

## **AID/АРОВЕС ЗАВИСИМЫЙ СОМАТИЧЕСКИЙ ГИПЕРМУТАГЕНЕЗ И ПЕРЕСТРОЙКИ ДНК ИММУНОГЛОБУЛИНОВЫХ И НЕИММУНОГЛОБУЛИНОВЫХ ГЕНОВ**

**Колотова Т. Ю.<sup>1</sup>, Макаренко В.Д.<sup>2</sup>, Сорокоумова  
Л.К.<sup>3</sup>, Давиденко М. Б.<sup>1</sup>**

<sup>1</sup> ГУ «Институт микробиологии и иммунологии  
им. И. И. Мечникова Национальной академии  
медицинских наук Украины», г. Харьков;

<sup>2</sup> Харьковская медицинская академия  
последипломного образования МЗ Украины,  
г. Харьков;

<sup>3</sup> Винницкий национальный медицинский  
университет имени Н. И. Пирогова МЗ Украины,  
г. Винница

Редактирование индуцируемой активацией дезаминазой (AID) ДНК лежит в основе соматического гипермутагенеза Ig генов, создающего разнообразие антител, и рекомбинации переключения классов Ig. Действие дезаминаз AID/АРОВЕС семейства не ограничивается только Ig локусами. В геноме обнаружено множество генетически незапрограммированных локусов-мишеней, в которых образуются AID/АРОВЕС зависимые повреждения и двойные разрывы ДНК, ведущие к перестройкам генома и соматическому мутагенезу.

В статье обсуждаются результаты исследований, в которых изучались механизмы локализации активности AID в Ig локусах и AID/АРОВЕС в не-Ig-локусах. Эпигенетические модификации гистонов и ДНК, цис-элементы ДНК и связывающиеся с ним транскрипционные факторы, образующие регуляторные кластеры и суперэнхансеры, синтез некодирующей РНК, особенности режима транскрипции участвуют в локус-специфическом привлечении AID дезаминазы. Однако не один из изученных факторов не является специфическим для локусов-мишеней. Таким образом, до настоящего момента нет достаточной ясности в вопросе о том, что же определяет генетически запрограммированную активность AID в Ig локусах, и генетически незапрограммированную активность дезаминаз AID/АРОВЕС семейства в не-Ig-локусах.

Изучение AID/АРОВЕС зависимого соматического гипермутагенеза и перестроек ДНК Ig и не-Ig-генов имеет как фундаментальное общебиологическое значение, так и прикладное медицинское. Фундаментальное значение заключается в том, что механизмы генетических перестроек в соматических клетках могут дать ключ к пониманию закономерностей эволюционной изменчивости живых организмов. Прикладное значение состоит в том, что на основе полученных данных о геномной нестабильности не-Ig-локусов могут быть разработаны методы, подавляющие нестабильность генома при канцерогенезе и замедляющие скорость формирования устойчивых к химиопрепаратам вариантов.

### **Введение**

Разнообразие антител формируется по крайней мере в результате трех основных процессов: V(D)J рекомбинации, рекомбинации переключения классов иммуноглобулиновых (Ig) генов и соматического гипермутагенеза Ig генов. Встреча с антигеном активирует В клетки и запускает процессы соматического гипермутагенеза и рекомбинации переключения классов Ig генов. В результате гипермутагенеза Ig гены приобретают множество мутаций, что позволяет возникнуть высокоаффинным антителам. При рекомбинации переключения образуются классы Ig, специализирующиеся на выполнении различных функций.

Ферменты семейства, которое называется индуцированная активацией цитозинная дезаминаза/аполипротеин В редактирующий комплекс (AID/АРОВЕС), относятся к суперсемейству дезаминаз, катализирующих дезаминирование азотистых оснований нуклеотидов и нуклеиновых кислот. К ферментам суперсемейства относятся цитидиндезаминазы, дезоксицитидилат монофосфат дезаминазы, гуаниновые дезаминазы, вовлеченные в метаболизм пиримидинов и пуринов у бактерий, эукариот и фагов. Члены AID/АРОВЕС семейства и некоторые другие семейства дезаминаз катализируют дезаминирование оснований в РНК и ДНК. Дезаминирование оснований запускает биохимические процессы, которые приводят к мутагенезу/редактированию молекул ДНК и РНК.

При встрече с антигеном и получении всех необходимых дополнительных кофакторов в В клетках активируется экспрессия AID дезаминазы, которая запускает гипермутагенез Ig генов. Активация В клеток параллельно с соматическим гипермутагенезом индуцирует процесс рекомбинации, вызывающий переключение классов Ig.

На первом этапе гипермутагенеза Ig генов AID дезаминирует цитозиновые (C) остатки и образуются дезоксиуридиновые остатки (dU). Субстратом для AID являются одноцепочечные участки молекулы ДНК. Далее в результате репликации U/G неправильно спаренных пар оснований возникают транзиции в районе C/G пар, C заменяется на тимин, а гуанин на аденин. Однако на этом процесс гипермутагенеза не заканчивается. dU является субстратом для эксцизионной репарации оснований (BER) и привлекает фермент урацил-ДНК гликозидазу (UNG), которая вырезает неправильный нуклеотид и создает одноцепочечную брешь. Репликация транслейзионными полимеразми через одноцепочечную брешь приводит к включению напротив любого из четырех нуклеотидов. Точный паттерн нуклеотидных замен, по-видимому, зависит от используемых транслейзионных полимераз.

U:G пара может распознаваться также системой репарации неправильно спаренных оснований (MMR). MMR узнает неправильно спаренные основания и вырезает множество нуклеотидов, иногда до нескольких тысяч нуклеотидов. В результате образуются протяженные участки одноцепочечной ДНК, которые затем

восстанавливаются с помощью транслейзионных полимераз. При использовании точных транслейзионных полимераз мутаций не возникает. Однако могут использоваться неточные транслейзионные полимеразы, вносящие мутации при синтезе ДНК, в том числе и мутации АТ пар.

При рекомбинации переключения дезаминирование С остатков приводит к появлению dU, процессуемого UNG гликозидазой с образованием одноцепочечной брши, которая привлекает апуриновую эндонуклеазу, вносящую одноцепочечный разрыв в ДНК. Одиночные разрывы ДНК возникают на обеих цепях ДНК в районе сайтов переключения (S) и, соответственно, формируются двойные разрывы ДНК (DSB) с выступающими концами. Выступающие концы либо съедаются экзонуклеазами, либо заполняются транслейзионными полимеразами, что и приводит к появлению мутаций. И, наконец, посредством репарации путем соединения негомологичных концов DSB разрывы сшиваются с образованием делеции.

АРОВЕС дезаминируют цитозин как в молекулах РНК, так и в одноцепочечной ДНК. Соответственно, АРОВЕС дезаминазы могут редактировать матричную РНК и РНК вирусов и вносить мутации в одноцепочечную ДНК. Считается, что основными мишенями АРОВЕС дезаминаз являются ретровирусы и ретротранспозоны. При обратной транскрипции ретровирусов и ретротранспозонов фермент осуществляет дезаминирование цитозиновых остатков и превращение С остатков минус цепи ДНК в dU. Помимо ретровирусов АРОВЕС3 дезаминазы изменяют геномы герпесвирусов, парвовирусов, папилломавирусов, вируса гепатита В [1-3]. В то же время АРОВЕС дезаминазы могут привлекаться к генетически незапрограммированным мишеням и вызывать в них соматический мутагенез, катагиз и перестройки генома [4].

#### **Происхождение дезаминаз, редактирующих молекулы РНК и ДНК**

Ранее считалось, что AID/АРОВЕС семейство дезаминаз возникает в эволюции в момент формирования адаптивной иммунной системы у позвоночных [5]. Наиболее древним охарактеризованным ферментом этой группы до последнего времени считались две цитозиновые дезаминазы у бесчерепных миног и миксин, которые осуществляют соматические перестройки генов переменных лимфоцитарных рецепторов (VLR) [6]. Однако по мере накопления данных стало ясно, что это не так.

Гены семейства AID/АРОВЕС дезаминаз в настоящее время обнаружены у различных представителей метазой, диктиостелиума и водорослей. Среди метазой гены семейства выявлены у оболочников, иглокожих, членистоногих, моллюсков, брахиопод, кольчатых червей, книдарий, губок [7].

Каково же происхождение генов AID/АРОВЕС семейства? Бактерии обладают

большим и разнообразным семейством дезаминаз, выполняющих функцию токсинов, вовлеченных в межвидовую конкуренцию [8]. Токсины представляют собой мультидоменные белки. Модули очень вариабельны и обычно содержат нуклеазные домены различных семейств, которые разрезают ДНК или РНК молекулы в клетке мишени [8]. Содержат токсины и дезаминазные домены. Именно дезаминазные домены бактериальных токсинов согласно современным данным и явились предшественниками генов AID/АРОВЕС семейства и некоторых других редактирующих нуклеиновые кислоты семейств [8, 9].

