

Химический факультет

На правах рукописи



Чугунова Анастасия Андреевна

**Митохондриальный пептид L116 является
посредником между клеточным дыханием и
липидным метаболизмом**

Специальность 02.00.10 – биоорганическая химия

**Диссертация на соискание ученой степени
кандидата химических наук**

Научные руководители:

д. х. н., профессор,
академик РАН **Донцова О.А.**

д. х. н., профессор РАН **Сергиев П.В.**

Москва – 2019

Оглавление

ОГЛАВЛЕНИЕ	2
СПИСОК ИСПОЛЪЗУЕМЫХ СОКРАЩЕНИЙ	4
ВВЕДЕНИЕ	6
ОБЗОР ЛИТЕРАТУРЫ.....	11
Методы детекции КОРС-пептидов	11
<i>Вычислительные методы</i>	<i>11</i>
<i>Рибосомное профилирование</i>	<i>18</i>
<i>Масс-спектрометрия</i>	<i>24</i>
Различные виды КОРС.....	28
5'-НТО	29
Кодирующая область мРНК.....	31
3'-НТО.....	32
Длинные некодирующие РНК.....	32
Циркулярные РНК.....	36
Пре-микроРНК.....	36
Рибосомная РНК.....	37
Методы исследования и характеристики КОРС-пептидов.....	37
Проверка кодирующего потенциала КОРС	38
Поиск биологической функции КОРС-пептида.....	41
ЗАКЛЮЧЕНИЕ	45
МАТЕРИАЛЫ И МЕТОДЫ.....	46
Реактивы и биопрепараты	46
Буферы и растворы	47
Праймеры, использованные для клонирования и секвенирования.....	50
Методики, использованные в работе.....	52
ПЦР	52
Приготовление компетентных клеток.....	54
Трансформация компетентных клеток.....	54
Выделение плазмидной ДНК.....	54
Разделение фрагментов ДНК в агарозном геле	56
Выделение фрагментов ДНК из агарозного геля.....	56
Секвенирование	56
Клонирование гидовых РНК в рх458.....	56
Клонирование кДНК Cyb5r3 в вектор pUC19.....	58
Клонирование конструкций pSBtet_L116, pSBtet_HA116, pSBtet_L116mCherry.....	59
Клонирование конструкций pSBtet_Cyb5r3, pSBtet_Cyb5r3HA, pSBtet_Cyb5r3eGFP.....	60
Клонирование матрицы для гомологичной рекомбинации для замены Gly2 на Ala2 Cyb5r3	62
Проверка эффективности CRISPR-Cas9 с помощью T7-эндонуклеазы	64
Белковый электрофорез по Леммли	65
Окрашивание белковых гелей нитратом серебра.....	65
Трипсинолиз белковых образцов для анализа методом MALDI-MS.....	65
Вестерн-блоттинг.....	66
Культивирование эукариотических клеток	67
Получение модифицированных клеточных линий с помощью CRISPR-Cas9 системы	67
Получение модифицированных клеточных линий с помощью транспозиции.....	68
Проверка моноклонов.....	68
Иммунопреципитация.....	69

<i>Иммуноцитохимия и конфокальная микроскопия.....</i>	<i>69</i>
<i>Измерение клеточного дыхания</i>	<i>71</i>
РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ.....	72
Постановка задачи.....	72
Доказательство экспрессии и определение локализации L116 в клетках	76
Создание клеток с инактивированным геном 1500011k16R1k	78
Изучение фенотипа клеток с инактивированным 1500011k16R1k	80
Поиск белковых партнеров L116	85
Изучение фенотипа клеточной линии с нарушением митохондриальной локализации Cyb5r3	87
Изучение функциональности взаимодействия L116 и Cyb5r3	88
Изучение влияния инактивации L116 на Cyb5r3	89
Изучение влияния инактивации Cyb5r3 на L116	91
Изучение окислительно-восстановительного баланса клетки	93
Изучение липидного состава нокаутных клеток.....	94
Обсуждение результатов.....	98
ВЫВОДЫ	102
СПИСОК ЛИТЕРАТУРЫ	103

Список используемых сокращений

НА - Human influenza hemagglutinin, гемагглютинин человеческого гриппа

мРНК - Матричная РНК

НТО - Нетранслируемая область

ОРС - Открытая рамка считывания

кОРС – Короткая открытая рамка считывания

рРНК - Рибосомальная РНК

днРНК – Длинная некодирующая РНК

гРНК – Гидовая РНК

CRISPR - Clustered regularly interspaced short palindromic repeats, короткие палиндромные повторы, регулярно расположенные группами

РФ – Рибосомные футпринты

НАД – Никотинамидадениндинуклеотид

ROS – Reactive oxygen species, реактивные формы кислорода

TAG – Triglyceride, триглицериды

PC – Phosphatidylcholine, фосфатидилхолин

ак – Аминокислоты

нк – Нуклеотиды

ДМСО – Диметилсульфоксид

ЭДТА - Этилендиаминтетрауксусная кислота

ТЕМЕД - N,N,N',N'-тетраметилэтилендиамид

SDS - Sodium dodecyl sulfate, додецилсульфат натрия

ПААГ - Полиакриламидный гель

Введение

Актуальность проблемы

Пептиды и маленькие белки – это класс биомолекул, играющих важную роль в клетке [1,2]. Например, существуют пептиды и маленькие белки, регулирующие сон (орексин) [3,4], стресс (CFR) [5], метаболизм (лептин) [6] и другие процессы [2]. Поиск новых биологически активных пептидов привел к открытию огромного числа ранее неаннотированных пептидов, кодируемых короткими открытыми рамками считывания. Подобные пептиды, синтезируемые рибосомой сразу короткими, следует отличать от известных нейропептидов и гормонов, которые протеолитически вырезаются из более длинного предшественника. Последние исследования показали, что кОРС–пептиды являются важными регуляторами во многих жизненных процессах, таких как метаболизм [7,8], эндоцитоз [9], развитие [10,11] и апоптоз [12].

Степень разработанности темы

Хотя исследования в этой области начались несколько десятилетий назад, именно сейчас стало очевидным, что разнообразие данных биологически активных молекул было недооценено. За последние несколько лет стало понятно, что около 5-10% генов могут кодировать кОРС-пептиды. Большое число таких пептидов было идентифицировано в различных организмах от бактерий до человека [7,8,12-19] и функция многих из кОРС-пептидов остается не изученной.

Данная работа посвящена изучению 56 ак пептида, кодируемого геном, ранее ошибочно аннотированным как длинная не кодирующая РНК. Данный пептид локализуется в митохондриях и связан с несколькими функциями этих органелл.

Цели и задачи исследования

Целью данного проекта являлось изучение функциональной роли высоко-консервативного пептида, кодируемого геном 1500011k15Rik.

Для достижения цели были поставлены следующие задачи:

1. Определить, синтезируется ли данный пептид в клеточных линиях и живом организме.
2. Определить внутриклеточную локализацию L116 пептида.
3. Инактивировать ген 1500011k15Rik в линиях клеток и исследовать фенотип нокаутных клеток.
4. Найти белковых партнеров исследуемого пептида.

Объект исследования

Механизмы регуляции клеточных процессов кОПС-пептидами.

Предмет исследования

Функция митохондриального пептида, кодируемого геном 1500011k16Rik.

Научная новизна исследования

Впервые проведено исследование функциональной роли гена 1500011k16Rik. Результаты экспериментов, полученные при использовании нескольких разных взаимодополняющих методов, свидетельствуют об участии пептида, кодируемого данным геном, в регуляции активности комплекса I дыхательной цепи и липидного метаболизма. В результате более тщательного анализа последовательности, мы обнаружили открытую рамку считывания, которая кодировала консервативный среди хордовых пептид длиной 56 ак,. Для доказательства того, что предполагаемый пептид действительно синтезируется в клетках, с помощью CRISPR/Cas9 системы были создана клеточная линия с точным встраиванием участка, кодирующего флюоресцентный белок mCherry непосредственно перед стоп кодоном гена 1500011k16Rik.

Теоретическая значимость исследования

В представленной работе нам удалось показать, что пептид локализуется в митохондриях. При исследовании нокаутных клеточных линий мы обнаружили, что активность НАДН-убихинон оксиредуктазы, комплекса I дыхательной цепи, снижена, однако это снижение не обусловлено тем, что пептид является его

компонентом. Анализ с помощью ко-иммунопреципитации показал, что пептид взаимодействует с цитохром b5 редуктазой 3 (Cyb5r3). Мы предположили, что пептид может регулировать Cyb5r3 и таким образом влиять на активность комплекса I дыхательной цепи. Впоследствии, мы показали, что митохондриальный нокаут Cyb5r3 обладает таким же фенотипом, что и нокаут по данному пептиду.

Практическая значимость исследования

Результаты данной работы способствуют пониманию функциональной роли в клетке пептида, кодируемого геном 1500011k16Rik. Также, данная работа показывает важность изучения неисследованного класса биомолекул - кОПС-пептидов, играющих важную роль в различных клеточных процессах.

Методология диссертационного исследования

В данной работе использовались современные методы исследований, такие как точная модификация генома с помощью CRISPR/Cas9 системы, биосенсоры и анализ липидного состава клеток. Для изучения функции пептида было создано множество клеточных линий с модифицированным геномом, например, клеточные линии нокаутные по данному пептиду или митохондриальной форме Cyb5r3. Для получения клеток нокаутных и с точной мутацией в геноме использовали CRISPR/Cas9 систему. Также для ко-иммунопреципитации были получены линии с суперэкспрессией тагированного пептида. Для экзогенной экспрессии применялась транспозоновая система Sleeping Beauty. Чтобы понять, как пептид влияет на функцию митохондрий, было исследовано митохондриальное дыхание в культуре клеток. Процессы клеточного дыхания и окислительного фосфорилирования исследовали с помощью прибора Оксиграф (Hansatech instruments). Для изучения липидного состава клеток была использована жидкостная хроматография с последующей масс-спектрометрией и биоинформатическим анализом большого объема данных.

Основные положения, выносимые на защиту

- Ген 1500011k16Rik кодирует пептид, который локализуется в митохондриях и экспрессируется как в клеточных линиях, так и в живом организме.
- Пептид L116 влияет на функцию Cyb5r3 через прямое взаимодействие.

- Удаление пептида L116 и митохондриальной формы Cyb5r3 приводит к нарушению активности комплекса I дыхательной цепи.
- Регулируя функцию Cyb5r3 в митохондриях, L116 влияет на липидный состав клеток, что скорее всего, является причиной уменьшения активности комплекса I.

Степень достоверности результатов

Воспроизводимость экспериментов и статистическая обработка данных подтверждает достоверность результатов диссертационного исследования. Все экспериментальные процедуры соответствуют поставленным целям и задачам. Результаты получены на современном научном оборудовании и с использованием реактивов, произведенных ведущими мировыми компаниями.

Апробация работы

Результаты были представлены на следующих научных мероприятиях: The 42nd FEBS Congress 2017, Иерусалим, Израиль 07.09.-10.09.2017; “Shaping the future: big data, biomedicine and frontier technologies-2017” ежегодная Skoltech-MIT конференция, Москва, Россия, 25.04.-26.04.2017. Также результаты работы обсуждались на семинарах лаборатории химии нуклеопротеидов с 2014 по 2018 гг.

Публикации

- 1) **Chugunova A.**, Loseva E., Mazin P., Mitina A., Navalayeu T., Bilan D., Vishnyakova P., Marey M., Golovina A., Serebryakova M., Pletnev P., Rubtsova M., Mair W., Vanyushkina A., Khaitovich P., Belousov V., Vysokikh M., Sergiev P. and Dontsova O. LINC00116 codes for a mitochondrial peptide linking respiration and lipid metabolism. // *Proc Natl Acad Sci U S A*. 2019. V. 116. pp. 4940-4945. IF 9, 504.
- 2) Smirnova V., Shestakova E., Bikmetov D., **Chugunova A.**, Osterman I., Serebryakova M., Sergeeva O., Zatsepin T., Shatsky I., Terenin I. eIF4G2 balances its own mRNA translation via a PCBP2-based feedback loop. // *RNA*. 2019. V. 25(7) pp. 757-767. IF 3,949
- 3) Petrova O., Mantsyzov A., Rodina E., Efimov S., Hackenberg C., Hakanpaa J., Klochkov V., Lebedev A., **Chugunova A.**, Malyavko A., Zatsepin T., Mishin A., Zvereva M., Lamzin V., Dontsova O. and Polshakov V. Structure and function of the N-terminal domain of the yeast telomerase reverse transcriptase. // *Nucleic Acids Res*. 2018. V. 46. pp. 1525-1540. IF 11,567.

4) **Chugunova A.**, Navalayeu T., Dontsova O., Sergiev P. Mining for Small Translated ORFs. // J Proteome Res. 2018. V. 17. pp. 1-11. IF 3,950.

5) **Chugunova A.**, Dontsova O. and Sergiev P. Methods of Genome Engineering: a New Era of Molecular Biology. // Biochemistry (Mosc). 2016. V. 81. pp. 662-677. IF 1,724.

Личный вклад автора

Личный вклад автора в проведенное исследование заключался в анализе литературных данных для постановки задач и целей проекта, планировании и проведении экспериментальных процедур, анализе и оформлении полученных данных, представлении результатов на научных мероприятиях и подготовке публикаций в научные журналы.

Структура и объем диссертации

Диссертация состоит из списка обозначений, введения, обзора литературы, материалов и методов, результатов и обсуждения, заключения, выводов, списка литературы, который включает 221 источник. Объем диссертации 125 страниц, 29 рисунков и 2 таблицы.

Обзор литературы

кОРС-пептиды оставались долгое время малоизученными из-за несовершенства компьютерных методов достоверного предсказания их генов и экспериментальных поисков к их обнаружению. Сейчас становится очевидным, что пептиды могут составлять значительную часть протеома клетки. Предметом данного обзора станут вычислительные и биохимические методы, используемые для идентификации кОРС-пептидов, а также методы их изучения и классификация.

Методы детекции кОРС-пептидов

На данный момент, идентификация кОРС-пептидов основана на трех подходах: вычислительные методы, рибосомное профилирование и масс-спектрометрия. Все подходы будут рассмотрены в этой главе.

Вычислительные методы

Анализ геномной последовательности на предмет наличия открытых рамок считывания является первым шагом в поиске новых кОРС-пептидов. Данный метод не обладает специфичностью к действительно транслируемым открытым рамкам считывания, так как потенциальная кОРС может быть обнаружена в любой нуклеотидной последовательности совершенно случайно. Таким образом, бесчисленное количество кОРС могут быть найдены в любом геноме, хотя большинство из них не кодируют функциональных пептидов. Одна из проблем вычислительных методов для поиска кодирующей последовательности – это какое ограничение длины использовать как пороговое? Ранее использовали ограничение в 100 ак [20], так как лишь несколько известных функциональных белков меньше 100 ак [21]. Использование данного порогового значения не оптимально, так как кОРС могут состоять всего лишь из двух кодонов [22] и тогда большая часть функциональных пептидов не будут идентифицированы. Вторая проблема, с которой столкнулись исследователи, заключается в том, что некоторые функциональные белки начинаются не с AUG-кодона [23]. Таким образом, если допустить прочтение рамок, начинающихся не с AUG, число закодированных в

геноме гипотетических, пептидов многократно возрастает. Если ограничиваться только рамками считывания, начинающимися с AUG, многие пептиды будут упущены. На данный момент существует несколько вычислительных подходов для поиска функциональных пептидов (Таблица 1).

Таблица 1

Программа	Ссылка	Принцип работы
CRITICA (coding region identification tool invoking comparative analysis)	[24]	Алгоритм предсказания генов, который интегрирует анализ эволюционного отбора и нуклеотидного состава для идентификации потенциально-кодирующих ОРС. Последовательность ДНК потенциальной кОРС выравнивается с похожими последовательностями из других организмов. Аналогичная процедура проводится с аминокислотными последовательностями гипотетических продуктов трансляции. В случае если продукт трансляции обладает большим процентом идентичности, чем нуклеотидная последовательность, это служит доказательством кодирующего потенциала ОРС.

CPC (Coding Potential Calculator)	[25]	Биоинформатическая программа, использующая шесть особенностей последовательностей, кодирующих белки, для их выявления в геноме. Три особенности связаны с качеством ORC. Это длина ORC, покрытие, т.е. соотношение длины ORC к длине РНК и целостность, т.е. расположение рамки считывания в одной молекуле РНК. Другие три особенности, анализируемые с помощью BLASTX, основаны на схожести последовательности с уже известными, каталогизированными в UniProt Reference Clusters (число хитов, их качество и распределение).
sORFinder	[26]	Инструмент, позволяющий идентифицировать кОРС с кодирующим потенциалом на основе сходства их нуклеотидного состава с <i>bona fide</i> кодирующими генами, используя скрытые цепи Маркова.
PhastCons	[27]	Программа предсказывает консервативные элементы во множественных выравниваниях.

		Она основана на статистике скрытой филогенетической модели Маркова (phylo-HMM), которая принимает во внимание вероятность нуклеотидной замены в каждом месте генома и как эта вероятность изменяется от одного сайта к другому.
PhyloCSF	[28]	Сравнительный метод, анализирующий множественные выравнивания нуклеотидных последовательностей, используя статистическую модель замен кодонов родственных белков, для определения кодирующего потенциала консервативных последовательностей.
micPDP	[29]	Метод, который оценивает соотношение синонимичных замен к несинонимичным (Ka/Ks). Программа фильтрует выравнивания последовательностей-кандидатов на основе их консервативности, покрытия (соотношение длины ОРС и РНК), а затем метод PhyloCSF применяется для определения их кодирующего потенциала.
uPEPperoni	[30]	Программа позволяет

		идентифицировать консервативные пептиды в 5'-НТО, сравнивая нуклеотидную последовательность кОПС с последовательностями мРНК в базе данных NCBI RefSeq.
--	--	---

В основном предсказательные программы используют комбинацию различных характеристик кодирующих последовательностей, для того чтобы отличить их от не кодирующих (Рисунок 1). Большинство из них основано на одних и тех же принципах [31]:

1) идентификация консервативной кОПС с помощью сравнения гомологичных последовательностей различных видов (Рисунок 1а);

2) анализ частоты использования кодонов и других характерных особенностей кодирующей последовательности (Рисунок 1б);

3) оценка сходства с уже известными белками или функциональными доменами (Рисунок 1в-г).

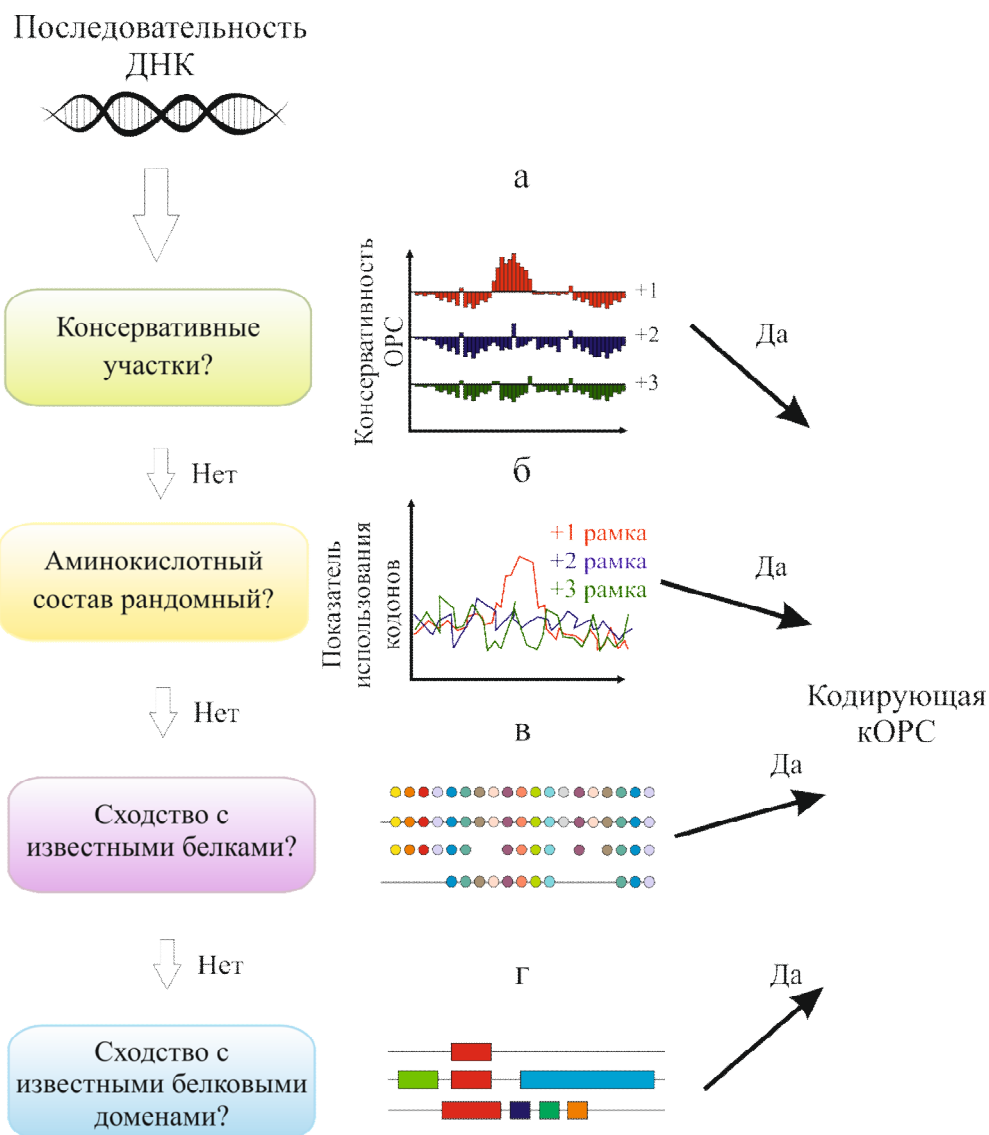


Рисунок 1. Характеристики нуклеотидной и предполагаемой аминокислотной последовательности, используемые для обнаружения предполагаемых кОРС. (а) Пример распределения показателя консервативности в трех различных рамках считывания. Показатель консервативности может включать соотношение синонимичных замен к несинонимичным (dS/dN) или параметр, который учитывает вероятность обнаружения конкретной замены в кодирующем регионе на определенных межвидовых эволюционных расстояниях по сравнению с некодирующим регионом. (б) Пример вероятностного распределения показателя для трех различных рамок считывания. В качестве параметра может быть использовано либо предпочтение использования кодона или дикодона (гексамера), либо 3-х нуклеотидная периодичность в нуклеотидном составе. (в) Пример выравнивания белковых последовательностей предполагаемого пептида со сходными последовательностями, найденными у известных белков. Цветными шариками показаны различные аминокислоты, линия обозначает пропуск. (г) Поиск домена внутри

предполагаемого белка, который бы был похож на какой-либо консервативный домен известного белка. Прямоугольниками обозначены различные белковые домены.

При поиске функциональных пептидов консервативность - это один из самых важных показателей. В большинстве случаев, функциональные пептиды и белки являются продуктом эволюции и как следствие их последовательность должна обладать консервативностью по сравнению с некодирующей последовательностью (Рисунок 1а). Для определения кодирующей последовательности могут быть использованы различные показатели, начиная от консервативности нуклеотидной последовательности [27] до анализа соотношения синонимичных замен к несинонимичным, так называемый dN/dS показатель [32,33]. Принимая во внимание вероятность замены кодонов определенного типа на данном филогенетическом расстоянии, можно более точно предсказывать кодирующие последовательности [34]. Сходство с известными белковыми доменами (Рисунок 1б) или известными белками (Рисунок 1в) также служит важным показателем. Стоит отметить, что оба подхода имеют явное ограничение, заключающееся в том, что поиск гомологов может не иметь смысла для только недавно появившихся функциональных пептидов, с короткой эволюционной историей [29]. Также стоит отметить, что полногеномные последовательности и выравнивания доступны не для всех родственных видов. Поиск сходства с известными доменами или белками ограничен длиной ОРС, так как для этого метода важно количество консервативных позиций [35].

Другой подход для предсказания функциональных пептидов заключается в анализе нуклеотидного и кодового состава [36,37]. Кодирующие последовательности отличаются от некодирующих таким простым параметром, как частота нуклеотидов, например, для многих организмов кодирующие последовательности более GC-богаты [38]. Также вероятность того, что последовательность является кодирующей, можно оценить с помощью периодичности частоты нуклеотидов, так как все три позиции кодона отличаются по составу в среднем [37].

Сравнительные геномные подходы могут быть комбинированы с нуклеотидной и кодовой периодичностью для повышения точности

предсказания кодирующего потенциала последовательности [24]. Однако на данный момент самым убедительным доказательством функциональности кОРС-пептидов является эволюционная консервативность среди различных видов [39]. Биохимические методы необходимы для подтверждения существования предсказанных с помощью вычислительных методов пептидов.

Рибосомное профилирование

На данный момент рибосомное профилирование является одной из важных технологий, позволивших качественно и количественно охарактеризовать трансляцию всей совокупности мРНК в клетке [40]. Рибосомное профилирование использует классический метод рибосомного футпринтинга [41], где активно транслирующиеся в клетках мРНК, выделяемые в виде полисом, обрабатываются нуклеазой *in vitro* для получения фрагментов РНК, защищенных рибосомами [40]. Таким образом после обработки получают «футпринты» длиной примерно 30 нк, которые могут быть соотнесены с последовательностью мРНК и показать точное место транслирующей рибосомы (Рисунок 2а). Благодаря данному методу было показано, что многие РНК, считавшиеся не кодирующими, на самом деле активно транслируются [42,43].

Несмотря на все преимущества данного метода, исследователи столкнулись с рядом проблем. При первичном анализе данных рибосомного профайла стало казаться, что практически любая РНК может транслироваться [40,42,44-47]. Похожий феномен наблюдается при анализе транскриптома, когда кажется, что большая часть генома транскрибируется, но при этом многие РНК не обладают никакой функцией. По аналогии, можно предположить, что продукты «полнотранскриптомной» трансляции не являются функциональными. Фактически, большая часть из них крайне нестабильна и деградирует в течение нескольких минут после синтеза [48]. Более того оказалось, что футпринты могут быть сигналом не только транслирующей рибосомы, но и белков, ассоциированных или совыделяющихся с рибосомами. Например, сканирующая рибосомная субъединица 40S также вносит вклад в футпринты [49].

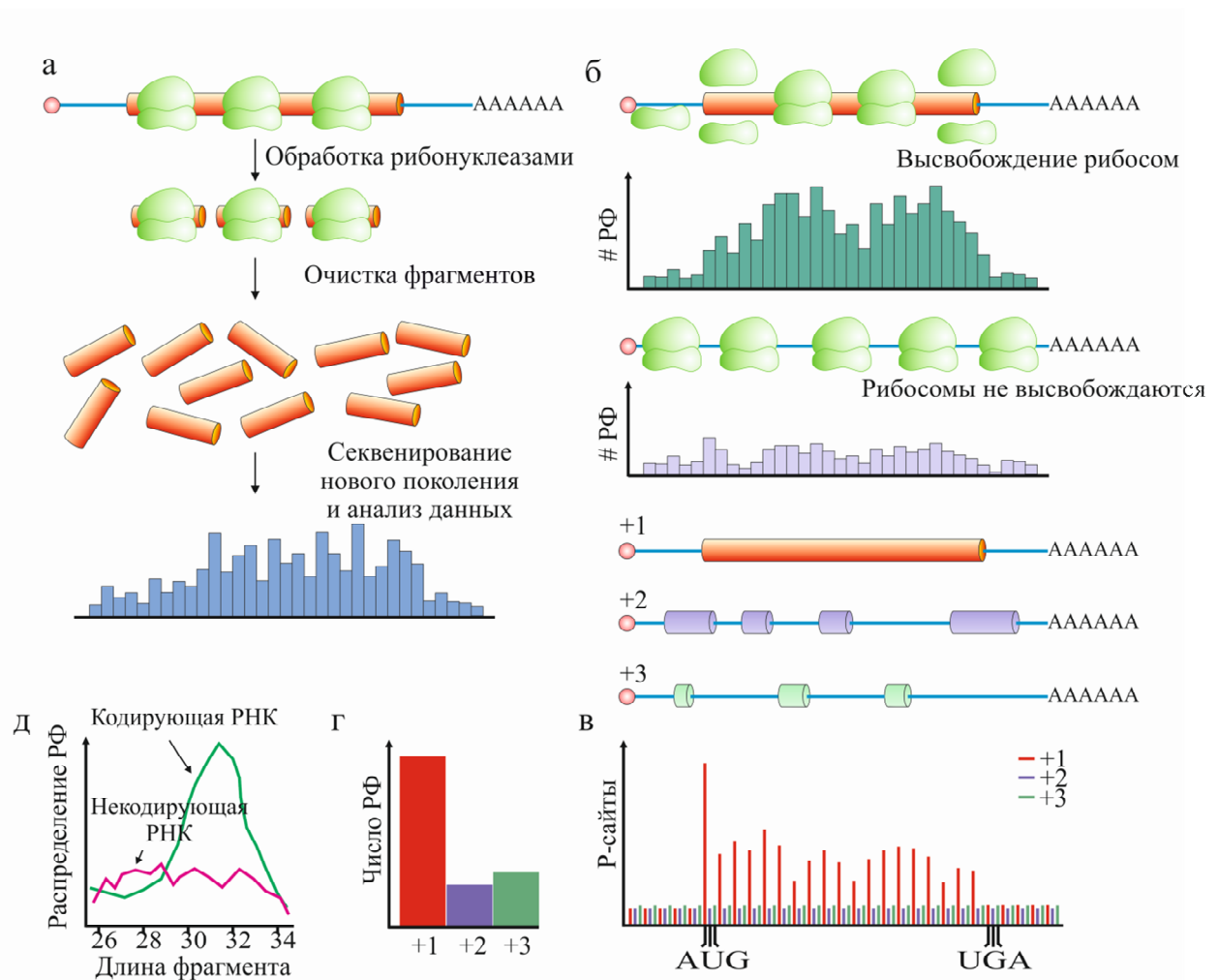


Рисунок 2. (а) Схема эксперимента по рибосомному профилированию. Сперва мРНК, связанная с рибосомами, выделяется из клеток. После обработки нуклеазой остаются только те участки РНК, которые были связаны с рибосомами. Последующее выделение маленьких кусочков РНК и их секвенирование, позволяет установить местоположение рибосом на каждой РНК. (б-д) Различные биоинформатические принципы анализа данных рибосомного профилирования: (б) *Ribosome release score*. Голубая линия обозначает мРНК, а красный цилиндр - ОРС. Столбцы показывают плотность рибосом для каждой позиции транскрипта. В случае кодирующей мРНК, происходит резкое уменьшение количества рибосом после стоп-кодона по сравнению с некодирующей РНК. (в) *RiboTaper*. Три возможные рамки считывания показаны как +1 (красная), +2 (голубая) and +3 (зеленая). В данном примере только рамка +1 кодирует функциональный пептид. AUG и UGA обозначают старт- и стоп-кодоны соответственно. В случае если ОРС транслируется, будет наблюдаться 3-х нуклеотидная периодичность футпринтов. (г) *ORF score*. Кодирующая ОРС содержит наибольшее число рибосомных футпринтов. (д) *FLOSS*. Картинка схематически

показывает распределение длины рибосомных футпринтов для кодирующей и не кодирующей РНК. Футпринты рибосом, а не РНК-связывающих белков, обычно имеют длину около 30-32 нк. Для всех изображений красный шарик обозначает 5'-кэп РНК, РФ (рибосомные футпринты), ОРС (открытая рамка считывания).

Для того чтобы отличить футпринты реально транслирующих рибосом от футпринтов, обусловленных артефактами, были предприняты попытки улучшить экспериментальную базу метода и вычислительные алгоритмы для обработки экспериментальных данных.

С момента создания метода было разработано большое число вычислительных алгоритмов для анализа данных (Таблица 2). Данные подходы основаны на определённых свойствах рибосомных футпринтов. Например, 3-х нуклеотидная периодичность, являющаяся результатом движения рибосомы от одного кодона к другому во время трансляции, наиболее часто используемое свойство [29,50-54]. Такой подход позволяет детектировать даже сдвиг рамки считывания и трансляцию перекрывающихся ОРС [55].

Таблица 2

Программа	Ссылка	Принцип работы
Ribosome Release Score (RRS)	[56]	Данный параметр рассчитывает соотношение между общим количеством ридов кодирующей области к общему числу ридов, картированных в 3'-НТО, таким образом, измеряя диссоциацию рибосом на стоп-кодоне предположительно кодирующей области.
Translated ORF Classifier (TOC)	[42]	Программа использует четыре свойства для определения кодирующего потенциала: 1. Эффективность трансляции (соотношение плотности рибосомных футпринтов по всей длине ОРС к экспрессии РНК, полученной из РНК секвенирования).

		2. Соотношение среднего покрытия рибосомными футпринтами внутри ОРС к покрытию вне ОРС. Под покрытием подразумевается число прочтений, содержащих данный конкретный нуклеотид. 3. Доля длины ОРС (длина ОРС, деленная на полную длину транскрипта). 4. Показатель высвобождения (число футпринтов внутри ОРС, деленное на их число после стоп-кодона).
Fragment Length Organization Similarity Score (FLOSS)	[44]	Этот метод основан на том, что распределение длин рибосомных футпринтов отличается от распределения длин фрагментов не кодирующих РНК. Измеряет кодирующий потенциал на основе сравнения распределения длины прочтений рибосомных футпринтов для исследуемого гена с таким же параметром для известных кодирующих генов.
ORFscore	[29]	Подход основан на том, что транслирующая рибосома движется от кодона к кодону. Таким образом, футпринты рибосомного профайлинга обладают трех-нуклеотидной периодичностью на кодирующей ОРС. Данный метод анализирует лишь футпринты с длиной 28-29 нк, соответствующих средней длине прочтений футпринтов транслирующих рибосом.
ORF Regression Algorithm for	[57]	Данный подход рассчитывает вероятность трансляции ОРС из данных рибосомного профайлинга, сравнивая паттерны заполняемости

Translational Evaluation of RPFs (ORF-RATER)		рибосомами (на стадии инициации, терминации и элонгации) с кодирующими ОРС. Алгоритм использует модель линейной регрессии, которая позволяет объединение самой разнообразной информации и оценивает каждую ОРС в соответствии с окружающим контекстом.
RibORF Classifier	[51]	Отличает кодирующие последовательности от не кодирующих на основе следующих параметров: 1. Трех-нуклеотидная периодичность рибосомных футпринтов. 2. Однородность распределения футпринтов по всей длине исследуемой ОРС.
RiboTaper	[52]	Работает по аналогии ORFscore, но использует многоканальный спектральный анализ для получения трех-нуклеотидной периодичности из первичных данных рибосомного профайлинга. Это позволяет рассчитать паттерны ОРС, используя футпринты различной длины, при условии, что Р-сайт определен для каждой длины.
SPECtre	[53]	Идентифицирует кодирующие ОРС, используя трех-нуклеотидную периодичность футпринтов.
RiboGalaxy	[54]	Интегрированная платформа для анализа и визуализации данных рибосомного профайлинга.
Proteoformer	[58]	Платформа, объединяющая данные рибосомного профилирования и масс-спектрометрии.

RiboHMM	[50]	Использует скрытые цепи Маркова для анализа рибосомных футпринтов. Принимает во внимание ожидаемую относительную плотность футпринтов в кодирующей, 5'-НТО и 3'-НТО областях и прилегающей к границам кодирующей и нетранслируемой области, а также распределение футпринтов по положениям внутри триплетов.
---------	------	--

Другой алгоритм основан на том факте, что длина рибосомного футпринта составляет 28-31 нк. Распределение длины футпринтов для исследуемой ОРС помогает различить транслируемые участки РНК от тех, которые лишь сканируются, связаны с пост-терминирующей рибосомой или другими белками [44].

Наконец, рибосомные футпринты обладают характерным распределением по всему транскрипту. А именно, мы ожидаем резкое уменьшение количества футпринтов после стоп-кодона. Такой подход использует ряд программ для определения транслируемых последовательностей [42,56]. Данный алгоритм неэффективен для предсказания перекрывающихся ОРС (Рисунок 2б, в, г, д). Многие исследования учитывают соотношение покрытия футпринтами между ОРС и нетранслируемыми областями [42,51].

Одним из усовершенствований экспериментальной базы является использование антибиотиков харрингтонина и лактимидомицина для остановки рибосомы на старт-кодоне [59,60], что позволяет определить стартовые точки трансляции [44,61,62]. Для того, чтобы решить проблему с неспецифическими футпринтами, обусловленными сканирующей 40S субъединицей и РНК-связывающими белками, был предложен метод Poly-Ribo-Seq. Данный метод предполагает анализировать только РНК, связанные с полисомами, в составе которых мРНК активно транслируется [63].

Также был изобретен метод TCP-seq, который предполагает использование формальдегидной сшивки. Данный подход позволяет выделить малые рибосомные субъединицы, связанные с участками мРНК, помимо обычных 80S связанных футпринтов, что помогает получить информацию об инициации и элонгации трансляции [64].

