

# НЕКОДИРУЮЩИЕ РНК И СТРУКТУРА ХРОМАТИНА

©2007 г. С. А. ЛАВРОВ, М. В. КИБАНОВ

*Институт молекулярной генетики РАН, Москва*

I. Введение. II. Структура хроматина. Краткое описание. III. Системы дозовой компенсации. IV. РНК-интерференция и структура хроматина. V. Заключение.

## I. ВВЕДЕНИЕ

Помимо матричных РНК (мРНК), содержащих воспроизводимую в ходе трансляции информацию о первичной структуре белка, в клетках эукариот также синтезируется ряд молекул РНК, не кодирующих белки. Подобные РНК могут быть структурными компонентами органелл (к примеру, рибосомальные РНК), участвовать в синтезе белка (транспортные РНК), обладать ферментативной активностью или выполнять регуляторные функции.

Участие некодирующих РНК в регуляции транскрипции было обнаружено сравнительно недавно [1]. Наиболее известные и изученные системы – инактивация X-хромосомы млекопитающих и гиперактивация X-хромосомы дрозофилы при дозовой компенсации.

---

*Принятые сокращения:* HMG – high mobility group, АДФ – аденозин-дифосфат, НАТ – histone acetyltransferase, HDAC – histone deacetylase, HMT – histone methyltransferase, PC – Polycomb, PcG – Polycomb group, HP1 – heterochromatic protein 1, п.н – пар нуклеотидов, DMT – DNA methyltransferase, MBD – methyl-cytosine binding domain, *Xic* – X-chromosome inactivation center, *Xist* – X-inactive specific transcript, *Xite* – X-inactivation intergenic transcription element, DCC – dosage compensation complex, RNAi – RNA interference, siRNA – short interfering RNA, дцРНК – двухцепочечная РНК, RISC – RNA-induced silencing complex, RITS – RNA-induced transcriptional silencing, RDRC – RNA-directed RNA polymerase complex, RdDM – RNA-dependent DNA methylation, DDM1 – Decrease in DNA Methylation 1.

*Адрес для корреспонденции:* С.А.Лавров, ИМГ РАН, slavrov@img.ras.ru

Обзор написан при поддержке гранта РФФИ 07-04-07054-д, Программы фундаментальных исследований Президиума РАН «Молекулярная и клеточная биология», гранта Президента РФ ведущим научным школам НШ-6113.2006.4.

Кроме того, в последние годы получены данные о роли коротких некодирующих РНК (siRNA, компонента системы *РНК-интерференции*) в регуляции транскрипции генов на уровне структуры хроматина. Имеются указания на существование подобных схем регуляции у дрожжей, инфузории *Tetrahymena thermophila* (в последнем случае транскрипционная репрессия оказывается предварительным этапом перед физической элиминацией гетерохроматина [2, 3]), высших растений, млекопитающих [4–6], насекомых [7].

В обзоре основное внимание будет уделено механизмам достижения дозовой компенсации у дрозофилы и млекопитающих, а также двум связанным с РНК-интерференцией явлениям – РНК-зависимому метилированию ДНК (RdDM) у растений и зависимому от РНК формированию гетерохроматина у дрожжей. Причина в том, что эти процессы к настоящему времени наиболее подробно исследованы на молекулярном уровне – хотя многие детали остаются непонятными и даже труднообъяснимыми.

## II. СТРУКТУРА ХРОМАТИНА. КРАТКОЕ ОПИСАНИЕ

Хроматин можно определить как комплекс ДНК и связанных с ней белков. У всех эукариот структурной единицей хроматина является нуклеосома – образование из 8 молекул гистонов, формирующих ядро, вокруг которого обвит участок ДНК длиной примерно 145 п.н. ДНК делает около 1,7 оборота вокруг ядра. Соседние нуклеосомы могут находиться на разных расстояниях друг от друга, а плотность и регулярность укладки зависит от функционального состояния участка хроматина. В транскрипционно-активных районах плотность укладки низка – хроматин *декомпактизован*. Напротив, в *гетерохроматинизированных* районах (о них будет рассказано ниже) нуклеосомы разделены участками ДНК более-менее постоянной длины (около 40 п.н.) и расположены регулярно.

