

ПОРТРЕТ ТРЁХ БИЦИСТРОННЫХ мРНК МЛЕКОПИТАЮЩИХ, КОТОРЫЕ ПРОИСХОДЯТ ИЗ ГЕНОВ *ASNSD1*, *SLC35A4* И *MIEF1*

Мини-обзор

© 2025 Д.Е. Андреев^{1,2*}, И.Н. Шатский²

¹ Институт биоорганической химии имени академиков М.М. Шемякина и Ю.А. Овчинникова РАН,
117997 Москва, Россия; электронная почта: cycloheximide@yandex.ru

² Московский государственный университет имени М.В. Ломоносова,
НИИ физико-химической биологии имени А.Н. Белозерского, 119992 Москва, Россия

Поступила в редакцию 04.10.2024

После доработки 25.11.2024

Принята к публикации 07.12.2024

Последние достижения в области функциональной геномики позволили идентифицировать тысячи коротких открытых рамок считывания (sORF) в 5'-лидерах мРНК млекопитающих. Хотя большинство sORF, по всей видимости, не кодирует функциональные белки, установлено, что небольшое их количество эволюционирует как белок-кодирующие гены. В результате десятки таких sORF уже аннотированы как белок-кодирующие гены. мРНК, которые содержат и белок-кодирующую sORF, и аннотированную белок-кодирующую открытую рамку считывания, называют бицистронными. В этой работе рассматриваются три гена, *ASNSD1*, *SLC35A4* и *MIEF1*, из которых происходят соответствующие бицистронные мРНК. Обсуждается функциональная роль соответствующих белковых продуктов и регуляция трансляции этих бицистронных мРНК, а также то, как такая необычная организация генов могла возникнуть в процессе эволюции.

КЛЮЧЕВЫЕ СЛОВА: инициация трансляции, реинициация, проскальзывающее сканирование, двойное кодирование, бицистронная мРНК.

DOI: 10.31857/S0320972525010034 EDN: CQDSNX

ВВЕДЕНИЕ

Ранние работы Marylin Kozak позволили установить основные принципы регуляции трансляции в клетках эукариот, которые основаны на сканировании рибосомой 5'-лидеров мРНК [1]. В соответствии с моделью сканирования, т⁷G-кеп, расположенный на 5'-конце всех молекул мРНК, узнаётся фактором инициации eIF4F. Этот фактор привлекает 43S рибосомный комплекс на мРНК за счёт взаимодействия между eIF4F и eIF3. Затем 43S рибосомный комплекс начинает сканирование 5'-лидера мРНК до тех пор, пока не происходит узнавание стартового кодона (обычно AUG). После этого происходит диссоциация инициаторных факторов, присоединение 60S субчастицы, и начинается синтез полипептида (см. соответствующие обзоры [2–5]).

Последовательности 5'-лидеров мРНК оказывают значительное влияние на процесс сканирования, поскольку препятствия на пути сканирующей рибосомы могут снижать количество рибосомных комплексов, которые достигают стартового кодона. Интересно отметить, что примерно половина мРНК млекопитающих содержит как минимум один AUG-кодон в своих 5'-лидерах, что приводит к возникновению коротких открытых рамок считывания, uORF [6–10]. Метод рибосомного профилирования, который позволяет детектировать транслирующие рибосомы с одноклеотидным разрешением на полногеномном уровне [11], позволил обнаружить множественные трансляционные события в 5'-лидерах мРНК. Этот результат делает устаревшим ранее используемый термин «5'-нетранслируемая область» (5'-НТО).

Принятые сокращения: ISR – интегрированный стрессовый ответ; PFDL – префолдин-подобный модуль; sORF – короткая открытая рамка считывания; uORF – открытая рамка считывания в 5'-лидере мРНК.

* Адресат для корреспонденции.

Функциональная роль uORF интенсивно изучается. Множество работ посвящено изучению регуляторной функции uORF в регуляции трансляции продукта основной рамки считывания, однако мы не будем подробно останавливаться на этом аспекте, поскольку он подробно обсуждается в недавних обзорных статьях [4, 12–21]. Однако некоторые uORF кодируют функциональные белки. Бицистронными мРНК называют те мРНК, которые кодируют два белка в двух разделённых открытых рамках считывания. Такая организация позволяет регулировать соотношение двух белковых продуктов на трансляционном уровне. В данной работе обсуждаются три наиболее хорошо изученные бицистронные мРНК млекопитающих и приводятся данные об их трансляционной регуляции. Также в данной работе выдвинут ряд предположений о том, как изменение соотношения двух белков, закодированных в бицистронных мРНК, может влиять на клеточную физиологию.

ASDURF–ASNSD1

Генетический локус *ASNSD1* кодирует слабо охарактеризованный белок, который предположительно имеет глутамин-гидролизирующую активность. Кроме того, в этом локусе закодирован короткий белок-кодирующий ген *ASDURF* (ENSG00000286053), который может экспрессироваться или как отдельная открытая рамка считывания за счёт использования альтернативного сайта полиаденилирования, расположенного в экзоне 4 (внутри кодирующей последовательности *ASNSD1*), или как вышеупомянутая открытая рамка считывания. По данным CAGE (cap analysis gene expression), полученным в ходе выполнения проекта FANTOM5 [22] (анализ проводился в геномном браузере Zenbu [23]), эти два гена имеют общий единственный промотор. Таким образом, *ASNSD1* может транслироваться только с бицистронной мРНК, которая также содержит вышеупомянутый *ASDURF* (рис. 1, *a*, рис. 1П Приложения, *a*).

Белок *ASDURF*, также известный как *ASNSD1-SEP*, изначально был идентифицирован Slavoff et al. [28] при использовании пептидомного подхода, оптимизированного для идентификации коротких полипептидов. С использованием масс-спектрометрии разведения изотопов (*isotope dilution mass spectrometry* [29]) было показано, что внутриклеточная концентрация *ASDURF* в клетках K562 составляет 386 молекул на клетку [28]. Функция *ASDURF* оставалась неизвестной до 2020 года, когда Cloutier et al. показали, что он является субъединицей префолдин-подобного модуля (PFDL), входящего в состав комплекса PAQosome [30].

PAQosome, или Rvb1–Rvb2–Tah1–Pih1/prefoldin-like (R2TP/PFDL)-комплекс, является специфическим шапероном, отвечающим за сборку и созревание многих ключевых мультисубъединичных белковых комплексов в клетках млекопитающих [31]. У млекопитающих этот комплекс состоит из двух модулей, R2TP и PFDL. Cloutier et al. сначала применили метод биотинилирования от сближения (*proximity-dependent biotinylation*, BioID), чтобы идентифицировать белки, взаимодействующие с двумя субъединицами PAQosome, PIH1D1 и UXT. Оказалось, что эндогенный *ASDURF* был обогащён в обоих экспериментах. Затем авторы применили аффинную очистку с помощью FLAG-тэга для двух других субъединиц, RPAP3 и URI1, и провели детекцию белковых партнёров с помощью масс-спектрометрии. В этом случае *ASDURF* также был обнаружен среди обогащённых белковых партнёров. Более того, очистка с помощью *ASDURF*-FLAG позволила идентифицировать другие субъединицы PAQosome, что подтверждает тот факт, что *ASDURF* является интегральным компонентом PFDL-модуля. Сборка *in vitro* комплекса PFDL из очищенных компонентов показала, что *ASDURF* является шестой субъединицей гетерогексамерного комплекса PFDL, который также содержит субъединицы UXT1, PFDN6, PDRG1, URI1 и PFDN2. Взаимодействие *ASDURF* с другими субъединицами PFDN было независимо подтверждено Hofman et al. [32].

PAQosome вовлечена в регуляцию активности многих мультисубъединичных и рибонуклеопротеиновых комплексов, таких как рибонуклеопротеины семейства L7Ae (snoRNP (малые ядрышковые рибонуклеопротеины) box C/D и H/ACA, snRNP (малые ядерные рибонуклеопротеины) U4, теломеразная и сelenобелковая мРНП), snRNP U5, киназные комплексы Phosphatidylinositol 3-kinase-related (PIKKs) (ATM, ATR, DNA-PKcs, mTOR, SMG-1 и TRRAP) и РНК-полимеразы I, II и III [33–38]. Если *ASDURF* необходим для функционирования PAQosome, то его отсутствие может повлиять на жизнеспособность клетки. Недавно, используя скринирование с помощью CRISPR–Cas9, Hofman et al. [32] показали, что инактивация *ASDURF* снижает жизнеспособность клеток MYC-направляемой медуллобластомы. Важно отметить, что возобновление экспрессии *ASDURF* возвращало жизнеспособность клеток. *In vivo* нокаут *ASDURF* увеличивал общую выживаемость мышей с ксенографтами D425 MYC-направляемых клеток медуллобластомы. Интересно отметить, что нокаут *ASDURF* *in vitro* снижал включение бромоуридина в S-фазе клеточного цикла, что говорит о возможной активности *ASDURF*/PFDL в контроле клеточного цикла.