Например, у различных позвоночных и губок обнаружен новый класс дезаминаз, которые содержат N-концевые сигнальные пептиды [7]. Наличие пептида позволяет предположить, что дезаминазы этого семейства секретируются и могут выполнять какую-то функцию в эндоплазматической сети или вне клетки. Они могут вызывать мутагенез внеклеточных паразитов или даже патогенных клеток, например, трансмиссивных опухолей [10]. Дезаминазы этого семейства были названы новые секретируемые AID/АРОВЕС дезаминазы (SNAD).

Геномы различных видов динофлагеллат содержит от 1 до 39 генов DYW семейства редактирующих молекулы РНК дезаминаз. Предполагается, что дезаминазы динофлагеллат участвуют в процессах редактирования РНК в хлоропластах. Однако не исключено, что они участвуют и в защите от вирусов и транспозонов.

Таким образом, существует несколько семейств эукариотических дезаминаз, которые, по всей вероятности, произошли от дезаминазных доменов бактериальных токсинов. Латеральный перенос объясняет внезапное эволюционное появление и неравномерное распределение семейств дезаминаз среди эукариот.

Например, АРОВЕС4 подсемейство дезаминаз выявлено у четвероногих, лопастеперых рыб (sarcopterygians) и бесчелюстных, но часто отсутствует у лучеперых (actinopterygians). При этом АРОВЕС4 обнаружен у книдарий и у некоторых водорослей. Такой паттерн распределения генов позволяет предположить, что АРОВЕС4 подсемейство возникло на ранних стадиях развития метазой, а затем распространилось посредством латерального переноса генов.

О том, что у эукариот AID/АРОВЕС дезаминазы редактируют молекулы ДНК и РНК свидетельствуют следующие данные.

У представителей иглокожих Strongylocentrotus purpuratus и брахиопод Lingula anatine обнаружены гены, кодирующие белки, сходные по последовательности и активности с AID/АРОВЕС дезаминазами позвоночных [5]. Экспрессия этих генов повышена в коеломотах, которые постоянно и непосредственно взаимодействуют с микробами. Повидимому, AID/АРОВЕС дезаминазы редактируют гены или транскрипты, участвующие в иммунном ответе. Однако это пока точно не доказано.

У улиток Biomphalaria лектины плазмы или так называемые фибриногенподобные белки (FREPs) в

соматических клетках диверсифицируются. Так же как и Ig FREP белки узнают антигенные вариации паразитов улиток, таких, например, как трематод. Гены, кодирующие FREP белки, соматически диверсифицируются с помощью соматического гипермутагенеза и генной конверсии [11]. Брюхоногие, в том числе *Biomphalaria* и *Aplysia californica*, содержат гены AID/APOBEC дезаминаз [7]. Предполагается, что дезаминазы моллюсков являются ключевыми ферментами механизма соматического мутагенеза FREP белков.

Таким образом, по всей вероятности, зависимые от дезаминаз механизмы соматического гипермутагенеза и перестроек ДНК возникали в течение эволюции многоклеточных независимо и множество раз.

#### **Генетически запрограммированные мишени AID дезаминазы**

Как было отмечено выше, генетически запрограммированными мишенями для AID дезаминазы являются Ig локусы. Какие же условия необходимы и достаточны для локус-специфического привлечения AID и AID зависимого мутагенеза и рекомбинации переключения?

Предпочтительной последовательностью для AID дезаминазы является RGYW/WRCY (где R=пурин, Y=пиримидин и W=A или T), которая встречается приблизительно с частотой одна последовательность на 36 пары оснований в геноме человека. Очевидно, что наличие таких последовательностей в Ig локусах не является достаточным условием для специфического присоединения и действия AID дезаминазы.

Хорошо известно, что для соматического гипермутагенеза и рекомбинации переключения необходимы энхансеры и связывающиеся с ними транскрипционные факторы [12, 13]. В настоящее время с использованием очень чувствительных методов удалось еще раз подтвердить тот факт, что энхансеры и энхансероподобные элементы локализует соматический гипермутагенез [14]. Энхансеры кооперируются для регуляции и локализации соматического гипермутагенеза и для проявления их активности нужны сайты связывания с транскрипционными факторами. Предположили, что энхансеры локализует соматический гипермутагенез с помощью привлечения AID дезаминазы, сборки мутагеназного комплекса и/или превращения Ig локуса в подходящий субстрат для активности AID [14]. Однако точно роль регуляторных элементов и транскрипционных факторов в организации гипермутагенеза и рекомбинации переключения не установлена.

ZMYND8 транскрипционный фактор необходим для гипермутагенеза кодирующего гены тяжелой цепи иммуноглобулинов (IgH) локуса, а также транскрипции S районов и физиологического уровня рекомбинации переключения [15]. ZMYND8 связывается с B клеточными суперэнхансерами, в том числе с расположенным на 3' конце IgH локуса 3'RR суперэнхансером. При связывании он снижает

транскрипцию и активность самого 3'RR энхансера, но при этом повышается как уровень гипермутагенеза, так и рекомбинации переключения.

3'RR суперэнхансер регулирует рекомбинацию переключения, включая синтез первичных герминальных транскриптов (GLT) донорного и акцепторного S районов, паузу РНК II полимеразного комплекса, присоединение AID и образование DSB [13, 14]. В то же время энхансер не влияет на репарацию образовавшихся DSB разрывов.

Согласно некоторым экспериментальным данным AID непосредственно может связываться с транскрипционными факторами. Так AID присоединяется к HLH домену E2A транскрипционного фактора и PD домену PAX5 фактора [16]. Образование комплекса между транскрипционными факторами и AID регулируется сигнальными системами. В свою очередь комплексы контролируют гипермутагенез и рекомбинацию переключения. Другими словами такие комплексы направляют процессы мутагенеза и рекомбинации.

Помимо регуляторных элементов и связывающихся с ними транскрипционных факторов для соматической изменчивости нужен определенный паттерн эпигенетических модификаций Ig локусов. Как донорные, так и акцепторные S регионы специфически обогащены такими эпигенетическими модификациями гистонов как H3K9/K14ac, H3K27ac, H4K8ac и H3K4me3, которые контролируют рекомбинацию переключения [17]. Ig локусы обогащены эпигенетическими маркерами активных энхансеров, элонгации транскрипции, а также активных промоторов в B клетках, но не в фибробластах, в которых они не подвергаются соматической изменчивости [18].

В свою очередь Ig локусы приобретают паттерны эпигенетических модификаций в результате активности суперэнхансеров и взаимодействующих с ними транскрипционных факторов [13].

По всей вероятности, регуляторные элементы и устанавливаемые ими эпигенетические модификации локусов служат для настройки особенного режима транскрипции локуса, который позволяет сформироваться мутагеназному комплексу, а, возможно, регулируют и активность мутагеназного комплекса.

Какие же особенности режима транскрипции Ig локусов делают их мишенями для гипермутагенеза и рекомбинации? Для действия AID нужна активная транскрипция локуса, в результате которой образуется субстрат для дезаминазы - одноцепочечная ДНК. Однако транскрипция сама по себе не может объяснить тот факт, что запрограммированными мишенями для дезаминаз являются Ig гены, поскольку большинство транскрибируемых генов не являются мишенями для AID.

Вторым хорошо установленным фактором, необходимым для привлечения AID к Ig локусам, является кофактор транскрипции Spt5 [19]. Spt5 белок в составе DSIF комплекса присоединяется к РНК II полимеразному комплексу и подавляет элонгацию транскрипции *in vitro* [20]. *In vivo* показана строгая

корреляция между районами, с которыми связывается Spt5, и районами, в которых РНК полимеразы II делает паузу [21]. В активированных В клетках в Ig локусах плотно упакованные Spt5 и Pol II молекулы занимают несколько тысяч пар нуклеотидов. В свою очередь Spt5 связывается с AID и таким образом присоединяет дезаминазу к остановившемуся РНК II полимеразному комплексу [19]. Spt5 локализует привлечение AID дезаминазы в районы начала транскрипции (transcription start sites) (TSS).

Однако сайты связывания Spt5 с AID концентрируются поблизости от TSS районов не только Ig локусов, но и ряда других генов [22]. Spt5 связывается в В клетках с ~60% экспрессируемых генов [19]. И пауза транскрипции характерна для более чем 30% транскрибируемых генов [23]. Поэтому Spt5 белок сам по себе не может быть решающим фактором, способствующим присоединению к локусу дезаминазы.

Еще один белок - связывающийся с одноцепочечной ДНК репликационный белок А (RPA) – также нужен для при влечения AID к Ig локусам [22, 24]. RPA белок играет важную роль в репликации, рекомбинации и репарации. В течение этих процессов, а также транскрипции белок стабилизирует одноцепочечные нити ДНК. Но возникающих при различных процессах одноцепочечных нитей ДНК много, и каким образом фактор способствует специфическому привлечению дезаминазы к Ig генам также непонятно.

Помимо Spt5 фактора к остановившемуся РНК полимеразному комплексу присоединяются экзосомы. Экзосомы нужны для физиологического уровня гипермутационного и рекомбинации переключения Ig генов. Субстратом для экзосом часто являются различные виды некодирующей РНК.