Регуляторные области генов обычно либо свободны от нуклеосом, либо содержат так называемые «позиционированные» нуклеосомы – то есть нуклеосомы с фиксированным положением относительно последовательности гена [8]. Это необходимо для обеспечения доступа транскрипционных факторов и компонентов транскрипционного комплекса к регуляторным участкам и промотору. Существует группа белков, способных с затратами энергии АТФ перемещать нуклеосомы на нити ДНК. Это так называемые факторы, моделирующие структуру хроматина (chromatin remodeling factors) [9]. В ряде случаев они необходимы для инициации транскрипции

либо, напротив, транскрипционной репрессии. В качестве примера можно привести белки семейства *swi/snf*, названные так по первому представителю, дрожжевому SWI2/SNF2. SWI2/SNF2 у дрожжей функционирует как активатор транскрипции, аналогичную роль играет ортолог у дрозофилы – BRAHMA. BRAHMA входит в состав комплекса из примерно 10 белков, обладающего моделирующей активностью [10]. Функционально близкие белки обнаруживаются у большинства эукариот [11–13].

Ядро нуклеосомы состоит из гистонов четырех типов (по два каждого) H2A, H2B, H3 и H4. Гистон H1 участвует во взаимодействии нуклеосом между собой и нужен для их укладки в компактную структуру более высокого порядка – так называемую 30-нм нить. Гистоны представляют собой небольшие основные белки, высоко консервативные у всех эукариот. Молекула гистона содержит глобулярную C-концевую часть и аморфный (не имеющий четко выраженной вторичной структуры) N-конец. N-концы гистонов выступают за пределы ядра и могут взаимодействовать с другими белками хроматина. Помимо гистонов основных типов, в геноме могут присутствовать минорные варианты, выполняющие специфические функции. В состав хроматина входят также негистоновые структурные белки – HMG, компоненты кинетохора, белки ядерной оболочки, топоизомераза II, и многие другие.

Гистоны не только обеспечивают укладку нити ДНК в ядре, но и активно участвуют в регуляции транскрипции. Это свойство связано с упомянутыми выше N-концами, основания в которых подвергаются ковалентным модификациям. Гистоны могут быть ацетилированы по различным основаниям, метилированы, фосфорилированы, АДФ-рибозилированы и убиквитинилированы. Разнообразие вариантов ковалентных модификации позволяет предполагать наличие уникального гистонного кода, определяющего возможность и уровень экспрессии генов [14–16]. На рис. 1 представлена структура нуклеосомы и варианты ковалентных модификаций гистонов. Число возможных сочетаний модификаций очень велико – достаточно, чтобы все нуклеосомы в геноме оказались уникальными.

Рассмотрение всех модификаций гистонов и их сочетаний выходит за рамки обзора, однако на некоторых вариантах следует остановиться – из-за высокой консервативности и универсальной роли таких модификаций.

*Ацетилирование.* Наиболее исследовано ацетилирование остатков лизина N-концов гистонов H3 и H4. Осуществляется ферментами ацетилтрансферазами гистонов (HAT). Ацетилирование

