

ОБЗОРЫ

©Коллектив авторов

ПОДХОДЫ К ПОИСКУ КРИТИЧЕСКИ ВАЖНЫХ ГЕНОВ В РАКОВОМ ГЕНОМЕ

М.А. Пятницкий^{1,2}, Д.С. Карпов^{1,3}, С.А. Мошковский^{1,4}*

¹НИИ биомедицинской химии им. В.Н. Ореховича,

119121, Москва, ул. Погодинская, д. 10, стр. 8; эл. почта: mpyat@bioinformatics.ru

²Национальный исследовательский университет “Высшая школа экономики”, 101000, Москва

³Институт молекулярной биологии им. В.А. Энгельгардта РАН, 119991, Москва

⁴Российский национальный исследовательский университет им. Н.И. Пирогова Минздрава РФ, 117997, Москва

Понятие критически важных генов, чья утрата функции приводит к гибели клетки, является одной из основополагающих концепций генетики и имеет важное значение для фундаментальных и прикладных исследований. Особенно перспективно изучение данной области применительно к онкологии, поскольку поиск генетических уязвимостей раковых клеток позволяет выявить новые потенциальные мишени для противоопухолевой терапии. Современные возможности биотехнологий позволяют проводить крупномасштабные проекты по секвенированию соматических мутаций в опухолях, а также напрямую вмешиваться в работу генетического аппарата раковых клеток. Благодаря этому накоплен значительный массив знаний о встречающихся генетических вариантах и их фенотипических проявлениях. Не вызывает сомнения, что в недалеком будущем эти знания найдут применение в клинической практике. В настоящем обзоре описаны основные экспериментальные и вычислительные подходы к поиску критически важных генов, концентрируясь на применении этих методов в области молекулярной онкологии.

Ключевые слова: раковый геном; критически важные гены; CRISPR; РНК-интерференция; синтетические летали

DOI: 10.18097/PBMC20186404303

ВВЕДЕНИЕ

Ранние единичные [1, 2], а затем и полногеномные [3, 4] исследования по генетическому нокауту на модельных организмах показали, что в геноме существуют гены, являющиеся незаменимыми для выживания организма. Утрата функций таких критически важных генов (КВГ) приводит либо к гибели (в случае одноклеточного организма), либо к стерильности и невозможности стать полноценной взрослой особью (в случае многоклеточного организма). Особенно значимым представляется изучение генов, чье функционирование необходимо для пролиферации или выживания опухолевых клеток. Такие гены (или ген) в дальнейшем мы будем называть критически важными опухолевыми генами (КВОГ), не делая различий между геном и кодируемыми им белками.

Наиболее очевидный подход к терапии злокачественных заболеваний состоит в том, чтобы находить и специфически ингибировать только КВОГ, не затрагивая при этом КВГ здоровых клеток. Например, клетки опухоли в силу своей повышенной пролиферации обычно более чувствительны к лекарственным препаратам, повреждающим ДНК или нарушающим митотическую активность. Однако клиническая практика показывает, что реализация этого подхода сталкивается со значительными трудностями – многие типы опухолей характеризуются быстрым развитием резистентности к подобным цитотоксичным препаратам [5].

Для некоторых типов опухолей характерен так называемый феномен “онкогенной зависимости”,

состоящий в том, что выживание раковых клеток зависит от активности определённого КВОГ или ассоциированного с ним сигнального пути [6]. Такой КВОГ, часто связанный с регуляцией клеточного роста и дифференцировки, и активированный вследствие различных мутаций (изменение копийности, транслокация, и т.д.), называется онкогеном.

Первые экспериментальные свидетельства того, что пролиферация и выживаемость раковых клеток могут зависеть от наличия активированных онкогенов, были получены в работе [7], где было показано, что нокаут мутировавшего *KRAS* методом гомологичной рекомбинации приводит к замедлению роста опухоли *in vitro* и *in vivo*. Скрининговые эксперименты по поиску КВОГ на культурах клеток были начаты в работе [8], где с помощью технологии РНК-интерференции изучали 21 ген. К настоящему моменту благодаря ряду рассмотренных ниже крупномасштабных исследований [9-12] достигнут существенный прогресс в данном направлении.

Феномен онкогенной зависимости успешно транслирован в клиническую практику. Классическим примером является зависимость клеток хронического миелоидного лейкоза от повышенной активности киназы ABL, которая обусловлена транслокацией её гена под промотор гена *BCR*. Разработка селективного ингибитора BCR-ABL, иматиниба, произвела революцию в эффективности лечения хронического миелоидного лейкоза [13]. Другим примером может служить чувствительность рака молочной железы с гиперэкспрессией рецептора ERBB2 (HER2) к ингибиторам HER2, включая лапатиниб и трастузумаб [14].

* - адресат для переписки

Помимо изучения простых примеров онкогенной зависимости вида “опухоль с биомаркером *X* уязвима к его ингибированию”, возможно сформулировать и задачу следующего уровня – “существуют ли опухоли с биомаркером *X*, уязвимые к ингибированию гена *Y*?”. Такой подход называется поиском синтетических леталей, и имеет существенное достоинство в виде расширения диапазона потенциальных генов-мишеней. Ценой за это является необходимость экспериментальной проверки множества сочетаний биомаркер-ингибитор, количество которых комбинаторно возрастает. Благодаря широкомасштабному скринингу выявлен ряд синтетических леталей, которые потенциально могут быть использованы в клинике: клеточные линии с мутантным геном *ARID1A*, участвующим в перестройке хроматина, чувствительны к ингибированию его гомолога *ARID1B* [15], а клетки рака молочной железы с амплифицированным *MYC* чувствительны к ингибированию компонентов сплайсосомы [16].

В настоящем обзоре предпринята попытка описать различные подходы к поиску КВОГ.

1. ЭКСПЕРИМЕНТАЛЬНЫЕ МЕТОДЫ

Основной экспериментальной стратегией определения КВОГ является поочередное “выключение” гена в какой-либо клеточной линии. В скрининговых исследованиях наиболее часто ген инактивируют либо с помощью РНК-интерференции, либо с помощью системы CRISPR/Cas9. Далее оба эти метода будут рассмотрены подробнее.

1.1. РНК-интерференция

Одна из наиболее распространенных технологий по выключению гена основана на использовании феномена РНК-интерференции – стандартного механизма контроля экспрессии на посттранскрипционном уровне. В клетку вводят

синтетические двухцепочечные РНК (так называемые малые интерферирующие РНК, миРНК), которые иницируют молекулярные каскады, приводящие к деградации целевой мРНК. Для проведения скринингового эксперимента используют две возможные стратегии (рис. 1). Согласно первой стратегии, в каждой лунке планшета клетки подвергаются воздействию только одной гено-специфичной миРНК, а потенциальные КВОГ определяются по измерению сигнала люминисценции с определенной лунки. В другом варианте гено-специфичные реагенты вносятся в культуру клеток в виде пула векторов, содержащих последовательность малых шпилечных РНК (мшРНК) ко всем инактивируемым генам. В клетке мшРНК процессируются в активную миРНК, а каждый вектор помечен уникальным штрих-кодом. После культивирования трансцифированные клетки секвенируют, и по относительному недостатку прочтений со штрих-кодом определяют потенциальные КВОГ [17].

В ходе пилотного исследования технология миРНК была протестирована для изучения 21 потенциального КВОГ на клетках HeLa [8]. Через три года было опубликовано первое крупномасштабное скрининговое исследование [18], в котором благодаря использованию 5305 миРНК было найдено 37 КВОГ, необходимых для деления клеток.