На данный момент, существует несколько полезных ресурсов, таких, как sORF.org [65] и GWIPS-viz [66] для исследования кОРС. sORF.org (<http://www.sorfs.org>) - удобная база данных кОРС, идентифицированных с помощью рибосомного профайла в трех разных клеточных линиях HCT116 (человек), E14_mESC (мышь) and S2 (дрозофила). The GWIPS-viz (<http://gwips.ucc.ie>) содержит данные рибосомного профилирования из различных организмов и позволяет исследователю анализировать альтернативные протеоформы. Также недавно был разработан Ribogalaxy [67], который объединяет весь процесс обработки данных рибосомного профилирования.

Несмотря на все преимущества рибосомного профилирования, его точность зависит от точности экспериментального набора данных [68]. Например, существуют примеры, когда кОРС, охарактеризованные как кодирующие по одному протоколу [27], не имеет рибосомных футпринтов по другому [11,42]. Не менее важными оказываются способы обработки данных, разница в которых может вести к противоположным заключениям [42,56].

Рибосомное профилирование эффективно, однако и его необходимо сочетать с другими методами исследования.

Масс-спектрометрия

Масс-спектрометрия – это эффективный метод прямой детекции и оценки количества белков и пептидов и является золотым стандартом в протеомных исследованиях. Данный метод отличается от рибосомного профилирования тем, что можно детектировать пептиды, кодируемые кОРС, и таким образом напрямую подтвердить то, что данная РНК транслируется. На данный момент наиболее популярным подходом для изучения белкового состава клетки является метод скорострельной протеомики. Он основан на сочетании жидкостной хроматографии и масс-спектрометрии и полезен для детекции как природных пептидов, так и

полученных с помощью гидролиза полноразмерных белков [69] (Рисунок 3а). В протеомных исследованиях успешная идентификация пептида зависит от полноты базы данных белковых последовательностей. Выбор базы данных определяет чувствительность и специфичность идентификации пептида, а также время и вычислительные ресурсы необходимые для анализа данных. Определение пептидов осуществляется путем сравнения полученных MS/MS спектров с теоретическими спектрами всех пептидов, представленных в одной базе данных, например, Ensembl, RefSeq, UniProtKB [70] (Рисунок 3б, в). Большинство пептидов, идентифицированных в протеомных исследованиях, это пептиды, которые можно сопоставить с уже известными белок-кодирующими генами. У эукариот многие детектируемые пептиды локализируются внутри экзонов, оставшиеся пептиды (менее 20%) [71] находятся в области соединения экзонов. Короткие пептиды могут отсутствовать в некоторых базах данных, или быть пропущенными из-за вариативности генома или отсутствия информации об альтернативно-сплайсированных транскриптах.

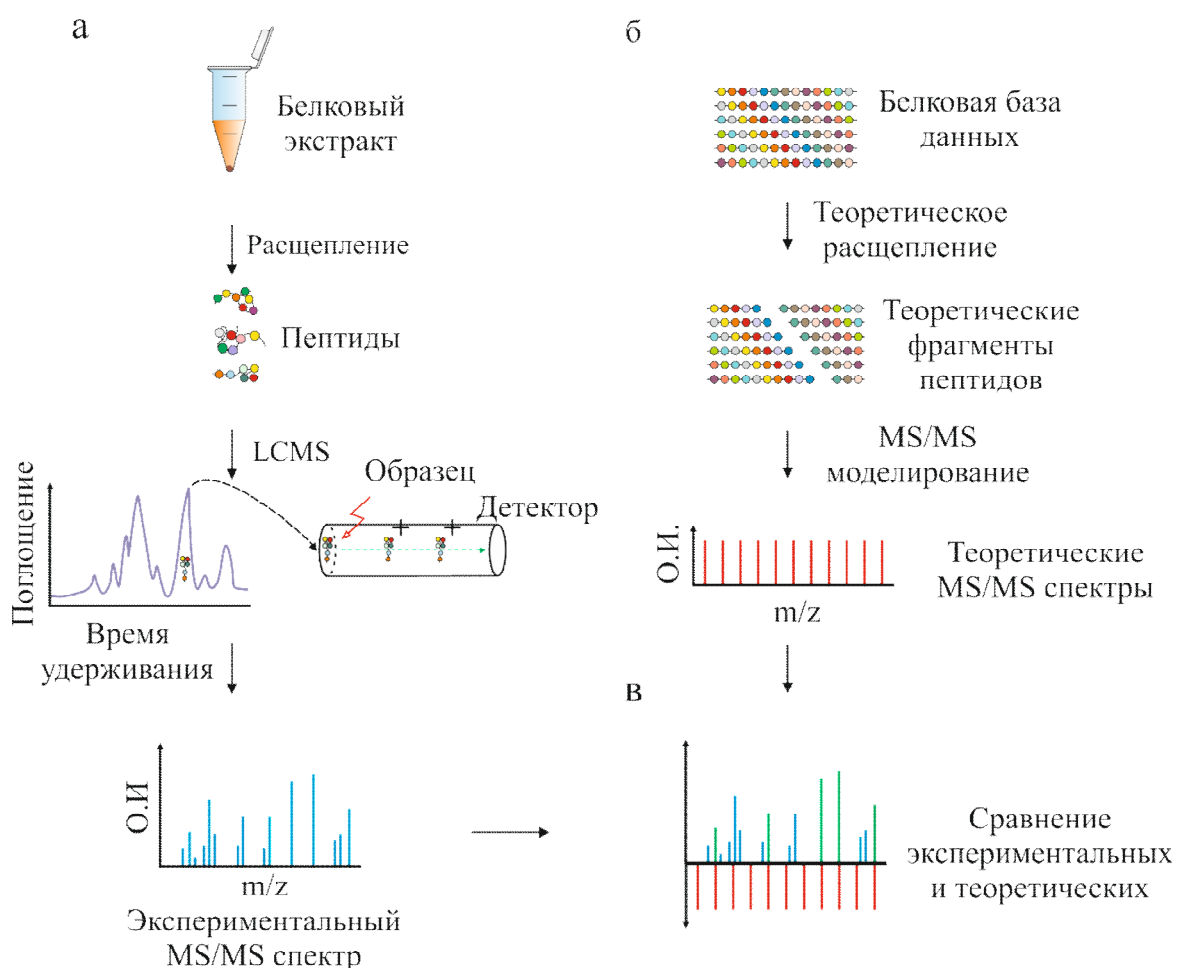


Рисунок 3. Идентификация пептидов с помощью скорострельной протеомики. (а) Используя трипсин или другую протеазу, белковый экстракт расщепляется на пептиды, которые затем разделяются с помощью жидкостной хроматографии с последующей tandemной масс-спектрометрией. (б) Белковые последовательности в базе данных расщепляются на фрагменты теоретически, учитывая специфичность используемой протеазы. Затем моделируются теоретические MS/MS спектры. (в) Сравнение теоретических и экспериментальных спектров позволяет идентифицировать пептиды. LCMS (liquid chromatography-mass spectrometry); MS/MS (tandem mass spectrum); О.И. (относительная интенсивность).

Создание специализированных под конкретную задачу баз данных может решить данную проблему. Существует несколько подходов создания баз данных, например, трансляция всего генома. К сожалению, такой массив информации сложно использовать, в связи с его чрезвычайно большим размером и, как следствие, присутствием большого числа несуществующих белковых последовательностей [72,73]. К примеру, в результате трансляции человеческого генома (UCSC v.19) получается база данных 3,2 Гб, что соответствует 70 кратному увеличению по сравнению с 45 Мб Ensembl [73]. Несколько вычислительных стратегий могут быть использованы для того, чтобы исключить менее вероятные последовательности на основании отсутствия гомологии с известными кодирующими последовательностями и ограничения длины 30 аминокислотными остатками [74]. Однако, чем больше последовательностей в базе данных, тем больше вероятность, что наилучшее совпадение со спектром неверно и становится труднее различить между истинным и ложным определением [75]. В результате, поиск MS/MS спектров в большой базе может привести к обнаружению новых пептидов, однако общее число достоверно идентифицированных пептидов значительно упадет [74,76].

Кодирующие последовательности возможно определить, используя трансляцию базы данных EST (Expressed Sequence Tag) в трех рамках считывания. Преимущество данной базы данных состоит в том, что она предоставляет экспериментальное доказательство транскрипции, включая информацию о структуре участков интрон-экзон и сплайсинге [77-79]. Однако, как и предыдущая

база данных, она обладает значительным большим размером по сравнению со стандартными, что усложняет ее использование.

На данный момент, когда секвенирование стало более доступным, исследователи получили возможность создавать персонализированные базы данных белковых последовательностей на основании секвенирования экспрессирующихся РНК [80,81] или данных рибосомного профилирования [82]. Транскрипты транслируются в трех рамках считывания, процесс может быть автоматизирован с помощью нескольких биоинформатических программ. Используя систему Galaxy-P, пользователи могут конвертировать исходные данные РНК секвенирования в три разные базы данных для анализа: содержащую белковые последовательности с вариациями в аминокислотной последовательности; новые пептиды в области соединения экзонов; только белки, кодируемые транскриптами с определенным уровнем экспрессии [83]. Также недавно был разработан PROTEOFORMER [58], который автоматически анализирует данные рибосомного профилирования и визуализирует количество и положение рибосом по всему геному. Данный инструмент включает информацию о расположении стартовых точек трансляции, что позволяет разграничивать ОРС всех продуктов трансляции. Таким образом полная база данных белковых последовательностей, основанная на данных рибосомного профилирования может быть использована как референс-база данных для масс-спектрометрических исследований. Такой подход увеличивает долю идентифицированных пептидов на 3% и 11% для человека и мыши соответственно и позволяет детектировать ряд 5'-удлиненных форм белков и кОРС-пептидов.

Также как и в случае рибосомного профилирования, пробоподготовка для масс-спектрометрии может значительно повлиять на результат. Стратегия использования изоэлектрофокусировки пептидов позволила идентифицировать большое число пептидов (13,078 для человека и 10,637 для мыши) [84], в связи с тем, что MS/MS спектры из фракции с определённой изоэлектрической точкой будут сравниваться с пептидами, обладающими ожидаемой изоэлектрической точкой. Также исследователи продемонстрировали, что использование технологии ERLIC (Electrostatic Repulsion Hydrophilic Interaction Chromatography) позволяет

определить в одном и том же образце 90-94 кОПС-пептидов вместо 13-19 без данной технологии [85].

Несмотря на несомненный успех секвенирования РНК и схожих технологий в изучении разнообразия генома, протеомные исследования играют ключевую роль в обнаружении функциональных транскриптов. Анализ данных, полученных с помощью рибосомного профилирования и масс-спектрометрии, могут служить убедительным доказательством существования белкового продукта.

Различные виды кОПС

Ранее считалось, что белок не может быть меньше 100 а.к.д. длиной. Это привело к тому, что многие мРНК, кодирующие пептиды меньше этого порогового значения, были аннотированы как не кодирующие. Таким же образом была отброшена возможность существования таких пептидов в других якобы не кодирующих областях. В частности, традиционное представление о том, что мРНК кодирует только один белок и содержит 5'- и 3'- не кодирующие области было подвергнуто сомнению. Появилось большое число экспериментально доказанных примеров того, что аннотированные белки не являются единственным продуктом трансляции некоторых мРНК. В особенности, масс-спектроскопические исследования и данные рибосомного профилирования показали наличие альтернативных ОПС и их продуктов. В данной главе мы классифицируем кОПС-пептиды по расположению кодирующих их последовательностей для того, чтобы показать их распространенность и разнообразие (Рисунок 4).

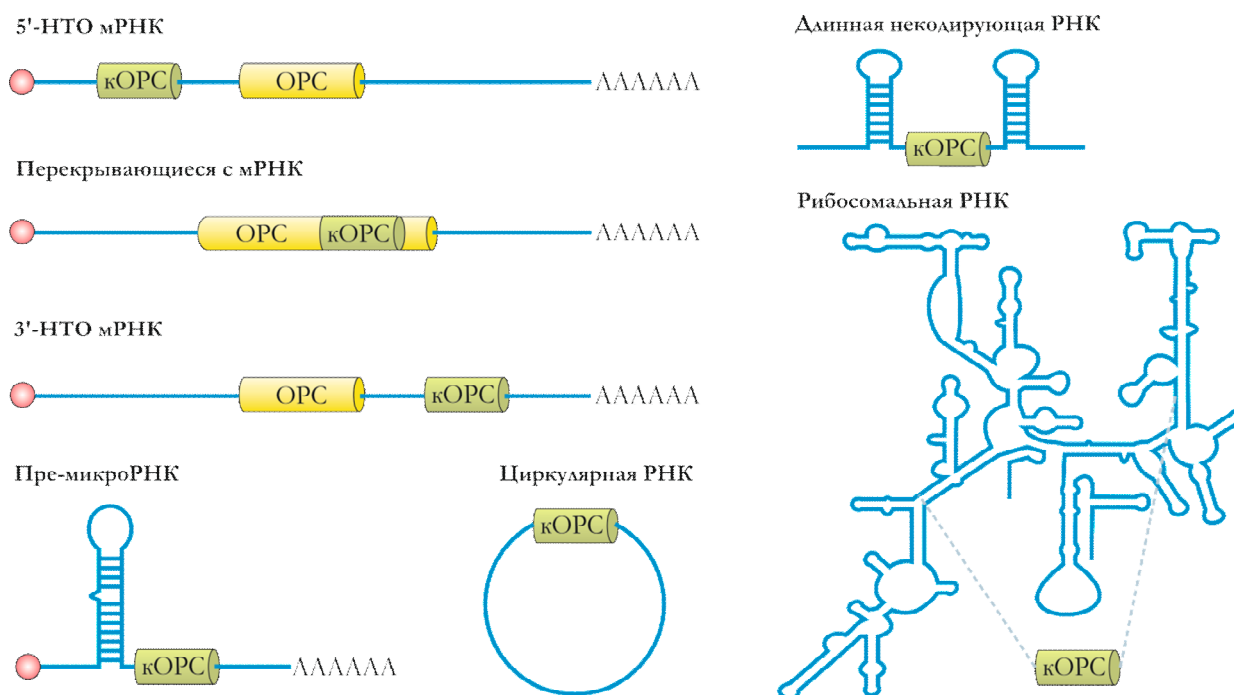


Рисунок 4. Обзор РНК, который могут кодировать кОРС-пептиды. Красный шарик обозначает 5'-кэп РНК; ОРС (открытая рамка считывания); НТО (нетранслируемая область).

5'-НТО

кОРС, кодируемые в 5'-нетранслируемой области мРНК являются чаще всего регуляторным элементом, контролирующим экспрессию основной рамки считывания. Около 40% 5'-НТО млекопитающих содержит данные ОРС, что говорит об их важности [86]. Некоторые из них консервативны среди различных видов, предполагая эволюционный отбор функциональных кОРС [87-91]. К примеру, гены *CD36*, *MDM2*, *ERBB2*, *SOC1* и *RARB* содержат консервативные и функционально охарактеризованные кОРС, регулирующие трансляцию [91]. 5'-НТО кОРС являются характерным элементом некоторых классов мРНК, трансляция которых не должна быть эффективной, например, кодирующих онкогены и другие важные белки для дифференцировки, регуляции клеточного цикла и ответа на стресс [92-95]. Мутации в таких кОРС ведут к различным заболеваниям, что было показано для нескольких генов человека [89,96-100]. В большинстве случаев [91,101], но не всегда [102], 5'-НТО кОРС вовлечены в регуляцию трансляции нижележащего гена. Они могут уменьшать количество белка через модулирование эффективности трансляции [103] или вызывать

деградацию мРНК [104-106]. Напротив, в условиях стресса они могут положительно регулировать уровень экспрессии белка [107,108]. Существует класс 5'-НТО кОПС, которые влияют на трансляцию в ответ на малые молекулы, вызывая остановку рибосомы на стоп-кодоне [109-113].

В большинстве случаев для таких кОПС значение имеет процесс трансляции, а не ее продукт. Однако существуют примеры, когда пептиды, кодируемые 5'-НТО кОПС, имеют функциональное значение помимо регуляции трансляции. Например, 5'-нетранслируемая область РНК SgrS (sugar transport-related sRNA) бактерии кодирует короткий пептид SgrT, ингибирующий потребление глюкозы через прямое взаимодействие с транспортером глюкозы [13] (Рисунок 5). Другой пример – это высококонсервативные пептиды uMKKS1 и uMKKS2 [114]. Пептиды транслируются с короткого транскрипта и затем направляются в митохондриальную мембрану, однако их функция до сих пор непонятна. Также еще одним примером является пептид, закодированный 5'-лидерной последовательности гена *AT1aR* (angiotensin type 1a receptors). Пептид препятствует работе сигнальных путей, несвязанных с G белком и активированных ангиотензином II. Было предположено, что он взаимодействует с аллостерическим сайтом AT1aR, который отличается от сайта связывания ангиотензина II, и затем разрушается после опосредованного рецептором эндоцитоза [9]. В целом, 5'-НТО кОПС не только регулируют трансляцию последующей рамки считывания, но и являются источником пептидов, которые обладают независимой функцией в клетке.

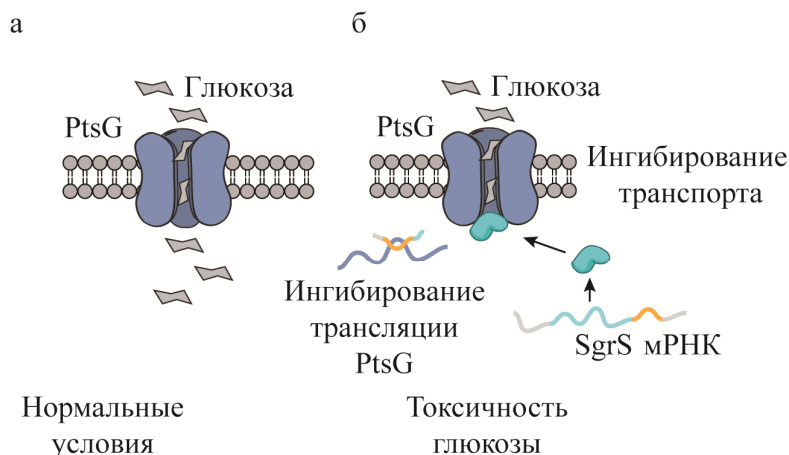


Рисунок 5. Бактериальный кОПС-пептид с метаболической функцией [115]. (а) Бактерии используют транспортер PtsG для импорта глюкозы при нормальных условиях. (б) Приток большого количества глюкозы очень токсичен для клетки. Чтобы справиться с этим, бактерии используют PHK SgrS, кодирующую кОПС-пептид SgrT. Ингибирование поглощения глюкозы происходит двумя путями. SgrS ингибирует трансляцию с помощью антисенс взаимодействия с мРНК ptsG, а пептид SgrT, взаимодействуя с транспортером PtsG, ингибирует его активность.

Кодирующая область мРНК

За последнее время стало понятно, что одна мРНК может быть транслирована в различных рамках, в результате чего получатся совершенно разные пептиды или белки [116,117]. Еще в 1996 году было показано, что антиген меланомы gr75 кодирует два разных полипептида, белок gr75 и пептид длиной 24 ак. Он служит как опухолеспецифический антиген, участвующий в реакции отторжения опухоли, и узнается Т клетками [118]. С тех пор было показано, что несколько человеческих антигенов отторжения опухоли продуцируются с альтернативных ОРС [119-121]. Интересно, что данный феномен наблюдается не только для раковых клеток. Существует несколько примеров парных генов, когда помимо основного ОРС есть еще перекрывающийся с ним. К ним можно отнести такие гены как *INK4a/ARF* [122], *гистон H4/OGP* [123], *XLalphas/ALEX* [124,125], *PrP/altPrP* [126], *ATXN1/altATXN1* [127]. Биоинформатический анализ показал, что

около 41% человеческого генома содержит как минимум одну альтернативную ОРС внутри основной, и большинство из них кодируют пептиды длиной меньше 90 ак [128]. Более того, для генов *BDH2*, *NIPAI*, *SCARB2*, *LGALS3BP*, *VEGFC* и *p53* было показано с помощью масс-спектрометрии, что продукты трансляции их альтернативной ОРС действительно существуют. Исследователям также удалось показать, что многие из них обладают определенной локализацией, что говорит в пользу их функциональности [128].

3'-НТО

В отличие 5'-НТО пептидов, 3'-НТО мало изучены, так как считалось, что эта область мРНК не транслируется [129]. Было показано, что рибосомы иногда остаются ассоциированными с 3'-НТО мРНК, однако ранее это считалось результатом пропуска стоп-кодона [130,131] или задержкой диссоциации комплекса рибосом и мРНК после терминации [132]. Несколько исследований приводят доводы в пользу трансляции 3' области мРНК. С помощью рибосомного профилирования и масс-спектрометрии было обнаружено, что некоторые 3'-НТО транслируются и продукты их трансляции могут быть обнаружены в клетке [29,85,128,129,133,134]. Например, 3'-НТО мРНК гена гистосовместимости H60 транслируется и молекулы МНС I представляют получающийся пептид иммунной системе [135,136]. Взаимодействие этого комплекса с цитотоксическими Т клетками способствует их толерантности к антигенам собственного организма. Также недавние исследования открыли, что 3'-НТО гена *MRV11* кодирует пептид, который локализуется и взаимодействует с *BRCA1*, хотя роль данного взаимодействия остается неисследованной [128]. Предсказано существование многих пептидов, закодированных в 3'-НТО, а вышеупомянутые примеры показывают, что этот класс молекул заслуживает большего внимания, чем считалось ранее.

Длинные некодирующие РНК

РНК длиннее 200 нк, но в то же время не имеющие ОРС длиннее 100 ак, ранее аннотировались как длинные некодирующие РНК (днРНК), представляющие на данный момент огромный класс РНК в клетке. Вследствие их длины, они могут кодировать большое число неаннотированных биоактивных пептидов. И

действительно, было показано, что многие днРНК кодируют пептиды с различными [10,137-140] функциями в клетке.

Одним из примеров является РНК *pgc*, которая содержится в полярных гранулах дрозофилы и ранее считалась некодирующей. Однако, когда секвенирование генома дрозофилы выявило ошибку в последовательности, стало понятно, что *pgc* содержит ОРС, кодирующую высококонсервативный среди 12 видов рода *Drosophila* пептид длиной 71 аа [141]. Иммуноокрашивание показало, что Pgc пептид локализуется в полярных клетках, предшественниках половых клеток. Репрессия РНКП II-зависимой транскрипции в первичных половых клетках – это общий механизм для поддержания этих клеток и предотвращения их дифференцировки. Транскрипционная репрессия ассоциирована с потерей фосфорилирования С-концевого домена РНК полимеразы II. Оказалось, что Pgc пептид взаимодействует с P-TEFb (positive elongation factor b), который принимает участие в фосфорилировании С-концевого домена РНК полимеразы II, тем самым активируя ее. Удаление Pgc пептида ведет к тому, что клетки не могут предотвратить реакцию фосфорилирования [141,142]. Было найдено, что пептид принимает участие в ингибировании перемещения Р-TEFb к сайтам транскрипции, таким образом, предотвращая модификацию С-концевого домена РНК полимеразы II.

Еще один пример такого рода это *pri* РНК, которая экспрессируется во время раннего развития. Мутантные по данному гену мушки умирают во время эмбриогенеза. РНК содержит пять маленьких и высококонсервативных ОРС, и лишь только первые 4 транслируются [18]. Исследователям удалось показать, что эти пептиды регулируют транскрипционный фактор *shavenbaby (svb)*, или *Ovo* [10], однако только недавно стало понятно, каким образом это происходит (Рисунок 6). *Ovo* экспрессируется как удлинённая неактивная форма белка, и Pri пептиды принимают участие в укорочении его N-концевого участка, чтобы получить более короткую активную форму белка. Полногеномный скрининг с помощью РНК интерференции показал, что убиквитин-протеасомная система возможно участвует в Pri-зависимом процессинге белка *Ovo*. N-концевой домен транскрипционного фактора подвергается Ubr3-опосредованному

убиквитинилированию и последующему процессингу до активной формы. Пептиды связываются с Ubr3, тем самым способствуя ассоциации между Ubr3 и N-концевым доменом Ovo. [143].

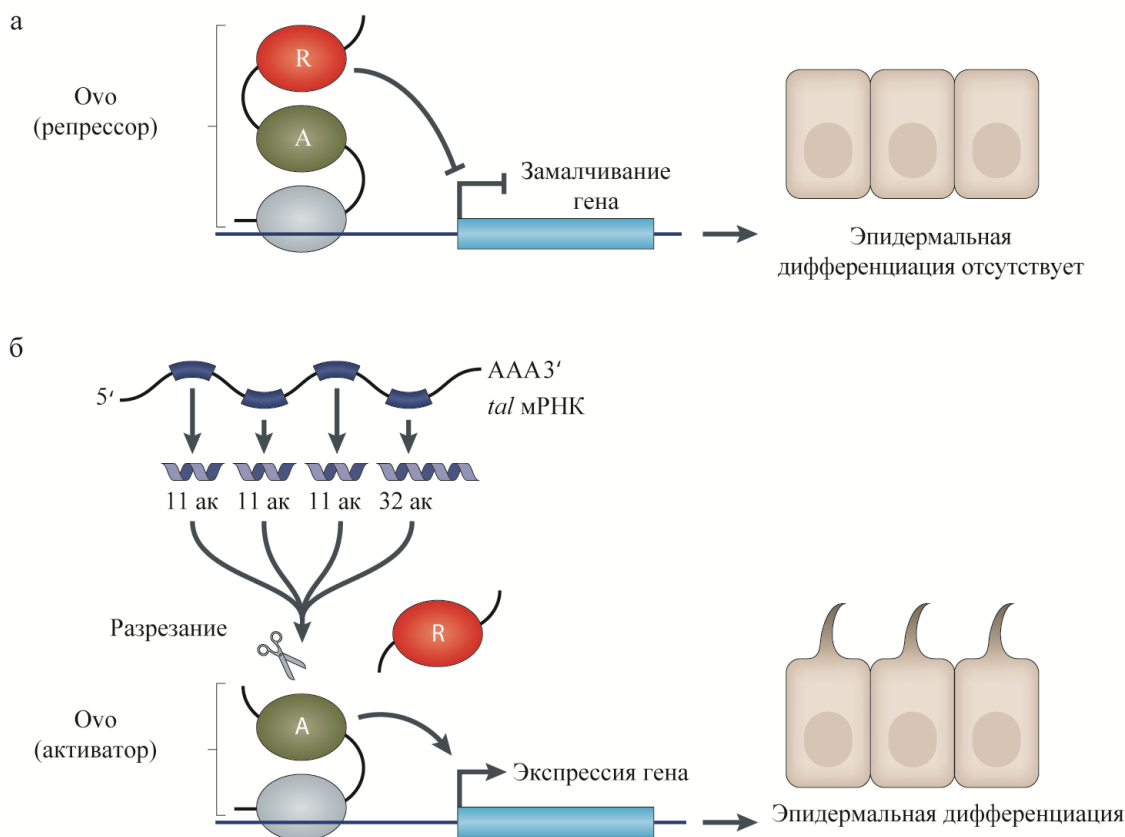


Рисунок 6. Пептиды *Tal* регулируют транскрипционный фактор *Ovo* [144]. (а) Полноразмерный белок *Ovo*, который является транскрипционным фактором, связывается с промотором целевого гена и репрессирует его транскрипцию. Отсутствие экспрессии данного гена препятствует дифференциации эпидермальных клеток личинки и образованию трихом. (б) *Tal* РНК кодирует четыре кОПС-пептида, которые экспрессируются в эпидермальных клетках. Эти пептиды вызывают разрезание белка *Ovo*, что приводит к удалению N-концевого репрессорного домена. Укороченная форма *Ovo* активирует экспрессию целевого гена, что приводит к образованию трихом. R (N-концевой репрессионный домен).

Для Ca^{2+} -АТФазы (SERCA), находящейся в саркоплазматическом ретикулуме, было найдено несколько пептидов, регулирующих ее функцию. Транскрипты миорегулин (MLN) и DWORF изначально были аннотированы как днРНК. Однако оказалось, что они кодируют пептиды, обладающие

трансмембранным доменом, длиной 46 [138] и 34 ак [145] соответственно. MLN пептид очень схож с трансмембранными пептидами фосфоламбан (PLN) и сарколипин (SLN), которые взаимодействуют с SERCA и регулируют ее активность переноса Ca^{2+} , что очень важно для работы мышц [146]. Также как и гомологичные ему пептиды, MLN также локализуется в саркоплазматической мембране и взаимодействует с SERCA, таким образом регулируя концентрацию Ca^{2+} через ингибирование активности Ca^{2+} -АТФазы. Пептид DWORF наоборот активирует SERCA. Недавние биоинформатические исследования обнаружили два новых ингибирующих SERCA пептида, эндорегулин (ELN) и новый регулин (ANL), которые кодируются также днРНК [147]. Также как и в случае MLN, PLN, SLN пептидов, ELN и ALN имеют SERCA связывающий мотив (Рисунок 7). Однако в отличие от первых трех, эти два пептиды представлены в не мышечных тканях, предполагая их участие в общем механизме регуляции внутриклеточной динамики Ca^{2+} .

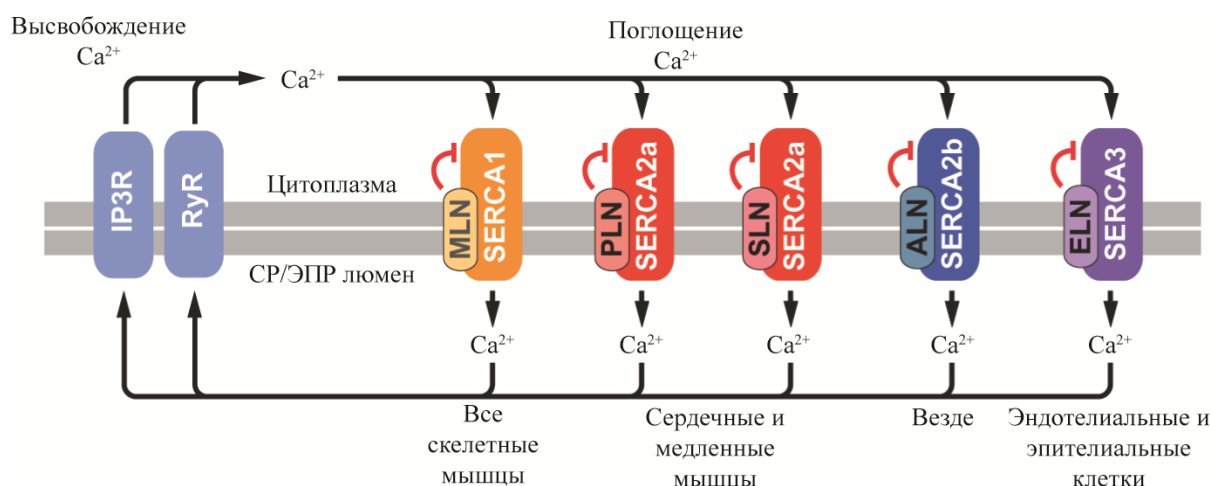


Рисунок 7. Семейство пептидов, ингибирующих SERCA [147]. Модель показывает паттерны экспрессии основных видов SERCA и кОПС-пептидов, которые ингибируют соответствующий вид транспортера, в различных типах мышц и других тканях позвоночных.

Существует множество доказательств того, что многие транскрипты, считавшиеся некодирующими, на самом деле кодируют пептиды с важными клеточными функциями. Но все эти примеры демонстрируют, что днРНК являются источником многих еще не открытых функциональных пептидов.

Циркулярные РНК

Недавние исследования показали, что в эукариотических клетках существует большое количество циркулярных РНК (цРНК) [148-151]. Этот новый класс РНК принимает участие в регуляции экспрессии генов, связывая микроРНК и регулируя транскрипцию и сплайсинг [152]. Функция цРНК не ограничивается вышеперечисленным, так как недавно было показано, что цРНК способны кодировать белки [153,154]. Было замечено, что ряд цРНК ассоциированы с рибосомами, и существование продуктов трансляции удалось показать экспериментально. Одним из таких примеров является пептид, кодируемый цРНК *circMbl1* [153]. Пептид направляется в синапсы, указывая на то, что взаимодействие между нейронами может затрагивать еще неизученные механизмы, основанные на кОПС пептидах. Более того, трансляция этой цРНК индуцируется при голодании и активации различных сигнальных путей, вовлеченных в старение.

Пре-микроРНК

Поскольку микроРНК транскрибируются как большой первичный транскрипт с последующим превращением в активную форму, длинный предшественник возможно обладает кодирующим потенциалом. Также на это может указывать тот факт, что микроРНК транскрибируются РНК полимеразой II. Возможность трансляции предшественников микроРНК была доказана. Пре-микроРНК *miR171b*, *miR165a* и *TAS* различных видов растений, как оказалось, кодируют пептиды [155,156]. Пептиды *miPER171b* и *miPER165a* усиливают накопление соответствующей активной формы микроРНК, в результате чего подавляется экспрессия генов, вовлеченных в развитие корневой системы. Аналогичный эффект исследователи получили, когда ввели экзогенные пептиды *miPER171b* и *miPER165a* в растения. Накопление микроРНК *miR171b* и *miR165a* приводило к ингибированию развития бокового корня и стимуляции роста основного корня. Для пяти других пре-микроРНК была показана способность кодировать пептиды, что показывает распространенность этого явления у растений. Наличие подобного явления в клетках млекопитающих пока остается неизученным [155].

Несколько исследований обнаружили интересный факт, что рибосомная РНК (рРНК), в особенности митохондриальная, обладает кодирующим потенциалом. Полипептид, высококонсервативный среди различных видов, закодирован в митохондриальной 16S рРНК [12]. До сих пор неясно, транслируется ли он в цитоплазме или митохондриях, но было показано, что он биологически активен при синтезе с использованием как митохондриального, так и цитоплазматического кода [157,158]. Это первый митохондриальный пептид, описанный как антиапоптотический. Взаимодействуя с белком Вах (Bcl-2-associated X), пептид предотвращает активацию Вах и смерть клетки [159,160]. С тех пор для этого пептида было показано, что он вовлечен в различные биологические процессы, такие как апоптоз, иммунный ответ, выживаемость клеток, метаболизм и ответ на окислительный стресс [161,162]. Стоит отметить, что с годами количество этого пептида уменьшается, что служит фактором риска развития многих заболеваний, включая болезнь Альцгеймера, атеросклероз, ишемическую болезнь сердца и инсульт, диабет II типа.

Позже было доказано, что митохондриальная 12S рРНК также кодирует пептид. Пептид был назван MOTS-c (mitochondrial open reading frame of the 12S rRNA-c) [8]. Он принимает участие в регуляции чувствительности к инсулину и энергетического гомеостаза, влияя на фолатный цикл и синтез пуринов *de novo*.

Методы исследования и характеристики кОРС-пептидов

Недавние исследования показали, что функционально важные пептиды представлены во многих организмах. Большинство из них высококонсервативны, а область их действия различна. Они выполняют как внутри-, так и внеклеточные функции. Несколько хорошо изученных пептидов демонстрируют, что кОРС-пептиды являются важными регуляторами многих фундаментальных процессов (Рисунок 8), таких, как развитие [11,19,163,164], репарация ДНК [165], декепирование РНК [166], гомеостаз кальция [138,145,147], метаболизм [8], стресс [139], клеточная смерть [12,159].

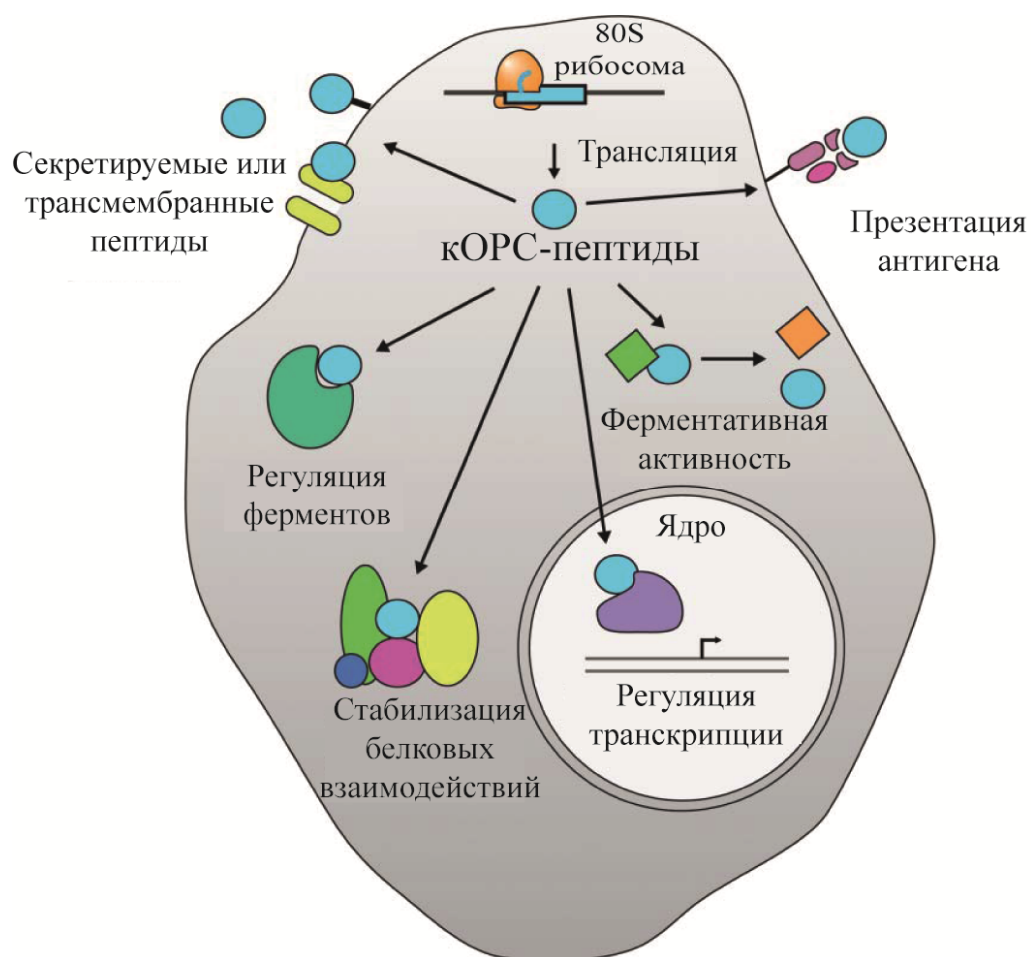


Рисунок 8. Различные функции, выполняемые кОРС-пептидами [167]. кОРС кодируют пептиды с различными внутриклеточными и внеклеточными функциями. Они могут принимать участие в процессах регуляции активности ферментов, стабилизации белок-белковых взаимодействий в цитоплазме и регуляции транскрипции. Также кОРС-пептиды принимают участие в иммунных процессах и передаче сигнала между клетками. Синим шариком обозначен кОРС-пептид.

