T</sub> могут быть чувствительны к некоторым мутагенам, связанным со сроком беременности или другими особенностями жизненного цикла. В связи с наличием положительных корреляций между продолжительностью беременности, размерами тела и продолжительности жизни, которые, в свою очередь, связаны со временем генерации поколения и уровнем митохондриального метаболизма, можно ожидать различий в мутагенезе мтДНК между видами с разными параметрами [77–79]. Чтобы жизненных циклов проверить ЭТУ гипотезу были проанализированы наборы данных соматических мутаций в мтДНК мышей и людей.

Для наиболее проверки потенциальной чувствительности двух распространенных транзиций к параметрам жизненного цикла позвоночных, был проанализирован недавний набор de novo соматических мутаций в мтДНК, полученных с помощью метода дуплексного секвенирования. В работе Санчес-Контрерас [52] были представлены три группы образцов: молодые мыши (4-5 мес), старые мыши (26 мес) и люди (10–90 лет), что позволило авторам доказать, что мутационный градиент усиливается с возрастом, поэтому было решено сосредоточиться на сравнении интенсивности градиентов Ц<sub>Т</sub>>T<sub>T</sub> и A<sub>T</sub>>Г<sub>т</sub>. Прежде всего, в сравнительных целях были построены градиенты Ц<sub>T</sub>>T<sub>T</sub> и A<sub>T</sub>>Г<sub>T</sub> в одном и том же масштабе (Рисунок 3.2) и использованы модели линейной регрессии, описывающие частоты двух наиболее распространенных мутаций в зависимости от ВПДОС (Таблица 3.1).

Образец	Тип мутации	Интерсепта	Наклон	Р-значение
Молодые мыши	Цт>Тт	20.273***	0.355	<0.001***
Молодые мыши	$A_T > \Gamma_T$	1.736***	0.021	<0.001***

	2	1	ъ		v	U
Габлица	i - <b>1</b>		Pegy	VILTATLI	пинеиных	молепеи
таолица	5	• 1	TCJ	yJIDIAIDI	JIMICHIDIA	моделен

Старые мыши	Щт>Тт	83.267***	0.703	<0.001***
Старые мыши	$A_T > \Gamma_T$	3.915***	0.127	<0.001***
Человек	Ц <sub>Т</sub> >Т <sub>Т</sub>	331.135***	4.866	<0.001***
Человек	$A_T > \Gamma_T$	100.536***	1.482	<0.001***

Примечание: A<sub>T</sub>>Г<sub>T</sub> — наиболее чувствительная к возрасту транзиция: увеличение наклонов от молодых мышей к старым мышам и к людям является самым сильным (от 0,021 до 0,127 и до 1,482).

Как и ожидалось, и как уже было показано в работе Санчес-Контрерас, градиенты обеих транзиций демонстрируют увеличение точек пересечения с осью Y и наклонов с возрастом.

Далее, были проанализированы наклоны регрессионных прямых как аппроксиматоры чувствительности мутаций к ВПДОС. Относительная крутизна наклонов сравнивалась между молодыми мышами, старыми мышами и человеком (Рисунок 3.3). Было обнаружено, что крутизна наклонов выше для  $A_T > \Gamma_T$  по сравнению с  $U_T > T_T$ : она в 6,05 (1,98 для  $U_T > T_T$ ) раз выше у старых мышей по сравнению с молодыми мышами, в 11,7 (6,9 для  $U_T > T_T$ ) раз выше у людей по сравнению с остарыми мышами и в 70,86 (13,7 для  $U_T > T_T$ ) раз выше у людей по сравнению с молодыми мышами. Дополнительно, проведя бутстреп с тысячью итераций (методом генерации повторной выборки) каждого процентиля ВПДОС с последующим пересчетом шести уровней крутизны наклонов регрессионных прямых (упомянутых выше) значимость полученных результатов была подтверждена: увеличение крутизны наклона  $A_T > \Gamma_T$  с возрастом сильнее, чем увеличение крутизны наклона  $U_T > T_T$  (Рисунок 3.4, все р-значения < 10<sup>16</sup>, Uкритерий Манна–Уитни).



Рисунок 3.2 — Градиенты мутаций Ц<sub>Т</sub>>Т<sub>т</sub> и А<sub>Т</sub>>Г<sub>т</sub> вдоль большой дуги мтДНК Примечание: градиенты более выражены у людей по сравнению со старыми мышами и у старых мышей по сравнению с молодыми мышами: как точки пересечения с осью Y, так и наклоны увеличиваются с возрастом выборки.



Рисунок 3.3 — Изменение крутизны наклонов регрессионных прямых в зависимости от возрастной группы и типа мутации.



Рисунок 3.4 Соотношение крутизны наклонов, полученных методом генерации повторной выборки (ресемплинг)

Примечание: «\*\*\*» означает на рисунке р-значение <0,001.

Далее, был проведен глубокий ресемплинг каждой соматической мутации по данным дуплексного секвенирования для подтверждения значимости полученных результатов, который показал, что частота транзиций A<sub>T</sub>>Г<sub>T</sub> действительно увеличивается стремительнее с возрастом по сравнению с частотой Ц<sub>Т</sub>>Т<sub>Т</sub>. Учитывая высокую вариабельность транзиций числа наблюдаемых мутантных молекул и молекул "дикого типа" для каждого типа мутаций ( $A_T > \Gamma_T$  и  $\coprod_T > T_T$ ), каждого процентиля вдоль большой дуги репликации мтДНК и каждого образца (молодые мыши, старые мыши и человек) был проведен дополнительный анализ методом генерации повторной выборки. Основная цель этого анализа состояла в том, чтобы продемонстрировать устойчивость описанных результатов к стохастичности из-за различных частот наблюдаемых мутаций. В каждом процентиле каждого образца были отобраны 400 000 случайных молекул (часть мутантных и часть "дикого типа"). 400 000 было числом выбора, поскольку оно было меньше минимального числа наблюдаемых молекул каждого типа в каждом процентиле каждого образца. Дополнительное тестирование показало, что это число не должно качественно повлиять на результат. Затем в каждом образце для каждого процентиля для каждого типа замещения на основе 400 000 случайно отобранных молекул была реконструирована частота этой мутации (разделив количество случайно выбранных мутантных молекул на количество случайно выбранных молекул "дикого типа"). Далее была построена модель линейной регрессии, в которой частота мутаций каждого типа была описана как функция ВПДОС, которая была аппроксимирована последовательным расположением перцентилей вдоль большой дуги репликации мтДНК. В итоге были получены шесть значений крутизны наклонов (два типа замещения \* три вида образца), описывающих частоту мутаций как функцию ВПДОС. Дополнительно, 500 раз была проведена процедура ресемплинга и получены распределения этих шести наклонов (Рисунок 3.5, Таблица 3.2). Как и прежде, было обнаружено, что медианные значения крутизны наклонов увеличиваются с возрастом образца.

Образец	$A_T > \Gamma_T * (10^{-9})$	$\Pi_{T} > T_{T} * (10^{-9})$
Молодые мыши 1.98 34.2		34.29
Старые мыши	11.99	69.16
Человек 149.62		482.91

Таблица 3.2 — Медианные значения шести наклонов



Рисунок 3.5 — Распределение шести наклонов

Примечание: пунктирные вертикальные линии отмечают соответствующие медианы.

Далее, крутизна наклонов более старых образцов была разделена на крутизну наклонов более молодых образцов для одного и того же типа замещения с получением относительных увеличений крутизны наклонов. Например, относительное увеличение крутизны наклона между старыми и молодыми мышами составляет 6,05 для  $A_T > \Gamma_T$  (11,99/1,98 ~ 6,05), а для  $U_T > T_T$  всего 2,02 (69,16/34,29 ~ 2,02). Дополнительные сравнения показывают такую же тенденцию. Все эти результаты полностью согласуются с выводами, полученными без процедуры ресемплинга. Чтобы получить р-значения для различий между относительными увеличении крутизны наклона между  $A_T > \Gamma_T$  и  $U_T > T_T$ , сравнивались соответствующие распределения (Таблица 3.3, Рисунок 3.6).

Сравниваемые образцы	<b>Α</b> <sub>T</sub> >Γ <sub>T</sub>	Цт>Тт	Наклон А <sub>Т</sub> >Гт круче наклона Ц <sub>Т</sub> >Тт (р-значения U-критерий Манна– Уитни)
Старые мыши/молодые мыши	6.05	2.02	p < 1.109e-11
Человек/старые мыши	12.47	6.98	p < 2.2e-16
Человек/молодые мыши	75.55	14.08	p < 2.2e-16

Таблица 3.3 — Относительные увеличения крутизны наклонов



Рисунок 3.6 — Относительные увеличения крутизны наклонов

Таким образом, ожидалось, что мутации  $A_T > \Gamma_T$  из-за более быстрого увеличения крутизны наклонов с возрастом будут вносить пропорционально больший вклад в образцы от старых организмов, чем  $U_T > T_T$ . Для этого был проанализирован мутационный спектр молодых мышей, старых мышей и людей и продемонстрировано, что общая фракция  $A_T > \Gamma_T$  действительно увеличивается с возрастом (Рисунок 3.7, все р-значения < 1,583е-08, U-критерий Манна-Уитни), тогда как фракция  $U_T > T_T$  не демонстрирует монотонного увеличения с возрастом (Рисунок 3.8, все р-значения < 1,583е-05, U-критерий Манна-Уитни).



Рисунок 3.7 — Частота  $A_T > \Gamma_T$  в общем мутационном спектре увеличивается с

#### возрастом

Примечание: «\*\*\*» означает на рисунке р-значение <0,001.





В целом, используя наборы данных соматических мутаций мтДНК, полученные с помощью высокочувствительного метода дуплексного секвенирования, было обнаружено, что A<sub>T</sub>>Г<sub>T</sub> более чувствителен к возрасту по сравнению с Ц<sub>T</sub>>T<sub>T</sub>.

Предполагая сходство мутагенеза мтДНК в соматических и тканях зародышевой линии, ожидалось увидеть избыток  $A_T > \Gamma_T$  в старых образцах тканей зародышевой линии. Действительно, недавнее глубокое секвенирование de novo мутаций мтДНК в ооцитах старых и молодых мышей подтвердило, что самым сильным признаком старения ооцитов является увеличение доли замен  $A_T > \Gamma_T$  [80]. Поэтому, дополнительно были проанализировали de novo мутации мтДНК человека в зависимости от репродуктивного возраста женщины, который является показателем возраста ооцитов. Было показано, что количество de novo

мутаций мтДНК у детей увеличивается с возрастом матери [81,82], однако никаких возрастных изменений в мутационных спектрах мтДНК пока не зарегистрировано из-за небольшого размера выборки de novo мутаций. Поэтому, повторно были проанализировали все de novo мутации зародышевой линии из двух недавних исследований пар мать-ребенок. Из-за небольшого доступного размера выборки все анализы de novo мутаций в парах мать-ребенок были лишь наводящими на размышления, но постоянно демонстрировали тенденцию увеличения доли A<sub>T</sub>>Г<sub>T</sub> с возрастом ооцита. Анализ данных, полученных у Реболледо-Джарамилло и др. 2014 (Rebolledo-Jaramillo et al. 2014), показал, что фракции  $A_T > \Gamma_T$ увеличение характеризуется наивысшим возрастом оплодотворения (средний возраст оплодотворения 34,1, N=4) по сравнению с Ц<sub>Т</sub>>Т<sub>т</sub> (средний возраст оплодотворения 29,9, N=4) или всеми остальными заменами (средний возраст оплодотворения = 32,1, N=12). Из-за небольшого размера выборки (всего 16 мутаций) эта тенденция не была значимой (p = 0,19, если сравнивать A<sub>T</sub>>Г<sub>T</sub> со всеми другими заменами; p=0,0956, если сравнивать А<sub>Т</sub>>Г<sub>Т</sub> с Ц<sub>Т</sub>>Т<sub>Т</sub>; односторонний U-критерий Манна-Уитни). При анализе данных, полученных у Вей и др. 2019 (Wei et al. 2019), доступа к репродуктивному возрасту не было, поэтому использовалась такую метрику, как частота вариантного аллеля (ЧВА, англ. variant allele frequency) каждого de novo варианта мтДНК в зависимости от возраста ооцита. Ожидалось, что мутации A<sub>T</sub>>Г<sub>T</sub> встречаются преимущественно старых ооциты И, V таким образом, характеризуются более низкой частотой по сравнению с другими заменами. Чтобы уменьшить потенциальный эффект отбора для анализа брались только редкие (с ЧВА <10%) замены. Далее, создав подмножество потомков с как минимум двумя редкими de novo транзициями мтДНК, одна из которых представляет собой  $A_T > \Gamma_T$ , а другая не  $A_T > \Gamma_T$  (всего 23 потомка), было показано, что  $A_T > \Gamma_T$  в среднем имеет частоту на 1% ниже по сравнению с пациентами без  $A_T > \Gamma_T$ . Статистическая разница между  $\text{ЧВА}(A_T > \Gamma_T)$  и  $\text{ЧВА}(\text{нет } A_T > \Gamma_T)$  была незначительно значима (p=0,05015, парный односторонний U-критерий Манна-Уитни). Однако, основываясь на всех полученных результатах, был сделан вывод что повышение частоты de novo мутаций  $A_T > \Gamma_T$  характеризовалось самым высоким возрастом репродукции. Полученные результаты показывают, что частота de novo мутаций  $A_T > \Gamma_T$  увеличивается с возрастом как в соматических тканях, так и в тканях зародышевой линии млекопитающих.

Ожидается, что внутривидовые сравнения (старые и молодые мыши или старые и молодые люди) более надежные, поскольку мутагенез мтДНК внутри одного и того же вида, скорее всего, очень похож. Однако, предполагая, что основные правила мутагенеза мтДНК достаточно стабильны для всех видов млекопитающих, было принято решение расширить логику и провести сравнительный анализ видов, как описано в следующем разделе.

# 3.1.2 А<sub>Т</sub>>Г<sub>Т</sub> наиболее часты у млекопитающих с большим временем генерации поколения.

Вариации в мутационных спектрах мтДНК между разными видами [10,11] ранее не имели общего объяснения. Полученные результаты, ровно как и литературные данные [80] позволяют предположить, что эти вариации и особенно транзиции Ат>Гт могут быть связаны со старением. Таким образом, предполагалось, что видоспецифичные мутационные спектры мтДНК зависят от времени генерации поколения, которое, в свою очередь, является хорошим млекопитающих. Поскольку показателем возраста ооцитов y ооциты млекопитающих замирают от рождения до полового созревания, что занимает недели для мышей или десятилетия для людей [83], было решено использовать видоспецифичное время генерации поколения (англ. generation length) в качестве естественного показателя возраста ооцитов у различных видов млекопитающих. Кроме того, поскольку ооциты являются единственной клеточной линией, через которую мтДНК передается (3a редкими исключениями отцовского наследования) из поколения в поколение у млекопитающих [84], ожидалось увидеть корреляцию между видоспецифичными свойствами мтДНК и временем генерации поколения (аппроксиматор возраста ооцитов) у разных видов млекопитающих. Время генерации поколения определяется как средний возраст

родителей текущей популяции [71,85]; оно доступно для подавляющего большинства видов млекопитающих, а также ассоциировано с многочисленными экологическими (масса тела, размер помета, эффективная численность популяции) и физиологическими (базальный уровень метаболизма) параметрами млекопитающих [86,87].

Многочисленные последовательности мтДНК, полученные в результате экологических, эволюционных и популяционных генетических исследований различных видов животных [88] являются ценным источником полиморфизмов мтДНК, используемых в данной работе. На основе уникальной разработки (программный комплекс, см. Материалы и методы) был реконструирован мутационный спектр сотен видов млекопитающих. Мутационный спектр определяется, как вероятность того, что каждый нуклеотид мутирует в другой, основанная на наблюдаемых и нормализованных частотах двенадцати типов нуклеотидных замен в четырехкратно вырожденных синонимичных сайтах всех доступных внутривидовых полиморфизмов генов, кодирующих белки мтДНК. Вкратце, были (i) скачаны все доступные нуклеотидные последовательности митохондриальных генов млекопитающих, кодирующих митохондриальные белки, (ii) были получены множественные выравнивания кодонов для каждого гена каждого вида, (iii) было укоренено внутривидовое митохондриальное дерево по последовательности ближайшего вида, (iv) было реконструированы предковые последовательности в каждом внутреннем узле, (v) был получен список поляризованных однонуклеотидных замен, (vi) которые были нормализованы по частоте предковых нуклеотидов. Сосредоточив внимание на наиболее нейтральных (здесь и далее имеются ввиду замены, происходящие в четырехкратно вырожденных синонимических позициях) 70 053 заменах, был 611 реконструирован нейтральный мутационный спектр для видов Средний мутационный спектр всех доступных млекопитающих. видов млекопитающих (Рисунок 3.9) демонстрирует сильный избыток замен A<sub>T</sub>>Г<sub>T</sub> и  $\coprod_{T} > T_{T}$ , что было показано в предыдущих исследованиях [4,20].

34



Рисунок 3.9 — Средний мутационный спектр мтДНК видов млекопитающих (N = 611)

Примечание: на правой панели отражено содержание нуклеотидов в четырехкратно вырожденных синонимических сайтах всех доступных белков мтДНК.

После нормализации (см. Материалы и методы) 12 типов замен, изображенных на левой панели, по содержанию нуклеотидов, изображенному на правой панели был получен нормализованный спектр (Рисунок 3.10), используемый в последующих анализах. Например, скорость замещения T<sub>T</sub>>Ц<sub>T</sub> значительно снизилась после нормализации на высокое содержание нуклеотидов T<sub>T</sub>, в то время как фракция Ц<sub>T</sub>>T<sub>T</sub> увеличилась после нормализации на очень небольшое количество нуклеотида Ц<sub>T</sub>.



Рисунок 3.10 — Средний нормализованный мутационный спектр мтДНК видов млекопитающих (N = 611)

Чтобы сосредоточиться на видоспецифичной изменчивости мутационных спектров, был подробно проанализирован ген ЦИТЬ, который был наиболее распространенным в полученной ранее базе данных митохондриаьных геномов: он содержал 56% всех извлеченных замен (39 112 из 70 053, использованных для построения рисунка 3.9). Сравнение спектра, полученного с помощью СҮТВ, между видами позволило устранить эффект градиента (см. раздел 3.1.1) и сосредоточиться на потенциальном влиянии особенностей жизненного цикла. В качестве простейшей метрики мутационного спектра для каждого вида сначала было рассчитано соотношение транзиций/трансверсий (Ts/Tv) как сумму частот всех транзиций, деленную на сумму частот всех трансверсий. Для 424 видов млекопитающих с реконструированным Ts/Tv и известным временем генерации поколения обнаружилась положительная корреляцию между ними (Рисунок 3.11, ро-коэффициент Спирмена = 0.23, р-значение = 2.021e-06, N = 424).


Рисунок 3.11 — Положительная корреляция между временем генерации поколения и Ts/Tv у млекопитающих

Чтобы доказать достоверность увеличения Ts/Tv с временем генерации поколения, все млекопитающие были разделены на несколько групп: (i) по квартилям времени генерации поколения (Рисунок 3.12, все значения р <0,05, U-критерий Манна-Уитни) и (ii) по медианам времени генерации поколения (Рисунок 3.13, р-значение = 1,687е-06, U-критерий Манна-Уитни).



Рисунок 3.12 — Виды млекопитающих с большим временем генерации поколения демонстрируют повышенное Ts/Tv (группы разделены по квартилям)



Рисунок 3.13 — Виды млекопитающих с большим временем генерации поколения демонстрируют повышенное Ts/Tv (группы разделены по квартилям)

Увеличение Ts/Tv с увеличением продолжительности поколения также выражено на уровне семейств (Таблица 3.4, Рисунок 3.14). Для каждого семейства млекопитающих, содержащего не менее 3 видов в нашем наборе данных, было оценено медианное значение Ts/Tv и медианное значение времени генерации поколения. Используя ранговую корреляцию Спирмена была подтверждена положительная корреляция между этими значениями (р = 0,0016, ро-коэффициент Спирмена = 0,85, N = 11). В таблице ниже представлена одиннадцати двойные описательная статистика семей, a боксплоты визуализируют распределение Ts/Tv и времени генерации поколения в каждой семье.

Таблица 3.4 — Описательная статистика на основе семейств анализируемого набора данных: медиана Ts/Tv и медианное время генерации поколения. Сокращения, представленные в таблице, используются на Рисунке 3.14.

Семейство	Медианная Ts/Tv	Медианное время генерации поколения	Сокращение	Ν
Insectivora	8.81	427	Ins	41
Rodentia	7.46	601	Rod	120
Didelphimorphia	7.08	643	Did	18
Lagomorpha	6.55	1047	Lag	19
Chiroptera	10.75	2064	Chi	81
Carnivora	13.21	2435	Car	34
Cervidae	13.33	2555	Cer	7
Suidae	10.57	2606	Sui	4
Bovidae	13.34	2870	Bov	16
Primates	14.14	3806	Pri	49
Cetacea	27.74	5158	Cet	6





Примечание: Боксплоты были получены с использованием библиотеки R «boxplotdbl». Заполненная поверхность для каждой семьи отражает интерквартильный диапазон, перекрестие линий представляет медианы, а усы обозначают минимум и максимум, исключая выбросы (некоторые выбросы на графике не показаны). Положение трехбуквенной аббревиатуры для каждого семейства соответствует медиане времени генерации поколения и Ts/Tv соответствующего семейства.

Хотя несколько приведенных выше анализов продемонстрировали надежность ассоциации между Ts/Tv и временем генерации поколения для подтверждения достоверности результатов была использована обобщенная регрессия методом наименьших квадратов (PGLS). Из-за необходимости объединить набор данных с доступным филогенетическим деревом, сохранилось менее половины наших данных по видам (N = 211), но, несмотря на это, основная

Во-первых, тенденция была значительна. было показано, Ts/Tvчто положительно коррелирует с временем генерации поколения (модель 1А в Таблице 3.5). Во-вторых, в модель было добавлено количество мутаций как число реконструированных синонимичных четырехкратных вырожденных ЦИТЬ, которые были использованы реконструкции замен В для видоспецифичного мутационного спектра (модель 1В в Таблице 3.5). Несмотря на то, что количество мутаций отрицательно коррелирует с Ts/Tv (вероятно, изза большего размера выборки короткоживущих млекопитающих), влияние времени генерации поколения на Ts/Tv оставалось значимым. В-третьих, отметив, что интерсепта незначительно отклоняется от нуля в модели В та же модель была построена через начало координат (фиксируя точку пересечения на нуле), в результате чего была получена сильная и значимая корреляция между Ts/Tv и временем генерации поколения (модель С в Таблице 3.5). В финальной модели 1С, время генерации поколения связано с Ts/Tv гораздо более значимо по сравнению с «количеством мутаций».

Таблица 3.4 — Модели PGLS, описывающие взаимосвязи между Ts/Tv и временем генерации поколения.