Благодаря успеху отдельных исследований, сообществом были предприняты усилия для выявления КВОГ среди сотен клеточных линий. Наиболее масштабным экспериментом такого рода стал проект Achilles [12]. Первоначально были проанализированы 216 раковых линий, причём каждый из протестированных 11000 генов являлся мишенью для нескольких мшРНК, в среднем пяти мшРНК на один ген. Для учета результатов от воздействия на ген нескольких мшРНК был разработан алгоритм ATARiS [19], который позволяет охарактеризовать эффект супрессии

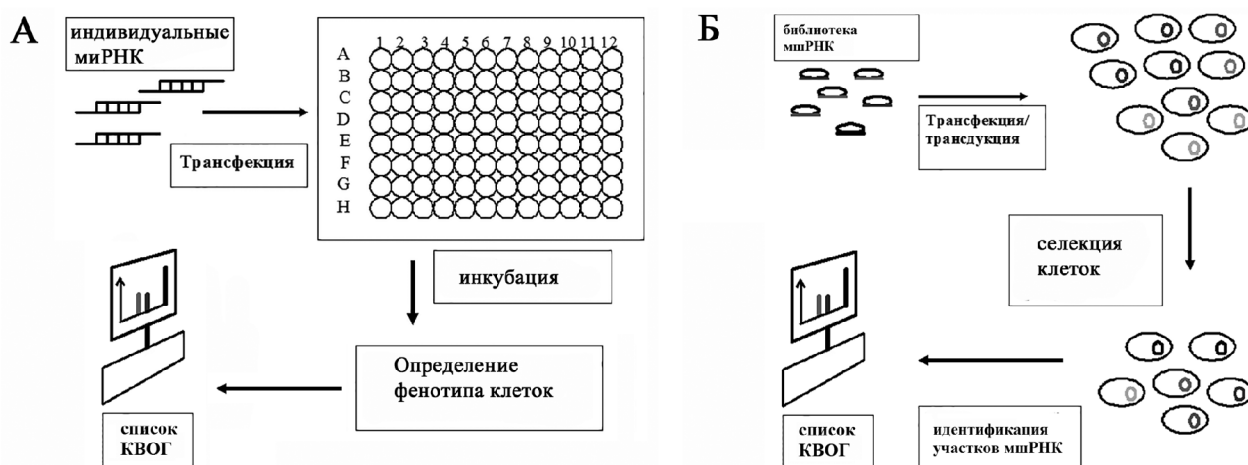


Рисунок 1. Идентификация КВОГ с помощью РНК-интерференции. (А) миРНК скрининг. Индивидуальные миРНК трансфицируются в клетки в отдельных лунках культурального планшета. После инкубации, определяется фенотип клеток или детектируется сигнал репортерного белка, что и используется в установлении КВОГ. (Б) мшРНК-скрининг. Библиотека мшРНК вводится в клетки. После селекции на антибиотике и селекции в присутствии лекарственного препарата из выживших клеток выделяется геномная ДНК и секвенируются фрагменты, кодирующие мшРНК. По относительному количеству мшРНК определяют КВОГ.

каждого гена одной суммарной величиной. Недавно проект Achilles был расширен, итоговое количество протестированных клеточных линий достигло 501 [20]. Также была разработана модификация алгоритма ATARiS, получившая название DEMETER. В новой версии для оценки эффекта супрессии учитывается последовательность нуклеотидов в таргетном гене и последовательность мшРНК. С помощью алгоритма DEMETER было найдено 769 генов, которые с высокой степенью вероятности являются критически важными, по крайней мере, в одной клеточной линии, причём 20% из кодируемых ими белков могут быть использованы как мишени для лекарственного воздействия. Только 10 белков являются такими мишенями в 58% из 501 клеточной линии.

Технология РНК-интерференции является первым методом, который был использован для широкомасштабного определения КВОГ. Однако она не свободна от недостатков, основным из которых является связывание миРНК с нецелевыми генами. Также по своей природе РНК-интерференция приводит к временному подавлению функции гена (генетический нокдаун), не инактивируя его полностью (генетический нокаут). Поэтому помимо проблемы ложно-положительных находок, характерной для многих “омных” методик, для технологии РНК-интерференции существует опасность и ложно-негативных срабатываний, когда для проявления фенотипического эффекта требуется полное выключение функции гена. Таким образом для полноценной инактивации гена при помощи РНК-интерференции требуется воздействие нескольких миРНК. Ещё одним источником ошибок является предпочтительное использование стратегии, основанной только на миРНК или мшРНК – было показано, что некоторые гены

определяются как КВОГ исключительно благодаря использованию той или иной платформы [21].

1.2. CRISPR/Cas9

Бактериальная система адаптивного ответа против чужеродной ДНК CRISPR/Cas9 была впервые использована в редактировании генома человека пять лет назад и уже стала революционной технологией в геномной инженерии [22, 23]. В отличие от РНК-интерференции, система CRISPR/Cas9 позволяет вводить наследуемые изменения в заданные участки генома и модифицировать нетранскрибируемые области.

В настоящее время наиболее широко используется система CRISPR/SpyCas9 ПА подтипа из *S. pyogenes*. Как правило, она состоит из двух компонентов: неспецифичной РНК-зависимой ДНК-эндонуклеазы SpyCas9 и химерной направляющей РНК (хнРНК), рисунок 2А. Эндонуклеаза SpyCas9 образует комплекс с хнРНК, что переводит ее в конформацию, компетентную для поиска мишени в геноме. Направляющая РНК включает в себя участок длиной 20 нуклеотидов – спейсер, который комплементарен целевому участку генома – протоспейсеру. Спейсер направляющей РНК и определяет специфичность действия SpyCas9. Если образуется стабильный гибрид спейсера и протоспейсера РНК:ДНК, то SpyCas9 вносит двухцепочечный разрыв в ДНК, который служит сигналом к инициации репарации. Чаще всего в клетках высших эукариот двухцепочечные разрывы ДНК репарируются мутагенной системой NHEJ (Non-homologous End Joining), результатом чего становятся вставки или делеции в области разрыва. В случае кодирующих областей генов это чаще всего приводит к сдвигу рамки считывания или разрушению сайта сплайсинга и в итоге к синтезу дефектного белка.

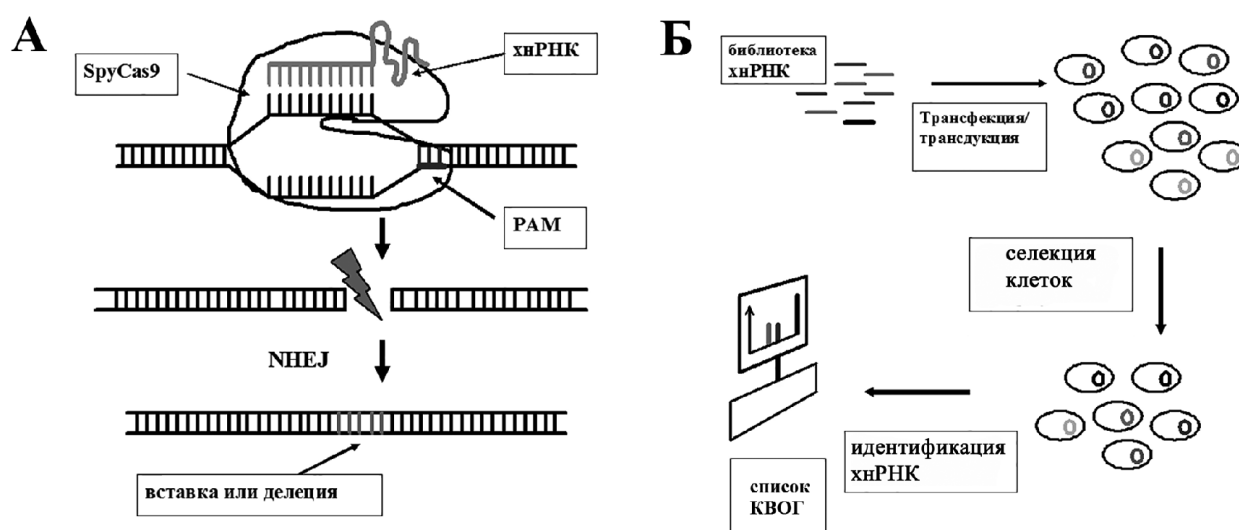


Рисунок 2. Идентификация КВГ/КВОГ с помощью системы CRISPR/Cas9. (А) Схема внесения изменений в геном. РНК-зависимая ДНК-эндонуклеаза Cas9 *S. pyogenes* в комплексе с хнРНК распознаёт РАРМ и протоспейсер мишени ДНК. Затем Cas9 вносит двухцепочечный разрыв ДНК, активирующий мутагенную систему репарации ДНК с соединением негомологичных концов, что приводит к возникновению вставок и делеций в участке мишени. (Б) хнРНК-скрининг. Библиотека хнРНК вводится в клетки. После селекции на антибиотике и селекции в присутствии лекарственного препарата из выживших клеток выделяется геномная ДНК и секвенируются фрагменты, содержащие спейсеры хнРНК. По относительному количеству спейсеров определяют КВОГ.

Для более эффективного нарушения функций гена систему CRISPR/Cas нацеливают на начало кодирующей области гена. Однако небольшие делеции, не приводящие к сдвигу рамки считывания в малозначимом домене, могут не повлиять на активность белка. Более эффективным подходом можно считать нацеливание системы CRISPR/Cas на участки, кодирующие важные каталитические домены [9]. В полногеномных скринингах при помощи системы CRISPR установлено, что эффективность повреждения фрагмента ДНК падает с увеличением числа его копий в геноме [9, 10, 24]. Вместе с этим обнаружена цитотоксичность системы CRISPR/Cas, нацеленной против полиплоидных локусов, не зависящая от последовательности этих локусов [10], что может служить источником ложно-положительных результатов. Для борьбы с этим явлением разработан алгоритм CERES, использующий линейную регрессионную модель зависимости цитотоксичности хнРНК от числа копий гена и выявляющий вклад именно ген-специфичного эффекта [25].