Хотя последние достижения в вычислительной биологии и экспериментальных методах позволили идентифицировать огромное число кОРС-пептидов, это не означает, что все они обладают биологической функцией. Для того чтобы определить физиологическую роль идентифицированного пептида, в каждом случае необходимо отдельное экспериментальное исследование.

Проверка кодирующего потенциала кОРС

Для того чтобы показать, что кОРС действительно кодирует пептид существует несколько методов. Первым методом, наиболее часто применяемым, является *in vitro* трансляция (Рисунок 9). Для этого эксперимента полноразмерная

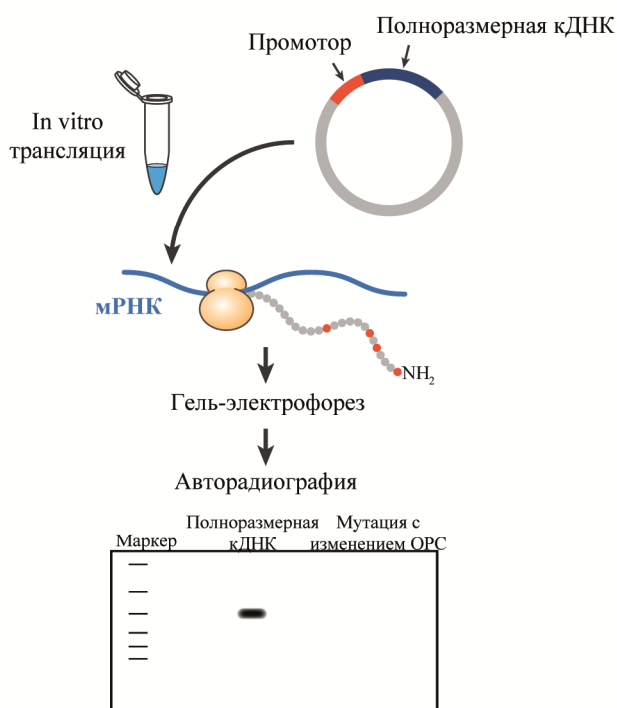
кДНК исследуемого гена клонируется в вектор и затем экспрессия данной конструкции оценивается, используя бесклеточную трансляционную систему в присутствии ^{35}S -метионина. Белковые продукты затем анализируются геле-электрофорезом и с помощью авторадиографии визуализируются радиоактивно-меченные продукты предсказанной длины [138,168,169]. Этот метод основан на том, что рибосома может различить кодирующие и не кодирующие транскрипты, данный факт подтверждается рядом исследований [56,170]. Хотя данный подход может быть информативным и полезным для большого скрининга потенциально кодирующих кОРС, результаты необходимо интерпретировать с аккуратностью, так как трансляция последовательности *in vitro* не говорит о том, что она будет транслироваться *in vivo*. Если введение мутации в исследуемую кОРС приводит к исчезновению пептида, это убедительно доказывает существования продукта трансляции именно данного гена.

Для подтверждения существования пептида можно получить и использовать для детекции специфические антитела [137,139,141,145]. Процесс создания эффективных антител против маленьких пептидов является сложной задачей по нескольким причинам. Во-первых, из-за их размера, что приводит к маленькому набору пептидов оптимальной антигенности. Во-вторых, многие кОРС-пептиды обладают трансмембранным доменом, который скрывает относительно большую часть и так короткой последовательности, что также ограничивает область дизайна оптимального эпитопа. И, конечно, большинство исследователей сталкивались с проблемой недостаточной чувствительности антител. Например, если ген экспрессируется на очень низком уровне, даже самых высокоаффинных антител может быть недостаточно для детекции.

Наиболее элегантный способ доказать существование пептида – модифицировать кодирующую последовательность в геноме, включив в нее подходящий эпитоп, например, с помощью CRISPR-Cas9 системы (Рисунок 10). Эпитоп может быть интегрирован в геном благодаря гомологичной рекомбинации с использованием ДНК матрицы, содержащий соответствующую конструкцию. Данная стратегия активно применяется в научных исследованиях кОРС-пептидов и хорошо зарекомендовала себя на практике [18,138,153]. Более того, в случаях,

когда мРНК кодирует несколько кОПС, данный подход позволяет идентифицировать предпочтительный продукт трансляции. Модифицированные клетки, полученные таким образом, могут в дальнейшем использоваться для последующих экспериментов по иммунопреципитации, иммуноцитохимии и вестерн-блоттингу.

а



б

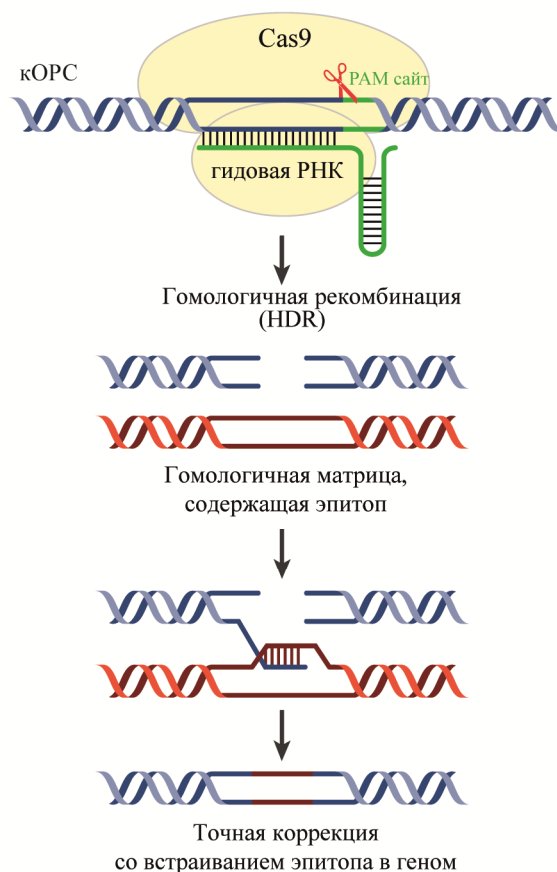


Рисунок 9. Методы доказательства экспрессии кОПС-пептида [171]. (а) Кодированный потенциал кОПС может быть проанализирован с помощью *in vitro* трансляции. Полноразмерная кДНК пептида должна быть клонирована в вектор, содержащий сильный промотор. Трансляция осуществляется в присутствии радиоактивного ³⁵S-метионина, который встраивается в транслирующийся пептид. После чего продукты трансляции анализируются с помощью гель-электрофореза с последующей авторадиграфией. В качестве контроля может быть использована кДНК с преждевременным стоп-кодоном. (б) Схема показывает принцип работы CRISPR-Cas9 системы для таргетного и точного изменения генома. В данном примере ввода эпитопа в геном. Cas9 эндонуклеаза делает двуцепочечный разрыв в месте генома, которое определяется последовательностью гРНК. После этого двуцепочечный разрыв может

быть репарирован двумя путями: с помощью гомологичной рекомбинации, либо путем негомологичного соединения концов. Клетки с точным встраиванием эпитопа будут продуцировать исследуемый белок с довеском, который можно будет детектировать с помощью вестерн-блоттинга.

Хотя данный подход имеет массу преимуществ, существует несколько трудностей, с которыми могут столкнуться исследователи. Во-первых, необходимо получить и протестировать N- и C-тагированные формы пептида. Эпитоп на C-конце с меньшей вероятностью будет мешать сигнальным последовательностям на N-конце, но эпитоп на N-конце часто приводит к лучшей растворимости белка [172]. Внутренние эпитопы также могут быть использованы, особенно когда N- и C-тагированные формы белка ведут себя по-разному. Кроме положения эпитопа, также стоит учитывать его аминокислотный состав. В некоторых случаях эпитоп может быть больше, чем исследуемый пептид. В этих случаях следует учитывать длину и биохимические свойства тага (заряд и гидрофобность).

Поиск биологической функции кОРС-пептида

Доказав существование пептида, можно переходить к поиску его биологической роли. На данный момент, исследования в этой области показывают, что пептиды, как правило, являются регуляторами больших белков или целых комплексов, таким образом, ключевым моментом в изучении их функции является идентификация белок-белковых взаимодействий. Идентификация белковых партнеров позволяет исследователю понять, в каком окружении находится кОРС-пептид, таким образом можно предположить его возможную функцию. Основным методом, который применяется для данных целей, это ко-иммунопреципитация. Функциональная протеомика была успешно использована для определения белков, взаимодействующих с кОРС-пептидами [139,166]. Ко-иммунопреципитация может быть осуществлена с помощью высокоаффинных антител на исследуемый пептид. Однако исследователи часто сталкиваются с тем, что антитела, которые хорошо работают для вестерн-блоттинга, могут не работать в экспериментах с ко-иммунопреципитацией. Существует несколько подходов для решения данной проблемы. Самый простой – это экзогенная экспрессия кОРС-пептида с подходящим тагом. В данном случае иммунопреципитация может быть проведена с использованием коммерчески доступных антител, что значительно упрощает

задачу исследователя. Однако, суперэкспрессия белка любого размера может привести к изменению его локализации и набора взаимодействующих партнеров и как следствие мешать определению его природной функции в клетке [173]. Другой подход подразумевает использование CRISPR-Cas9 системы для редактирования последовательности гена пептида в его природном геномном контексте. Как было выше сказано, с помощью данной системы можно вставить необходимый эпитоп в локус генома, кодирующий кОПС-пептид, что также позволит проводить иммунопреципитацию, используя коммерчески доступные антитела.

Недавно был разработан ряд эффективных методов, позволяющих осуществлять выявление белок-белковых взаимодействий в живой клетке, в том числе слабых, с помощью маркировки белков, опосредованной каким-нибудь ферментом [174-177]. Применение может варьироваться от изучения высокоспецифических белок-белковых взаимодействий до оценки полного интерактома исследуемого белка. В данном подходе могут быть использованы следующие ферменты BirA лигаза [174,176,177], APEX [174,175,178], HRP [179,180], которые осуществляют реакцию ковалентного присоединения биотина к близлежащим белкам в клетке (Рисунок 10). Помеченные биотином белки затем могут быть выделены и идентифицированы с помощью масс-спектрометрии или вестерн-блоттинга. Таким образом, в отличие от ко-иммунопреципитации, где взаимодействие изучается после лизиса клеток, данный экспериментальный подход позволяет изучать как сильные, так и слабые взаимодействия белков в интактных клетках, где локализация белков сохраняется.

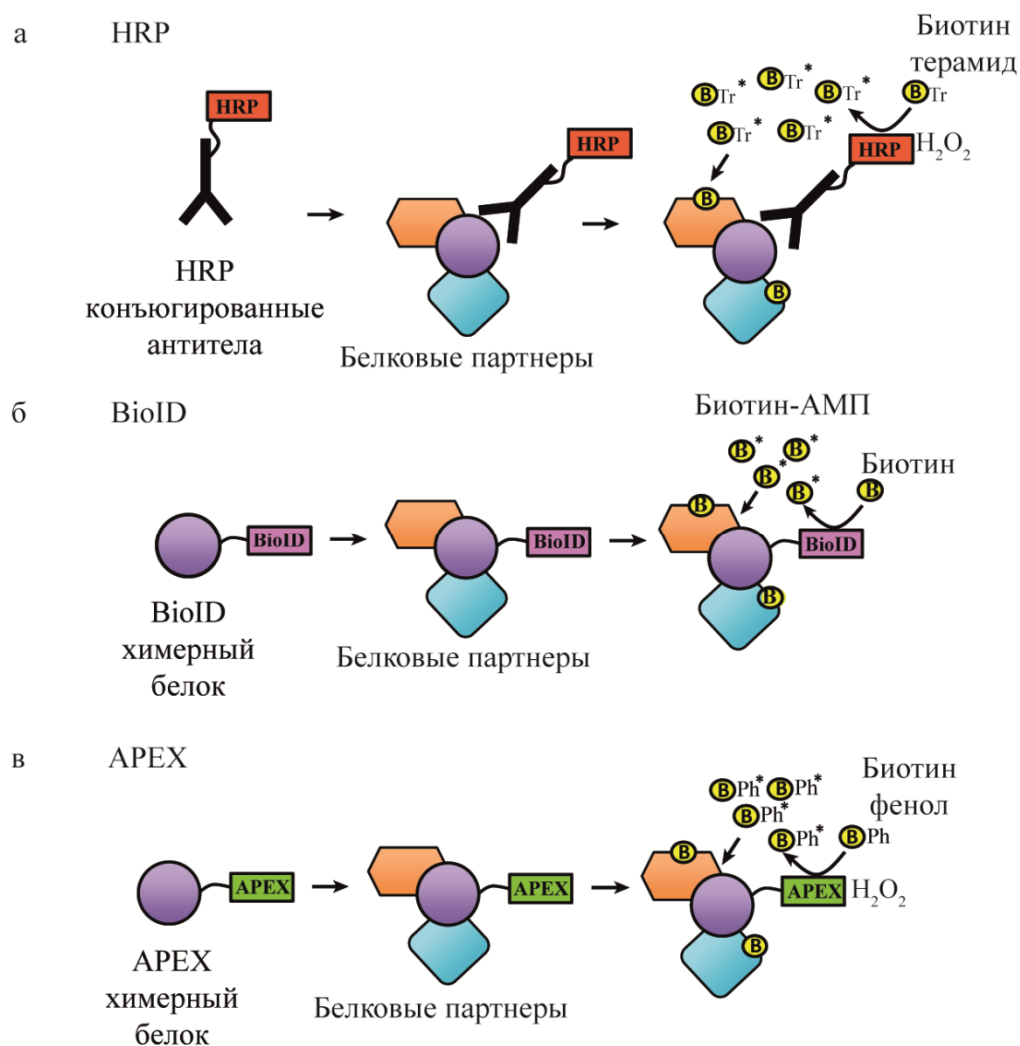


Рисунок 10. Методы маркировки биотином [181]. (а) HRP. Антитела на исследуемый белок конъюгированы с пероксидазой хрена (HRP), которая при добавлении биотин-терамида модифицирует биотином близлежащие белки. (б) BioID. Белок является мутантной версией бактериального белка *BirA*, он может быть использован для эндогенного модифицирования биотином всех близлежащих белков. Химерная форма исследуемого белка и BioID экспрессируется в клетках, которые инкубируются с биотином. Реактивная форма биотина-АМФ, полученная с помощью BioID, модифицирует взаимодействующие белки. (в) APEX. Аскорбат пероксидаза (APEX) также может быть использована для эндогенного модифицирования близлежащих белков. В данном случае ген гибрида исследуемого белка с APEX экспрессируется в клетках. После добавления биотин-фенола и H_2O_2 к клеткам, аскорбат пероксидаза начинает продуцировать реактивную форму биотин-фенола, которая быстро взаимодействует с белками партнерами. Биотин-помеченные белки могут быть выделены и проанализированы с помощью масс-спектрометрии.

Для более детального исследования парных взаимодействий может быть использована простейшая форма маркировки белков биотином, опосредованная ферментом BirA (Рисунок 12).

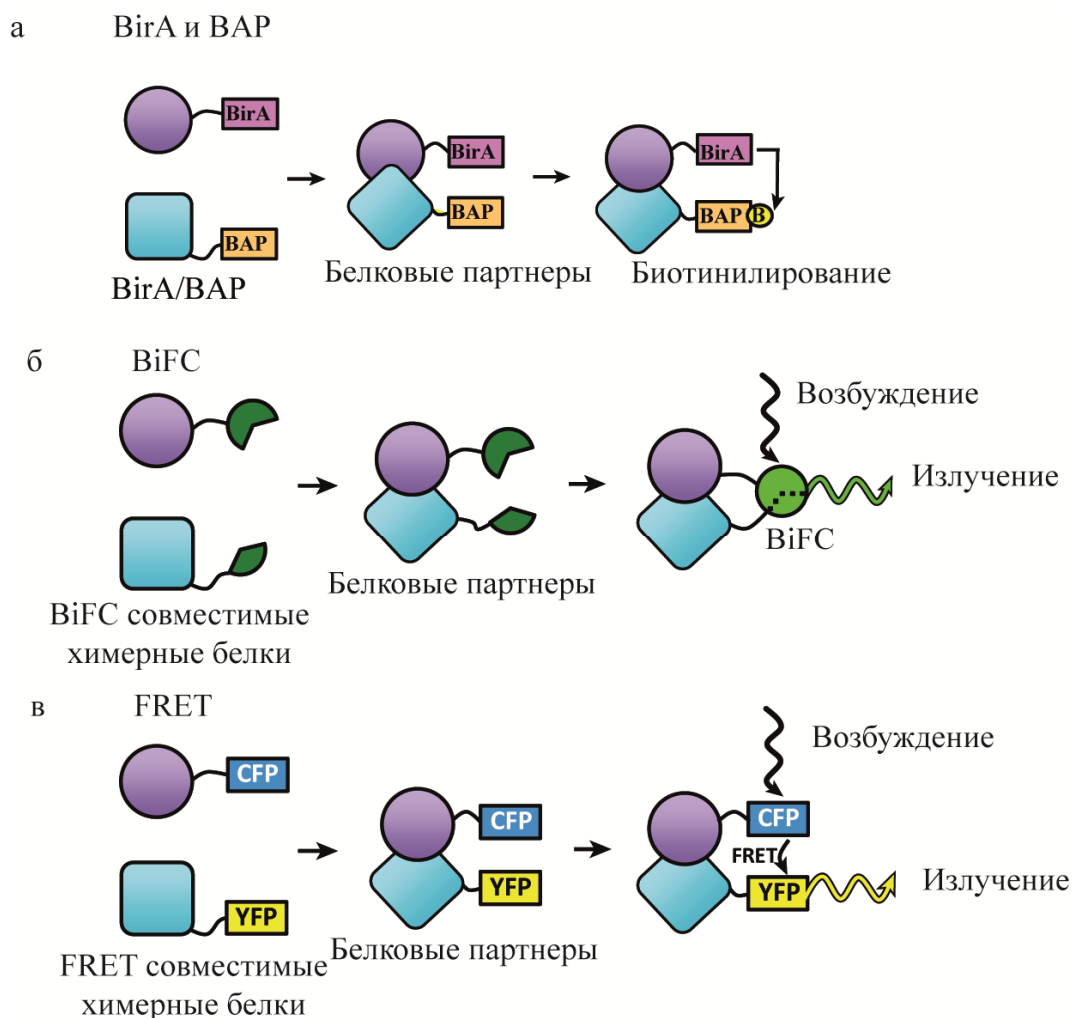


Рисунок 11. Методы подтверждения взаимодействия двух белков [181]. (а) *BirA* и *BAP*. Биотинилаза *BirA* может быть использована для подтверждения взаимодействия двух определенных белков. Один из исследуемых белков должен быть конъюгирован с *BirA*, а другой с пептидом *BAP*, или так называемым *AviTag*. При взаимодействии двух исследуемых белков *BirA* модифицирует *BAP* конъюгированный белок, который затем может быть выделен и проанализирован вестерн-блоттингом. (б) *BiFC*. Данный метод использует комплементацию флюорофоров для подтверждения взаимодействия двух белков. Гены гибридов исследуемых белков, сшитых с нефлюоресцентными фрагментами *BiFC* экспрессируются в клетке. В результате взаимодействия двух белков, фрагменты *BiFC* взаимодействуют и флюоресцируют, что может быть детектировано с помощью микроскопии или цитометра. (в) *FRET*. Может быть применен для исследования

взаимодействия белков в клетках. Гены гибридов исследуемых белков, сшитых с подходящей парой флюорофоров экспрессируются в клетке. При взаимодействии двух белков, близость донора и акцептора приводит к резонансному переносу энергии флюоресценции, приводящему к уменьшению эмиссии донора и увеличению эмиссии акцептора.

Для этого в клетках экспрессируется ген исследуемого белка, слитый с ферментом BirA, и гибрида предполагаемого белкового партнера с акцепторным пептидом биотина BAP (biotin acceptor peptide) [182,183]. Также для исследования взаимодействия пары белков можно использовать такие методы как BiCF (biomolecular fluorescence complementation), где два нефлюоресцентных белка, взаимодействуя друг с другом, образуют флюорофор [184], и FRET (Förster resonance energy transfer), который основан на переносе энергии от возбужденного донора-флюорофора к близлежащему акцептору-флюорофору [185] (Рисунок 11).

Также не менее важным является исследование локализации пептида, которая поможет определить дальнейшее направление исследования. Например, если пептид обладает митохондриальной локализацией, то он вероятно будет каким-то образом влиять на функцию митохондрий. Локализацию пептида можно определить как с помощью антител на пептид, так и с помощью экспрессии тагированной формы пептида.

Заключение

кОПС-пептиды представляют новый класс регуляторных биомолекул. Они играют важную роль в различных процессах от регуляции трансляции и метаболизма, до эмбрионального развития и старения. кОПС кодируются различными видами РНК, в том числе ранее считавшимися некодирующими. На данный момент исследователи обладают рядом инструментов, позволяющих осуществлять изучение кОПС-пептидов. Такие мощные методы как рибосомное профилирование и CRISPR-Cas9 система дают возможность находить новые еще неизученные кОПС-пептиды и изучать их биологическую роль в клетке.

Материалы и методы

Реактивы и биопрепараты

В работе были использованы следующие реактивы и препараты:

Merck: SDS, $MgCl_2$, $CaCl_2$.

Serva: ТЕМЕД, Triton X-100.

Химмед: HCl, CH_3COOH .

Ферейн: C_2H_5OH .

Helicon $NaCl$, $NaOAc$, ЭДТА, Hepes, $(NH_4)_2S_2O_8$, бромфеноловый синий, ксиленцианол, бромистый этидий, Tween 20, акриламид, N,N'-метиленбисакриламид, бакто-агар, агароза для электрофореза, бычий сывороточный альбумин.

Sigma: $MnCl_2$, ампициллин, доксициклин, НА-агароза, НА-пептид, антимицин А, глутаминовая кислота, яблочная кислота, янтарная кислота, малоновая кислота, KCN.

Amresco: среда LB, мочевиная, ДМСО, ДТТ, Tris, таблетки PBS на 100 и 500 мл.

Thermo Scientific: T4 ДНК-лигаза, T4 PNK, PCR Master Mix, генецитин G418, эндонуклеаза рестрикции SfiI и BbsI.

NEB: Q5 ДНК полимеразы.

Евроген: набор реагентов для выделения ДНК из агарозного геля Cleanup mini, набор реагентов для выделения плазмид Plasmid miniprep и олигодезоксирибонуклеотиды.

Promega: набор реагентов для выделения плазмид PureYield Plasmid Midiprep System.

Roche: Complete EDTA-free protease inhibitor cocktail.

Calbiochem: Phosphatase Inhibitor Cocktail.

Gibco: среда DMEM/F12, антибиотики PenStrep, 2.5% трипсин (10x), глутамин GlutaMax (100x).

Serva: дигитонин, ротенон

Буферы и растворы

В работе использовали следующие среды для культивирования бактериальных и эукариотических клеток, буферы и растворы.

Микробиологические среды

LB	1% бакто-триптона, 0.5% дрожжевого экстракта, 171 mM NaCl
Агаризованная среда LB	1.5% бакто-агар, 1% бакто-триптона, 0.5% дрожжевого экстракта, 171 mM NaCl
TB	10 mM PIPES pH 6.7, 15 mM CaCl ₂ , 250 mM KCl, 55 mM MnCl ₂

Растворы для белкового электрофореза

Разделяющий гель	12% смеси полиакриламида и бисакриламида (в соотношении 1:29), 0.375 M Tris-HCl pH 8.8, 0.1% SDS, 0.1% персульфат аммония, 0.1% TEMED
Концентрирующий гель	5% смеси полиакриламида и бисакриламида (в соотношении 1:29), 0.125 M Tris-HCl pH 6.8, 0.1% SDS, 0.1% персульфат аммония, 0.1% TEMED
Трис-глициновый буфер	25 mM Tris, 250 mM глицин, 0.1% SDS
4x буфер для нанесения белковых образцов (Лэммли)	60 mM Tris-HCl pH 6.8, 0.7 M β -меркаптоэтанол, 25% глицерин, 2%

	SDS, 0.1% бромфеноловый синий
--	-------------------------------

Растворы для серебряного окрашивания гелей

Фиксаж	10% уксусная кислота, 20% - этанол
Раствор для сенсibilизации	$\text{Na}_2\text{S}_2\text{O}_3 \times 5\text{H}_2\text{O}$ 0.3 г/л
Раствор нитрата серебра	AgNO_3 1 г/л, формалин 3.7%
Проявитель	Na_2CO_3 40 г/л, $\text{Na}_2\text{S}_2\text{O}_3 \times 5\text{H}_2\text{O}$ 6 мг/л, формалин 3.7%
Раствор для остановки реакции	10% уксусная кислота

Растворы для подготовки образцов для масс-спектрометрии

Буфер для гидролиза	100 мМ NH_4HCO_3 в воде для анализа методом HPLC
Буфер для отмывок	50 мМ NH_4HCO_3 , 40% CH_3CN

Растворы для электрофореза ДНК

ТВЕ	100 мМ Tris-HCl, 100 мМ H_3BO_3 , 2 мМ ЭДТА
Агарозный гель	1% агароза, 100 мМ Tris-HCl, 100 мМ H_3BO_3 , 2 мМ ЭДТА, 0.5 мкг/мл этидийбромид
Буфер для нанесения образцов в агарозный гель	50% глицерин, ТВЕ, 0.1% бромфеноловый синий, 0.1% ксиленцианол

Буферы для иммунопреципитации

Лизирующий буфер	20 mM Hepes pH 7.5, 100 mM NaCl, 0,05% TRITON, 1mM ДТТ, Complete
------------------	--

	EDTA-free protease inhibitor cocktail (1 таблетка на 50 мл), Phosphatase inhibitor cocktail (100x)
Промывочный буфер (PBS)	140 mM NaCl, 2.7 mM KCl, 10 mM Na ₂ HPO ₄ , 1,8 mM KH ₂ PO ₄ , pH 7,3
Элюирующий раствор	300 мкг/мл НА-пептид в PBS

Буферы для ферментов

10x буферный раствор для T4 ДНК-лигазы (Thermo Scientific)	400 mM Tris-HCl pH 7.8, 100 mM MgCl ₂ , 100 mM ДТТ, 5 mM АТФ
2x PCR Master Mix (Thermo Scientific)	0.05 ед./мкл Taq ДНК полимеразы, буфер для реакции, 4 mM MgCl ₂ , 0.4 mM каждого нуклеотида
10x буферный раствор «G» (Thermo Scientific)	100 mM Tris-HCl pH 7.5, 100 mM MgCl ₂ , 500 mM NaCl, 1 мг/мл БСА
10x буферный раствор «A» (Thermo Scientific)	500 mM Tris-HCl pH 7.6°C, 100 mM MgCl ₂ , 50 mM ДТТ, 1 mM спермидин
10x буферный раствор Cut Smart Buffer (NEB)	500 mM KOAc, 200 mM Tris-ацетат, 100 mM Mg(OAc) ₂ , 1 мг/мл БСА

Остальные растворы

TBST	50 mM Tris-HCl pH 7.6, 150 mM NaCl, 0.05% Tween 20
Буфер для выделения ДНК	0.45% Tween 20, 0.45% Triton X100, 2.5 mM MgCl ₂ , 50 mM KCl, 10 mM Tris-HCl pH 8.3, 100мкг/мл протеиназа К
Буфер для пермеабилзации	PBS + 1% Triton X100
Буфер для измерения дыхания	0.137 M NaCl, 5 mM KCl, 0.7 mM NaH ₂ PO ₄ , 1 mM MgCl ₂ , 25 mM Tris-HCl pH 7.5

Праймеры, использованные для клонирования и секвенирования

Праймеры, для внесения и проверки мутаций в старт-кодоне для гена 1500011k16Rik:

Название праймера	Последовательность (5'-3')
L116_ATG_F	CACCGGTGCTGAGTGGTTGCAATGG
L116_ATG_R	AAACCCATTGCAACCACTCAGCACC
Human_U6_Seq_Fw_Insert	ACTATCATATGCTTACCGTAAC
L116_ATG_check_F	GAGTCAGGGAAGTCTGCTTCCTTT
L116_ATG_check_R	CTCAGGCCAGGTCCAGCTTTTTC

Праймеры, для внесения и проверки мутаций в стоп-кодоне для гена 1500011k16Rik:

Название праймера	Последовательность (5'-3')
L116_stop_F	CACCGTCGGGCTGCAGCGCGTGCTC
L116_stop_R	AAACGAGCACGCGCTGCAGCCCGAC
L116_CRend_check_F	GGTTGCAATGGCGGACGTGTCT
L116_CRend_check_R	CTATTTTCTTTCTGAAGCAAGCCATGTTC

Праймеры, для внесения и проверки мутаций в старт-кодоне для гена Cyb5r3:

Название праймера	Последовательность (5'-3')
Cyb5r3_CR1_F	CACCGTCTCTTCTTCGCCACCATGG
Cyb5r3_CR1_R	AAACCCATGGTGGCGAAGAAGAGAC
Cyb5r3_ATG_check_F	CCACTAGCCAGGGCGTGA
Cyb5r3_ATG_check_R	GTCTTCACAAGCGGAAGTGG

Праймеры, для клонирования и секвенирования гомологичной матрицы для замены Gly2 на Ala2 у Cyb5r3:

Название праймера	Последовательность (5'-3')
Cyb5r3_ATG_F	CCACTAGCCAGGGCGTGA

Cyb5r3_ATG_R	TCCGAGGTGTTGAGACCGA
Cyb5r3_ATG_GlytoAla_F	CTTCTTCGCCACCATGGCTGCC
Cyb5r3_ATG_GlytoAla_R	AGACGGTGGCCTCACGCACGG
pUC_Fw	GTAAAACGACGGCCAGT
pUC_Rev	CAGGAAACAGCTATGAC

Праймеры, для получения кДНК Cyb5r3:

Название праймера	Последовательность (5'-3')
Cyb5r3_F	CAGCGTCCAGCCATCAGAA
Cyb5r3_R	AGAGGATGGCACTGGAGTCTG

Праймеры, для клонирования и секвенирования плазмиды pSBtet_L116, pSBtet_HAL116 и pSBtet_L116mCherry:

Название праймера	Последовательность (5'-3')
L116_F	AAAAAAAGGCCTCTGAGGCCATGGCGGACG TGTCTGAGAGGACG
L116_R	AAAAAAAGGCCTGACAGGCCTCAGGCCAGG TCCAGCTTTTTCTG
HAL116_F	AAAAAAAGGCCTCTGAGGCCATGTATCCAT ATGATGTTCCAGATTATGCT
HAL116_R	AAAAAAAGGCCTGACAGGCCTCAGGCCAGG TCCAGCTTTTTCT
L116mCh_F	AAAAAAAGGCCTCTGAGGCCATGGCGGACG TGTCTGAGAG
L116mCh_R	AAAAAAAGGCCTGACAGGCCTTACTTGTAC AGCTCGTCCATGCC
Seq_SB_F	TTATACCTACTTTCCGTACCAC
Seq_SB_R	CCAAACTCATCAATGTATCTTATC

Праймеры, для клонирования и секвенирования плазмиды pSBtet_Cyb5r3, pSBtet_Cyb5r3HA и pSBtet_Cyb5r3eGFP:

Название праймера	Последовательность (5'-3')
-------------------	----------------------------

Cyb5r3_F	AAAAAAAGGCCTCTGAGGCCACATGGGGGC CCAGCTGA
Cyb5r3_R	AAAAAAAGGCCTGACAGGCCTCAGAAGGTG AAGCATCGCTCCTT
Cyb5r3HA_pUC_F	TTCCAGATTATGCTTGAGGGTACCGAGCTCG AATTCAGTGG
Cyb5r3HA_pUC_R	CATCATATGGATAGAAGGTGAAGCATCGCTC CTTGGGATGGCC
Cyb5r3HA_pSB_F	AAAAAAAGGCCTCTGAGGCCACATGGGGGC CCAGCTGA
Cyb5r3HA_pSB_R	AAAAAAAGGCCTGACAGGCCTCAAGCATAA TCTGGAACATCATATGGATAGAAG
Cyb5r3eGFP_pUC_F	GCAGGTCGACTCTAGAGGATCCCCTTAGAAT TCCTTGTACAGCTCGTCCATG
Cyb5r3eGFP_pUC_R	CCAAGGAGCGATGCTTCACCTTCGTGAGCAA GGGCGAGGAG
Cyb5r3eGFP_pSB_F	AAAAAAAGGCCTCTGAGGCCACATGGGGGC CCAGCTGA
Cyb5r3eGFP_pSB_R	AAAAAAAGGCCTGACAGGCCTTAGAATTCC TTGTACAGCTCGTCCATG
Cyb5r3_seq_F	CATTGAATTCCGGGGTCCC
Cyb5r3_seq_R	CTGGTAGACCAGTAGCCCA

Методики, использованные в работе

ПЦР

Составляли следующую смесь для ПЦР с Taq-полимеразой:

Раствор	Конечная концентрация
2x PCR Master Mix	1x
Праймер 1	0,5 мкМ
Праймер 2	0,5 мкМ
Геномная или плазмидная ДНК	100-500 или 1-2 нг
H ₂ O	До нужного объема

Время необходимое для элонгации рассчитывали при условии, что скорость синтеза ДНК Taq полимеразой составляет 1000 нуклеотидов в минуту.

Параметры ПЦР:

Шаг	Температура	Время
Предварительная денатурация	95°C	5 мин
25–35 циклов	95°C от 50 до 72°C 72°C	30 с 30 с 1 мин/кб
Дополнительное время	72°C	10 мин
Конец	4–10°C	

Составляли следующую смесь для ПЦР с Q5-полимеразой:

Раствор	Конечная концентрация
5x Reaction buffer	1x
дНТФ	200 мкМ
5X Q5 High GC Enhancer	1x
Q5 High-Fidelity DNA полимеразы	0.02 U/мкл
Праймер 1	0,5 мкМ
Праймер 2	0,5 мкМ
геномная или плазмидная ДНК	100-500 или 1-2 нг
H ₂ O	До нужного объема

Время необходимое для элонгации рассчитывали при условии, что скорость синтеза ДНК Q5 полимеразой составляет 1000 нуклеотидов за 30 с.

Параметры ПЦР:

Шаг	Температура	Время
Предварительная денатурация	98°C	30 с
25–35 циклов	98°C от 50 до 72°C 72°C	5–10 с 10–30 с 20–30 с/кб

Дополнительное время	72°C	2 мин
Конец	4–10°C	

Приготовление компетентных клеток

Колонию клеток *E. coli* JM109 помещали в 100 мл среды LB и инкубировали при температуре 18°C и перемешивании (125 об/мин) до оптической плотности $A_{600} \sim 0,5$. После этого клетки инкубировали на льду в течение 10 мин и затем скручивали на 5000 об/мин, при температуре 4°C в течение 10 мин. Супернатант сливали и клетки ресуспендировали в 40 мл охлажденного до 4°C буферного раствора ТВ. Данную суспензию инкубировали на льду в течение 10 мин. Затем клетки снова осаждали центрифугированием при 5000 об/мин, при температуре 4°C 10 мин. Далее процедура промывки клеток буфером ТВ повторялась. На последней стадии клетки ресуспендировали в 5 мл буферного раствора ТВ, после чего добавляли к ним до 7% (350 мкл) ДМСО и инкубировали на льду в течение 10 мин. Суспензию клеток алликовтили по 100-200 мкл в 1,5 мл стерильные пробирки и быстро замораживали в жидком азоте. Клетки хранили при температуре –70°C.