Для образования DSB в S районах переключения разрывы должны образоваться как на матричной, так и нематричной цепях ДНК. Но субстратом для AID дезаминазы является только одноцепочечная ДНК. Поэтому матричная цепь ДНК недоступна для AID. При рекомбинации переключения S районы транскрибируются с образованием GLT транскриптов. GLT транскрипты остаются связанными с ДНК, образуя ДНК-РНК гибриды, которые называются R-петли. Экзосомы за счет 3'-5' экзорибонуклеолитической активности деградируют гибридирующиеся с ДНК молекулы некодирующей РНК, делая тем самым матричную цепь доступной для образования разрывов ДНК и давая возможность дезаминировать цитозин на обеих цепях ДНК.

Помимо некодирующих GLT транскриптов субстратом для экзосом являются молекулы xTSS-RNA [25], транскрипция которых стартует вверх по течению в среднем на расстоянии 110 п.н. от TSS сайтов мРНК, т.е. дивергентно по отношению к транскрипции мРНК. Следовательно, транскрибируются xTSS-RNA в антисмысловом направлении по отношению к транскрипции гена [26]. Размер xTSS-RNA достигает 600 п.н.

При дивергентной транскрипции позади РНК полимеразных комплексов возникает отрицательная суперспирализация ДНК, способствующая формированию протяженных ДНК-РНК гибридов или R-петель. К некодирующим молекулам РНК в R-петлях присоединяются экзосомы, разрушающие R-петли и освобождающие одноцепочечную ДНК, служащую субстратом для AID дезаминазы.

В В лимфоцитах xTSS-RNA транскрибируется с Ig генов и играет важную роль в привлечении AID дезаминазы к Ig локусам [25]. Но поскольку на 80% активных промоторов млекопитающих РНК синтезируется дивергентно, то дивергентная транскрипция не является решающим фактором локус-специфичности гипермутационного и рекомбинации переключения [26]. Помимо этого, дивергентно транскрибируется РНК на энхансерах. В результате транскрипции энхансеров образуется eРНК.

Обсуждаемые экспериментальные данные имеют важное биологическое значение. Они доказывают функциональную значимость синтезируемой на энхансерах и промоторах РНК, поскольку РНК служит для локального привлечения ферментов и факторов к ДНК. Связавшиеся с синтезируемой на регуляторных элементах комплексы могут как осуществлять транскрипцию генов, так и при некоторых условиях изменять ДНК. Juan Chen и соавторы считают, что : такая регуляторная парадигма может быть применимой к различным биологическим процессам [27]. Это положение очень важно для выявления механизмов регуляции мутагенеза.

Хорошо известно, что молекулы некодирующей РНК могут действовать ин транс. У цилиат комплементарные взаимодействия между РНК, синтезирующейся в старом макроядре, и ДНК нового макроядра локализуют перестройки генома формирующегося макроядра.

Оказалось, что GLT транскрипт может действовать не только ин цис, как было описано выше, но и ин транс [28]. GLT транскрипты содержат как S районы, так и экзоны. S районы, являющиеся интронами, вырезаются в результате сплайсинга и образуют РНК, имеющую форму лассо. Затем лассообразные молекулы РНК с помощью специального фермента превращаются в линейные формы. AID дезаминазы связываются с линейными формами, узнавая G-квартетные структуры которыми обогащены молекулы сплайсированной интронной РНК. Комплекс AID-РНК комплементарно связывается с S районом, локально присоединяя к S району AID дезаминазу [28].

Процессированный GLT транскрипт действительно может обеспечить локус-специфическое привлечение AID. Однако нельзя исключить способность молекул длинных некодирующих РНК, синтезируемых на других локусах и обогащенных остатками гуанина, при определенных условиях выполнять такую же функцию локального привлечения AID дезаминазы.

Белок ROD1 является еще одним фактором, необходимым для рекомбинации переключения и соматического гипермутационного Ig локуса. У мышей,

нокаутированных по этому гену, значительно снижаются уровни гипермутации и рекомбинации переключения [27].

ROD1 является членом семейства РНК связывающих белков, которые узнают и взаимодействуют с высокой аффинностью с полипиримидиновыми последовательностями. ROD1 связывается с синтезирующимися транскриптами в области локусов, транскрибирующихся в двух направлениях, и рекрутирует к этим мишеням AID. Но, как уже отмечалось выше, дивергентная транскрипция характерна для многих промоторов и энхансеров. Поэтому специфичность гипермутации этот фактор не объясняет.

ROD1 селективно связывается с полипиримидиновыми последовательностями РНК, что могло бы быть фактором селективного привлечения дезаминазы. Однако транскрипция S районов дает обогащенные гуанином молекулы РНК. Транскрипция переменных районов Ig генов также не дает обогащенные пиримидинами молекулы РНК.

Таким образом, для локализации соматического мутагенеза и рекомбинации Ig локусов нужен регуляторный аппарат, состоящий из цис-элементов ДНК и взаимодействующих с ними транскрипционных факторов. Какую функцию может выполнять регуляторный аппарат? Во-первых, он способствует приобретению локусами эпигенетических модификаций, необходимых для мутагенеза и рекомбинации [13].

Во-вторых, совместно регуляторный аппарат и эпигенетические модификации контролируют инициацию транскрипции [29]. Однако, исходя из современных данных, вполне реалистичным является предположение, согласно которому два типа регуляции кооперируются не только для инициации транскрипции, но и для того, чтобы настроить режим активности Ig локусов, позволяющий привлечь AID, собрать мутагеназный или рекомбиназный комплексы и настроить работу комплексов в режиме образования мутаций или рекомбинации [13, 14].

Для подтверждения этой модели нужны дополнительные исследования взаимосвязей между компонентами регуляторного аппарата и эпигенетическими модификациями с одной стороны и компонентами режима транскрипции и редактирования Ig локусов с другой.

По-прежнему остается не ясным, каким образом режим транскрипции Ig локусов делает их специфическими мишенями для соматического гипермутации и рекомбинации. Из выявленного на данный момент множества условий и факторов транскрипции, необходимых для привлечения AID дезаминазы, ни один не является специфическим для Ig локусов.

Поэтому можно предположить, что особенное сочетание общих для многих локусов факторов транскрипции делает Ig локусы мишенями для AID зависимого мутагенеза и рекомбинации. В свою очередь создавать особое сочетание качественных и количественных параметров режима транскрипции

могут регуляторные элементы и эпигенетические модификации локусов.

### *Незапрограммированные генетически мишени дезаминаз AID/APOBEC семейства*

Помимо гипермутации Ig генов, которые являются генетически запрограммированными мишенями действия AID дезаминазы, в клетках, особенно раковых, существует множество других незапрограммированных мишеней соматического мутагенеза и транслокаций [30, 31]. В то же время скорость мутагенеза незапрограммированных мишеней на несколько порядков ниже, чем скорость мутагенеза Ig локуса [14, 31].

О том, что гены являются мишенью для AID дезаминазы, судят по их связыванию с AID дезаминазой. Однако связывание сайта с AID автоматически не означает, что в сайте образуются повреждения ДНК, двойные разрывы ДНК и мутации. В В клетках AID связывается с 5910 генами [22]. Но только в приблизительно в 300 возникают AID индуцируемые повреждения ДНК [32, 33].

В свою очередь образование двойных разрывов не означает, что locus является мишенью для мутаций, поскольку образование двойных разрывов может быть связано с физиологическими процессами перепрограммирования режима активности локуса, например, с процессом деметилирования цитозинов [34].

Согласно последним опубликованным в доступной для нас литературе данным удалось выявить 291 регион генома, содержащих 275 генов, в которых рекуррентно образуются AID зависимые мутации в Ung<sup>-/-</sup>Msh2<sup>-/-</sup> В клетках герминальных центров мышей [35]. Однако в Ung<sup>-/-</sup>Msh2<sup>-/-</sup> клетках отсутствуют ключевые ферменты систем репарирующих дезоксиридин. Поэтому в нормальных условиях часть этих генов могут не быть мишенями для зависимого мутагенеза.

Картирование двойных разрывов, соматических мутаций и хромосомных перестроек позволяют выделить среди незапрограммированных мишеней AID дезаминазы два типа генов. К первому типу относятся протоонкогены кодирующие белки пролиферации и апоптоза такие как Muc, Pim1, Jund, Bcl2, а ко второй группе относятся гены дифференцировки и активации В клеток: Pax5, Cd79b, Aicda, Irf8, Bach2, Nfkb [4].

Незапрограммированный соматический мутагенез и транслокации затрагивают не только гены, но и регуляторные элементы генома [36]. Из 236 мишеней AID зависимых DSB разрывов в активированных В лимфоцитах мышей 33 локализованы в дистальных энхансерах [33]. Потенциально в этих энхансерах могут возникать транслокации и/или мутагенеза. Однако, как будет показано ниже, нельзя исключить, что DSB разрывы имеют физиологическое значение.