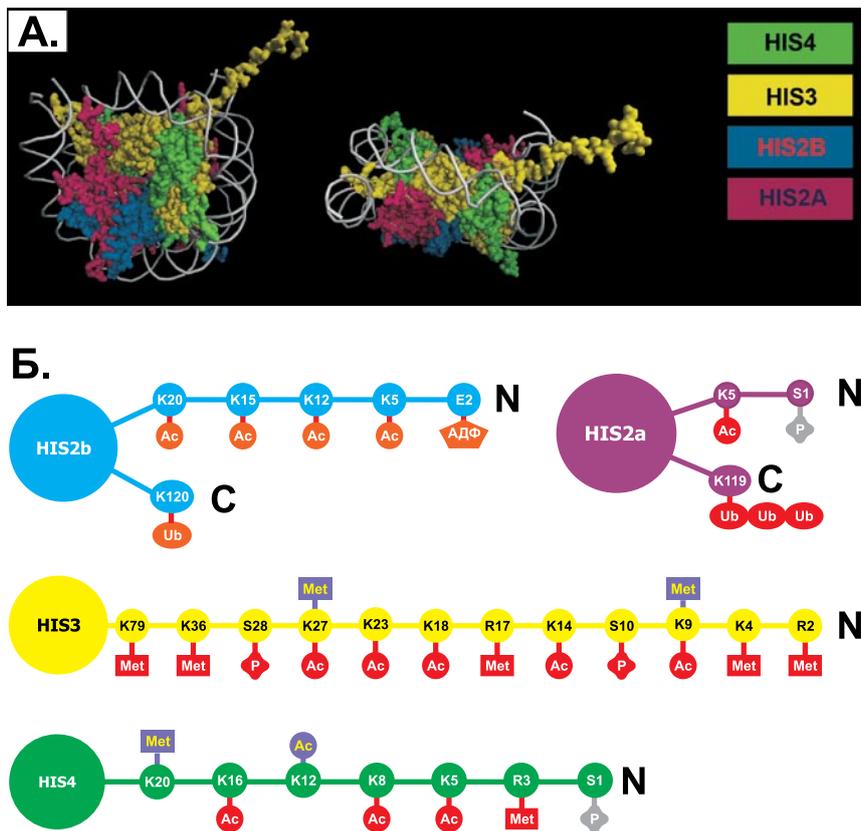


Рис. 1. Структура нуклеосомы и варианты ковалентных модификаций гистонов.

А. Модель нуклеосомы по данным рентгеноструктурного анализа в двух проекциях. Гистоны H2A, H2B, H3 и H4 в составе ядра нуклеосомы обозначены разными цветами.

Б. Варианты ковалентных модификаций гистонов. Подлежащие модификации аминокислоты в концевых районах гистонов обозначены однобуквенным кодом и номером, соответствующим положению начиная с N-конца молекулы. Ac – ацетилирование, Met – метилирование, P – фосфорилирование, Ub – убиквитинилирование, АДФ – АДФ-рибозилирование. Метилированы могут быть лизин и аргинин, ацетируется лизин, фосфорилируется серин, АДФ-рибозилируется глутаминовая кислота. Синим цветом обозначены модификации, характерные для репрессированного хроматина, красным – для активного. Серым цветом отмечены модификации, связанные с конденсацией хромосом при митозе либо гаметогенезе.

ассоциируется с активацией транскрипции, хотя, возможно, не для всех вариантов модификации [17]. Ацетилирование по лизину в 16 положении гистона H4 специфически связано с гиперактивацией транскрипции при дозовой компенсации (далее). Деацетилирование гистонов осуществляется ферментами деацетилазами (HDAC) и связано с транскрипционной репрессией.

*Метилирование.* Метилируются остатки лизина, причем каждый остаток может быть моно-, ди- и триметилирован. Метилирование лизина в 9 и 27 положениях гистона H3 связано с транскрипционной репрессией, а метилирование лизина в положении 4 – с активацией транскрипции. Метилирование осуществляют метилтрансферазы гистонов (HMT), деметилирование исследовано слабо [18].

Гистон H3, метилированный по лизину в **положении 9** у всех эукариот является маркером *гетерохроматина*.

Гетерохроматин определялся изначально как часть хроматина, остающаяся конденсированной на протяжении всего клеточного цикла (в противоположность эухроматину). В основном это прицентромерные и теломерные области хромосом. Структурно гетерохроматин состоит в основном из повторов различных типов и имеет характерный набор белков и модифицированных гистонов, из которых типичен упомянутый выше гистон H3, метилированный по лизину-9. Обычные эухроматиновые гены при переносе в гетерохроматиновое окружение репрессируются на уровне транскрипции. Процесс приобретения участка эухроматина в силу тех или иных причин свойств гетерохроматина называется *гетерохроматинизация*.

С гетерохроматином ассоциирован ряд белков, взаимодействующих между собой и обеспечивающих поддержание характерного компактного состояния. Наиболее подробно структурные компоненты гетерохроматина исследованы у дрозофилы [19–21], однако и у других эукариот имеются сходные компоненты.

Характерными белками гетерохроматина являются:

*HDAC1* – деацетилаза гистонов. Деацетилазы имеются у всех эукариот (например, у дрожжей CLR3, у растений HDA6)[15, 22].