Технология CRISPR/Cas9 активно используется для выполнения полногеномных скринингов на клеточных линиях человека (рис. 2Б). Так, в работе [26] использована библиотека GeCKO, состоящая из 64751 хнРНК против 18080 генов, с целью выявления КВОГ для линии A375 меланомы человека и плюрипотентных стволовых клеток HUES62. В этой же работе показано, что дисфункция генов *NF1*, *MED12*, *NF2*, *CUL3*, *TADA2B*, и *TADA1* связана с устойчивостью к ингибитору BRAF – вемурафенибу. В работе [24] использовали библиотеку из 73151 хнРНКs против 7114 генов для идентификации генов, нарушение функций которых повышает устойчивость хронической миелогенной лейкемии к 6-тиогуанину, а у клеточной линии лимфомы Беркитта HL60 – устойчивость к ингибитору ДНК топоизомеразы IА – этопозиду. В первом случае обнаружены гены репарации *MSH2*, *MSH6*, *MLH1* и *PMS2*, а во втором – гены *TOP2A* и *CDK6*.

В работе [27] авторы создали библиотеку высокого разрешения, состоящую из 176500 направляющих РНК против 17661 генов человека, и провели скрининг на КВОГ у клеток различных линий колоректального рака, рака шейки матки, глиобластомы, и hTERT-иммортилизованного эпителия сетчатки. Результаты были объединены с полученными ранее данными по меланоме и проанализированы с помощью алгоритма BAGEL. В итоге выявлено более 2000 важных генов, включая 1580 КВОГ в соответствии с моделью “ромашки” [28]. Результаты скрининга подтверждены данными ингибиторного анализа. Интересным результатом стало обнаружение различий между близкими линиями колоректального рака. Так, линия DLD1 избирательно подавляется ингибитором сигнального пути EGFR – эрлотинибом, тогда как HCT116 избирательно подавляется метформином – ингибитором комплекса окислительного фосфорилирования ETC I.

В работе [29] использовали библиотеку из 278000 хнРНК для профилирования менее

охарактеризованных областей генома, включающих 19000 генов Refseq и 3900 гипотетических генов. В результате был обнаружен 2280 КВОГ, причём 486 из них перекрывались с предшествующими работами [24, 27, 30].

Технология CRISPR/Cas9 не лишена недостатков, главным из которых является наличие внецелевой активности Cas9, которая зависит от последовательности хнРНК [31].

1.3. Проблемы и трудности

Экспериментальные методы представляются наиболее естественным способом поиска КВОГ, однако, единое мнение относительно “золотого стандарта” по их определению пока не выработано.

Скрининговые исследования КВГ/КВОГ были проведены для многих модельных организмов и человеческих клеточных линий. Были предприняты попытки идентифицировать набор КВОГ, общих для всех типов клеток. Например, в трёх последовательных работах по РНК-интерференции [11, 12, 32] сообщалось приблизительно о 250-300 таких КВОГ. В то же время в недавних работах по CRISPR в среднем идентифицировано от 1500 до 2200 КВОГ, которые связаны с метаболизмом, репликацией ДНК, транскрипцией, сплайсингом и биосинтезом белка [24, 27, 30]. Прямое сравнение CRISPR и РНК-интерференции пока даёт неоднозначные результаты. Например, согласно данным [33], CRISPR/Cas9 показывает лучшие результаты по сравнению с РНК-интерференцией, в то время как в работе [34] сделан вывод о сопоставимости обеих технологий. При работе с пятью клеточными линиями только 3% генов были детектированы как КВОГ согласно обоим методам [9]. Поэтому оптимальной стратегией, по-видимому, является одновременное использование обоих подходов. Например, в [29] разработан алгоритм RANKS, который позволяет работать с обоими типами данных, вычисляя уровень значимости эффекта подавления отдельного гена.

Одним из возможных объяснений различий результатов является переменная эффективность генетического нокдауна с помощью РНК-интерференции [17]. Например, ген может быть критически важным в случае полного подавления его экспрессии, но может оказывать различное влияние на жизнеспособность клеток, когда ингибирован только частично. В таком случае, ген будет детектирован как универсальный КВОГ с использованием нокаутных методов (CRISPR/Cas9), или как контекст-зависимый КВОГ с использованием генетического нокдауна (РНК-интерференция).

Существуют и общебиологические трудности по идентификации КВОГ. Исследователи ограничены экспериментами на человеческих клеточных линиях, поскольку по понятным причинам генетические вмешательства *in vivo* проводить невозможно. Однако работа с клеточными линиями не позволяет смоделировать межклеточные взаимодействия и микроокружение опухоли [35]. Свойства клеточной линии могут адаптироваться под условия

культивирования в лаборатории, включая накопление новых мутаций. Все эти факторы могут быть важны для достоверного определения КВОГ и затруднять трансляцию полученных результатов в практику.

2. ВЫЧИСЛИТЕЛЬНЫЕ ПРЕДСКАЗАНИЯ

Существенной проблемой, сдерживающей экспериментальное определение КВОГ методом “грубой силы” является необходимость перебора множества комбинаций из подавляемого гена, клеточной линии и используемой технологии. Дополнительным фактором является низкая вероятность “успеха” – обнаружения КВОГ. Например, согласно работе [28], число КВОГ для человека равно 291, что составляет 1,4% от общего количества генов. Поэтому экспериментальные подходы естественным образом дополняются биоинформатическими методами *in silico*. На рисунке 3 схематично проиллюстрированы основные подходы к предсказанию КВОГ, которые детально описываются ниже.

2.1. Эволюционные методы

Четыре десятилетия назад в работе [36] было высказано предположение, что белки, необходимые для жизнеспособности или фертильности организма, должны эволюционировать с более низкой скоростью по сравнению с заменимыми белками, то есть находиться под давлением естественного отбора. Поэтому одним из направлений по предсказанию КВОГ является поиск генов, для которых существуют признаки очищающего (негативного) отбора при клональной эволюции опухоли.

Для определения давления отбора на уровне гена часто используют величину dN/dS (известную также как соотношение Ka/Ks) [37], которая представляет собой нормализованное соотношение между несинонимичными и синонимичными мутациями. Если ген находится под действием негативного отбора, то данная величина меньше единицы, поскольку несинонимичные мутации в таком гене имеют меньший шанс закрепиться в популяции. Тест dN/dS широко используется в молекулярной

эволюции, а благодаря современному прогрессу онкогеномики этот метод возможно применять и для определения отбора в раковых клетках [38, 39]. В работе [40] данный тест был использован в качестве основы для предсказания 92 КВОГ генов меланомы.

Однако недавние исследования показали, что непосредственное применение теста dN/dS для данных по соматическим мутациям приводит к заниженным оценкам этой величины, тем самым завышая долю генов, которые якобы находятся под давлением негативного отбора [41, 42]. Это обусловлено существованием определённых паттернов мутаций. Было обнаружено, что доли разных типов однонуклеотидных замен (C>A, C>G, C>T, T>A, T>C, T>G) с учётом контекста соседних нуклеотидов существенно различаются в рамках генома, причём для каждого рака характерен определённый набор таких паттернов (сигнатура) [43]. Считается, что появление паттернов является следствием различных мутационных процессов [44]. Например, для меланомы характерна сигнатура с преобладанием индуцированных УФ-излучением замен C>T.

За последнее время разработано несколько альтернативных подходов к оценке давления отбора, учитывающие контекстно-зависимые мутационные процессы. Для этого используются как модификации теста dN/dS [41, 42], так и байесовский подход [45]. Результаты этих работ свидетельствуют о том, что сигнал отрицательного отбора в раке весьма незначительный. Так, согласно [41], только 0,14% генов находятся под отрицательным отбором, для небольшой доли генов (2,2%) характерен положительный отбор, в то время как подавляющее большинство генов аккумулирует мутации нейтрально – соответствующие величины dN/dS приблизительно равны единице. Единственной категорией, для которой удалось показать связь между пониженным значением величины dN/dS и эссенциальностью гена – это nonsense-мутации в гемизиготных участках генома, что согласуется с данными работы [46].