Активность AID дезаминазы является основным фактором мутагенеза при множественной миеломе (ММ). Секвенирование геномов 765 пациентов с ММ позволило обнаружить регуляторные

регионы, мутации которых значительно изменяют экспрессию генов [36]. Среди них вариации количества копий цис-регуляторных элементов гена MYC и одиночные замены в энхансере гена PAX5. В большинстве случаев паттерн изменений нуклеотидных последовательностей свидетельствует об участии AID дезаминазы в формировании мутаций.

Очень интересным является сделанное в ходе исследования наблюдение, согласно которому многие гены, входящие в повреждаемые метаболические пути, приобретают мутации как в кодирующей области, так и в области регуляторных цис-элементов [36].

В результате активности APOBEC дезаминаз могут возникать новые энхансеры. Рекуррентная C-T замена в некодирующей последовательности, расположенной на расстоянии 4 т.п.н. вниз по течению от сайта начала транскрипции LMO1 онкогена в первичных T клеточных острых лимфобластных лейкозах, образуется благодаря активности дезаминаз APOBEC семейства [37]. Мутация приводит к появлению нового сайта связывания с транскрипционным фактором MYB и к аномально высокой моноаллельной экспрессии онкогена LMO1.

Активность AID/APOBEC дезаминаз приводит к образованию в незапрограммированных мишенях не только одиночных мутаций и транслокаций, но и вызывает катаегиз.

В большинстве проанализированных геномах раковых клеток обнаружены кластеры множественных мутаций размером менее 10 т.п.н. [38], которые возникают в результате катаегиза. Катаегиз часто сопряжен с геномными перестройками, в том числе с локализованным хромотрипсисом, при котором возникает множество DSB. ДНК рвется на фрагменты, соединяющиеся в новом порядке [39]. При хромотрипсисе происходит одномоментное образование от нескольких десятков до нескольких сотен двойных разрывов ДНК на определенных хромосомах.

На основании ряда наблюдений предположили, что механизмом катаегиза является дезаминирование цитозинов на одноцепочечной ДНК, появляющейся при репарации двойных разрывов [4]. В не-B-клеточных опухолях дезаминирование цитозина, осуществляется в основном APOBEC3A и APOBEC3B дезаминазами, а в B клеточных лимфомах AID дезаминазой.

Механизм катаегиза отличается от механизма соматического гипермутации. [4]. Предполагается, что катаегиз в лимфомах состоит из двух стадий. На первой стадии AID дезаминирует одноцепочечную ДНК, которая появляется при транскрипции незапрограммированных мишеней в G1 фазе клеточного цикла. Дальнейшая обработка повреждений системами репарации BER и MMR приводит к образованию DSB, которые в S или G2M фазах клеточного цикла лигируются с помощью гомологичной рекомбинации. В течение гомологичной рекомбинации или после нее AID может дезаминировать образующуюся в ходе рекомбинации одноцепочечную ДНК, тем самым вызывая

образование кластера мутаций. Второй этап катаегиза не зависит от транскрипции.

Каким же образом неиммноглобулиновые гены становятся генетически незапрограммированными мишенями для AID дезаминазы? Ответ на этот вопрос может решить фундаментальную проблему биологии о направленности мутаций.

### **Локализация незапрограммированных мишеней дезаминаз AID/APOBEC семейства эпигенетическими модификациями**

Незапрограммированные мишени AID являются активно транскрибируемыми генами, в которых РНК II полимеразы делает паузу, присоединяются Spt5 фактор и экзосомы, происходит дивергентная транскрипция и синтез некодирующей РНК [19, 25, 32, 40]. Однако не одно из этих условий, как было показано выше необходимых и для мутагенеза Ig генов, само по себе недостаточно для локализации горячих точек транслокаций. Например, в B лимфоцитах и эмбриональных фибробластах мышей (MEF) некоторые гены одинаково транскрибируются, но являются мишенью для AID только в одном типе клеток [18].

Изучение паттерна вызываемых AID дезаминазой транслокаций в B лимфоцитах и MEF клетках показало, что горячие точки транслокаций в двух типах клеток практически не совпадают [18]. Незапрограммированные мишени AID дезаминазы в B клетках мышей и человека также совпадают менее чем 50% [33]. Следовательно, клеточный контекст определяет паттерн перестроек генома, вызываемых AID дезаминазой. Одним из факторов, который может определять клеточную специфичность мишеней, является паттерн эпигенетических модификаций локусов в данном типе клеток. Вторым - специфическое для клеток сочетание активных цис-регуляторных элементов и связывающихся с ними транскрипционных факторов.

Мутации при хронической лимфоцитарной лейкемии локализуются в промоторах и энхансерах [41]. В регуляторных элементах с увеличенным содержанием эпигенетического маркера гистонов H3K27Ac наблюдается трехкратное повышение количества мутаций, образующихся за счет AID зависимого соматического гипермутации. Эти наблюдения позволили предположить, что H3K27Ac является элементом эпигенетического кода, способствующего привлечению к локусу AID дезаминазы и проявлению ее активности.

Горячие локусы транслокаций в MEF фибробластах и B лимфоцитах обогащены эпигенетическими маркерами активных энхансеров (H3K27Ac, H3K4me1 и Med12) только в том типе клеток, в котором они являются мишенью для AID дезаминазы [18]. Такая же корреляция обнаружена для маркеров элонгации транскрипции (H3K36me3 и H3K79me2), а также маркеров активных промоторов (зарождающаяся РНК, H3K4me3 и H3K9Ac).

В то же время сам по себе ни один из этих маркеров не определяет специфические для типа

клеток горячие локусы транслокаций AID дезаминазы. Только комбинация трех эпигенетических маркеров, один из которых H3K27Ac характерен для активного энхансера, другой H3K36me3 – для элонгации транскрипции, и третий, зарождающаяся РНК, – для активного промотора, локализует мишени транслокаций в двух типах клеток. Таким же паттерном эпигенетических модификаций обладает Ig locus, кодирующий тяжелые цепи иммуноглобулинов IgH, в В клетках, но не в MEF клетках.

В то же время вполне сопоставимый уровень эпигенетических модификаций в IgH локусе и незапрограммированных мишенях ставит вопрос, почему мутации и рекомбинация переключения в IgH локусе происходят гарантированно и с большей скоростью. Помимо этого, некоторые гены обладают необходимым паттерном эпигенетических модификаций, тем не менее, данные о том, что они являются мишенью для вызываемых AID дезаминазой транслокаций отсутствуют [18]. Таким образом, выявленный паттерн эпигенетических модификаций не является достаточным условием для локус-специфического проявления редактирующей активности AID дезаминазы.

Не только эпигенетические модификации гистонов формируют мишени для AID дезаминазы. Метилированные цитозины ДНК непосредственно могут становиться мишенями вызываемого AID дезаминазой мутагенеза и направлять генетические изменения в неиммуноглобулиновых генах.

Субстратом для AID дезаминазы является не только цитозин, но и метилированный цитозин в CpG динуклеотидах. AID дезаминазы деметируют ДНК с помощью дезаминирования метилцитозина с образованием тимина и последующего привлечения системы репарации неправильно спаренных оснований.

При созревании В клеток в герминальных центрах локус-специфически деметируется ДНК и увеличивается разнообразие паттернов метилирования [34]. Оба процесса отменяются у животных с делецией гена AID. Причем дифференциально метилированными цитозинами в В клетках герминальных центров и в неактивированных В клетках обогащены гены, являющиеся мишенями для незапрограммированного соматического гипермутагенеза, и эти гены формируют сети, необходимые для развития и пролиферации В клеток.

Другими словами, гены, метилирование которых снижается и становится разнообразным за счет активности AID, являются мишенями для незапрограммированного соматического мутагенеза. Это наблюдение позволяет предположить, что сбои в процессе физиологического деметилирования цитозина могут приводить к AID зависимому патологическому мутагенезу не-Ig генов.

Вызываемый AID-дезаминазой мутагенез метилированных цитозинных часто встречается в опухолях. Спектр мутаций при фолликулярной лимфоме указывает на их образование в результате деметилирующей активности AID дезаминазы [42].

Такой же паттерн соматического мутагенеза обнаружен и в геномах многих других опухолей [43].

Практически у всех больных с EGFR-зависимым (рецептор эпидермального фактора роста) раком легких, которых лечат тирозинкиназным ингибитором EGFR рецептора TKI, развивается устойчивость за счет точечной мутации EGFR гена. Является ли мутация индуцируемой TKI, или отбирается из предшествующих обработке ингибитором клонов? Оказалось, что TKI активирует NFκB сигнальный путь, который индуцирует экспрессию AID дезаминазы. AID дезаминирует 5-метилцитозин с образованием тимина, и в результате появляется необходимая для формирования устойчивости мутация [44]. Подавление NFκB сигнального пути снижает частоту, а нокаутирование гена AID дезаминазы полностью предотвращает возникновение мутации.