Значительно меньше известно про *ASNSD1*, который закодирован во втором цистроне.

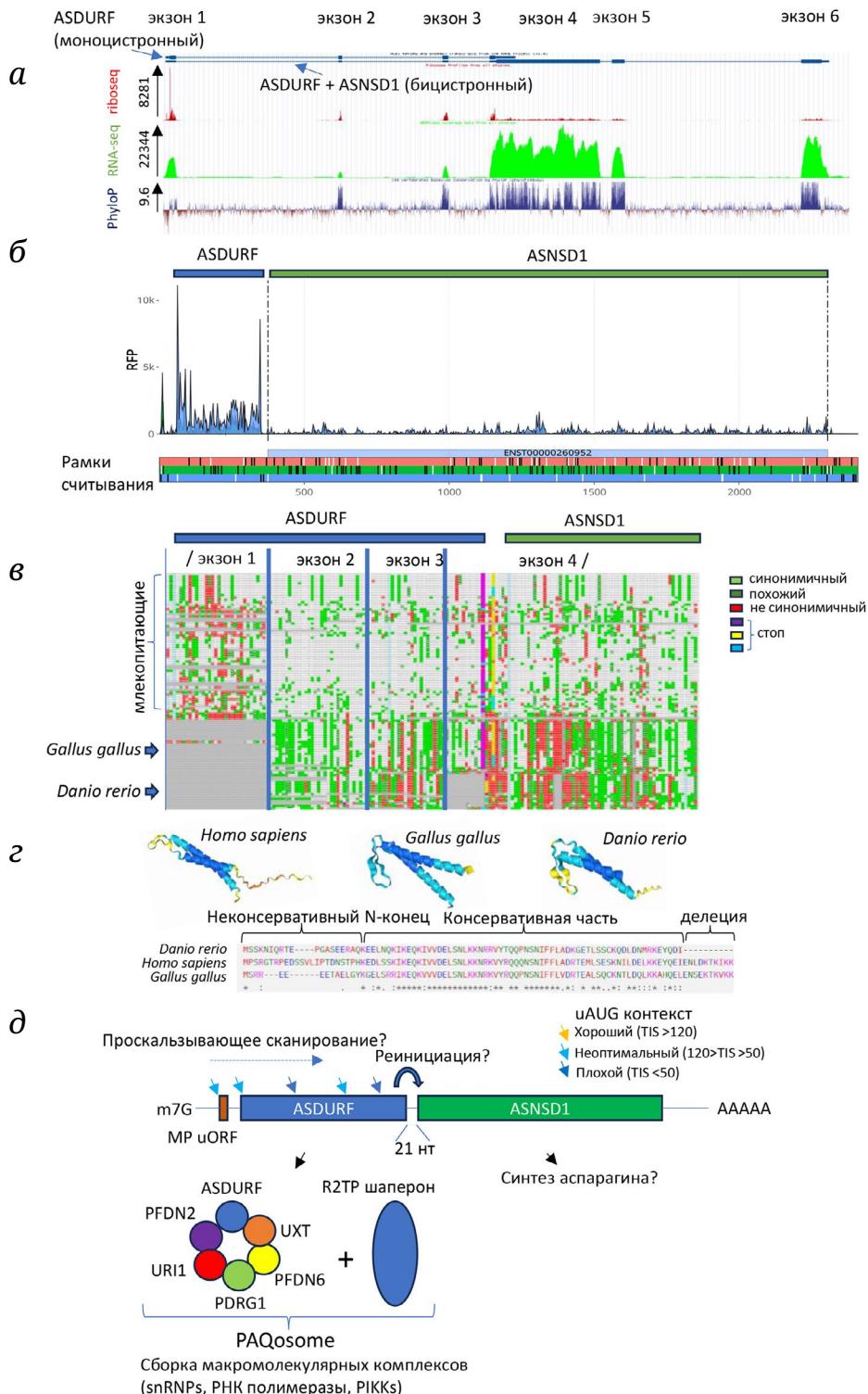


Рис. 1. ASDURF-ASNSD1. *а* – Данные Riboseq и RNAseq для гена *ASNSD1*, полученные с помощью геномного браузера GWIPS-Viz [24]. Эволюционная консервативность нуклеотидов у 100 позвоночных показана как сигнал PhyloP (phyloP100way) [25]. *б* – Данные Riboseq из транскриптомного браузера Ribocrypt (ribocrypt.org, подготавливается к публикации), рибосомные футприанты окрашены в цвета, соответствующие трём открытым рамкам считывания, показаны внизу диаграммы. *в* – Множественное выравнивание последовательности, проведённое с помощью инструмента CodAlignView («CodAlignView: a tool for exploring signatures of protein-coding evolution in an alignment», I. Jungreis, M. Lin, M. Kellis, *in preparation*), использован набор выравниваний hg30_100. *г* – Структуры и выравнивание белков ASDURF *Homo sapiens*, *Gallus gallus* и *Danio rerio*. Структурные модели получены с помощью AlphaFold3 [26]. Выравнивание последовательностей проводилось с помощью ClustalW [27]. *д* – Схематическое представление бицистронной мРНК ASDURF-ASNSD1. AUG-кодоны, расположенные перед вторым цистроном, отмечены цветными стрелками. Функциональная роль соответствующих белковых продуктов показана на схеме внизу

ASNSD1 имеет *N*-концевой домен class-II glutamine amidotransferase и *C*-концевой домен asparagine synthase B, принадлежащий к суперсемейству adenine nucleotide alpha hydrolase (AANH). Этот белок, скорее всего, вовлечён в синтез аспарагина.

Анализ данных рибосомного профилирования человеческих клеточных линий показывает наличие трёх транслируемых участков: короткой uORF, кодирующую двухаминокислотный MP-пептид (расположена на расстоянии 46 нуклеотидов от первого цистрона), ASDURF (первый цистрон) и ASNSD1 (второй цистрон) (рис. 1, б). Триплетная периодичность сигнала рибосомного профилирования соответствует трансляции этих трёх открытых рамок считывания. ASDURF транслируется на порядок более эффективно по сравнению с ASNSD1, что, по всей видимости, говорит о том, что первый цистрон подавляет трансляцию второго цистрона (рис. 1, б). Однако интересно отметить, что в экспериментах с репортерными конструкциями мутация стартового кодона ASDURF активировала трансляцию люциферазы, слитой со стартовым кодоном второго цистрона, всего в 2 раза [39], что меньше ожидаемого эффекта. По всей видимости, в случае, когда стартовый AUG-кодон ASDURF мутирован, рибосомы начинают инициировать на тех AUG-кодонах внутри открытой рамки считывания ASDURF, которые в норме не узнаются.

Филогенетический анализ показал, что последовательность ASDURF, которая соответствует первому экзону, консервативна у млекопитающих, в то же время последовательности 2 и 3 экзонов содержат множественные синонимичные замены также в геномах позвоночных (рис. 1, в). Анализ ASNSD1-генов *Gallus gallus* и *Danio rerio* показал, что эти животные имеют альтернативные первые экзоны, содержащие AUG-кодон в рамке считывания ASDURF. Из-за этого *N*-концевая часть ASDURF является наиболее вариабельной. Однако, в соответствии с моделированием с помощью AlphaFold 3, эти белки тем не менее могут иметь похожую структурную организацию, соответствующую структуре b-префолдина (рис. 1, г).

Каким образом может регулироваться трансляция бицистронной мРНК ASDURF-ASNSD1? Чтобы исследовать эффективность инициации на AUG-кодонах, мы исследовали данные о количественном анализе трансляционной эффективности AUG-кодонов со всеми возможными комбинациями нуклеотидного контекста в положении от -6 до +5 [40]. Эффективность контекстов стартовых кодонов находится в диапазоне от 12 (самые неэффективные) до 150 (самые эффективные). Интересно отметить, что все AUG-кодоны, расположенные перед ASNSD1, включая стартовый кодон ASDURF, имеют средние или слабые кон-

тексты (рис. 1, д). Основываясь на этих данных, мы предполагаем, что трансляция ASNSD1 может регулироваться или в условиях изменения эффективности узнавания стартовых кодонов, или при регуляции скорости элонгации трансляции. Усиление инициации на неоптимальных AUG-кодонах или снижение скорости рибосом, транслирующих первый цистрон ASDURF, должно, в свою очередь, значительно снижать трансляцию второго цистрона ASNSD1.