Трансформация компетентных клеток

Пробирку с компетентными клетками инкубировали во льду до полного размораживания. Затем к клеткам добавляли 1–5 мкл раствора плазмидной ДНК (~0,1 мкг) или лигазную смесь и инкубировали во льду в течение 15-30 мин. Данную смесь прогревали в термостате при 42°C в течение 45 секунд и затем резко охлаждали во льду. После этого добавляли 700 мкл среды LB и инкубировали при 37°C в течение 1 часа. В случае плазмидной ДНК 100–200 мкл, а в случае лигазной смеси полный объем трансформационной смеси скручивали, супернатант удаляли и в малом объеме LB ресуспендировали клетки. Затем высевали на чашку Петри с твердой средой LB, содержащей соответствующий антибиотик. Инкубировали чашку 16 часов при 37°C.

Выделение плазмидной ДНК

Плазмидная ДНК выделялась с помощью набора реагентов компании Евроген согласно протоколу производителя.

Колонию клеток JM109 *E. coli* помещали в 5 мл среды LB с соответствующим антибиотиком и инкубировали при перемешивании 16 часов при 37°C. Полученную культуру клеток центрифугировали при 6000 об/мин 5 мин. Осадок клеток ресуспендировали в 250 мкл раствора Ресуспендирующего буфера (Евроген), затем клетки лизировали с помощью 250 мкл Лизис буфера (Евроген) и перемешивали до образования вязкого раствора. К полученной суспензии добавляли 350 мкл раствора Нейтрализующего раствора (Евроген), перемешивали до появления белого осадка. Суспензию центрифугировали 10 мин при 14000 об/мин, отбирали супернатант и пропускали его через колонку для выделения ДНК (Евроген). После этого колонку промывали 2 раза по 300 мкл Промывочным раствором (Евроген). Элюировали с ДНК с помощью 30 мкл деионизированной воды.

Большие количества плазмиды выделялись с помощью набора реагентов PureYield Plasmid Midiprep (Promega) и вакуумной системы Vac-Man Vacuum Manifold (Promega). Для этого клетки, содержащие плазмиду, помещали в 50-100 мл среды LB с соответствующим антибиотиком и инкубировали при перемешивании 16 часов при 37°C. Затем клетки центрифугировали при 5000 g 10 мин. Осадок ресуспендировали в 3 мл Ресуспендирующего раствора (Promega), затем лизировали 3 мл Лизирующего раствора (Promega), перемешивали и инкубировали в течение 3 мин при комнатной температуре. В полученный раствор добавляли 5 мл Нейтрализующего раствора (Promega). Получившийся осадок центрифугировали при 15000g в течение 10 мин при комнатной температуре. Далее все растворы пропускали через колонки под вакуумом. В Очищающую колонку (Promega) помещали в Связывающую колонку (Promega) и пропускали через них супернатант. После того как супернатант прошел обе колонки, Очищающую колонку удаляли. Далее Связывающую колонку промывали 5 мл Раствора для удаления эндотоксинов (Promega), а после 20 мл Промывочного раствора (Promega). Мембрану оставляли под вакуумом в течение 1 мин, чтобы дать ей высохнуть. ДНК элюировали 400-600 мкл деионизированной воды, выдерживая под вакуумом используя Eluator Vacuum Elution Device (Promega).

Разделение фрагментов ДНК в агарозном геле

Агарозный гель готовился с учетом длины фрагмента ДНК. Для этого смешивали нужное количество агарозы с соответствующим объемом раствора 1xTBE (1% гель - 1 г агарозы на 100 мл раствора 1xTBE). Далее раствор прогревали до полного растворения агарозы. Расплавленную агарозу охлаждали до ~ 40–50°C и добавляли необходимое количество (до 0.01 %) водного раствора бромистого этидия с концентрацией 10 мг/мл, перемешивали, выливали в специальную форму с гребенками. После застывания геля образцы наносились в ячейки и проводили электрофорез при силе тока ~ 100 мА. Изображение геля получали с помощью камеры с УФ лампой (прибор ChemiDoc (Bio-Rad)).

Выделение фрагментов ДНК из агарозного геля

Выделение фрагментов ДНК из агарозного геля осуществлялось с помощью набора реагентов фирмы Евроген согласно протоколу с использованием фирменных реагентов. Часть агарозного геля, содержащую нужный фрагмент ДНК, растворяли при 55°C в Связывающем буфере (Евроген) в соотношении на 100 мкг геля 300 мкл буфера. Затем получившийся раствор пропускали через колонку для выделения ДНК (Евроген). Колонку промывали 750 мкл Промывочного раствора (Евроген). ДНК элюировали 30 мкл деионизированной воды.

Секвенирование

Все нуклеотидные последовательности анализировали с помощью автоматического секвенатора ABI prism 3100-Avant genetic Analyzer (4-х капиллярный). Для проведения анализа секвенирования образцы отдавали в Центр Коллективного Пользования «Геном» (Институт молекулярной биологии им. В.А. Энгельгардта РАН).

Клонирование гидовых РНК в рх458

На первом этапе необходимо было разрезать 1мкг плазмиды рх458 рестриктазой BbsI, 37°C, 30 мин:

Раствор	Количество
рх458	1 мкг

Рестриктаза BbsI	1 мкл
10X Буфер «G»	2 мкл
H ₂ O	До 20 мкл

Затем разрезанную плазмиду чистили через гель, вырезали из геля нужную зону и выделяли с помощью набора реагентов (Евроген). Праймеры, кодирующие последовательность соответствующей геновой РНК, необходимо было фосфорилировать и отжечь. Для фосфорилирования составляли следующую смесь:

Раствор	Количество
олиг 1 (100 μM)	1 мкл
олиг 2 (100 μM)	1 мкл
10X T4 Ligation Buffer	1 мкл
H ₂ O	6.5 мкл
T4 PNK	0.5 мкл

Далее отжиг проводили по следующей программе:

37°C 30 мин

95°C 5 мин, после охлаждение до 25°C со скоростью 5°C/мин

Смесь для лигирования:

Раствор	Количество
px458, порезанная BbsI	50 нг
Дуплекс праймеров (разбавленный 1:250)	1 мкл
10X T4 Ligation Buffer	1 мкл
T4 DNA Ligase	1 мкл
H ₂ O	до 10 мкл

Лигазную смесь инкубировали при комнатной температуре в течение 5 мин. Затем трансформировали лигазной смесью компетентные клетки JM109 *E. coli*. Полученные колонии скалывали и растили ночь в 2 мл среды LB. Затем клетки центрифугировали, выделяли плазмиды и проверяли встраивание нужной последовательности гидовой РНК секвенированием.

Клонирование кДНК Cyb5r3 в вектор pUC19

Для получения кДНК гена ставили ПЦР с тотальной кДНК клеточной линии NIH₃T₃ с соответствующими праймерами. Проверяли на геле наличие продукта, а затем выделяли ПЦР продукт набором реагентов (Евроген). Полученный продукт кинировали:

Раствор	Количество
ПЦР-продукт	10 мкл
АТФ (10 мМ)	2 мкл
10X Reaction Buffer A for T4 PNK	2 мкл
H ₂ O	3 мкл
T4 PNK	1 мкл

Смесь инкубировали в течение 20 мин при 37°C. Для инактивации фермента смесь инкубировали в течение 10 мин при 75°C. Затем вставку лигировали в вектор pUC по тупым концам. Смесь для лигирования включала следующие компоненты:

Раствор	Количество
pUC19, порезанная SmaI	20-60 нг
Вставка	3х–10х мольный избыток к вектору
10X T4 Ligation Buffer	2 мкл
T4 DNA Ligase	1 мкл

H ₂ O	до 20 мкл
------------------	-----------

Полученную смесь инкубировали при комнатной температуре 5 мин, а затем ею трансформировали компетентные клетки JM109 *E. coli*.

*Клонирование конструкций pSBtet_L116, pSBtet_HAL116,
pSBtet_L116mCherry.*

Ранее в нашей лаборатории были клонированы конструкции, кодирующие HAL116 и L116mCherry в pUC19. Данные последовательности было необходимо переклонировать в вектор для транспозиции pSBtet-Neo (Addgene). Для этого с помощью праймеров, содержащих на концах сайты рестрикции SfiI, были получены ПЦР продукты для L116, HAL116 и L116mCherry.

Обрабатывали исходную плазмиду pSBtet-Neo рестриктазой SfiI. Для этого составили смесь, содержащую следующие компоненты:

Раствор	Количество
pSBtet-Neo	1 мкг
10X Cut Smart Buffer	5 мкл
SfiI	1 мкл
H ₂ O	до 50 мкл

Инкубировали 3-4 ч при 50°C.

Расщепленную эндонуклеазой плазмиду очищали электрофорезом в агарозном геле и выделяли из агарозы с помощью набора реагентов (Евроген). ПЦР продукты, содержащие сайты рестрикции и необходимые химерные последовательности, также обработали рестриктазой SfiI для клонирования по липким концам.

Раствор	Количество
ПЦР продукты	500 нг - 1 мкг

10X Cut Smart Buffer	5 мкл
SfiI	1 мкл
H ₂ O	до 50 мкл

Инкубировали 3-4 ч при 50°C.

Линеаризованный плазмидный вектор и ПЦР продукты лигировали. В лигазную смесь входили следующие компоненты:

Раствор	Количество
pSBtet-Neo, порезанная SfiI	20-60 нг
ПЦР продукт	3х–10х мольный избыток к вектору
10X T4 Ligation Buffer	2 мкл
T4 DNA Ligase	1 мкл
H ₂ O	до 20 мкл

Лигазную смесь инкубировали в течение 5 мин при комнатной температуре, а затем трансформировали компетентные клетки JM109 *E. coli*. Из полученных колоний выделяли плазмиды, анализировали наличие необходимой вставки ПЦР. Отсутствие нежелательных мутаций контролировали с помощью секвенирования.

Клонирование конструкций pSBtet_Cyb5r3, pSBtet_Cyb5r3HA, pSBtet_Cyb5r3eGFP.

Для получения плазмиды, кодирующей химерную последовательность Cyb5r3HA, плазмиду pUC, содержащую кДНК гена Cyb5r3, проводили ПЦР с соответствующими праймерами, которые содержали в качестве довесков последовательность HA-эпитопа. Линеаризованную плазмиду затем замыкали с помощью Т4 ДНК лигазы:

Раствор	Количество
---------	------------

pUC19_Cyb5r3HA	50 нг
10X T4 Ligation Buffer	2 мкл
T4 DNA Ligase	1 мкл
H ₂ O	до 20 мкл

Далее смесью трансформировали компетентные клетки JM109 *E. coli*. Плазмиды, выделенные из клеток, выращенных из полученных колоний, проверяли на присутствие необходимой вставки. Мутации в нуклеотидной последовательности контролировались секвенированием.

Плазмиду, кодирующую химерный белок Cyb5r3eGFP, было решено получить безрестриктным способом клонирования, который был освоен и применен мною в другом исследовании [186]. Для этого ген eGFP был амплифицирован ПЦР с помощью специальных праймеров, содержащих в качестве довесков последовательность исходной плазмиды pUC_Cyb5r3HA. Данный ПЦР продукт почистили с помощью набора реагентов (Евроген). ПЦР продукт был использован в качестве праймера для получения плазмиды pUC_Cyb5r3eGFP с помощью последующей ПЦР. После ПЦР полученную смесь обработали ферментом DpnI для удаления исходной плазмиды и трансформировали ею компетентные клетки JM109 *E. coli*. Плазмиды, выделенные из клеток, выращенных из полученных колоний, проверяли на присутствие необходимой вставки. Мутации в нуклеотидной последовательности контролировались секвенированием.

Далее эти последовательности было необходимо переклонировать в вектор для транспозиции pSBtet-Neo (Addgene). Для этого, с помощью праймеров, содержащих на концах сайты узнавания эндонуклеазой рестрикции SfiI, были получены ПЦР продукты для Cyb5r3, Cyb5r3HA и Cyb5r3eGFP.

Исходную плазмиду и ПЦР продукты обработали рестриктазой SfiI. Для этого составили смесь, содержащую следующие компоненты:

Раствор	Количество
---------	------------

ДНК	500 нг - 1 мкг
10X Cut Smart Buffer	5 мкл
SfiI	1 мкл
H ₂ O	до 50 мкл

Инкубировали 3-4 ч при 50°C.

Порезанную плазмиду очищали через агарозный гель и выделяли с помощью набора реагентов (Евроген), ПЦР продукты также очищали с помощью набора реагентов (Евроген). Порезанную плазмиду и ПЦР продукты лигировали. В лигазную смесь входили следующие компоненты:

Раствор	Количество
pSBtet-Neo, порезанная SfiI	20-60 нг
ПЦР продукт	3х–10х мольный избыток к вектору
10X T4 Ligation Buffer	2 мкл
T4 DNA Ligase	1 мкл
H ₂ O	до 20 мкл

Лигазную смесь инкубировали в течение 5 мин при комнатной температуре, а затем трансформировали ею компетентные клетки JM109 *E. coli*. Из клеток, выращенных из полученных колоний выделяли плазмиды, анализировали наличие необходимой вставки ПЦР. Отсутствие мутаций контролировали с помощью секвенирования.

Клонирование матрицы для гомологичной рекомбинации для замены Gly2 на Ala2 Cyb5r3

Сперва, с помощью ПЦР амплифицировали участок генома, включающий в себя старт-кодон гена Cyb5r3 и плечи гомологии. Полученный ПЦР продукт

очистили с помощью гель электрофореза и набора реагентов (Евроген). Далее ПЦР продукт кинировали:

Раствор	Количество
ПЦР-продукт	10 мкл
АТФ (10 мМ)	2 мкл
10X Reaction Buffer A for T4 PNK	2 мкл
H ₂ O	3 мкл
T4 PNK	1 мкл

Смесь инкубировали в течение 20 мин при 37°C. Для инактивации фермента смесь инкубировали в течение 10 мин при 75°C. Затем вставку лигировали в вектор pUC19 по тупым концам. Смесь для лигирования включала следующие компоненты:

Раствор	Количество
pUC19, порезанная SmaI	20-60 нг
Вставка	3х–10х мольный избыток к вектору
10X T4 Ligation Buffer	2 мкл
T4 DNA Ligase	1 мкл
H ₂ O	до 20 мкл

Полученную смесь инкубировали при комнатной температуре 5 мин, а затем трансформировали ею компетентные клетки JM109 *E. coli*.

Для того чтобы мутировать второй кодон, в праймеры, с которыми был проведен ПЦР, была внесена мутация. Линеаризованную плазмиду затем замыкали с помощью Т4 ДНК лигазы:

Раствор	Количество
pUC19_Cyb5r3_Gly2toAla2	50 нг
10X T4 Ligation Buffer	2 мкл
T4 DNA Ligase	1 мкл
H ₂ O	до 20 мкл

Далее смесью трансформировали компетентные клетки JM109 *E. coli* и плазмиды, выделенные из клеток, выращенных из полученных колоний, проверяли на присутствие необходимой вставки. Мутации в нуклеотидной последовательности контролировались секвенированием.

Проверка эффективности CRISPR-Cas9 с помощью T7-эндонуклеазы

Для проверки эффективности разрезания геномной ДНК системой CRISPR-Cas9 ставили ПЦР с праймерами, окружающими место разреза. Полученный ПЦР продукт регидризовали по следующей схеме:

Шаг	Температура	Скорость охлаждения	Время
Денатурация	95°C		5 минут
Отжиг	95-85°C	-2°C/секунду	
	85-25°C	-0,1°C/секунду	
Хранение	12°C		

После этого смешивали следующие компоненты:

Раствор	Количество
10x MgCl ₂	3 мкл
ПЦР продукт	26 мкл
T7-эндонуклеаза	1 мкл

Инкубировали в течение 40 мин при 37°C. Реакцию останавливали добавлением 1,5 мкл 0,25 М ЭДТА и наносили на 3% агарозный гель.

Белковый электрофорез по Леммли

Белковые образцы анализировали методом электрофореза в денатурирующем полиакриламидном геле (в присутствии 0,1% SDS). В кассете для заливки ПААГ Mini-Protean II (Bio-Rad) сначала полимеризовали разделяющий гель, процентность которого варьировали от 8% до 12% в зависимости от размера исследуемого белка. Затем поверх разделяющего геля в камеру заливали концентрирующий гель, в котором формировались ячейки для образцов с помощью “гребенки”. К образцам предварительно добавляли 5х буфер для нанесения и выдерживали их при 95°C в течение 7 мин. На первом этапе электрофорез проводили в трис-глициновом буфере при 70 В, как только образцы пересекали границу концентрирующего и разделяющего геля – при 130 В. Электрофорез останавливали, когда бромфеноловый синий достигал конца геля. После проведения электрофореза гель окрашивали нитратом серебра либо делали Вестерн-блоттинг.

Окрашивание белковых гелей нитратом серебра

Гель предварительно инкубировали в фиксаже не менее часа. Затем его отмывали в деионизованной воде 5 раз по 15 мин. После промывок добавляли раствор для сенсibilизации и инкубировали в течение 2 мин. Далее гель промывали деионизованной водой 3 раза по 1 мин. В растворе нитрата серебра гель инкубировали в течение 10-15 мин. Затем раствор нитрата серебра удаляли и промывали гель деионизованной водой 3 раза по 30 с. После добавления проявителя инкубировали до начала фоновое окрашивания. Затем проявитель быстро удаляли и останавливали реакцию добавлением 10% уксусной кислоты. Все процедуры проводили при постоянном перемешивании.

Трипсинолиз белковых образцов для анализа методом MALDI-MS

Зону, содержащую белок нужной молекулярной массы, вырезали из геля, окрашенного серебром, и помещали в чистую пробирку. Кусочек геля промывали буфером для отмывки несколько раз, инкубируя при комнатной температуре 20 мин. После этого к кусочку геля добавляли ацетонитрил и инкубировали в течение

5 мин при комнатной температуре. Затем ацетонитрил убирали, а гель высушивали на воздухе. Образец инкубировали с трипсином в течение 4 ч при 37°C в буфере для гидролиза. Реакцию останавливали 0,5 % ТФУ

Масс-спектрометрический анализ образцов выполняла д.х.н. Серебрякова М. В.

Вестерн-блоттинг

После ПААГ электрофореза в денатурирующих условиях белковые образцы переносили на PVDF мембрану методом полусухого переноса. В качестве буфера для переноса использовали трис-глициновый буфер с добавлением 20% этанола. Затем мембрану инкубировали в буфере TBST с 5% БСА при комнатной температуре в течение 1-2 часов при постоянном перемешивании. С первичными антителами мембрану инкубировали ночь при 4°C. Далее делали 3 промывки по 10 мин буфером TBST. Инкубацию со вторичными антителами делали при перемешивании при комнатной температуре 1-2 часа в TBST, содержащем 5% БСА и рекомендованное производителем или экспериментально подобранное количество антител. Промывали мембрану от не связавшихся антител 5 раз по 2 мин буфером TBST. Для визуализации зон анализируемых белков добавляли к мембране хемилюминесцентный агент-субстрат для пероксидазы (Amersham ECL reagents, компания GE Healthcare Life Sciences). После этого мембрану переносили на стекло и фиксировали люминесценцию с помощью прибора ChemiDoc XRS System (Bio-Rad).

Следующие антитела были использованы в работе:

1. кроличьи поликлональные anti-L116 (Eurogentec);
2. кроличьи поликлональные anti-Cyb5r3 (Santa Cruz Biotech, sc-12040);
3. кроличьи поликлональные anti-Tom20 (Sigma-Aldrich, HPA011562);
4. козьи поликлональные anti-β-актин (Abcam, ab8229);
5. крысиные моноклональные anti-НА (клон 3F10) (Sigma, 11867423001);
6. кроличьи поликлональные anti-mCherry (Abcam, ab183628);
7. кроличьи поликлональные anti-Park7 (Abcam, ab18257).

Культивирование эукариотических клеток

Эукариотические клетки культивировались в среде, содержащей следующие компоненты:

1. среда DMEM/F12 (Gibco) 450 мл для NIH3T3 и RPMI 1640 (Gibco) для NSO;
2. эмбриональная телячья сыворотка FBS (Gibco) 50 мл (финальная концентрация 10%);
3. раствор стрептомицина (5 мг/мл) с пенициллином (5000 ед/мл) (Gibco) 10 мл;
4. раствор GlutaMAX 100x (Gibco) 5 мл.

Клетки выращивались в CO₂-инкубаторе (Sanyo) при 37°C и содержании CO₂ 5%. При пересеве NIH3T3 клетки сперва промывались PBS, а затем обрабатывались трипсином, содержащим ЭДТА, в течение 5-8 минут. Затем к клеткам добавляли среду DMEM/F12, как только они открепилась от поверхности флакона. Для пересева NSO клетки отбирали и разводили в свежей среде в соотношении 1:5 или 1:10. Количество клеток в объеме считали помощью камеры Горяева или прибора Scepter 2.0 (Millipore).

Получение модифицированных клеточных линий с помощью CRISPR-Cas9 системы

Для получения нокаутных клеточных линий использовали отработанный ранее метод [187]. Клетки трансфецировали плазмидами рх458, кодирующими Cas9 и гидовую РНК для N- и/или С-конца ОРС. Для получения клеточных линий с точным встраиванием или мутацией, клетки трансфецировали плазмидой рх458 и гомологичной матрицей, содержащей необходимую мутацию. В качестве трансфецирующего реагента использовали Lipofectamine 3000 (ThermoFisher). Адгезивные клетки за день до эксперимента высевали на 6-ти луночную плашку так, чтобы конfluence на следующий день составляла 70%. Для трансфекции готовили две смеси:

1. смешивали 120 мкл Opti-Mem (Gibco) и 5 мкл Lipofectamine 3000;
2. 120 мкл Opti-Mem (Gibco), 5 мкл реагента P3000 и 2500 нг соответствующей плазмиды.

Эти два раствора смешивали в соотношении 1:1 и инкубировали 5-20 мин при комнатной температуре. Далее клеткам меняли среду на Opti-Mem без антибиотика, содержащий 10% FBS. Затем к клеткам добавляли ДНК-липидный комплекс и инкубировали ночь при 37°C в CO₂-инкубаторе.

На следующий день клетки анализировали на наличие зеленой флюоресценции с помощью клеточного сортера Becton Dickinson FACSAria III. Клетки сортировали на моноклоны в 96-луночные плашки. Сортировку выполняла к.х.н. Рубцова М. П.

Получение модифицированных клеточных линий с помощью транспозиции

Клетки трансфецировали плазмидами pSBtet-Neo, содержащими необходимую конструкцию, и pCMV(CAT)T7100x (Addgene). Адгезивные клетки за день до эксперимента высевали на 6-ти луночную плашку так, чтобы конфлюэнтность на следующий день составляла 70%. Для трансфекции готовили две смеси:

1. смешивали 120 мкл Opti-Mem (Gibco) и 5 мкл Lipofectamine 3000;
2. 120 мкл Opti-Mem (Gibco), 5 мкл реагента P3000 и 2500 нг соответствующей плазмиды.

Эти два раствора смешивали в соотношении 1:1 и инкубировали 5-20 мин при комнатной температуре. Далее клеткам меняли среду на Opti-Mem без антибиотика, содержащий 10% FBS. Затем к клеткам добавляли ДНК-липидный комплекс и инкубировали ночь при 37°C в CO₂-инкубаторе. Затем клетки, встроившие конструкцию в геном, отбирали выращивая на среде с антибиотиком G418 в течение 7 дней.

Проверка моноклонов

Часть клеток отбирали на выделение ДНК, а другую часть высевали в свежую среду до анализа результатов. Клетки центрифугировали при 4°C в течение 10 минут при 1000 об/мин. Затем промывали их PBS и центрифугировали при 4°C в течение 10 минут при 1000 об/мин. Затем добавляли к ним буфер для выделения ДНК и сначала инкубировали при 65°C в течение 15 мин, а затем для инактивации протеиназы К инкубировали при 98°C 2 мин. Полученный раствор

использовали для ПЦР. ПЦР продукты анализировали либо электрофорезом в агарозном геле, либо секвенированием.

Иммунопреципитация

Для иммунопреципитации клетки растили на 15см чашках в 30 мл питательной среды DMEM/F12, содержащей все необходимые дополнительные компоненты. Для эксперимента убирали среду, а затем промывали клетки 10 мл PBS. После чего добавляли 3 мл нового раствора PBS и с помощью специального шпателя соскребали клетки и собирали их. Затем клетки центрифугировали при 4°C в течение 10 минут при 1000 об/мин. Далее клетки ресуспендировали в лизирующем буфере (в соотношении 1 грамм клеток/5 мл буфера) и инкубировали в течение 30 мин. Далее осаждали центрифугированием дебрис на максимальных оборотах при 4°C в течение 30 мин. Получившийся супернатант переносили в чистую пробирку и измеряли концентрацию белка методом Бредфорд. Небольшую часть экстракта отбирали для анализа на денатурирующем ПААГ.

На 1 г клеток брали 50-100 мкл хорошо перемешанной суспензии НА-агарозы. Смолу промывали 2 раз в 1 мл раствора PBS и 2 раза 1 мл Лизирующего буфера. Супернатант, полученный из лизата клеток, смешивали с необходимым количеством промытой смолы и инкубировали в течение ночи при 4°C и постоянном перемешивании. На следующий день инкубационную смесь центрифугировали для осаждения смолы. Все, что не связалось со смолой, отбирали, а смолу промывали 5 раз Промывочным буфером. Белки, связавшиеся со смолой, элюировали сначала раствором НА-пептида в течение 10 минут при комнатной температуре, а потом Буфером для нанесения при 95°C, 5 минут.

Иммуноцитохимия и конфокальная микроскопия

Для иммуноцитохимии клетки высевали на покровные стёкла в 6-ти луночную плашку и оставляли на ночь в инкубаторе при 37 °C и содержании CO₂ 5%. На следующий день клетки, промывали 3 раза PBST и фиксировали 4% раствором параформальдегида в PBS при комнатной температуре в течение 10 мин. Далее, промывали стёкла PBST 3 раза по 5 мин, после пермеабилizировали клетки с помощью буфера для пермеабилizации и инкубировали 15 мин при комнатной температуре. Затем также 3 раза промывали стекла PBST по 5 минут.

Затем клетки инкубировали с раствором, содержащем 3% БСА и 0.1% Triton X-100 в PBS, в течение 1 ч при комнатной температуре. Окрашивание антителами проводили по следующей схеме:

1. инкубировали стёкла с первичными антителами, растворенными в PBS с 0.1% Triton X-100, в течение ночи при 4°C;
2. далее промывали стёкла от первичных антител: 0.1% Triton X-100 в PBS, 5 раз по 5 мин;
3. инкубировали стёкла со вторичными антителами, растворенными в PBS с 0.1% Triton X-100 в течение 1-2 ч при 4°C;
4. стекла промывали от вторичных антител: 0.1% Triton X-100 в PBS, 5 раз по 5 мин;
5. для окрашивания ДНК стекла с клетками инкубировали с 100 мМ DAPI в PBS 5 мин;
6. стекла промывали от DAPI: 0.1% Triton X-100 в PBS 3 раза по 5 мин;
7. в конце заключали стёкла в раствор Moviol, высушивали ночь.

Иммуноцитохимические препараты анализировали с помощью флуоресцентного микроскопа Nikon C2.

Для конфокальной микроскопии клетки высевали на специальные 35 мм чашки со стеклянным дном (MatTek) за 1-2 дня до эксперимента. Затем клетки промывали средой Opti-Mem (Gibco) без красного фенолового. Далее инкубировали клетки в среде Opti-Mem, содержащей необходимый краситель, в CO₂-инкубаторе при 37°C в течение 15 мин.

В работе были использованы следующие красители:

1. для окрашивания митохондрий в зеленый цвет - MitoTracker Green (ThermoFisher);
2. для окрашивания митохондрий в красный цвет - MitoTracket DeepRed (ThermoFisher);
3. для окрашивания ядра в синий цвет - Dapi (ThermoFisher).

Изображение клеток получали и анализировали с помощью конфокального микроскопа Carl Zeiss LSM700.

Измерение клеточного дыхания

Для эксперимента клетки растили на 15см чашках в 30 мл питательной среды DMEM/F12, содержащей все необходимые дополнительные компоненты. Для эксперимента убирали среду, а затем промывали клетки 10 мл PBS. После чего добавляли 3 мл нового раствора PBS и с помощью специального шпателя соскребали клетки и собирали их. Количество клеток подсчитывали с помощью камеры Горяева. Затем клетки центрифугировали при 4°C в течение 10 минут при 1000 об/мин. Далее клетки ресуспендировали в буфере для измерения клеточного дыхания так, чтобы в 1 мл содержалось $4 \cdot 10^6$ клеток.

Перед измерением клетки инкубировали в течение 5 мин при комнатной температуре с дигитонином (10 мг/мл), конечная концентрация 0.02 мг/мл. Затем клетки помещали в ячейку прибора Hansatech Oxygraph Plus System и измеряли скорость поглощения кислорода по следующему протоколу:

Шаг	Концентрация	Время
Na-соль глутаминовой и яблочной кислоты	10 mM/2.5 mM	3-5 мин
Ротенон	2 мкМ	3-5 мин
Na-соль янтарной кислоты	10 mM	3-5 мин
Na-соль малоновой кислоты	2 mM	3-5 мин
Глицерол-3-фосфат	30 mM	3-5 мин
Антимицин А	2 мкМ	3-5 мин
Na-соль аскорбиновой кислоты	0.5 mM	3-5 мин
KCN	1 mM	3-5 мин

Результаты и обсуждение

Постановка задачи

Короткие пептиды, кодируемые в геномах большинства организмов, играют важную роль в различных клеточных процессах. В связи с их малым размером они зачастую пропускаются при автоматической аннотации генома. Таким образом, многие РНК, которые кодируют кОРС-пептиды, были ошибочно аннотированы как некодирующие. В данной работе объектом исследования был выбран ген 1500011k16Rik, который был аннотирован как днРНК и содержал небольшой консервативный участок [188]. В результате биоинформатического анализа было выяснено, что короткий консервативный участок может являться кОРС, кодирующей пептид длиной 56 ак. Белковая последовательность данного пептида высококонсервативна среди хордовых [189] (Рисунок 12) и по предсказаниям биоинформатических программ содержит трансмембранный домен, на основании чего можно предположить, что он может быть локализован в мембранных органеллах.

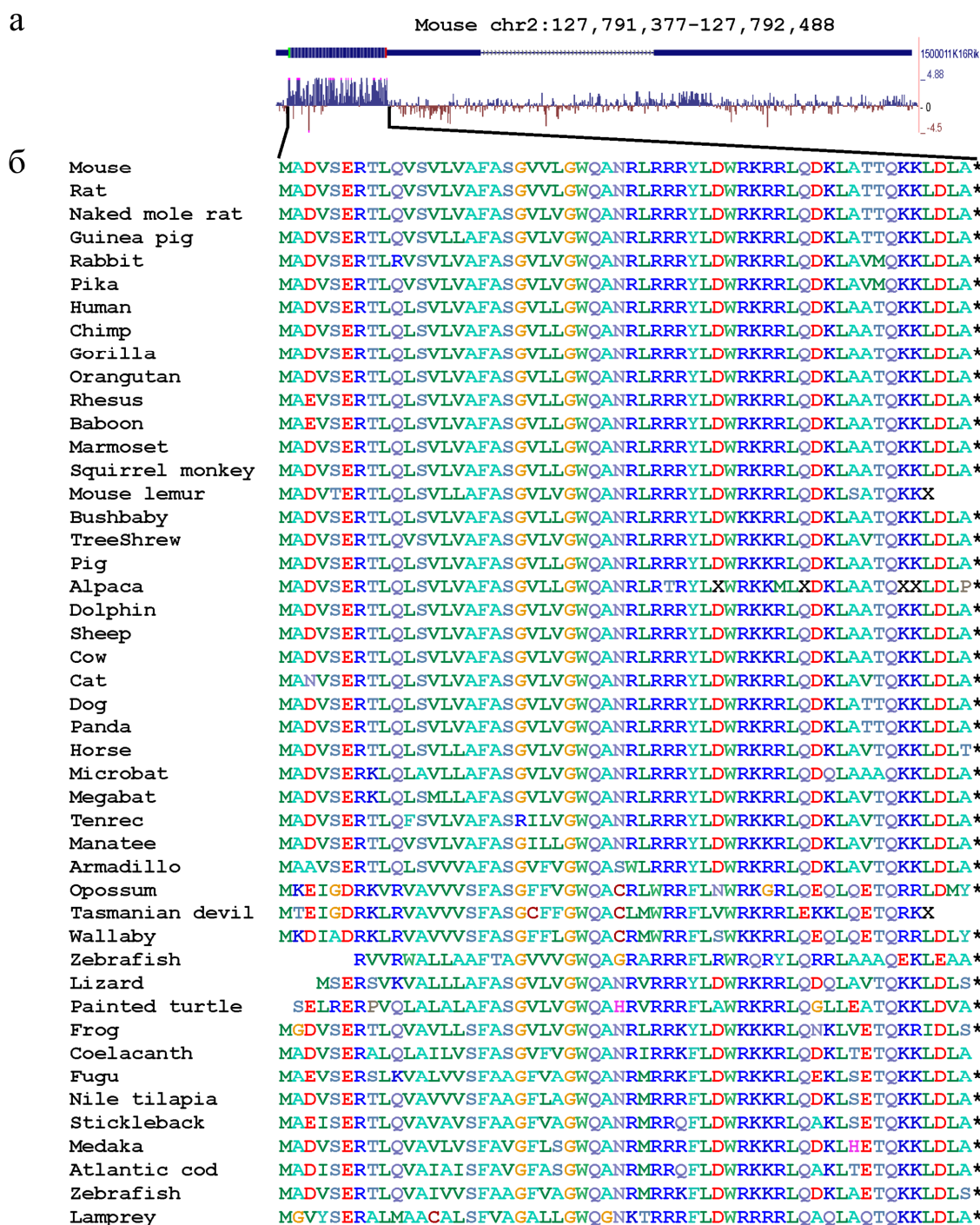


Рисунок 12. Анализ консервативности L116. (а) На рисунке представлен участок генома, кодирующий 1500011k16Rik. Старт-кодон кОПС отмечен зеленой полоской, стоп-кодон – красной. Ниже показана карта консервативности нуклеотидной последовательности гена L116 среди хордовых [189]. (б) Выравнивание белковых последовательностей, кодируемых данной кОПС среди хордовых.

Других известных белковых доменов или сигнальных последовательностей обнаружить не удалось. Дополнительным доказательством кодирующего

потенциала данного гена являются данные рибосомного профилирования [190], которые выявили значительное покрытие рибосомными футпринтами участка кОРС (Рисунок 13а). Выравнивание предполагаемого продукта трансляции кОРС показало высокую степень консервативности белкового состава с высоким соотношением синонимичных к несинонимичным заменам и отсутствию преждевременных стоп-кодонов (Рисунок 13б, в). В результате данного анализа мы заключили, что ген 1500011k16Rik, скорее всего кодирует пептид длиной 56 аа, который мы назвали L116. Целью настоящей работы стало изучение биологической роли пептида L116 в клетке.

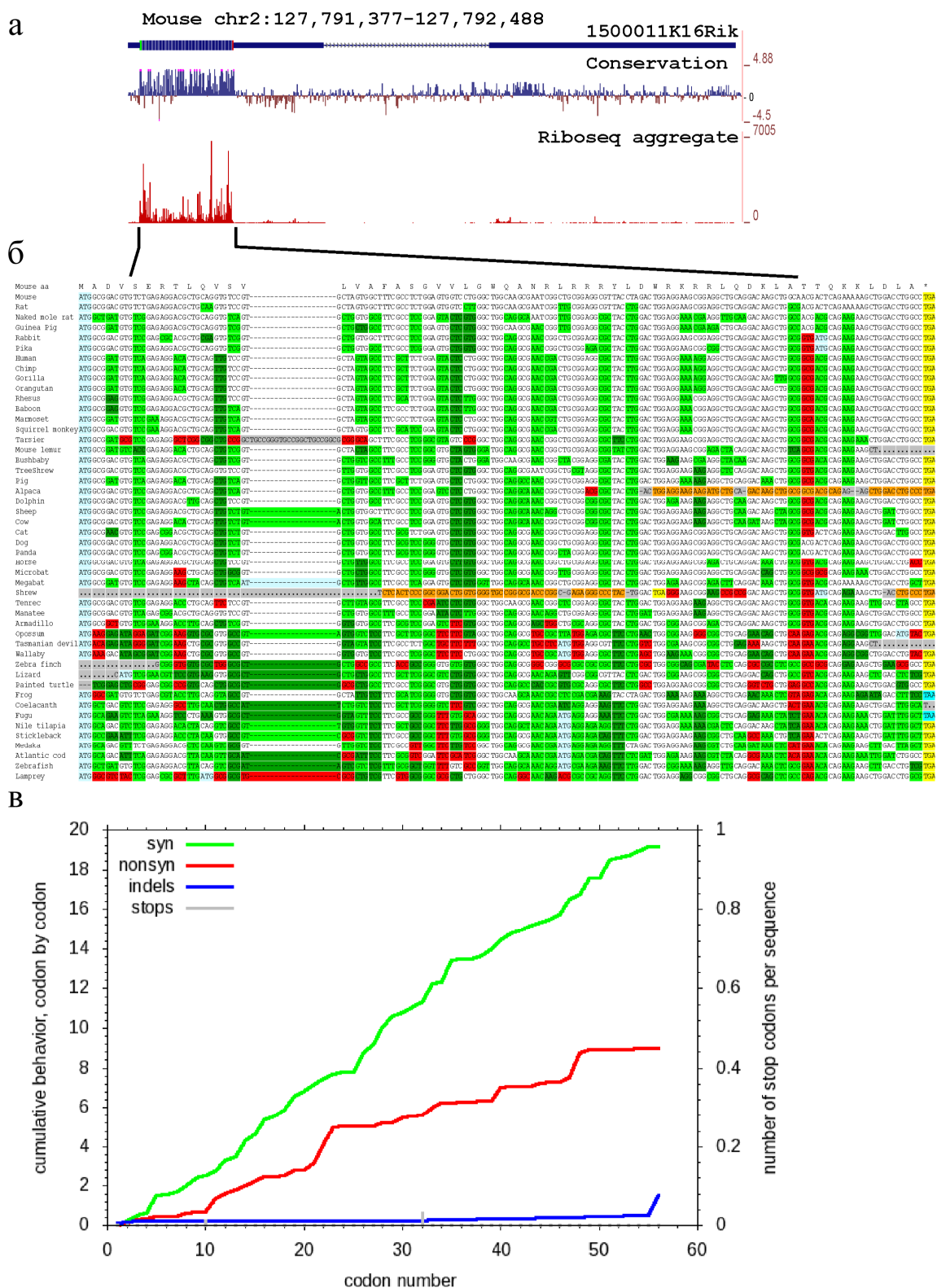


Рисунок 13. Сведения, подтверждающие то, что ген 1500011k16Rik и гомологичные кодируют функциональный пептид. (а) На рисунке представлен участок генома мыши, кодирующий 1500011k16Rik. Старт-кодон кОПС отмечен зеленой полоской, стоп-кодон – красной. (б, в) Ниже показана карта консервативности среди

хордовых [189] и диаграмма покрытия рибосомными футпринтами [190] соответственно.