В примордиальных герминальных клетках (PGC) на ранних стадиях эмбриогенеза происходит эпигенетическое перепрограммирование, включающее в себя деметилирование ДНК. В дефицитных по AID PGC клетках уровень метилирования в три раза выше, чем в клетках дикого типа [45]. Эти данные свидетельствуют об участии AID дезаминазы в деметилировании генома PGC клеток, а, следовательно, можно допустить существование AID зависимого мутагенеза в половых клетках.

Итак, экспериментальные данные дают основания для предположения о существовании эпигенетического кода, т.е. системы команд, регулирующих присоединение AID дезаминазы и ее дальнейшую активность в локусах-мишенях. С помощью кода накопленная в локусе эпигенетическая информация может транслироваться в мутации и перестройки ДНК.

### ***Роль суперэнхансеров в образовании мишеней соматической изменчивости***

Как и при гипермутагенезе и рекомбинации переключения Ig генов регуляторные элементы играют ключевую роль в образовании AID/APOBEC-зависимых DSB разрывов, мутагенеза и перестроек в генетически незапрограммированных мишенях.

Как обсуждалось выше, горячие локусы транслокаций обогащены эпигенетическими маркерами активных энхансеров и промоторов, что свидетельствует об участии регуляторных элементов в локализации мишеней AID зависимых DSB разрывов и транслокаций [18]. Ниже мы рассмотрим более подробно экспериментальные данные, подтверждающие это предположение.

AID дезаминаза активна в интерфазе клеточного цикла. Интерфазный хроматин разделяется на эухроматические А компартменты и гетерохроматические В компартменты. 96% всех мишеней AID зависимых DSB (233 из 236) в про-В клетках мышей локализованы в активно транскрибируемых А компартментах [33].

86% мишеней AID зависимых DSB в активированных мышьяных В лимфоцитах располагаются в проксимальных к промоторам

последовательностях и энхансерах, связанных друг с другом с образованием регуляторных кластеров [33]. В некоторых случаях в регуляторных кластерах находятся несколько AID мишеней. Например, в Pax5 локусе проксимальная к промотору последовательность, являющаяся мишенью для AID дезаминазы, связана с тремя энхансерами, один из которых располагается на расстоянии 250 т.п.н. и также является мишенью для AID. Промоторы Lyba, Lybe и Rohema генов формируют регуляторный кластер размером 100 т.п.н. и являются мишенями для AID.

A и B компартменты разделены на субдомены, которые называются топологически ассоциированные домены (TAD) [46]. Внутри TAD доменов часто формируются контакты между различными элементами. В то же время элементы TAD домена взаимодействуют с элементами других доменов гораздо реже, чем друг с другом внутри доменов. Размеры TAD в среднем 880 т.п.н. (от десятка т.п.н. до несколько миллионов п.н.). 84% мишеней AID дезаминазы в активированных B лимфоцитах мышей локализованы в кластерах, образованных регуляторными элементами, расположенными внутри одного и того же TAD домена [33].

Таким образом, мишенями для AID зависимых DSB являются преимущественно области проксимальные по отношению к промоторам и энхансерам, связанные друг с другом и формирующие регуляторные кластеры внутри одного TAD домена.

Поскольку кластеры регуляторных элементов часто образуют суперэнхансеры, возник вопрос, не являются ли суперэнхансеры предпочтительными мишенями для AID [33].

Суперэнхансеры это регуляторные элементы, состоящие из множества взаимодействующих энхансеров и промоторов, с которыми специфически связываются транскрипционные факторы. Энхансеры физически связываются друг с другом и с промоторами генов мишеней, активность которых они контролируют. Суперэнхансеры регулируют транскрипцию генов, определяющих тип клетки. Клетки млекопитающих обладают тысячами энхансеров. Но обычно содержат только несколько сотен суперэнхансеров.

В активированных B клетках мышей 174 мишени AID (76%) находятся в суперэнхансерных доменах. На всех трех локусах, кодирующих Ig гены, AID зависимые DSB образуются в суперэнхансерах, элементы которых связаны друг с другом в пространстве.

Всего в B клетках идентифицировано 1003 суперэнхансера. Из них 824 не являются мишенями для AID [33]. Следовательно, быть суперэнхансером не является достаточным условием для того, чтобы стать мишенью для AID. Суперэнхансеры, являющиеся мишенями для AID дезаминазы, отличаются от других суперэнхансеров большей доступностью, более высоким содержанием H3K27Ac, они крупнее по размеру, и с взаимодействующих с ними промоторов стартует активная транскрипция [33]. Кроме того, они больше вовлечены в пространственные

взаимодействия и сами активно транскрибируются с образованием ePHK. Например, из двух энхансеров, расположенных вверх по течению от гена Pax5, только на одном интенсивно транскрибируется ePHK, и именно в этом энхансере образуются DSB, хромосомные транслокации и соматические мутации.

Данные еще одного исследования подтверждают, что суперэнхансеры, являющиеся мишенями для AID зависимых транслокаций, обогащены H3K27Ac по сравнению с суперэнхансерами, не являющимися мишенями [47]. Более того, внутри самого суперэнхансера горячие точки транслокаций также обогащены H3K27Ac по сравнению с остальным суперэнхансерным доменом, что подтверждает наличие эпигенетического кода, локализуемого мишени образуемых AID повреждений и нестабильности генома.

Пространственно связанные элементы суперэнхансеров действуют синергически при привлечении AID дезаминазы и образовании DSB и мутаций в локусах-мишенях [33]. Регуляторные элементы и связывающиеся с ними транскрипционные факторы вносят каждый свой вклад в установлении режима активности суперэнхансеров и контролируемых суперэнхансерами генов, позволяющего привлечь AID дезаминазу и образовать DSB разрывы.

Суперэнхансеры локализуют не только образование AID зависимых DSB, но мутаций и транслокаций.

Горячая точка транслокации BCL6 гена [48] располагается внутри интрона в суперэнхансере, на котором синтезируется длинная некодирующая РНК (lncPHK). Границы lncPHK точно совпадают с границами зоны транслокации. Транскрибируется lncPHK в противоположном направлении по отношению к транскрипции гена BCL6 гена. То есть в локусе происходит дивергентная транскрипция, в ходе которой временно возникают протяженные районы одноцепочечной ДНК. AID активно дезаминирует цитозин в зоне транслокации.

47 из 51 незапрограммированных мишеней AID зависимых транслокаций в B клетках, активированных II-4, лежат внутри суперэнхансеров, которые перекрываются с телами белок-кодирующих генов [47].

В суперэнхансерах, являющихся мишенями для AID зависимых транслокаций, обнаружен более высокий уровень конвергентной транскрипции, чем в суперэнхансерах, также перекрывающихся с генами, но не являющихся мишенью для AID. При конвергентной транскрипции с суперэнхансера синтезируется антисмысловая РНК. Антисмысловый РНК является по отношению к РНК, синтезируемой с перекрывающихся с суперэнхансерами генов. На основании этих данных предполагают, что конвергентная транскрипция является очень важным фактором локализации транслокаций в незапрограммированных мишенях. Однако в Ig локусах конвергентная транскрипция не обнаружена.

Эктопическая экспрессия AID дезаминазы в MEF фибробластах индуцирует образование

хромосомных транслокаций преимущественно между геномными сайтами, которые регулируются активно транскрибируемыми суперэнхансерными доменами [33].

Значительная часть (57%–70%) мишеней мутагенеза в лимфоме Бэркитта и диффузной крупноклеточной В-клеточной лимфоме (DLBCL) располагается в доменах, содержащих суперэнхансеры и регуляторные кластеры [33]. В регуляторном кластере BCL7A гена в лимфоме Бэркитта соматическому гипермутагенезу подвергается как промотор, так и энхансер. В В лимфоцитах DLBCL лимфомы катаегиз обнаружен только в проксимальных к промотору последовательностях гена BCL6, в то же время среди 26 геномов других первичных лимфом человека катаегиз обнаружен также в лежащем вверх по течению энхансере.

Формирование активных суперэнхансеров тканеспецифично. В соответствии с этим более чем 50% суперэнхансерных доменов, являющихся мишенями для AID зависимого образования DSB, присутствуют только в стимулированных В клетках, но не в 18 других типах мышечных клеток [33]. Причем даже в про-В клетках мишени только на 32% совпадают с активированными В клетками.

Тканеспецифичность суперэнхансеров определяет и специфический паттерн мишеней AID зависимой соматической изменчивости. С помощью картирования индуцируемых AID транслокаций выявлено 29 и 43 горячих локусов в MEF фибробластах и В клетках соответственно. Но только три горячие точки оказались общими для двух типов клеток [33]. Генные локусы, в которых локализовались горячие точки, интенсивно транскрибируются и перекрываются с суперэнхансерами в том типе клеток, в которых происходила транслокация. Например, гены Flnb и Pim1 экспрессируются в двух типах клеток, но являются мишенями для транслокаций только в том типе клеток, в котором экспрессия гена регулируется суперэнхансером. Интересно, что транслокации в MEF фибробластах затрагивали те суперэнхансерные домены, которые контролируют гены пролиферации и дифференцировки фибробластов.