SLC35A4URF–SLC35A4

Генетический локус SLC35A4 кодирует белок, для которого изначально была предсказана активность в трансмембранным транспорте пиrimидиновых нуклеотидов, связанных с сахарами. Кроме того, в этом локусе находится короткий белок-кодирующий ген ENSG00000293600, который расположен перед SLC35A4. По аналогии с описанным ранее случаем мы предлагаем называть этот ген SLC35A4URF. SLC35A4URF может транслироваться либо как отдельная открытая рамка считывания, если используется альтернативный сайт полиденилирования между SLC35A4 и SLC35A4URF, или в составе бицистронной мРНК. На основании анализа данных CAGE, полученных в ходе выполнения проекта FANTOM5 [22] (анализ проводился в геномном браузере Zenbu [23]), можно заключить, что SLC35A4 может транслироваться только с бицистронной мРНК (рис. 2, а, рис. П1, б).

В 2015 г. мы предсказали, что вышерасположенная короткая открытая рамка считывания в гене SLC35A4 может кодировать функциональный белок [41]. Функциональная роль SLC35A4URF, или SLC35A4-MP, изучалась в двух недавних работах, где было показано, что этот белок связан с митохондриями [42, 43]. Yang et al. [43] показали, что белок SLC35A4URF локализован в митохондриях. Используя ко-иммунопреципитацию и последующую масс-спектрометрию, а также анализ с помощью Вестерн-блоттинга, авторы показали, что изучаемый белок взаимодействует с белками внешней мембраны митохондрий, в частности, с VDAC1 и VDAC3. Кроме того, эксперименты по совместному окрашиванию показали, что SLC35A4URF колокализуется с маркером внешней митохондриальной мембраны TOMM20. Rocha et al. [42] создали кроличьи поликлональные антитела против SLC35A4URF и также показали, что этот белок локализован в митохондриях. Однако на основании экспериментов с протеиназой K был сделан вывод о том, что SLC35A4URF является белком не внешней, а внутренней мембраны митохондрий. Эти результаты согласуются с наличием однопроходного трансмембранныго домена,

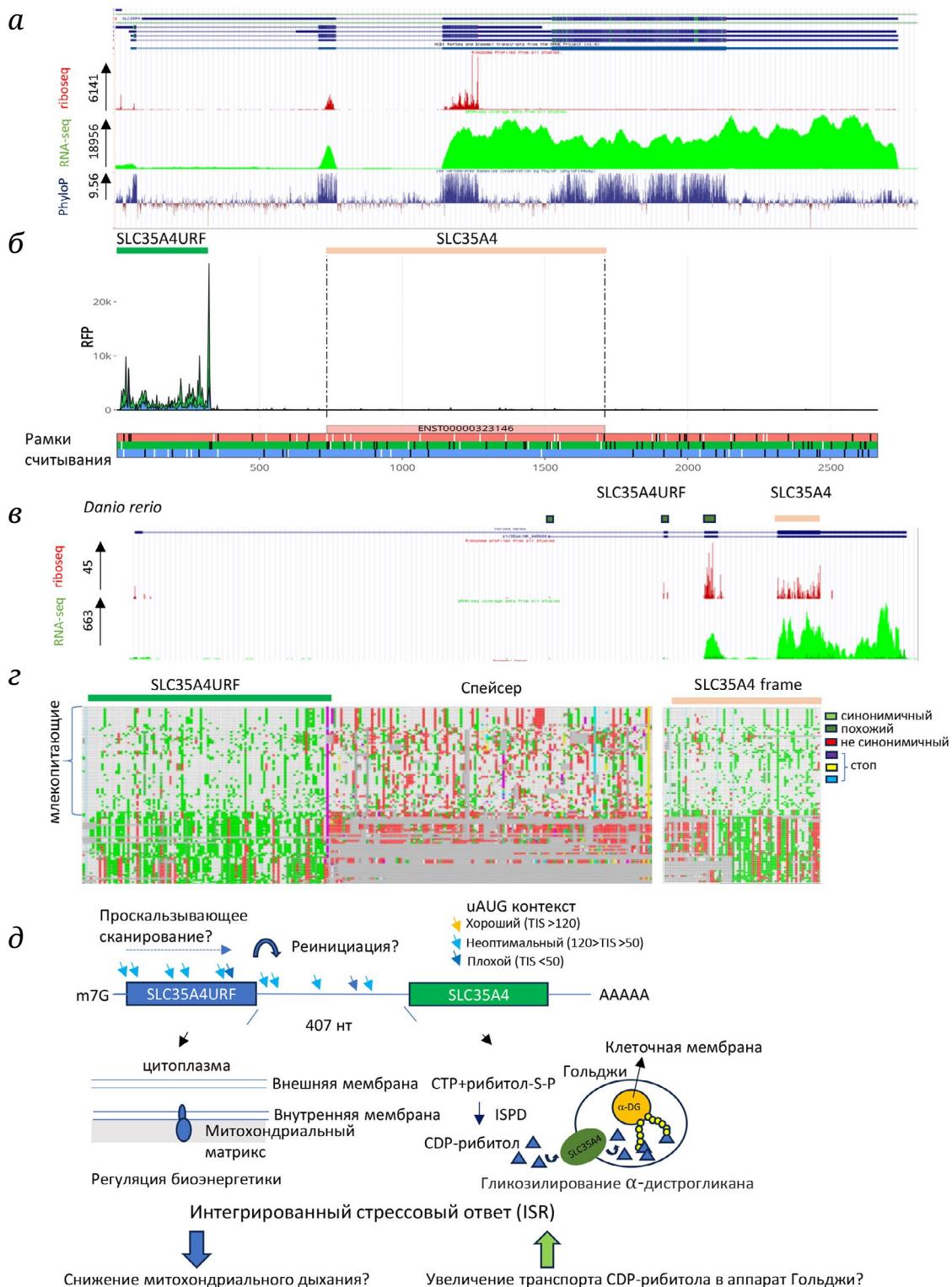


Рис. 2. *SLC35A4URF-SLC35A4*. а – Данные Riboseq и RNAseq для гена *SLC35A4*, полученные с помощью геномного браузера GWIPS-Viz [24]. Эволюционная консервативность нуклеотидов у 100 позвоночных показана как сигнал PhyloP (phyloP100way) [25]. б – Данные Riboseq из транскриптомного браузера Ribocrypt (ribocrypt.org, подготовливается к публикации), рибосомные футпринты окрашены в цвета, соответствующие трем открытым рамкам считывания, показаны внизу диаграммы. в – Данные Riboseq и RNAseq для гена *SLC35A4* *Danio rerio*, полученные с помощью геномного браузера GWIPS-Viz [24]. г – Множественное выравнивание последовательности, проведённое с помощью CodAlignView («CodAlignView: a tool for exploring signatures of protein-coding evolution in an alignment», I. Jungreis, M. Lin, M. Kellis, in preparation), использован набор выравниваний hg30_100 E. д – Схематическое представление бицистронной мРНК *SLC35A4URF-SLC35A4*. AUG-кодоны, расположенные перед вторым цистроном, отмечены цветными стрелками. Функциональная роль соответствующих белковых продуктов показана на схеме внизу

расположенного между аминокислотами 62 и 84 в SLC35A4URF [42].

Было показано, что функциональная роль SLC35A4URF связана с дыхательной цепью митохондрий. Используя редактирование генома, Rocha et al. создали соответствующую нокаутную линию HEK293T и показали, что потеря SLC35A4URF приводит к значительному подавлению базального и максимального дыхания, а также дыхания при утечке протонов. Интересно отметить, что трёхкратное повышение экспрессии SLC35A4URF в клетках MCF7 нарушает клеточный рост, что сопровождается потерей мембранныго потенциала митохондрий. Кроме того, происходит активация транскрипции генов, вовлечённых в путь, связанные с активностью митохондрий, такие как ответ на гипоксию и негативная регуляция мембранныго потенциала митохондрий. Эти результаты говорят о том факте, что количество SLC35A4URF должно строго регулироваться для поддержания оптимального клеточного роста.

Функциональная роль продукта второго цистрона, SLC35A4, связана с регуляцией гликозилирования в аппарате Гольджи. SLC35A4 и его паралог SLC35A1 могут быть вовлечены в транспорт CDP-рибитола (CDP-ribitol) из цитоплазмы в аппарат Гольджи [44]. CDP-рибитол, который синтезируется в цитоплазме с помощью фермента ISPD, используется для сборки сложного гликана на гликобелке α-дистрогликане (α -dystroglycan) при помощи ферментов аппарата Гольджи FKTN и FKRP [45]. Эта модификация с помощью остатков рибитола необходима для правильного взаимодействия α -дистрогликана с белками внеклеточного матрикса, такими как ламинин. Локализация экзогенного белка SLC35A4 в аппарате Гольджи была показана Rocha et al. [42].