Модель	Переменная	Коэффициент	р-значение
1A. Ts/Tv ~ log <sub>2</sub> (время генерации поколения); N=211, lambda [ML] = 0.000	Интерсепта	-26.4047	0.094182
	log <sub>2</sub> (время генерации поколения)	4.1771	0.004472
1В. Ts/Tv ~ log <sub>2</sub> (время генерации поколения) + log <sub>2</sub> (Общее количество мутаций); N=211, lambda [ML] = 0.000	Интерсепта	12.2181	0.558976
	log <sub>2</sub> (время генерации поколения)	2.9679	0.048715
	log2(Общее количество мутаций)	-4.3222	0.006407

1C. Ts/Tv ~ 0 + log <sub>2</sub> (время генерации поколения) +	log <sub>2</sub> (время генерации поколения)	3.75742	2.454e-08
log <sub>2</sub> (Общее количество мутаций); N=211, lambda [ML] = 0.000	log2(Общее количество мутаций)	-3.70511	0.001629

Чтобы понять, какой тип замены преимущественно формировал наблюдаемую корреляцию между Ts/Tv и временем генерации поколения, были проведены двенадцать анализов попарной ранговой корреляции между каждым типом замены и временем генерации поколения. Было показано, что только частота  $A_T > \Gamma_T$  положительно коррелировала с временем генерации поколения (Рисунок 3.15, ро-коэффициент Спирмена = 0,252, номинальное р-значение = 1,188e-07, N=424), тогда как несколько редких трансверсий показали слабую и отрицательную корреляцию ( $T_T > A_T$ ,  $T_T > \Gamma_T$ ,  $\coprod_T > A_T$  и  $\Gamma_T > T_T$ : все значения рокоэффициент Спирмена < -0,17, все номинальные значения P < 0,0003).



Рисунок 3.15 —  $A_T > \Gamma_T$  - тип замен, частота которых сильнее всего коррелирует

с временем генерации поколения

Включив все пять типов этих замен в множественную линейную модель, было показано, что A<sub>T</sub>>Г<sub>T</sub> является единственной трансверсией, сильнее всего связанной со временем генерации поколения (Формула 3.1):

$$log_2$$
(Время генерации поколения) ~ (Формула 3.1)  
10.39+0.26\*(A<sub>T</sub>> $\Gamma_T$ )-0.18\*(T<sub>T</sub>>A<sub>T</sub>)-0.16\*(Ц<sub>T</sub>>A<sub>T</sub>),  
N = 424, R<sup>2</sup>=0.104,

общее p-значение = 1,249е-10; p-значения трансверсий <0,002, p-значение A<sub>T</sub>>Г<sub>T</sub> = 6,42е-06, представленные коэффициенты масштабированы.

Далее, все пять типов замен были включены в исходную линейную модель. Следуя логике пошаговой множественной модели (удаляя на каждом шаге одну из самых незначимых переменных) было определено, что возможны три типа замен были независимо связаны с временем генерации поколения. Сравнивая масштабированные коэффициенты регрессии, было обнаружено, что самый сильный (самый высокий абсолютный размер эффекта), а также наиболее значимый (наименьшее p-значение) эффект был связан с транзициями  $A_T > \Gamma_T$ . Положительная корреляция времени генерации поколения с частотой транзиции  $A_T > \Gamma_T$  и отрицательная корреляция с частотами  $U_T > A_T$  и  $T_T > A_T$ , как и ожидалось, приводят к увеличению Ts/Tv у долгоживущих видов, описанных выше. Важно отметить, что включение в линейную модель (Формула 3.1) общего количества мутаций, используемых для оценки видоспецифичного мутационного спектра (синонимические мутации в четырехкратно вырожденных сайтах гена ЦИТb) не влияет на эти результаты количественно (Формула 3.2):

общее p-значение = 2,464e-14; p-значение  $A_T > \Gamma_T = 1,66e-05$ ; p-значение  $T_T > A_T = 0,0103$ ; p-значение  $U_T > A_T = 0,00249$ ; p-значение Общее количество мутаций = 5,29e-06; представленные коэффициенты масштабированы.

Видно, что общее количество мутаций отрицательно коррелирует с временем генерации поколения, вероятно, из-за того, что мутации короткоживущих мелких млекопитающих лучше изучены, и в GenBank хранится больше последовательностей. Аналогично проведенному выше анализу филогенетической инерции были использованы три модели, анализирующие ассоциацию A<sub>T</sub>>Г<sub>T</sub> с временем генерации поколения. Итоговая модель 1F (Таблица 3.5) показывает значительное положительное влияние времени генерации поколения на частоту A<sub>T</sub>>Г<sub>T</sub> и отсутствие влияния «общего количества мутаций».

Модель	Переменная	Коэффициент	р-значение
1D. А <sub>Т</sub> >Г <sub>Т</sub> ~ log <sub>2</sub> (время генерации поколения); N=211, lambda [ML] = 0.336	Интерсепта	0.0202212	0.8184
	log2(время генерации поколения)	0.0115248	0.1488
1Е. А <sub>Т</sub> >Г <sub>Т</sub> ~ 0+log <sub>2</sub> (время генерации поколения); N=211, lambda [ML] = 0.317	log <sub>2</sub> (время генерации поколения)	0.0133441	3.173e-08
1F. $A_T > \Gamma_T \sim 0$ + log <sub>2</sub> (время генерации поколения) + log <sub>2</sub> (Общее количество мутаций); N=211, lambda [ML] = 0.329	log <sub>2</sub> (время генерации поколения)	0.0123511	0.0003097
	log2(Общее количество мутаций)	0.0018587	0.6942430

Таблица 3.5 — Анализ филогенетической инерции (PGLS), описывающий взаимосвязи между частотой замен А<sub>Т</sub>>Г<sub>т</sub> и временем генерации поколения.

Чтобы доказать независимость полученных результатов от общего количества входных полиморфных мутаций, используемых для реконструкции 12-компонентных мутационных спектров, набор данных был разделен на несколько подмножеств с разными порогами минимального количества мутаций: 12 х 3, 5, 7, 9, 11, 13. Ниже приведены результаты ассоциаций между A<sub>T</sub>>Г<sub>T</sub> и временем генерации поколения с использованием корреляционных тестов (Таблица 3.6) и PGLS (Таблица 3.7). Во всех анализах была выявлена положительная связь между А<sub>Т</sub>>Г<sub>Т</sub> и временем генерации поколения. Интересно, что параллельно с ожидаемым увеличением р-значений с уменьшением размера 132 выборки в замены характеризуется порог самыми высокими коэффициентами в обоих типах тестов, что может быть связано с уменьшенным шумом и более надежной оценкой мутационного спектра.

Таблица 3.6 — Ранговые корреляции Спирмена A<sub>T</sub>>Г<sub>T</sub> в зависимости от времени генерации поколения, оцененные для разных порогов минимального количество мутаций (синонимичных четырежды вырожденных в гене ЦИТb), используемых для реконструкции мутационного спектра.

Минимальное количество мутаций, используемых для реконструкции мутационного спектра	N (количество видов, доступных для анализа)	Ро-коэффициент Спирмена	Номинальное р-значение
36	309	0.2431854	1.542e-05
60	211	0.2474482	0.0002838
84	159	0.2426535	0.002058
108	122	0.2375639	0.008419
132	88	0.3203167	0.002346
156	68	0.2740958	0.02371

Примечание: для всех протестированных порогов номинальные р-значения значимы, а коэффициенты положительны.

Таблица 3.7 — Анализ филогенетической инерции (PGLS), описывающий взаимосвязи между A<sub>T</sub>>Г<sub>T</sub> и временем генерации поколения, оцененные для разных порогов минимального количества мутаций, используемых для реконструкции мутационного спектра.

Минимальное количество мутаций, используемых для реконструкции мутационного спектра	N (количество видов, доступных для анализа)	Ро- коэффициент Спирмена	Номинальное р-значение
36	156	0.0124924	3.042e-05
60	97	0.0133371	0.0001337
84	73	0.0133791	0.001515
108	58	0.0115877	0.02612
132	41	0.015109	< 2.2e-16
156	32	0.014024	2.505e-12

Примечание: для всех протестированных порогов номинальные р-значения значимы, а коэффициенты положительны. Во всех моделях было замечено, что интерсепта незначительно отклоняется от нуля (как в модели А Таблицы 3.4), поэтому результаты анализа PGLS представлены после построения модели через начало координат (как в модели В Таблицы 3.4): A<sub>T</sub>>Г<sub>T</sub> ~ 0 + log<sub>2</sub>(время генерации поколения).

Чтобы проанализировать мутационные спектры мтДНК млекопитающих дополнительно нестандартным способом, было проведено снижение размерности параметров путем реконструкции принципиальных компонент (PCA) мутационных спектров 424 видов млекопитающих. Мы заметили, что первая компонента (PC1) в основном обусловлена наиболее распространенными заменами Ц<sub>Т</sub>>Т<sub>т</sub>, тогда как вторая (PC2) в основном обусловлена заменами A<sub>T</sub>>Г<sub>T</sub> (Рисунок 3.16). После корреляции первых десяти принципиальных компонент со временем генерации поколения было показано, что только вторая компонента существенно коррелировала Корреляция второй С ним. принципиальной компоненты со временем генерации была поколения существенно сильнее (ро-коэффициент = -0,39, р-значение < 2,2e-16, N = 424) по сравнению с единственным эффектом A<sub>T</sub>>Г<sub>T</sub> (ро-коэффициент Спирмена = 0,252, p-значение = 1,188e-07, N = 424) и единственным эффектом Ts/Tv (poкоэффициент Спирмена = 0,23, р-значение = 2,021e-06; N = 424) из чего можно сделать вывод, что вторая компонента может отражать сложную сигнатуру конкретного мутагена, связанного со временем генерации поколения.



Рисунок 3.16 — Ц<sub>Т</sub>>Т<sub>Т</sub> имеет самую высокую нагрузку на первую принципиальную компоненту, а А<sub>Т</sub>>Г<sub>Т</sub> имеет самую высокую нагрузку на вторую принципиальную компоненту.

Примечание: на график нанесена отрицательная PC2, чтобы отразить положительную корреляцию со временем генерации поколения. Правая панель: вторая принципиальная компонента коррелирует со временем генерации поколения млекопитающих. Время генерации поколения имеет цветовую маркировку от темно-зеленого (самое короткое время генерации поколения) до светло-зеленого (самое длинное время генерации поколения).

Действительно, обе трансверсии  $T_T > A_T$  и  $U_T > A_T$ , отрицательно связанные со временем генерации поколения (Формулы 3.2), имеют сильное влияние на вторую принципиальную компоненту и, как и ожидалось, направлены в противоположную сторону по сравнению с  $A_T > \Gamma_T$ .

Чтобы выяснить, влияет ли на вторую принципиальную компоненту общее количество мутаций, использованных для реконструкции мутационного спектра, было проведено несколько анализов. Было показано, что вторая принципиальная компонента положительно коррелирует с количеством мутаций (рокоэффициент Спирмена = 0,2247093, р-значение = 2,964е-06). Однако во множественной линейной модели, в которой вторая принципиальная компонента была функцией как времени генерации поколения, так и общего количества мутаций, влияние времени генерации поколения сильнее и значительнее (Формула 3.3):

PC2 ~ -0.34\*(Время генерации поколения) + (Формула 3.3)
0.16\*(Общее количество мутаций),
N = 424, R 2=0.109,

общее p-значение = 1,121e-11; p-значение время генерации поколения = 7,07e-09; p-значение общее количество мутаций = 0,006; представленные коэффициенты масштабированы.

Первая принципиальная компонента, поддержанная в основном наиболее распространенной транзицией Ц<sub>Т</sub>>Т<sub>т</sub>, по-видимому, представляет собой неизвестный источник вариаций, который не удалось связать с какими-либо особенностями жизненного цикла, такими как время генерации поколения, размер тела, температура тела, масса тела, скорость метаболизма или общее количество мутаций, используемых для получения мутационного спектра.

Дополнительные анализы, учитывающие влияние потенциальных ошибок секвенирования и содержания нуклеотидов, подтвердили достоверность наших результатов.

Поскольку редкие генетические варианты могут возникать из-за повреждения ДНК во время секвенирования [89–91] полученные результаты были повторно воспроизведены, приняв во внимание только две наиболее распространенные транзиции:  $A_T > \Gamma_T$  и  $U_T > T_T$ . Анализируя фракцию  $A_T > \Gamma_T$  (полученную как  $A_T > \Gamma_T/(A_T > \Gamma_T + U_T > T_T)$ ), положительная корреляция этой фракции со временем генерации поколения сохранилась (ро-коэффициент Спирмена = 0,21, р-значение = 8,441e-06, n = 424).

Аналогично анализу филогенетической инерции, который проводился ранее, были использованы три модели, демонстрирующие связь  $A_T > \Gamma_T / (A_T > \Gamma_T + U_T > T_T)$  со временем генерации поколения. Модель 11 (Таблица 3.8) показывает значительное положительное влияние времени генерации поколения на  $A_T > \Gamma_T / (A_T > \Gamma_T + U_T > T_T)$  и отсутствие влияния «общего количества мутаций».

Таблица 3.8 — Модели PGLS, описывающие взаимосвязи между A<sub>T</sub>> $\Gamma_T/(A_T>\Gamma_T+$ Ц<sub>T</sub>> $T_T$ ) и временем генерации поколения.

Модель	Переменная	Коэффициент	р-значение
1G. А <sub>Т</sub> >Г <sub>Т</sub> /(А <sub>Т</sub> >Г <sub>Т</sub> + Ц <sub>Т</sub> >Т <sub>Т</sub> ) ~ log <sub>2</sub> (время генерации поколения); N=211, lambda [ML] = 0.000	Интерсепта	-0.0353585	0.6318952
	log <sub>2</sub> (время генерации поколения)	0.0239113	0.0005575
1H. $A_T > \Gamma_T / (A_T > \Gamma_T + \Pi_T > T_T) \sim 0 + \log_2$ (время генерации поколения); N=211, lambda [ML] = 0.000	log <sub>2</sub> (время генерации поколения)	0.0206577	< 2.2e-16
1І. $A_T > \Gamma_T / (A_T > \Gamma_T + $ $I_T > T_T) \sim 0 + $ log <sub>2</sub> (время генерации	log <sub>2</sub> (время генерации поколения)	0.0225146	6.665e-12

log <sub>2</sub> (Общее	-0.0034229	0.5373
количество		
мутаций)		
	log <sub>2</sub> (Общее количество мутаций)	log <sub>2</sub> (Общее -0.0034229 количество мутаций)

Доля  $A_T > \Gamma_T$  в каждом видоспецифичном мутационном спектре зависит как от количества наблюдаемых A<sub>T</sub>>Г<sub>T</sub> замен внутри вида, так и количества нуклеотидов Ат в четырехкратно вырожденных синонимических позициях в гене ЦИТЬ Чтобы данного вида, используемых ДЛЯ нормализации. продемонстрировать, что наблюдаемые выше результаты были обусловлены не только изменением частоты предковых нуклеотидов (т.е. знаменатель в процессе нормализации) была пересчитана доля Ат>Гт для каждого вида, используя только замены из A<sub>T</sub>: A<sub>T</sub>> $\Gamma_T$  / (A<sub>T</sub>> $\Gamma_T$  + A<sub>T</sub>> $\amalg_T$  + A<sub>T</sub>> $\Upsilon_T$ ). В данном подходе содержание нуклеотидов в генах разных видов не учитывается. В итоге была получена положительная корреляция между A<sub>T</sub>>Г<sub>T</sub> и временем генерации поколения (ро-коэффициент Спирмена = 0,164, р-значение = 0,0007, N = 424), что позволяет предположить, что A<sub>T</sub>>Г<sub>T</sub> выше у долгоживущих млекопитающих независимо от содержания нуклеотидов.

Подобно анализам филогенетической инерции, проведенным ранее, были использованы модели, демонстрирующие ассоциацию  $A_T > \Gamma_T/(A_T > \Gamma_T + A_T > U_T + A_T > T_T)$  со временем генерации поколения. Модель 1К (Таблица 3.9) показывает значительное положительное влияние времени генерации поколения на  $A_T > \Gamma_T/(A_T > \Gamma_T + A_T > U_T + A_T > T_T)$  и отсутствие влияния «общего количества мутаций».

Таблица 3.9 — Модели PGLS, описывающие взаимосвязи между A<sub>T</sub>> $\Gamma_T/(A_T>\Gamma_T + A_T>U_T + A_T>T_T)$  и временем генерации поколения.

Модель	Переменная	Коэффициент	р-значение
1J. $A_T > \Gamma_T / (A_T > \Gamma_T +$	Интерсепта	0.6214988	1.759e-12

$A_T> U_T + A_T> T_T) \sim log_2$ (время генерации поколения); N=211, lambda [ML] = 0.000	log <sub>2</sub> (время генерации поколения)	0.0208230	0.007154
1К. $A_T > \Gamma_T / (A_T > \Gamma_T + A_T > \Pi_T + A_T > \Pi_T + A_T > T_T) \sim$ log <sub>2</sub> (время генерации поколения) + log <sub>2</sub> (Общее	Интерсепта log <sub>2</sub> (время генерации поколения)	0.6581961 0.0196741	1.655e-08 0.01515
количество мутации); N=211, lambda [ML] = 0.000	log <sub>2</sub> (Общее количество мутаций)	-0.0041067	0.62634

В целом, анализ внутривидовых четырехкратно вырожденных синонимических полиморфизмов у сотен видов млекопитающих продемонстрировал устойчивую связь мутационного спектра мтДНК (в основном за счет частоты  $A_T > \Gamma_T$ ) с видоспецифичным временем генерации поколения.

# 3.1.3 МтДНК млекопитающих с большим временем генерации поколения более бедна Ат и богата Гт вследствие интенсивного мутагенеза в направлении А<sub>Т</sub>>Г<sub>т</sub>.

Ожидается, что мутационный сдвиг (англ. mutational bias), если он сильнее отбора, в долгосрочной перспективе изменит содержание нуклеотидов во всем геноме. Во-первых, чтобы смоделировать возможное влияние мутационного нуклеотидный состав, было использовано слвига на компьютерное моделирование, которое вывело ожидаемый нейтральный нуклеотидный состав основе ВХОДНОГО 12-компонентного мутационного спектра. Данное на моделирование было проведено отдельно для млекопитающих с очень коротким (меньше нижнего дециля: 554 дня, N = 27) и очень длинным (больше верхнего дециля: 5221 дней, N = 25) временем генерации поколения (Рисунок 3.17, Таблица 3.10); средний мутационный спектр млекопитающих с очень большим временем поколения характеризовался генерации почти двукратным увеличением частоты A<sub>T</sub>>Г<sub>T</sub>.



Рисунок 3.17 — Ожидаемый нуклеотидный состав мтДНК млекопитающих с очень коротким и очень длинным временем генерации поколения в результате

#### моделирования

Ожидаемое содержание нуклеотидов при композиционном равновесии было получено с использованием моделирования на языке R. Основной целью моделирования было получение равновесного нуклеотидного состава (частоты четырех нуклеотидов), полученного в соответствии с заданным мутационным спектром. Моделирование началось co случайных частот четырех ИЗ нуклеотидов (левая часть графиков) и со временем частоты нуклеотидов менялись по мутационному спектру, достигая насыщения (правая часть графиков). Кривые, соответствующие частотам нуклеотидов, сходятся в ходе моделирования к одному значению. Эти равновесные значения использовались наших последующих анализах в качестве ожидаемого В содержания нуклеотидов. Для видов млекопитающих с очень коротким временем генерации поколения равновесные частоты были следующими:  $A_T \sim 0,161, T_T \sim 0,588, \Gamma_T \sim 0$ 0,233, Цт ~ 0,017. Для видов млекопитающих с очень большим временем генерации поколения равновесные частоты были следующими: A<sub>T</sub> ~ 0,096, T<sub>T</sub> ~ 0,635,  $\Gamma_T \sim 0,233$ ,  $\mu_T \sim 0,043$ . Виды с очень большой продолжительностью поколений характеризуются пониженной частотой A<sub>T</sub>.

Таблица 3.10 — Средние мутационные спектры млекопитающих с очень коротким и очень длинным временем генерации поколения, использованные для моделирования.

Средний мутационный спектр млекопитающих с очень коротким временем генерации поколения (меньше нижнего дециля: 554 дня, N = 27)		Средний мутационный спектр млекопитающих с очень длинным временем генерации поколения (больше верхнего дециля: 5221 дня, N =25)	
$T_T > A_T$	0.0060787	$T_T > A_T$	0.0026855
Тт>Цт	0.0174578	Тт>Цт	0.0399171
$T_T > \Gamma_T$	0.0049765	$T_T > \Gamma_T$	0.0030254
A <sub>T</sub> >T <sub>T</sub>	0.0191651	A <sub>T</sub> >T <sub>T</sub>	0.0123239
$A_T > \Gamma_T$	0.0028298	$A_T > \Gamma_T$	0.0031279
Α <sub>T</sub> >Γ <sub>T</sub>	0.1239227	Α <sub>T</sub> >Γ <sub>T</sub>	0.2173696
Щт>Тт	0.6664678	Цт>Тт	0.5693657
Цт>Ат	0.0392862	Цт>Ат	0.0317306
Щ <sub>T</sub> >Γ <sub>T</sub>	0.0189419	Щ <sub>T</sub> >Γ <sub>T</sub>	0.0132412
$\Gamma_T > T_T$	0.0155279	$\Gamma_{T} > T_{T}$	0.0169743
$\Gamma_T > A_T$	0.0832090	$\Gamma_T > A_T$	0.0871418
Гт>Цт	0.0021366	Гт>Цт	0.0030971

Примечание: Как и ожидалось, A<sub>T</sub>>Г<sub>T</sub> почти в два раза выше у видов с очень длинным временем генерации поколения.

Результаты этого моделирования показали, что ожидаемый нуклеотидный состав млекопитающих с большим временем генерации поколения характеризуется пониженной частотой A<sub>T</sub>. Для дальнейшей проверки

результатов моделирования была использована аналитическая модель, основанная на системе дифференциальных уравнений. Используя этот подход, были получены точные равновесные частоты, очень близкие к полученным путем моделирования. Для видов млекопитающих с очень коротким временем генерации поколения равновесные частоты были следующие: A<sub>T</sub> ~ 0,159, T<sub>T</sub> ~  $0,598, \Gamma_T \sim 0,228, \Pi_T \sim 0,016.$  Для видов млекопитающих с очень большой длиной поколения равновесные частоты были следующими:  $A_T \sim 0.093$ ,  $T_T \sim 0.651$ ,  $\Gamma_T \sim 0.651$ ,  $\Gamma$  $0,212, \ \text{Ц}_{\text{T}} \sim 0,044$ . Кроме того, используя этот подход было продемонстрировано, что это равновесие не зависит от начальных условий (стартовые частоты четырех нуклеотидов) и зависит исключительно от мутационного спектра. После сравнения ожидаемых нуклеотидных частот (полученных с помощью компьютерного моделирования и аналитических моделей) с наблюдаемыми нейтральными частотами в гене ЦИТЬ было показано, что ожидаемые и наблюдаемые значения достаточно близки друг к другу (Рисунок 3.18). Далее были использованы наблюдаемые нейтральные частоты нуклеотидов всех 12 генов, кодирующих белки (кроме НД6, так как он кодируется на другой цепи), и получены аналогичные результаты (Рисунок 3.19).



Рисунок 3.18 – Ожидаемое и наблюдаемое содержание нейтральных нуклеотидов в гене ЦИТЬ у видов млекопитающих с коротким и очень длинным временем генерации поколения

Примечание: Красные точки обозначают виды с очень длинным временем генерации поколения; синие точки обозначают виды с очень коротким временем генерации поколения.

Было показано, что наблюдаемый нуклеотидный состав довольно похож на ожидаемый, а это означает, что анализируемые виды достаточно близки к композиционному равновесию и продолжают сходиться к равновесию. Расположение ожидаемых и наблюдаемых частот нуклеотидов относительно диагональной линии статично, так как невозможно реконструировать прошлые и будущие эволюционные траектории без дополнительной информации.