Следовательно, незапрограммированные мишени AID зависимого образования DSB и соматической изменчивости представляют собой не определенный набор генов. Мишени формируются временно и тканеспецифично в результате условий жизнедеятельности клетки, которые активируют определенный паттерн суперэнхансеров и связавшихся с ними транскрипционных факторов.

Таким образом, в В клетках мышей и человека запрограммированными и незапрограммированными мишенями для AID зависимых DSB, мутаций, катаегиза и транслокаций становятся интенсивно контролируемые локусы, в которых взаимодействующие друг с другом регуляторные элементы образуют активно транскрибируемые суперэнхансеры и регуляторные кластеры. Суперэнхансеры и регуляторные кластеры взаимодействуют с транскрипционными факторами и синергически устанавливают такой режим активности

локуса, который способствует привлечению AID, образованию DSB разрывов и появлению мутаций и перестроек ДНК, причем в ряде случаев сами регуляторные элементы становятся горячими точками генетических перестроек. Следовательно, регуляторные элементы могут контролировать не только изменчивость генов, но и свою собственную изменчивость. Однако механизм, посредством которого суперэнхансеры регулируют не только транскрипцию локуса, но и его генетическую изменчивость, по-прежнему остается загадкой.

### ***Возможная роль транскрипционных факторов в образовании AID/APOBEC зависимых физиологических DSB и соматическом мутагенезе.***

DSB разрывы не всегда являются патологическими, ведущими к образованию мутаций и перестроек ДНК. Экспериментально показано образование DSB при деметилировании генома, активации генов и регуляции транскрипции [49].

Регулируемое образование DSB разрывов топоизомеразы в некоторых промоторах стимулирует транскрипцию генов, контролирующей специфические для клеток функции [50, 51]. А каспаза 3, которая фрагментирует ДНК при апоптозе, способствует образованию DSB при дифференцировке различных клеток [52]. В мышечных клетках каспаза 3 стимулирует активируемую каспазой ДНКазу (CAD), которая образует двойные разрывы, в том числе промотора гена p21, увеличивая его экспрессию [53].

AID/APOBEC-зависимые DSB разрывы могут иметь физиологическое значение при перепрограммировании метилирования ДНК, активации энхансеров и транскрипции генов. Как рассматривалось выше, AID зависимое перепрограммирование метилирования цитозина при созревании В лимфоцитов является физиологическим и генетически запрограммированным процессом [34]. В то же время в ходе деметилирования образуются повреждения ДНК, которые, при пока неизвестных условиях, могут становиться патологическими, превращаясь в мутации и перестройки генома.

Известно по крайней мере три механизма, с помощью которых эстрогеновый рецептор (ER) вызывает образования DSB в местах связывания с ДНК и посредством разрывов активирует транскрипцию.

При транскрипции энхансеров возникает торсионный стресс, который затрудняет элонгацию транскрипции и подавляет инициацию транскрипции. Снять стресс могут топоизомеразы, формирующие разрывы ДНК.

Действительно, при индуцируемой ER и андрогеновым (AR) рецепторами транскрипции под действием топоизомераз образуются временные разрывы ДНК. В результате торсионное напряжение снимается, хроматин эпигенетически модифицируется и транскрипция активируется [50, 54, 55].

Привлечение к локусу и активация гистоновой деметилазы LSD1 ER рецепторами приводит к деметилированию H3 гистона по лизинового остатку в 9 положении как в энхансерах, так и промоторах. При деметилировании гистонов образуется перекись

водорода, которая локально модифицирует ДНК с образованием 8-оксогуанина. К 8-оксогуанину присоединяется 8-оксогуанин-ДНК гликозидаза и факторы BER системы репарации, вызывающие образование DSB. В течение репарации двойных разрывов хроматин эпигенетически модифицируется, что приводит к активации транскрипции генов [55].

Наконец был открыт третий механизм активации транскрипции регулируемых ER рецепторами генов, согласно которому DSB разрывы образуются в результате действия APOBEC3B дезаминазы [56]. ER связываются с цис-регуляторными элементами и привлекают к местам присоединения APOBEC3B дезаминазу, которая превращает цитозин в урацил, тем самым активируя BER систему репарации ДНК. В ходе репарации образуются двойные разрывы ДНК, которые в свою очередь репарируются NHEJ системой. Репарация сопровождается эпигенетической перестройкой хроматина, что активирует экспрессию генов в клетках рака груди.

Однако описанные процессы могут привести к мутагенезу локусов, направляемого ER рецептором и осуществляемого APOBEC3B дезаминазой

Рак молочной железы вызывается постоянно повышенным уровнем эстрогенов в крови [57, 58]. Поэтому можно предположить, связь между повышенным уровнем гормона и мутациями, вызывающими опухоль.

С помощью секвенирования 560 геномов рака молочной железы и сопоставления с картой сайтов связывания с ER рецептором (ERBS) обнаружено обогащение соматическими мутациями ERBS сайтов, в том числе и рекуррентными мутациями [59]. Как интенсивность связывания регуляторного элемента с ER рецептором, так и частота связывания коррелируют с повышением скорости соматических мутаций. Наиболее часто используемые цис-элементы подвергаются мутагенезу с большей скоростью, чем менее используемые элементы. Скорость мутагенеза регионов генома, с которыми ER рецептор связывается конститутивно, также достоверно повышена. Паттерн мутаций свидетельствует о действии APOBEC3B дезаминазы.

Другие транскрипционные факторы также увеличивают мутагенез в специфически связывающихся с ними сайтах, но в меньшей степени [59].

Подавляющее большинство мутирующих сайтов (98%) находятся в некодирующих регуляторных элементах. Причем регуляторные элементы, подвергающиеся соматическому мутагенезу, чаще образуют контакты с другими регуляторными элементами, а гены, ассоциированные с мутирующими ERBS в рамках одного TAD домена, активно экспрессируются. Среди ERBS, мутирующих с наибольшей скоростью, только три перекрываются с кодирующими районами генов, являющихся драйверами опухолевого процесса - FOXA1, CBFB и CDH1.

Мутации ERBS изменяют экспрессию ряда дистальных генов посредством изменения связывания

с ними других, не являющихся ER, транскрипционных факторов и влияния на пространственную структуру ДНК. Например, рекуррентно мутирующий ERBS в LRRC3C-GSDMA локусе изменяет экспрессию множества дистальных генов. А рекуррентная мутация промотора гена ZNF143 снижает связывание с транскрипционным фактором ZBTB7A, увеличивая формирование хроматиновых петель и повышая экспрессию трех других дистально расположенных генов.

Еще один рецептор стероидных гормонов AR рецептор способствует образованию DSB разрывов, в результате которых возникают транслокации в клетках рака простаты [60].

Генотоксический стресс, вызываемый радиацией, и дигидротестостерон индуцируют транслокации в клеточной линии рака простаты [60]. Связывание расположенных в интронах генов TMPRSS2, ERG и ETV1 регуляторных элементов с AR рецептором вначале способствует интра- или интерхромосомному соединению вступающих в транслокацию локусов. AR локально изменяет структуру хроматина таким образом, что локус становится доступен для двух ферментов: AID дезаминазы и кодируемой LINE-1 элементами ORF2 эндонуклеазы. Ферменты синергически образуют DSB в местах транслокации, которые затем лигируются системой репарации NHEJ. Более того, сайты интронов, связанные с AR рецептором, становятся мишенями для соматического гипермутагенеза.

Таким образом, транскрипционные факторы, связываясь с регуляторными сайтами энхансеров, суперэнхансеров и промоторов, могут привлекать AID/APOBEC3B дезаминазы, образующие DSB разрывы. В свою очередь повреждения ДНК могут выполнять физиологическую роль в активации транскрипции локусов, а могут при неизвестных дополнительных условиях превращаться в мутации и транслокации.

Взаимодействие транскрипционных факторов с цис-регуляторными элементами ДНК может направлять процесс мутагенеза, локально привлекая не только AID/APOBEC3B дезаминазы, но и ORF2 эндонуклеазу. Более того, транскрипционные факторы могут увеличивать чувствительность сайта к мутагенам.

При раке кожи, вызываемом ультрафиолетовым облучением, образуются горячие точки мутагенеза. Они расположены в некоторых промоторах, которые мутируют у многих больных меланомой пациентов. Оказалось, что параллельные мутации в промоторах это не только результат отбора, но и направленности самого процесса мутагенеза [61]. Повышение уровня мутагенеза в области промоторов возникает вследствие связывания с цис-элементами ДНК транскрипционных факторов, в основном ETS транскрипционного фактора [62]. ETS в значительной мере формирует горячие точки образования пиримидиновых димеров. Расположение горячих точек коррелирует с рекуррентными мутациями, несмотря на интенсивную репарацию повреждений.

Но основании этих данных, можно предположить, что транскрипционные факторы,

соединяясь с регуляторными элементами, с помощью привлечения AID/APOBEC дезаминаз направляют процессы соматической изменчивости как самих регуляторных элементов, так и генов. Более того, транскрипционные факторы, связываясь с регуляторными элементами делают их горячими точками не только для AID/APOBEC зависимого мутагенеза, но и для других механизмов геномной нестабильности.