Данные Riboseq для бицистронной мРНК SLC35A4URF-SLC35A4 показывают, что первый цистрон транслируется эффективно, при этом трансляция второго цистрона практически не детектируется (рис. 2, а и б). После открытой рамки считывания первого цистрона, имеющей длину 312 нт, следует спейсер длиной 407 нт, который содержит несколько AUG-кодонов (всего перед вторым цистроном находятся 11 AUG). Филогенетический анализ показал, что спейсерный регион является наименее консервативным в геномах позвоночных (рис. 2, г). Интересно отметить, что данные рибосомного профилирования для *Danio rerio*, у которой межцистронный спейсерный участок гораздо короче по сравнению с человеком, показывают трансляцию обоих цистронов (рис. 2, в).

Регуляция трансляции бицистронной мРНК SLC35A4URF-SLC35A4 при интегрированном стрессовом ответе (ISR) была обнаружена в двух неза-

висимых исследованиях [41, 46]. При развитии ISR трансляция SLC35A4 значительно увеличивается, что сопровождается подавлением трансляции первого цистрона. Andreev et al. [41] использовали репортерные конструкции, в которых второй цистрон был заменён геном люциферазы светлячка. В нормальных условиях трансляция копированной бицистронной мРНК давала очень низкие, но детектируемые значения люциферазной активности. Эти значения были на два порядка ниже значений для репортерной мРНК с коротким 5'-лидером. После обработки арсенитом, который вызывает фосфорилирование фактора инициации eIF2, трансляция второго цистрона увеличивалась, в то время как трансляция контрольной репортерной мРНК подавлялась примерно в 10 раз. Таким образом, последовательность мРНК перед вторым цистроном SLC35A4 достаточна для приятия устойчивости к стрессу. Анализ последовательности показывает, что все 11 AUG-кодонов, расположенных перед вторым цистроном, имеют неоптимальные нуклеотидные контексты (рис. 2, д). Вероятно, трансляция второго цистрона возможна благодаря проскальзывающему сканированию.

Подводя итог, можно предположить, что трансляционный контроль SLC35A4URF-SLC35A4 может связывать митохондриальное дыхание и межклеточный матрикс. При ISR трансляция первого цистрона снижается, что должно приводить к снижению активности дыхательной цепи митохондрий. В то же время стресс-зависимая активация трансляции второго цистрона, SLC35A4, может стимулировать созревание α -дистрогликана. Стоит отметить, что трансляция SLC35A1, который также вовлечён в транспорт CDP-рибитола, значительно подавляется при стрессе, вызванном обработкой арсенитом [41]. Интересно, что при оверэкспрессии SLC35A4 в клетках, в которых отсутствует SLC35A1, гликозилирование α -дистрогликана восстанавливается, однако молекулярный вес его гликозилированной формы значительно снижается [44]. По всей видимости, SLC35A4 способен менять паттерн гликозилирования своего субстрата. Можно предположить, что в условиях стресса паттерн гликозилирования α -дистрогликана может меняться из-за изменения соотношения белков SLC35A4 и SLC35A1.

MIURF–MIEF1

На основании филогенетического анализа вышеупомянутой открытой рамки считывания было выдвинуто предположение, что генетический локус MIEF1 может кодировать бицистронную мРНК [41]. Продукт трансляции второго

цистрона, MIEF1, играет роль в регуляции деления и слияния митохондрий. Кроме того, в первом цистроне (ENSG00000285025) закодирован небольшой белок MIURF. Этот белок локализован в митохондриях и, возможно, также способен регулировать активность MIEF1. На основании анализа данных CAGE, полученных в ходе выполнения проекта FANTOM5 [22] (анализ проводился в геномном браузере Zenbu [23]), было выяснено, что *MIURF* и *MIEF1* имеют общий промотор (рис. 1П, в).

MIEF1 (также известен как MID51) и *MIEF2* (также известен как MID49) являются двумя схожими белками, открытыми в 2011 году. Эти белки вовлечены в регуляцию динамики митохондрий, поскольку при их подавлении или оверэкспрессии меняется морфология митохондрий – происходит их деление или слияние [47]. Белки семейства *MIEF* являются белками внешней митохондриальной мембраны. Эти белки содержат однопроходный трансмембранный домен на своём *N*-конце, при этом большая часть белка ориентирована в цитозоль. Белки семейства *MIEF* являются рецепторами, связывающими как белки, активирующие слияние митохондрий (*MFN1* и *MFN2*), так и белок *DRP1*, активирующий деление митохондрий [48]. Было показано, что белки семейства *MIEF* могут образовывать олигомеры, а также взаимодействовать со своими цитоплазматическими доменами, что может приводить к сближению и последующему слиянию соседних митохондрий [48].

Белок *MIURF*, закодированный в первом цистроне, имеет две различные активности: он регулирует трансляцию в митохондриальном матриксе, а также регулирует активность *MIEF1*, который локализован на цитоплазматической стороне внешней мембраны митохондрий. *MIURF* был впервые обнаружен в составе интермедиата сборки большой рибосомной субчастицы митохондрий. Незрелые рибосомные субъединицы митохондрий были выделены из клеточной линии, полученной на основе HEK293S, и проанализированы с помощью криоэлектронной микроскопии. Авторы обнаружили дополнительную электронную плотность рядом с белком *uL14m*. Используя алгоритмы предсказания пространственной структуры на основе электронной плотности, а также проведя анализ белков методом масс-спектрометрии, исследователи идентифицировали три белка: *MALSU*, *mt-ACP* и *MIURF* (на момент публикации этот белок назывался *L0R8F8*). Было предположено, что этот белковый модуль предотвращает преждевременную ассоциацию рибосомных субчастиц [49]. Rathore et al. [50] получили дополнительное подтверждение того, что *MIURF* вовлечен в регуляцию митохондриальной трансляции. Используя разные подходы, авторы показали, что *MIURF* взаимодействует с митохон-

дриальной рибосомой, и что *MIURF* стимулирует митохондриальную трансляцию. Важно отметить: в этих двух работах было показано, что митохондриальные рибосомы взаимодействуют с эндогенным белком *MIURF*, что является серьёзным аргументом, подтверждающим вовлечённость этого белка в процесс трансляции в митохондриях. До конца не ясно, является ли *MIURF* только фактором биогенеза рибосом, или он также способен напрямую участвовать в процессе трансляции.

Другая активность *MIURF* была обнаружена в 2020 году. Chen et al. [51] применили полногеномный скрининг CRISPR, чтобы идентифицировать неканонические открытые рамки считывания, которые влияют на клеточный рост, и в результате *MIURF* попал в список наиболее значимых обнаружений. Далее, авторы показали, что экзогенный *MIURF*, слитый с *Teg*, способен физически взаимодействовать с продуктом второго цистрона *MIEF1*. Важно отметить, что и тэгированный *MIEF1* позволил выделить *MIURF*. Интересно отметить, что оверэкспрессия *MIURF* приводит к делению митохондрий, а нокаут *MIURF* приводит к повышенному слиянию митохондрий. Данный эффект может быть устранён при экзогенной экспрессии *MIURF*. Эти результаты показывают, что *MIURF* может влиять на активность *MIEF1* за счёт регуляции взаимодействий с белками, регулирующими деление и слияние митохондрий.

Для того чтобы понять, как может регулироваться трансляция бицистронной мРНК *MIURF-MIEF1*, мы проанализировали данные рибосомного профилирования (рис. 3, а и б) и провели филогенетический анализ (рис. 3, в). Данные Riboseq говорят о трансляции обоих цистронов в бицистронной мРНК. У человека перед вторым цистроном находятся 5 AUG-кодоны. С третьего AUG-кодона в неоптимальном нуклеотидном контексте начинается трансляция *MIURF*. Анализ геномов позвоночных говорит о том, что открытая рамка считывания *MIURF* консервативна у позвоночных, в то же время спейсер между двумя цистронами менее консервативен. Поскольку все пять AUG-кодонов имеют неоптимальные нуклеотидные контексты, можно предположить, что трансляция второго цистрона может регулироваться за счёт проскальзывающего сканирования (рис. 3, г).