Подводя итог, следует отметить крайне незначительный уровень отрицательного отбора в раковом геноме [41, 45, 46]. Частично это можно

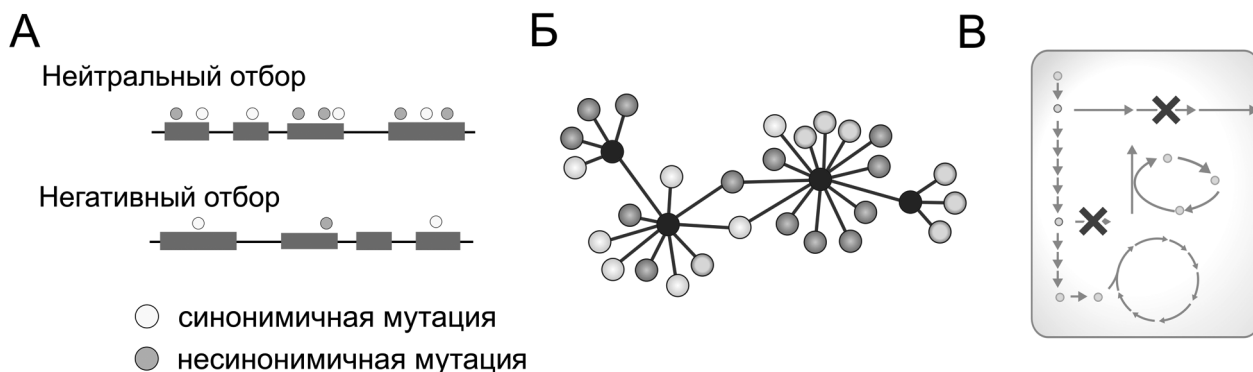


Рисунок 3. Различные подходы к вычислительному предсказанию КВГ/КВОГ. (А) Эволюционные методы - поиск генов, для которых существуют признаки негативного отбора (доля несинонимичных мутаций меньше доли синонимичных мутаций); (Б) Интерактомные методы - анализ топологии сетей белок-белковых взаимодействий; (В) Моделирование внутриклеточного метаболизма - определение реакций, в наибольшей степени необходимых для роста и пролиферации опухоли.

объяснить недостатком имеющихся данных по соматическим мутациям. Влияние негативного отбора по сравнению с положительным отбором сложнее зафиксировать, поскольку статистическая мощность для достоверной детекции дефицита мутаций меньше, чем для детекции избытка мутаций. Помимо упоминавшихся специфических паттернов однонуклеотидных замен, другими осложняющими факторами является различие в интенсивности мутагенеза вдоль генома, потенциальная “загрязнённость” данных герминативными мутациями и глобальная гетерогенность генетической картины опухоли.

2.2. Интерактомика

Клеточная функция каждого белка реализуется через взаимодействие с другими белками: субъединицами, если белок входит в состав молекулярного комплекса, или прочими участниками биохимического процесса, если белок участвует в метаболическом или сигнальном каскаде. В настоящее время сети таких белок-белковых взаимодействий охарактеризованы для большинства модельных организмов. В результате было показано, что КВГ склонны иметь больше партнеров в сети белок-белковых взаимодействий по сравнению с другими белками – так называемая концепция “центральность-летальность” [47].

Данное наблюдение привело к формированию подхода по предсказанию эссенциальности белков, опираясь на топологические свойства сети белок-белковых взаимодействий [48]. Белок-белковые взаимодействия в клетке формально описываются в виде графа, состоящего из вершин и соединяющих их рёбер. Это позволяет применять в биологии широко разработанный математический аппарат теории графов. Существуют различные подходы к оценке величины центральности (“важности”) каждой вершины графа: количество входящих и исходящих рёбер (центральность по степени), частота прохождения через данную вершину кратчайших путей между всеми парами вершин (центральность по посредничеству), отношение числа других узлов графа к сумме расстояний между данным узлом и всеми другими (центральность по близости) и т.д. Например, в работе [49] для идентификации КВГ на примере *S. cerevisiae* сравниваются 16 различных определений центральности, причем максимальная точность предсказания составила почти 70%. Благодаря введению новых топологических мер (локальная плотность центральности) этот результат был улучшен почти на 10% [50, 51]. Также используются и другие специфичные дескрипторы: число генов, которые регулирует данный КВОГ, количество регуляторов верхнего уровня (для транскрипционных факторов), число ферментов, участвующих в реакциях с метаболитами, получающимися в ходе реакции (для ферментов) [52].

В последнее время фокус исследований в области интерактомики смещается в сторону изучения контролируемости сетей. Сеть считается контролируемой, если, воздействуя на определённые

узлы, её возможно привести в желаемое состояние за конечное время [53]. Недавно было опубликовано программное обеспечение CytoCtrlAnalyser [54], в котором реализовано девять алгоритмов, позволяющих изучать белок-белковые взаимодействия с точки зрения контролируемости сети.

В работе [55] определили минимальный набор белков, воздействие на которые позволяет контролировать динамику белок-белковых взаимодействий. Такие контролируемые белки являются перекрестками для различных клеточных процессов и обогащены онко-ассоциированными белками, а также достоверно чаще кодируются экспериментально определенными КВОГ [56]. Это позволяет рассматривать анализ контролируемости как новый метод по предсказанию уязвимых точек в геноме и протеоме.

Анализ контролируемости сетей белок-белковых взаимодействий применительно к раковым клеткам представлен в работе [57]. В ней было показано, что для управления полной интерактомной сетью требуется контроль над приблизительно 70% всех белков, в то время как для управления сетью КВОГ требуется контролировать только 6% всех узлов сети. В ходе такого анализа было выявлено несколько белков-мишеней существующих лекарств, которые пока не используются в противораковой терапии.

2.3. Вычислительный анализ метаболизма

Ещё одним подходом к предсказанию КВГ является вычислительный анализ внутриклеточного метаболизма. Подобные попытки успешно предпринимались для одноклеточных модельных организмов как *E. coli* [58], *S. cerevisiae* [59] и *M. genitalium* [60]. Известно, что для раковых клеток характерны специфические особенности метаболизма, необходимые для ускоренного роста и пролиферации клеток [61], а репрограммирование биоэнергетики (аэробный гликолиз, так называемый эффект Варбурга) отмечено как отличительная черта опухолей [62].

Существуют попытки описать все протекающие в клетке метаболические реакции в едином математическом формализме, так называемые полногеномные модели метаболизма [63]. Много усилий было затрачено на реконструкцию метаболических сетей в тканях человека. Это привело к созданию модели метаболизма Recon 1, в которой интегрированы данные из более чем 1500 литературных источников [64]. В текущей версии Recon 2.2 содержится данные о 5324 метаболитах, 7785 реакциях и 1675 генах [65]. Накладывая различные ограничения на полногеномную модель, возможно реконструировать метаболизм специфичных клеточных типов. С этой целью разработаны такие алгоритмы как MBA [66], mCADRE [67], INIT [68] и tINIT [69]. В результате получается математическая модель, позволяющая симулировать протекание метаболических реакций в нескольких внутриклеточных компартментах. Дополнительная информация о специфике метаболизма той или иной ткани вносится в модель в виде результатов высокопроизводительных экспериментов,

например, учитываются транскриптомные данные о тканеспецифичности экспрессии ферментов. Также в большинство моделей могут быть добавлены данные протеомики и метаболомики.

Полученные модели внутриклеточного метаболизма анализируются для определения реакций, в наибольшей степени необходимых для роста и пролиферации опухоли. Для этого используется так называемый анализ баланса метаболических потоков. В данном методе изучается распределение метаболических потоков (стационарных скоростей превращения веществ в метаболическом пути) и их относительный вклад в образование того или иного метаболита [70]. Это позволяет вычислить необходимые скорости реакций для оптимизации выхода какого-либо продукта, например общей биомассы. В работе [71] проиллюстрировано применение этого метода для моделей метаболизма светлоклеточной почечной карциномы и аденокарциномы простаты с целью определения КВОГ. Для рака почки авторам удалось предсказать и экспериментально подтвердить 5 КВОГ (*AGPAT6*, *GALT*, *GCLC*, *GSS*, *RRM2B*), связанных с метаболизмом глицеролипидов, галактозы, глутамата и 3-уреидопропионата. В случае же рака простаты предсказания метода не отличались от случайного шума. Для обоснования неудачи авторы выдвигают ряд возможных объяснений. Во-первых, в используемой модели гены классифицируются как КВОГ, если их продукты участвуют в метаболических реакциях по продуцированию биомассы. Однако для выживаемости раковых клеток важна продукция и других метаболитов, в частности NADP и антиоксидантов [72], что не учитывалось в модели. Во-вторых, как обсуждалось ранее, экспериментальные технологии по определению КВОГ также отличаются высоким уровнем шума и слабо согласуются между собой, и, следовательно, не могут служить “золотым стандартом”. Наконец, неполнота модели – отсутствующие ключевые реакции и метаболиты – также может приводить к недопредсказанию КВОГ [73].

В работе [74] изучали точность предсказания КВОГ при помощи построения тканеспецифичных моделей метаболизма для шести различных методов. Авторы использовали данные по четырём раковым клеточным линиям, а в качестве золотого стандарта были выбраны результаты технологии CRISPR. Для оптимизирования модели по продуцированию биомассы использовали анализ баланса метаболических потоков, причём ген классифицировали как КВОГ только в том случае, если максимальная скорость роста после виртуального нокаута падала в 100 и более раз по сравнению с диким типом. Было показано, что тканеспецифичные модели метаболизма имеют лучшую предсказательную силу по сравнению с полногеномными моделями, а точность методов MBA[66], mCADRE [67] и INIT [68] находится приблизительно на одном уровне. Также было установлено, что важным фактором, влияющим на точность модели, является величина порога на генную экспрессию для классификации гена

как активного или неактивного в данной ткани. Показано, что более строгие значения данного порога ведут к более корректным предсказаниям КВОГ.