Рисунок 3.19 – Ожидаемое и наблюдаемое содержание нейтральных нуклеотидов во всех 12 белок-кодирующих генах у видов млекопитающих с

коротким и очень длинным временем генерации поколения

При условии, что мутагенез является основной движущей силой этого процесса, ожидается, что все точки медленно приближаются к диагонали: частота T<sub>T</sub> увеличивается, тогда как A<sub>T</sub>, Ц<sub>T</sub> и Г<sub>T</sub> уменьшаются во времени. Однако, если наблюдаемая закономерность обусловлена отбором (что невозможно полностью исключить) некоторые из этих точек могут двигаться против равновесия: частота T<sub>T</sub> может уменьшаться, в то время как A<sub>T</sub>, Ц<sub>T</sub> и Г<sub>T</sub>

могут увеличиваться со временем. Чтобы восстановить потенциальное направление частоты нуклеотида, предковое состояние всех позвоночных было аппроксимировано довольно наивным подходом – просто усредняя нейтральное содержание нуклеотидов (в 12 генах, кодирующих белки, кроме НД6) во всем наборе данных митохондриальных геномов всех позвоночных, за исключением млекопитающих: Chondrichthyes + Cyclostomata (N=126), Actinopterygii (N=1770), Amphibia (N=205), Aves (N=432), Reptilia (N=269). В сравнении медианных частот нуклеотидов у млекопитающих (N=788) по сравнению со всеми остальными группами, было продемонстрировано, что частота T<sub>T</sub> у млекопитающих выше (T<sub>T</sub> у млекопитающих составляет 0,462 против 0,402 у всех остальных), тогда как Ат, Цт и Гт ниже (Ат: 0,187 против 0,190, Гт: 0,295 против 0,338, Цт: 0,045 против 0,061). Результаты этого пилотного теста, хотя и требуют дискуссии и должны быть подтверждены в будущих работах путем надлежащей реконструкции предкового состояния всех позвоночных, поддерживают мутационное объяснение Рисунков 3.18 и 3.19, где все точки, скорее всего, движутся к равновесию, а не против него. Дополнительный аргумент, что короткоживущие млекопитающие ближе к диагонали по сравнению с долгоживущие, делает рабочую гипотезу более надежной, поскольку ожидается, что короткоживущие млекопитающие будут иметь более высокую скорость молекулярной эволюции в единицу времени и, следовательно, ожидается, что они быстрее придут к равновесию. В целом предполагается, что наблюдаемые в настоящее время отклонения между наблюдаемыми и ожидаемыми частотами нуклеотидов со временем уменьшаются, а мтДНК млекопитающих, хотя уже довольно близка к равновесию, продолжает эволюционировать в этом направлении. Чтобы визуализировать эту рабочую гипотезу, на рисунки были добавлены пунктирные стрелки, соединяющие наблюдаемые частоты нуклеотидов с ожидаемыми в будущем.

В целом было обнаружено, что нейтральный (синонимичный четырехкратно вырожденный) нуклеотидный состав млекопитающих близок к

их композиционному равновесию, и поэтому ожидается увидеть влияние мутационного сдвига на содержание нуклеотидов в мтДНК млекопитающих.

Для этого было протестировано, приведет ли увеличение  $A_T > \Gamma_T$  у видов, с большим временем генерации поколения, к снижению частот  $A_T$  и увеличению частот  $\Gamma_T$  в соответствующих последовательностях. Поскольку время генерации поколения коррелирует с увеличением  $A_T > \Gamma_T$ , ожидается, что оно должно демонстрировать положительную корреляцию с  $\Gamma_T$  и отрицательную с  $A_T$ . Проверяя все четыре парные корреляции между видоспецифичным временем генерации поколения и содержанием нуклеотидов ( $A_T$ ,  $T_T$ ,  $\Gamma_T$ ,  $\Pi_T$ ), были показаны две самые сильные корреляции: отрицательная с  $A_T$  и положительная с  $\Gamma_T$  (Рисунок 3.20, для  $A_T$  ро-коэффициент Спирмена = -0.31, р-значение = 1.287e-15; для  $\Gamma_T$  ро-коэффициент Спирмена = 0.47, р-значение < 2.2e-16; N = 650).



Рисунок 3.20 – Частоты нуклеотидов в нейтральных сайтах всех 13 белоккодирующих генов в зависимости от времени генерации поколения Примечание: доля А<sub>т</sub> уменьшается, а доля Г<sub>т</sub> увеличивается (N = 650).

Включение всех четырех типов частот нуклеотидов в множественную линейную модель подтвердило важность только А<sub>т</sub> и Г<sub>т</sub>, эффект которых также был устойчив к филогенетической инерции (Таблица 3.11).

Таблица 3.11 – Пошаговая обратная множественная модель

Модель	Переменная	Коэффициент	р-значение
1L. Время генерации поколения ~ А <sub>T</sub> + Г <sub>T</sub> ; N=650	Интерсепта	11.06	< 2.2e-16
	A <sub>T</sub>	-0.11	0.023
	Γ	0.46	< 2.2e-16

Примечание: Рассмотрение филогенетической инерции посредством филогенетически независимых контрастов показало тот же тренд (доля  $A_T$ : ро-коэффициент Спирмена = - 0,09, р-значение = 0,025; доля  $\Gamma_T$ : ро-коэффициент Спирмена = 0,09, р-значение = 0,021).

Таким образом, был сделан вывод, что мтДНК млекопитающих с длинным временем генерации поколения по сравнению с видами с коротким временем генерации поколения более бедна A<sub>T</sub> и богата Г<sub>T</sub>, что согласуется с более интенсивным мутагенезом A<sub>T</sub>>Г<sub>T</sub> в первом случае.

Наконец, было протестироано, определяет ли избыток  $\Gamma_T$  и дефицит  $A_T$  у долгоживущих видов положительный  $\Gamma_T A_T$ -сдвиг. Асимметрия нуклеотидов  $\Gamma_T A_T$  аппроксимирует уровень асимметрии в распределении этих двух нуклеотидов. На основании нуклеотидного состава в нейтральных позициях 12 белок-кодирующих генов (кроме НДб), была оценена асимметрия  $\Gamma_T A_T$  для каждого вида млекопитающих и скоррелирована со временем генерации поколения. Как и ожидалось, была продемонстрирована положительная корреляция (PGLS, коэффициент = 0.13, р-значение = 2.9 \*10<sup>-4</sup>). Чтобы визуализировать контраст в асимметрии  $\Gamma_T A_T$  между самыми короткоживущими и самыми долгоживущими видами в наборе данных, на график были нанесены фракции  $A_T$  и  $\Gamma_T$  вдоль большой дуги репликации мтДНК медового опоссума

(время генерации поколения 341 день) и кита (время генерации поколения 18980 дней) (Рисунок 3.21). Очевидно, что в среднем мтДНК медового опоссума имеет избыток A<sub>T</sub> (красный цвет на рисунке), тогда как кит имеет избыток Г<sub>T</sub> (серый цвет на рисунке).



Рисунок 3.21 – структура мтДНК двух видов млекопитающих с предельными значениями времени генерации поколений: медового опоссума и кита

Примечание: Верхняя панель: каждый столбец представляет частоту нуклеотидов в окне из 20 нуклеотидов. У обоих млекопитающих снижается A<sub>T</sub> и увеличивается Г<sub>т</sub> вдоль большой дуги репликации мтДНК: от нижнего левого

(начало репликации легкой цепи) к верхнему правому (начало репликации тяжелой цепи). Однако, помимо градиента, мтДНК кита имеет целостный общегеномный дефицит  $A_T$  и избыток  $\Gamma_T$  – признак увеличенного времени генерации поколения. Нижняя панель: тепловые карты визуализируют асимметрию использования кодонов 12 генов, кодирующих белок (кроме НДб). Кит более контрастен, чем медовый опоссум, с точки зрения асимметрии, обусловленной заменами, ассоциированными с продолжительностью жизни,  $T_{\Lambda}>$ Ц $_{\Lambda}$  ( $A_T>\Gamma_T$ ). Тепловые карты обоих видов одинаково контрастируют с точки зрения асимметрии, обусловленной заменами  $\Gamma_{\Lambda}>$   $A_{\Lambda}$  (Ц $_T>T_T$ ), которые имеют высокую и одинаковую (не связанную с возрастом) скорость замен у обоих видов.

Дополнительно, было решено проанализировать, как мутагенез  $A_T > \Gamma_T$ влияет на асимметрию использования кодонов. Если  $A_T > \Gamma_T$  представляет собой сильную и однородную мутационную силу, ожидается, что большинство аминокислот будут демонстрировать дефицит XXA<sub>T</sub> и избыток кодонов XX $\Gamma_T$ .

Поскольку легкая цепь мтДНК эквивалентна мРНК 12 генов, кодирующих белок (кроме НДб), использование кодонов анализируется с точки зрения обозначения легкой цепи, и, таким образом, ожидается, что замены  $A_T > \Gamma_T$ уменьшат частоту кодонов XXT<sub>Л</sub> (поскольку  $T_Л$  комплементарен  $A_T$ ) и увеличивают частоту кодонов XXЦ<sub>Л</sub> (поскольку  $U_Л$  комплементарен  $\Gamma_T$ ). Аналогичным образом ожидается, что мутации  $U_T > T_T$  будут увеличивать частоты XXA<sub>Л</sub> и уменьшать частоты кодонов XXГ<sub>Л</sub>. Чтобы визуализировать этот эффект, были построены тепловые карты использования кодонов опоссума и кита. Было продемонстрировано, что асимметрия XXЦ<sub>Л</sub> (см. Материалы и методы) действительно сильнее у кита (0,70) по сравнению с медовым опоссумом (0,52). Интересно, что асимметрия XXA<sub>Л</sub> одинаково высока у обоих этих видов (кит: 0,93, опоссум: 0,95). Количественно оценивая асимметрию XXЦ<sub>Л</sub> и асимметрию XXA<sub>Л</sub> у всех видов млекопитающих с полными митохондриальными геномами, было отмечено, что (i) как XXЦ<sub>Л</sub>, так и XXA<sub>Л</sub> значительно превышают нулевое ожидание 0,5 (Рисунок 3.22); (ii) асимметрия XXA<sub>л</sub> намного сильнее по сравнению с асимметрией XXЦ<sub>л</sub> (Таблица 3.12), что подтверждает, что скорость замещения  $U_T > T_T$  намного выше по сравнению с  $A_T > \Gamma_T$  и (iii) асимметрия XXЦ<sub>л</sub> сильно коррелирует и положительно со временем генерации поколения млекопитающих, в то время как XXA<sub>л</sub> демонстрирует слабую отрицательную корреляцию, что не подтверждается статистикой, учитывающей филогению (Таблица 3.13), подтверждающей, что только замены  $A_T > \Gamma_T$  (те, которые влияют на асимметрию XXЦ<sub>л</sub>) связаны со временем генерации поколения.



Рисунок 3.22 – Распределение асимметрии использования кодонов XXAл

#### и ХХЦл

Примечание: Парный U-тест Манна-Уитни (N = 648) показал более выраженную асимметрию XXA<sub>л</sub> по сравнению с XXЦ<sub>л</sub> (p-значение < 2,2e-16), хотя оба они значительно отклоняются от ожидаемого 0,5 (p-значения < 2,2e-16, критерий Уилкоксона с mu = 0,5). Таблица 3.12 – Ранговые парные корреляции асимметрии кодонов с временем генерации поколения.

Ν	Метод	Асимметрия	Ро-коэффициент Спирмена	р-значение
648	парные корреляции	ХХЦл	0.3437	2.2e-16
	Спирмена	$XXA_{\Lambda}$	-0.2233	9.164e-09

Таблица 3.13 – Стандартные линейные модели и модели филогенетической инерции (PGLS) ассоциации асимметрии кодонов со временем генерации поколения.

Модель	Переменная	Коэффициент	р-значение
1M. XXA <sub>л</sub> ~ log <sub>2</sub> (время генерации поколения);	Интерсепта	0.285261	< 2.2e-16
N=648	log2(время генерации поколения)	0.029795	< 2.2e-16
1N. XXЦл ~ log <sub>2</sub> (время генерации поколения):	Интерсепта	0.993573	< 2.2e-16
N=648	log <sub>2</sub> (время генерации поколения)	-0.007203	5.15e-09
10. scale(XXA <sub>Л</sub> ) ~ log <sub>2</sub> (время генерации	Интерсепта	1.001718	0.27865
поколения); N=648, lambda [ ML] = 0.998	log <sub>2</sub> (время генерации поколения)	-0.084755	0.09565
1P. scale(XXЦ <sub>л</sub> ) ~l оg <sub>2</sub> (время генерации	Интерсепта	-1.416211	0.03242
поколения); N=648, lambda [ ML] = 0.998	log <sub>2</sub> (время генерации поколения)	0.076297	0.02704

Примечание: Только XXЦл демонстрирует надежную связь со временем генерации поколения.

В целом было продемонстрировано, что мутагенез A<sub>T</sub>>Г<sub>T</sub>, который более выражен у долгоживущих видов, сильно формирует его полногеномные последовательности мтДНК: содержание нуклеотидов (низкая частота A<sub>T</sub> и высокая частота Г<sub>T</sub>), нуклеотидный сдвиг (сильный положительный Г<sub>T</sub>A<sub>T</sub>-сдвиг) и использование кодонов (положительная ХХЦ<sub>л</sub> асимметрия).

## **3.1.4** ГтАт -сдвиг является функцией как ВПДОС, так и времени генерации поколения.

Ранее было показано, что частота замен A<sub>T</sub>> Г<sub>T</sub> зависит от того, сколько времени родительская тяжелая цепь находилась в одноцепочечном (ВПДОС) состоянии во время асинхронной репликации мтДНК. Гены, расположенные близко к началу репликации легкой цепи (OL), такие как ЦОГ1, проводят минимальное время в одноцепочечном состоянии и демонстрируют низкую частоту  $A_T > \Gamma_T$ , тогда как гены, расположенные далеко от  $O_L$ , проводят больше времени в одноцепочечном состоянии. и демонстрируют соответственно более высокие частоты A<sub>T</sub>>Г<sub>T</sub> [52,75]. Таким образом, ожидается, что эффективно нуклеотидный нейтральный состав мтДНК является функцией как геноспецифичного ВПДОС, так и видоспецифичного времени генерации поколения. Чтобы проверить это, для каждого гена каждого вида была рассчитана асимметрия ГтАт-сдвига после разделения всех вилов млекопитающих на виды с коротким и длинным временем генерации поколения в соответствии с медианой (медиана = 2190 дней, короткоживущие N = 325, долгоживущие N = 319). Затем был построен график распределения  $\Gamma_{T}A_{T}$ -сдвига у млекопитающих с коротким и длинным временем генерации поколения для каждого гена, ранжируя их по большой дуге репликации мтДНК от ЦОГ1 (ранг равен 1) до ЦИТЬ (ранг равен 10), что соответствует увеличению ВПДОС. Как и ожидалось ГтАт-сдвиг увеличивается как с геноспецифичным ВПДОС, так и с видоспецифичным временем генерации поколения (Рисунок 3.23). Запустив несколько линейных моделей, в которых Г<sub>Т</sub>А<sub>Т</sub>-сдвиг является функцией как ВПДОС, так и времени генерации поколения, подтвердилось, что оба фактора влияют на сдвиг, причем в очень одинаковой степени (Таблица 3.14). Модели с

масштабированными переменными и с фиктивными переменными позволяют сравнивать размеры эффекта ВПДОС и времени генерации поколения: насколько сильно они влияют на Г<sub>т</sub>А<sub>т</sub>-сдвиг. Хотя коэффициент времени генерации поколения меньше, чем коэффициент ВПДОС, они сопоставимы, что позволяет предположить, что время генерации поколения так же важно, как и положение гена в большой дуге репликации мтДНК.



Рисунок 3.23 – Изменение содержания нуклеотидов мтДНК (Г<sub>т</sub>А<sub>т</sub>-сдвиг) коротко- и долгоживущих млекопитающих.

Примечание: Все гены (кроме НД6), расположенные в большой дуге репликации мтДНК, ранжированы по ВПДОС: от ЦОГ1 до ЦИТЬ. Г<sub>т</sub>А<sub>т</sub>-сдвиг увеличивается как с геноспецифичным ВПДОС, так и с видоспецифичным временем генерации поколения.

Таблица 3.14 – Г<sub>т</sub>А<sub>т</sub>-сдвиг является положительной функцией как ВПДОС, так и времени генерации поколения.

Модель	Переменная	Коэффициент	р-значение
1Q. Г <sub>Т</sub> А <sub>Т</sub> -сдвиг ~	Интерсепта	-0.8498809	<2e-16

ВПДОС + log <sub>2</sub> (время генерации поколения); N=648	впдос	0.0724620	<2e-16
	log2(время генерации поколения)	0.0387322	<2e-16
1R. Г <sub>Т</sub> А <sub>Т</sub> -сдвиг ~ scale(ВПЛОС) +	Интерсепта	0.164223	<2e-16
scale(время генерации поколения); N=648	scale(ВПДОС )	0.069142	<2e-16
	scale(время генерации поколения)	0.111258	<2e-16
1S. Г <sub>Т</sub> А <sub>Т</sub> -сдвиг ~ ВПЛОС * log <sub>2</sub> (время	Интерсепта	-0.8619907	<2e-16
генерации поколения); N=648	впдос	0.0735574	<2e-16
	log2(время генерации поколения)	0.0409340	2.66e-05
	log2(время генерации поколения):В ПДОС	-0.0001992	0.82
1Т. Г <sub>т</sub> А <sub>т</sub> -сдвиг ~ Фиктивная(длинное ВПДОС) + Фиктивная(Долгоживущ	Интерсепта	0.0002864	0.955
ий вид)	Фиктивная(д линное ВПДОС)	0.1385817	<2e-16
	Фиктивная(Д олгоживущи й вид)	0.1903651	<2e-16

Примечание: Фиктивная (Долгоживущий вид), равная 1 для млекопитающих со

временем генерации поколения выше медианного и 0 для остальных;

Фиктивная(длинное ВПДОС), равная 1 для генов ЦОГ1, ЦОГ2, АТФ6, АТФ8, ЦОГ3 и 0 для НД3, НД4Л, НД5, ЦИТЬ были введены в модели, чтобы показать надежность ассоциаций.

В линейных моделях не наблюдалось существенного взаимодействия между ВПДОС и временем генерации поколения, что позволяет предположить, что либо эти факторы влияют на состав нуклеотидов независимо друг от друга, либо сигнал взаимодействия слишком слаб, чтобы быть значимым с нашим размером выборки. Чтобы это проверить, на график было нанесено содержание нуклеотидов в нейтральных сайтах каждого из 13 генов, сопоставленное с расположением генов, ранжированных от самого короткого (ЦОГ1) до самого длительного времени (ЦИТb), проведенного в одноцепочечном состоянии (Рисунок 3.23). Было продемонстрировано, что частота Г<sub>т</sub> увеличивалась с ВПДОС (ро-коэффициент Спирмена  $\geq 0.62$ , р-значения  $\leq 0.028$ ). Учитывая тот факт, что этот анализ был ограничен небольшим количеством генов (13), был проведен более чувствительный анализ, разделящий геном каждого вида на 50 окон, содержащих 25 нейтральных нуклеотидов и с использованием только генов, расположенных на одной цепи большой дуги репликации мтДНК (все гены, кроме НД1, НД2 и НД6). Для каждого окна были оценены частоты четырех нуклеотидов и для каждого вида рассчитан ро-коэффициент Спирмена и рзначение. Было продемонстрировано, что частота Г<sub>т</sub> действительно имеет самое сильное смещение в сторону положительных значений ро-коэффициента и, кроме того, второе по силе смещение принадлежит Ат, которое уменьшалось по геному. Чтобы дополнительно проверить, являются ли некоторые из этих градиентов более выраженными у долгоживущих млекопитающих по сравнению с короткоживущими, была оценена доля долгоживущих млекопитающих (со временем генерации поколения больше медианы, 1101,1 дня) среди значимых (со р-значениями <= 0,01) ро-коэффициентов. Только для нуклеотидов Г<sub>т</sub> был обнаружен значительный избыток долгоживущих млекопитающих среди значимых ро-коэффициентов (отношение шансов = 1,61, p-значение = 0,003521,

тест Фишера). Это говорит о том, что скорость увеличения частоты  $\Gamma_{T}$  выше у долгоживущих млекопитающих. Более быстрые изменения (более сильные градиенты) состава  $A_{T}$  и  $\Gamma_{T}$  вдоль большой дуги репликации мтДНК у долгоживущих млекопитающих можно интерпретировать как взаимодействие между ВПДОС и временем генерации поколения, как будто скорость замещения  $A_{T}>\Gamma_{T}$  увеличивается быстрее в зависимости от ВПДОС в случае большего времени генерации поколения.



Рисунок 3.23 – Долгосрочный эффект мутационного сдвига: содержание нейтральных нуклеотидов у млекопитающих

Примечание: Все гены, расположенные в большой дуге репликации мтДНК, ранжированы по ВПДОС. Гены НД2 и НД1 не реплицируются внутри большой дуги. Нижняя панель: градиент изменения нуклеотидов со ВПДОС (увеличение Г<sub>т</sub>) более выражен у долгоживущих млекопитающих. Гистограммы значений

Ро-коэффициента Спирмена показывают, что A<sub>T</sub> уменьшается со ВПДОС, в то время как Г<sub>T</sub> увеличивается. Мозаичный график отражает избыток долгоживущих млекопитающих среди всех видов со значительным Ро-коэффициентом для Г<sub>T</sub>, что отражает более высокую скорость увеличения частоты Г<sub>T</sub> в мтДНК долгоживущих млекопитающих.

В целом полученные результаты показывают, что содержание нуклеотидов, сформированное мутационным сдвигом A<sub>T</sub>>Г<sub>T</sub>, положительно и сильно зависит как от ВПДОС, так и от времени генерации поколения (Рисунок 3.24).



Рисунок 3.24 – Графический вывод

Примечание: Скорость замены  $A_T > \Gamma_T$  (отмечена красным градиентом) увеличивается как с геноспецифичным ВПДОС, так и с видоспецифичным временем генерации поколения. Размер эффекта времени генерации поколения сопоставим с размером эффекта ВПДОС. Скорость замещения  $U_T > T_T$  (отмечена серым градиентом) чувствительна только к ВПДОС.

3.2 Мутационные спектры экзотермическими животных: замена Ат>Гт ассоциирована с температурой тела.

### 3.2.1 Видоспецифичный мутационный спектр мтДНК рыб связан с температурой окружающей среды посредством повышенной асимметрии А<sub>T</sub>>Г<sub>т</sub>.

Чтобы изучить потенциальное влияние температуры на мутагенез мтДНК, были проанализированы мутационные спектры рыб (Actinopterygii N=180; Chondrichthyes N=8), которые являются экзотермическими животными, обитающими в широком диапазоне температур окружающей среды. Используя полиморфизмы внутривидовые мтДНК. были синонимические реконструированы 12-компонентный мутационные спектры для каждого из 188 видов рыб для каждого митохондриального гена. В среднем мутационном спектре рыб наблюдался явный избыток транзиций, причем наиболее частыми были  $U_T > T_T$ , а вторыми по частоте были  $A_T > \Gamma_T$  (Рисунок 3.25). Эта картина очень напоминала спектры, наблюдаемые у млекопитающих и в злокачественных опухолях человека[4,20]. Далее были собраны данные о среднегодовой температуре воды для каждого вида рыб (см. Материалы и методы) и проанализирована ее связь с 12 типами замен в мутационном спектре мтДНК.



Рисунок 3.25 – Средний мутационный спектр проанализированных видов рыб

Было продемонстрировано, что два типа замен,  $A_T > \Gamma_T$  и  $U_T > T_T$ , имели значительную положительную корреляцию с температурой (Таблица 3.15, Рисунок 3.26). Чтобы определить, увеличивается ли асимметрия этих замен с температурой, были проанализированы соответствующие соотношения  $A_T > \Gamma_T / T_T > U_T$  и  $U_T > T_T / \Gamma_T > A_T$ .