Интересно отметить, что трансляция *MIURF* и *MIEF1* дифференциально регулируется в процессе развития ISR, индуцированного обработкой клеток арсенитом [41]. В условиях стресса значительно снижается соотношение трансляции первого цистрона *MIURF* и второго цистрона *MIEF1*. Этот результат говорит о возможности того, что трансляционный контроль *MIURF* и *MIEF1*

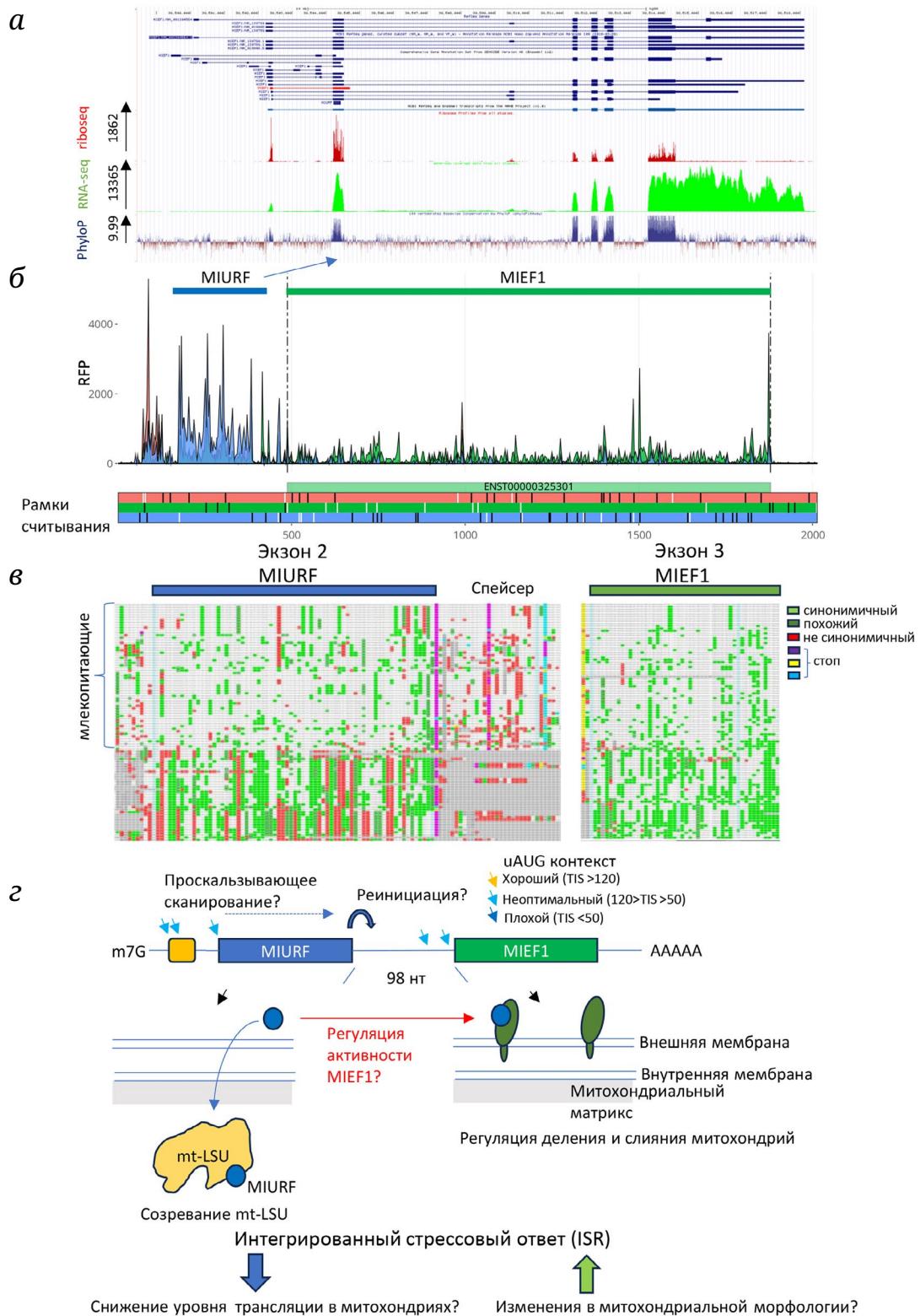


Рис. 3. MIURF–MIEF1. *а* – Данные Riboseq и RNAseq для гена *MIEF1*, полученные с помощью геномного браузера GWIPS-Viz [24]. Эволюционная консервативность нуклеотидов у 100 позвоночных показана как сигнал PhyloP (phyloP100way) [25]. *б* – Данные Riboseq из транскриптомного браузера Riboscrypt (riboscrypt.org, подготавливается к публикации), рибосомные футпринты окрашены в цвета, соответствующие трём открытым рамкам считывания, показаны внизу диаграммы. *в* – Множественное выравнивание последовательности, проведённое с помощью CodAlignView («CodAlignView: a tool for exploring signatures of protein-coding evolution in an alignment», I. Jungreis, M. Lin, M. Kellis, *in preparation*), использован набор выравниваний hg30_100. *г* – Схематическое представление бицистронной мРНК *MIURF–MIEF1*. AUG-кодоны, расположенные перед вторым цистроном, отмечены цветными стрелками. Функциональная роль соответствующих белковых продуктов показана на схеме внизу

может оказывать влияние на различные митохондриальные процессы. При нормальных условиях *MIURF* продуцируется больше, чем *MIEF1*. Протеомный анализ, проведённый в работе Delcourt et al. [52], показал, что отношение белков *MIURF/MIEF* при нормальных условиях роста для клеток HEK293, HeLa и клеток кишечника человека составляет 2,71, 5,73 и 2,62 соответственно. Поскольку *MIURF* синтезируется в цитоплазме, он способен в первую очередь связаться с *MIEF1*, который локализован на поверхности внешней мембраны митохондрий. Избыток молекул *MIURF*, по-видимому, затем транспортируется в митохондриальный матрикс для участия в созревании митохондриальных рибосом. При стрессе отношение количества *MIURF* и *MIEF1* снижается. Если все имеющиеся молекулы *MIURF* окажутся связаны в комплексе с *MIEF1*, биогенез рибосом в митохондриях будет заблокирован. Эта гипотеза требует экспериментального подтверждения.

ЗАКЛЮЧЕНИЕ

В данной работе проведен анализ данных о трансляционной регуляции и функциональной роли трёх бицистронных мРНК, которые эволюционно консервативны у позвоночных. Во всех трёх случаях два цистрона играют различные функциональные роли. В связи с этим возникает вопрос о том, каким образом данные бицистронные мРНК могли возникнуть в ходе эволюции. Стоит отметить, что все три гена имеют паралоги, соответствующие их вторым цистронам (*ASNS*; *SLC35A1*, *A2*, *A3*, *A5* и *MIEF2*), при этом у таких паралогов нет соответствующих вышерасположенных открытых рамок считывания. На основании этого можно предположить, что *ASNSD1*, *SLC35A4* и *MIEF1* могли образоваться в процессе дупликации генов, при котором произошло объединение с первыми цистронами, которые до этого функционировали индивидуально. Такие слияния привели к образованию регуляторных систем, которые контролируются на трансляционном уровне, что по какой-то причине оказалось положительное влияние на жизнеспособность организма и сохранилось в процессе эволюции. Интересно отметить, что межцистронные последовательности, которые должны оказывать значительное влияние на регуляцию трансляции, в целом менее консервативны по сравнению с цистронами. Вероятно, эти последовательности эволюционировали по-разному у разных животных, чтобы обеспечивать точный трансляционный контроль. В качестве примера можно привести ген *SLC35A4*, в котором межцистронный спейсер у человека имеет значительно большую длину и

больше AUG-кодонов (5 и 1 AUG) по сравнению с *Danio rerio*.

Необычная бицистронная организация мРНК предполагает наличие специфических механизмов трансляционного контроля. Изменения в активности компонентов трансляционного аппарата способны привести к быстрому изменению трансляции белков, закодированных в первых и вторых цистронах. Действительно, для двух из трёх описанных бицистронных мРНК было показано, что их трансляция дифференциально регулируется в условиях ISR, когда концентрация тройного комплекса eIF2–Met-tRNAi–GTP (tRNAi – инициаторная тРНК) становится лимитирующей. Вероятно, что регуляция трансляции этих мРНК не ограничивается ISR. Мы обнаружили, что во всех трёх случаях AUG-кодоны, расположенные перед вторым цистроном, имеют неоптимальные нуклеотидные контексты. Исходя из этого, можно предположить, что регуляция трансляции может осуществляться в условиях, при которых меняется эффективность выбора стартовых кодонов. Ключевые факторы инициации, влияющие на эффективность выбора стартовых кодонов – это eIF1 и eIF5 [53–59]. Повышенный уровень eIF1 снижает инициацию на неоптимальных стартовых кодонах, eIF5 имеет противоположный эффект. Мы предполагаем, что эти факторы инициации способны регулировать трансляцию бицистронных мРНК, обсуждаемых в рамках данной работы. eIF4G2, также известный как DAP5 – ещё один белок, способный регулировать трансляцию бицистронных мРНК. В нескольких недавних работах показано, что eIF4G2 может регулировать трансляцию мРНК, имеющих короткие открытые рамки считывания в своих 5'-лидерах. Эта регуляция осуществляется либо за счёт проскальзывающего сканирования, либо за счёт реинициации [60–65]. В частности, было показано, что eIF4G2 регулирует трансляцию мРНК *POLG*, которая кодирует два различных белка, *POLG* и *POLGARF* [66]. Очень интересно выяснить, какие факторы отвечают за дифференциальную трансляцию двух цистронов в случае *ASNSD1*, *SLC35A4* и *MIEF1*.