Моделирование метаболизма опухоли открывает возможности не только по поиску КВОГ, но и позволяет предсказывать потенциальные антиметаболиты – структурные аналоги эндогенных метаболитов, которые вызывают гибель раковой клетки путём одновременного ингибирования нескольких ферментов. Антиметаболиты являются одними из наиболее распространённых противораковых препаратов: антифолаты (метотрексат), антипиримидины (цитарабин, 5-фторурацил) и антипурины (6-меркаптопурин). В работе [69] были построены персональные модели метаболизма для шести пациентов с гепатоцеллюлярной карциномой. В результате работы алгоритма tINIT был предсказан 101 антиметаболит, который может быть эффективным в остановке роста опухоли у всех пациентов, причём 22 из них уже используются при различных методах лечения рака. В частности, для всех шести пациентов в качестве потенциально эффективных антиметаболитов для лечения гепатоцеллюлярной карциномы были предсказаны аналоги левокарнитина и метаболиты, связанные с его синтезом. В качестве экспериментальной проверки авторы показали ослабление роста клеточной линии HepG2 вследствие применения функционального аналога левокарнитина – пергекселина, являющегося ингибитором карнитин-пальмитоилтрансферазы. В работе [75] реконструировали модели метаболизма и предсказали 85 антиметаболитов, которые могут ингибировать рост или “убить” любую из проанализированных одиннадцати клеточных линий, и в то же время не токсичны для 83 типов клеток здоровой ткани человека. Подтверждая результаты работы [69], один из предсказанных антиметаболитов, пергекселин, был успешно провалидирован на клеточной линии рака простаты PC-3 и эпидермоидной карциномы A-431.

2.4. Проблемы и трудности

Применение вычислительных методов предсказания КВОГ сталкивается с рядом проблем, которые, в первую очередь, связаны с недостаточной экспериментальной основой. Например, в случае интерактивных методов анализируются сети белок-белковых взаимодействий, однако последние недостаточно изучены в случае раковых клеток, а большинство результатов получены на модельных организмах (в первую очередь для *S. cerevisiae*). Также никак не учитывается динамика процесса, то есть зависимость белок-белковых взаимодействий от времени. Все описанные вычислительные методы для предсказания КВОГ трудно верифицировать из-за отсутствия единого “золотого стандарта”, хотя в последнее время наметилась тенденция считать таковыми результаты технологии CRISPR/cas9.

В настоящее время точность согласия предсказания и эксперимента невелика. Например, в работе [45] величина площади под кривой чувствительности-специфичности составила 0,62,

что незначительно отличается от величины 0,5, полученной для случайного предсказания. Недавно в рамках проекта “DREAM Challenges” было проведено соревнование 48 научных групп по предсказанию КВОГ [76]. Исходные данные по РНК-интерференции из проекта Achilles [12] были расширены добавлением еще 149 клеточных линий, а общее количество генов составило 17000. Участникам требовалось предсказать вероятность классификации каждого гена как КВОГ в каждой клеточной линии. Условно истинным значением считался результат обработки экспериментальных данных алгоритмом DEMETER [20]. Дополнительные данные включали в себя результаты генотипирования, количество копий и экспрессию генов, а также уровни метилирования CpG-сайтов. В результате наибольшая величина корреляции между предсказанием и условной истиной составила всего 0,2, что представляется весьма скромным достижением. При этом лучшие команды использовали биологическое знание в виде известных сигнальных путей и опирались на транскриптомные данные, в то время как сведения о мутациях и копийности генов практически не были задействованы.

3. СИНТЕТИЧЕСКИЕ ЛЕТАЛИ

Естественным обобщением понятия КВГ/КВОГ на несколько генов является концепция синтетической летальности. Термин “синтетическая леталь” был введен в 40-х годах прошлого века Ф. Добржанским [77] в результате анализа предыдущих наблюдений над дрозофилами, когда выяснилось, что определенные сочетания мутаций – но не каждая мутация в отдельности – являются летальными для организма. За последующее время была проведена значительная работа по идентификации синтетических леталей модельных организмов, прежде всего для *S. cerevisiae*. Например, в работе [78] представлены результаты масштабного проекта, в ходе которого было исследовано 23 миллиона попарных комбинаций и выявлено 500000 синтетических летальных межгеновых взаимодействий. Помимо попарных летальных взаимодействий также существуют работы по изучению комбинаций из трёх генов [79]. Понятие синтетической летали проиллюстрировано на рисунке 4.

Идея целенаправленного поиска синтетических леталей, необходимых для выживания злокачественных опухолей, была высказана всего 20 лет назад в работе [80]. Поэтому объем экспериментальных данных, полученных для раковых клеточных линий, относительно невелик по сравнению с исследованиями, проведенными на модельных организмах. Однако благодаря современным высокопроизводительным технологиям, включая RNAi и CRISPR/cas9, общий массив данных по изучению синтетических леталей в раке быстро растёт [81, 82].

Как уже отмечалось, для поиска синтетических леталей необходимо осуществлять перебор

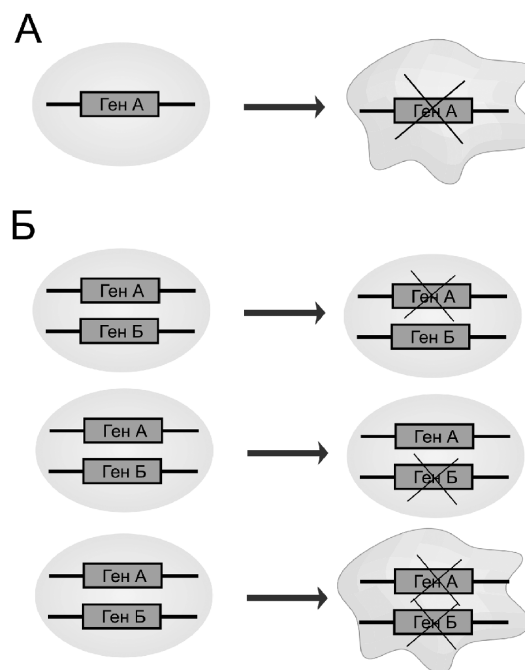


Рисунок 4. Понятие синтетической летальности. (А) Потеря функции КВГ/КВОГ приводит к немедленной гибели клеток. (Б) Потеря функции гена А или Б не обладает фенотипическим эффектом. Одновременная потеря функции обоих генов А и Б приводит к синтетическому летальному взаимодействию - гибели клетки

значительного количества сочетаний генов и клеточных линий. Поэтому здесь на первый план выходят биоинформатические методы, позволяющие сузить круг поисков. Например, алгоритм DAISY [83] предсказывает синтетические летали, сочетая анализ экспрессии и копийности генов с результатами генетического нокдауна с помощью мшРНК. При этом величина площади под кривой чувствительности-специфичности составила 0,78. Также удалось экспериментально подтвердить предсказанную эффективность для 6 из 9 препаратов, применявшихся для супрессии *VHL*-мутантной линии светлоклеточного рака почки. Недавно был предложен алгоритм MiSL [84], благодаря которому было экспериментально подтверждено синтетически летальное взаимодействие между мутантным геном *IDH1* и *ACACA* при остром миелоидном лейкозе.

Разработаны методы, которые, используя информацию об ортологах, позволяют “перенести” экспериментальные данные о синтетических летальных модельных организмах на другие виды [85]. Например, в работе [86] на клеточной линии рабдомиосаркомы сумели воспроизвести исходно открытую на *S. cerevisiae* синтетическую леталь между гистоновой деацетилазой RPD3 и ферментом репарации тирозил-ДНК фосфодиэстеразой TDP1. Другим примером переноса на основе эволюционной консервативности является синтетическая леталь между серинтреониновой протеинкиназой PLK1 и белком CKS1B, который регулирует циклин-зависимые протеинкиназы и часто гиперэкспрессирован при опухолях лёгких и печени [87].

Существуют методы, использующие концепции синтетической летальности для предсказания оптимальной стратегии борьбы с опухолью [88]. В работе [89] разработан алгоритм Statius, анализирующий данные о копийности генов для поиска метаболических уязвимостей. Для каждого образца определяются гены ферментов, которые подверглись гомозиготной делеции. Далее производится попытка определить изоферменты, которые возможно заблокировать, используя известный набор лекарственных средств.