Доказательство экспрессии и определение локализации L116 в клетках

Для того, чтобы доказать существование пептида L116 в культивируемой мышинной клеточной линии, мы применили метод описанный ранее в литературном обзоре. Для встраивания флуоресцентного белка mCherry на С-конец пептида мы использовали систему CRISPR/Cas9 [191,192]. Данная система позволяет направленно вносить двухцепочечные разрывы в ДНК, которые впоследствии будут восстановлены клеточным аппаратом репарации. Существует два пути репарации. Негомологичное соединение концов – это превалирующий путь в клетке [193], в результате, в месте разрыва могут образоваться вставки или делеции. В данном случае нам было необходимо, чтобы репарация пошла по пути гомологичной рекомбинации. Вариант последовательности, который нужно внести в геном, содержится на плазмиде, так называемой матрице для гомологичной рекомбинации.

С помощью сервиса CRISPR design (<http://crispr.mit.edu>) были подобраны несколько гидовых РНК к области ДНК, соответствующей С-концу пептида, но при этом не разрезали гомологичную матрицу, содержащую вставку. Гидовые РНК были клонированы в вектор px458, после чего необходимо было проверить эффективность разрезания ДНК с помощью Т7-эндонуклеазы. Самую эффективную гРНК мы выбрали для дальнейших экспериментов. Для получения клеточной линии с точным встраиванием эпитопа, мы отбирали после трансфекции клетки, обладающие красной флуоресценцией. Дальнейшая проверка с помощью ПЦР и секвенирования подтвердила успешное получение нами клеточной линии NIH3T3 с точным встраиванием флуоресцентного белка mCherry на С-конец пептида L116 (Рисунок 14а). Эндогенная экспрессия химерного белка L116-mCherry может быть детектирована как с помощью иммуноблоттинга (Рисунок 14б), так и с помощью флуоресцентной микроскопии (Рисунок 14в). Иммуноцитохимия также показала, что пептид действительно локализуется в мембранных органеллах, а не случайно распределен по всей клетке.

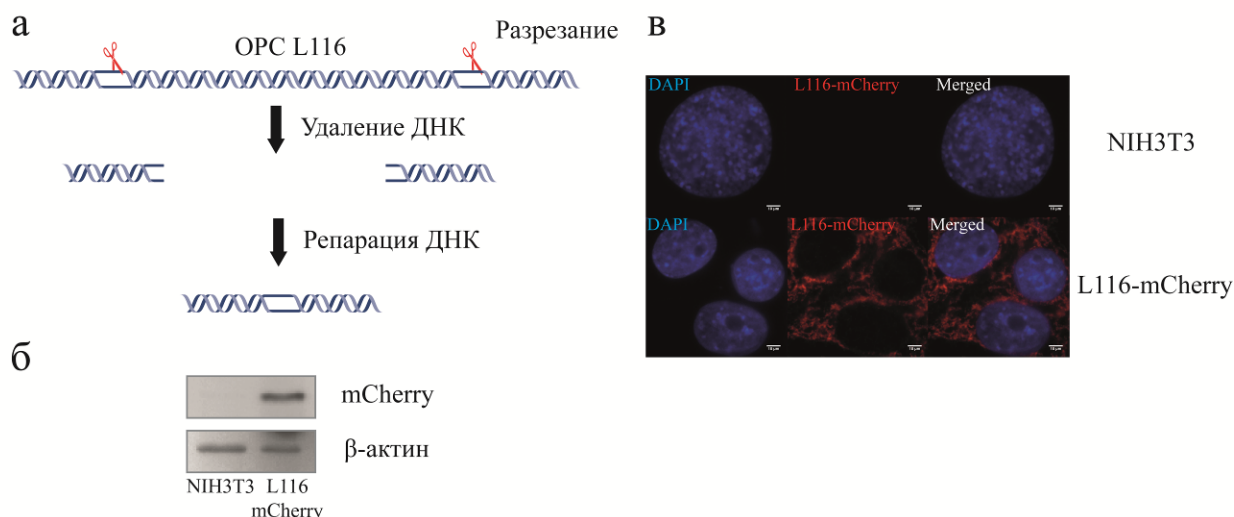


Рисунок 14. (а) Схема редактирования OPC L116 в геноме клеток. (б) Иммуноблоттинг антителами на mCherry клеточных экстрактов из NIH3T3 и клеток, экспрессирующих ген L116-mCherry в природном контексте гена L116. β-актин был использован как контроль нанесения. (в) Иммуноцитохимия клеток NIH3T3 и клеток, экспрессирующих ген L116-mCherry. Ядра были окрашены DAPI. Размерная шкала 10 мкм.

Несмотря на то, что эндогенный уровень экспрессии L116 был достаточен для того, чтобы показать сам факт синтеза белка L116-mCherry и его локализацию в каких-то мембранных органеллах клетки, низкая экспрессия эндогенной формы белка L116-mCherry мешала дальнейшему исследованию локализации белка. В связи с этим было принято решение создать клеточную линию с повышенной экспрессией химерного белка. Для этих целей мы воспользовались системой транспозиции Sleeping Beauty [194], которая позволяет встраивать экспрессирующую кассету в случайное место генома клеточной линии. Встраивание кассеты, кодирующей гибридный белок L116-mCherry под контролем гибридного доксициклин-индуцируемого промотора на основе промотора RPBSA, в геном происходило при котрансфекции созданной кассеты с плазмидой, кодирующей транспозазу SB100X. В результате использования линии клеток с повышенной экспрессией L116-mCherry было выяснено, что пептид L116 находится в митохондриях, о чем свидетельствует его ко-локализация со специальным красителем для митохондрий MitoTracker GreenFM (Рисунок 15).

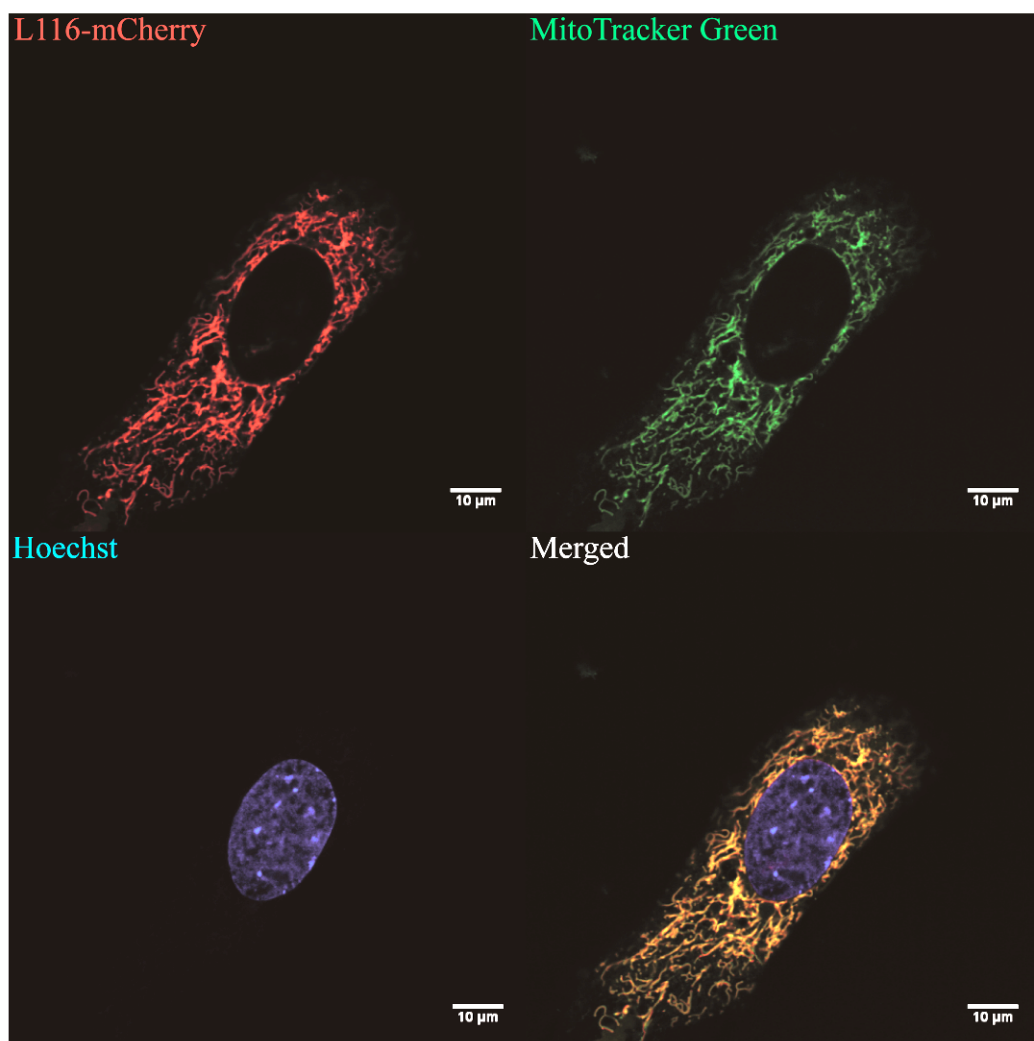


Рисунок 15. Конфокальная микроскопия клеток с повышенной экспрессией гена *L116-mCherry*. Митохондрии окрашены с помощью *MitoTracker Green FM*, ядра - *Hoechst 33342*. Размерная шкала 10 мкм.

Создание клеток с инактивированным геном 1500011k16Rik

Для создания нокаутных клеток мы также воспользовались системой CRISPR-Cas9. Было решено использовать две гидовые РНК, с участками комплементарности, соответствующими N- и C-концевому участку пептида. Отбор и проверка клонов осуществлялись способом, описанным выше. В результате нам удалось получить ряд клонов с полным удалением кОПС в гене 1500011k16Rik (Рисунок 16а, б).

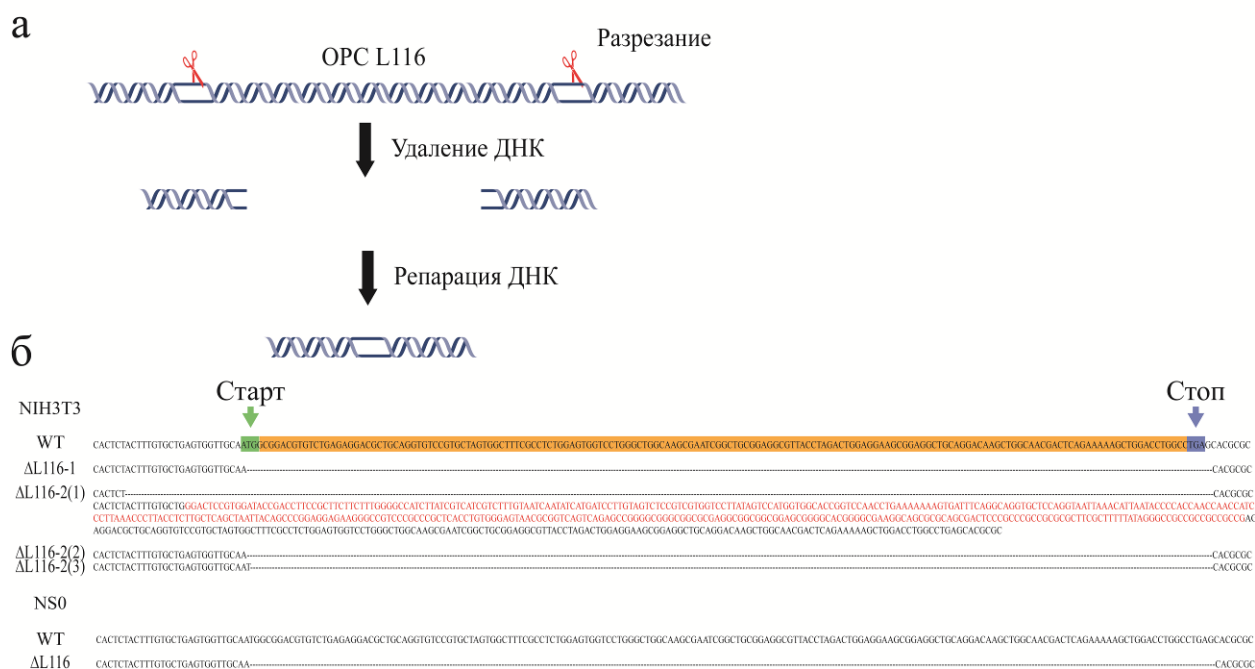


Рисунок 16. Генотипы клонов с инактивацией гена *1500011k16Rik*. Участок мышинного генома, кодирующий *L116*, в диком типе и клонах с мутациями, нарушающими *ОРС*. Кодирующая область отмечена оранжевым цветом, старт- и стоп-кодона окрашены в зеленый и голубой цвет соответственно. Делеции отмечены пунктиром, вставки показаны красным цветом.

Для подтверждения инактивации гена на белковом уровне были получены поликлональные антитела против *L116*. Получение антител осуществила по нашему заказу компания Eurogentec. При выполнении данной задачи мы столкнулись с проблемой низкой экспрессии эндогенного пептида в клеточной линии *Н1Н3Т3*. Пептид не удавалось детектировать вестерн-блоттингом после электрофоретического разделения полного клеточного лизата. Для решения этой проблемы было решено выделять митохондрии из клеток и проводить иммуноблоттинг после электрофоретического разделения митохондриальных экстрактов. В результате нам удалось показать, что пептид присутствует только в клеточных линиях дикого типа (*Н1Н3Т3* и *NS0*) и его экспрессия не детектируется в клетках с инактивированным геном (Рисунок 17а). Для того чтобы убедиться, что пептид представлен не только в клеточных линиях, но также и в живом организме, мы проанализировали различные мышинные органы с помощью вестерн-блоттинга на предмет экспрессии пептида (Рисунок 17б). Пептид *L116* хорошо экспрессируется в таких тканях, как почки и печень. Эти данные

подтверждают результаты, полученные с помощью количественного ПЦР (Рисунок 17в).

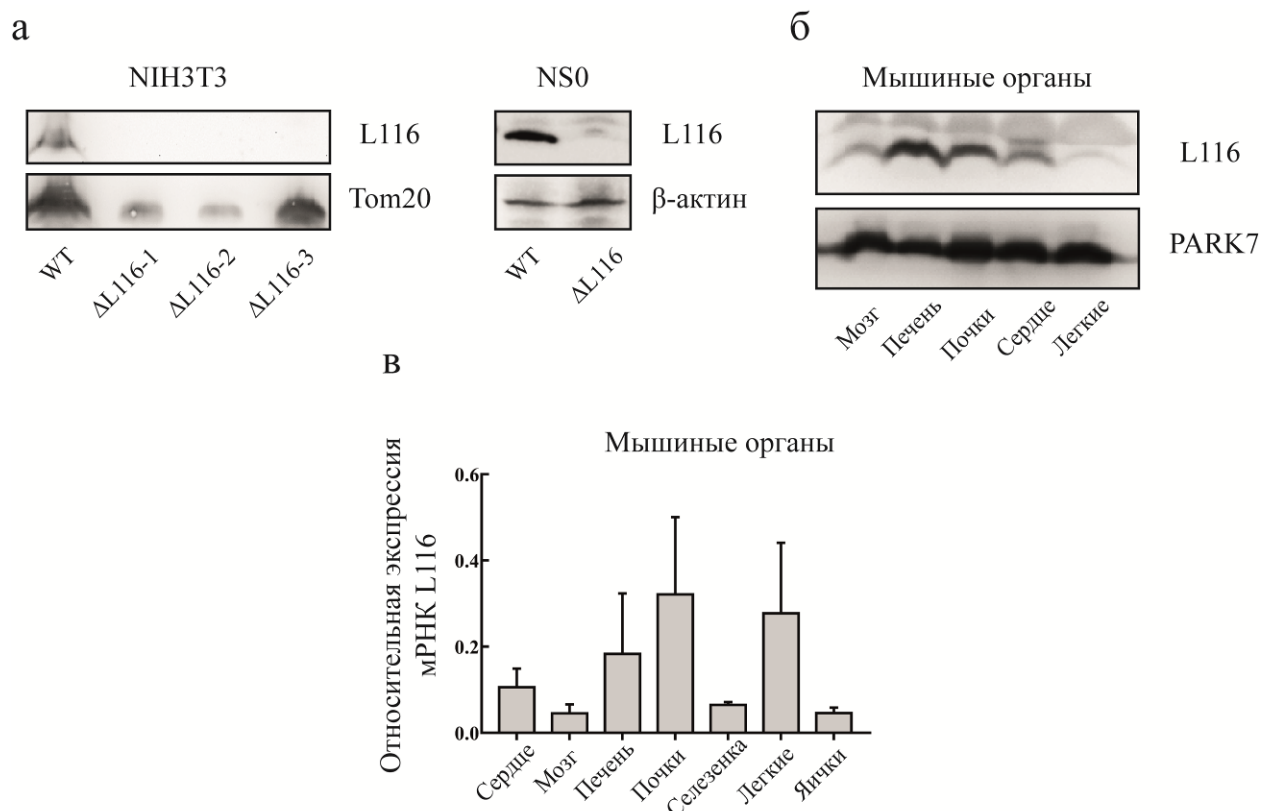


Рисунок 17. (а) Иммуноблоттинг продуктов электрофоретического разделения белков митохондрий, изолированных из клеток NIH3T3 и полных клеточных экстрактов из NS0 с помощью анти-L116 антител. Tom20 и β -актин были использованы как контроль нанесения для митохондриальных и клеточных экстрактов, соответственно. Δ L116-1, 2, 3 – это различные нокаутные клеточные линии, созданные на основе NIH3T3. (б) Иммуноблоттинг продуктов электрофоретического разделения экстрактов мышечных органов с помощью анти-L116 антител. PARK7 был использован как контроль нанесения [195]. (в) RT-qPCR анализ экспрессии мРНК L116 в различных мышечных органах. Для анализа были использованы три биологические реплики (по 3 технических). Столбцы показывают относительную экспрессию, нормализованную на GAPDH, данные представлены как среднее $\pm$ SD.

Изучение фенотипа клеток с инактивированным 1500011k16Rik

Работа выполнена совместно с Лабораторией молекулярных механизмов старения, Научно-исследовательский институт физико-химической биологии имени А.Н.Белозерского, МГУ имени М.В. Ломоносова (Высоких М.Ю.). Вследствие того, что пептид L116 локализуется в митохондриях, мы

предположили, что его удаление может влиять на функцию митохондрий. Для того, чтобы проверить данную гипотезу, мы решили измерить такие показатели функциональности митохондрий, как мембранный потенциал и продукцию активных форм кислорода (ROS, reactive oxygen species). Распределение клеток с высоким и низким митохондриальным потенциалом было проанализировано с помощью специального радиометрического красителя JC1 (Рисунок 18а). Общий уровень продукции ROS в клетках дикого типа и мутантных был оценен с использованием специального красителя DCFDA (Рисунок 18б). Результаты показывают, что никакой существенной разницы между клетками дикого типа и мутантными по обоим параметрам нет.

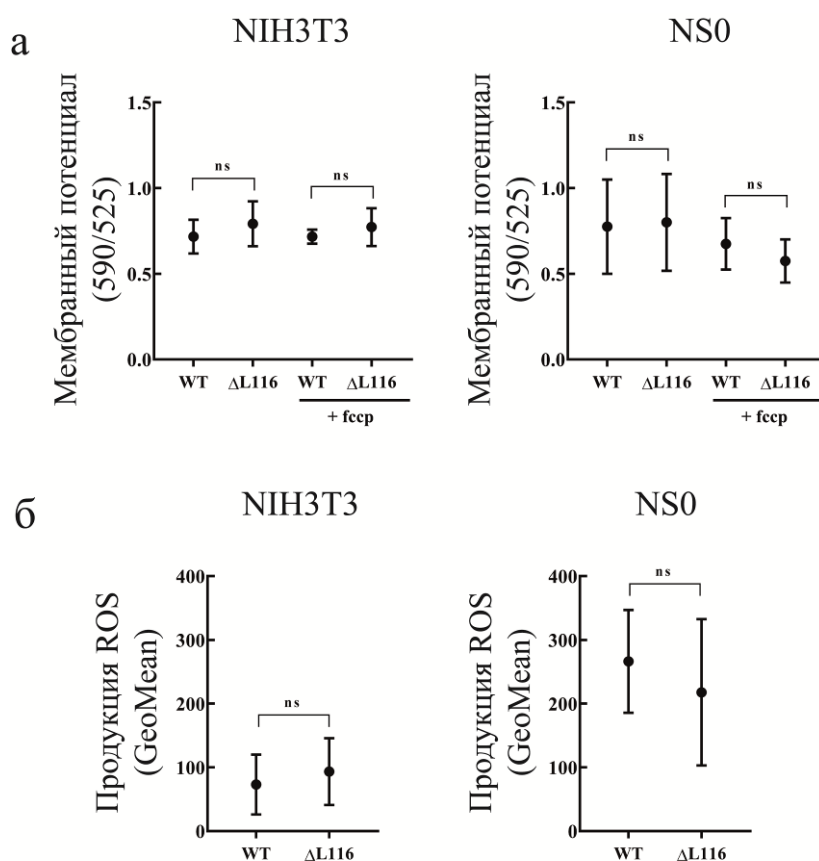


Рисунок 18. (а) Мембранный потенциал митохондрий, измеренный с помощью специального красителя JC1 (соотношение красной и зеленой флуоресценции) для клеток дикого типа и нокаутных (для NIH3T3 были измерены все три нокаутных линии $\Delta L116$ -1,2,3). Представленные значения являются средним для 6 независимых экспериментов для NIH3T3 и 4 независимых экспериментов для NS0. *t*-тест Стюдента (ns - незначительно). (б) Продукция ROS была проанализирована с помощью специального красителя DCFDA для клеток дикого типа и нокаутных (для NIH3T3 были измерены все

три нокаутных линии $\Delta L116$ -1,2,3). Представленные значения являются средним для 6 независимых экспериментов для *NIH3T3* и 4 независимых экспериментов для *NS0*. *t*-тест Стюдента (*ns* - незначительно).

Для того чтобы проверить, влияет ли пептид L116 на активность комплексов дыхательной цепи, было решено сравнить скорости поглощения кислорода клетками с интактным и инактивированным геном 1500011k16Rik. Скорость поглощения кислорода позволяет оценить поток электронов через комплексы I-IV дыхательной цепи (Рисунок 19а). Активность комплекса I (НАДН убихинон оксиредуктаза), измеряемая как ротенон чувствительное поглощение кислорода с использованием глутамата и малата в качестве субстратов, оказалась ниже в случае удаления L116 в обеих клеточных линиях.

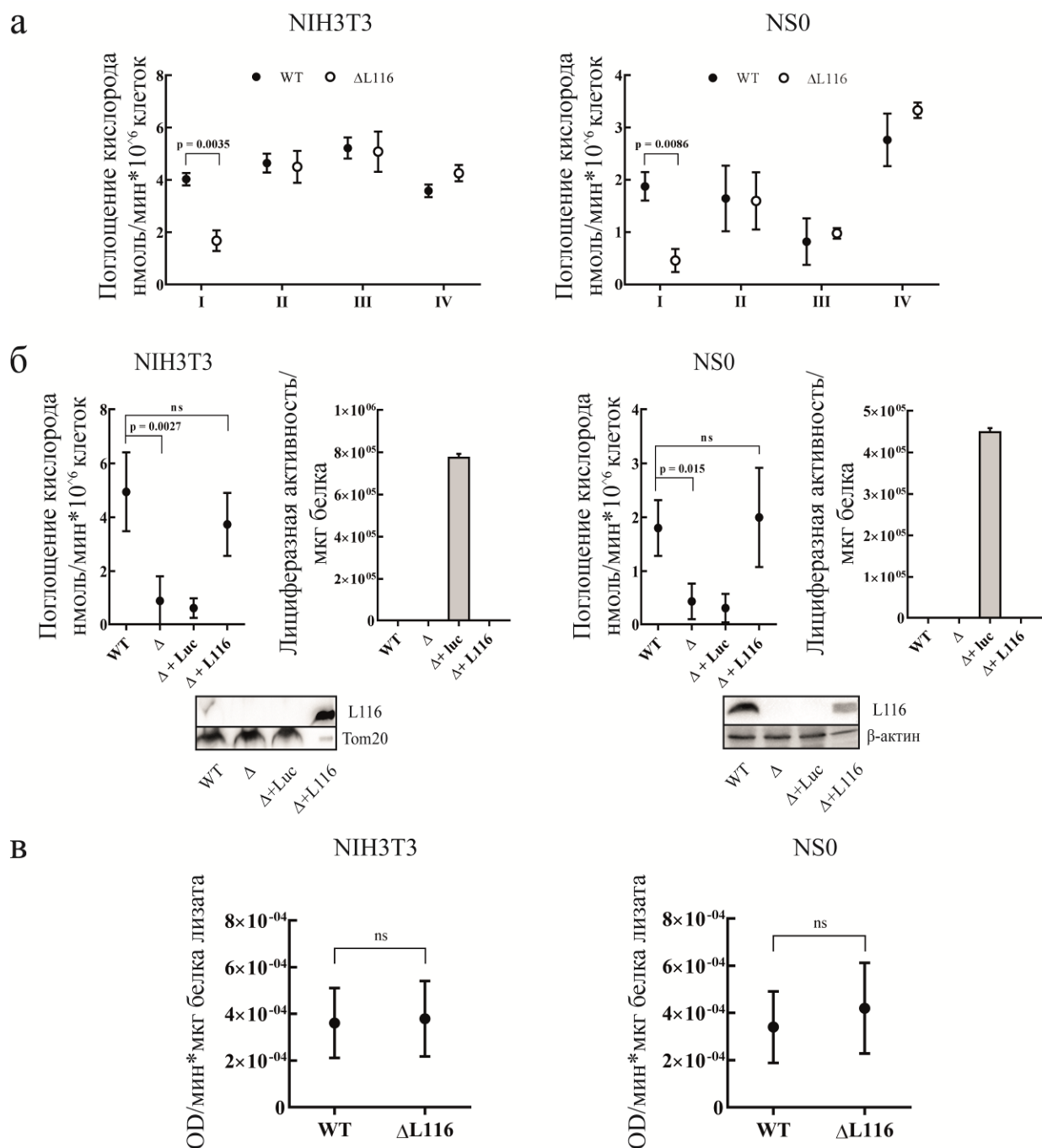


Рисунок 19. (а) Скорость поглощения кислорода для комплексов I-IV дыхательной цепи клеток дикого типа и нокаутных. Для NIH3T3 использована нокаутная линия $\Delta L116$ -3. Представленные значения являются средним для 3 независимых экспериментов (3 технические реплики каждый). Для статистически достоверного изменения указан уровень вероятности с поправкой на множественное сравнение ($p < 0.05$, t-тест Стюдента, поправка Холма-Сидака). (б) Скорость поглощения кислорода для комплекса I в клетках дикого типа и нокаутных, нокаутных, комплементированных конструкцией, кодирующей люциферазу, и нокаутных, комплементированных конструкцией, кодирующей L116 (для NIH3T3 использована нокаутная линия $\Delta L116$ -3). Представленные

значения являются средним для 3 независимых экспериментов (3 технические реплики каждый). Для статистически достоверного изменения указан уровень вероятности с поправкой на множественное сравнение ($p < 0.05$, однофакторный ANOVA, поправка Даннетта). Ниже представлен иммуноблоттинг продуктов электрофоретического разделения изолированных митохондрий из клеток NIH3T3 и клеточных экстрактов из NS0, который показывает экспрессию пептида L116. Том20 и β -актин были использованы как контроли нанесения для митохондрий и клеточных экстрактов соответственно. Люциферазный анализ показывает активность только в клетках, комплементированных конструкцией, кодирующей люциферазу. Во всех случаях данные представлены как среднее $\pm$ SD. (в) Активность изолированного комплекса I в клетках дикого типа и нокаутных (для NIH3T3 были измерены все три нокаутных линии клеток Δ L116-1,2,3). Представленные значения являются средним для 6 независимых экспериментов для NIH3T3 и 5 независимых экспериментов для NS0 (2 технические реплики каждый). t-тест Стюдента (ns - незначительно). Во всех случаях данные представлены как среднее $\pm$ SD.

В связи с тем, что данный эффект мог быть обусловлен побочными мутациями, вызванными CRISPR-Cas9 системой, мы решили сделать дополнительный контроль. Для этого в нокаутную клеточную линию мы ввели экспрессионную конструкцию на основе транспозонного вектора, кодирующую L116 или люциферазу, как отрицательный контроль.. Это позволило нам определить, является ли данный эффект следствием отсутствия пептида L116 в клетках. Действительно, именно экспрессия L116, но не люциферазы, приводила к восстановлению активности комплекса I (Рисунок 19б).

На следующем этапе работы, у нас возник вопрос, каким образом пептид может влиять на функцию комплекса I. НАДН убихинон оксиредуктаза – это большой мультисубъединичный комплекс [196], которому помимо собственно структурных компонентов, необходимо несколько дополнительных факторов сборки. Первым предположением было, что пептид L116 может являться одним из компонентов первого комплекса, либо нужен для его сборки. Для проверки данного предположения, мы выделили комплекс I из клеток дикого типа и нокаутных по L116 и сравнили его способность окислять НАДН (Рисунок 19в). Активность изолированного комплекса I оказалась одинаковой в обоих случаях, что говорит о том, что пептид L116 не является частью комплекса I и не принимает

участие в его сборке. Таким образом, можно сделать вывод о том, что пептид L116 влияет на функцию комплекса I опосредованно.

Поиск белковых партнеров L116

Так как удаление пептида L116 влияет на активность комплекса I опосредованно, мы предположили, что это может происходить через взаимодействие с другим белком или ферментом, который необходим для правильной работы комплекса I. Для поиска белков, с которыми взаимодействует L116, мы создали линию клеток, в которой экспрессировался ген пептида L116 с НА-эпитопом. На первом этапе работы было принято решение сделать клеточную линию с точным встраиванием НА-довеска на N-конец пептида с помощью CRISPR-Cas9 системы. Нам удалось получить ряд клонов с необходимой вставкой. Однако проверка экспрессии гена пептида с помощью иммуноблоттинга антителами на НА-эпитоп показала, что в природном геномном контексте L116 ген гибридного белка НА-L116 не экспрессируется. Для проведения опытов по иммунопреципитации L116 была получена клеточная линия с повышенной экспрессией НА-L116. Она была использована для поиска белковых партнеров L116. Единственный белок, который воспроизводимо и достоверно совыделялся с L116, оказался НАДН цитохром b5 редуктазой 3 (Cyb5r3) (Рисунок 20а, б). Для того, чтобы убедиться, что это взаимодействие специфично, а не обусловлено наличием НА-эпитопа, мы провели ко-иммунопреципитацию партнеров L116-mCherry из клеток, в геном которых была встроена соответствующая конструкция. Как и в случае с НА-L116, в результате иммунопреципитации L116 за довесок mCherry нам удалось выделить Cyb5r3 (Рисунок 20в). Чтобы проверить возможную колокализацию L116 и Cyb5r3, мы экспрессировали L116-mCherry и Cyb5r3-eGFP одновременно в клетках NIH3T3 (Рисунок 20г). С помощью этого опыта нам удалось показать, что L116 колокализуется с Cyb5r3, подтверждая то, что наблюдаемое взаимодействие этих белков не является артефактом разрушения клеток при выделении экстракта.

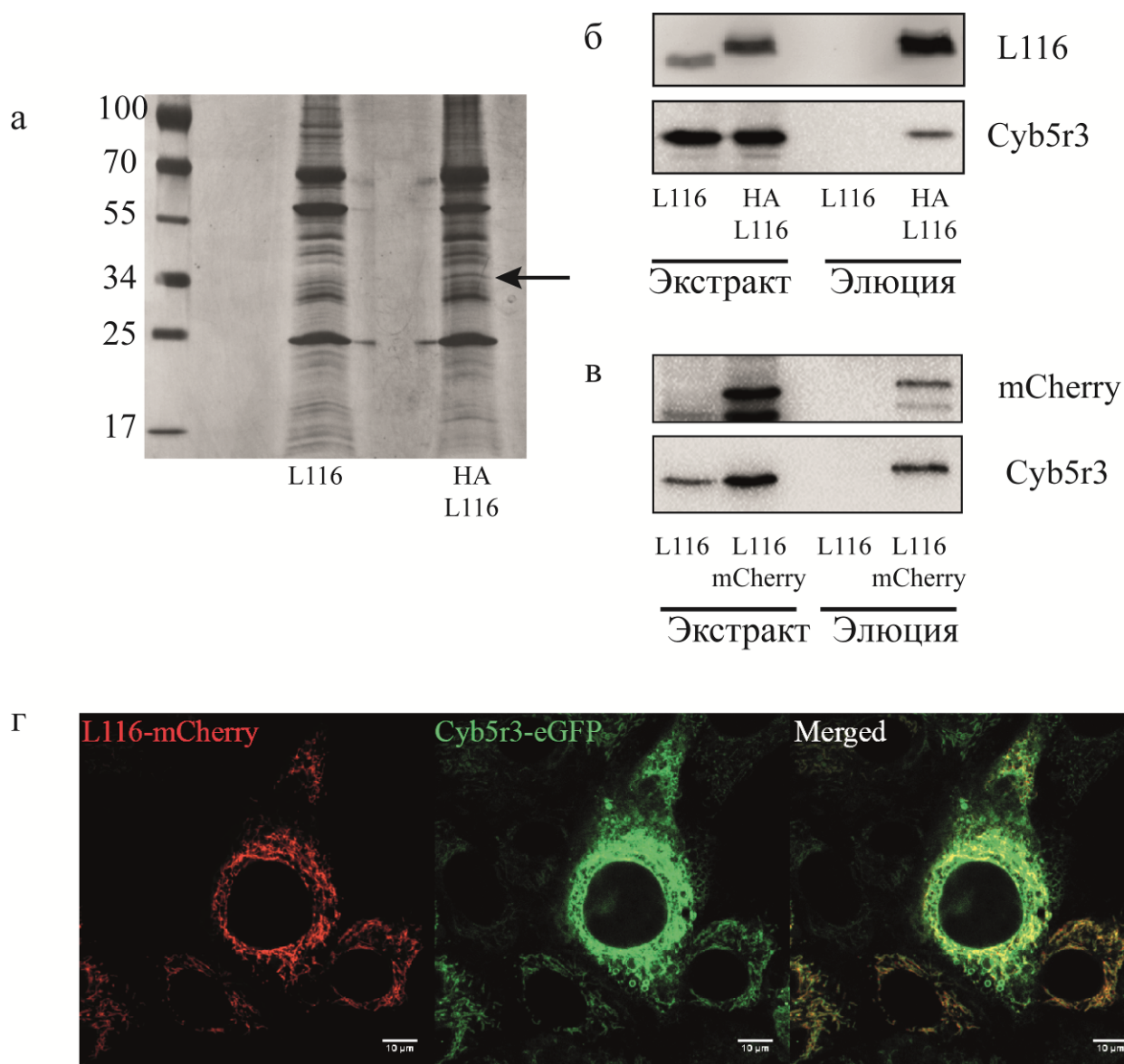


Рисунок 20. Окрашивание нитратом серебра продуктов электрофоретического разделения элюата иммунопреципитации L116 за HA-эпитоп из клеточной линии NIH3T3, с повышенной экспрессией гена HA-L116. Клетки, экспрессирующие ген L116 без довеска были использованы в качестве контроля. (б) Иммуноблоттинг продуктов электрофоретического разделения элюата иммунопреципитации L116 за HA-эпитоп из клеточной линии NIH3T3, с повышенной экспрессией гена HA-L116 антителами на L116 (верхняя панель) и антителами на Cyb5r3 (нижняя панель). Клетки, экспрессирующие L116 без тага были использованы в качестве контроля. (в) Иммуноблоттинг продуктов электрофоретического разделения элюата иммунопреципитации L116 за mCherry из клеточной линии NIH3T3, с повышенной экспрессией гена L116-mCherry антителами на mCherry (верхняя панель) и антителами на Cyb5r3 (нижняя панель). Клетки, экспрессирующие ген L116 без тага были использованы в качестве контроля. (г)

*Конфокальная микроскопия клеток экспрессирующих L116-mCherry и Cyb5r3- eGFP.
Шкала 10 мкм.*

Изучение фенотипа клеточной линии с нарушением митохондриальной локализации Cyb5r3

Чтобы понять, действительно ли L116 влияет на работу комплекса I через Cyb5r3, мы решили мутировать ген Cyb5r3 в клетках. Известно, что Cyb5r3 локализуется как в митохондриях, так и в ЭПР [197], где может участвовать в различных процессах. Cyb5r3 должен быть N-меристилизован по аминокислоте Gly2 для локализации в митохондриях [198]. Из анализа научной литературы [198] известно, что мутация Gly на Ala приводит к тому, что Cyb5r3 перестает интегрироваться в мембрану митохондрий. Поэтому было принято решение создать клеточную линию с точечной заменой кодона Gly на кодон Ala в геноме, что позволит нам нарушить митохондриальную локализацию Cyb5r3 (Δ Cyb5r3_{mito}). Для создания этой линии мы использовали систему CRISPR/Cas9 с гидовой РНК, направляющей разрезание генома в участке, соответствующем началу кодирующей области Cyb5r3 и матрицу для гомологичной рекомбинации. Нам удалось получить клон с точной заменой на всех хромосомах, что было подтверждено секвенированием продуктов ПЦР-амплификации соответствующего участка генома. Иммуноблоттинг продуктов электрофоретического разделения митохондриальных экстрактов из клеток дикого типа и мутантных Δ Cyb5r3_{mito} показал, что мутация приводит к снижению количества Cyb5r3 в митохондриях (Рисунок 21а). Не полное отсутствие белка в митохондриальных экстрактах может быть связано с частичным совыделением ЭПР фракции при очистке митохондрий [199]. Измерение скорости поглощения кислорода для мутантных Δ Cyb5r3_{mito} клеток (Рисунок 21б) показало, что нарушение митохондриальной локализации белка Cyb5r3 приводит к такому же фенотипу, что и инактивация пептида L116. Таким образом, можно заключить, что инактивация гена пептида L116 и нарушение митохондриальной локализации Cyb5r3 одинаково влияют на комплекс I.