Таблица 3.15 – Ранговые корреляции Спирмена двух наиболее распространенных транзиций и их асимметрия с температурой.

Ν	Переменная	Ро-коэффициент Спирмена	р-значение
188	$A_T > \Gamma_T$	0.2985488	3.162e-05
	$A_T > \Gamma_T / T_T > \amalg_T$	0.3889124	2.654e-07
	Щ <sub>Т</sub> >Т <sub>Т</sub>	0.1939811	0.007644
	$\coprod_T > T_T / \Gamma_T > A_T$	-0.08912368	0.2564

Примечание: Частоты всех остальных типов замен не показали значимой корреляции с температурой (р-значения > 0,1).



Рисунок 3.26 – Видоспецифичный мутационный спектр связан с температурой.

Примечание: левая панель: положительная корреляция A<sub>T</sub>>Г<sub>T</sub> с температурой воды; правая панель: положительная корреляция log2(A<sub>T</sub>>Г<sub>T</sub>/Ц<sub>T</sub>>T<sub>T</sub>) с температуройводы.

Логика анализа асимметрии состоит в том, что соотношение  $A_T > \Gamma_T/T_T > U_T$ представляет собой асимметрию замен  $A_T > \Gamma_T$ .  $A > \Gamma$  и T > U являются комплементарными эквивалентными заменами: замена, обозначенная как  $A > \Gamma$  в тяжелой цепи ( $A_T > \Gamma_T$ ), могла возникнуть как T > U в легкой цепи ( $T_T > U_T$ ), и наоборот. Если вероятность мутации  $A > \Gamma$  одинакова как в тяжелой ( $A_T > \Gamma_T$ ), так и в легкой ( $A_T > \Gamma_T = T_T > U_T$ ) цепях, ожидается увидеть симметрию скорости  $A_T > \Gamma_T$  относительно скорости  $T_T > U_T$  (оба показателя нормированы на частоту предковых нуклеотидов). Однако, было продемонстрировано, что обе асимметрии достаточно сильны:  $A_T > \Gamma_T$  была выше, чем  $T_T > U_T$  ( $A_T > \Gamma_T/T_T > U_T >$ 1), а  $U_T > T_T$  была выше, чем  $\Gamma_T > A_T$  ( $U_T > T_T/\Gamma_T > A_T >$  1), что указывает на наличие мутагенов, которые преимущественно индуцируют  $A > \Gamma$  и U > T в тяжелой цепи.

Затем было отмечено, что асимметрия  $A_T > \Gamma_T$ , но не асимметрия  $U_T > T_T$ , демонстрирует сильную положительную корреляцию с температурой (Рисунок 3.26, Таблица 3.16).

Таблица 3.16 – Линейная моде	ель, описывающая асим	лметрию А <sub>Т</sub> >Г <sub>Т</sub> как функци	łЮ
температуры.			

Модель	Переменная	Коэффициент	р-значение
2А. $A_T > \Gamma_T \sim$ scale(Температура);	Интерсепта	0.018959	< 2e-16 ***
N=188	scale(Температура)	0.007750	9.66e-05 ***
2В. $A_T > \Gamma_T / T_T > \Pi_T \sim$ scale(Температура):	Интерсепта	3.6584	<2e-16 ***
N=188	scale(Температура)	0.9647	0.0145 *
2С. Цт>Тт~	Интерсепта	0.068390	< 2e-16 ***

scale(Температура); N=188	scale(Температура)	0.015680	0.00638 **
2D. $U_T > T_T / \Gamma_T > A_T \sim$ scale(Температура):	Интерсепта	10.1405	<2e-16 ***
N=188	scale(Температура)	-1.0889	0.181

Повышенная асимметрия  $A_T > \Gamma_T$  у видов, обитающих в теплой и холодной воде, показывает, что этот мутаген чувствителен к температуре, т.е. индуцирует пропорционально больше  $A_T > \Gamma_T$  в тяжелой цепи по сравнению с легкой цепью с повышением температуры. Эта тенденция устойчива к филогенетической инерции (Таблица 3.17) и остается качественно схожей, когда анализируются данные по конкретным семействам и когда весь набор данных разделяется на несколько подгрупп Семейств в зависимости от температуры (Рисунок 3.27, рокоэффициент Спирмана = 0,5, р-значение = 0,02631, N = 19).

Таблица 3.17 – Анализ филогенетическо	й инерции (PGLS) асимметрии А <sub>Т</sub> >Г <sub>Т</sub> в
зависимости от температуры.	

Модель	Переменная	Коэффициент	р-значение
2E. log <sub>2</sub> (A <sub>T</sub> >Г <sub>T</sub> )~ scale(Температура);	Интерсепта	-8.91589	< 2.2e-16 ***
N=52, lambda [ ML] : 0.000	log <sub>2</sub> (Температура)	0.70605	0.0002791 ***
2F. log <sub>2</sub> (A <sub>T</sub> >Г <sub>T</sub> /T <sub>T</sub> >Ц <sub>T</sub> ) ~ scale(Температура);	Интерсепта	-0.59919	0.341474
N=52, lambda [ ML] : 0.000	log <sub>2</sub> (Температура)	0.44507	0.009124 **
2G. log <sub>2</sub> (Ц <sub>T</sub> >T <sub>T</sub> )~ scale(Температура);	Интерсепта	-5.39515	6.384e-07 ***
N= 52, lambda [ ML] : 0.667	log <sub>2</sub> (Температура)	0.39050	0.02828 *
2H. $log_2(U_T > T_T/\Gamma_T > A_T)$ ~ scale(Температура):	Интерсепта	3.25550	1.982e-07 ***
N=52, lambda [ ML] : 0.000	log <sub>2</sub> (Температура)	-0.12593	0.3795


Рисунок 3.27 – Спектр мтДНК рыб связан с температурой: анализ Семейств Примечание: Медиана асимметрии A<sub>T</sub>>Г<sub>T</sub> (A<sub>T</sub>>Г<sub>T</sub>/T<sub>T</sub>>Ц<sub>T</sub>) положительно коррелирует со средней температурой в Семействах.

В предыдущей главе было продемонстрировано, что A<sub>T</sub>>Г<sub>T</sub> у млекопитающих положительно коррелирует с временем генерации поколения. Чтобы проверить, существует ли подобная связь у рыб, был использован аналогичный показатель - время полового созревания, оцениваемое как возраст, в котором 50% популяции впервые нерестятся.

Был проведены ряд ранговых корреляций и множественный линейный регрессионный анализ, чтобы изучить влияние температуры и времени полового созревания на мутационный спектр мтДНК (N = 96), в частности, частоту  $A_T > \Gamma_T$ , асимметрию  $A_T > \Gamma_T$ , частоту  $U_T > T_T$ ,  $U_T > T_T$  асимметрию. Прежде всего, была

продемонстрирована отрицательная корреляция между температурой и временем созревания (ро-коэффициент Спирмана = -0.31, р-значение = 0.001076, N = 96), о которой также сообщалось ранее[92], что позволяет предположить, что тепловодные виды имеют тенденцию достигать половой зрелости быстрее.

Ни одна из четырех ключевых метрик мутационного спектра не продемонстрировала корреляции со временем созревания (Таблицы 3.18, 3.19. 3.20).

Таблица 3.18 – Ранговые корреляции Спирмена между временем созревания и ключевыми компонентами мутационного спектра.

Ν	Переменная	Ро-коэффициент Спирмена	р-значение
96	$A_T > \Gamma_T$	-0.1569224	0.1268
	Щт>Тт	0.005575674	0.957
	$A_T > \Gamma_T / T_T > \coprod_T$	-0.04573348	0.6229
	$\coprod_{T} > T_T / \Gamma_T > A_T$	-0.0616844	0.5051

Таблица 3.19 – Множественные линейные модели с асимметрией  $A_T > \Gamma_T$  и  $U_T > T_T$  и  $A_T > \Gamma_T$  ( $A_T > \Gamma_T / T_T > U_T$ ) и асимметрией  $U_T > T_T$  ( $U_T > T_T / \Gamma_T > A_T$ ) в зависимости от температуры и времени созревания.

Модель	Переменная	Коэффициент	р-значение
2I. A <sub>T</sub> >Г <sub>T</sub> ~ scale(Температура) *	Интерсепта	0.020323	2.45e-09 ***
scale(Время полового созревания); N=96	scale(Температура)	0.009526	0.0024 **
	scale(Время полового созревания)	-0.001981	0.5623
	scale(Время полового	-0.001094	0.7280

	созревания):scale(В ремя полового созревания)		
2J.A <sub>T</sub> >Г <sub>T</sub> ~ scale(Температура) +	Интерсепта	0.020647	2.68e-10 ***
scale(Время полового созревания); N=96	scale(Температура)	0.009511	0.00232 **
	scale(Время полового созревания)	-0.001431	0.63500
2K. A <sub>T</sub> >Г <sub>T</sub> /T <sub>T</sub> >Ц <sub>T</sub> ~ scale(Температура) *	Интерсепта	3.6506	5.35e-12 ***
scale(Время полового созревания); N=69	scale(Температура)	1.0256	0.0168 *
	scale(Время полового созревания)	0.5740	0.2578
	scale(Время полового созревания):scale(В ремя полового созревания)	0.7080	0.1184
2L. $A_T > \Gamma_T / T_T > \Pi_T \sim$ scale(Температура) +	Интерсепта	3.4250	8.86e-12 ***
scale(Время полового созревания); N=69	scale(Температура)	0.9908	0.0219 *
	scale(Время полового созревания)	0.1439	0.7377
2M. Ц <sub>T</sub> >T <sub>T</sub> ~ scale(Температура) *	Интерсепта	0.074275	1.47e-11 ***
scale(Время полового созревания); N=96	scale(Температура)	0.013601	0.159
	scale(Время полового созревания)	0.004570	0.670
	scale(Время полового	0.006202	0.530

	созревания):scale(В ремя полового созревания)		
2N. Ц <sub>T</sub> >T <sub>T</sub> ~ scale(Температура) +	Интерсепта	0.072437	5.25e-12 ***
scale(Время полового созревания); N=96	scale(Температура)	0.013687	0.155
	scale(Время полового созревания)	0.001455	0.878
2О. $\[ \]_T > T_T / \Gamma_T > A_T \sim \]$ scale(Температура) *	Интерсепта	9.5799	1.36e-10 ***
scale(Время полового созревания); N=69	scale(Температура)	0.2455	0.840
	scale(Время полового созревания)	1.0112	0.490
	scale(Время полового созревания):scale(В ремя полового созревания)	0.5267	0.686
2Р. $\coprod_T > T_T / \Gamma_T > A_T \sim$ scale(Температура) +	Интерсепта	9.4122	2.87e-11 ***
scale(Время полового созревания); N=69	scale(Температура)	0.2197	0.855
	scale(Время полового созревания)	0.6912	0.572

Таблица 3.20 – Множественные линейные модели с температурой и временем созревания в зависимости от A<sub>T</sub>>Г<sub>T</sub> и Ц<sub>T</sub>>T<sub>T</sub>.

Модель	Переменная	Коэффициент	р-значение
2Q. Температура ~ scale( $A_T > \Gamma_T$ ) + scale( $\coprod_T > T_T$ ); N=188	Интерсепта	17.12660	<2e-16 ***
	scale( $A_T > \Gamma_T$ )	2.05598	0.00544 **

	scale(U <sub>T</sub> >T <sub>T</sub> )	0.05291	0.94237
2R. Время полового созревания ~	Интерсепта	3.9298	<2e-16 ***
scale( $A_T > \Gamma_T$ ) + scale( $\coprod_T > T_T$ ); N=188	scale( $A_T > \Gamma_T$ )	-0.6647	0.101
	scale(U <sub>T</sub> >T <sub>T</sub> )	0.3372	0.382

В целом мы пришли к выводу, что мутационный спектр мтДНК рыб, в частности частота  $A_T > \Gamma_T$ , асимметрия  $A_T > \Gamma_T$  и менее выраженная частота  $U_T > T_T$ , связаны только с температурой.

## 3.2.2 Мутационный спектр мтДНК рыб влияет на их нейтральный нуклеотидный состав: тепловодные и долгоживущие виды, как правило, бедны А<sub>т</sub>Ц<sub>т</sub> и богаты Г<sub>т</sub>Т<sub>т</sub>.

На нуклеотидный состав может влиять мутационный спектр, как было показано в предыдущей главе: если мутационный сдвиг сильнее отбора и анализируемые виды близки к нуклеотидному композиционному равновесию, то мутационный спектр у тепловодных рыб в конечном итоге приводит к снижению Ат и увеличению Гт в течение эволюции. Чтобы определить, насколько рыбы близки к своему нуклеотидному композиционному равновесию, было проведено сравнение ожидаемого и наблюдаемого нуклеотидного составы в наиболее нейтральных, четырехкратно вырожденных положениях. Сначала были рассчитаны типичные мутационные спектры для холодноводных и тепловодных рыб, используя самые холодные (N=19, температура  $\leq 8,18$  °C) и самые теплые  $(N = 23, \text{температура} \ge 27 \,^{\circ}\text{C})$  децили анализируемых видов. Как и ожидалось, у тепловодных рыб наблюдалась повышенная частота A<sub>T</sub>>Г<sub>T</sub> и асимметрия A<sub>T</sub>>Г<sub>T</sub> (Таблица 3.21).

Таблица	3.21	—	Средние	митохондриальные	мутационные	спектры
холодново	одных	(N=1	9) и теплов	одных (N=23) рыб.		

Средний мутационный спектр холодноводных рыб (N = 19)		Средний мутационный спектр тепловодных рыб (N =23)	
$T_T > A_T$	0.00917	$T_T > A_T$	0.00486
Тт>Цт	0.08326	Тт>Цт	0.09990
$T_T > \Gamma_T$	0.00651	$T_T > \Gamma_T$	0.01177
A <sub>T</sub> >T <sub>T</sub>	0.01286	A <sub>T</sub> >T <sub>T</sub>	0.02306
$A_T > \Gamma_T$	0.00360	$A_T > \Gamma_T$	0.00987
$A_T > \Gamma_T$	0.10091	$A_T > \Gamma_T$	0.17954
Щ <sub>Т</sub> >Т <sub>Т</sub>	0.63431	$\coprod_T > T_T$	0.45972
Щт>Ат	0.00959	Щт>Ат	0.0147
Щ <sub>T</sub> >Γ <sub>T</sub>	0.0238	$\coprod_T > \Gamma_T$	0.0227
$\Gamma_{\rm T} > T_{\rm T}$	0.03175	$\Gamma_{\rm T} > T_{\rm T}$	0.02409
$\Gamma_{T} > A_{T}$	0.07851	$\Gamma_{\mathrm{T}} > \mathbf{A}_{\mathrm{T}}$	0.13772
Гт>Цт	0.00662	Гт>Цт	0.01207

Во-вторых, используя компьютерное моделирование и аналитическое моделирование, был получен ожидаемый состав нейтральных нуклеотидов для холодноводных и тепловодных рыб на основе этих мутационных спектров (Рисунок 3.28). Это позволило ответить на вопрос, каким был бы нуклеотидный состав в мтДНК в долгосрочной перспективе, если бы единственным фактором, влияющим на молекулярную эволюцию, был мутагенез (т.е. специфический мутационный спектр).

Для этого были использованы два подхода: компьютерное моделирование (в среде R) и дифференциальные уравнения. Моделирование начиналось со случайных частот четырех нуклеотидов (левая часть графиков) и со временем (с моделируемыми поколениями) частоты нуклеотидов менялись в соответствии с мутационным спектром, достигая насыщения (правая часть графиков).



Рисунок 3.28 – Ожидаемое содержание нуклеотидов при композиционном равновесии в мтДНК у рыб

Оба подхода продемонстрировали идентичный ожидаемый нуклеотидный состав (сплошные кривые соответствуют компьютерному моделированию, которые сходятся в ходе моделирования к одному значению, пунктирные линии соответствуют аналитическому моделированию, полученному как решение системы дифференциальных уравнений). Равновесные значения (4 частоты нуклеотидов для каждой температуры) использовались в последующих анализах

для получения ожидаемого содержания нуклеотидов в четырехкратно вырожденных позициях.

В-третьих, был получен наблюдаемый нуклеотидный состав, после расчёта нейтрального содержания нуклеотидов в 12 (кроме НДб) белоккодирующих генах мтДНК из самых холодных и самых теплых децилей, коорый был сравнен с ожидаемым составом (Рисунок 3.29). Было продемонстрировано, что наблюдаемый нуклеотидный состав аналогичен ожидаемому, что указывает на то, что анализируемые виды близки к композиционному равновесию.



Рисунок 3.29 – Наблюдаемое и ожидаемое композиционное равновесие нуклеотидного состава тепловодных и холодноводных рыб Примечание: Красные точки обозначают виды, обитающие в теплых водах; синие точки обозначают виды, обитающие в холодных водах.

Учитывая, что нейтральный нуклеотидный состав мтДНК рыб близок к нейтральному равновесию и отсутствует сильный отбор по четырехкратно

вырожденным синонимическим позициям в мтДНК, ожидается увидеть корреляцию между содержанием нуклеотидов и температурой окружающей среды. Анализируя нейтральное содержание нуклеотидов в 12 (кроме НДб) белок-кодирующих генах рыб, были продемонстрированы четыре тенденции: снижение  $A_T$  и  $U_T$  и увеличение  $\Gamma_T$  и  $T_T$  у тепловодных видов по сравнению с холодноводными видами. (Рисунок 3.30,  $A_T$ : ро-коэффициент Спирмена = -0,161, р-значение = 0,003163;  $T_T$ : ро-коэффициент Спирмена = 0,132, р-значение = 0,01605;  $\Gamma_T$ : ро-коэффициент Спирмена = 0,124, р-значение = 0,02414;  $U_T$ : ро-коэффициент Спирмена = -0,255, р-значение = 2,473е-06, N = 333).



Рисунок 3.30 – Содержание нейтральных нуклеотидов в геноме мтДНК связано с температурой

Важно отметить, что все эти изменения отражают последствия увеличения соотношений A<sub>T</sub>>Г<sub>T</sub> и Ц<sub>T</sub>>T<sub>T</sub> у тепловодных рыб. Чтобы подтвердить эти четыре

ассоциации, оны были объединены в Г<sub>т</sub>А<sub>т</sub>-сдвиг и Т<sub>т</sub>Ц<sub>т</sub>-сдвиг, после чего была продемонстрироана их связь с температурой и временем полового созревания (Таблица 3.22, Таблица 3.23).

Таблица 3.22 – Ранговые корреляции Спирмена между Г<sub>т</sub>А<sub>т</sub>-сдвигом, Т<sub>т</sub>Ц<sub>т</sub>сдвигом и температурой.

Ν	Переменная	Ро-коэффициент Спирмена	р-значение
333	ГтАт-сдвиг	0.1816663	0.0008673
	Т <sub>Т</sub> Ц <sub>Т</sub> -сдвиг	0.2429729	7.307e-06

Таблица 3.23 – Ранговые корреляции Спирмена между фракциями из 4 нуклеотидов, Г<sub>Т</sub>А<sub>Т</sub>-сдвигом и Т<sub>Т</sub>Ц<sub>Т</sub>-сдвигом и временем полового созревания.

Ν	Переменная	Ро-коэффициент Спирмена	р-значение
		Chinpmenu	
217	A <sub>T</sub>	-0.06058898	0.3744
	Тт	0.2597488	0.0001084
	Γ <sub>T</sub>	-0.06446041	0.3446
	Ц <sub>т</sub>	-0.1890223	0.005212
	ГтАт-сдвиг	0.1816663	0.0008673
	Т <sub>Т</sub> Ц <sub>Т</sub> -сдвиг	0.2429729	7.307e-06

Далее была введена специальная метрика, объединяющая все предыдущие наблюдения - S<sub>TT</sub>-S<sub>AU</sub> (см. Материалы и методы), отражающую асимметрию содержания нуклеотидов. Регрессия S<sub>TT</sub>-S<sub>AU</sub> в зависимости от температуры привела к сильной, ожидаемой положительной взаимосвязи (Рисунок 3.31, Таблица 3.24), которая устойчива к филогенетической инерции (Таблица 3.25) и

остается значимой в двух отдельных группах рыб: рано созревающие и поздно созревающие (Рисунок 3.32).



Рисунок 3.31 – S<sub>TГ</sub>-S<sub>АЦ</sub> чувствителен как к температуре, так и ко времени полового созревания, причем температура является наиболее сильным фактором

Примечание: Левая панель: положительная связь между S<sub>TF</sub>-S<sub>AЦ</sub> и температурой. Правая панель: аналогичные положительные наклоны между S<sub>TF</sub>-S<sub>AЦ</sub> и температурой для рано и поздно созревающих рыб, где поздно созревающие рыбы обозначены голубыми точками, а рано созревающие обозначены зелеными точками.

Все виды были разделены на 2 группы: рано-и поздносозревающие, в зависимости от медианы времени полового созревания (медиана = 3 года). Вовторых, каждая группа была дополнительно разделена на две подгруппы: холодноводные и тепловодные виды, исходя из медианы температуры (19,7°C и 15°C для рано- и поздносозревающих групп соответственно).



Рисунок 3.32 – Нуклеотидный состав мтДНК зависит от температуры как у рано-, так и у поздносозревающих групп рыб

Таблица 3.24 – Модели линейной регрессии между Г<sub>Т</sub>А<sub>Т</sub>-сдвиг, Т<sub>Т</sub>Ц<sub>Т</sub>-сдвиг, S<sub>ТГ</sub>-S<sub>АЦ</sub>, температурой и временем полового созревания.