Несмотря на всю привлекательность идеи о синтетической летальности с точки зрения расширения арсенала противораковой терапии, к настоящему времени существует единственный класс препаратов, основанный на этой концепции и одобренный для применения в клинике – так называемые PARP-ингибиторы. Данный класс лекарственных соединений ингибирует активность поли(ADP-рибоза)-полимераз (PARP), участвующих в репарации однонитевых разрывов ДНК. Если в клетке мутированы гены *BRCA1/2*, отвечающие за гомологичную репарацию двунитевых разрывов ДНК, то ингибирование PARP приводит к гибели клеток, хотя детали этого молекулярного механизма ещё не до конца изучены [90]. При этом даже для единственного клинически успешного примера PARP-ингибиторов не все пациенты одинаково реагируют на терапию, а в дальнейшем развивается резистентность [91]. Таким образом, проблема поиска в раковой клетке воспроизводимых и эффективных синтетических леталей остаётся крайне актуальной.

Незначительный прогресс в трансляции в практику может объясняться тем, что синтетические летальные взаимодействия между генами в значительной степени зависят от генетического контекста, тканеспецифичности и микроокружения опухоли. Поэтому необходимо разрабатывать методы отбора пациентов, для которых воздействие на синтетические летали будет эффективным. Также, синтетические летали являются редким событием, и для их идентификации требуются широкомасштабные скрининговые исследования для перебора парных сочетаний генов с учётом различных клеточных линий [90]. Наконец, для экспериментальных технологий характерны ложно-положительные находки. Например, исходное сообщение о том, что клетки с мутацией в гене *KRAS* зависят от активности серинтреониновой протеинкиназы *STK33* [92] было затем опровергнуто [93].

По-видимому, оптимальной стратегией для борьбы с ложно-положительными находками является использование мета-анализа, учёт побочных эффектов супрессии генов и валидирование с помощью других экспериментальных технологий. Так в работе [94] путём реанализа данных по мшРНК из проектов Achilles [12] и COLT-cancer [11] были найдены синтетические летали *PIK3CA-PKN3* и *PIK3CA-HMX3*, что впоследствии было воспроизведено с помощью технологии CRISPR/cas9.

4. НЕКОДИРУЮЩИЕ ОБЛАСТИ ГЕНОМА

Большинство упомянутых работ сфокусированы на изучении белок-кодирующих КВОГ. Однако за последнее время накапливается всё больше свидетельств в пользу того, что генетические варианты в некодирующих областях генома также играют важную роль в патогенезе заболеваний и могут быть связаны с выживаемостью клеток. Например, большинство маркеров, детектируемых в исследованиях по полногеномному поиску ассоциаций (GWAS), картируются именно на белок-некодирующие участки генома, включая регуляторные области, длинные некодирующие РНК, микроРНК [95].

Существует ряд трудностей при анализе некодирующих участков генома в силу их меньшей структурированности. Например, сложнее предсказывать биологический эффект генетического варианта, отсутствует понятие нонсенс-мутации. Тем не менее, разработан ряд методов предсказания степени патогенности мутации, включая CADD [96], DeepSEA [97], EIGEN [98]. Указанные алгоритмы опираются на результаты проекта ENCODE [99] и учитывают такие параметры как связывание транскрипционных факторов, наличие открытого хроматина, ассоциированного с гиперчувствительностью к ДНКазе I, уровень эволюционной консервативности и т.д.

В ряде публикаций продемонстрирована важность функционирования длинных некодирующих РНК (днРНК), этой “темной материи” генома [100]. Так, в работе [101] по результатам РНК-интерференции выявили 26 из 638 проанализированных днРНК, которые влияли на морфологию и клеточный цикл культуры HeLa. В частности, была подтверждена значимость днРНК LINC00152, которая экспрессируется во многих клеточных линиях, и играет ключевую роль в ряде опухолей, включая рак желудка [102] и светлоклеточную почечную карциному [103]. Для изучения днРНК также применяют и технологию CRISPR/cas9. По результатам скрининга 700 днРНК было обнаружено 43 днРНК, которые оказывали негативное влияние на рост клеточной линии рака печени [104].

Из методологических трудностей генетического нокаутирования днРНК можно отметить опасность повредить другие регуляторные элементы, включая энхансеры и микроРНК. Также некоторые днРНК находятся в интронах белок-кодирующих генов, а, следовательно, воздействие на них может нарушить процессы сплайсинга нецелевого гена. Это может снизить достоверность определения эссенциальности той или иной днРНК.

В настоящее время работы по планомерному изучению некодирующих участков генома, чья функциональность критически важна для онкогенеза, находятся на начальном этапе. Можно предсказать, что разработка этой до сих пор малоизученной области принесет плоды в виде новых типов мишеней для генной терапии онкологических заболеваний.

ЗАКЛЮЧЕНИЕ

В настоящем обзоре были рассмотрены различные подходы к определению генов, чье функционирование критически важно для выживания раковой клетки: экспериментальные технологии РНК-интерференции и CRISPR/cas9, вычислительные предсказания на основе онкогеномики, метаболомики и интерактомики, синтетические летали. Актуальность этой задачи не вызывает сомнения, поскольку каждый такой ген представляет собой потенциальную мишень для противоопухолевой терапии. Более того, благодаря активно развивающейся практике секвенирования раковых геномов и транскриптомов в клинике, существуют попытки применить указанные наработки на уровне отдельных пациентов [69]. Таким образом, изучение КВОГ теоретически позволяет значительно продвинуться на пути к прецизионной онкологии и персонализированному лечению.

В то же время существует ряд нерешённых фундаментальных и прикладных проблем, тормозящих развитие данной области. В первую очередь, это недостаток достоверных экспериментальных данных, слабое согласие между различными технологиями, и, как следствие, отсутствие общепризнанного “золотого стандарта” (например, аналога метода Сэнгера для генотипирования). Поскольку эксперименты на клеточных линиях не позволяют смоделировать межклеточные взаимодействия и микроокружение опухоли, то данные по КВОГ, полученные *in vitro*, могут не воспроизводиться при исследованиях *in vivo*. К настоящему моменту мало изучена значимость некодирующих областей генома для выживания раковой клетки. Наконец, всегда присутствует проблема гетерогенности опухоли, в результате которой снижается достоверность вновь получаемых данных о геноме и транскриптоме.

В то же время за последние пять лет изучение КВОГ получило мощный импульс в связи с внедрением технологии CRISPR/cas9. Значительно возрос объём доступных экспериментальных данных и количество публикуемых биоинформатических алгоритмов. Поэтому есть основания предполагать, что в обозримое время возможен переход количественных изменений в качественные, и работы в данной области смогут перешагнуть рубеж исключительно научных изысканий и будут применяться в клинике для борьбы с онкологическими заболеваниями.

БЛАГОДАРНОСТИ

Работа выполнена в рамках Программы фундаментальных научных исследований государственных академий наук на 2013-2020 годы.