*SUV39* – метилтрансфераза гистонов, специфично метилирующая лизин-9 гистона H3. Ортологи имеются у дрожжей (CLR4), растений (KYP), млекопитающих – и, видимо, у всех эукариот [20, 23].

*HP1* – гетерохроматиновый белок 1. У дрожжей функциональный эквивалент SWI6, имеются ортологи у растений и млекопитающих. Характерной особенностью HP1 является наличие *хромодомена*. Хромодомен – белковый домен, характерный для многих белков хроматина. Хромодомены различаются по структуре, способны

связываться с метилированными гистонами и, возможно, с РНК [24–27]. Хромодомен HP1 и аналогов специфически узнает гистон H3, метилированный по лизину-9 [28].

В формировании гетерохроматина участвует большое количество других белков – компоненты аппарата репликации ДНК [29], ДНК-связывающие белки (SUVAR 3–7, [30]), в геноме имеется несколько потенциальных ацетилаз, деацетилаз и метилтрансфераз гистонов.

Помимо гетерохроматиновых белков, у дрозофилы найдена еще одна система репрессии генов – группа *Polycomb* (PcG). Группа названа по одному из членов – содержащему хромодомену белку POLYCOMB (PC) и включает более 10 участников [31, 32]. В отличие от гетерохроматиновых белков, покрывающих районы генома размером в миллионы п.н., белки *Polycomb* осуществляют «точечную» реессию эухроматиновых генов. Мишени группы – ряд генов, отвечающих за развитие мухи, к примеру гены из кластеров *Bithorax* и *Antennapedia* [33, 34], общее число мишеней несколько сотен [35, 36]. Аналогичные системы обнаружены у растений, млекопитающих, нематоды и, видимо, имеются у всех высших эукариот [37–40]. Хромодомен PC и родственных белков отличается по специфичности от HP1 и узнает гистон H3, метилированный по лизину в **27-м положении** [24].

Возможны ковалентные модификации не только гистонов, но и ДНК. Это метилирование остатков цитозина с образованием метилцитозина. Подобная модификация ДНК обнаружена и играет большую роль в регуляции транскрипции у млекопитающих, растений, тогда как у дрожжей, нематоды *Caenorabditis elegans* не обнаружена. У дрозофилы метилирование ДНК незначительно и функции его не вполне ясны [41].

Метилируются обычно основания цитозина в последовательностях вида CpG (симметричное метилирование), CpNpG и CpHpH (несимметричное метилирование, Н – А, Т или С). В регуляторных областях генов имеются участки, обогащенные сайтами метилирования – так называемы «CpG-островки». Метилирование осуществляют ферменты ДНК-метилтрансферазы (DMT), представленные у растений и млекопитающих в большом разнообразии [42, 43]. Обычно метилирование последовательности гена ассоциировано с транскрипционной репрессией, а сама метилированная ДНК узнается белками, содержащими метилцитозин-связывающий домен (MBD) и входящими в состав репрессирующих комплексов [44–46].

### III. СИСТЕМЫ ДОЗОВОЙ КОМПЕНСАЦИИ

У многих организмов половые хромосомы (X и Y) отличаются по набору генов. К примеру, и у млекопитающих, и у плодовой мушки *Drosophila melanogaster* X-хромосома самок полноценная, содержит большое число генов и в целом не отличается от неполовых хромосом (аутосом), тогда как мужская Y-хромосома содержит мало генов и гетерохроматинизирована. У млекопитающих и *Drosophila melanogaster* набор половых хромосом самцов – XY, самок XX.

Организмы с разными половыми хромосомами сталкиваются с проблемой двукратного избытка генов в случае гомогаметного пола (XX) по сравнению с гетерогаметным (XY). Это несоответствие в количестве генов должно быть скомпенсировано для того, чтобы обеспечить одинаковый уровень экспрессии генов X-хромосомы в клетках обоих полов. Процесс выравнивания уровня транскрипции генов с двух X-хромосом самок и одной X-хромосомы самцов называется «дозовая компенсация».