Модель	Переменная	Коэффициент	р-значение
2S. Г <sub>Т</sub> А <sub>Т</sub> -сдвиг ~ Температура; N=333	Интерсепта	0.174308	8.01e-14 ***
1 71 7	Температура	0.002867	0.00802 **
2Т. Т <sub>т</sub> Ц <sub>т</sub> -сдвиг ~ Температура: N=333	Интерсепта	0.6054590	< 2e-16 ***
1 21 7	Температура	0.0036119	2.63e-05 ***

2U. S <sub>TГ</sub> -S <sub>АЦ</sub> ~ Температура; N=333	Интерсепта	0.3576532	< 2e-16 ***
1 11 /	Температура	0.0034870	3.22e-07 ***
2V. Г <sub>Т</sub> А <sub>Т</sub> -сдвиг ~ Время полового	Интерсепта	0.1876786	7.27e-16 ***
созревания; N=217	Время полового созревания	0.0007169	0.821
2W. Т <sub>Т</sub> Ц <sub>Т</sub> -сдвиг ~ Время полового	Интерсепта	0.659894	< 2e-16 ***
созревания; N=217	Время полового созревания	0.006972	0.00534 **
2X. S <sub>TГ</sub> -S <sub>АЦ</sub> ~ Время полового созревания;	Интерсепта	0.390449	<2e-16 ***
N=217	Время полового созревания	0.004200	0.0234 *
2AA. Температура ~ scale(ГтАт-слвиг) +	Интерсепта	19.2908	< 2e-16 ***
scale(Т <sub>Т</sub> Ц <sub>Т</sub> -сдвиг); N=131	scale(Г <sub>Т</sub> А <sub>Т</sub> -сдвиг)	1.5234	0.000278 ***
	scale(Т <sub>Т</sub> Ц <sub>Т</sub> -сдвиг)	2.0620	1.06e-06 ***
2AB. Время полового созревания ~	Интерсепта	4.4415	< 2e-16 ***
scale( $\Gamma_T A_T$ -сдвиг) + scale( $T_T \Pi_T$ -сдвиг);	scale(Г <sub>Т</sub> А <sub>Т</sub> -сдвиг)	0.4793	0.21477
N=131	scale(ТтЦт-сдвиг)	1.1892	0.00244 **

Таблица 3.25 – Ана	филогенетической ин	ерции (PGLS).
--------------------	---------------------	---------------

Модель	Переменная	Коэффициент	р-значение
2AC. Г <sub>Т</sub> А <sub>Т</sub> -сдвиг ~ Температура; N=131, lambda [ ML] : 1.000	Интерсепта	0.1158136	0.2910
	Температура	0.0015013	0.1548
$2AD T_T $ Ц $_T$ -сдвиг ~	Интерсепта	0.76067668	<2e-16 ***

Температура; N=131, lambda [ ML] : 0.979	Температура	0.00073197	0.5089
2AE. S <sub>TГ</sub> -S <sub>АЦ</sub> ~ Температура; lambda [ ML] : 0.991	Интерсепта	0.40049676	1.629e-09 ***
	Температура	0.00130582	0.09781.
2AF. Температура ~ scale( $\Gamma_T A_T$ -сдвиг) + scale( $T_T U_T$ -сдвиг); N=131, lambda [ ML] : 0.769	Интерсепта	14.73620	2.831e-05 ***
	scale(Г <sub>Т</sub> А <sub>Т</sub> -сдвиг)	2.10675	0.004665 **
	scale(Т <sub>Т</sub> Ц <sub>Т</sub> -сдвиг)	2.46107	0.003183 **
2AG. Время полового созревания ~ scale(Г <sub>Т</sub> А <sub>Т</sub> -сдвиг) + scale(Т <sub>Т</sub> Ц <sub>Т</sub> -сдвиг); N=131, , lambda [ ML] : 0.493	Интерсепта	8.754531	1.02e-06 ***
	scale(Г <sub>Т</sub> А <sub>Т</sub> -сдвиг)	1.078439	0.02547 *
	scale(Т <sub>Т</sub> Ц <sub>Т</sub> -сдвиг)	0.070535	0.89899

Было продемонстрировано, что S<sub>TT</sub>-S<sub>AU</sub> выше у рыб, обитающих в теплой воде, по сравнению с рыбами, обитающими в холодной воде, как у раносозревающих (р-значение = 0,0002245, U-критерий Манна-Уитни, 29 холодноводных видов по сравнению с 30 тепловодными видами), так и в группах с длительным созреванием (р-значение = 0,00105, U-критерий Манна-Уитни, 35 холодноводных и 37 тепловодных видов).

Учитывая повышенную долю A<sub>T</sub>>Г<sub>T</sub> у долгоживущих млекопитающих по сравнению с короткоживущими, продемонстрированную в предыдущей главе, мы исследовали потенциальную связь S<sub>TΓ</sub>-S<sub>AU</sub> как с температурой, так и со временем созревания у рыб (Формула 3.4):

все p-значения < 9,8e-05.

Оба фактора положительно влияют на S<sub>TF</sub>-S<sub>AU</sub>, однако температура оказывает большее влияние на S<sub>TF</sub>-S<sub>AU</sub> по сравнению со временем созревания (все коэффициенты стандартизированы, Таблица 3.26).

Таблица 3.26 – Модели линейной регрессии между Г<sub>т</sub>А<sub>т</sub>-сдвигом, Т<sub>т</sub>Ц<sub>т</sub>сдвигом, S<sub>TГ</sub>-S<sub>АЦ</sub> и температурой + временем созревания.

Модель	Переменная	Коэффициент	р-значение
2АН. Г <sub>Т</sub> А <sub>Т</sub> -сдвиг ~ scale(Температура+2) + scale(Время полового созревания); N=131	Интерсепта	0.20311	<2e-16 ***
	scale(Температура+2)	0.04612	0.00233 **
	scale(Время полового созревания)	0.01452	0.32360
2AI. Т <sub>Т</sub> Ц <sub>Т</sub> -сдвиг ~ scaleТемпература+2) + scale(Время полового созревания); N=131	Интерсепта	0.70170	< 2e-16 ***
	scale(Температура+2)	0.04000	0.000641 ***
	scale(Время полового созревания)	0.04201	0.000292 ***
2AJ. S <sub>TT</sub> -S <sub>AЦ</sub> ~ scaleTемпература+2) + scale(Время полового созревания); N=131	Интерсепта	0.421937	< 2e-16 ***
	scale(Температура+2)	0.047876	1.10e-08 ***
	scale(Время полового созревания)	0.031083	9.82e-05 ***

И температура, и время полового созревания незначительно влияют на S<sub>TT</sub>-S<sub>AU</sub> с учетом филогенетической инерции (Таблица 3.27). Следует отметить, что влияние времени созревания не было значимым в анализе 12-компонентного мутационного спектра. Было продемонстрировано, что и температура, и время созревания оказывают положительное влияние на S<sub>TT</sub>-S<sub>AU</sub>, где температура оказывает большее влияние по сравнению со временем созревания.

Модель	Переменная	Коэффициент	р-значение
2АК. Г <sub>Т</sub> А <sub>Т</sub> -сдвиг ~ log <sub>2</sub> (Температура+2) + log <sub>2</sub> (Время полового созревания); N=131, lambda [ ML] : 1.000	Интерсепта	0.0256617	0.83270
	log <sub>2</sub> (Температура+2)	0.0201862	0.08424 .
	log2(Время полового созревания)	0.0115187	0.10653
2AL. Т <sub>Т</sub> Ц <sub>Т</sub> -сдвиг ~ log <sub>2</sub> (Температура+2) + log <sub>2</sub> (Время полового созревания); N=131, lambda [ ML] : 0.975	Интерсепта	0.7126050	4.671e-12 ***
	log <sub>2</sub> (Температура+2)	0.0120913	0.3161
	log2(Время полового созревания)	0.0038918	0.5652
2AM. S <sub>TГ</sub> -S <sub>АЦ</sub> ~ log <sub>2</sub> (Температура+2) + log <sub>2</sub> (Время полового созревания); N=131, lambda [ ML] : 0.986	Интерсепта	0.3198276	1.491e-05 ***
	log <sub>2</sub> (Температура+2)	0.0184000	0.03428 *
	log2(Время полового созревания)	0.0097012	0.04748 *

Таблица 3.27 – Анализ филогенетической инерции (PGLS).

И температура, и время созревания оказывают незначительно значимое влияние на S<sub>TT</sub>-S<sub>AU</sub> с учетом филогенетической инерции. В заключение было продемонстрировано, что на нейтральное содержание нуклеотидов в мтДНК рыб влияют два фактора: сильный, зависящий от температуры мутаген и более слабый, связанный с продолжительностью жизни мутаген.

## 3.2.3 Мутационный спектр мтДНК чувствителен к температуре у всех классов позвоночных.

Температурная чувствительность мутационных спектров мтДНК у рыб позволяет предположить, что это явление может быть универсальным. Чтобы проверить это, было проведено сравнение мутационных спектров мтДНК пяти классов позвоночных, сгруппированных в категории хладнокровных (Actinopterygii, Amphibia, Reptilia) и теплокровных (Mammalia, Aves) (см.

Материалы и методы). Сравнивая мутационный спектр мтДНК ( $A_T > \Gamma_T$  и  $A_T > \Gamma_T / T_T > U_T$ ) между этими группами, была продемонстрирована ожидаемая тенденция: увеличение  $A_T > \Gamma_T$  и  $A_T > \Gamma_T$  асимметрии у теплокровных классов (Рисунок 3.34, W = 430270, р-значение = 5,878е-08), что указывает на то, что мутационный спектр мтДНК чувствителен к температуре у всех позвоночных.



Рисунок 3.33 – Сравнение между классами показывает повышение A<sub>T</sub>>Г<sub>T</sub> асимметрии у теплокровных и видов

Поскольку мутационный спектр мтДНК всех позвоночных животных зависит от температуры, ожидается, что содержание нейтральных нуклеотидов в геноме также может формироваться под влиянием температурно-зависимого мутационного сдвига: теплокровные виды (млекопитающие и птицы) могут быть более  $A_T$  бедные и богатые  $\Gamma_T$  по сравнению с хладнокровными. Используя содержание нуклеотидов в четырехкратно вырожденных сайтах 12 генов (кроме НД6), был продемонстрирован дефицит  $A_T$  (Рисунок 3.34, W = 488941, р-значение <2,2e-16) и избыток  $\Gamma_T$  у теплокровных животных.



Рисунок 3.34 – Сравнение между классами показывает избыток А<sub>т</sub> и дефицит Г<sub>т</sub> у теплокровных видов

Этот результат предполагает потенциально широкое влияние температуры (Рисунок 3.35, W = 14, р-значение < 2,2e-16), которое в ближайшем будущем может быть описано для других таксонов.



Рисунок 3.35 – Распределение температуры тела между классами

## 3.3 Мутационный спектр мтДНК в опухолевых тканях человека изменяется во время канцерогенеза и связан со скоростью пролиферации клеток в тканях предшественниках.

Результаты недавнего комплексного исследования соматических мутаций мтДНК в опухолевых тканях человека показали, что мутационные спектры мтДНК консервативны при различных типах рака с редкими трансверсиями и явным преобладанием транзиций Ц<sub>Т</sub>>Т<sub>Т</sub> и А<sub>Т</sub>>Г<sub>Т</sub> [4]. Однако незначительные изменения в мутационных спектрах могут указывать на стадию рака или скорость тканеспецифического метаболизма. Чтобы проверить, изменяется ли мутационный спектр мтДНК во время канцерогенеза, была использована коллекция из 7611 соматических митохондриальных мутаций из 37 типов опухолей. полученных ИЗ консорциума The Cancer Genome Atlas И сгруппированных в 21 группу по типу ткани предшественника.

Чтобы ранжировать все соматические мутации от ранних до поздних, были оценены ЧВА каждого варианта: высокие ЧВА являются признаком ранних мутаций (возникших на ранних стадиях канцерогенеза и, следовательно, присутствующих в большинстве раковых клеток), а низкие ЧВА признак поздних мутаций (возникших на последних стадиях канцерогенеза и, таким образом, представленных в субклоне раковых клеток). Далее проводилось сравнение специфичных для рака ЧВА между образцами с (N = 5436) и без (N=2265) соответствующих мутаций в нормальных тканях. Как и ожидалось, было продемонстрировано увеличение специфичных для рака ЧВА в образцах, где та же мутация, даже с низким ЧВА, наблюдалась в нормальной ткани (Рисунок 3.35, р-значение = 0,005, U-критерий Манна-Уитни).



Рисунок 3.35 – ЧВА опухолей приблизительно соответствует времени возникновения варианта

Примечание: Частота высока для ранних вариантов (наблюдается и в нормальных тканях, N = 5436) и низка для поздних вариантов (не наблюдается в нормальных тканях, N = 2265).

В целом этот анализ подтверждает, что имеется возможность приблизительно определить время возникновения мутаций, используя ЧВА: мутации с высокими показателями ЧВА возникли на ранних стадиях, а низкие

показатели ЧВА - на поздних стадиях. Разделив все варианты по медиане ЧВА (4,5%), было продемонстрировано значительное снижение Ts/Tv среди поздних вариантов с 14,9 до 10,3 (Рисунок 3.36, отношение шансов Фишера = 0,69, р-значение = 2,524e-05, N = 7611).



Рисунок 3.36 – Ts/Tv выше для ранних (с высокой ЧВА) и поздних (с низкой ЧВА) вариантов

Примечание: Левая панель: интегральный анализ включает все соматические варианты. Центральная панель: анализ конкретного типа опухоли учитывает тенденцию, специфичную для каждого типа опухоли; каждый тип опухоли отмечен сегментом, Детский рак мозга отмечен красным сегментом. Правая панель: анализ конкретного образца демонстрирует увеличение ЧВА(Ts) по сравнению с ЧВА(Tv) в каждом образце.

Этот результат устойчив к исключению потенциально некачественных вариантов: когда были удалены 25% вариантов с самыми низкими р-значениями (варианты с  $-\log_{10}(p$ -значение) <= 59) был продемонстрирован почти идентичный результат (медиана ЧВА = 9,4%, отношение шансов Фишера = 0,67, р-значение = 0,0002, N = 5709). Чтобы исключить потенциальное влияние различных типов опухолей на интегральный анализ, был проведен тест, существует ли увеличение

Ts/Tv с ЧВА для каждого типа рака индивидуально. Все варианты были разделены на редкие и распространенные (по медиане ЧВА) и для каждой группы были рассчитаны Ts/Tv и среднее ЧВА. Для визуализации были выделены сегменты графика для каждого типа опухоли, соединяющие Ts/Tv и ЧВА или редкие варианты с Ts/Tv и ЧВА. В целом эти сегменты демонстрируют положительную динамику (р-значение = 0,0006, парный U-критерий Манна-Уитни, N = 36), однако индивидуально только детский рак мозга продемонстрировал значительное увеличение Ts/Tv с ЧВА (медиана ЧВА = 10,1%, отношение шансов Фишера = 0,06, р-значение = 1,14e-06, N = 186). Удаление детского рака мозга из набора данных не влияет на результаты (рзначение = 0,001, парный U-тест Манна-Уитни, N = 35). Чтобы проанализировать вклад  $A_T > \Gamma_T$  по сравнению с  $\coprod_T > T_T$ , было проведено сравнение ЧВА( $A_T > \Gamma_T$ ) и ЧВА( $U_T > T_T$ ) внутри каждого образца и установлено, что  $A_T > \Gamma_T$  характеризуется несколько более высоким ЧВА по сравнению с  $\prod T > T_T$ (медиана различия между ЧВА (А<sub>Т</sub>>Г<sub>т</sub>) и ЧВА(Ц<sub>Т</sub>>Т<sub>т</sub>) составляет 0,1%, рзначение = 0,0097, парный критерий Манна-Уитни, N=983). Это означает, что  $A_T > \Gamma_T$  в среднем происходит раньше, чем  $\prod T > T_T$ . Поскольку отношения между этими двумя наиболее распространенными транзициями не зависят от редких трансверсий, этот результат независимо подтверждает надежность выводов об изменении характера мутагенеза при канцерогенезе.

Таким образом, было продемонстрировано, что мутационный спектр мтДНК меняется в ходе канцерогенеза: как Ts/Tv, так и A<sub>T</sub>>Г<sub>T</sub>/Ц<sub>T</sub>>T<sub>T</sub> уменьшаются от ранних вариантов к поздним, что позволяет предположить, что митохондриальное мутагенное микроокружение является динамичным на протяжении всего канцерогенеза.

Многие соматические митохондриальные мутации, отмеченные в опухолях человека, могли произойти до начала канцерогенеза и, таким образом, могут указывать на свойства здоровых тканей. Одним из наиболее выдающихся тканеспецифичных свойств, связанных как с мутагенезом[93], так и с митохондриальным метаболизмом, является скорость обновления стволовых клеток в данной ткани. На основании обширного поиска в литературе была получена приблизительная скорость пролиферации стволовых клеток в каждом из 21 образца нормальной ткани, являющейся предком анализируемых типов опухолей (Таблица 3.28)[93–95].

Таблица 3.28 – Скорость пролиферации стволовых клеток в различных тканях человека (в днях).

Опухолевая ткань: аннотация GTEx	Скорость пролиферации стволовых клеток данной ткани, дни
Матка	4
Толстая/прямая кишка	3.5
Желудок	5.5
Шейка матки	6.0
Пищевод	11.0
Миелоидная ткань	30.0
Лимфоидная ткань	30.0
Молочная железа	84.5
Простата	120.0
Кожа	147
Желчевыводящие пути	400.0
Поджелудочная железа	360
Печень	400.0
Почки	1000.0
Щитовидная железа	4138.0
Легкие	5143
Эпителий тонкой кишки	2-4
ЦНС	10000.0
Яичники	11000

Сердечная мускулатура	25300
Скелетная мускулатура	5510

Используя приблизительную скорость пролиферации стволовых клеток в каждом из 21 образца нормальной ткани, предшествующей проанализированным типам опухолей, все опухоли были разделены на три категории: быстро пролиферирующие – со скоростью обновления стволовых клеток менее месяца (матка, ободочная/прямая кишка, желудок, шейка матки, пищевод, голова/шея, лимфоидная и миелоидная ткани), промежуточные – со скоростью обновления стволовых клеток более месяца и менее десяти лет (простата, кожа, мочевой пузырь, желчевыводящие пути, поджелудочная железа, печень и почки) и медленно пролиферирующие – со скоростью обновления стволовых клеток более десяти лет (щитовидная железа, легкие, ЦНС, яичники).

Тs/Tv для каждой группы оценивалась после тысячи 50% повторных выборок мутаций с помощью jackknife, что подтвердило тенденцию увеличения Ts/Tv в группах тканей с медленной и быстрой пролиферацией. Оценивая мутационный спектр для каждой категории, было продемонстрировано увеличение отношения Ts/Tv от 10,0 в быстро пролиферирующей категории до 12,6 в промежуточной и до 14,5 в медленно пролиферирующей (Рисунок 3.37). Обе транзиции  $U_T > T_T$  и  $A_T > \Gamma_T$  независимо вносят вклад в эту тенденцию ( $A_T > \Gamma_T/Tv$ : 2,8, 3,8, 4,6;  $U_T > T_T/Tv$ : 5,0, 6,6, 7,5 для быстро, промежуточно и медленно пролиферирующих тканей).

Поскольку трансверсии являются очень низкочастотными вариантами и некоторые из них могут быть ошибками секвенирования[89] был проведен тест, сохраняется ли данная тенденция при анализе только транзиций. Была рассчитана доля  $A_T > \Gamma_T$  на все остальные транзиции ( $U_T > T_T$ ,  $T_T > U_T$ ,  $\Gamma_T > A_T$ ), а также доля  $U_T > T_T$  на все остальные транзиции ( $A_T > \Gamma_T$ ,  $T_T > U_T$ ,  $\Gamma_T > A_T$ ). Было продемонстрировано соответствующее увеличение фракции  $A_T > \Gamma_T$  ( $A_T > \Gamma_T/T_S$ : 0,39, 0,44, 0,46), но не фракции  $U_T > T_T$  ( $U_T > T_T/T_S$ : 1,00,1,12, 1,08).

Наконец, было проведено сравнение эффектов частых транзиций друг с другом и продемонстрировано увеличение фракции  $A_T > \Gamma_T$  по сравнению с  $U_T > T_T$  в медленно пролиферирующих тканях ( $A_T > \Gamma_T/U_T > T_T$ : 0,56, 0,58, 0,61 для быстро, промежуточно и медленно делящихся тканей).



Рисунок 3.37 – Медленно делящиеся ткани обогащены транзициями, прежде всего A<sub>T</sub>>Г<sub>T</sub>

Таким образом, Ts/Tv увеличивается от быстро пролиферирующих тканей к медленно пролиферирующим, и этот эффект связан с увеличением доли транзиций A<sub>T</sub>>Г<sub>т</sub>.

Объединив два наблюдения, описанные выше: повышенная доля  $A_T > \Gamma_T$  на ранних стадиях опухоли и повышенная доля  $A_T > \Gamma_T$  в медленно пролиферирующих тканях, ожидается, что вероятность транзиции  $A_T > \Gamma_T$  является функцией как (i) относительного времени возникновения каждой соматической мутации, приблизительно как 1 минус ЧВА, так и (ii) скорости тканеспецифической пролиферации. Чтобы проверить это, была использована множественная логистическая регрессия. Было продемонстрировано, что вероятность транзиции  $A_T > \Gamma_T$  действительно уменьшается как со временем

канцерогенеза, так и со скоростью пролиферации ткани (Рисунок 3.38, Таблица 3.29).

Таблица 3.29 – Результаты множественных логистических регрессий, связывающих биномиальную переменную A<sub>T</sub>>Г<sub>T</sub> (1 = да или 0 = нет) с независимыми переменными.

Модель	Переменная	Коэффициент	р-значение
3А. Основная модель, изображенная на рисунке 1 приложения. Зависимая переменная равна 1 (A <sub>T</sub> >Γ <sub>T</sub> ) или 0 (любые другие замены), N = 7611.	Интерсепта	-0.95	< 2e-16
	Скорость пролиферации	0.049	0.049
	ЧВА	0.088	0.0004
3B. Идентична модели 3A, но непрерывная переменная «Скорость пролиферации» заменена фиктивной переменной, которая равна 1 в случае быстро пролиферирующих образцов и 0 в случае промежуточной и медленной пролиферации.	Интерсепта	-0.95	< 2e-16
	Фиктивная(быстрая пролиферация)	-0.065	0.013
	ЧВА	0.093	0.0002
3C. Идентична модели 3B, но проанализированные мутации — это только транзиции, N = 7035.	Интерсепта	-0.842	< 2e-16
	Фиктивная(быстрая пролиферация)	-0.054	0.043
	ЧВА	0.071	0.0058

Примечание: в качестве независимых переменных были использованы скорость пролиферации тканеспецифических клеток; ЧВА каждой мутации в опухолевых образцах. Результаты были получены с помощью функции glm в R. Все переменные были масштабированы.



Рисунок 3.38 – Вероятность А<sub>т</sub>>Г<sub>т</sub> выше на ранних стадиях рака (когда клетки более нормоксичны и медленно делятся) и в тканях с низкой скоростью пролиферации

Примечание: Ось X отражает стадии канцерогенеза, аппроксимированные 1 – ЧВА. Ось Y отмечает скорость пролиферации стволовых клеток. Каждый кружок отражает одну мутацию. Размер кружков пропорционален вероятности того, что данная мутация является  $A_T > \Gamma_T$ , где вероятность получена из модели логистической регрессии и нормализована до диапазона 0-1. Красные кружки обозначают  $A_T > \Gamma_T$ , серые кружки — все остальные замены. Видно, что кружки (вероятность  $A_T > \Gamma_T$ ) уменьшаются слева направо, отражая усиление гипоксии на более поздних стадиях опухоли, и сверху вниз, отражая градиент тканей с низкой и высокой скоростью пролиферации.

Таким образом, замены A<sub>T</sub>>Г<sub>T</sub> чаще встречаются (i) на более ранних стадиях канцерогенеза и (ii) в опухолях, происходящих из медленно пролиферирующих тканей.

Изменения мутационных спектров в процессе канцерогенеза, а также между медленно и быстро пролиферирующими тканями могут быть связаны с различным уровнем аэробного метаболизма – от гипоксии до нормоксии. Было показано, например, что мтДНК при гипоксической опухоли толстой кишки

демонстрирует трехкратное снижение частоты мутаций по сравнению с нормоксичной прилегающей неопухолевой тканью [96]. Этот эффект был главным образом обусловлен уменьшением обеих распространенных транзиций ( $U_T > T_T$  и  $A_T > \Gamma_T$ ), которые, как ожидается, будут связаны с окислительным повреждением и, следовательно, станут более редкими при гипоксических опухолевых состояниях. Принимая во внимание, что практически все виды рака со временем становятся более гипоксическими [97], фракция  $A_T > \Gamma_T$  может быть интерпретирована как маркер окислительного повреждения в данный момент развития опухоли: от высокоаэробного, нормоксического, на ранней стадии (высокая фракция  $A_T > \Gamma_T$ ) до поздних, гипоксических стадий (низкая фракция  $A_T > \Gamma_T$ ).

Используя показатели гипоксии, полученные для многих отдельных образцов опухолей в рамках Консорциума ICGC/TCGA [98], была напрямую протестирована связь между гипоксией и фракцией  $A_T > \Gamma_T$ . Анализ всего набора данных показал лишь косвенные доказательства корреляции (ро-коэффициент Спирмена = -0,046, р-значение = 0,1869, N = 828), однако эта связь стала более очевидной, когда сравнивался верхний дециль наиболее гипоксических видов рака (частота  $A_T > \Gamma_T$  составляет 0,20, N = 94 с оценкой Buffa выше или равной 32) по сравнению со всеми другими, менее гипоксическими видами рака (частота  $A_T > \Gamma_T$  составляет 0,28, N = 734, оценка Buffa ниже 32) (Рисунок 3.39, р-значение = 0,0047, U-тест Манн-Уитни).