Количество аннотированных бицистронных мРНК продолжает увеличиваться. Недавно 25 коротких открытых рамок считывания в 5'-лидерах были аннотированы как новые белок-кодирующие гены [67]. Изучение функциональной роли этих новых белок-кодирующих генов позволит понять принципы эволюционирования геномов и установить новые случаи связанных различными клеточными процессов за счёт трансляционного контроля новых бицистронных мРНК. Для изучения этих случаев будут полезны приведённые в данной работе примеры: *ASNSD1*, *SLC35A4* и *MIEF1*.

Вклад авторов. Д.Е. Андреев – концепция исследования, анализ данных, написание рукописи; Д.Е. Андреев, И.Н. Шатский – редактирование рукописи.

Финансирование. Эта работа поддержана Российским научным фондом (грант 20-14-00121).

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Соблюдение этических норм. Настоящая статья не содержит описания каких-либо проведённых авторами исследований с участием людей или животных в качестве объектов.

СПИСОК ЛИТЕРАТУРЫ

1. Kozak, M. (1980) Evaluation of the “scanning model” for initiation of protein synthesis in eucaryotes, *Cell*, **22**, 7-8, [https://doi.org/10.1016/0092-8674\(80\)90148-8](https://doi.org/10.1016/0092-8674(80)90148-8).
2. Hinnebusch, A. G. (2011) Molecular mechanism of scanning and start codon selection in eukaryotes, *Microbiol. Mol. Biol. Rev.*, **75**, 434-467, first page of table of contents, <https://doi.org/10.1128/MMBR.00008-11>.
3. Hinnebusch, A. G. (2014) The scanning mechanism of eukaryotic translation initiation, *Annu. Rev. Biochem.*, **83**, 779-812, <https://doi.org/10.1146/annurev-biochem-060713-035802>.
4. Hinnebusch, A. G., Ivanov, I. P., and Sonenberg, N. (2016) Translational control by 5'-untranslated regions of eukaryotic mRNAs, *Science*, **352**, 1413-1416, <https://doi.org/10.1126/science.aad9868>.
5. Shirokikh, N. E., and Preiss, T. (2018) Translation initiation by cap-dependent ribosome recruitment: Recent insights and open questions, *Wiley Interdiscip. Rev. RNA*, **9**, e1473, <https://doi.org/10.1002/wrna.1473>.
6. Rogozin, I. B., Kochetov, A. V., Kondrashov, F. A., Koonin, E. V., and Milanesi, L. (2001) Presence of ATG triplets in 5' untranslated regions of eukaryotic cDNAs correlates with a ‘weak’ context of the start codon, *Bioinformatics*, **17**, 890-900, <https://doi.org/10.1093/bioinformatics/17.10.890>.
7. 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. (2000) Statistical analysis of the 5' untranslated region of human mRNA using “Oligo-Capped” cDNA libraries, *Genomics*, **64**, 286-297, <https://doi.org/10.1006/geno.2000.6076>.
8. Pesole, G., Gissi, C., Grillo, G., Licciulli, F., Liuni, S., and Saccone, C. (2000) Analysis of oligonucleotide AUG start codon context in eukaryotic mRNAs, *Gene*, **261**, 85-91, [https://doi.org/10.1016/s0378-1119\(00\)00471-6](https://doi.org/10.1016/s0378-1119(00)00471-6).
9. Davuluri, R. V., Suzuki, Y., Sugano, S., and Zhang, M. Q. (2000) CART classification of human 5' UTR sequences, *Genome Res.*, **10**, 1807-1816, <https://doi.org/10.1101/gr.gr-1460r>.
10. Andreev, D. E., Loughran, G., Fedorova, A. D., Mikhaylova, M. S., Shatsky, I. N., and Baranov, P. V. (2022) Non-AUG translation initiation in mammals, *Genome Biol.*, **23**, 111, <https://doi.org/10.1186/s13059-022-02674-2>.
11. Ingolia, N. T., Ghaemmaghami, S., Newman, J. R., and Weissman, J. S. (2009) Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling, *Science*, **324**, 218-223, <https://doi.org/10.1126/science.1168978>.
12. Dever, T. E., Ivanov, I. P., and Hinnebusch, A. G. (2023) Translational regulation by uORFs and start codon stringency, *Genes Dev.*, **37**, 474-489, <https://doi.org/10.1101/gad.350752.123>.
13. Tidu, A., and Martin, F. (2024) The interplay between cis- and trans-acting factors drives selective mRNA translation initiation in eukaryotes, *Biochimie*, **217**, 20-30, <https://doi.org/10.1016/j.biochi.2023.09.017>.
14. Renz, P. F., Valdivia-Francia, F., and Sendoel, A. (2020) Some like it translated: small ORFs in the 5'UTR, *Exp. Cell Res.*, **396**, 112229, <https://doi.org/10.1016/j.yexcr.2020.112229>.
15. Silva, J., Fernandes, R., and Romao, L. (2019) Translational regulation by upstream open reading frames and human diseases, *Adv. Exp. Med. Biol.*, **1157**, 99-116, https://doi.org/10.1007/978-3-030-19966-1_5.
16. Chen, H. H., and Tarn, W. Y. (2019) uORF-mediated translational control: recently elucidated mechanisms and implications in cancer, *RNA Biol.*, **16**, 1327-1338, <https://doi.org/10.1080/15476286.2019.1632634>.
17. Zhang, H., Wang, Y., and Lu, J. (2019) Function and evolution of upstream ORFs in eukaryotes, *Trends Biochem. Sci.*, **44**, 782-794, <https://doi.org/10.1016/j.tibs.2019.03.002>.
18. Cabrera-Quio, L. E., Herberg, S., and Pauli, A. (2016) Decoding sORF translation - from small proteins to gene regulation, *RNA biology*, **13**, 1051-1059, <https://doi.org/10.1080/15476286.2016.1218589>.
19. Young, S. K., and Wek, R. C. (2016) Upstream open reading frames differentially regulate gene-specific translation in the integrated stress response, *J. Biol. Chem.*, **291**, 16927-16935, <https://doi.org/10.1074/jbc.R116.733899>.
20. Wethmar, K. (2014) The regulatory potential of upstream open reading frames in eukaryotic gene expression, *Wiley Interdiscip. Rev. RNA*, **5**, 765-778, <https://doi.org/10.1002/wrna.1245>.
21. Somers, J., Poyry, T., and Willis, A. E. (2013) A perspective on mammalian upstream open reading frame function, *Int. J. Biochem. Cell Biol.*, **45**, 1690-1700, <https://doi.org/10.1016/j.biocel.2013.04.020>.
22. Lizio, M., Harshbarger, J., Shimoji, H., Severin, J., Kasukawa, T., Sahin, S., Abugessaisa, I., Fukuda, S., Hori, F., Ishikawa-Kato, S., Mungall, C. J., Arner, E., Baillie, J. K., Bertin, N., Bono, H., de Hoon, M., Diehl, A. D., Dimont, E.,