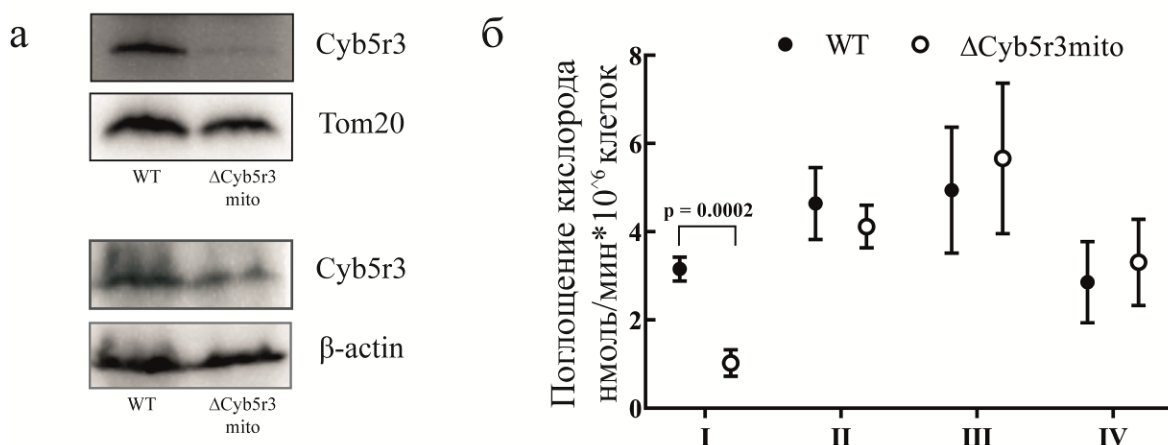


Рисунок 21. (а) Иммуноблоттинг антителами на Cyb5r3 продуктов электрофоретического разделения митохондриальных и полных клеточных экстрактов клеток дикого типа и мутантных $\Delta\text{Cyb5r3}_{\text{mito}}$. Антитела к Tom20 и β -актину были использованы для контроля идентичности количества нанесенных белков. (б) Активность комплексов I-IV в клетках с инактивированной митохондриальной Cyb5r3. Представленные значения являются средним для 4 независимых экспериментов (3 технические реплики каждый). Для статистически достоверного изменения указан уровень вероятности с поправкой на множественное сравнение ($p < 0.05$, t -тест Стюдента, поправка Холма-Сидака). Данные представлены как среднее $\pm$ SD.

Изучение функциональности взаимодействия L116 и Cyb5r3

Наша следующая задача состояла в том, чтобы установить функционируют ли L116 и Cyb5r3 в одном пути или в параллельных путях. Для этого мы решили создать линии с нокаутами по обоим белкам, нокаут Cyb5r3 с суперэкспрессией L116, нокаут L116 с суперэкспрессией Cyb5r3. Чтобы понять будет ли инактивация или суперэкспрессия L116 приводить к каким-то изменениям в фенотипе $\Delta\text{Cyb5r3}_{\text{mito}}$, мы сравнили потребление кислорода в данных клетках. Результаты эксперимента показали, что инактивация или суперэкспрессия L116 в $\Delta\text{Cyb5r3}_{\text{mito}}$ мутантных клетках не приводит ни к ухудшению, ни к улучшению фенотипа клеток. Таким образом, мы сделали вывод, что вероятно L116 выполняет свою функцию через Cyb5r3, и они функционируют в одном пути.

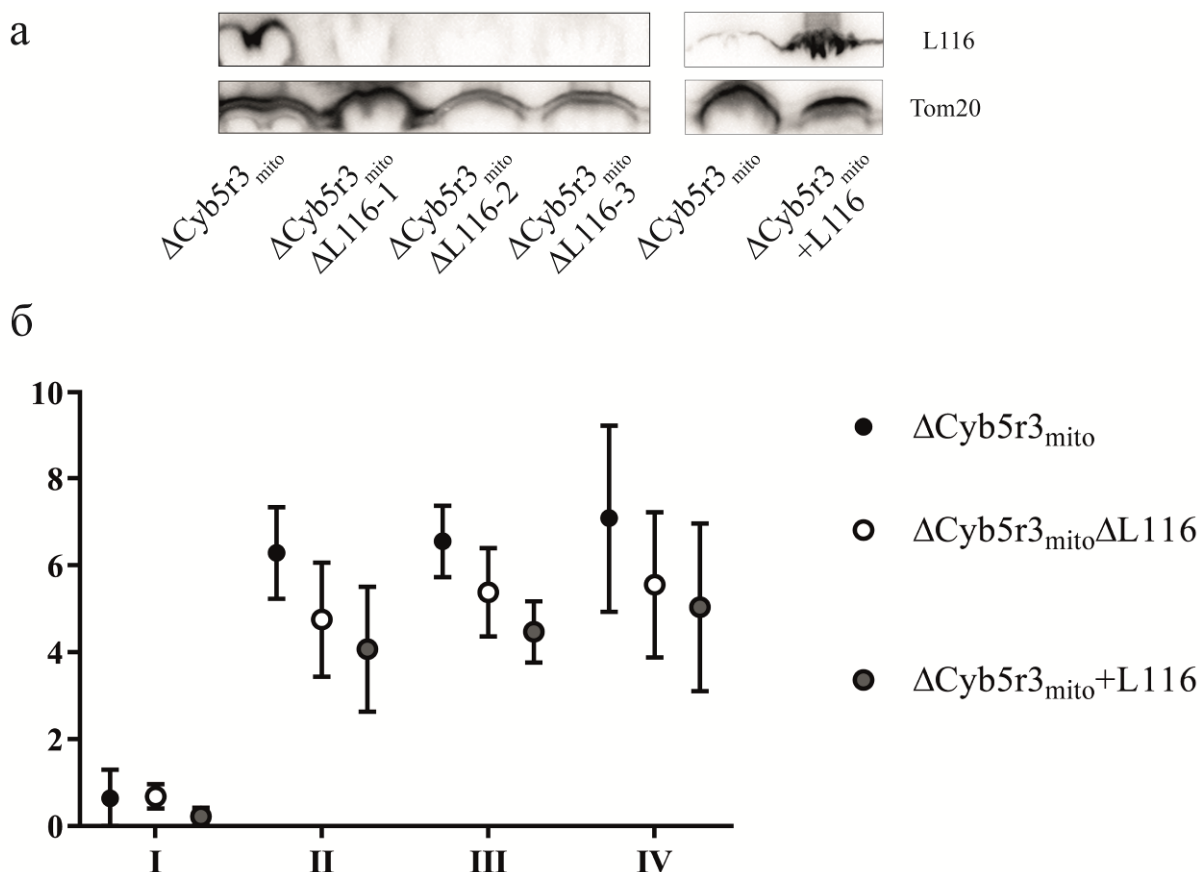


Рисунок 22. (а) Иммуноблоттинг продуктов электрофоретического разделения митохондриальных экстрактов клеток $\Delta\text{Cyb5r3}_{\text{mito}}$ и клеток $\Delta\text{Cyb5r3}_{\text{mito}}$ с инактивацией L116 антителами на L116. Антитела к Tom20 были использованы в качестве контроля нанесения. (б) Активность комплексов I-IV в мутантных клетках $\text{Cyb5r3}_{\text{mito}}$, $\text{Cyb5r3}_{\text{mito}}$ с инактивацией L116 и $\text{Cyb5r3}_{\text{mito}}$ с суперэкспрессией L116. Представленные значения являются средним для 4 независимых экспериментов (3 технические реплики каждый). Для статистически достоверного изменения указан уровень вероятности с поправкой на множественное сравнение ($p < 0.05$, t -тест Стюдента, поправка Холма-Сidak). Данные представлены как $\text{среднее} \pm \text{SD}$.

Изучение влияния инактивации L116 на Cyb5r3

На данный момент известно, что КОРС-пептиды могут принимать участие во многих клеточных процессах и выполнять различные функции. В данном случае, было решено проверить, влияет ли L116 на стабильность, локализацию или активность Cyb5r3. С помощью иммуноблоттинга нам удалось показать, что L116 не влияет на количество белка Cyb5r3 в клетках (Рисунок 23а).

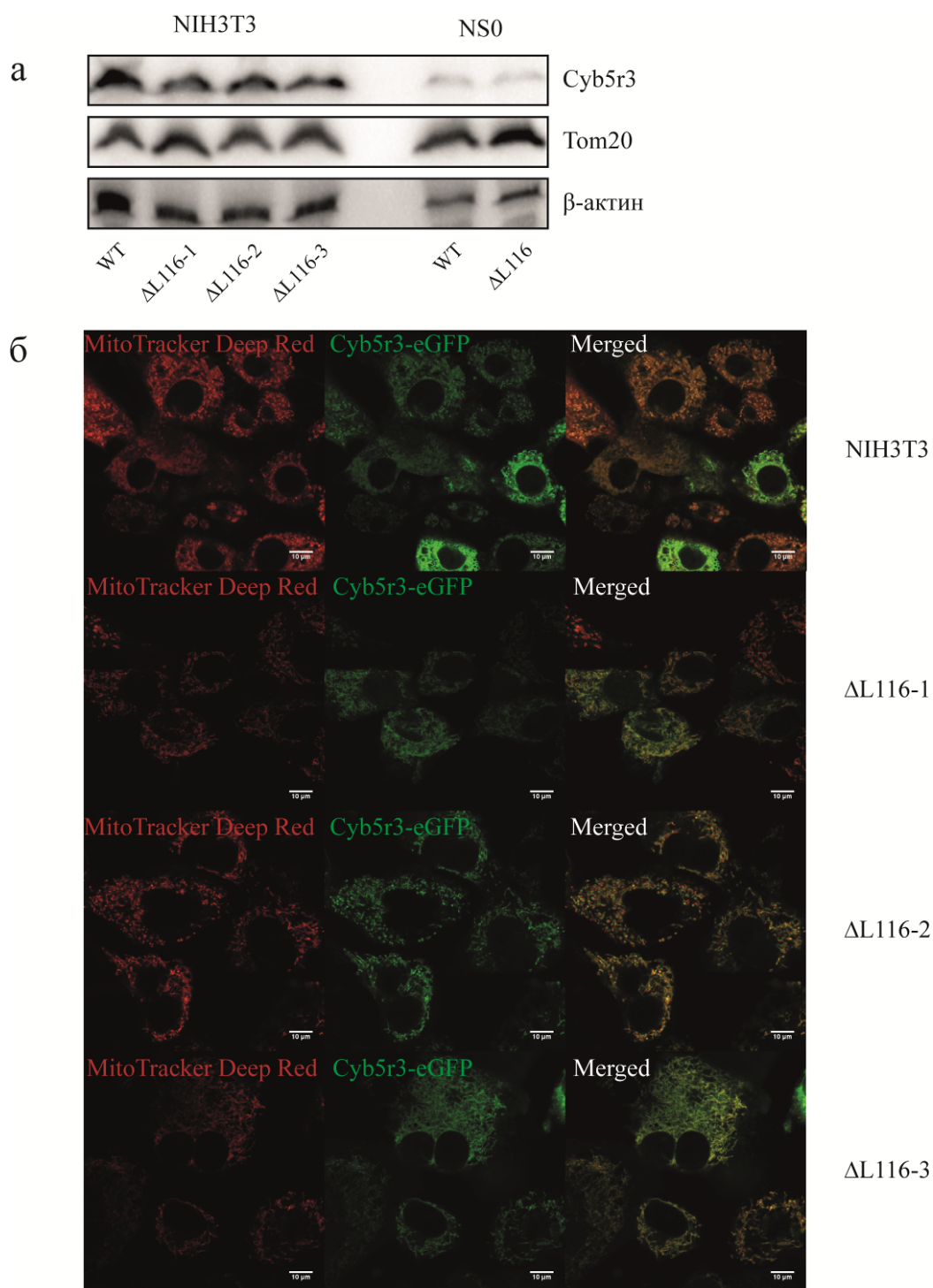


Рисунок 23. (а) Иммуноблоттинг продуктов электрофоретического разделения экстрактов из клеток дикого типа и различных нокаутных *L116* антителами на *Cyb5r3* и *Tom20*. β -актин был использован как контроль нанесения. (б) Конфокальные изображения клеток *NIH3T3* и различных производных этой линии, нокаутных по *L116* и экспрессирующих ген *Cyb5r3-eGFP*, и окрашенных специальным красителем для митохондрий *MitoTracker Deep Red*. Шкала 10 мкм.

Для того чтобы оценить возможное влияние пептида на локализацию Cyb5r3, мы ввели экзогенную форму Cyb5r3-eGFP в нокаутные по L116 клетки с помощью транспозиции (Рисунок 23б). После окрашивания митохондрий красителем MitoTracker Deep Red в полученных линиях мы использовали флуоресцентную микроскопию, чтобы оценить митохондриальную локализацию Cyb5r3 в клетках без L116. По результатам эксперимента оказалось, что локализация Cyb5r3 не меняется при инактивации гена пептида.

Для того, чтобы проверить, не влияет ли L116 на ферментативную активность Cyb5r3, ее НАДН дегидрогеназная активность была измерена с использованием ферроцианида калия в качестве акцептора электронов (Рисунок 24). Как оказалось, способность Cyb5r3 окислять НАДН не зависит от наличия L116. Таким образом, удаление пептида L116 не влияет на стабильность, локализацию и способность окислять НАДН белком Cyb5r3.

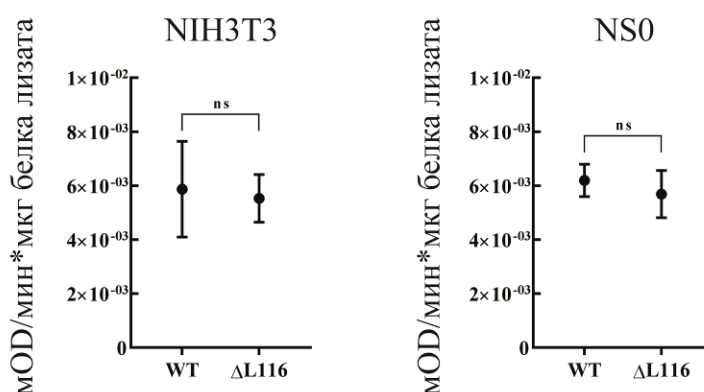


Рисунок 24. НАДН дегидрогеназная активность Cyb5r3, измеренная для клеток дикого типа и нокаутных по L116 (для NIH3T3 были измерены все три нокаута ΔL116-1,2,3). Представленные значения являются средним для 4 независимых экспериментов для NIH3T3 и 3 независимых экспериментов для NS0 (2 технические реплики каждый). *t*-тест Стюдента (*ns* - незначительно). Во всех случаях данные представлены как среднее ± SD.

Изучение влияния инактивации Cyb5r3 на L116

Мы предположили, что инактивация Cyb5r3 может влиять на стабильность и локализацию L116. Анализ экстрактов из изолированных митохондрий с помощью иммуоблоттинга выявил, что удаление Cyb5r3 не приводит к изменениям в уровне L116 (Рисунок 25а). Для того чтобы оценить возможное влияние Cyb5r3 на

локализацию пептида, мы ввели экзогенную форму L116-mCherry в мутантные по Cyb5r3 клетки с помощью транспозиции (Рисунок 25б). После окрашивания митохондрий красителем MitoTracker Green FM в полученных линиях мы использовали флуоресцентную микроскопию, чтобы оценить митохондриальную локализацию L116 в клетках без митохондриальной Cyb5r3. По результатам эксперимента оказалось, что локализация пептида не меняется при инактивации митохондриальной Cyb5r3.

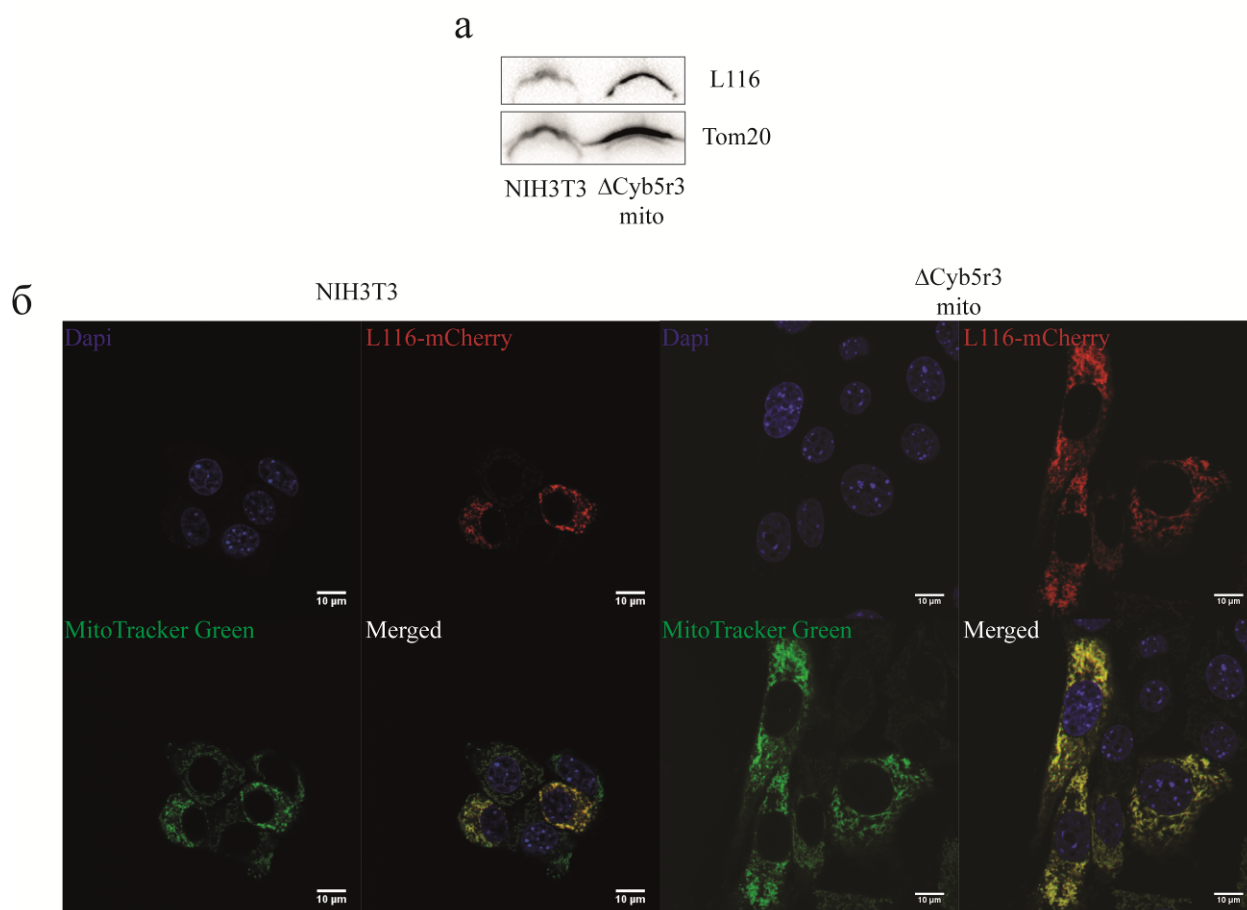


Рисунок 25. (а) Иммуноблоттинг продуктов электрофоретического разделения митохондриальных экстрактов из клеток дикого типа и клеток с инактивацией митохондриальной формы Cyb5r3 антителами на L116. Tom20 был использован как контроль нанесения. (б) Конфокальные изображения клеток NIH3T3 и клеток линии, в которой присутствует только мутантная форма Cyb5r3 (Gly2 to Ala2), экспрессирующих ген L116-mCherry и окрашенных красителем для митохондрий MitoTracker Green FM. Шкала 10 мкм.

Изучение окислительно-восстановительного баланса клетки

Работа выполнена совместно с Лабораторией молекулярных технологий, Институт биоорганической химии им. академиков М.М. Шемякина и Ю.А. Овчинникова Российской академии наук (Белоусов В.В.).

Уменьшение активности НАДН убихинон оксиредуктазы (комплекса I) может быть связано с уменьшением количества субстрата НАДН. Недавние исследования показали, что Sub5r3 может влиять на баланс $\text{НАД}^+/\text{НАДН}$ [200,201]. Поэтому, чтобы проверить возможное влияние инактивации L116 на соотношение $\text{НАД}^+/\text{НАДН}$, мы применили флуоресцентный сенсор SoNar [202]. Данный сенсор является фьюженом srYFP и НАДН-связывающего домена белка T-Rex. Он позволяет измерять $\text{НАД}^+/\text{НАДН}$ соотношение, благодаря тому, что он, связываясь с НАД^+ и НАДН, имеет разную длину волны возбуждения - 485 nm и 420 nm соответственно. В результате, нам удалось показать, что при нормальных условиях цитоплазматический баланс $\text{НАД}^+/\text{НАДН}$ в клетках дикого типа и с инактивацией L116 не отличается (Рисунок 26а). Более того, мы проанализировали, как меняется окислительно-восстановительный потенциал клеток дикого типа и нокаутных по L116 при добавлении различных метаболитов, влияющих на соотношение $\text{НАД}^+/\text{НАДН}$ (Рисунок 26б, в). Результаты данного эксперимента показали, что динамика изменения $\text{НАД}^+/\text{НАДН}$ при добавлении различных веществ также не отличается между клетками дикого типа и клетками без пептида L116. Таким образом, можно заключить, что цитоплазматический баланс $\text{НАД}^+/\text{НАДН}$ не зависит от наличия L116.

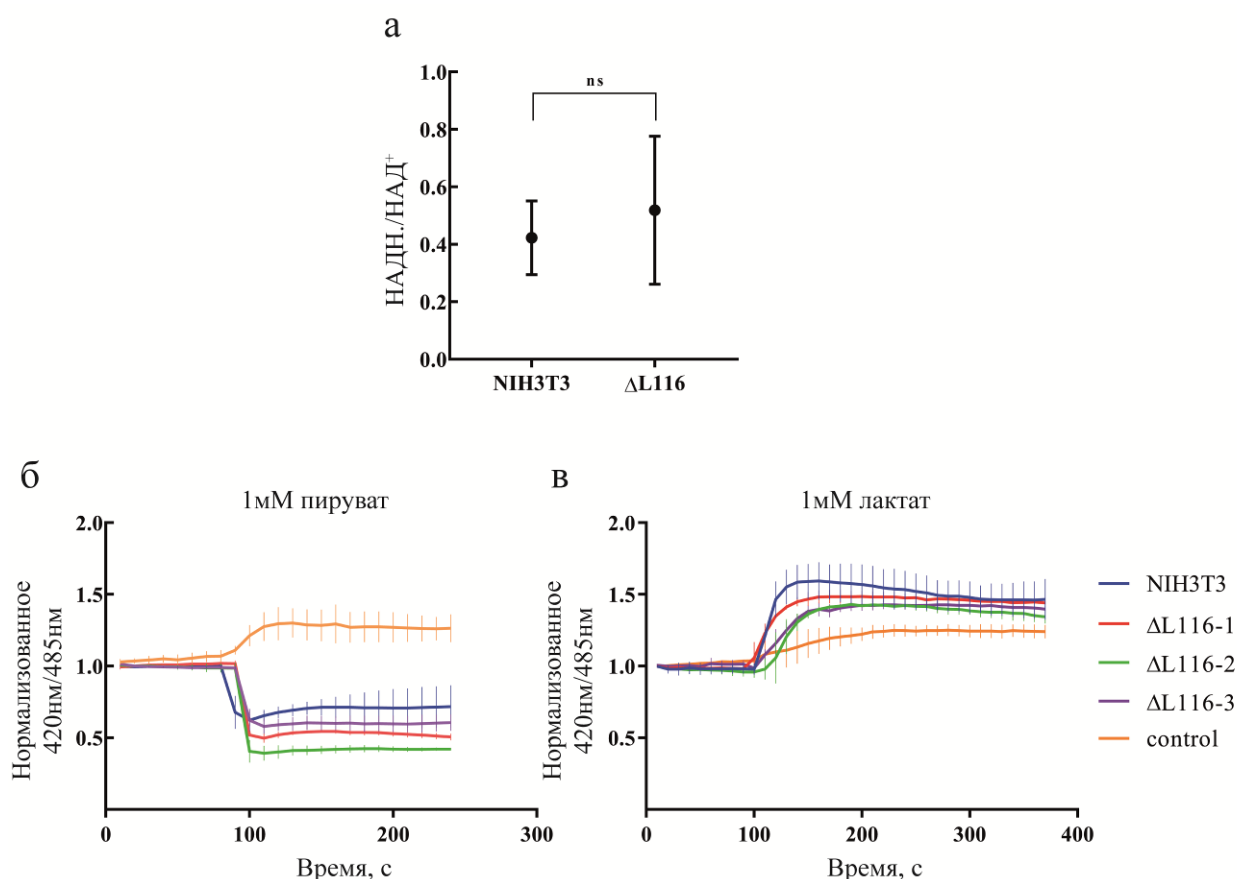


Рисунок 26. (а) НАДН/НАД⁺ соотношение в клетках NIH3T3 и нокаутных по L116. (б) Динамика изменения флуоресценции SoNar в клетках дикого типа и различных нокаутных по L116 линиях в ответ на добавление экзогенного 1мМ пирувата. (в) Динамика изменения флуоресценции SoNar в клетках дикого типа и различных нокаутных по L116 линиях в ответ на добавление экзогенного 1мМ лактата. В качестве контроля был использован iNarC сенсор [203]. Представленные значения являются средним для 4 независимых экспериментов. t-тест Стюдента (ns - незначительно). Во всех случаях данные представлены как среднее ± SD.

Изучение липидного состава нокаутных клеток

Существует два важных процесса, в которых принимает участие мембранная форма Cyb5r3: образование Δ9 двойной связи в гидрофобном хвосте жирных кислот [204] и биосинтез холестерина [205]. Для того, чтобы оценить влияние пептида L116 на активность Cyb5r3 в липидном метаболизме, было решено сравнить полный липидный состав клеток дикого типа и нокаутных по L116. Работа была выполнена совместно с научной группой профессора Хайтовича Сколковского института науки и технологии. В результате эксперимента были получены данные для представленности в клетках более чем тысячи липидов для

каждой клеточной линии. Количество около сотни липидов значительно (Рисунок 27а) и воспроизводимо (Рисунок 27б) меняются при удалении L116. Количество большинства триглицеридов (TAG) повышается, тогда как количество фосфолипидов снижается при инактивации L116. Эти результаты воспроизводятся также в клеточной линии NS0 при инактивации гена L116 (Рисунок 27в).

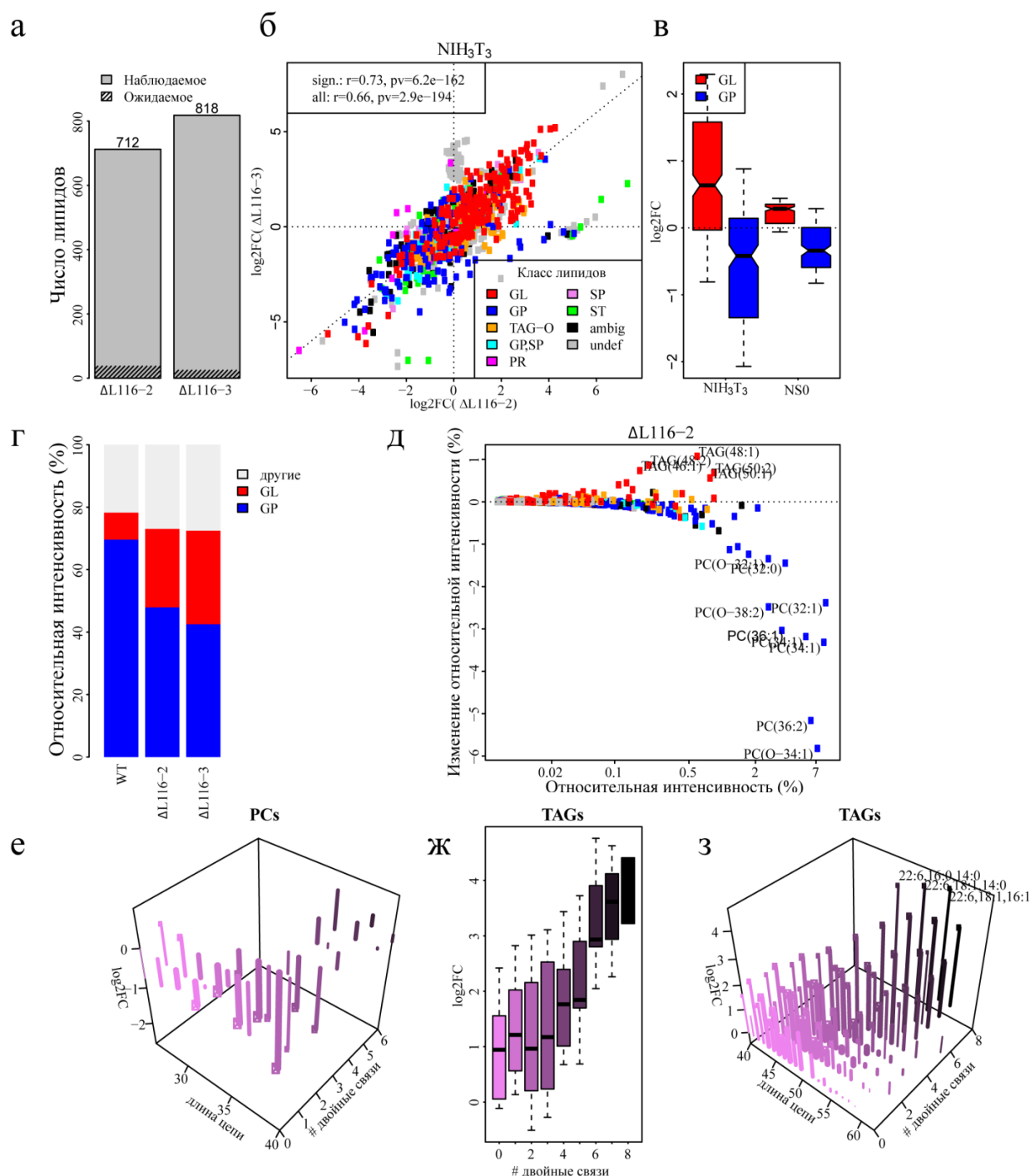


Рисунок 27. (а) Число липидов статистически достоверно изменяющихся в нокаутных по L116 линиях ΔL116-2 и ΔL116-3, случайное ожидание показано штриховкой. (б) Корреляция изменений липидного состава между ΔL116-2 и ΔL116-3.

Различные цвета обозначают различные классы липидов: глицеролипиды, глицерофосфолипиды, 1-алкил-2,3-ацилглицериды, сфинголипиды, пренолы, стеролы. Липиды, значительно изменяющиеся в обоих нокаутах, показаны квадратами. (в) Распределение представленности классов глицеролипидов (красный) и глицерофосфолипидов (синий) в клетках линий NIH3T3 и NS0 без L116 относительно контрольных изогенных линий с интактым L116. (г) Относительная интенсивность (количество) глицеролипидов (красный), глицерофосфолипидов (синий) и других липидов (серый) в NIH3T3 и нокаутных по L116 линиях. (д) Связь между относительной интенсивностью (количеством) липидов в диком типе (по оси x) и изменением относительной интенсивности (количества) в линии Δ L116-2 (ось y). (е и з) Изменение количества фосфатидилхолинов (е) и триглицеридов (з) в Δ L116-2 и Δ L116-3 (по оси z). Длина цепи и количество двойных связей соответствует координатам x и y. Точный состав цепей жирной кислоты показан для трех TAG с наибольшим изменением концентрации. Концентрация TAG и PC в диком типе показана шириной линии (в масштабе шкалы). Липиды, количество которых изменяется статистически значимо, отмечены кружком в конце линии. Статистическая значимость была определена с помощью t-теста Стюдента ($p < 0.05$, поправка Бенджамини-Хохберга). (ж) Распределение изменений количества триглицеридов (TAG) в зависимости от общего числа двойных связей.

Несмотря на то, что фосфатидилхолины (PC) являются одной из наиболее представленных групп липидов в используемых в нашей работе клеточных линиях, их количество значительно уменьшается (порядка 20%) при инактивации гена L116 (Рисунок 27г, д). Фосфатидилхолины, содержащие относительно длинные цепи с 1-3 ненасыщенными связями, изменяются больше всего (Рисунок 27е). В противоположность фосфатидилхолинам количество триглицеридов повышается в нокаутных по L116 клетках (Рисунок 27ж). Стоит отметить, что в основном это либо короткие, либо полиненасыщенные триглицериды. Больше всего изменяются триглицериды с числом атомов углерода и ненасыщенных связей 52:6, 54:7 и 56:8 (Рисунок 27з). MS-MS анализ показал, что все они содержат в качестве ацильной группы докозагексаеновую кислоту.

Для того чтобы проверить, что L116 влияет на липидный состав клеток через взаимодействие с Cytb5r3, мы сравнили клеточный состав липидов клеток

дикого типа, с инактивацией L116 и с инактивацией митохондриальной формы Cyb5r3 (Рисунок 28а).

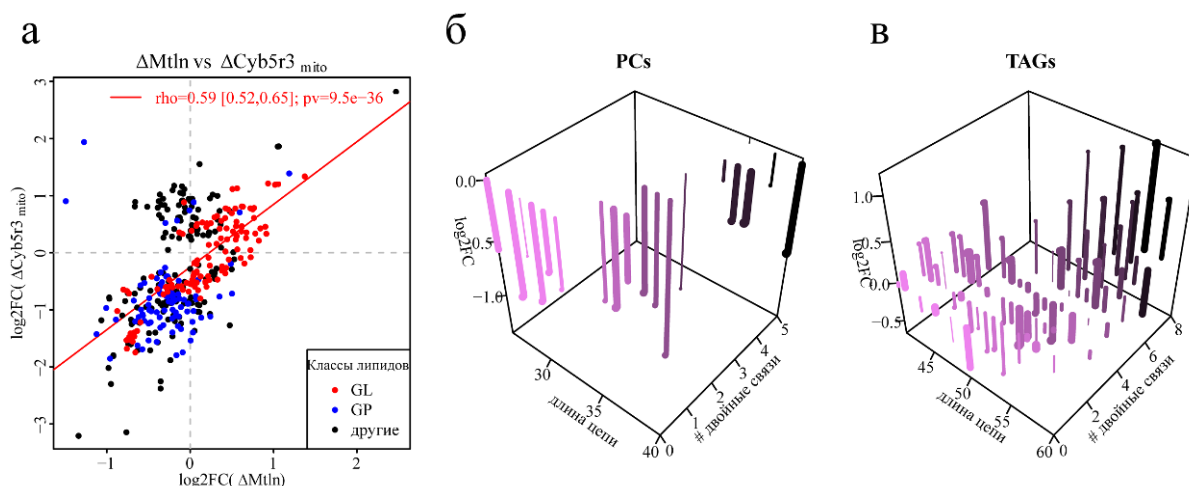


Рисунок 28. (а) Сходство изменений липидного состава клеток при инактивации L116 (ось x) и нарушении митохондриальной локализации Cyb5r3 (ось y). Каждая точка представляет один липид, глицеролипиды (GL), глицерофосфолипиды (GP) и другие классы липидов показаны красным, синим и черным соответственно. Линия регрессии наименьшего квадрата показана красным. Коэффициент корреляции Пирсона, 95% доверительный интервал и значение p (t -критерий) показаны в верхнем левом углу. Изменение количества различных видов фосфатидилхолина (б) или триглицеридов (в) в $\Delta Cyb5r3_{mito}$ (по оси z). Длина цепи и количество двойных связей соответствует координатам x и y. Концентрация TAG и PC в диком типе показана шириной линии (в масштабе шкалы). Липиды, количество которых меняется статистически значимо, отмечены кружком в конце линии. Статистическая значимость была определена с помощью t -теста Стюдента ($p < 0.05$, поправка Бенджамини-Хохберга).