Рисунок 3.39 – Вероятность A<sub>T</sub>>Г<sub>T</sub> снижается в гипоксических образцах

Примечание: Схема слева демонстрирует ожидаемые изменения уровня нормоксии в различных тканях (например, медленно делящиеся ооциты и быстро делящийся эпителий толстой кишки) и стадии опухоли (все типы рака со временем постепенно становятся более гипоксическими). Фракция  $A_T > \Gamma_T$  (красные звездочки) отражает окислительное повреждение в данный момент развития опухоли: от высокоаэробных, нормоксических (красный цвет фона) ранних стадий до поздних, гипоксических стадий (серый цвет фона). Гистограммы справа демонстрируют различия в частотах  $A_T > \Gamma_T$  между нормоксичными образцами (734 образца с оценкой Buffa ниже 32) и гипоксическими образцами (94 образца с оценкой Buffa выше или равными 32). Каждый столбик соответствует одному образцу, где красная часть отражает частоту  $A_T > \Gamma_T$  для конкретного образца.

В целом, предполагается, что мутационный спектр мтДНК чувствителен к уровню аэробного метаболизма посредством замен A<sub>T</sub>>Г<sub>т</sub>, подверженных окислительному повреждению, которое выше на ранних стадиях опухоли и в медленно пролиферирующих тканях.

## ГЛАВА 4. ОБСУЖДЕНИЕ

Используя эволюционный подход, мы было реконструировано большое количество мутаций мтДНК у сотен видов позвоночных. Этот набор данных позволил сравнить мутационные спектры между видами с разным жизненным циклом и физиологическими характеристиками. Хотя предыдущие исследования показали, что существуют различия в мутационном спектре мтДНК между видами [10], причина этих изменений не объяснена. В данной диссертации было предположено, что подавляющее большинство четырехкратно вырожденных синонимических замен фактически нейтральны и, следовательно, различия в мутационных спектрах отражают разные мутагены, а не отбор. Предполагается, что нет никаких доказательств того, что отбор действует на синонимичные кодоны в мтДНК позвоночных. Фактически, исследования показали, что мутационный сдвиг формирует использование синонимичных кодонов в мтДНК человека [20] и мтДНК других млекопитающих.

Было продемонстрировано, что скорость замен  $A_T > \Gamma_T$  и ее эффекты, такие как нейтральные частоты нуклеотидов,  $\Gamma_T A_T$ -скос и асимметрия использования кодонов, увеличиваются с продолжительностью жизни организма и возрастом. Этот результат был обнаружен в трех довольно разных эволюционно-временных масштабах: (i) месяцы и годы для соматических мутаций и de novo мутаций в тканях зародышевой линии; (ii) десятки или сотни тысяч лет для нейтральных внутривидовых полиморфизмов мтДНК (что представляет собой среднее время сегрегации нейтральных внутривидовых полиморфизмов мтДНК, зафиксированных между видами. Эта тенденция, универсальная во времени и для всех видов млекопитающих, требует специального объяснения.

Предполагается, что полученные результаты В первую очередь демонстрируют именно процесс мутагенеза мтДНК, а не естественного отбора. Во-первых, эффект отбора не ожидается в случае чрезвычайно редких, с ЧВА менее 1%, мутаций мтДНК, полученных в подходе дуплексного секвенирования. Во-вторых, из-за небольшого количества или отсутствия доказательств отбора мтДНК на синонимичных четырехкратно вырожденных сайтах В млекопитающих [100] были рассматрены полиморфные варианты.

На митохондриальный мутационный спектр, и особенно замены A<sub>T</sub>>Г<sub>т</sub>, могут влиять как ошибки гамма-ДНК-полимеразы, так и повреждения, связанные со средой внутри. Недавний элегантный эксперимент с мышами, гомозиготными по дефицитной экзонуклеазе гамма-ДНК-полимеразе, показал, что градиенты A<sub>T</sub>>Г<sub>T</sub> и Ц<sub>T</sub>>Т<sub>T</sub> в основном формируются эндогенным мутагеном, связанным с асинхронной репликацией ДНК, а не ошибками ДНК-полимеразы [52]. Учитывая, что активные формы кислорода являются основным вредным побочным продуктом аэробного метаболизма, предполагается, что ключевые результаты данной работы могут быть связаны с эффектами окислительного повреждения.

Данная гипотеза основана на высокой чувствительности  $A_T > \Gamma_T$  к ВПДОС: одноцепочечная ДНК более уязвима к агентам, повреждающим ДНК [101]. ДНК существует в одноцепочечной форме в процессе репликации, транскрипции и репарации ДНК. Мутационный градиент вдоль большой и малой дуг репликации мтДНК [52] предполагает, что ВПДОС во время репликации более мутагенно в случае мтДНК, чем ВПДОС во время транскрипции и репарации. Однако ненулевые точки пересечения, наблюдаемые для обеих распространенных транзиций (Рисунок 3.2), когда ВПДОС, управляемое репликацией, равно нулю, предполагают, что фоновый мутагенез, вероятно связанный со ВПДОС, управляемым транскрипцией или репарацией, может играть некоторую роль.

Старение связано с множественными изменениями клеточного метаболизма. Например, ткани старых мышей производят меньше АТФ и имеют более низкие уровни НАД+, чем соответствующие ткани молодых мышей [102]. Учитывая, что митохондрии являются центральным узлом энергетического метаболизма, их функционирование также меняется с возрастом [103]. Однако до сих пор неясно, существует ЛИ универсальная закономерность метаболических изменений во время старения. Здесь было показано, что мутагенез мтДНК зависит от возраста организма и видоспецифичного времени генерации поколения, что подразумевает некоторые общие изменения внутри митохондрий. Предполагается, что окислительное повреждение может быть универсальным фактором, который усугубляется с возрастом. Действительно, уровень карбонилирования белка в фибробластах пожилых людей выше, чем у молодых [104], а перекисное окисление липидов связано со многими возрастными заболеваниями [105]. Более того, предполагается, что митохондрии являются основным источником возрастных окислительных повреждений: экспрессия митохондриально-ориентированного антиоксидантного фермента каталазы — увеличивает продолжительность жизни мышей [106].

Важно подчеркнуть, что обнаруженная замена А>Г отличается от хорошо документированной характеристики активных форм кислорода: трансверсий Г>Т, которые являются результатом АФК-индуцированной модификации

гуанина - 8-оксогуанина. Ранее было показано, что этот ожидаемый признак окислительного повреждения, а именно замены Г>Т, не встречается в мтДНК: он встречается очень редко, не зависит от возраста и демонстрирует слабую связь с окислительным повреждением, если таковая имеется [4,34,45]. Интересно, что трансверсии Г>Т ( $\Gamma_T$ >T<sub>T</sub> и Ц<sub>Т</sub>>A<sub>T</sub>) в данной работе, несмотря на редкость и слабый эффект, демонстрируют тенденцию, противоположную A<sub>T</sub>> $\Gamma_T$ , что может отражать классический признак окислительного повреждения, более выраженный у короткоживущих видов из-за их более высокой относительной скорости основного обмена. Однако его эффект действительно слаб.

Здесь предполагается, что вместо замен  $\Gamma > T$  дезаминирование аденозина, являющееся основным источником замен A<sub>T</sub>>Г<sub>T</sub>, связано с окислительным повреждением одноцепочечной мтДНК. Возможно, что А>Г является новым маркером окислительного повреждения мтДНК, характерным ДЛЯ одноцепочечной ДНК. Полученные в данной работе результаты и более ранние публикации (обсуждаемые ниже) подтверждают эту гипотезу. Недавнее глубокое секвенирование de novo мутаций мтДНК в ооцитах старых и молодых мышей подтвердило, что повышенная доля замен A<sub>T</sub>>Г<sub>T</sub> является лучшим признаком старения мтДНК [80]. Это открытие также предполагает, что продолжительная фиксация мтДНК внутри спящих ооцитов может быть универсальной чертой всех видов млекопитающих – и, таким образом, вариации в мутационных спектрах мтДНК между разными видами могут отражать возрастные процессы у разных видов млекопитающих.

Избыток замен A:T > Г:Ц недавно наблюдался в аэробных и анаэробных условиях у Escherichia coli. Важно подчеркнуть, что эффект был обусловлен отстающей цепью, проводившей больше времени в одноцепочечных условиях [107]. Результаты этих экспериментов согласуются с гипотезой о том, что транзиции А>Г связаны с окислительным повреждением одноцепочечной ДНК.

Имеются доказательства того, что окислительный стресс может вызывать мутации А:Т > Г:Ц, полученные от эукариот, то есть дрожжей [108] и мышей [109]. В обоих исследованиях очевидное снижение окислительного стресса (путем наложения аэробных условий или использования антиоксидантов соответственно) приводило к специфическому снижению показателей А:Т > Г:Ц, подразумевая, что часть этих мутаций возникает в результате окислительного повреждения. Интересно, что в обоих исследованиях использовались системы с дефицитом репарации ошибочных спариваний (англ. MMR - DNA mismatch repair), что приводило к заметному увеличению количества мутаций  $A:T > \Gamma:U$ по сравнению с системами, владеющими MMR. Большая часть этого MMRзависимого увеличения должна быть зависима от окислительного стресса. Это потенциально подразумевает, что MMR, помимо своей традиционной роли в ошибок репликации, исправлении также участвует В восстановлении окислительных повреждений. Это согласуется с растущим объемом данных других исследований [110]. Эта интригующая область остается спорной и требует дополнительных исследований.

Было показано, что А>Г является наиболее асимметричной заменой в ядерном геноме человека, что является отличительной чертой мутагена, действующего через специфическое повреждение одноцепочечной ДНК [111]. Хотя ключевой мутаген до сих пор не определен, предполагается, что в случае мтДНК он может быть связан с окислительным повреждением.

Было показано, что доля замен A<sub>T</sub>>Г<sub>т</sub> в мтДНК положительно коррелирует с температурой окружающей среды у видов Actinopterygii. Поскольку более высокая температура связана с усилением аэробного метаболизма [112], предполагается, что A<sub>T</sub>>Г<sub>т</sub> является маркером окислительного повреждения.

Замены А>Г, связанные с окислительным повреждением, могут быть ключевым процессом, объясняющим давнюю эволюционную загадку повышенного содержания GC в аэробных по сравнению с анаэробными бактериями [113–115]. Согласно общепринятым знаниям, ожидается, что Г>Т будет выше у аэробных бактерий по сравнению с анаэробными, что делает геномы первых более бедными ГЦ и богатыми АТ. Однако многочисленные эмпирические данные свидетельствуют об обратном: у аэробных бактерий содержание ГЦ повышено по сравнению с анаэробными. Предполагая, что

окислительное повреждение оказывает одинаковое мутагенное воздействие как на митохондриальные, так и на бактериальные геномы, можно считать, что высокая скорость замещения А>Г у аэробных бактерий может сделать их геномы более Г-богатыми.

Было показано, например, что рак толстой кишки при переходе в более гипоксическую стадию снижает частоту мутаций мтДНК за счет замедления обоих распространенных транзиций: Ц<sub>Т</sub>>Т<sub>Т</sub> и A<sub>T</sub>>Г<sub>Т</sub> [96]. В данном исследовании были частично подтверждены выводы Эриксон и др. 2012 (Ericson et al. 2012), распространив их на все типы опухолей и продемонстрировав, что гипоксические состояния, чаще наблюдаемые в быстроделящихся клетках и на поздних стадиях опухолей, связаны с уменьшением доли транзиций A<sub>T</sub>>Г<sub>т</sub>.

Избыток Г<sub>т</sub> в мтДНК долгоживущих млекопитающих ранее был показан у Лехманн и др. 2006 (Lehmann et al. 2006) [9]. Они предложили объяснение этого наблюдения основанное на отборе, предполагая повышенную стабильность геномов богатых Γт, преимущество долгоживущим что может дать млекопитающим. Полученные результаты показывают, что избыток Гт у быть долгоживущих млекопитающих может нейтральным следствием мутагенеза A<sub>T</sub>>Г<sub>T</sub>, а не результатом отбора. Кроме того, низкий эффективный размер популяции долгоживущих млекопитающих увеличивает силу случайного генетического дрейфа и скорость фиксации слабо вредных вариантов в их мтДНК [116-118], что делает объяснение, основанное на отборе, еще менее Однако, чтобы оценить аргумент, основанный на отборе, вероятным. дополнительно были проанализированы корреляция нейтрального нуклеотидного состава мтДНК (А<sub>т</sub>, Т<sub>т</sub>, Г<sub>т</sub>, Ц<sub>т</sub>) со временем генерации поколения. Было продемонстрировано, что частоты нуклеотидов Ат и Гт в третьем наиболее нейтральном положении кодонов демонстрируют наиболее сильную корреляцию (отрицательную для A<sub>T</sub> и положительную для Г<sub>T</sub>) со временем генерации поколения. Хотя нельзя исключить аргумент, основанный на отборе (если отбор действует на общий нуклеотидный состав, увеличивая содержание ГЦ, можно ожидать, что наиболее сильный эффект будет в 3-й позиции кодона.

Однако необходимы дальнейшие исследования для разделения эффектов мутагенеза и отбора. Продемонстрированные результаты предполагают, что мутагенез также может объяснить наблюдаемый избыток Г<sub>т</sub> у долгоживущих млекопитающих.

Температурная чувствительность мутационного спектра может быть связана с усилением окислительного повреждения у теплых видов с высоким уровнем метаболизма [112]. Чтобы более непосредственно проверить влияние окислительного повреждения, дополнительно были проанализированы экспериментальные данные о потреблении кислорода (мг кислорода на килограмм рыбы в час при 20°C и стандартном уровне активности) для небольшого числа видов рыб с известными мутационными спектрами. Несмотря на чрезвычайно небольшой размер выборки, была продемонстрирована положительная тенденция между потреблением кислорода и A<sub>T</sub>>Г<sub>т</sub> (Рисунок 3.38, ро-коэффициент Пирсона = 0.8239473, р-значение = 1.448e-05, N = 19).



Рисунок 3.40 – Зависимость A<sub>T</sub>>Г<sub>т</sub> от уровня потребления кислорода

Таким образом, главный вывод проведенных анализов о том, что температура влияет на мутационный спектр, можно объяснить с помощью следующих логических шагов: (i) температура связана с уровнем аэробного метаболизма, (ii) окислительное повреждение является побочным продуктом аэробного метаболизма, (iii) A<sub>T</sub>>Г<sub>T</sub> увеличивается из-за окислительного повреждения. Недавние эксперименты по накоплению мутаций с Chironomus riparius и E. coli [59,60] показали, что доля транзиций увеличивается в условиях высоких температур. Однако необходимы дальнейшие экспериментальные чтобы более мутационные исследования, полно понять детальные характеристики температуры. Эффект температуры является плейотропным он влияет не только на уровень аэробного метаболизма и окислительного повреждения, но и на характеристики жизненного цикла, такие как время генерации поколения: действительно, высокая температура связана с короткой продолжительностью жизни. Таким образом, ожидается, что мутационная характеристика температуры будет как минимум двойной.

В текущей работе невозможно однозначно доказать, что замены A<sub>T</sub>>Г<sub>т</sub> вызваны окислительным повреждением, так как пока не существует четко установленной экспериментальной связи между окислительным повреждением и A>Г. Объединив несколько наблюдений, которые в совокупности подтверждают данные утверждения, была предложена рабочая гипотеза, которая поддается проверке и может быть оценена научным сообществом в будущем., поэтому мы хотели бы изложить предположение.

Интересно, что продуктом окислительного (или гидролитического) специфического дезаминирования аденозина одноцепочечной ДНК является инозин, который преимущественно спаривается с Ц и, следовательно, индуцирует мутации А>Г/Т>Ц [119]. Сообщалось, что устойчивые уровни инозина в ДНК составляют порядка 1 на 1 000 000 нуклеотидов [120]. Интересно, что этот уровень не является несовместимым с наблюдаемыми мутационными фракциями А>Г/Т>Ц в старых тканях, которые, как сообщается, составляют примерно 1 мутацию на 100 000 нуклеотидов [34]. Следует отметить, что при
равновесном уровне инозина 10<sup>-6</sup> ожидается, что частота мутаций составит 10<sup>-6</sup> на дупликацию ДНК (т. е. 10<sup>-5</sup> на 10 дупликаций), что более чем достаточно для объяснения наблюдаемых фракций мутантов, поскольку клетки линии в старых тканях претерпевают более 40 дупликаций. В действительности, на скорость мутаций влияет множество неизвестных повреждающих факторов, особенно в реплицирующемся пуле мтДНК. Также вышеуказанные уровни инозина были зарегистрированы для ядерной, а не митохондриальной ДНК. Тем не менее, эти приблизительные оценки подразумевают, ЧТО инозин можно считать многообещающим кандидатом на связь между окислительным стрессом и асимметричными мутациями А>Г/Т>Ц. Данная работа может послужить толчком к лальнейшим исследованиям в этой области. Прямое экспериментальное исследование влияния различных мутагенов [33] и особенно влияния окислительного повреждения [121] и химических модификаций нуклеотидов [122,123] на одноцепочечную мтДНК значительно улучшит понимание мутагенеза мтДНК.

В целом, полученные результаты предлагают ценную возможность значительно расширить способы использования мутационных спектров мтДНК. Например, соматические мутации мтДНК с низкой гетероплазмией из просто нейтрального маркера, используемого для отслеживания клеточных линий [12], могут быть преобразованы в показатель, связанный с клеточно-специфичным состоянием старения, что может быть особенно важно в высоко гетерогенных тканях, таких как рак. De novo мутации мтДНК могут предсказывать биологический возраст человеческих ооцитов, что можно использовать в методах экстракорпорального оплодотворения. Мутационные спектры мтДНК, реконструированные для немодельных видов позвоночных, могут помочь приблизительно оценить среднее время генерации поколения — показатель, который не всегда легко оценить эмпирически.

#### выводы

- Мутационный спектр мтДНК млекопитающих определяется повреждениями ассоциированными с продолжительностью жизни организма, при чём мутации A<sub>T</sub>>Г<sub>T</sub> могут являться специфичным для митохондрий признаком химического повреждения, связанного со старением.
- Мутационный спектр мтДНК рыб и других эктотермных позвоночных повреждениями ассоциированными с температурой среды, что подтверждается положительной корреляцией мутаций A<sub>T</sub>>Г<sub>T</sub> с видоспецифичной температурой окружающей среды и, как следствие, скоростью метаболизма.
- Мутационный спектр мтДНК человека в опухолевых тканях ассоциирован со скоростью пролиферации клеток и с уровнем нормоксии/гипоксии посредством мутаций A<sub>T</sub>>Г<sub>т</sub>.
- Предполагается, что мутационный спектр мтДНК позвоночных животных чувствителен к уровню аэробного метаболизма посредством транзиций A<sub>T</sub>>Г<sub>T</sub>, которые являются уникальным маркером окислительного повреждения клеток, тканей и организмов, специфического для митохондриального генома.

### СПИСОК ЛИТЕРАТУРЫ

- Tyshkovskiy A., Ma S., Shindyapina A.V., Tikhonov S., Lee S.-G., Bozaykut P., Castro J.P., Seluanov A., Schork N.J., Gorbunova V., Dmitriev S.E., Miller R.A., Gladyshev V.N. Distinct longevity mechanisms across and within species and their association with aging // Cell. 2023. Vol. 186, № 13. P. 2929– 2949.e20.
- 2. Haghani A., Li C.Z., Robeck T.R., Zhang J., Lu A.T., Ablaeva J., Acosta-Rodríguez V.A., Adams D.M., Alagaili A.N., Almunia J., Aloysius A., Amor N.M.S., Ardehali R., Arneson A., Baker C.S., Banks G., Belov K., Bennett N.C., Black P., Blumstein D.T., Bors E.K., Breeze C.E., Brooke R.T., Brown J.L., Carter G., Caulton A., Cavin J.M., Chakrabarti L., Chatzistamou I., Chavez A.S., Chen H., Cheng K., Chiavellini P., Choi O.-W., Clarke S., Cook J.A., Cooper L.N., Cossette M.-L., Day J., DeYoung J., Dirocco S., Dold C., Dunnum J.L., Ehmke E.E., Emmons C.K., Emmrich S., Erbay E., Erlacher-Reid C., Faulkes C.G., Fei Z., Ferguson S.H., Finno C.J., Flower J.E., Gaillard J.-M., Garde E., Gerber L., Gladyshev V.N., Goya R.G., Grant M.J., Green C.B., Hanson M.B., Hart D.W., Haulena M., Herrick K., Hogan A.N., Hogg C.J., Hore T.A., Huang T., Izpisua Belmonte J.C., Jasinska A.J., Jones G., Jourdain E., Kashpur O., Katcher H., Katsumata E., Kaza V., Kiaris H., Kobor M.S., Kordowitzki P., Koski W.R., Krützen M., Kwon S.B., Larison B., Lee S.-G., Lehmann M., Lemaître J.-F., Levine A.J., Li X., Li C., Lim A.R., Lin D.T.S., Lindemann D.M., Liphardt S.W., Little T.J., Macoretta N., Maddox D., Matkin C.O., Mattison J.A., McClure M., Mergl J., Meudt J.J., Montano G.A., Mozhui K., Munshi-South J., Murphy W.J., Naderi A., Nagy M., Narayan P., Nathanielsz P.W., Nguyen N.B., Niehrs C., Nyamsuren B., O'Brien J.K., Ginn P.O., Odom D.T., Ophir A.G., Osborn S., Ostrander E.A., Parsons K.M., Paul K.C., Pedersen A.B., Pellegrini M., Peters K.J., Petersen J.L., Pietersen D.W., Pinho G.M., Plassais J., Poganik J.R., Prado N.A., Reddy P., Rey B., Ritz B.R., Robbins J., Rodriguez M., Russell J., Rydkina E., Sailer L.L., Salmon A.B., Sanghavi A., Schachtschneider K.M., Schmitt D., Schmitt T., Schomacher L., Schook L.B., Sears K.E., Seifert A.W., Shafer A.B.A., Shindyapina A.V., Simmons M., Singh K., Sinha I., Slone J., Snell R.G., Soltanmohammadi E., Spangler M.L., Spriggs M., Staggs L., Stedman N., Steinman K.J., Stewart D.T., Sugrue V.J., Szladovits B., Takahashi J.S., Takasugi M., Teeling E.C., Thompson M.J., Van Bonn B., Vernes S.C., Villar D., Vinters H.V., Vu H., Wallingford M.C., Wang N., Wilkinson G.S., Williams R.W., Yan Q., Yao M., Young B.G., Zhang B., Zhang Z., Zhao Y., Zhao P., Zhou W., Zoller J.A., Ernst J., Seluanov A., Gorbunova V., Yang X.W., Raj K., Horvath S. DNA methylation networks underlying mammalian traits // Science. 2023. Vol. 381, № 6658. P. eabq5693.
- 3. Goncharov A.G., Immanuel Kant Baltic Federal University, Popad'in K.Y., Kozenkov I.I., Mikhailova A.G., Lobanova V.V., Tatarkina M.A., Ilyushchenko D.V., Reverchuk I.V., Tynterova A.M., Dvirsky A.A., Dzhigkaev A.K.,

Timofeeva E.V., Gunbin K.V. Methodological approaches to the study of the determinants of somatic mitochondrial heteroplasmy in the elderly // J. Ural Med. Acad. Sci. Bulletin of the Ural Medical Academie Science, 2022. Vol. 19,  $N_{\rm P}$  3. P. 322–332.