#### ДОЗОВАЯ КОМПЕНСАЦИЯ МЛЕКОПИТАЮЩИХ

У млекопитающих был выработан механизм дозовой компенсации, заключающийся в инактивации одной из двух X-хромосом в клетках самок на раннем этапе эмбрионального развития. Согласно имеющимся данным, инактивация происходит на стадии четырехклеточного зародыша [47, 48]. Этот процесс регулируется участком X-хромосомы, называемым «центр инактивации X-хромосомы» (*Xic*) [49]. Локус *Xic* необходим для процесса инактивации: он участвует в выборе X-хромосомы, которой предстоит быть выключенной, а также в инициации этого процесса. Именно от локуса *Xic* репрессия транскрипции распространяется далее в обе стороны вдоль X-хромосомы. Если удалить локус, то мутантная X-хромосома не инактивируется, а вставка дополнительной копии в одну из аутосом приводит к выключению соседних с местом инсерции генов [50–52].

С локуса *Xic* транскрибируются некодирующие РНК – *Xist*, длиной около 17 т.н., и *Tsix*, антисмысловая по отношению к *Xist* (отсюда произошло название *Xist* ↔ *tsiX*). С процессом дозовой компенсации у млекопитающих на самом деле связано два РНК-зависимых эффекта – инактивация одной X-хромосомы при участии РНК *Xist* и репрессия транскрипции самой *Xist* при участии РНК *Tsix* на второй (остающейся активной) X-хромосоме [1, 50, 53–56]. Оба процесса тесно увязаны друг с другом, сбой приводит к нежизнеспособности эмбриона в случае, если обе X-хромосомы будут выключены либо останутся включенными. Рассмотрим их подробнее.

*Xist. инактивация X-хромосомы*

Процесс инактивации начинается со значительного усиления транскрипции гена *Xist*. В результате, число копий РНК-продукта на клетку резко возрастает [57]. Образующиеся транскрипты накапливаются вблизи локуса *Xic* и распространяются в обе стороны от него вдоль X-хромосомы. В тех случаях, когда ген *Xist* не работает, X-хромосома не инактивируется [58].

Инактивация X-хромосомы с участием *Xist* сопровождается ковалентными модификациями гистонов. Имеет место деацетилирование гистонов H3 и H4, деметилирование по лизину в 4-м положении гистона H3 и метилирование 9-го и 27-го остатков лизина гистона H3 [59–61]. В дальнейшем происходит замена гистона H2A на его гетерохроматиновый вариант – тасгоH2A1, ДНК активно метилируется и инактивированная X-хромосома начинает реплицироваться в ходе клеточного цикла с задержкой [62, 63]. X-хромосома приобретает все основные черты гетерохроматина.

Механизм, с помощью которого *Xist*-РНК приводят к выключению соответствующей X хромосомы, до сих пор остается неясным. Факт накопления транскриптов в большом количестве на инактивируемой X-хромосоме позволяет предполагать, что, возможно, это способствует привлечению белков, участвующих в гетерохроматинизации X-хромосомы. В частности, было обнаружено что белки группы *Polycomb* (PcG) – EED и ENX1 (так же известный, как EZH2) – локализуются на X-хромосоме в зависимости от присутствия *Xist*-РНК, но только на самом начальном этапе инактивации [60, 64, 65]. Появление комплекса EED-ENX1 сопровождается метилированием лизина в положении 27 гистона H3. Как полагают, вносимые этими белками и ассоциированными с ними ферментными комплексами изменения в структуру хроматина способствуют поддержанию выключенного состояния X-хромосомы. На более позднем этапе инактивации, уже после формирования «закрытого» хроматина, происходит замена гистона H2A на его минорный вариант – тасгоH2A1, обычно встречающийся в гетерохроматине. Этот процесс также зависит от присутствия *Xist*-РНК [66, 67]. Интересно, что экспрессия гена *Xist* приводит к инактивации X-хромосомы только в клетках, находящихся в процессе дифференцировки, тогда как в дифференцированных клетках механизм дозовой компенсации не работает. Другими словами, существует определенный временной интервал в развитии, когда происходит инактивация одной из X хромосом, вне этих рамок дозовая компенсация невозможна [68]. Возможно, что временное «окно» связано с активностью комплекса EED-ENX1.