Результаты эксперимента показали, что изменения в липидном составе клеток нокаутных по L116 хорошо коррелируют с изменениями в мутантных клетках $\Delta Cyb5r3_{mito}$. В обоих случаях мы наблюдаем увеличение количества триглицеридов (Рисунок 28б) и снижение количества фосфатидилхолинов (Рисунок 28в). Для клеток с нарушением митохондриальной локализации Cyb5r3 не наблюдается значительного усиления или ослабления изменений липидного состава при дополнительной инактивации или суперэкспрессии гена L116 (Рисунок 29).

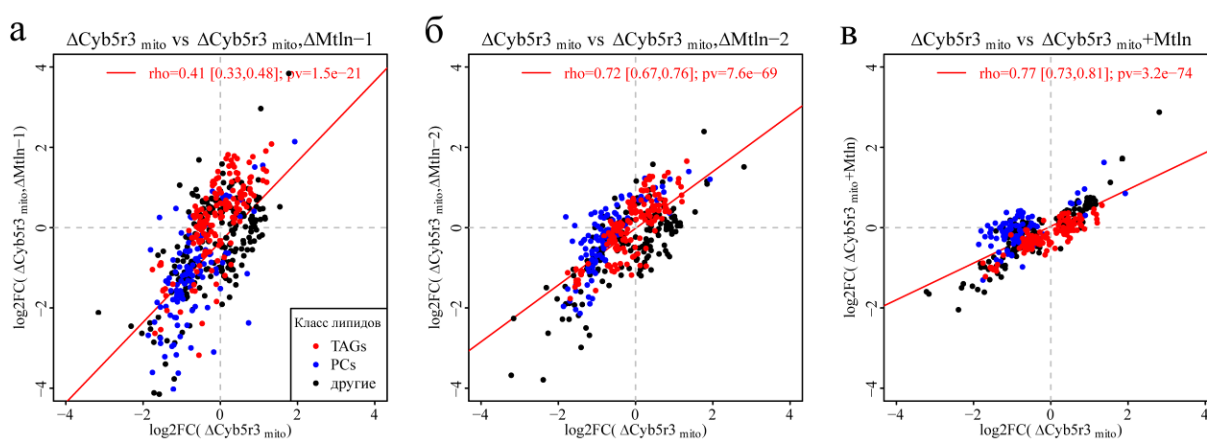


Рисунок 29. Отсутствие выраженного синтетического фенотипа инактивации или суперэкспрессии гена *L116* при нарушении митохондриальной локализации *Cyb5r3*. (а, б) Взаимосвязь между $\log_2$ изменениями липидного состава в клетках $\Delta Cyb5r3_{mito}$ (ось *x*) и одновременной инактивации *L116* и нарушении митохондриальной локализации *Cyb5r3*, $\Delta L116/\Delta Cyb5r3_{mito}$ (ось *y*). Для двойных мутантов $\Delta L116/\Delta Cyb5r3_{mito}$ представлены две независимых клеточных линии. (в) Взаимосвязь между $\log_2$ изменениями липидного состава в клетках $\Delta Cyb5r3_{mito}$ (ось *x*) и одновременной суперэкспрессии *L116* и нарушении митохондриальной формы локализации *Cyb5r3* (ось *y*).. Каждая точка представляет один липид, глицеролипиды (GL), глицерофосфолипиды (GP) и другие классы липидов показаны красным, синим и черным соответственно. Линия регрессии наименьшего квадрата показана красным. Коэффициент корреляции Пирсона, 95% доверительный интервал и значение *p* (*t*-критерий) показаны в верхнем левом углу.

Обсуждение результатов

Недавние исследования показали, что геномы высших организмов кодируют огромное число кОПС-пептидов, которые играют важную роль в различных клеточных процессах [144,171,206]. В данной работе мы исследовали функцию неизученного ранее пептида *L116*. В результате работы нам удалось выяснить, что данный высококонсервативный среди хордовых пептид действительно образуется в клетках, локализуется в митохондриях, принимает участие в клеточном метаболизме и необходим, хотя и не напрямую, для функционирования комплекса I дыхательной цепи.

Большинство кОПС-пептидов выполняют свою функцию посредством взаимодействия с большими белковыми комплексами [167], как было показано для

сарколипина и фосфоламбана [138,207], Toddler/ELABELA [11,163] and DWORF [145]. Мы предположили, что L116 пептид также, взаимодействуя с каким-то белком, регулирует активность комплекса I. Результаты ко-иммунопреципитации L116 за НА-эпитоп показали, что пептид взаимодействует с НАДН цитохром b5 редуктазой 3 (Cyb5r3). Нарушение митохондриальной локализации Cyb5r3 также негативно влияет на активность комплекса I, как и удаление L116. Таким образом, L116, вероятно, является регулятором Cyb5r3, а его действие на комплекс I опосредуется Cyb5r3.

Практически одновременно с нашим исследованием еще две научные группы [208,209] независимо показали, что изучаемый нами пептид существует в клетках и имеет митохондриальную локализацию. Работы независимых групп отчасти подтверждают, отчасти дополняют наши результаты [208]. Было показано, что удаление L116 приводит к уменьшению количества респирасом, содержащих комплексы I, III и IV, а также к накоплению свободного комплекса I, что в результате может приводить к снижению активности комплекса I. Взятые вместе с нашими данными эти результаты позволяют объяснить влияние L116 на активность комплекса I дыхательной цепи только в контексте интактных митохондриальных мембран, но не на активность отдельно выделенного комплекса I. Также опубликованные нашими зарубежными коллегами исследования показали, что L116 также взаимодействует с субъединицами МТР (mitochondrial trifunctional protein) HADHA и HADHB, который принимает участие в β -окислении [209], хотя это взаимодействие не наблюдалось в наших экспериментах. Может быть, подобные различия в спектре партнеров говорят о большем разнообразии функций L116, чем мы предполагаем, а возможно, отражают различия в выборе экспериментальной системы и клеточной линии. Удивительно при этом, что участие L116 в метаболизме жирных кислот является общим логическим выводом из нашей работы и работ наших коллег.

Пока остается открытым вопрос о том, каков конкретный молекулярный механизм влияния L116 на активность Cyb5r3. Cyb5r3 обладает множеством функций. Например, растворимая форма Cyb5r3 принимает участие в восстановлении метгемоглобина [210], мембран-связанная форма важна для

образования ненасыщенных жирных кислот [204] и биосинтеза холестерина [205]. Также недавние исследования показали, что Cyb5r3 может входить в состав амидоксимной системы, которая возможно вовлечена в липогенез [211-213] и метаболизм лекарств.

В результате экспериментов нам удалось выяснить, что L116 не влияет на локализацию или стабильность Cyb5r3. Измерение биологической активности Cyb5r3 является непростой задачей. В простейшей, весьма ограниченной системе, используя в качестве акцептора электронов ферроцианид калия, не удалось обнаружить зависимости от L116 в способности Cyb5r3 окислять НАДН. Однако данный метод является упрощением и не позволяет напрямую измерить активность митохондриальной Cyb5r3 в отношении природных субстратов. Из-за участия Cyb5r3 в метаболизме липидов, разумно было предположить, что взаимодействующий с этим ферментом пептид L116 может изменять ее каталитическую активность в реакциях, связанных с метаболизмом липидов. Поэтому мы и решили сравнить состав липидов в клетках дикого типа, нокаутных по гену L116, и в клетках с инактивацией митохондриальной формы Cyb5r3. Было обнаружено, что количество большого числа триглицеридов повышается, а фосфатидилхолинов понижается как при инактивации гена L116, так и при нарушении митохондриальной локализации Cyb5r3. Данные научной литературы показывают, что липидный состав может очень сильно влиять на функционирование дыхательной цепи, в особенности на активность комплекса I [214-217]. По нашим данным, инактивация L116 сильно уменьшила количество фосфатидилхолинов. В связи с этим, мы предполагаем, что именно снижение количества фосфатидилхолинов может приводить к уменьшению активности комплекса I. В одном из исследований было показано [217], что удаление холина из рациона крыс приводит к накоплению триглицеридов в печени и нарушению функции дыхательной цепи, особенно выраженное для НАДН убихинон оксиредуктазы. Более того важность фосфатидилхолинов для активности комплекса I подтверждается тем фактом, что добавление L-альфа-глицерилфосфорилхолина (GPC) увеличивает активность комплекса I [218]. Еще в одном исследовании утверждается, что активность комплекса I зависит от количества фосфатидилхолина и фосфатидилэтаноламина [219]. Согласно

литературным данным, снижение количества фосфатидилхолинов в митохондриях в результате инактивация фосфатидилхолинового транспортера приводит к снижению активности дыхательной цепи, нарушению формирования респирасом и структуры кристы [220]. Две исследовательские группы независимо показали, что нокаут L116 приводит к нарушению структуры кристы [209] и формирования дыхательных суперкомплексов [208]. Известно, что респирасомы располагаются в мембране кристы [221], внутренняя мембрана которой состоит на 80% из фосфатидилхолинов [222]. Не исключено, что на формирование суперкомплексов может влиять форма кристы, которая, в свою очередь, зависит от состава липидов. Таким образом, данные наблюдения вместе с полученными нами результатами позволяют предположить, что нарушение активности комплекса I при инактивации L116, может определяться изменением активности взаимодействующей с L116 оксидоредуктазы Cytb5r3, что в свою очередь приводит к изменениям в составе митохондриальной мембраны.

Выводы

- Пептид L116 присутствует как в клеточных линиях, так и в живом организме.
- L116 локализуется в митохондриях.
- L116 взаимодействует с цитохром b5 редуктазой 3.
- При инактивации гена пептида L116, и нарушении митохондриальной локализации Cyb5r3, активность комплекса I в клетке уменьшается.
- В нокаутных по гену L116 клетках происходят изменения в липидном составе, что, скорее всего, является причиной уменьшения активности комплекса I.

Список литературы

1. Kastin, A.J. and Pan, W. Concepts for biologically active peptides. // *Curr Pharm Des.* 2010. V. 16. pp. 3390-3400.
2. *Handbook of Biologically Active Peptides.* 2nd ed.
3. de Lecea, L., Kilduff, T.S., Peyron, C., Gao, X., Foye, P.E., Danielson, P.E., Fukuhara, C., Battenberg, E.L., Gautvik, V.T., Bartlett, F.S., 2nd, Frankel, W.N., van den Pol, A.N., Bloom, F.E., Gautvik, K.M. and Sutcliffe, J.G. The hypocretins: hypothalamus-specific peptides with neuroexcitatory activity. // *Proc Natl Acad Sci U S A.* 1998. V. 95. pp. 322-327.
4. Sakurai, T., Amemiya, A., Ishii, M., Matsuzaki, I., Chemelli, R.M., Tanaka, H., Williams, S.C., Richardson, J.A., Kozlowski, G.P., Wilson, S., Arch, J.R., Buckingham, R.E., Haynes, A.C., Carr, S.A., Annan, R.S., McNulty, D.E., Liu, W.S., Terrett, J.A., Elshourbagy, N.A., Bergsma, D.J. and Yanagisawa, M. Orexins and orexin receptors: a family of hypothalamic neuropeptides and G protein-coupled receptors that regulate feeding behavior. // *Cell.* 1998. V. 92. pp. 573-585.
5. Vale, W., Spiess, J., Rivier, C. and Rivier, J. Characterization of a 41-residue ovine hypothalamic peptide that stimulates secretion of corticotropin and beta-endorphin. // *Science.* 1981. V. 213. pp. 1394-1397.
6. Zhang, Y., Proenca, R., Maffei, M., Barone, M., Leopold, L. and Friedman, J.M. Positional cloning of the mouse obese gene and its human homologue. // *Nature.* 1994. V. 372. pp. 425-432.
7. Dong, X., Wang, D., Liu, P., Li, C., Zhao, Q., Zhu, D. and Yu, J. Zm908p11, encoded by a short open reading frame (sORF) gene, functions in pollen tube growth as a profilin ligand in maize. // *J Exp Bot.* 2013. V. 64. pp. 2359-2372.
8. Lee, C., Zeng, J., Drew, B.G., Sallam, T., Martin-Montalvo, A., Wan, J., Kim, S.J., Mehta, H., Hevener, A.L., de Cabo, R. and Cohen, P. The mitochondrial-derived peptide MOTS-c promotes metabolic homeostasis and reduces obesity and insulin resistance. // *Cell Metab.* 2015. V. 21. pp. 443-454.
9. Yosten, G.L., Liu, J., Ji, H., Sandberg, K., Speth, R. and Samson, W.K. A 5'-upstream short open reading frame encoded peptide regulates angiotensin type 1a receptor production and signalling via the beta-arrestin pathway. // *J Physiol.* 2016. V. 594. pp. 1601-1605.
10. Kondo, T., Plaza, S., Zanet, J., Benrabah, E., Valenti, P., Hashimoto, Y., Kobayashi, S., Payre, F. and Kageyama, Y. Small peptides switch the transcriptional activity of Shavenbaby during *Drosophila* embryogenesis. // *Science.* 2010. V. 329. pp. 336-339.
11. Pauli, A., Norris, M.L., Valen, E., Chew, G.L., Gagnon, J.A., Zimmerman, S., Mitchell, A., Ma, J., Dubrulle, J., Reyon, D., Tsai, S.Q., Joung, J.K., Saghatelian, A. and Schier, A.F. Toddler: an embryonic signal that promotes cell movement via Apelin receptors. // *Science.* 2014. V. 343. pp. 1248636.
12. Hashimoto, Y., Niikura, T., Tajima, H., Yasukawa, T., Sudo, H., Ito, Y., Kita, Y., Kawasumi, M., Kouyama, K., Doyu, M., Sobue, G., Koide, T., Tsuji, S., Lang, J., Kurokawa, K. and Nishimoto, I. A rescue factor abolishing neuronal cell death by

- a wide spectrum of familial Alzheimer's disease genes and Abeta. // *Proc Natl Acad Sci U S A*. 2001. V. 98. pp. 6336-6341.
13. Wadler, C.S. and Vanderpool, C.K. A dual function for a bacterial small RNA: SgrS performs base pairing-dependent regulation and encodes a functional polypeptide. // *Proc Natl Acad Sci U S A*. 2007. V. 104. pp. 20454-20459.
 14. Casson, S.A., Chilley, P.M., Topping, J.F., Evans, I.M., Souter, M.A. and Lindsey, K. The POLARIS gene of Arabidopsis encodes a predicted peptide required for correct root growth and leaf vascular patterning. // *Plant Cell*. 2002. V. 14. pp. 1705-1721.
 15. Rohrig, H., Schmidt, J., Miklashevichs, E., Schell, J. and John, M. Soybean ENOD40 encodes two peptides that bind to sucrose synthase. // *Proc Natl Acad Sci U S A*. 2002. V. 99. pp. 1915-1920.
 16. Kastenmayer, J.P., Ni, L., Chu, A., Kitchen, L.E., Au, W.C., Yang, H., Carter, C.D., Wheeler, D., Davis, R.W., Boeke, J.D., Snyder, M.A. and Basrai, M.A. Functional genomics of genes with small open reading frames (sORFs) in *S. cerevisiae*. // *Genome Res*. 2006. V. 16. pp. 365-373.
 17. Gleason, C.A., Liu, Q.L. and Williamson, V.M. Silencing a candidate nematode effector gene corresponding to the tomato resistance gene Mi-1 leads to acquisition of virulence. // *Mol Plant Microbe Interact*. 2008. V. 21. pp. 576-585.
 18. Kondo, T., Hashimoto, Y., Kato, K., Inagaki, S., Hayashi, S. and Kageyama, Y. Small peptide regulators of actin-based cell morphogenesis encoded by a polycistronic mRNA. // *Nat Cell Biol*. 2007. V. 9. pp. 660-665.
 19. Galindo, M.I., Pueyo, J.I., Fouix, S., Bishop, S.A. and Couso, J.P. Peptides encoded by short ORFs control development and define a new eukaryotic gene family. // *PLoS Biol*. 2007. V. 5. pp. e106.
 20. Basrai, M.A., Hieter, P. and Boeke, J.D. Small open reading frames: beautiful needles in the haystack. // *Genome Res*. 1997. V. 7. pp. 768-771.
 21. Frith, M.C., Forrest, A.R., Nourbakhsh, E., Pang, K.C., Kai, C., Kawai, J., Carninci, P., Hayashizaki, Y., Bailey, T.L. and Grimmond, S.M. The abundance of short proteins in the mammalian proteome. // *PLoS Genet*. 2006. V. 2. pp. e52.
 22. Tanaka, M., Sotta, N., Yamazumi, Y., Yamashita, Y., Miwa, K., Murota, K., Chiba, Y., Hirai, M.Y., Akiyama, T., Onouchi, H., Naito, S. and Fujiwara, T. The Minimum Open Reading Frame, AUG-Stop, Induces Boron-Dependent Ribosome Stalling and mRNA Degradation. // *Plant Cell*. 2016. V. 28. pp. 2830-2849.
 23. Ivanov, I.P., Firth, A.E., Michel, A.M., Atkins, J.F. and Baranov, P.V. Identification of evolutionarily conserved non-AUG-initiated N-terminal extensions in human coding sequences. // *Nucleic Acids Res*. 2011. V. 39. pp. 4220-4234.
 24. Badger, J.H. and Olsen, G.J. CRITICA: coding region identification tool invoking comparative analysis. // *Mol Biol Evol*. 1999. V. 16. pp. 512-524.
 25. Kong, L., Zhang, Y., Ye, Z.Q., Liu, X.Q., Zhao, S.Q., Wei, L. and Gao, G. CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. // *Nucleic Acids Res*. 2007. V. 35. pp. W345-349.
 26. Hanada, K., Akiyama, K., Sakurai, T., Toyoda, T., Shinozaki, K. and Shiu, S.H. sORF finder: a program package to identify small open reading frames with high coding potential. // *Bioinformatics*. 2010. V. 26. pp. 399-400.
 27. Siepel, A., Bejerano, G., Pedersen, J.S., Hinrichs, A.S., Hou, M., Rosenbloom, K., Clawson, H., Spieth, J., Hillier, L.W., Richards, S., Weinstock, G.M., Wilson,

- R.K., Gibbs, R.A., Kent, W.J., Miller, W. and Haussler, D. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. // *Genome Res.* 2005. V. 15. pp. 1034-1050.
28. Lin, M.F., Jungreis, I. and Kellis, M. PhyloCSF: a comparative genomics method to distinguish protein coding and non-coding regions. // *Bioinformatics.* 2011. V. 27. pp. i275-282.
 29. Bazzini, A.A., Johnstone, T.G., Christiano, R., Mackowiak, S.D., Obermayer, B., Fleming, E.S., Vejnar, C.E., Lee, M.T., Rajewsky, N., Walther, T.C. and Giraldez, A.J. Identification of small ORFs in vertebrates using ribosome footprinting and evolutionary conservation. // *EMBO J.* 2014. V. 33. pp. 981-993.
 30. Skarszewski, A., Stanton-Cook, M., Huber, T., Al Mansoori, S., Smith, R., Beatson, S.A. and Rothnagel, J.A. uPEPPER: an online tool for upstream open reading frame location and analysis of transcript conservation. // *BMC Bioinformatics.* 2014. V. 15. pp. 36.
 31. Housman, G. and Ulitsky, I. Methods for distinguishing between protein-coding and long noncoding RNAs and the elusive biological purpose of translation of long noncoding RNAs. // *Biochim Biophys Acta.* 2016. V. 1859. pp. 31-40.
 32. Nei, M. and Gojobori, T. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. // *Mol Biol Evol.* 1986. V. 3. pp. 418-426.
 33. Yang, Z. Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. // *Mol Biol Evol.* 1998. V. 15. pp. 568-573.
 34. Lin, M.F., Jungreis, I. and Kellis, M. PhyloCSF: a comparative genomics method to distinguish protein coding and non-coding regions. // *Bioinformatics.* 2011. V. 27. pp. i275-282.
 35. Couso, J.P. Finding smORFs: getting closer. // *Genome Biol.* 2015. V. 16. pp. 189.
 36. Gelfand, M.S. Prediction of function in DNA sequence analysis. // *J Comput Biol.* 1995. V. 2. pp. 87-115.
 37. Alioto, T. and Guigó, R. (2008) In Frishman, D. and Valencia, A. (eds.), *Modern Genome Annotation: the BioSapiens Network*. Springer, Vienna, pp. 7–40.
 38. Louie, E., Ott, J. and Majewski, J. Nucleotide frequency variation across human genes. // *Genome Res.* 2003. V. 13. pp. 2594-2601.
 39. Eddy, S.R. A model of the statistical power of comparative genome sequence analysis. // *PLoS Biol.* 2005. V. 3. pp. e10.
 40. Ingolia, N.T., Ghaemmaghami, S., Newman, J.R. and Weissman, J.S. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. // *Science.* 2009. V. 324. pp. 218-223.
 41. Wolin, S.L. and Walter, P. Ribosome pausing and stacking during translation of a eukaryotic mRNA. // *EMBO J.* 1988. V. 7. pp. 3559-3569.
 42. Chew, G.L., Pauli, A., Rinn, J.L., Regev, A., Schier, A.F. and Valen, E. Ribosome profiling reveals resemblance between long non-coding RNAs and 5' leaders of coding RNAs. // *Development.* 2013. V. 140. pp. 2828-2834.
 43. Smith, J.E., Alvarez-Dominguez, J.R., Kline, N., Huynh, N.J., Geisler, S., Hu, W., Collier, J. and Baker, K.E. Translation of small open reading frames within unannotated RNA transcripts in *Saccharomyces cerevisiae*. // *Cell Rep.* 2014. V. 7. pp. 1858-1866.
 44. Ingolia, N.T., Brar, G.A., Stern-Ginossar, N., Harris, M.S., Talhouarne, G.J., Jackson, S.E., Wills, M.R. and Weissman, J.S. Ribosome profiling reveals

- pervasive translation outside of annotated protein-coding genes. // *Cell Rep.* 2014. V. 8. pp. 1365-1379.
45. Ingolia, N.T., Lareau, L.F. and Weissman, J.S. Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. // *Cell.* 2011. V. 147. pp. 789-802.
 46. Brar, G.A., Yassour, M., Friedman, N., Regev, A., Ingolia, N.T. and Weissman, J.S. High-resolution view of the yeast meiotic program revealed by ribosome profiling. // *Science.* 2012. V. 335. pp. 552-557.
 47. Kim, M.S., Pinto, S.M., Getnet, D., Nirujogi, R.S., Manda, S.S., Chaerkady, R., Madugundu, A.K., Kelkar, D.S., Isserlin, R., Jain, S., Thomas, J.K., Muthusamy, B., Leal-Rojas, P., Kumar, P., Sahasrabudde, N.A., Balakrishnan, L., Advani, J., George, B., Renuse, S., Selvan, L.D., Patil, A.H., Nanjappa, V., Radhakrishnan, A., Prasad, S., Subbannayya, T., Raju, R., Kumar, M., Sreenivasamurthy, S.K., Marimuthu, A., Sathe, G.J., Chavan, S., Datta, K.K., Subbannayya, Y., Sahu, A., Yelamanchi, S.D., Jayaram, S., Rajagopalan, P., Sharma, J., Murthy, K.R., Syed, N., Goel, R., Khan, A.A., Ahmad, S., Dey, G., Mudgal, K., Chatterjee, A., Huang, T.C., Zhong, J., Wu, X., Shaw, P.G., Freed, D., Zahari, M.S., Mukherjee, K.K., Shankar, S., Mahadevan, A., Lam, H., Mitchell, C.J., Shankar, S.K., Satishchandra, P., Schroeder, J.T., Sirdeshmukh, R., Maitra, A., Leach, S.D., Drake, C.G., Halushka, M.K., Prasad, T.S., Hruban, R.H., Kerr, C.L., Bader, G.D., Iacobuzio-Donahue, C.A., Gowda, H. and Pandey, A. A draft map of the human proteome. // *Nature.* 2014. V. 509. pp. 575-581.
 48. Baboo, S. and Cook, P.R. "Dark matter" worlds of unstable RNA and protein. // *Nucleus.* 2014. V. 5. pp. 281-286.
 49. Wilson, B.A. and Masel, J. Putatively noncoding transcripts show extensive association with ribosomes. // *Genome Biol Evol.* 2011. V. 3. pp. 1245-1252.
 50. Raj, A., Wang, S.H., Shim, H., Harpak, A., Li, Y.I., Engelmann, B., Stephens, M., Gilad, Y. and Pritchard, J.K. Thousands of novel translated open reading frames in humans inferred by ribosome footprint profiling. // *Elife.* 2016. V. 5.
 51. Ji, Z., Song, R., Regev, A. and Struhl, K. Many lncRNAs, 5'UTRs, and pseudogenes are translated and some are likely to express functional proteins. // *Elife.* 2015. V. 4. pp. e08890.
 52. Calviello, L., Mukherjee, N., Wyler, E., Zauber, H., Hirsekorn, A., Selbach, M., Landthaler, M., Obermayer, B. and Ohler, U. Detecting actively translated open reading frames in ribosome profiling data. // *Nat Methods.* 2016. V. 13. pp. 165-170.
 53. Chun, S.Y., Rodriguez, C.M., Todd, P.K. and Mills, R.E. SPECtre: a spectral coherence--based classifier of actively translated transcripts from ribosome profiling sequence data. // *BMC Bioinformatics.* 2016. V. 17. pp. 482.
 54. Dunn, J.G. and Weissman, J.S. Plastid: nucleotide-resolution analysis of next-generation sequencing and genomics data. // *BMC Genomics.* 2016. V. 17. pp. 958.
 55. Michel, A.M., Choudhury, K.R., Firth, A.E., Ingolia, N.T., Atkins, J.F. and Baranov, P.V. Observation of dually decoded regions of the human genome using ribosome profiling data. // *Genome Res.* 2012. V. 22. pp. 2219-2229.
 56. Guttman, M., Russell, P., Ingolia, N.T., Weissman, J.S. and Lander, E.S. Ribosome profiling provides evidence that large noncoding RNAs do not encode proteins. // *Cell.* 2013. V. 154. pp. 240-251.

57. Fields, A.P., Rodriguez, E.H., Jovanovic, M., Stern-Ginossar, N., Haas, B.J., Mertins, P., Raychowdhury, R., Hacohen, N., Carr, S.A., Ingolia, N.T., Regev, A. and Weissman, J.S. A Regression-Based Analysis of Ribosome-Profiling Data Reveals a Conserved Complexity to Mammalian Translation. // *Mol Cell*. 2015. V. 60. pp. 816-827.
58. Crappe, J., Ndah, E., Koch, A., Steyaert, S., Gawron, D., De Keulenaer, S., De Meester, E., De Meyer, T., Van Crielinge, W., Van Damme, P. and Menschaert, G. PROTEOFORMER: deep proteome coverage through ribosome profiling and MS integration. // *Nucleic Acids Res*. 2015. V. 43. pp. e29.
59. Lee, S., Liu, B., Lee, S., Huang, S.X., Shen, B. and Qian, S.B. Global mapping of translation initiation sites in mammalian cells at single-nucleotide resolution. // *Proc Natl Acad Sci U S A*. 2012. V. 109. pp. E2424-2432.
60. Gao, X., Wan, J., Liu, B., Ma, M., Shen, B. and Qian, S.B. Quantitative profiling of initiating ribosomes in vivo. // *Nat Methods*. 2015. V. 12. pp. 147-153.
61. Heiman, M., Schaefer, A., Gong, S., Peterson, J.D., Day, M., Ramsey, K.E., Suarez-Farinas, M., Schwarz, C., Stephan, D.A., Surmeier, D.J., Greengard, P. and Heintz, N. A translational profiling approach for the molecular characterization of CNS cell types. // *Cell*. 2008. V. 135. pp. 738-748.
62. Sanz, E., Yang, L., Su, T., Morris, D.R., McKnight, G.S. and Amieux, P.S. Cell-type-specific isolation of ribosome-associated mRNA from complex tissues. // *Proc Natl Acad Sci U S A*. 2009. V. 106. pp. 13939-13944.
63. Aspden, J.L., Eyre-Walker, Y.C., Phillips, R.J., Amin, U., Mumtaz, M.A., Brocard, M. and Couso, J.P. Extensive translation of small Open Reading Frames revealed by Poly-Ribo-Seq. // *Elife*. 2014. V. 3. pp. e03528.
64. Archer, S.K., Shirokikh, N.E., Beilharz, T.H. and Preiss, T. Dynamics of ribosome scanning and recycling revealed by translation complex profiling. // *Nature*. 2016. V. 535. pp. 570-574.
65. Olexiouk, V., Crappe, J., Verbruggen, S., Verhegen, K., Martens, L. and Menschaert, G. sORFs.org: a repository of small ORFs identified by ribosome profiling. // *Nucleic Acids Res*. 2016. V. 44. pp. D324-329.
66. Michel, A.M., Ahern, A.M., Donohue, C.A. and Baranov, P.V. GWIPS-viz as a tool for exploring ribosome profiling evidence supporting the synthesis of alternative proteoforms. // *Proteomics*. 2015. V. 15. pp. 2410-2416.
67. Michel, A.M., Mullan, J.P., Velayudhan, V., O'Connor, P.B., Donohue, C.A. and Baranov, P.V. RiboGalaxy: A browser based platform for the alignment, analysis and visualization of ribosome profiling data. // *RNA Biol*. 2016. V. 13. pp. 316-319.
68. O'Connor, P.B., Andreev, D.E. and Baranov, P.V. Comparative survey of the relative impact of mRNA features on local ribosome profiling read density. // *Nat Commun*. 2016. V. 7. pp. 12915.
69. Bantscheff, M., Lemeer, S., Savitski, M.M. and Kuster, B. Quantitative mass spectrometry in proteomics: critical review update from 2007 to the present. // *Anal Bioanal Chem*. 2012. V. 404. pp. 939-965.
70. Nesvizhskii, A.I. Proteogenomics: concepts, applications and computational strategies. // *Nat Methods*. 2014. V. 11. pp. 1114-1125.
71. Nesvizhskii, A.I. Proteogenomics: concepts, applications and computational strategies. // *Nat Methods*. 2014. V. 11. pp. 1114-1125.

72. Fermin, D., Allen, B.B., Blackwell, T.W., Menon, R., Adamski, M., Xu, Y., Ulintz, P., Omenn, G.S. and States, D.J. Novel gene and gene model detection using a whole genome open reading frame analysis in proteomics. // *Genome Biol.* 2006. V. 7. pp. R35.
73. Khatun, J., Yu, Y., Wrobel, J.A., Risk, B.A., Gunawardena, H.P., Secrest, A., Spitzer, W.J., Xie, L., Wang, L., Chen, X. and Giddings, M.C. Whole human genome proteogenomic mapping for ENCODE cell line data: identifying protein-coding regions. // *BMC Genomics.* 2013. V. 14. pp. 141.
74. Blakeley, P., Overton, I.M. and Hubbard, S.J. Addressing statistical biases in nucleotide-derived protein databases for proteogenomic search strategies. // *J Proteome Res.* 2012. V. 11. pp. 5221-5234.
75. Nesvizhskii, A.I. A survey of computational methods and error rate estimation procedures for peptide and protein identification in shotgun proteomics. // *J Proteomics.* 2010. V. 73. pp. 2092-2123.
76. Krug, K., Carpy, A., Behrends, G., Matic, K., Soares, N.C. and Macek, B. Deep coverage of the Escherichia coli proteome enables the assessment of false discovery rates in simple proteogenomic experiments. // *Mol Cell Proteomics.* 2013. V. 12. pp. 3420-3430.
77. Tanner, S., Shen, Z., Ng, J., Florea, L., Guigo, R., Briggs, S.P. and Bafna, V. Improving gene annotation using peptide mass spectrometry. // *Genome Res.* 2007. V. 17. pp. 231-239.
78. Choudhary, J.S., Blackstock, W.P., Creasy, D.M. and Cottrell, J.S. Interrogating the human genome using uninterpreted mass spectrometry data. // *Proteomics.* 2001. V. 1. pp. 651-667.
79. Edwards, N.J. Novel peptide identification from tandem mass spectra using ESTs and sequence database compression. // *Mol Syst Biol.* 2007. V. 3. pp. 102.
80. Wang, X., Slebos, R.J., Wang, D., Halvey, P.J., Tabb, D.L., Liebler, D.C. and Zhang, B. Protein identification using customized protein sequence databases derived from RNA-Seq data. // *J Proteome Res.* 2012. V. 11. pp. 1009-1017.
81. Slavoff, S.A., Mitchell, A.J., Schwaid, A.G., Cabili, M.N., Ma, J., Levin, J.Z., Karger, A.D., Budnik, B.A., Rinn, J.L. and Saghatelian, A. Peptidomic discovery of short open reading frame-encoded peptides in human cells. // *Nat Chem Biol.* 2013. V. 9. pp. 59-64.
82. Menschaert, G., Van Crielinge, W., Notelaers, T., Koch, A., Crappe, J., Gevaert, K. and Van Damme, P. Deep proteome coverage based on ribosome profiling aids mass spectrometry-based protein and peptide discovery and provides evidence of alternative translation products and near-cognate translation initiation events. // *Mol Cell Proteomics.* 2013. V. 12. pp. 1780-1790.
83. Sheynkman, G.M., Johnson, J.E., Jagtap, P.D., Shortreed, M.R., Onsongo, G., Frey, B.L., Griffin, T.J. and Smith, L.M. Using Galaxy-P to leverage RNA-Seq for the discovery of novel protein variations. // *BMC Genomics.* 2014. V. 15. pp. 703.
84. Branca, R.M., Orre, L.M., Johansson, H.J., Granholm, V., Huss, M., Perez-Bercoff, A., Forshed, J., Kall, L. and Lehtio, J. HiRIEF LC-MS enables deep proteome coverage and unbiased proteogenomics. // *Nat Methods.* 2014. V. 11. pp. 59-62.
85. Ma, J., Ward, C.C., Jungreis, I., Slavoff, S.A., Schwaid, A.G., Neveu, J., Budnik, B.A., Kellis, M. and Saghatelian, A. Discovery of human sORF-encoded

- polypeptides (SEPs) in cell lines and tissue. // *J Proteome Res.* 2014. V. 13. pp. 1757-1765.
86. Young, S.K. and Wek, R.C. Upstream Open Reading Frames Differentially Regulate Gene-specific Translation in the Integrated Stress Response. // *J Biol Chem.* 2016. V. 291. pp. 16927-16935.
 87. Neafsey, D.E. and Galagan, J.E. Dual modes of natural selection on upstream open reading frames. // *Mol Biol Evol.* 2007. V. 24. pp. 1744-1751.
 88. Crowe, M.L., Wang, X.Q. and Rothnagel, J.A. Evidence for conservation and selection of upstream open reading frames suggests probable encoding of bioactive peptides. // *BMC Genomics.* 2006. V. 7. pp. 16.
 89. Calvo, S.E., Pagliarini, D.J. and Mootha, V.K. Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. // *Proc Natl Acad Sci U S A.* 2009. V. 106. pp. 7507-7512.
 90. Suzuki, Y., Ishihara, D., Sasaki, M., Nakagawa, H., Hata, H., Tsunoda, T., Watanabe, M., Komatsu, T., Ota, T., Isogai, T., Suyama, A. and Sugano, S. Statistical analysis of the 5' untranslated region of human mRNA using "Oligo-Capped" cDNA libraries. // *Genomics.* 2000. V. 64. pp. 286-297.
 91. Iacono, M., Mignone, F. and Pesole, G. uAUG and uORFs in human and rodent 5'untranslated mRNAs. // *Gene.* 2005. V. 349. pp. 97-105.
 92. Morris, D.R. and Geballe, A.P. Upstream open reading frames as regulators of mRNA translation. // *Mol Cell Biol.* 2000. V. 20. pp. 8635-8642.
 93. Spriggs, K.A., Bushell, M. and Willis, A.E. Translational regulation of gene expression during conditions of cell stress. // *Mol Cell.* 2010. V. 40. pp. 228-237.
 94. Kozak, M. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. // *Nucleic Acids Res.* 1987. V. 15. pp. 8125-8148.
 95. Kozak, M. An analysis of vertebrate mRNA sequences: intimations of translational control. // *J Cell Biol.* 1991. V. 115. pp. 887-903.
 96. Kondo, S., Schutte, B.C., Richardson, R.J., Bjork, B.C., Knight, A.S., Watanabe, Y., Howard, E., de Lima, R.L., Daack-Hirsch, S., Sander, A., McDonald-McGinn, D.M., Zackai, E.H., Lammer, E.J., Aylsworth, A.S., Ardinger, H.H., Lidral, A.C., Pober, B.R., Moreno, L., Arcos-Burgos, M., Valencia, C., Houdayer, C., Bahuau, M., Moretti-Ferreira, D., Richieri-Costa, A., Dixon, M.J. and Murray, J.C. Mutations in IRF6 cause Van der Woude and popliteal pterygium syndromes. // *Nat Genet.* 2002. V. 32. pp. 285-289.
 97. Witt, H., Luck, W., Hennies, H.C., Classen, M., Kage, A., Lass, U., Landt, O. and Becker, M. Mutations in the gene encoding the serine protease inhibitor, Kazal type 1 are associated with chronic pancreatitis. // *Nat Genet.* 2000. V. 25. pp. 213-216.
 98. Liu, L., Dilworth, D., Gao, L., Monzon, J., Summers, A., Lassam, N. and Hogg, D. Mutation of the CDKN2A 5' UTR creates an aberrant initiation codon and predisposes to melanoma. // *Nat Genet.* 1999. V. 21. pp. 128-132.
 99. Cazzola, M. and Skoda, R.C. Translational pathophysiology: a novel molecular mechanism of human disease. // *Blood.* 2000. V. 95. pp. 3280-3288.
 100. Barbosa, C., Peixeiro, I. and Romao, L. Gene expression regulation by upstream open reading frames and human disease. // *PLoS Genet.* 2013. V. 9. pp. e1003529.
 101. Calvo, S.E., Pagliarini, D.J. and Mootha, V.K. Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. // *Proc Natl Acad Sci U S A.* 2009. V. 106. pp. 7507-7512.