- Yuan Y., Ju Y.S., Kim Y., Li J., Wang Y., Yoon C.J., Yang Y., Martincorena I., Creighton C.J., Weinstein J.N., Xu Y., Han L., Kim H.-L., Nakagawa H., Park K., Campbell P.J., Liang H., PCAWG Consortium. Comprehensive molecular characterization of mitochondrial genomes in human cancers // Nat. Genet. 2020. Vol. 52, № 3. P. 342–352.
- 5. Falkenberg M., Gustafsson C.M. Mammalian mitochondrial DNA replication and mechanisms of deletion formation // Crit. Rev. Biochem. Mol. Biol. 2020. Vol. 55, № 6. P. 509–524.
- 6. Wallace D.C. A mitochondrial paradigm of metabolic and degenerative diseases, aging, and cancer: a dawn for evolutionary medicine // Annu. Rev. Genet. 2005. Vol. 39. P. 359–407.
- Taylor R.W., Turnbull D.M. Mitochondrial DNA mutations in human disease // Nat. Rev. Genet. 2005. Vol. 6, № 5. P. 389–402.
- Wallace D.C., Chalkia D. Mitochondrial DNA genetics and the heteroplasmy conundrum in evolution and disease // Cold Spring Harb. Perspect. Biol. 2013. Vol. 5, № 11. P. a021220.
- Lehmann G., Budovsky A., Muradian K.K., Fraifeld V.E. Mitochondrial genome anatomy and species-specific lifespan // Rejuvenation Res. 2006. Vol. 9, № 2. P. 223–226.
- Belle E.M.S., Piganeau G., Gardner M., Eyre-Walker A. An investigation of the variation in the transition bias among various animal mitochondrial DNA // Gene. 2005. Vol. 355. P. 58–66.
- 11. Montooth K.L., Rand D.M. The spectrum of mitochondrial mutation differs across species // PLoS Biol. 2008. Vol. 6, № 8. P. e213.
- Ludwig L.S., Lareau C.A., Ulirsch J.C., Christian E., Muus C., Li L.H., Pelka K., Ge W., Oren Y., Brack A., Law T., Rodman C., Chen J.H., Boland G.M., Hacohen N., Rozenblatt-Rosen O., Aryee M.J., Buenrostro J.D., Regev A., Sankaran V.G. Lineage Tracing in Humans Enabled by Mitochondrial Mutations and Single-Cell Genomics // Cell. 2019. Vol. 176, № 6. P. 1325–1339.e22.
- Shen J., Khan N., Lewis L.D., Armand R., Grinberg O., Demidenko E., Swartz H. Oxygen consumption rates and oxygen concentration in molt-4 cells and their mtDNA depleted (rho0) mutants // Biophys. J. 2003. Vol. 84, № 2 Pt 1. P. 1291– 1298.
- 14. Slupphaug G., Kavli B., Krokan H.E. The interacting pathways for prevention and repair of oxidative DNA damage // Mutat. Res. 2003. Vol. 531, № 1-2. P. 231–251.
- Akbari M., Skjelbred C., Følling I., Sagen J., Krokan H.E. A gel electrophoresis method for detection of mitochondrial DNA mutation (3243 tRNALeu (UUR)) applied to a Norwegian family with diabetes mellitus and hearing loss // Scand. J. Clin. Lab. Invest. Taylor & Francis, 2004. Vol. 64, № 2. P. 86–92.

- 16. Muraresku C.C., McCormick E.M., Falk M.J. Mitochondrial Disease: Advances in clinical diagnosis, management, therapeutic development, and preventative strategies // Curr. Genet. Med. Rep. 2018. Vol. 6, № 2. P. 62–72.
- Galtier N., Nabholz B., Glémin S., Hurst G.D.D. Mitochondrial DNA as a marker of molecular diversity: a reappraisal // Mol. Ecol. 2009. Vol. 18, № 22. P. 4541–4550.
- 18. Ballard J.W.O., Whitlock M.C. The incomplete natural history of mitochondria // Mol. Ecol. 2004. Vol. 13, № 4. P. 729–744.
- Lareau C.A., Ludwig L.S., Muus C., Gohil S.H., Zhao T., Chiang Z., Pelka K., Verboon J.M., Luo W., Christian E., Rosebrock D., Getz G., Boland G.M., Chen F., Buenrostro J.D., Hacohen N., Wu C.J., Aryee M.J., Regev A., Sankaran V.G. Massively parallel single-cell mitochondrial DNA genotyping and chromatin profiling // Nat. Biotechnol. 2021. Vol. 39, № 4. P. 451–461.
- Ju Y.S., Alexandrov L.B., Gerstung M., Martincorena I., Nik-Zainal S., Ramakrishna M., Davies H.R., Papaemmanuil E., Gundem G., Shlien A., Bolli N., Behjati S., Tarpey P.S., Nangalia J., Massie C.E., Butler A.P., Teague J.W., Vassiliou G.S., Green A.R., Du M.-Q., Unnikrishnan A., Pimanda J.E., Teh B.T., Munshi N., Greaves M., Vyas P., El-Naggar A.K., Santarius T., Collins V.P., Grundy R., Taylor J.A., Hayes D.N., Malkin D., ICGC Breast Cancer Group, ICGC Chronic Myeloid Disorders Group, ICGC Prostate Cancer Group, Foster C.S., Warren A.Y., Whitaker H.C., Brewer D., Eeles R., Cooper C., Neal D., Visakorpi T., Isaacs W.B., Bova G.S., Flanagan A.M., Futreal P.A., Lynch A.G., Chinnery P.F., McDermott U., Stratton M.R., Campbell P.J. Origins and functional consequences of somatic mitochondrial DNA mutations in human cancer // Elife. 2014. Vol. 3.
- 21. Turrens J.F. Mitochondrial formation of reactive oxygen species // J. Physiol. 2003. Vol. 552, № Pt 2. P. 335–344.
- Muller F.L., Lustgarten M.S., Jang Y., Richardson A., Van Remmen H. Trends in oxidative aging theories // Free Radic. Biol. Med. 2007. Vol. 43, № 4. P. 477– 503.
- 23. Alexeyev M.F. Is there more to aging than mitochondrial DNA and reactive oxygen species? // FEBS J. 2009. Vol. 276, № 20. P. 5768–5787.
- 24. Pakendorf B., Stoneking M. Mitochondrial DNA and human evolution // Annu. Rev. Genomics Hum. Genet. 2005. Vol. 6. P. 165–183.
- Reyes A., Gissi C., Pesole G., Saccone C. Asymmetrical directional mutation pressure in the mitochondrial genome of mammals // Mol. Biol. Evol. 1998. Vol. 15, № 8. P. 957–966.
- Hulbert A.J., Pamplona R., Buffenstein R., Buttemer W.A. Life and death: metabolic rate, membrane composition, and life span of animals // Physiol. Rev. 2007. Vol. 87, № 4. P. 1175–1213.
- 27. Sohal R.S., Weindruch R. Oxidative stress, caloric restriction, and aging // Science. 1996. Vol. 273, № 5271. P. 59–63.
- Beckman K.B., Ames B.N. The free radical theory of aging matures // Physiol. Rev. 1998. Vol. 78, № 2. P. 547–581.

- 29. Richter C., Park J.W., Ames B.N. Normal oxidative damage to mitochondrial and nuclear DNA is extensive // Proc. Natl. Acad. Sci. U. S. A. 1988. Vol. 85, № 17. P. 6465–6467.
- 30. Van Houten B., Woshner V., Santos J.H. Role of mitochondrial DNA in toxic responses to oxidative stress // DNA Repair . 2006. Vol. 5, № 2. P. 145–152.
- Fraga C.G., Shigenaga M.K., Park J.W., Degan P., Ames B.N. Oxidative damage to DNA during aging: 8-hydroxy-2'-deoxyguanosine in rat organ DNA and urine // Proc. Natl. Acad. Sci. U. S. A. 1990. Vol. 87, № 12. P. 4533–4537.
- 32. Alexandrov L.B., Nik-Zainal S., Wedge D.C., Aparicio S.A.J.R., Behjati S., Biankin A.V., Bignell G.R., Bolli N., Borg A., Børresen-Dale A.-L., Boyault S., Burkhardt B., Butler A.P., Caldas C., Davies H.R., Desmedt C., Eils R., Eyfjörd J.E., Foekens J.A., Greaves M., Hosoda F., Hutter B., Ilicic T., Imbeaud S., Imielinski M., Jäger N., Jones D.T.W., Jones D., Knappskog S., Kool M., Lakhani S.R., López-Otín C., Martin S., Munshi N.C., Nakamura H., Northcott P.A., Pajic M., Papaemmanuil E., Paradiso A., Pearson J.V., Puente X.S., Raine K., Ramakrishna M., Richardson A.L., Richter J., Rosenstiel P., Schlesner M., Schumacher T.N., Span P.N., Teague J.W., Totoki Y., Tutt A.N.J., Valdés-Mas R., van Buuren M.M., van 't Veer L., Vincent-Salomon A., Waddell N., Yates L.R., Australian Pancreatic Cancer Genome Initiative, ICGC Breast Cancer Consortium, ICGC MMML-Seq Consortium, ICGC PedBrain, Zucman-Rossi J., Futreal P.A., McDermott U., Lichter P., Meyerson M., Grimmond S.M., Siebert R., Campo E., Shibata T., Pfister S.M., Campbell P.J., Stratton M.R. Signatures of mutational processes in human cancer // Nature. 2013. Vol. 500, № 7463. P. 415-421.
- Kucab J.E., Zou X., Morganella S., Joel M., Nanda A.S., Nagy E., Gomez C., Degasperi A., Harris R., Jackson S.P., Arlt V.M., Phillips D.H., Nik-Zainal S. A Compendium of Mutational Signatures of Environmental Agents // Cell. 2019. Vol. 177, № 4. P. 821–836.e16.
- 34. Kennedy S.R., Salk J.J., Schmitt M.W., Loeb L.A. Ultra-sensitive sequencing reveals an age-related increase in somatic mitochondrial mutations that are inconsistent with oxidative damage // PLoS Genet. 2013. Vol. 9, № 9. P. e1003794.
- Selman C., McLaren J.S., Himanka M.J., Speakman J.R. Effect of long-term cold exposure on antioxidant enzyme activities in a small mammal // Free Radic. Biol. Med. 2000. Vol. 28, № 8. P. 1279–1285.
- 36. Ku H.H., Sohal R.S. Comparison of mitochondrial pro-oxidant generation and anti-oxidant defenses between rat and pigeon: possible basis of variation in longevity and metabolic potential // Mech. Ageing Dev. 1993. Vol. 72, № 1. P. 67–76.
- Lewis K.N., Buffenstein R. Chapter 6 The Naked Mole-Rat: A Resilient Rodent Model of Aging, Longevity, and Healthspan // Handbook of the Biology of Aging (Eighth Edition) / ed. Kaeberlein M.R., Martin G.M. San Diego: Academic Press, 2016. P. 179–204.
- 38. Rong Z., Tu P., Xu P., Sun Y., Yu F., Tu N., Guo L., Yang Y. The

Mitochondrial Response to DNA Damage // Front Cell Dev Biol. 2021. Vol. 9. P. 669379.

- 39. Larsen N.B., Rasmussen M., Rasmussen L.J. Nuclear and mitochondrial DNA repair: similar pathways? // Mitochondrion. 2005. Vol. 5, № 2. P. 89–108.
- 40. Alexeyev M., Shokolenko I., Wilson G., LeDoux S. The maintenance of mitochondrial DNA integrity--critical analysis and update // Cold Spring Harb. Perspect. Biol. 2013. Vol. 5, № 5. P. a012641.
- 41. Bohr V.A., Stevnsner T., de Souza-Pinto N.C. Mitochondrial DNA repair of oxidative damage in mammalian cells // Gene. 2002. Vol. 286, № 1. P. 127–134.
- 42. Karahalil B., de Souza-Pinto N.C., Parsons J.L., Elder R.H., Bohr V.A. Compromised incision of oxidized pyrimidines in liver mitochondria of mice deficient in NTH1 and OGG1 glycosylases // J. Biol. Chem. 2003. Vol. 278, № 36. P. 33701–33707.
- 43. Hazra T.K., Kow Y.W., Hatahet Z., Imhoff B., Boldogh I., Mokkapati S.K., Mitra S., Izumi T. Identification and characterization of a novel human DNA glycosylase for repair of cytosine-derived lesions // J. Biol. Chem. 2002. Vol. 277, № 34. P. 30417–30420.
- 44. de Souza-Pinto N.C., Eide L., Hogue B.A., Thybo T., Stevnsner T., Seeberg E., Klungland A., Bohr V.A. Repair of 8-oxodeoxyguanosine lesions in mitochondrial dna depends on the oxoguanine dna glycosylase (OGG1) gene and 8-oxoguanine accumulates in the mitochondrial dna of OGG1-defective mice // Cancer Res. 2001. Vol. 61, № 14. P. 5378–5381.
- 45. Zsurka G., Peeva V., Kotlyar A., Kunz W.S. Is There Still Any Role for Oxidative Stress in Mitochondrial DNA-Dependent Aging? // Genes . 2018. Vol. 9, № 4.
- 46. Wang J., Xiong S., Xie C., Markesbery W.R., Lovell M.A. Increased oxidative damage in nuclear and mitochondrial DNA in Alzheimer's disease // J. Neurochem. 2005. Vol. 93, № 4. P. 953–962.
- 47. Hazra T.K., Das A., Das S., Choudhury S., Kow Y.W., Roy R. Oxidative DNA damage repair in mammalian cells: a new perspective // DNA Repair . 2007. Vol. 6, № 4. P. 470–480.
- 48. Hegde M.L., Hazra T.K., Mitra S. Early steps in the DNA base excision/singlestrand interruption repair pathway in mammalian cells // Cell Res. 2008. Vol. 18, № 1. P. 27–47.
- 49. Wallace S.S. Biological consequences of free radical-damaged DNA bases // Free Radic. Biol. Med. 2002. Vol. 33, № 1. P. 1–14.
- 50. Kamiya H. Mutagenic potentials of damaged nucleic acids produced by reactive oxygen/nitrogen species: approaches using synthetic oligonucleotides and nucleotides: survey and summary // Nucleic Acids Res. 2003. Vol. 31, № 2. P. 517–531.
- Kamiya H., Murata-Kamiya N., Koizume S., Inoue H., Nishimura S., Ohtsuka E. 8-Hydroxyguanine (7,8-dihydro-8-oxoguanine) in hot spots of the c-Ha-ras gene: effects of sequence contexts on mutation spectra // Carcinogenesis. 1995. Vol. 16, № 4. P. 883–889.

- 52. Sanchez-Contreras M., Sweetwyne M.T., Kohrn B.F., Tsantilas K.A., Hipp M.J., Schmidt E.K., Fredrickson J., Whitson J.A., Campbell M.D., Rabinovitch P.S., Marcinek D.J., Kennedy S.R. A replication-linked mutational gradient drives somatic mutation accumulation and influences germline polymorphisms and genome composition in mitochondrial DNA // Nucleic Acids Res. 2021. Vol. 49, № 19. P. 11103–11118.
- 53. Sanchez-Contreras M., Sweetwyne M.T., Tsantilas K.A., Whitson J.A., Campbell M.D., Kohrn B.F., Kim H.J., Hipp M.J., Fredrickson J., Nguyen M.M., Hurley J.B., Marcinek D.J., Rabinovitch P.S., Kennedy S.R. The multitissue landscape of somatic mtDNA mutations indicates tissue-specific accumulation and removal in aging // Elife. 2023. Vol. 12.
- 54. Gillooly J.F., Brown J.H., West G.B., Savage V.M., Charnov E.L. Effects of size and temperature on metabolic rate // Science. 2001. Vol. 293, № 5538. P. 2248–2251.
- 55. Timoféeff-Ressovsky N.W. Qualitativer Vergleich der Mutabilität vonDrosophila funebris undDrosophila melanogaster // Z. Indukt. Abstamm. Vererbungsl. 1936. Vol. 71, № 1. P. 276–280.
- 56. Gillooly J.F., Allen A.P., West G.B., Brown J.H. The rate of DNA evolution: effects of body size and temperature on the molecular clock // Proc. Natl. Acad. Sci. U. S. A. 2005. Vol. 102, № 1. P. 140–145.
- 57. Belfield E.J., Brown C., Ding Z.J., Chapman L., Luo M., Hinde E., van Es S.W., Johnson S., Ning Y., Zheng S.J., Mithani A., Harberd N.P. Thermal stress accelerates Arabidopsis thaliana mutation rate // Genome Res. 2021. Vol. 31, № 1. P. 40–50.
- 58. Matsuba C., Ostrow D.G., Salomon M.P., Tolani A., Baer C.F. Temperature, stress and spontaneous mutation in Caenorhabditis briggsae and Caenorhabditis elegans // Biol. Lett. 2013. Vol. 9, № 1. P. 20120334.
- 59. Waldvogel A.-M., Pfenninger M. Temperature dependence of spontaneous mutation rates // Genome Res. 2021. Vol. 31, № 9. P. 1582–1589.
- 60. Chu X.-L., Zhang B.-W., Zhang Q.-G., Zhu B.-R., Lin K., Zhang D.-Y. Temperature responses of mutation rate and mutational spectrum in an Escherichia coli strain and the correlation with metabolic rate // BMC Evol. Biol. 2018. Vol. 18, № 1. P. 126.
- 61. Shamanskiy V., Mikhailova A.A., Tretiakov E.O., Ushakova K., Mikhailova A.G., Oreshkov S., Knorre D.A., Ree N., Overdevest J.B., Lukowski S.W., Gostimskaya I., Yurov V., Liou C.-W., Lin T.-K., Kunz W.S., Reymond A., Mazunin I., Bazykin G.A., Fellay J., Tanaka M., Khrapko K., Gunbin K., Popadin K. Secondary structure of the human mitochondrial genome affects formation of deletions // BMC Biol. 2023. Vol. 21, № 1. P. 103.
- 62. Koh G., Degasperi A., Zou X., Momen S., Nik-Zainal S. Mutational signatures: emerging concepts, caveats and clinical applications // Nat. Rev. Cancer. 2021. Vol. 21, № 10. P. 619–637.
- 63. Zou X., Koh G.C.C., Nanda A.S., Degasperi A., Urgo K., Roumeliotis T.I., Agu C.A., Badja C., Momen S., Young J., Amarante T.D., Side L., Brice G., Perez-

Alonso V., Rueda D., Gomez C., Bushell W., Harris R., Choudhary J.S., Genomics England Research Consortium, Jiricny J., Skarnes W.C., Nik-Zainal S. A systematic CRISPR screen defines mutational mechanisms underpinning signatures caused by replication errors and endogenous DNA damage // Nat Cancer. 2021. Vol. 2, № 6. P. 643–657.

- 64. Harris K., Pritchard J.K. Rapid evolution of the human mutation spectrum // Elife. 2017. Vol. 6.
- Moorjani P., Amorim C.E.G., Arndt P.F., Przeworski M. Variation in the molecular clock of primates // Proc. Natl. Acad. Sci. U. S. A. 2016. Vol. 113, № 38. P. 10607–10612.
- 66. Raina S.Z., Faith J.J., Disotell T.R., Seligmann H., Stewart C.-B., Pollock D.D. Evolution of base-substitution gradients in primate mitochondrial genomes // Genome Res. 2005. Vol. 15, № 5. P. 665–673.
- 67. Krishnan N.M., Seligmann H., Raina S.Z., Pollock D.D. Detecting gradients of asymmetry in site-specific substitutions in mitochondrial genomes // DNA Cell Biol. 2004. Vol. 23, № 10. P. 707–714.
- 68. Baker K.T., Nachmanson D., Kumar S., Emond M.J., Ussakli C., Brentnall T.A., Kennedy S.R., Risques R.A. Mitochondrial DNA Mutations are Associated with Ulcerative Colitis Preneoplasia but Tend to be Negatively Selected in Cancer // Mol. Cancer Res. 2019. Vol. 17, № 2. P. 488–498.
- 69. Hoekstra J.G., Hipp M.J., Montine T.J., Kennedy S.R. Mitochondrial DNA mutations increase in early stage Alzheimer disease and are inconsistent with oxidative damage // Ann. Neurol. 2016. Vol. 80, № 2. P. 301–306.
- 70. Efimenko B., Popadin K., Gunbin K. NeMu: a comprehensive pipeline for accurate reconstruction of neutral mutation spectra from evolutionary data // Nucleic Acids Res. 2024.
- Di Marco M., Pacifici M., Santini L., Baisero D., Francucci L., Grottolo Marasini G., Visconti P., Rondinini C. Generation length for mammals // Nat. Conserv. Pensoft Publishers, 2013. Vol. 5. P. 89–94.
- 72. Froese R., Pauly D. FishBase 2000: Concepts Designs and Data Sources. WorldFish, 2000. 344 p.
- 73. Budovsky A., Craig T., Wang J., Tacutu R., Csordas A., Lourenço J., Fraifeld V.E., de Magalhães J.P. LongevityMap: a database of human genetic variants associated with longevity // Trends Genet. 2013. Vol. 29, № 10. P. 559–560.
- 74. McNab B.K. An Analysis of the Body Temperatures of Birds // Condor. Oxford Academic, 1966. Vol. 68, № 1. P. 47–55.
- 75. Faith J.J., Pollock D.D. Likelihood analysis of asymmetrical mutation bias gradients in vertebrate mitochondrial genomes // Genetics. 2003. Vol. 165, № 2. P. 735–745.
- 76. Tanaka M., Ozawa T. Strand asymmetry in human mitochondrial DNA mutations // Genomics. 1994. Vol. 22, № 2. P. 327–335.
- Polishchuk L.V., Tseitlin V.B. Scaling of Population Density on Body Mass and a Number-Size Trade-Off // Oikos. [Nordic Society Oikos, Wiley], 1999. Vol. 86, № 3. P. 544–556.

- 78. Damuth J. Population density and body size in mammals // Nature. Springer Science and Business Media LLC, 1981. Vol. 290, № 5808. P. 699–700.
- 79. White C.R., Seymour R.S. Allometric scaling of mammalian metabolism // J. Exp. Biol. 2005. Vol. 208, № Pt 9. P. 1611–1619.
- 80. Arbeithuber B., Hester J., Cremona M.A., Stoler N., Zaidi A., Higgins B., Anthony K., Chiaromonte F., Diaz F.J., Makova K.D. Age-related accumulation of de novo mitochondrial mutations in mammalian oocytes and somatic tissues // PLoS Biol. 2020. Vol. 18, № 7. P. e3000745.
- Rebolledo-Jaramillo B., Su M.S.-W., Stoler N., McElhoe J.A., Dickins B., Blankenberg D., Korneliussen T.S., Chiaromonte F., Nielsen R., Holland M.M., Paul I.M., Nekrutenko A., Makova K.D. Maternal age effect and severe germline bottleneck in the inheritance of human mitochondrial DNA // Proc. Natl. Acad. Sci. U. S. A. 2014. Vol. 111, № 43. P. 15474–15479.
- 82. Wei W., Tuna S., Keogh M.J., Smith K.R., Aitman T.J., Beales P.L., Bennett D.L., Gale D.P., Bitner-Glindzicz M.A.K., Black G.C., Brennan P., Elliott P., Flinter F.A., Floto R.A., Houlden H., Irving M., Koziell A., Maher E.R., Markus H.S., Morrell N.W., Newman W.G., Roberts I., Sayer J.A., Smith K.G.C., Taylor J.C., Watkins H., Webster A.R., Wilkie A.O.M., Williamson C., NIHR BioResource–Rare Diseases, 100,000 Genomes Project–Rare Diseases Pilot, Ashford S., Penkett C.J., Stirrups K.E., Rendon A., Ouwehand W.H., Bradley J.R., Raymond F.L., Caulfield M., Turro E., Chinnery P.F. Germline selection shapes human mitochondrial DNA diversity // Science. 2019. Vol. 364, № 6442.
- Von Stetina J.R., Orr-Weaver T.L. Developmental control of oocyte maturation and egg activation in metazoan models // Cold Spring Harb. Perspect. Biol. 2011. Vol. 3, № 10. P. a005553.
- 84. Sato K., Sato M. Multiple ways to prevent transmission of paternal mitochondrial DNA for maternal inheritance in animals // J. Biochem. 2017. Vol. 162, № 4. P. 247–253.
- 85. Tacutu R., Craig T., Budovsky A., Wuttke D., Lehmann G., Taranukha D., Costa J., Fraifeld V.E., de Magalhães J.P. Human Ageing Genomic Resources: integrated databases and tools for the biology and genetics of ageing // Nucleic Acids Res. 2013. Vol. 41, № Database issue. P. D1027–D1033.
- 86. Peters R.H., Peters R.H. The Ecological Implications of Body Size. Cambridge University Press, 1986. 329 p.
- 87. Damuth J. Interspecific allometry of population density in mammals and other animals: the independence of body mass and population energy-use // Biol. J. Linn. Soc. Lond. Oxford Academic, 2008. Vol. 31, № 3. P. 193–246.
- 88. Hebert P.D.N., Cywinska A., Ball S.L., deWaard J.R. Biological identifications through DNA barcodes // Proc. Biol. Sci. 2003. Vol. 270, № 1512. P. 313–321.
- Chen L., Liu P., Evans T.C. Jr, Ettwiller L.M. DNA damage is a pervasive cause of sequencing errors, directly confounding variant identification // Science. 2017. Vol. 355, № 6326. P. 752–756.
- 90. Chen L., Liu P., Evans T.C. Jr, Ettwiller L.M. Response to Comment on "DNA damage is a pervasive cause of sequencing errors, directly confounding variant

identification" // Science. 2018. Vol. 361, № 6409.