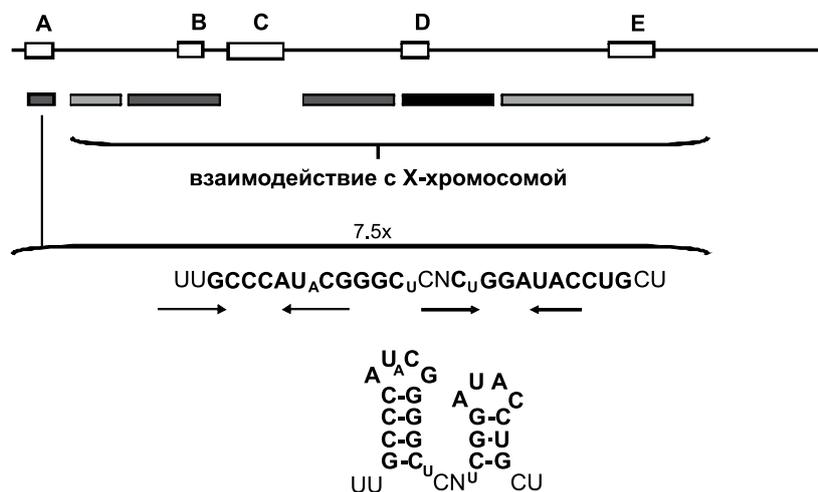


Рис. 2. Структура некодирующей РНК *Xist*.

Обозначены консервативные участки (А–Е), встречающиеся у различных млекопитающих. Участок А на 5'-конце РНК необходим для инактивации X-хромосомы. Последовательность участка повторяется 7,5 раз и состоит из инвертированных повторов, способных образовывать две шпилечные структуры (подстрочным шрифтом указаны наиболее редко встречающиеся основания). Остальные районы *Xist* отвечают за ее ассоциацию с X-хромосомой (чем сильнее окрашивание, тем больше влияние на взаимодействие).

С помощью делеционного анализа было продемонстрировано, что за локализацию на X-хромосоме и за выключение генов отвечают разные участки *Xist*-РНК [67] (рис. 2). Первые разбросаны по всей длине молекулы, они не имеют общих мотивов и, скорее всего, представляют собой низкоаффинные сайты, с которыми кооперативно связываются пока не известные факторы. В то же время на 5'-конце *Xist*-РНК был локализован консервативный участок, отвечающий за установление репрессии (А-участок). Делеция этого района не сказывается на накоплении *Xist*-РНК на X хромосоме, однако инактивации транскрипции при этом также не происходит. Участок содержит инвертированные повторы, которые, согласно результатам компьютерного моделирования, способны образовывать шпилечные структуры, возможно, являющиеся сайтами связывания для белков-компонентов репрессорных комплексов. Делеции остальных консервативных районов не оказывали никакого влияния на процесс дозовой компенсации [69, 70]. В настоящее время одной из задач исследователей является поиск белков, взаимодействующих с функциональными доменами *Xist*-РНК.

РНК *Xist* не обнаруживается в свободном виде в нуклеоплазме [67]. Следовательно, молекулы *Xist* либо быстро разрушаются вне комплекса с инактивируемой X-хромосомой, либо связываются с X-хромосомой котранскрипционно и в нуклеоплазме не диффундируют. Возможно, данный факт является важным для ограничения «поля деятельности» *Xist*-РНК X-хромосомой и предупреждения ошибочной инактивации аутосом и (или) второй X-хромосомы.

#### *Tsix. инактивация Xist*

Инактивация X-хромосомы происходит в два этапа. Первичная инактивация может произойти уже на стадии четырехклеточного зародыша [47, 48], при этом выключается X-хромосома отцовского происхождения. Если на этом этапе дозовая компенсация в силу тех или иных причин окажется нарушенной, эмбрион будет нежизнеспособен. Далее следует предимплантационный период развития и имплантация. После имплантации X-хромосома от отца может быть реактивирована а затем одна из X-хромосом снова репрессируется случайным образом [48, 71]. Инактивация X-хромосомы отцовского происхождения – пример *геномного импринтинга*, одного из вариантов *эпигенетического наследования*. Под эпигенетическим наследованием подразумевается наследование признаков, не связанное с изменениями первичной структуры ДНК.