ЛИТЕРАТУРА

- Brenner S. (1974) *Genetics*, **77**, 71-94.
- Goebel M.G., Petes T.D. (1986) *Cell*, **46**, 983-992.
- Kamath R.S., Fraser A.G., Dong Y., Poulin G., Durbin R., Gotta M., Kanapin A., Le Bot N., Moreno S., Sohrmann M., Welchman D.P., Zipperlen P., Ahringer J. (2003) *Nature*, **421**, 231-237.
- Winzeler E.A., Shoemaker D.D., Astromoff A., Liang H., Anderson K., Andre B., Bangham R., Benito R., Boeke J.D., Bussey H. et al. (1999) *Science*, **285**, 901-906.
- Rabik C.A., Dolan M.E. (2007) *Cancer Treat. Rev.*, **33**, 9-23.
- Pagliarini R., Shao W., Sellers W.R. (2015) *EMBO Rep.*, **16**, 280-296.
- Shirasawa S., Furuse M., Yokoyama N., Sasazuki T. (1993) *Science*, **260**, 85-88.
- Harborth J., Elbashir S.M., Bechert K., Tuschl T., Weber K. (2001) *J. Cell Sci.*, **114**, 4557-4565.
- Munoz D.M., Cassiani P.J., Li L., Billy E., Korn J.M., Jones M.D., Golji J., Ruddy D.A., Yu K., McAllister G. et al. (2016) *Cancer Discov.*, **6**, 900-913.
- Aguirre A.J., Meyers R.M., Weir B.A., Vazquez F., Zhang C.Z., Ben-David U., Cook A., Ha G., Harrington W.F., Doshi M.B. et al. (2016) *Cancer Discov.*, **6**, 914-929.
- Marcotte R., Brown K.R., Suarez F., Sayad A., Karamboulas K., Krzyzanowski P.M., Sircoulomb F., Medrano M., Fedyszyn Y., Koh J.L.Y. et al. (2012) *Cancer Discov.*, **2**, 172-189.
- Cowley G.S., Weir B.A., Vazquez F., Tamayo P., Scott J.A., Rusin S., East-Seletsky A., Ali L.D., Gerath W.F., Pantel S.E. et al. (2014) *Sci. Data*, **1**, 140035.
- Druker B.J. (2008) *Blood*, **112**, 4808-4817.
- Hynes N.E., Lane H.A. (2005) *Nat. Rev. Cancer*, **5**, 341-354.
- Helming K.C., Wang X., Wilson B.G., Vazquez F., Haswell J.R., Manchester H.E., Kim Y., Kryukov G.V., Ghandi M., Aguirre A.J., Jagani Z., Wang Z., Garraway L.A., Hahn W.C., Roberts C.W. (2014) *Nat. Med.*, **20**, 251-254.
- Hsu T.Y., Simon L.M., Neill N.J., Marcotte R., Sayad A., Bland C.S., Echeverria G.V., Sun T., Kurley S.J., Tyagi S. et al. (2015) *Nature*, **525**, 384-388.
- Housden B.E., Muhar M., Gemberling M., Gersbach C.A., Stainier D.Y., Seydoux G., Mohr S.E., Zuber J., Perrimon N. (2017) *Nat. Rev. Genet.*, **18**, 24-40.
- Kittler R., Putz G., Pelletier L., Poser I., Heninger A.K., Drechsel D., Fischer S., Konstantinova I., Habermann B., Grabner H., Yaspo M.L., Himmelbauer H., Korn B., Neugebauer K., Pisabarro M.T., Buchholz F. (2004) *Nature*, **432**, 1036-1040.
- Shao D.D., Tsherniak A., Gopal S., Weir B.A., Tamayo P., Stransky N., Schumacher S.E., Zack T.I., Beroukhim R., Garraway L.A., Margolin A.A., Root D.E., Hahn W.C., Mesirov J.P. (2013) *Genome Res.*, **23**, 665-678.
- Tsherniak A., Vazquez F., Montgomery P.G., Weir B.A., Kryukov G., Cowley G.S., Gill S., Harrington W.F., Pantel S., Krill-Burger J.M. et al. (2017) *Cell*, **170**, 564-576.
- Bhinder B., Djaballah H. (2013) *Comb. Chem. High Throughput Screen.*, **16**, 665-681.
- Cong L., Ran F.A., Cox D., Lin S., Barretto R., Habib N., Hsu P.D., Wu X., Jiang W., Marraffini L.A., Zhang F. (2013) *Science*, **339**, 819-823.
- Jinek M., Chylinski K., Fonfara I., Hauer M., Doudna J.A., Charpentier E. (2012) *Science*, **337**, 816-821.
- Wang T., Birsoy K., Hughes N.W., Krupczak K.M., Post Y., Wei J.J., Lander E.S., Sabatini D.M. (2015) *Science*, **350**, 1096-1101.
- Meyers R.M., Bryan J.G., McFarland J.M., Weir B.A., Sizemore A.E., Xu H., Dharia N.V., Montgomery P.G., Cowley G.S., Pantel S. et al. (2017) *Nat. Genet.*, **49**, 1779-1784.

26. Shalem O., Sanjana N.E., Hartenian E., Shi X., Scott D.A., Mikkelsen T., Heckl D., Ebert B.L., Root D.E., Doench J.G., Zhang F. (2014) *Science*, **343**, 84-87.
27. Hart T., Chandrashekar M., Aregger M., Steinhart Z., Brown K.R., MacLeod G., Mis M., Zimmermann M., Fradet-Turcotte A., Sun S. et al. (2015) *Cell*, **163**, 1515-1526.
28. Hart T., Brown K.R., Sircoulomb F., Rottapel R., Moffat J. (2014) *Mol. Syst. Biol.*, **10**, 733.
29. Bertomeu T., Coulombe-Huntington J., Chatr-Aryamontri A., Bourdages K.G., Coyaud E., Raught B., Xia Y., Tyers M. (2018) *Mol. Cell Biol.*, **38**, e00302-17.
30. Blomen V.A., Majek P., Jae L.T., Bigenzahn J.W., Nieuwenhuis J., Staring J., Sacco R., van Diemen F.R., Olk N., Stukalov A., Marceau C., Janssen H., Carrette J.E., Bennett K.L., Colinge J., Superti-Furga G., Brummelkamp T.R. (2015) *Science*, **350**, 1092-1096.
31. Peng R., Lin G., Li J. (2016) *FEBS J.*, **283**, 1218-1231.
32. Luo B., Cheung H.W., Subramanian A., Sharifnia T., Okamoto M., Yang X., Hinkle G., Boehm J.S., Beroukhim R., Weir B.A. et al. (2008) *Proc. Natl. Acad. Sci. USA*, **105**, 20380-20385.
33. Evers B., Jastrzebski K., Heijmans J.P., Grernrum W., Beijersbergen R.L., Bernards R. (2016) *Nat. Biotechnol.*, **34**, 631-633.
34. Morgens D.W., Deans R.M., Li A., Bassik M.C. (2016) *Nat. Biotechnol.*, **34**, 634-636.
35. Bissell M.J., Hines W.C. (2011) *Nat. Med.*, **17**, 320-329.
36. Wilson A.C., Carlson S.S., White T.J. (1977) *Ann. Rev. Biochem.*, **46**, 573-639.
37. Nei M., Gojobori T. (1986) *Mol. Biol. Evol.*, **3**, 418-426.
38. Ostrow S.L., Barshir R., DeGregori J., Yeager-Lotem E., Hershsberg R. (2014) *PLoS Genet.*, **10**, e1004239.
39. Ovens K., Naugler C. (2012) *Theor. Biol. Med. Model.*, **9**, 44.
40. Pyatnitskiy M., Karpov D., Poverennaya E., Lisitsa A., Moshkovskii S. (2015) *PLoS One*, **10**, e0142819.
41. Martincorena I., Raine K. M., Gerstung M., Dawson K.J., Haase K., Van Loo P., Davies H., Stratton M.R., Campbell P.J. (2017) *Cell*, **171**, 1029-1041.
42. Van den Eynden J., Larsson E. (2017) *Front Genet.*, **8**, 74.
43. Alexandrov L.B., Nik-Zainal S., Wedge D.C., Aparicio S.A., Behjati S., Biankin A.V., Bignell G.R., Bolli N., Borg A., Borresen-Dale A.L. et al. (2013) *Nature*, **500**, 415-421.
44. Alexandrov L.B., Stratton M.R. (2014) *Curr. Opin. Genet. Dev.*, **24**, 52-60.
45. Weghorn D., Sunyaev S. (2017) *Nat. Genet.*, **49**, 1785-1788.
46. Van den Eynden J., Basu S., Larsson E. (2016) *PLoS Genet.*, **12**, e1006506.
47. Jeong H., Mason S.P., Barabasi A.L., Oltvai Z.N. (2001) *Nature*, **411**, 41-42.
48. Yu H.Y., Greenbaum D., Lu H.X., Zhu X.W., Gerstein M. (2004) *Trends Genet.*, **20**, 227-231.
49. del Rio G., Koschutski D., Coello G. (2009) *BMC Syst. Biol.*, **3**, 102.
50. Luo J., Qi Y. (2015) *PLoS One*, **10**, e0131418.
51. Qin C., Sun Y., Dong Y. (2016) *PLoS One*, **11**, e0161042.
52. da Silva J.P.M., Acencio M.L., Mornbach J.C.M., Vieira R., da Silva J.C., Lemke N., Sinigaglia M. (2008) *Physica A*, **387**, 1049-1055.
53. Liu Y.Y., Slotine J.J., Barabasi A.L. (2011) *Nature*, **473**, 167-173.
54. Wu L., Li M., Wang J., Wu F.X. (2018) *Bioinformatics*, **34**, 1428-1430.
55. Wuchty S. (2014) *Proc. Natl. Acad. Sci. USA*, **111**, 7156-7160.
56. Khuri S., Wuchty S. (2015) *BMC Bioinformatics*, **16**, 109.
57. Kanhaiya K., Czeizler E., Gratie C., Petre I. (2017) *Sci. Rep.*, **7**, 10327.
58. Edwards J.S., Palsson B.O. (2000) *Proc. Natl. Acad. Sci. USA*, **97**, 5528-5533.
59. Heavner B.D., Smallbone K., Price N.D., Walker L.P. (2013) *Database (Oxford)*, **2013**, bat059.
60. Karr J.R., Sanghvi J.C., Macklin D.N., Gutschow M.V., Jacobs J.M., Bolival B., Jr., Assad-Garcia N., Glass J.I., Covert M.W. (2012) *Cell*, **150**, 389-401.
61. Maximchik P.V., Kulikov A.V., Zhivotovsky B.D., Gogvadze V.G. (2016) *Biochemistry (Moscow)*, **81**, 65-79.
62. Hanahan D., Weinberg R.A. (2011) *Cell*, **144**, 646-674.
63. Nilsson A., Nielsen J. (2017) *Metab. Eng.*, **43**, 103-112.
64. Duarte N.C., Becker S.A., Jamshidi N., Thiele I., Mo M.L., Vo T.D., Srivas R., Palsson B.O. (2007) *Proc. Natl. Acad. Sci. USA*, **104**, 1777-1782.
65. Swainston N., Smallbone K., Hefzi H., Dobson P.D., Brewer J., Hanscho M., Zielinski D.C., Ang K.S., Gardiner N.J., Gutierrez J.M. et al. (2016) *Metabolomics*, **12**, 109.
66. Jerby L., Shlomi T., Ruppin E. (2010) *Mol. Syst. Biol.*, **6**, 401.
67. Wang Y., Eddy J.A., Price N.D. (2012) *BMC Syst. Biol.*, **6**, 153.
68. Agren R., Bordel S., Mardinoglu A., Pornputtapong N., Nookaew I., Nielsen J. (2012) *PLoS Comput. Biol.*, **8**, e1002518.
69. Agren R., Mardinoglu A., Asplund A., Kampf C., Uhlen M., Nielsen J. (2014) *Mol. Syst. Biol.*, **10**, 721.
70. Orth J.D., Thiele I., Palsson B.O. (2010) *Nat. Biotechnol.*, **28**, 245-248.
71. Gatto F., Miess H., Schulze A., Nielsen J. (2015) *Sci. Rep.*, **5**, 10738.
72. Cairns R.A., Harris I.S., Mak T.W. (2011) *Nat. Rev. Cancer*, **11**, 85-95.
73. Becker S.A., Palsson B.O. (2008) *BMC Syst. Biol.*, **2**, 14.
74. Opdam S., Richelle A., Kellman B., Li S., Zielinski D.C., Lewis N.E. (2017) *Cell Syst.*, **4**, 318-329.
75. Ghaffari P., Mardinoglu A., Asplund A., Shoaie S., Kampf C., Uhlen M., Nielsen J. (2015) *Sci. Rep.*, **5**, 8183.
76. Gonen M., Weir B.A., Cowley G.S., Vazquez F., Guan Y., Jaiswal A., Karasuyama M., Uzunangelov V., Wang T., Tsherniak A. et al. (2017) *Cell Syst.*, **5**, 485-497.
77. Dobzhansky T. (1946) *Genetics*, **31**, 269-290.
78. Costanzo M., VanderSluis B., Koch E.N., Baryshnikova A., Pons C., Tan G., Wang W., Usaj M., Hanchard J., Lee S.D., Pelechano V. et al. (2016) *Science*, **353**(6306), aaf1420.
79. Haber J.E., Braberg H., Wu Q., Alexander R., Haase J., Ryan C., Lipkin-Moore Z., Franks-Skiba K.E., Johnson T., Shales M., Lenstra T.L., Holstege F.C., Johnson J.R., Bloom K., Krogan N.J. (2013) *Cell Rep.*, **3**, 2168-2178.
80. Hartwell L.H., Szankasi P., Roberts C.J., Murray A.W., Friend S.H. (1997) *Science*, **278**, 1064-1068.
81. Dhanjal J.K., Radhakrishnan N., Sundar D. (2017) *Methods*, **131**, 66-73.
82. Thompson J.M., Nguyen Q.H., Singh M., Razorenova O.V. (2015) *Yale J. Biol. Med.*, **88**, 145-155.
83. Jerby-Aronson L., Pfetzer N., Waldman Y.Y., McGarry L., James D., Shanks E., Seashore-Ludlow B., Weinstock A., Geiger T., Clemons P.A., Gottlieb E., Ruppin E. (2014) *Cell*, **158**, 1199-1209.
84. Sinha S., Thomas D., Chan S., Gao Y., Brunen D., Torabi D., Reinisch A., Hernandez D., Chan A., Rankin E.B., Bernards R., Majeti R., Dill D.L. (2017) *Nat. Commun.*, **8**, 15580.
85. Jacunski A., Dixon S.J., Tatonetti N.P. (2015) *PLoS Comput. Biol.*, **11**, e1004506.