102. Rogers, G.W., Jr., Edelman, G.M. and Mauro, V.P. Differential utilization of upstream AUGs in the beta-secretase mRNA suggests that a shunting mechanism regulates translation. // *Proc Natl Acad Sci U S A*. 2004. V. 101. pp. 2794-2799.
103. Morris, D.R. and Geballe, A.P. Upstream open reading frames as regulators of mRNA translation. // *Mol Cell Biol*. 2000. V. 20. pp. 8635-8642.
104. Mendell, J.T., Sharifi, N.A., Meyers, J.L., Martinez-Murillo, F. and Dietz, H.C. Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. // *Nat Genet*. 2004. V. 36. pp. 1073-1078.
105. Yepiskoposyan, H., Aeschmann, F., Nilsson, D., Okoniewski, M. and Muhlemann, O. Autoregulation of the nonsense-mediated mRNA decay pathway in human cells. // *RNA*. 2011. V. 17. pp. 2108-2118.
106. Ruiz-Echevarria, M.J. and Peltz, S.W. The RNA binding protein Pub1 modulates the stability of transcripts containing upstream open reading frames. // *Cell*. 2000. V. 101. pp. 741-751.
107. Spriggs, K.A., Bushell, M. and Willis, A.E. Translational regulation of gene expression during conditions of cell stress. // *Mol Cell*. 2010. V. 40. pp. 228-237.
108. Andreev, D.E., O'Connor, P.B., Fahey, C., Kenny, E.M., Terenin, I.M., Dmitriev, S.E., Cormican, P., Morris, D.W., Shatsky, I.N. and Baranov, P.V. Translation of 5' leaders is pervasive in genes resistant to eIF2 repression. // *Elife*. 2015. V. 4. pp. e03971.
109. Wiese, A., Elzinga, N., Wobbes, B. and Smeekens, S. A conserved upstream open reading frame mediates sucrose-induced repression of translation. // *Plant Cell*. 2004. V. 16. pp. 1717-1729.
110. Fang, P., Wang, Z. and Sachs, M.S. Evolutionarily conserved features of the arginine attenuator peptide provide the necessary requirements for its function in translational regulation. // *J Biol Chem*. 2000. V. 275. pp. 26710-26719.
111. Hanfrey, C., Elliott, K.A., Franceschetti, M., Mayer, M.J., Illingworth, C. and Michael, A.J. A dual upstream open reading frame-based autoregulatory circuit controlling polyamine-responsive translation. // *J Biol Chem*. 2005. V. 280. pp. 39229-39237.
112. Law, G.L., Raney, A., Heusner, C. and Morris, D.R. Polyamine regulation of ribosome pausing at the upstream open reading frame of S-adenosylmethionine decarboxylase. // *J Biol Chem*. 2001. V. 276. pp. 38036-38043.
113. Raney, A., Law, G.L., Mize, G.J. and Morris, D.R. Regulated translation termination at the upstream open reading frame in s-adenosylmethionine decarboxylase mRNA. // *J Biol Chem*. 2002. V. 277. pp. 5988-5994.
114. Akimoto, C., Sakashita, E., Kasashima, K., Kuroiwa, K., Tominaga, K., Hamamoto, T. and Endo, H. Translational repression of the McKusick-Kaufman syndrome transcript by unique upstream open reading frames encoding mitochondrial proteins with alternative polyadenylation sites. // *Biochim Biophys Acta*. 2013. V. 1830. pp. 2728-2738.
115. Saghatelian, A. and Couso, J.P. Discovery and characterization of smORF-encoded bioactive polypeptides. // *Nat Chem Biol*. 2015. V. 11. pp. 909-916.
116. Kochetov, A.V. Alternative translation start sites and hidden coding potential of eukaryotic mRNAs. // *Bioessays*. 2008. V. 30. pp. 683-691.
117. Mouilleron, H., Delcourt, V. and Roucou, X. Death of a dogma: eukaryotic mRNAs can code for more than one protein. // *Nucleic Acids Res*. 2016. V. 44. pp. 14-23.

118. Wang, R.F., Parkhurst, M.R., Kawakami, Y., Robbins, P.F. and Rosenberg, S.A. Utilization of an alternative open reading frame of a normal gene in generating a novel human cancer antigen. // *J Exp Med*. 1996. V. 183. pp. 1131-1140.
119. Rosenberg, S.A., Tong-On, P., Li, Y., Riley, J.P., El-Gamil, M., Parkhurst, M.R. and Robbins, P.F. Identification of BING-4 cancer antigen translated from an alternative open reading frame of a gene in the extended MHC class II region using lymphocytes from a patient with a durable complete regression following immunotherapy. // *J Immunol*. 2002. V. 168. pp. 2402-2407.
120. Ronsin, C., Chung-Scott, V., Poullion, I., Aknouche, N., Gaudin, C. and Triebel, F. A non-AUG-defined alternative open reading frame of the intestinal carboxyl esterase mRNA generates an epitope recognized by renal cell carcinoma-reactive tumor-infiltrating lymphocytes in situ. // *J Immunol*. 1999. V. 163. pp. 483-490.
121. Huang, J., El-Gamil, M., Dudley, M.E., Li, Y.F., Rosenberg, S.A. and Robbins, P.F. T cells associated with tumor regression recognize frameshifted products of the CDKN2A tumor suppressor gene locus and a mutated HLA class I gene product. // *J Immunol*. 2004. V. 172. pp. 6057-6064.
122. Quelle, D.E., Zindy, F., Ashmun, R.A. and Sherr, C.J. Alternative reading frames of the INK4a tumor suppressor gene encode two unrelated proteins capable of inducing cell cycle arrest. // *Cell*. 1995. V. 83. pp. 993-1000.
123. Bab, I., Smith, E., Gavish, H., Attar-Namdar, M., Chorev, M., Chen, Y.C., Muhlrads, A., Birnbaum, M.J., Stein, G. and Frenkel, B. Biosynthesis of osteogenic growth peptide via alternative translational initiation at AUG85 of histone H4 mRNA. // *J Biol Chem*. 1999. V. 274. pp. 14474-14481.
124. Klemke, M., Kehlenbach, R.H. and Huttner, W.B. Two overlapping reading frames in a single exon encode interacting proteins--a novel way of gene usage. // *EMBO J*. 2001. V. 20. pp. 3849-3860.
125. Abramowitz, J., Grenet, D., Birnbaumer, M., Torres, H.N. and Birnbaumer, L. XLalphas, the extra-long form of the alpha-subunit of the Gs G protein, is significantly longer than suspected, and so is its companion Alex. // *Proc Natl Acad Sci U S A*. 2004. V. 101. pp. 8366-8371.
126. Vanderperre, B., Staskevicius, A.B., Tremblay, G., McCoy, M., O'Neill, M.A., Cashman, N.R. and Roucou, X. An overlapping reading frame in the PRNP gene encodes a novel polypeptide distinct from the prion protein. // *FASEB J*. 2011. V. 25. pp. 2373-2386.
127. Bergeron, D., Lapointe, C., Bissonnette, C., Tremblay, G., Motard, J. and Roucou, X. An out-of-frame overlapping reading frame in the ataxin-1 coding sequence encodes a novel ataxin-1 interacting protein. // *J Biol Chem*. 2013. V. 288. pp. 21824-21835.
128. Vanderperre, B., Lucier, J.F., Bissonnette, C., Motard, J., Tremblay, G., Vanderperre, S., Wisztorski, M., Salzert, M., Boisvert, F.M. and Roucou, X. Direct detection of alternative open reading frames translation products in human significantly expands the proteome. // *PLoS One*. 2013. V. 8. pp. e70698.
129. Mackowiak, S.D., Zauber, H., Bielow, C., Thiel, D., Kutz, K., Calviello, L., Mastrobuoni, G., Rajewsky, N., Kempa, S., Selbach, M. and Obermayer, B. Extensive identification and analysis of conserved small ORFs in animals. // *Genome Biol*. 2015. V. 16. pp. 179.

130. Dunn, J.G., Foo, C.K., Belletier, N.G., Gavis, E.R. and Weissman, J.S. Ribosome profiling reveals pervasive and regulated stop codon readthrough in *Drosophila melanogaster*. // *Elife*. 2013. V. 2. pp. e01179.
131. Arribere, J.A., Cenik, E.S., Jain, N., Hess, G.T., Lee, C.H., Bassik, M.C. and Fire, A.Z. Translation readthrough mitigation. // *Nature*. 2016. V. 534. pp. 719-723.
132. Miettinen, T.P. and Bjorklund, M. Modified ribosome profiling reveals high abundance of ribosome protected mRNA fragments derived from 3' untranslated regions. // *Nucleic Acids Res*. 2015. V. 43. pp. 1019-1034.
133. Gascoigne, D.K., Cheetham, S.W., Cattenoz, P.B., Clark, M.B., Amaral, P.P., Taft, R.J., Wilhelm, D., Dinger, M.E. and Mattick, J.S. Pinstripe: a suite of programs for integrating transcriptomic and proteomic datasets identifies novel proteins and improves differentiation of protein-coding and non-coding genes. // *Bioinformatics*. 2012. V. 28. pp. 3042-3050.
134. Prabakaran, S., Hemberg, M., Chauhan, R., Winter, D., Tweedie-Cullen, R.Y., Dittrich, C., Hong, E., Gunawardena, J., Steen, H., Kreiman, G. and Steen, J.A. Quantitative profiling of peptides from RNAs classified as noncoding. // *Nat Commun*. 2014. V. 5. pp. 5429.
135. Malarkannan, S., Shih, P.P., Eden, P.A., Horng, T., Zuberi, A.R., Christianson, G., Roopenian, D. and Shastri, N. The molecular and functional characterization of a dominant minor H antigen, H60. // *J Immunol*. 1998. V. 161. pp. 3501-3509.
136. Schwab, S.R., Li, K.C., Kang, C. and Shastri, N. Constitutive display of cryptic translation products by MHC class I molecules. // *Science*. 2003. V. 301. pp. 1367-1371.
137. Magny, E.G., Pueyo, J.I., Pearl, F.M., Cespedes, M.A., Niven, J.E., Bishop, S.A. and Couso, J.P. Conserved regulation of cardiac calcium uptake by peptides encoded in small open reading frames. // *Science*. 2013. V. 341. pp. 1116-1120.
138. Anderson, D.M., Anderson, K.M., Chang, C.L., Makarewich, C.A., Nelson, B.R., McAnally, J.R., Kasaragod, P., Shelton, J.M., Liou, J., Bassel-Duby, R. and Olson, E.N. A micropeptide encoded by a putative long noncoding RNA regulates muscle performance. // *Cell*. 2015. V. 160. pp. 595-606.
139. Matsumoto, A., Pasut, A., Matsumoto, M., Yamashita, R., Fung, J., Monteleone, E., Saghatelian, A., Nakayama, K.I., Clohessy, J.G. and Pandolfi, P.P. mTORC1 and muscle regeneration are regulated by the LINC00961-encoded SPAR polypeptide. // *Nature*. 2017. V. 541. pp. 228-232.
140. Nelson, B.R., Makarewich, C.A., Anderson, D.M., Winders, B.R., Troupes, C.D., Wu, F., Reese, A.L., McAnally, J.R., Chen, X., Kavalali, E.T., Cannon, S.C., Houser, S.R., Bassel-Duby, R. and Olson, E.N. A peptide encoded by a transcript annotated as long noncoding RNA enhances SERCA activity in muscle. // *Science*. 2016. V. 351. pp. 271-275.
141. Hanyu-Nakamura, K., Sonobe-Nojima, H., Tanigawa, A., Lasko, P. and Nakamura, A. *Drosophila* Pgc protein inhibits P-TEFb recruitment to chromatin in primordial germ cells. // *Nature*. 2008. V. 451. pp. 730-733.
142. Martinho, R.G., Kunwar, P.S., Casanova, J. and Lehmann, R. A noncoding RNA is required for the repression of RNAPolII-dependent transcription in primordial germ cells. // *Curr Biol*. 2004. V. 14. pp. 159-165.
143. Zanet, J., Benrabah, E., Li, T., Pelissier-Monier, A., Chanut-Delalande, H., Ronsin, B., Bellen, H.J., Payre, F. and Plaza, S. Pri sORF peptides induce

- selective proteasome-mediated protein processing. // *Science*. 2015. V. 349. pp. 1356-1358.
144. Andrews, S.J. and Rothnagel, J.A. Emerging evidence for functional peptides encoded by short open reading frames. // *Nat Rev Genet*. 2014. V. 15. pp. 193-204.
 145. Nelson, B.R., Makarewich, C.A., Anderson, D.M., Winders, B.R., Troupes, C.D., Wu, F., Reese, A.L., McAnally, J.R., Chen, X., Kavalali, E.T., Cannon, S.C., Houser, S.R., Bassel-Duby, R. and Olson, E.N. A peptide encoded by a transcript annotated as long noncoding RNA enhances SERCA activity in muscle. // *Science*. 2016. V. 351. pp. 271-275.
 146. Kranias, E.G. and Hajjar, R.J. Modulation of cardiac contractility by the phospholamban/SERCA2a regulatome. // *Circ Res*. 2012. V. 110. pp. 1646-1660.
 147. Anderson, D.M., Makarewich, C.A., Anderson, K.M., Shelton, J.M., Bezprozvannaya, S., Bassel-Duby, R. and Olson, E.N. Widespread control of calcium signaling by a family of SERCA-inhibiting micropeptides. // *Sci Signal*. 2016. V. 9. pp. ra119.
 148. Salzman, J., Gawad, C., Wang, P.L., Lacayo, N. and Brown, P.O. Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types. // *PLoS One*. 2012. V. 7. pp. e30733.
 149. Jeck, W.R., Sorrentino, J.A., Wang, K., Slevin, M.K., Burd, C.E., Liu, J., Marzluff, W.F. and Sharpless, N.E. Circular RNAs are abundant, conserved, and associated with ALU repeats. // *RNA*. 2013. V. 19. pp. 141-157.
 150. Memczak, S., Jens, M., Elefsinioti, A., Torti, F., Krueger, J., Rybak, A., Maier, L., Mackowiak, S.D., Gregersen, L.H., Munschauer, M., Loewer, A., Ziebold, U., Landthaler, M., Kocks, C., le Noble, F. and Rajewsky, N. Circular RNAs are a large class of animal RNAs with regulatory potency. // *Nature*. 2013. V. 495. pp. 333-338.
 151. Salzman, J., Chen, R.E., Olsen, M.N., Wang, P.L. and Brown, P.O. Cell-type specific features of circular RNA expression. // *PLoS Genet*. 2013. V. 9. pp. e1003777.
 152. Chen, L.L. The biogenesis and emerging roles of circular RNAs. // *Nat Rev Mol Cell Biol*. 2016. V. 17. pp. 205-211.
 153. Pamudurti, N.R., Bartok, O., Jens, M., Ashwal-Fluss, R., Stottmeister, C., Ruhe, L., Hanan, M., Wyler, E., Perez-Hernandez, D., Ramberger, E., Shenzen, S., Samson, M., Dittmar, G., Landthaler, M., Chekulaeva, M., Rajewsky, N. and Kadener, S. Translation of CircRNAs. // *Mol Cell*. 2017. V. 66. pp. 9-21 e27.
 154. Legnini, I., Di Timoteo, G., Rossi, F., Morlando, M., Briganti, F., Sthandier, O., Fatica, A., Santini, T., Andronache, A., Wade, M., Laneve, P., Rajewsky, N. and Bozzoni, I. Circ-ZNF609 Is a Circular RNA that Can Be Translated and Functions in Myogenesis. // *Mol Cell*. 2017. V. 66. pp. 22-37 e29.
 155. Lauressergues, D., Couzigou, J.M., Clemente, H.S., Martinez, Y., Dunand, C., Becard, G. and Combier, J.P. Primary transcripts of microRNAs encode regulatory peptides. // *Nature*. 2015. V. 520. pp. 90-93.
 156. Yoshikawa, M., Iki, T., Numa, H., Miyashita, K., Meshi, T. and Ishikawa, M. A Short Open Reading Frame Encompassing the MicroRNA173 Target Site Plays a Role in trans-Acting Small Interfering RNA Biogenesis. // *Plant Physiol*. 2016. V. 171. pp. 359-368.

157. Yen, K., Lee, C., Mehta, H. and Cohen, P. The emerging role of the mitochondrial-derived peptide humanin in stress resistance. // *J Mol Endocrinol*. 2013. V. 50. pp. R11-19.
158. Yamagishi, Y., Hashimoto, Y., Niikura, T. and Nishimoto, I. Identification of essential amino acids in Humanin, a neuroprotective factor against Alzheimer's disease-relevant insults. // *Peptides*. 2003. V. 24. pp. 585-595.
159. Guo, B., Zhai, D., Cabezas, E., Welsh, K., Nouraini, S., Satterthwait, A.C. and Reed, J.C. Humanin peptide suppresses apoptosis by interfering with Bax activation. // *Nature*. 2003. V. 423. pp. 456-461.
160. Zhai, D., Luciano, F., Zhu, X., Guo, B., Satterthwait, A.C. and Reed, J.C. Humanin binds and nullifies Bid activity by blocking its activation of Bax and Bak. // *J Biol Chem*. 2005. V. 280. pp. 15815-15824.
161. Lee, C., Yen, K. and Cohen, P. Humanin: a harbinger of mitochondrial-derived peptides? // *Trends Endocrinol Metab*. 2013. V. 24. pp. 222-228.
162. Gong, Z., Tas, E. and Muzumdar, R. Humanin and age-related diseases: a new link? // *Front Endocrinol (Lausanne)*. 2014. V. 5. pp. 210.
163. Chng, S.C., Ho, L., Tian, J. and Reversade, B. ELABELA: a hormone essential for heart development signals via the apelin receptor. // *Dev Cell*. 2013. V. 27. pp. 672-680.
164. Kondo, T., Hashimoto, Y., Kato, K., Inagaki, S., Hayashi, S. and Kageyama, Y. Small peptide regulators of actin-based cell morphogenesis encoded by a polycistronic mRNA. // *Nat Cell Biol*. 2007. V. 9. pp. 660-665.
165. Slavoff, S.A., Heo, J., Budnik, B.A., Hanakahi, L.A. and Saghatelian, A. A human short open reading frame (sORF)-encoded polypeptide that stimulates DNA end joining. // *J Biol Chem*. 2014. V. 289. pp. 10950-10957.
166. D'Lima, N.G., Ma, J., Winkler, L., Chu, Q., Loh, K.H., Corpuz, E.O., Budnik, B.A., Lykke-Andersen, J., Saghatelian, A. and Slavoff, S.A. A human microprotein that interacts with the mRNA decapping complex. // *Nat Chem Biol*. 2017. V. 13. pp. 174-180.
167. Cabrera-Quio, L.E., Herberg, S. and Pauli, A. Decoding sORF translation - from small proteins to gene regulation. // *RNA Biol*. 2016. V. 13. pp. 1051-1059.
168. Raney, A., Baron, A.C., Mize, G.J., Law, G.L. and Morris, D.R. In vitro translation of the upstream open reading frame in the mammalian mRNA encoding S-adenosylmethionine decarboxylase. // *J Biol Chem*. 2000. V. 275. pp. 24444-24450.
169. Akimoto, C., Sakashita, E., Kasashima, K., Kuroiwa, K., Tominaga, K., Hamamoto, T. and Endo, H. Translational repression of the McKusick-Kaufman syndrome transcript by unique upstream open reading frames encoding mitochondrial proteins with alternative polyadenylation sites. // *Biochim Biophys Acta*. 2013. V. 1830. pp. 2728-2738.
170. Banfai, B., Jia, H., Khatun, J., Wood, E., Risk, B., Gundling, W.E., Jr., Kundaje, A., Gunawardena, H.P., Yu, Y., Xie, L., Krajewski, K., Strahl, B.D., Chen, X., Bickel, P., Giddings, M.C., Brown, J.B. and Lipovich, L. Long noncoding RNAs are rarely translated in two human cell lines. // *Genome Res*. 2012. V. 22. pp. 1646-1657.
171. Makarewich, C.A. and Olson, E.N. Mining for Micropeptides. // *Trends Cell Biol*. 2017. V. 27. pp. 685-696.

172. Dyson, M.R., Shadbolt, S.P., Vincent, K.J., Perera, R.L. and McCafferty, J. Production of soluble mammalian proteins in *Escherichia coli*: identification of protein features that correlate with successful expression. // *BMC Biotechnol.* 2004. V. 4. pp. 32.
173. Gibson, T.J., Seiler, M. and Veitia, R.A. The transience of transient overexpression. // *Nat Methods.* 2013. V. 10. pp. 715-721.
174. Kim, D.I. and Roux, K.J. Filling the Void: Proximity-Based Labeling of Proteins in Living Cells. // *Trends Cell Biol.* 2016. V. 26. pp. 804-817.
175. Rhee, H.W., Zou, P., Udeshi, N.D., Martell, J.D., Mootha, V.K., Carr, S.A. and Ting, A.Y. Proteomic mapping of mitochondria in living cells via spatially restricted enzymatic tagging. // *Science.* 2013. V. 339. pp. 1328-1331.
176. Roux, K.J., Kim, D.I., Raida, M. and Burke, B. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. // *J Cell Biol.* 2012. V. 196. pp. 801-810.
177. Varnaite, R. and MacNeill, S.A. Meet the neighbors: Mapping local protein interactomes by proximity-dependent labeling with BioID. // *Proteomics.* 2016. V. 16. pp. 2503-2518.
178. Lam, S.S., Martell, J.D., Kamer, K.J., Deerinck, T.J., Ellisman, M.H., Mootha, V.K. and Ting, A.Y. Directed evolution of APEX2 for electron microscopy and proximity labeling. // *Nat Methods.* 2015. V. 12. pp. 51-54.
179. Li, X.W., Rees, J.S., Xue, P., Zhang, H., Hamaia, S.W., Sanderson, B., Funk, P.E., Farndale, R.W., Lilley, K.S., Perrett, S. and Jackson, A.P. New insights into the DT40 B cell receptor cluster using a proteomic proximity labeling assay. // *J Biol Chem.* 2014. V. 289. pp. 14434-14447.
180. Higashiyama, S. Lighting up by EMARS. // *J Biochem.* 2012. V. 152. pp. 381-382.
181. Lonn, P. and Landegren, U. Close Encounters - Probing Proximal Proteins in Live or Fixed Cells. // *Trends Biochem Sci.* 2017. V. 42. pp. 504-515.
182. Roux, K.J. Marked by association: techniques for proximity-dependent labeling of proteins in eukaryotic cells. // *Cell Mol Life Sci.* 2013. V. 70. pp. 3657-3664.
183. Fernandez-Suarez, M., Chen, T.S. and Ting, A.Y. Protein-protein interaction detection in vitro and in cells by proximity biotinylation. // *J Am Chem Soc.* 2008. V. 130. pp. 9251-9253.
184. Miller, K.E., Kim, Y., Huh, W.K. and Park, H.O. Bimolecular Fluorescence Complementation (BiFC) Analysis: Advances and Recent Applications for Genome-Wide Interaction Studies. // *J Mol Biol.* 2015. V. 427. pp. 2039-2055.
185. Piston, D.W. and Kremers, G.J. Fluorescent protein FRET: the good, the bad and the ugly. // *Trends Biochem Sci.* 2007. V. 32. pp. 407-414.
186. Petrova, O.A., Mantsyzov, A.B., Rodina, E.V., Efimov, S.V., Hackenberg, C., Hakanpaa, J., Klochkov, V.V., Lebedev, A.A., Chugunova, A.A., Malyavko, A.N., Zatsepin, T.S., Mishin, A.V., Zvereva, M.I., Lamzin, V.S., Dontsova, O.A. and Polshakov, V.I. Structure and function of the N-terminal domain of the yeast telomerase reverse transcriptase. // *Nucleic Acids Res.* 2018. V. 46. pp. 1525-1540.
187. Smirnova, V.V., Shestakova, E.D., Bikmetov, D.V., Chugunova, A.A., Osterman, I.A., Serebryakova, M.V., Sergeeva, O.V., Zatsepin, T.S., Shatsky, I.N. and Terenin, I.M. eIF4G2 balances its own mRNA translation via a PCBP2-based feedback loop. // *RNA.* 2019. V. 25. pp. 757-767.

188. Chugunova, A., Loseva, E., Mazin, P., Mitina, A., Navalayeu, T., Bilan, D., Vishnyakova, P., Marey, M., Golovina, A., Serebryakova, M., Pletnev, P., Rubtsova, M., Mair, W., Vanyushkina, A., Khaitovich, P., Belousov, V., Vysokikh, M., Sergiev, P. and Dontsova, O. LINC00116 codes for a mitochondrial peptide linking respiration and lipid metabolism. // *Proc Natl Acad Sci U S A*. 2019. V. 116. pp. 4940-4945.
189. Blanchette, M., Kent, W.J., Riemer, C., Elnitski, L., Smit, A.F., Roskin, K.M., Baertsch, R., Rosenbloom, K., Clawson, H., Green, E.D., Haussler, D. and Miller, W. Aligning multiple genomic sequences with the threaded blockset aligner. // *Genome Res*. 2004. V. 14. pp. 708-715.
190. Michel, A.M., Fox, G., A, M.K., De Bo, C., O'Connor, P.B., Heaphy, S.M., Mullan, J.P., Donohue, C.A., Higgins, D.G. and Baranov, P.V. GWIPS-viz: development of a ribo-seq genome browser. // *Nucleic Acids Res*. 2014. V. 42. pp. D859-864.
191. Doudna, J.A. and Charpentier, E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. // *Science*. 2014. V. 346. pp. 1258096.
192. Chugunova, A.A., Dontsova, O.A. and Sergiev, P.V. Methods of Genome Engineering: a New Era of Molecular Biology. // *Biochemistry (Mosc)*. 2016. V. 81. pp. 662-677.
193. Mansour, W.Y., Rhein, T. and Dahm-Daphi, J. The alternative end-joining pathway for repair of DNA double-strand breaks requires PARP1 but is not dependent upon microhomologies. // *Nucleic Acids Res*. 2010. V. 38. pp. 6065-6077.
194. Izsvak, Z. and Ivics, Z. Sleeping beauty transposition: biology and applications for molecular therapy. // *Mol Ther*. 2004. V. 9. pp. 147-156.
195. Wisniewski, J.R. and Mann, M. A Proteomics Approach to the Protein Normalization Problem: Selection of Unvarying Proteins for MS-Based Proteomics and Western Blotting. // *J Proteome Res*. 2016. V. 15. pp. 2321-2326.
196. Sharma, L.K., Lu, J. and Bai, Y. Mitochondrial respiratory complex I: structure, function and implication in human diseases. // *Curr Med Chem*. 2009. V. 16. pp. 1266-1277.
197. Martin-Montalvo, A., Sun, Y., Diaz-Ruiz, A., Ali, A., Gutierrez, V., Palacios, H.H., Curtis, J., Siendones, E., Ariza, J., Abulwerdi, G.A., Sun, X., Wang, A.X., Pearson, K.J., Fishbein, K.W., Spencer, R.G., Wang, M., Han, X., Scheibye-Knudsen, M., Baur, J.A., Shertzer, H.G., Navas, P., Villalba, J.M., Zou, S., Bernier, M. and de Cabo, R. Cytochrome b5 reductase and the control of lipid metabolism and healthspan. // *NPJ Aging Mech Dis*. 2016. V. 2. pp. 16006.
198. Borgese, N., Aggujaro, D., Carrera, P., Pietrini, G. and Bassetti, M. A role for N-myristoylation in protein targeting: NADH-cytochrome b5 reductase requires myristic acid for association with outer mitochondrial but not ER membranes. // *J Cell Biol*. 1996. V. 135. pp. 1501-1513.
199. Sadowski, P.G., Groen, A.J., Dupree, P. and Lilley, K.S. Sub-cellular localization of membrane proteins. // *Proteomics*. 2008. V. 8. pp. 3991-4011.
200. Hyun, D.H. and Lee, G.H. Cytochrome b5 reductase, a plasma membrane redox enzyme, protects neuronal cells against metabolic and oxidative stress through maintaining redox state and bioenergetics. // *Age (Dordr)*. 2015. V. 37. pp. 122.
201. de Cabo, R., Siendones, E., Minor, R. and Navas, P. CYB5R3: a key player in aerobic metabolism and aging? // *Aging (Albany NY)*. 2009. V. 2. pp. 63-68.

202. Zhao, Y., Hu, Q., Cheng, F., Su, N., Wang, A., Zou, Y., Hu, H., Chen, X., Zhou, H.M., Huang, X., Yang, K., Zhu, Q., Wang, X., Yi, J., Zhu, L., Qian, X., Chen, L., Tang, Y., Loscalzo, J. and Yang, Y. SoNar, a Highly Responsive NAD⁺/NADH Sensor, Allows High-Throughput Metabolic Screening of Anti-tumor Agents. // *Cell Metab.* 2015. V. 21. pp. 777-789.
203. Tao, R., Zhao, Y., Chu, H., Wang, A., Zhu, J., Chen, X., Zou, Y., Shi, M., Liu, R., Su, N., Du, J., Zhou, H.M., Zhu, L., Qian, X., Liu, H., Loscalzo, J. and Yang, Y. Genetically encoded fluorescent sensors reveal dynamic regulation of NADPH metabolism. // *Nat Methods.* 2017. V. 14. pp. 720-728.
204. Keyes, S.R. and Cinti, D.L. Biochemical properties of cytochrome b5-dependent microsomal fatty acid elongation and identification of products. // *J Biol Chem.* 1980. V. 255. pp. 11357-11364.
205. Reddy, V.V., Kupfer, D. and Caspi, E. Mechanism of C-5 double bond introduction in the biosynthesis of cholesterol by rat liver microsomes. // *J Biol Chem.* 1977. V. 252. pp. 2797-2801.
206. Chugunova, A., Navalayeu, T., Dontsova, O. and Sergiev, P. Mining for Small Translated ORFs. // *J Proteome Res.* 2018. V. 17. pp. 1-11.
207. Magny, E.G., Pueyo, J.I., Pearl, F.M., Cespedes, M.A., Niven, J.E., Bishop, S.A. and Couso, J.P. Conserved regulation of cardiac calcium uptake by peptides encoded in small open reading frames. // *Science.* 2013. V. 341. pp. 1116-1120.
208. Stein, C.S., Jadiya, P., Zhang, X., McLendon, J.M., Abouassaly, G.M., Witmer, N.H., Anderson, E.J., Elrod, J.W. and Boudreau, R.L. Mitoregulin: A lncRNA-Encoded Microprotein that Supports Mitochondrial Supercomplexes and Respiratory Efficiency. // *Cell Rep.* 2018. V. 23. pp. 3710-3720 e3718.
209. Makarewich, C.A., Baskin, K.K., Munir, A.Z., Bezprozvannaya, S., Sharma, G., Khemtong, C., Shah, A.M., McAnally, J.R., Malloy, C.R., Szveda, L.I., Bassel-Duby, R. and Olson, E.N. MOXI Is a Mitochondrial Micropeptide That Enhances Fatty Acid beta-Oxidation. // *Cell Rep.* 2018. V. 23. pp. 3701-3709.
210. Passon, P.G. and Hultquist, D.E. Soluble cytochrome b 5 reductase from human erythrocytes. // *Biochim Biophys Acta.* 1972. V. 275. pp. 62-73.
211. Neve, E.P., Nordling, A., Andersson, T.B., Hellman, U., Diczfalussy, U., Johansson, I. and Ingelman-Sundberg, M. Amidoxime reductase system containing cytochrome b5 type B (CYB5B) and MOSC2 is of importance for lipid synthesis in adipocyte mitochondria. // *J Biol Chem.* 2012. V. 287. pp. 6307-6317.
212. Plitzko, B., Havemeyer, A., Bork, B., Bittner, F., Mendel, R. and Clement, B. Defining the Role of the NADH-Cytochrome-b5 Reductase 3 in the Mitochondrial Amidoxime Reducing Component Enzyme System. // *Drug Metab Dispos.* 2016. V. 44. pp. 1617-1621.
213. Neve, E.P., Kofeler, H., Hendriks, D.F., Nordling, A., Gogvadze, V., Mkrtchian, S., Naslund, E. and Ingelman-Sundberg, M. Expression and Function of mARC: Roles in Lipogenesis and Metabolic Activation of Ximelagatran. // *PLoS One.* 2015. V. 10. pp. e0138487.
214. Vazquez-Memije, M.E., Cardenas-Mendez, M.J., Tolosa, A. and Hafidi, M.E. Respiratory chain complexes and membrane fatty acids composition in rat testis mitochondria throughout development and ageing. // *Exp Gerontol.* 2005. V. 40. pp. 482-490.
215. Ellis, C.E., Murphy, E.J., Mitchell, D.C., Golovko, M.Y., Scaglia, F., Barcelo-Coblijn, G.C. and Nussbaum, R.L. Mitochondrial lipid abnormality and electron

- transport chain impairment in mice lacking alpha-synuclein. // *Mol Cell Biol*. 2005. V. 25. pp. 10190-10201.
216. Hagopian, K., Weber, K.L., Hwee, D.T., Van Eenennaam, A.L., Lopez-Lluch, G., Villalba, J.M., Buron, I., Navas, P., German, J.B., Watkins, S.M., Chen, Y., Wei, A., McDonald, R.B. and Ramsey, J.J. Complex I-associated hydrogen peroxide production is decreased and electron transport chain enzyme activities are altered in n-3 enriched fat-1 mice. // *PLoS One*. 2010. V. 5. pp. e12696.
 217. Hensley, K., Kotake, Y., Sang, H., Pye, Q.N., Wallis, G.L., Kolker, L.M., Tabatabaie, T., Stewart, C.A., Konishi, Y., Nakae, D. and Floyd, R.A. Dietary choline restriction causes complex I dysfunction and increased H₂O₂ generation in liver mitochondria. // *Carcinogenesis*. 2000. V. 21. pp. 983-989.
 218. Striffler, G., Tuboly, E., Gorbe, A., Boros, M., Pecz, D. and Hartmann, P. Targeting Mitochondrial Dysfunction with L-Alpha Glycerylphosphorylcholine. // *PLoS One*. 2016. V. 11. pp. e0166682.
 219. Sharpley, M.S., Shannon, R.J., Draghi, F. and Hirst, J. Interactions between phospholipids and NADH:ubiquinone oxidoreductase (complex I) from bovine mitochondria. // *Biochemistry*. 2006. V. 45. pp. 241-248.
 220. Horibata, Y., Ando, H., Zhang, P., Vergnes, L., Aoyama, C., Itoh, M., Reue, K. and Sugimoto, H. StarD7 Protein Deficiency Adversely Affects the Phosphatidylcholine Composition, Respiratory Activity, and Cristae Structure of Mitochondria. // *J Biol Chem*. 2016. V. 291. pp. 24880-24891.
 221. Cogliati, S., Enriquez, J.A. and Scorrano, L. Mitochondrial Cristae: Where Beauty Meets Functionality. // *Trends Biochem Sci*. 2016. V. 41. pp. 261-273.
 222. Ikon, N. and Ryan, R.O. Cardiolipin and mitochondrial cristae organization. // *Biochim Biophys Acta Biomembr*. 2017. V. 1859. pp. 1156-1163.