Похоже, мы сталкиваемся на данных примерах с закономерным явлением, которое заключается в том, что регуляторные элементы и транскрипционные факторы контролируют не только режим транскрипции локуса, но и их генетическую изменчивость, в том числе и собственную изменчивость.

### Выводы

Насколько же случайны AID/APOBEC зависимые мутации и перестройки генома?

Во-первых, AID/APOBEC зависимые мутации и перестройки ДНК не являются случайными в пространстве: они локализованы в локусах-мишенях генома. Организация локуса, эпигенетические модификации, режим транскрипции делают его мишенью для AID/APOBEC мутаций и перестроек генома. Тканеспецифическое формирование в локусе активных суперэнхансеров в значительной мере способствует его превращению в локус-мишень для AID/APOBEC.

Во-вторых, AID/APOBEC зависимая нестабильность генома не случайна во времени. В Ig локусах гипермутагенез и рекомбинация переключения активируются в В клетках при встрече с антигеном и получении всех необходимых дополнительных сигналов от других клеток иммунной системы.

Геномная нестабильность незапрограммированных локусов-мишеней также может быть локализована во времени активацией или сверхактивацией механизма мутагенеза.

Например, ингибитор тирозинкиназы активирует сигнальную систему, которая индуцирует экспрессию AID, что означает включение механизма геномной нестабильности в клетках рака легких [44]. В результате в локусе-мишени рекуррентно возникает мутация, формирующая устойчивость к ингибитору, т.е. адаптивная мутация.

Увеличение уровня экспрессии AID в результате мутации приводит к образованию Muc-Igh транслокаций [64].

Но только ли активация механизма AID/APOBEC зависимого мутагенеза проявляет потенциальные сайты-мишени. Исходя из экспериментальных данных, можно предположить, что, в-третьих, соматический мутагенез и перестройки ДНК не случайны по отношению к метаболической ситуации в клетке.

Метаболизм клетки влияет на установление эпигенетических модификаций, на паттерн активных регуляторных элементов и связывающихся с ними транскрипционных факторов в потенциальных

локусах-мишенях AID/APOBEC-дезаминаз. Изменения метаболической ситуации в клетке могут повлечь за собой перепрограммирование активности самих регуляторных элементов и регулируемых ими генов в локусах-мишенях. В результате может установиться такой функциональный режим локуса, отражающий метаболическую ситуацию в клетке, который превратит вызываемые дезаминазами физиологические временные и легко репарируемые повреждения ДНК в мутации и перестройки генома.

В пользу этого предположения свидетельствует роль APOBEC дезаминаз в мутагенезе ERBS сайтов. Нормальный уровень эстрогенов через посредство эстрогеновых рецепторов привлекает к ERBS сайтам связывания APOBEC3B дезаминазу, которая вызывает физиологические DSB [56]. Повышенный уровень эстрогенов при связывании с ERBS сайтами увеличивает вероятность образования мутаций в регуляторных сайтах, а в некоторых случаях и в регулируемых ими генах [59].

Следовательно, AID/APOBEC зависимые мутации и перестройки генома могут быть неслучайными в пространстве, времени и по отношению к метаболической ситуации, складывающейся в клетке.

### References

1. Stavrou S., Ross S. R. APOBEC3 proteins in viral immunity. 2015. *J Immunol.* V. 195. P. 4565–4570.
2. Refsland E. W., Harris R. S. The APOBEC3 family of retroelement restriction factors. *Curr Top Microbiol Immunol.* 2013. V.371. P.1–27.
3. Duggal N. K., Emerman M. Evolutionary conflicts between viruses and restriction factors shape immunity. *Nat Rev Immunol.* 2012. V.12. P. 687–695.
4. Casellas R., Basu U., Yewdell W.T. [et. al.] Mutations, kataegis and translocations in B cells: understanding AID promiscuous activity. *Nat Rev Immunol.* 2016. V.16. P. 164–240.
5. Liu M. C., Liao W. Y., Buckley K. M. [et. al.] AID/APOBEC-like cytidine deaminases are ancient innate immune mediators in invertebrates. *Nat Commun.* 2018. V.16, P. 1948.
6. Boehm T. [et al.] VLR-based adaptive immunity. *Annu Rev Immunol.* 2012. V.30. P. 203–220.
7. Krishnan A., Iyer L. M., Holland S.J. [et. al.] Diversification of AID/APOBEC-like deaminases in metazoa: multiplicity of clades and widespread roles in immunity. *Proc Natl Acad Sci USA.* 2018. V.115. P. E3201–E3210.
8. Zhang D., Iyer L. M., Aravind L. A novel immunity system for bacterial nucleic acid degrading toxins and its recruitment in various eukaryotic and DNA viral systems. *Nucleic Acids Res.* 2011. V.39. P. 4532–4552.

9. Iyer L. M., Zhang D., Rogozin I. B., Aravind L. Evolution of the deaminase fold and multiple origins of eukaryotic editing and mutagenic nucleic acid deaminases from bacterial toxin systems. *Nucleic Acids Res.* 2011. V.39. P. 9473–9497.
10. Metzger M. J. [et al.] Widespread transmission of independent cancer lineages within multiple bivalve species. *Nature.* 2016. V.534. P. 705–709.
11. Hanington P. C. [et al.] Role for a somatically diversified lectin in resistance of an invertebrate to parasite infection. *Proc Natl Acad Sci USA.* 2010. V.107. P. 21087–21092.
12. Rouaud P., Vincent-Fabert C., Saintamand A., Fiancette R. [et al.] The IgH 3' regulatory region controls somatic hypermutation in germinal center B cells. *J. Exp. Med.* 2013. V.210. P. 1501–1507.
13. Saintamand A., Rouaud P., Saad F., Rios G. [et al.] Elucidation of IgH 3' region regulatory role during class switch recombination via germline deletion. *Nat. Commun.* 2015. V.6. P. 7084.
14. Buerstedde J. M., Alinikula J., Arakawa H. [et al.] Targeting of somatic hypermutation by immunoglobulin enhancer and enhancer-like sequences. *PLoS Biol.* 2014. V.12. P. e1001831.
15. Delgado-Benito V., Rosen D. B., Wang Q. [et al.] The Chromatin Reader ZMYND8 Regulates Igh Enhancers to Promote Immunoglobulin Class Switch Recombination. *Mol Cell.* 2018., V.72. P. 636–649.e8.
16. Hauser J., Grundström C., Kumar R. [et al.] Regulated localization of an AID complex with E2A, PAX5 and IRF4 at the Igh locus. *Mol Immunol.* 2016. V.80. P. 78–90.
17. Sheppard E. C., Morrish R. B., Dillon M. J. [et al.] Epigenomic Modifications Mediating Antibody Maturation. *Front Immunol.* 2018. V.26. P. 9–355.
18. Wang Q., Oliveira T., Jankovic M. [et al.] Epigenetic targeting of activation-induced cytidine deaminase. *Proc Natl Acad Sci U S A.* 2014. V. 111. P. 18667-18759.
19. Pavri R. [et al.] Activation-induced cytidine deaminase targets DNA at sites of RNA polymerase II stalling by interaction with Spt5. *Cell.* 2010. V. 143. P. 122–133.
20. Yamaguchi Y., Takagi T., Wada T. [et al.] NELF, a multisubunit complex containing RD, cooperates with DSIF to repress RNA polymerase II elongation. *Cell.* 1999. V. 97. P. 41–51.
21. Rahl P. B., Lin C. Y., Seila A. C. [et al.] c-Myc Regulates Transcriptional Pause Release. *Cell.* 2010. V.141. P. 432–445.
22. Yamane A., Resch W., Kuo N. [et al.] Deep-sequencing identification of the genomic targets of the cytidine deaminase AID and its cofactor RPA in B lymphocytes. *Nat Immunol.* 2011. V.12. P. 62–69.
23. Adelman K., Lis J. T. Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nat Rev Genet.* 2012. V.13. P. 720–751.
24. Chaudhuri J., Khuong C., Alt F. W. Replication protein A interacts with AID to promote deamination of somatic hypermutation targets. *Nature.* 2004. V.430. P. 992–998.
25. Pefanis E. [et al.] Noncoding RNA transcription targets AID to divergently transcribed loci in B cells. *Nature.* 2014. V.514. P. 389–393.
26. Core L. J. [et al.] Analysis of nascent RNA identifies a unified architecture of initiation regions at mammalian promoters and enhancers. *Nat. Genet.* 2014. V.46. P. 1311–1320.
27. Chen J., Cai Z., Bai M. [et al.] The RNA-binding protein ROD1/PTBP3 cotranscriptionally defines AID-loading sites to mediate antibody class switch in mammalian genomes. *Cell Research.* 2018. V.28. P. 981–995.
28. Zheng S., Vuong B. Q., Vaidyanathan B. [et al.] Non-coding RNA generated following lariat-debranching mediates targeting of AID to DNA. *Cell.* 2015. V.161. P. 762–773.
29. Chandra V., Bortnick A., Murre C. AID targeting: old mysteries and new challenges. *Trends Immunol.* 2015. V.36. P. 527–562.
30. Pasqualucci L. [et al.] BCL-6 mutations in normal germinal center B cells: evidence of somatic hypermutation acting outside Ig loci. *Proc Natl Acad Sci USA.* 1998. V.95. P. 11816–11821.
31. Liu M. [et al.] Two levels of protection for the B cell genome during somatic hypermutation. *Nature.* 2008. V.451. P. 841–845.
32. Staszewski O. [et al.] Activation-induced cytidine deaminase induces reproducible DNA breaks at many non-Ig Loci in activated B cells. *Mol. Cell.* 2011. V.41. P. 232–242.
33. Qian J. [et al.] B cell super-enhancers and regulatory clusters recruit AID tumorigenic activity. *Cell.* 2014. V.159. P. 1524–1537.
34. Dominguez P. M., Teater M., Chambwe N. [et al.] Methylation Dynamics of Germinal Center B Cells Are Mediated by AID. *Cell Rep.* 2015. V. 29, P. 2086–1284.
35. Álvarez-Prado Á. F., Pérez-Durán P., Pérez-García A. [et al.] A broad atlas of somatic hypermutation allows prediction of activation-induced deaminase targets. *J Exp Med.* 2018. V. 215, P. 761–771.