- Freeman, T. C., Fujieda, K., et al. (2015) Gateways to the FANTOM5 promoter level mammalian expression atlas, *Genome Biol.*, **16**, 22, <https://doi.org/10.1186/s13059-014-0560-6>.
23. Severin, J., Lizio, M., Harshbarger, J., Kawaji, H., Daub, C. O., Hayashizaki, Y., Consortium, F., Bertin, N., and Forrest, A. R. (2014) Interactive visualization and analysis of large-scale sequencing datasets using ZENBU, *Nat. Biotechnol.*, **32**, 217-219, <https://doi.org/10.1038/nbt.2840>.
 24. 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. (2013) Peptidomic discovery of short open reading frame-encoded peptides in human cells, *Nat. Chem. Biol.*, **9**, 59-64, <https://doi.org/10.1038/ncchembio.1120>.
 25. Keshishian, H., Addona, T., Burgess, M., Kuhn, E., and Carr, S. A. (2007) Quantitative, multiplexed assays for low abundance proteins in plasma by targeted mass spectrometry and stable isotope dilution, *Mol. Cell. Proteomics*, **6**, 2212-2229, <https://doi.org/10.1074/mcp.M700354-MCP200>.
 26. Cloutier, P., Poitras, C., Faubert, D., Bouchard, A., Blanchette, M., Gauthier, M. S., and Coulombe, B. (2020) Upstream ORF-encoded ASDURF is a novel prefoldin-like subunit of the PAQosome, *J. Proteome Res.*, **19**, 18-27, <https://doi.org/10.1021/acs.jproteome.9b00599>.
 27. Houry, W. A., Bertrand, E., and Coulombe, B. (2018) The PAQosome, an R2TP-based chaperone for quaternary structure formation, *Trends Biochem. Sci.*, **43**, 4-9, <https://doi.org/10.1016/j.tibs.2017.11.001>.
 28. Hofman, D. A., Ruiz-Orera, J., Yannuzzi, I., Murugesan, R., Brown, A., Clauser, K. R., Condurat, A. L., van Dinter, J. T., Engels, S. A. G., Goodale, A., van der Lugt, J., Abid, T., Wang, L., Zhou, K. N., Vogelzang, J., Ligon, K. L., Phoenix, T. N., Roth, J. A., Root, D. E., Hubner, N., et al. (2024) Translation of non-canonical open reading frames as a cancer cell survival mechanism in childhood medulloblastoma, *Mol. Cell*, **84**, 261-276, e218, <https://doi.org/10.1016/j.molcel.2023.12.003>.
 29. Michel, A. M., Fox, G., Kiran, A. M., De Bo, C., O'Connor, P. B., Heaphy, S. M., Mullan, J. P., Donohue, C. A., Higgins, D. G., and Baranov, P. V. (2014) GWIPS-viz: development of a ribo-seq genome browser, *Nucleic Acids Res.*, **42**, D859-864, <https://doi.org/10.1093/nar/gkt1035>.
 30. Pollard, K. S., Hubisz, M. J., Rosenbloom, K. R., and Siepel, A. (2010) Detection of nonneutral substitution rates on mammalian phylogenies, *Genome Res.*, **20**, 110-121, <https://doi.org/10.1101/gr.097857.109>.
 31. Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A. J., Bambrick, J., Bodenstein, S. W., Evans, D. A., Hung, C. C., O'Neill, M., Reiman, D., Tunyasuvunakool, K., Wu, Z., Zemgulyte, A., Arvaniti, E., Beattie, C., et al. (2024) Accurate structure prediction of biomolecular interactions with AlphaFold 3, *Nature*, **630**, 493-500, <https://doi.org/10.1038/s41586-024-07487-w>.
 32. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, *Nucleic Acids Res.*, **22**, 4673-4680, <https://doi.org/10.1093/nar/22.22.4673>.
 33. Boulon, S., Marmier-Gourrier, N., Pradet-Balade, B., Wurth, L., Verheggen, C., Jady, B. E., Rothe, B., Pescia, C., Robert, M. C., Kiss, T., Bardoni, B., Krol, A., Branlant, C., Allmang, C., Bertrand, E., and Charpentier, B. (2008) The Hsp90 chaperone controls the biogenesis of L7Ae RNPs through conserved machinery, *J. Cell Biol.*, **180**, 579-595, <https://doi.org/10.1083/jcb.200708110>.
 34. Cloutier, P., Poitras, C., Durand, M., Hekmat, O., Fiola-Masson, E., Bouchard, A., Faubert, D., Chabot, B., and Coulombe, B. (2017) R2TP/Prefoldin-like component RUVBL1/RUVBL2 directly interacts with ZNHIT2 to regulate assembly of U5 small nuclear ribonucleoprotein, *Nat. Commun.*, **8**, 15615, <https://doi.org/10.1038/ncomms15615>.
 35. Malinova, A., Cvackova, Z., Mateju, D., Horejsi, Z., Abeza, C., Vandermoere, F., Bertrand, E., Stanek, D., and Verheggen, C. (2017) Assembly of the U5 snRNP component PRPF8 is controlled by the HSP90/R2TP chaperones, *J. Cell Biol.*, **216**, 1579-1596, <https://doi.org/10.1083/jcb.201701165>.
 36. Horejsi, Z., Takai, H., Adelman, C. A., Collis, S. J., Flynn, H., Maslen, S., Skehel, J. M., de Lange, T., and Boulton, S. J. (2010) CK2 phospho-dependent binding of R2TP complex to TEL2 is essential for mTOR and SMG1 stability, *Mol. Cell*, **39**, 839-850, <https://doi.org/10.1016/j.molcel.2010.08.037>.
 37. Kim, S. G., Hoffman, G. R., Poulogiannis, G., Buel, G. R., Jang, Y. J., Lee, K. W., Kim, B. Y., Erikson, R. L., Cantley, L. C., Choo, A. Y., and Blenis, J. (2013) Metabolic stress controls mTORC1 lysosomal localization and dimerization by regulating the TTT-RUVBL1/2 complex, *Mol. Cell*, **49**, 172-185, <https://doi.org/10.1016/j.molcel.2012.10.003>.
 38. Boulon, S., Pradet-Balade, B., Verheggen, C., Molle, D., Boireau, S., Georgieva, M., Azzag, K., Robert, M. C., Ahmad, Y., Neel, H., Lamond, A. I., and Bertrand, E. (2010) HSP90 and its R2TP/Prefoldin-like cochaperone are involved in the cytoplasmic assembly of RNA polymerase II, *Mol. Cell*, **39**, 912-924, <https://doi.org/10.1016/j.molcel.2010.08.023>.
 39. Nelde, A., Flotto, L., Jurgens, L., Szymik, L., Hubert, E., Bauer, J., Schliemann, C., Kessler, T., Lenz, G., Rammensee, H. G., Walz, J. S., and Wethmar, K. (2022) Upstream open reading frames regulate translation of cancer-associated transcripts and encode HLA-presented immunogenic tumor antigens, *Cell. Mol. Life Sci.*, **79**, 171, <https://doi.org/10.1007/s00018-022-04145-0>.