- 91. Stewart C., Leshchiner I., Hess J., Getz G. Comment on "DNA damage is a pervasive cause of sequencing errors, directly confounding variant identification" // Science. 2018. Vol. 361, № 6409.
- Keil G., Cummings E., de Magalhães J.P. Being cool: how body temperature influences ageing and longevity // Biogerontology. 2015. Vol. 16, № 4. P. 383– 397.
- 93. Tomasetti C., Vogelstein B. Cancer etiology. Variation in cancer risk among tissues can be explained by the number of stem cell divisions // Science. 2015. Vol. 347, № 6217. P. 78–81.
- 94. Seim I., Ma S., Gladyshev V.N. Gene expression signatures of human cell and tissue longevity // NPJ Aging Mech Dis. 2016. Vol. 2. P. 16014.
- 95. Wang C., Ross W.T., Mysorekar I.U. Urothelial generation and regeneration in development, injury, and cancer // Dev. Dyn. 2017. Vol. 246, № 4. P. 336–343.
- 96. Ericson N.G., Kulawiec M., Vermulst M., Sheahan K., O'Sullivan J., Salk J.J., Bielas J.H. Decreased mitochondrial DNA mutagenesis in human colorectal cancer // PLoS Genet. 2012. Vol. 8, № 6. P. e1002689.
- 97. Rosario S.R., Long M.D., Affronti H.C., Rowsam A.M., Eng K.H., Smiraglia D.J. Pan-cancer analysis of transcriptional metabolic dysregulation using The Cancer Genome Atlas // Nat. Commun. 2018. Vol. 9, № 1. P. 5330.
- 98. Bhandari V., Li C.H., Bristow R.G., Boutros P.C., PCAWG Consortium. Divergent mutational processes distinguish hypoxic and normoxic tumours // Nat. Commun. 2020. Vol. 11, № 1. P. 737.
- 99. Atkinson Q.D., Gray R.D., Drummond A.J. mtDNA variation predicts population size in humans and reveals a major Southern Asian chapter in human prehistory // Mol. Biol. Evol. 2008. Vol. 25, № 2. P. 468–474.
- 100. Uddin A., Chakraborty S. Synonymous codon usage pattern in mitochondrial CYB gene in pisces, aves, and mammals // Mitochondrial DNA A DNA Mapp Seq Anal. 2017. Vol. 28, № 2. P. 187–196.
- 101. Saini N., Gordenin D.A. Hypermutation in single-stranded DNA // DNA Repair . 2020. Vol. 91-92. P. 102868.
- 102. Gomes A.P., Price N.L., Ling A.J.Y., Moslehi J.J., Montgomery M.K., Rajman L., White J.P., Teodoro J.S., Wrann C.D., Hubbard B.P., Mercken E.M., Palmeira C.M., de Cabo R., Rolo A.P., Turner N., Bell E.L., Sinclair D.A. Declining NAD+ Induces a Pseudohypoxic State Disrupting Nuclear-Mitochondrial Communication during Aging // Cell. Elsevier, 2013. Vol. 155, № 7. P. 1624–1638.
- 103. Bellanti F., Romano A.D., Giudetti A.M., Rollo T., Blonda M., Tamborra R., Vendemiale G., Serviddio G. Many faces of mitochondrial uncoupling during age: damage or defense? // J. Gerontol. A Biol. Sci. Med. Sci. 2013. Vol. 68, № 8. P. 892–902.
- 104. Stadtman E.R. Protein oxidation and aging // Free Radic. Res. 2006. Vol. 40, № 12. P. 1250–1258.
- 105. Ademowo O.S., Dias H.K.I., Burton D.G.A., Griffiths H.R. Lipid (per)

oxidation in mitochondria: an emerging target in the ageing process? // Biogerontology. 2017. Vol. 18, № 6. P. 859–879.

- 106. Schriner S.E., Linford N.J., Martin G.M., Treuting P., Ogburn C.E., Emond M., Coskun P.E., Ladiges W., Wolf N., Van Remmen H., Wallace D.C., Rabinovitch P.S. Extension of murine life span by overexpression of catalase targeted to mitochondria // Science. 2005. Vol. 308, № 5730. P. 1909–1911.
- 107. Shewaramani S., Finn T.J., Leahy S.C., Kassen R., Rainey P.B., Moon C.D. Anaerobically Grown Escherichia coli Has an Enhanced Mutation Rate and Distinct Mutational Spectra // PLoS Genet. 2017. Vol. 13, № 1. P. e1006570.
- 108. Earley M.C., Crouse G.F. The role of mismatch repair in the prevention of base pair mutations in Saccharomyces cerevisiae // Proc. Natl. Acad. Sci. U. S. A. 1998. Vol. 95, № 26. P. 15487–15491.
- 109. Shin C.Y., Turker M.S. A:T --> G:C base pair substitutions occur at a higher rate than other substitution events in Pms2 deficient mouse cells // DNA Repair . 2002. Vol. 1, № 12. P. 995–1001.
- 110. Bridge G., Rashid S., Martin S.A. DNA mismatch repair and oxidative DNA damage: implications for cancer biology and treatment // Cancers . 2014. Vol. 6, № 3. P. 1597–1614.
- 111. Seplyarskiy V.B., Akkuratov E.E., Akkuratova N., Andrianova M.A., Nikolaev S.I., Bazykin G.A., Adameyko I., Sunyaev S.R. Error-prone bypass of DNA lesions during lagging-strand replication is a common source of germline and cancer mutations // Nat. Genet. 2019. Vol. 51, № 1. P. 36–41.
- 112. Martin A.P., Palumbi S.R. Body size, metabolic rate, generation time, and the molecular clock // Proc. Natl. Acad. Sci. U. S. A. 1993. Vol. 90, № 9. P. 4087– 4091.
- 113. Naya H., Romero H., Zavala A., Alvarez B., Musto H. Aerobiosis increases the genomic guanine plus cytosine content (GC%) in prokaryotes // J. Mol. Evol. 2002. Vol. 55, № 3. P. 260–264.
- 114. Romero H., Pereira E., Naya H., Musto H. Oxygen and guanine-cytosine profiles in marine environments // J. Mol. Evol. 2009. Vol. 69, № 2. P. 203–206.
- 115. Aslam S., Lan X.-R., Zhang B.-W., Chen Z.-L., Wang L., Niu D.-K. Aerobic prokaryotes do not have higher GC contents than anaerobic prokaryotes, but obligate aerobic prokaryotes have // BMC Evol. Biol. 2019. Vol. 19, № 1. P. 35.
- 116. Popadin K., Polishchuk L.V., Mamirova L., Knorre D., Gunbin K. Accumulation of slightly deleterious mutations in mitochondrial protein-coding genes of large versus small mammals // Proc. Natl. Acad. Sci. U. S. A. 2007. Vol. 104, № 33. P. 13390–13395.
- 117. Nikolaev S.I., Montoya-Burgos J.I., Popadin K., Parand L., Margulies E.H., National Institutes of Health Intramural Sequencing Center Comparative Sequencing Program, Antonarakis S.E. Life-history traits drive the evolutionary rates of mammalian coding and noncoding genomic elements // Proc. Natl. Acad. Sci. U. S. A. 2007. Vol. 104, № 51. P. 20443–20448.
- 118. Popadin K.Y., Nikolaev S.I., Junier T., Baranova M., Antonarakis S.E. Purifying selection in mammalian mitochondrial protein-coding genes is highly effective

and congruent with evolution of nuclear genes // Mol. Biol. Evol. 2013. Vol. 30, № 2. P. 347–355.

- 119. Alseth I., Dalhus B., Bjørås M. Inosine in DNA and RNA // Curr. Opin. Genet. Dev. 2014. Vol. 26. P. 116–123.
- 120. Pang B., Zhou X., Yu H., Dong M., Taghizadeh K., Wishnok J.S., Tannenbaum S.R., Dedon P.C. Lipid peroxidation dominates the chemistry of DNA adduct formation in a mouse model of inflammation // Carcinogenesis. 2007. Vol. 28, № 8. P. 1807–1813.
- 121. Degtyareva N.P., Saini N., Sterling J.F., Placentra V.C., Klimczak L.J., Gordenin D.A., Doetsch P.W. Mutational signatures of redox stress in yeast single-strand DNA and of aging in human mitochondrial DNA share a common feature // PLoS Biol. 2019. Vol. 17, № 5. P. e3000263.
- 122. Koh C.W.Q., Goh Y.T., Toh J.D.W., Neo S.P., Ng S.B., Gunaratne J., Gao Y.-G., Quake S.R., Burkholder W.F., Goh W.S.S. Single-nucleotide-resolution sequencing of human N6-methyldeoxyadenosine reveals strand-asymmetric clusters associated with SSBP1 on the mitochondrial genome // Nucleic Acids Res. 2018. Vol. 46, № 22. P. 11659–11670.
- 123. Hao Z., Wu T., Cui X., Zhu P., Tan C., Dou X., Hsu K.-W., Lin Y.-T., Peng P.-H., Zhang L.-S., Gao Y., Hu L., Sun H.-L., Zhu A., Liu J., Wu K.-J., He C. N6-Deoxyadenosine Methylation in Mammalian Mitochondrial DNA // Mol. Cell. 2020. Vol. 78, № 3. P. 382–395.e8.

# СПИСОК РАБОТ, ОПУБЛИКОВАННЫХ ПО ТЕМЕ ДИССЕРТАЦИИ

# Статьи в научных журналах и изданиях, входящих в перечень рецензируемых научных изданий, рекомендованных ВАК МинОбрНауки РФ для опубликования основных результатов диссертационных работ:

- Mikhailova A. G. [и др.]. A mitochondria-specific mutational signature of aging: increased rate of A > G substitutions on the heavy strand // Nucleic acids research. 2022. № 18 (50). P. 10264–10277.
- Shamanskiy V., Mikhailova A.A., Tretiakov E.O., Ushakova K., Mikhailova A. G., Oreshkov S. [и др.]. Secondary structure of the human mitochondrial genome affects formation of deletions // BMC biology. 2023. № 1 (21). Р. 103.
- Гончаров А.Г., Попадьин К.Ю., Козенков И.И., Михайлова А.Г., Лобанова В.В., Татаркина М.А., Ильющенко Д.В., Реверчук И.В., Тынтерова А.М., Двирский А.А., Джигкаев А.Х., Тимофеева Е.В., Гунбин К.В. Methodological approaches to the study of the determinants of somatic mitochondrial heteroplasmy in the elderly // Journal of Ural Medical Academic Science. 2022. № 3 (19). С. 322–332.

#### Препринты статей:

- Mikhaylova A. G. [и др.]. Mammalian mitochondrial mutational spectrum as a hallmark of cellular and organismal aging // bioRxiv. 2021. doi.org/10.1101/589168
- Mikhailova A. G. [и др.]. A>G substitutions on a heavy chain of mitochondrial genome marks an increased level of aerobic metabolism in warm versus cold vertebrates // bioRxiv. 2023. doi.org/10.1101/2020.07.25.221184
- Iliushchenko D., Efimenko B., Mikhailova A.G., Shamanskiy V., Saparbaev M.K., Mazunin I., Knorre D., Kunz W.S., Kapranov P., Denisov S., Fellay J., Khrapko K., Gunbin K., Popadin K. Mitochondrial mutation spectrum in Chordates: damage versus replication signatures, causes, and dynamics // bioRxiv. 2024. doi.org/10.1101/2023.12.08.570826

#### Тезисы докладов в материалах конференций:

- Mikhailova A. G. [и др.]. Deamination of adenine to guanine is a signature of oxidative damage in mitochondrial DNA // Abstracts from the 55th European Society of Human Genetics (ESHG) Conference: Hybrid Posters. Eur J Hum Genet 31 (Suppl 1). 2023. P. 345–709.
- Mikhailova A. G. [и др.]. Mitochondrial mutational spectrum in human cancers is sensitive to cellular hypoxia // Abstracts from the 54th European Society of Human Genetics (ESHG) Conference: e-Posters. Eur J Hum Genet 30 (Suppl 1). 2022. P. 88–608.
- 3. Гусаров Ю.С., Михайлова А.Г., Орешков С.С., Ефименко Б.Э., Гунбин К.В., Бурская В.О., Попадьин К.Ю. Усиление метаболизма у птиц с различными типами питания и экозонами провоцирует мутагенез мтднк вследствие урона от АФК // Сборник тезисов конференции ХИМБИОSEASONS 2022. 2022. С. 15.
- Shamanskiy V., Mikhailova A. A., Ushakova K., Mikhailova A. G., Oreshkov S., Knorre D., Tretiakov E. O. [и др.]. // Abstracts from the 54th European Society of Human Genetics (ESHG) Conference: e-Posters. Eur J Hum Genet 30 (Suppl 1). 2022. P. 88–608.
- Mikhaylova A.G. [и др.]. Mitochondrial mutational spectrum in poikilothermic versus homeothermic vertebrates: effects of the temperature // Bioinformatics of genome regulation and structure/systems biology (BGRS/SB-2020). The Twelfth International Multiconference Abstracts. 2020. P. 224-225.
- 6. Ushakova K., Shamanskiy V., Mikhailova A.A., Mikhailova A.G., Tretiakov E., Mazunin I., Popadin K., Gunbin K. A global human mitochondrial tree as a resource for population and evolutionary studies // Bioinformatics of genome regulation and structure/systems biology (BGRS/SB-2020). The Twelfth International Multiconference Abstracts. 2020. P. 263.
- 7. Mikhaylova A.G. [и др.]. Transition transversion ratio in mtdna is higher in long-versus short-lived mammalians: effects of ros and replication? //

Biodiversity: Genomics and Evolution (BioGenEvo-2018). Symposium. Abstracts. 2018. P. 29.

- Ushakova K., Mikhaylova A.A., Mikhaylova A.G., Knorre D., Mazunin I., Reymond A., Gunbin K., Popadin K. Variability in Gibbs energy of trna molecules in mitochondrial genomes of chordates: neutral selection or evolution towards optimization of translation? // Systems Biology and Bioinformatics (SBB-2018). The tenth International young Scientists School. Abstracts. 2018. P. 45.
- Mikhaylova A.A., Mikhaylova A.G., Ushakova K., Knorre D., Mazunin I., Reymond A., Gunbin K., Popadin K. Selfish elements drive mitochondrial and nuclear genome size in opposite directions // Bioinformatics of Genome Regulation and Structure\Systems Biology (BGRS\SB-2018). The Eleventh International Conference. 2018. P. 61.
- Sokol A., Ushakova K., Mikhaylova A.A., Mikhaylova A.G., Knorre D., Mazunin I., Reymond A., Gunbin K., Popadin K. HSP as a long-term buffer of the genome-wide mutation burden // Bioinformatics of Genome Regulation and Structure\Systems Biology (BGRS\SB-2018). The Eleventh International Conference. 2018. P. 81.

#### Доклады на конференциях и симпозиумах (без публикации тезисов):

- Михайлова А. [и др.]. Митохондриальные мутационные спектры: от эволюции позвоночных к онкологии и медицине [Электронный pecypc]: Proceedings of 11th Moscow Conference on Computational Molecular Biology MCCMB'23, August 3 - 6, 2023 Москва, Россия. М., ИППИ РАН, 2023.1 электрон. опт. диск (CD-ROM).
- 2. Гусаров Ю., Михайлова А., Ефименко Б., Ильющенко Д., Гунбин К., Бурская В., Попадьин К. Мутация Ah->Gh митохондриальной ДНК птиц ассоциирована со способностью к полету и дайвингу [Электронный pecypc]: Proceedings of 11th Moscow Conference on Computational Molecular

Biology MCCMB'23, August 3 - 6, 2023 Москва, Россия. М., ИППИ РАН, 2023.1 электрон. опт. диск (CD-ROM).

- 3. Кошель А., Осадчий Г., Михайлова А., Ефименко Б., Гунбин К., Попадьин К. Различия мутационного спектра мтДНК термитных и не-термитных тараканов ассоциированы с продолжительностью жизни [Электронный pecypc]: Proceedings of 11th Moscow Conference on Computational Molecular Biology MCCMB'23, August 3 6, 2023 Москва, Россия. М., ИППИ РАН, 2023.1 электрон. опт. диск (CD-ROM).
- Mikhailova A. G. [и др.]. A billion-year trend of amino acid substitutions in the mitochondrial genome [Электронный ресурс] // SMBE2023: abstracts book. URL:

<u>https://www.smbe2023.org/\_files/ugd/bccffb\_7ecd1df08896478aa99d806cb17</u> <u>d372c.pdf</u> (дата обращения: 03.08.2023)

- 5. Efimenko B., Mikhailova A. G., Popadin K. Mitochondria-specific signature of oxidative damage in human tissues: an excess of A>G on a heavy chain of mitochondrial genome in normoxic tissues [Электронный ресурс] // SMBE2023: abstracts book. URL: https://www.smbe2023.org/\_files/ugd/bccffb\_7ecd1df08896478aa99d806cb17 d372c.pdf (дата обращения: 03.08.2023)
- Mikhailova A. G. [и др.]. Deamination of adenine to guanine is a signature of oxidative damage in mitochondrial DNA [Электронный ресурс] // Mitochondria 2022 Workshop: On-line posters program. URL: https://program.eventact.com/Program/ivoYAAA/iiDY/en (дата обращения: 03.08.2023)
- Mikhailova A. G. [и др.]. Mitochondrial mutational spectrum in human cancers is sensitive to cellular hypoxia. [Электронный pecypc] // SMBEv2021: Abstract Book. URL: https://www.smbe.org/smbe/Portals/0/SMBEv2021%20-%20Updated%20Abstacts%2007072021.pdf (дата обращения: 03.08.2023)
- 8. Mikhailova A. G. [и др.]. A billion-year trend of amino acid substitutions in the mitochondrial genome [Электронный ресурс] // Proceedings of 10th Moscow

Conference on Computational Molecular Biology MCCMB'21, July 30th — August 2nd, 2021 Москва, Россия. М., ИППИ РАН, 2021. 1 электрон. опт. диск (CD-ROM).

- Galieva A., Mikhailova A. A., Mikhailova A. G., Shamanskiy V., Lobanova V., Ushakova K., Gunbin K., Popadin K. Asymmetrical mutagenesis drives aminoacid composition of the human mitochondrial genome [Электронный pecypc] // Proceedings of 10th Moscow Conference on Computational Molecular Biology MCCMB'21, July 30th — August 2nd, 2021 Москва, Россия. М., ИППИ РАН, 2021. 1 электрон. опт. диск (CD-ROM).
- 10.Gusarov Y., Mikhailova A. G., Oreshkov S., Mikhailova A. A., Knorre D., Polishchuk L. V., Kuptsov A., Gunbin K., BurskayaV., Popadin K. Mitochondrial mutational spectrum in birds: evidence of increased oxidative damage in species with high level of metabolism [Электронный ресурс] // Proceedings of 10th Moscow Conference on Computational Molecular Biology MCCMB'21, July 30th — August 2nd, 2021 Москва, Россия. М., ИППИ РАН, 2021. 1 электрон. опт. диск (CD-ROM).
- 11.Khaibulin E., Mikhailova A. G., Shamanskiy V., Popadin K. mtDNA codon usage of Chordata is primarily shaped by mutational spectra [Электронный pecypc] // Proceedings of 10th Moscow Conference on Computational Molecular Biology MCCMB'21, July 30th — August 2nd, 2021 Москва, Россия. М., ИППИ РАН, 2021. 1 электрон. опт. диск (CD-ROM).
- 12.Mikhailova A. G. [и др.] Mitochondrial mutational spectrum in vertebrates is shaped by temperature and generation time. [Электронный pecypc] // Proceedings of 9th Moscow Conference on Computational Molecular Biology MCCMB'19, July 27-30, 2019 Москва, Россия. М., ИППИ РАН. 2019. 1 электрон. опт. диск (CD-ROM).
- 13.Mikhailova A. A., Ushakova K., Mikhaylova A. G., Lobanova V., Mazunin I., Kravchenko P., Knorre D., Reymond A., Gunbin K., Popadin K. Tandem repeats are selfish elements which mark the level of hidden recombination in animal mitochondrial genomes [Электронный ресурс] // Proceedings of 9th Moscow

Conference on Computational Molecular Biology MCCMB'19, July 27-30, 2019 Москва, Россия. М., ИППИ РАН. 2019. 1 электрон. опт. диск (CD-ROM).

- 14. Mikhailova A. G. [и др.] Transition transversion ratio in mitochondrial genome is higher in long- versus short-lived mammalian species: effects of ROS and replication timing? [Электронный ресурс] // SMBE2018: Poster Group B. URL: <u>https://evolgen.biol.se.tmu.ac.jp/SMBE2018/POB.pdf</u> (дата обращения: 03.08.2023)
- 15.Ushakova K., Mikhailova A. A., Mikhailova A. G., Knorre D., Mazunin I., Reymond A., Gunbin G., Popadin K. Variability in Gibbs energy of tRNA molecules in mitochondrial genomes of Chordates: neutral selection or evolution towards optimization of translation? [Электронный ресурс] // SMBE2018: Poster Group B. URL: https://evolgen.biol.se.tmu.ac.jp/SMBE2018/POB.pdf (дата обращения: 03.08.2023)

## БЛАГОДАРНОСТИ

Выражаю искреннюю благодарность своему научному руководителю Попадьину Константину Юрьевичу за безусловную веру, вдохновение и удивительные дискуссии (и рекомендацию песни "100 лет одиночества").

Благодарю весь коллектив авторов рукописей, в особенности Гунбина Константина Владимировича и Макеева Всеволода Юрьевича за консультирование и неотъемлемый вклад в проект.

А также хочу поблагодарить лабораторию "Центр геномных исследований БФУ им. Канта" и всех своих коллег за дружескую поддержку и сотрудничество.

Наконец, особую благодарность хочу выразить своей семье, без которой моя научная деятельность не представилась бы возможной.