36. Hoang P. H., Dobbins S. E., Cornish A. J. [et al.] Whole-genome sequencing of multiple myeloma reveals oncogenic pathways are targeted somatically through multiple mechanisms. *Leukemia*. 2018. V.32. P. 2459–2470.
37. Li Z., Abraham B. J., Berezovskaya A. [et al.] APOBEC signature mutation generates an oncogenic enhancer that drives LMO1 expression in T-ALL. *Leukemia*. 2017. V.31. P. 2057–2064.
38. Alexandrov L. B. [et al.] Signatures of mutational processes in human cancer. *Nature*. 2013. V.500. P. 415–421.
39. Stephens P. J. [et al.] Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell*. 2011. V.144. P. 27–40.
40. Basu U. [et al.] The RNA exosome targets the AID cytidine deaminase to both strands of transcribed duplex DNA substrates. *Cell*. 2011. V.144. P. 353–363.
41. Beekman R., Chapaprieta V., Russiñol N. [et al.] The reference epigenome and regulatory chromatin landscape of chronic lymphocytic leukemia. *Nat Med*. 2018. V. 24. P. 868–880.
42. Rogozin I. B., Lada A. G., Goncarencu A. [et al.] Activation induced deaminase mutational signature overlaps with CpG methylation sites in follicular lymphoma and other cancers. *Sci Rep*. 2016. V. 7. P. 38133.
43. Rogozin I. B., Roche-Lima A., Lada A. G., Belinky F. [et al.] Nucleotide Weight Matrices Reveal Ubiquitous Mutational Footprints of AID/APOBEC Deaminases in Human Cancer Genomes. *Cancers (Basel)*. 2019. V. 12, pii: E211.
44. El Kadi N., Wang L., Davis A. [et al.] The EGFR T790M Mutation Is Acquired through AICDA-Mediated Deamination of 5-Methylcytosine following TKI Treatment in Lung Cancer. *Cancer Res*. 2018. V. 15, P. 6728–6735.
45. Popp C., Dean W., Feng S. [et al.] Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. *Nature*. 2010. V.463. P. 1101–1105.
46. Gibcus J. H., Dekker J. The hierarchy of the 3D genome. *Mol Cell*. 2013. V.49. P. 773–782.
47. Meng F. L., Du Z., Federation A. [et al.] Convergent transcription at intragenic super-enhancers targets AID-initiated genomic instability. *Cell*. 2014. V. 159, P. 1538–1586.
48. Lu Z., Pannunzio N. R., Greisman H. A. [et al.] Convergent BCL6 and lncRNA promoters demarcate the major breakpoint region for BCL6 translocations. *Blood*. 2015. V.126. P. 1730–1731.
49. Puc J., Aggarwal A. K., Rosenfeld M. G. Physiological functions of programmed DNA breaks in signal-induced transcription. *Nat Rev Mol Cell Biol*. 2017. V. 18. P. 471–476.
50. Ju B. G., Lunyak V. V., Perissi V. [et al.] A topoisomerase IIbeta-mediated dsDNA break required for regulated transcription. *Science*. 2006. N 312. P.1798–1802.
51. Madabhushi R. [et al.] Activity-Induced DNA Breaks Govern the Expression of Neuronal Early-Response Genes. *Cell*. 2015. V.161. P. 1592–1605.
52. Fernando P., Megeney L. A. Is caspase-dependent apoptosis only cell differentiation taken to the extreme? *FASEB J*. 2007. V.21. P. 8–17.
53. Larsen B. D. [et al.] Caspase 3/caspase-activated DNase promote cell differentiation by inducing DNA strand breaks. *Proc Natl Acad Sci U S A*. 2010. V.107. P. 4230–4235.
54. Puc J., Kozbial P., Li W. [et al.] Ligand-dependent enhancer activation regulated by topoisomerase-I activity. *Cell*. 2015. V.160. P. 367–380.
55. Perillo B., Ombra M. N., Bertoni A. [et al.] DNA oxidation as triggered by H3K9me2 demethylation drives estrogen-induced gene expression. *Science*. 2008. V. 319. P. 202–206.
56. Periyasamy M., Patel H., Lai C. F. [et al.] APOBEC3B-mediated cytidine deamination is required for estrogen receptor action in breast cancer. *Cell Rep*. 2015. V. 13. P. 108–129.
57. Deroo B. J., Korach K. S. Estrogen receptors and human disease. *J Clin Invest*. 2006. V.116. P. 561–570.
58. Yager J. D., Davidson N. E. Estrogen carcinogenesis in breast cancer. *N Engl J Med*. 2006. V. 354. P. 270–352.
59. Yang J., Wei X., Tufan T. [et al.] Recurrent mutations at estrogen receptor binding sites alter chromatin topology and distal gene expression in breast cancer. *Genome Biol*. 2018. V.19. P. 190.
60. Lin C., Yang L., Tanasa B. [et al.] Nuclear receptor-induced chromosomal proximity and DNA breaks underlie specific translocations in cancer. *Cell*. 2009. V.139. P. 1069–1083.
61. Elliott K., Boström M., Filges S. [et al.] Elevated pyrimidine dimer formation at distinct genomic bases underlies promoter mutation hotspots in UV-exposed cancers. *PLoS Genet*. 2018. V. 14. P. e1007849.

62. Mao P., Brown A. J., Esaki S. [et al.] ETS transcription factors induce a unique UV damage signature that drives recurrent mutagenesis in melanoma. *Nat Commun.* 2018. V.9. P. 26.

63. Dorsett Y. [et al.] MicroRNA-155 suppresses activation-induced cytidine deaminase-mediated Myc-Igh translocation. *Immunity.* 2008. V. 28. P. 630–638.

**AID/APOBEC–dependent somatic hypermutation and DNA rearrangements of immunoglobulin and non-immunoglobulin genes**

**Kolotova T. Yu., Makarenko V.D., Sorokoumova L.K., Davidenko M. B.**

Editing Ig genes by activation induced deaminase (AID) initiates the antibody diversification process in B lymphocytes. In mammalian B cells, this process includes somatic hypermutation (SHM) and class switch recombination (CSR). The activity of AID is largely confined to switch regions and Ig variable regions but now it is well established that AID/APOBEC-dependent damage and double DNA breaks leading to genome rearrangements and somatic mutagenesis are pervasive throughout the B cell genome. In this review, we focus on the molecular mechanisms that guide AID/APOBEC mutator to physiological and non-physiological targets. Epigenetic modifications, such as histone post-translational modifications and DNA methylation, cis-elements of DNA and the transcription factors that bind to it, which form regulatory clusters and superenhancers, synthesis of non-coding RNA and features of the transcription mode are involved in targeting of AID/APOBEC deaminases to Ig and non-Ig locus. However, not one of the studied factors is specific for the AID target loci. Thus, to date, there is insufficient clarity on the question of what determines the genetically programmed activity of AID in Ig loci and the genetically unprogrammed activity of AID/APOBEC family deaminases in non-Ig loci. The study of AID/APOBEC-dependent somatic hypermutation and rearrangements of Ig and non-Ig locus have both fundamental biological and applied medical significance. The fundamental significance is that the mechanisms of genetic rearrangements in somatic cells can provide a key to understanding the patterns of evolutionary variability of living organisms. The applied value is that based on the data obtained on the genomic instability of non-Ig loci, methods can be developed that suppress the instability of the genome during carcinogenesis and slow down the formation of chemotherapy resistant variants.

**Key words:** AID/APOBEC family of deaminases; DNA editing; Ig genes somatic hypermutation; Ig class switching recombination gene; AID/APOBEC off-targets; double-stranded DNA breaks; superenhancers; transcription factors; histone epigenetic modifications; non-coding RNA.