40. Noderer, W. L., Flockhart, R. J., Bhaduri, A., Diaz de Arce, A. J., Zhang, J., Khavari, P. A., and Wang, C. L. (2014) Quantitative analysis of mammalian translation initiation sites by FACS-seq, *Mol. Systems Biol.*, **10**, 748, <https://doi.org/10.15252/msb.20145136>.
41. 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. (2015) Translation of 5' leaders is pervasive in genes resistant to eIF2 repression, *eLife*, **4**, e03971, <https://doi.org/10.7554/eLife.03971>.
42. Rocha, A. L., Pai, V., Perkins, G., Chang, T., Ma, J., De Souza, E. V., Chu, Q., Vaughan, J. M., Diedrich, J. K., Ellisman, M. H., and Saghatelian, A. (2024) An inner mitochondrial membrane microprotein from the SLC35A4 upstream ORF regulates cellular metabolism, *J. Mol. Biol.*, **436**, 168559, <https://doi.org/10.1016/j.jmb.2024.168559>.
43. Yang, H., Li, Q., Stroup, E. K., Wang, S., and Ji, Z. (2024) Widespread stable noncanonical peptides identified by integrated analyses of ribosome profiling and ORF features, *Nat. Commun.*, **15**, 1932, <https://doi.org/10.1038/s41467-024-46240-9>.
44. Ury, B., Potelle, S., Caligiore, F., Whorton, M. R., and Bommer, G. T. (2021) The promiscuous binding pocket of SLC35A1 ensures redundant transport of CDP-ribitol to the Golgi, *J. Biol. Chem.*, **296**, 100789, <https://doi.org/10.1016/j.jbc.2021.100789>.
45. Gerin, I., Ury, B., Breloy, I., Bouchet-Seraphin, C., Bolsee, J., Halbout, M., Graff, J., Vertommen, D., Muccioli, G. G., Seta, N., Cuisset, J. M., Dabaj, I., Quijano-Roy, S., Grahn, A., Van Schaftingen, E., and Bommer, G. T. (2016) ISPD produces CDP-ribitol used by FKTN and FKRP to transfer ribitol phosphate onto alpha-dystroglycan, *Nat. Commun.*, **7**, 11534, <https://doi.org/10.1038/ncomms11534>.
46. Sidrauski, C., McGeachy, A. M., Ingolia, N. T., and Walter, P. (2015) The small molecule ISRib reverses the effects of eIF2alpha phosphorylation on translation and stress granule assembly, *eLife*, **4**, <https://doi.org/10.7554/eLife.05033>.
47. Palmer, C. S., Osellame, L. D., Laine, D., Koutsopoulos, O. S., Frazier, A. E., and Ryan, M. T. (2011) MiD49 and MiD51, new components of the mitochondrial fission machinery, *EMBO Rep.*, **12**, 565-573, <https://doi.org/10.1038/embor.2011.54>.
48. Yu, R., Liu, T., Jin, S. B., Ankarcrona, M., Lendahl, U., Nister, M., and Zhao, J. (2021) MIEF1/2 orchestrate mitochondrial dynamics through direct engagement with both the fission and fusion machineries, *BMC Biol.*, **19**, 229, <https://doi.org/10.1186/s12915-021-01161-7>.
49. Brown, A., Rathore, S., Kimanius, D., Aibara, S., Bai, X. C., Rorbach, J., Amunts, A., and Ramakrishnan, V. (2017) Structures of the human mitochondrial ribosome in native states of assembly, *Nat. Struct. Mol. Biol.*, **24**, 866-869, <https://doi.org/10.1038/nsmb.3464>.
50. Rathore, A., Chu, Q., Tan, D., Martinez, T. F., Donaldson, C. J., Diedrich, J. K., Yates, J. R., 3rd, and Saghatelian, A. (2018) MIEF1 microprotein regulates mitochondrial translation, *Biochemistry*, **57**, 5564-5575, <https://doi.org/10.1021/acs.biochem.8b00726>.
51. Chen, J., Brunner, A. D., Cogan, J. Z., Nunez, J. K., Fields, A. P., Adamson, B., Itzhak, D. N., Li, J. Y., Mann, M., Leonetti, M. D., and Weissman, J. S. (2020) Pervasive functional translation of noncanonical human open reading frames, *Science*, **367**, 1140-1146, <https://doi.org/10.1126/science.aay0262>.
52. Delcourt, V., Brunelle, M., Roy, A. V., Jacques, J. F., Salzet, M., Fournier, I., and Roucou, X. (2018) The protein coded by a short open reading frame, not by the annotated coding sequence, is the main gene product of the dual-coding gene MIEF1, *Mol. Cell. Proteomics*, **17**, 2402-2411, <https://doi.org/10.1074/mcp.RA118.000593>.
53. Grossley, R., Alvarado, C., Ivanov, I. P., Nicholson, O. B., Puglisi, J. D., Dever, T. E., and Lapointe, C. P. (2024) eIF1 and eIF5 dynamically control translation start site fidelity, *bioRxiv*, <https://doi.org/10.1101/2024.07.10.602410>.
54. Ly, J., Xiang, K., Su, K. C., Sissoko, G. B., Bartel, D. P., and Cheeseman, I. M. (2024) Nuclear release of eIF1 globally increases stringency of start-codon selection to preserve mitotic arrest physiology, *bioRxiv*, <https://doi.org/10.1101/2024.04.06.588385>.
55. Loughran, G., Sachs, M. S., Atkins, J. F., and Ivanov, I. P. (2012) Stringency of start codon selection modulates autoregulation of translation initiation factor eIF5, *Nucleic Acids Res.*, **40**, 2898-2906, <https://doi.org/10.1093/nar/gkr1192>.
56. Ivanov, I. P., Loughran, G., Sachs, M. S., and Atkins, J. F. (2010) Initiation context modulates autoregulation of eukaryotic translation initiation factor 1 (eIF1), *Proc. Natl. Acad. Sci. USA*, **107**, 18056-18060, <https://doi.org/10.1073/pnas.1009269107>.
57. Llacer, J. L., Hussain, T., Saini, A. K., Nanda, J. S., Kaur, S., Gordiyenko, Y., Kumar, R., Hinnebusch, A. G., Lorsch, J. R., and Ramakrishnan, V. (2018) Translational initiation factor eIF5 replaces eIF1 on the 40S ribosomal subunit to promote start-codon recognition, *eLife*, **7**, <https://doi.org/10.7554/eLife.39273>.
58. Martin-Marcos, P., Cheung, Y. N., and Hinnebusch, A. G. (2011) Functional elements in initiation factors 1, 1A, and 2beta discriminate against poor AUG context and non-AUG start codons, *Mol. Cell. Biol.*, **31**, 4814-4831, <https://doi.org/10.1128/MCB.05819-11>.

59. Nanda, J. S., Cheung, Y. N., Takacs, J. E., Martin-Marcos, P., Saini, A. K., Hinnebusch, A. G., and Lorsch, J. R. (2009) eIF1 controls multiple steps in start codon recognition during eukaryotic translation initiation, *J. Mol. Biol.*, **394**, 268-285, <https://doi.org/10.1016/j.jmb.2009.09.017>.
60. Hacisuleyman, E., Hale, C. R., Noble, N., Luo, J. D., Fak, J. J., Saito, M., Chen, J., Weissman, J. S., and Darnell, R. B. (2024) Neuronal activity rapidly reprograms dendritic translation via eIF4G2:uORF binding, *Nat. Neurosci.*, **27**, 822-835, <https://doi.org/10.1038/s41593-024-01615-5>.
61. Shestakova, E. D., Tumbinsky, R. S., Andreev, D. E., Rozov, F. N., Shatsky, I. N., and Terenin, I. M. (2023) The roles of eIF4G2 in leaky scanning and reinitiation on the human dual-coding POLG mRNA, *Int. J. Mol. Sci.*, **24**, 17149, <https://doi.org/10.3390/ijms242417149>.
62. She, R., Luo, J., and Weissman, J. S. (2023) Translational fidelity screens in mammalian cells reveal eIF3 and eIF4G2 as regulators of start codon selectivity, *Nucleic Acids Res.*, **51**, 6355-6369, <https://doi.org/10.1093/nar/gkad329>.
63. Weber, R., Kleemann, L., Hirschberg, I., Chung, M. Y., Valkov, E., and Igredja, C. (2022) DAP5 enables main ORF translation on mRNAs with structured and uORF-containing 5' leaders, *Nat. Commun.*, **13**, 7510, <https://doi.org/10.1038/s41467-022-35019-5>.
64. David, M., Olender, T., Mizrahi, O., Weingarten-Gabbay, S., Friedlander, G., Meril, S., Goldberg, N., Savidor, A., Levin, Y., Salomon, V., Stern-Ginossar, N., Bialik, S., and Kimchi, A. (2022) DAP5 drives translation of specific mRNA targets with upstream ORFs in human embryonic stem cells, *RNA*, **28**, 1325-1336, <https://doi.org/10.1261/rna.079194.122>.
65. Smirnova, V. V., Shestakova, E. D., Nogina, D. S., Mishchenko, P. A., Prikazchikova, T. A., Zatsepina, T. S., Kulakovskiy, I. V., Shatsky, I. N., and Terenin, I. M. (2022) Ribosomal leaky scanning through a translated uORF requires eIF4G2, *Nucleic Acids Res.*, **50**, 1111-1127, <https://doi.org/10.1093/nar/gkab1286>.
66. Loughran, G., Zhdanov, A. V., Mikhaylova, M. S., Rozov, F. N., Datskevich, P. N., Kovalchuk, S. I., Serebryakova, M. V., Kiniry, S. J., Michel, A. M., O'Connor, P. B. F., Papkovsky, D. B., Atkins, J. F., Baranov, P. V., Shatsky, I. N., and Andreev, D. E. (2020) Unusually efficient CUG initiation of an overlapping reading frame in POLG mRNA yields novel protein POLGARF, *Proc. Natl. Acad. Sci. USA*, **117**, 24936-24946, <https://doi.org/10.1073/pnas.2001433117>.
67. Mudge, J. M., Ruiz-Orera, J., Prensner, J. R., Brunet, M. A., Calvet, F., Jungreis, I., Gonzalez, J. M., Magrane, M., Martinez, T. F., Schulz, J. F., Yang, Y. T., Alba, M. M., Aspden, J. L., Baranov, P. V., Bazzini, A. A., Bruford, E., Martin, M. J., Calviello, L., Carvunis, A. R., Chen, J., et al. (2022) Standardized annotation of translated open reading frames, *Nat. Biotechnol.*, **40**, 994-999, <https://doi.org/10.1038/s41587-022-01369-0>.

A PORTRAIT OF THREE MAMMALIAN BICISTRONIC mRNA TRANSCRIPTS, DERIVED FROM THE GENES *ASNSD1*, *SLC35A4*, and *MIEF1*

Mini-Review

D. E. Andreev^{1,2*} and I. N. Shatsky²

¹*Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences,
117997 Moscow, Russia; e-mail: cycloheximide@yandex.ru*

²*Belozersky Institute of Physico-Chemical Biology, Lomonosov Moscow State University,
119992 Moscow, Russia*

Recent advances in functional genomics have allowed identification of thousands of translated short open reading frames (sORFs) in the 5' leaders of mammalian mRNA transcripts. While most sORFs are unlikely to encode functional proteins, a small number have been shown to have evolved as protein-coding genes. As a result, dozens of these sORFs have already been annotated as protein-coding ORFs. mRNAs that contain both a protein-coding sORF and an annotated coding sequence (CDS) are referred to as bicistronic transcripts. In this study, we focus on three genes – *ASNSD1*, *SLC35A4*, and *MIEF1* – which give rise to bicistronic mRNAs. We discuss recent findings regarding functional investigation of the corresponding polypeptide products, as well as how their translation is regulated, and how this unusual genetic arrangement may have evolved.

Keywords: translation initiation, reinitiation, leaky scanning, dual coding, bicistronic mRNA