86. Duffy S., Fam H.K., Wang Y.K., Styles E.B., Kim J.H., Ang J.S., Singh T., Larionov V., Shah S.P., Andrews B., Boerkoel C.F., Hieter P. (2016) *Proc. Natl. Acad. Sci. USA*, **113**, 9967-9976.
87. Reid R.J., Du X., Sunjevaric I., Rayannavar V., Dittmar J., Bryant E., Maurer M., Rothstein R. (2016) *Genetics*, **204**, 807-819.
88. Megchelenbrink W., Katzir R., Lu X., Ruppini E., Notebaart R.A. (2015) *Proc. Natl. Acad. Sci. USA*, **112**, 12217-12222.
89. Aksoy B.A., Demir E., Babur O., Wang W., Jing X., Schultz N., Sander C. (2014) *Bioinformatics*, **30**, 2051-2059.
90. O'Neil N.J., Bailey M. L., Hieter P. (2017) *Nat. Rev. Genet.*, **18**, 613-623.
91. Bouwman P., Jonkers J. (2014) *Clin. Cancer Res.*, **20**, 540-547.
92. Scholl C., Frohling S., Dunn I.F., Schinzel A.C., Barbie D.A., Kim S.Y., Silver S.J., Tamayo P., Wadlow R.C., Ramaswamy S., Dohner K. et al. (2009) *Cell*, **137**, 821-834.
93. Babij C., Zhang Y., Kurzeja R.J., Munzli A., Shehabeldin A., Fernando M., Quon K., Kassner P.D., Ruefli-Brasse A.A., Watson V.J. et al. (2011) *Cancer Res.*, **71**, 5818-5826.
94. Jaiswal A., Peddinti G., Akimov Y., Wennerberg K., Kuznetsov S., Tang J., Aittokallio T. (2017) *Genome Med.*, **9**, 51.
95. Zhang F., Lupski J.R. (2015) *Hum. Mol. Genet.*, **24**, R102-R110.
96. Kircher M., Witten D.M., Jain P., O'Roak B.J., Cooper G.M., Shendure J. (2014) *Nat. Genet.*, **46**, 310-315.
97. Zhou J., Troyanskaya O.G. (2015) *Nat. Methods*, **12**, 931-934.
98. Ionita-Laza I., McCallum K., Xu B., Buxbaum J.D. (2016) *Nat. Genet.*, **48**, 214-220.
99. ENCODE Project Consortium (2012) *Nature*, **489**, 57-74.
100. Hu X., Sood A.K., Dang C.V., Zhang L. (2018) *Curr. Opin. Genet. Dev.*, **48**, 8-15.
101. Notzold L., Frank L., Gandhi M., Polycarpou-Schwarz M., Gross M., Gunkel M., Beil N., Erfle H., Harder N., Rohr K. et al (2017) *Sci. Rep.*, **7**, 2265.
102. Zhao J., Liu Y., Zhang W., Zhou Z., Wu J., Cui P., Zhang Y., Huang G. (2015) *Cell Cycle*, **14**, 3112-3123.
103. Wu Y., Tan C., Weng W.W., Deng Y., Zhang Q.Y., Yang X.Q., Gan H.L., Wang T., Zhang P.P., Xu M.D., Wang Y.Q., Wang C.F. (2016) *Am. J. Cancer Res.*, **6**, 285-299.
104. Zhu S., Li W., Liu J., Chen C. H., Liao Q., Xu P., Xu H., Xiao T., Cao Z., Peng J., Yuan P., Brown M., Liu X.S., Wei W. (2016) *Nat. Biotechnol.*, **34**, 1279-1286.

Поступила: 20. 05. 2018.
Принята к печати: 27. 06. 2018.

SEARCHING FOR ESSENTIAL GENES IN CANCER GENOMES

M.A. Pyatnitskiy^{1,2}, D.S. Karpov^{1,3}, S.A. Moshkovskii^{1,4}

¹Institute of Biomedical Chemistry,
10 Pogodinskaya str., Moscow, 119121 Russia; e-mail: mpyat@bioinformatics.ru

²Higher School of Economics, Moscow, Russia

³Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, Russia

⁴Pirogov Russian National Research Medical University, Moscow, Russia

The concept of essential genes, whose loss of functionality leads to cell death, is one of the fundamental concepts of genetics and is important for fundamental and applied research. This field is particularly promising in relation to oncology, since the search for genetic vulnerabilities of cancer cells allows us to identify new potential targets for antitumor therapy. The modern biotechnology capacities allow carrying out large-scale projects for sequencing somatic mutations in tumors, as well as directly interfering the genetic apparatus of cancer cells. They provided accumulation of a considerable body of knowledge about genetic variants and corresponding phenotypic manifestations in tumors. In the near future this knowledge will find application in clinical practice. This review describes the main experimental and computational approaches to the search for essential genes, concentrating on the application of these methods in the field of molecular oncology.

Key words: cancer genome; essential genes; CRISPR; RNAi; synthetic lethals