На хромосоме, остающейся активной, экспрессия гена *Xist* подавляется. Происходит это, судя по всему, при участии бессмысленного транскрипта, так как в тех случаях, когда ген *Tsix* не экспрессируется, например, в результате нарушения промоторной области, наблюдается накопление *Xist*-РНК на хромосоме и ее последующая инактивация, тогда как продолжающаяся транскрипция гена *Tsix* предотвращает образование *Xist*-РНК [54, 55, 72, 73]. Если X-хромосома с поврежденным промотором гена *Tsix* наследуется от отца, то развитие протекает нормально, при этом в организме самок выключается исключительно «отцовская» хромосома. Напротив, большинство эмбрионов обоих полов гибнут, если наследуют поврежденную X хромосому от матери, так как в этом случае выключается единственная X-хромосома у самцов и обе X-хромосомы у самок. Таким образом, хромосома с поврежденным локусом *Tsix* инактивируется независимо от своего происхождения, то есть родителя, от которого она наследуется. Исходя из приведенных данных, было высказано предположение, что транскрипция гена *Tsix*, скорее всего, и является тем фактором импринтинга, который обеспечивает инактивацию «отцовских» X-хромосом и оставляет активными

X-хромосомы, наследуемые от матери. Однако экспрессия гена *Tsix* в норме в начале эмбриогенеза не детектируется. Следовательно, несмотря на то, что ген *Tsix* необходим для поддержания активного состояния X хромосомы, он вряд ли отвечает за репрессию гена *Xist* на «материнской» X хромосоме в самом начале развития, и должен существовать какой-то другой механизм. Были предложены две модели, объясняющие механизм импринтинга. Согласно одной из них, X-хромосома, наследуемая от отца, находится уже в «выключенном» состоянии [47], согласно другой, более вероятной, она быстро инактивируется после активации зиготы [48], при этом X-хромосома материнского происхождения не инактивируется, так как её ген *Xist* репрессирован эпигенетически. Данный механизм не зависит от гена *Tsix* или его продукта, по крайней мере, на ранних стадиях развития, однако, судя по всему, ген *Tsix* играет определенную роль в его поддержании, так как мутация в локусе *Tsix* в конечном итоге приводит к активации транскрипции гена *Xist* и инактивации «материнской» X-хромосомы. Природа импринтинга в этом случае до сих пор остается не выясненной, однако известно, что устанавливается он в процессе созревания ооцитов в профазе I мейоза [74]. Ранее сообщалось о том, что, возможно, метилирование промотора гена *Xist* является основой импринтинга [75–77], однако дальнейшие исследования не подтвердили эту точку зрения [78, 79].

Из приведенных данных следует, что процесс инактивации X хромосомы находится под контролем генов *Xist* и *Tsix* и кодируемых ими РНК, в то же время сам ген *Xist* негативно регулируется экспрессией антисмыслового транскрипта. Молекулярные механизмы активности *Tsix*, по имеющимся данным, схожи с механизмами работы *Xist*, за исключением того, что мишень *Xist* – хроматин X-хромосомы, а мишень *Tsix* – сам *Xist*. Например, на эмбриональных стволовых клетках было показано, что транскрипция гена *Tsix* приводит к формированию «закрытой» структуры хроматина вокруг промотора гена *Xist*, как полагают, в результате привлечения белковых комплексов, которые устанавливают соответствующие модификации N-концов гистонов и метилируют ДНК. В частности, в экспериментах по иммунопреципитации хроматина *Tsix*-РНК была обнаружена в комплексе с ДНК-метилтрансферазой DNMT3a, которая обладает способностью метилировать ДНК *de novo* [80]. Из более ранних работ известно, что промотор гена *Xist* по-разному метилирован в дифференцированных XX клетках, причем уровень метилирования коррелирует с транскрипционным статусом аллелей: неактивный промотор метилирован значительно сильнее по сравнению с активным промотором [81]. В то же время, в недиф-

ференцированных XX клетках оба промотора частично модифицированы в равной степени [50, 82]. Это свидетельствует о появлении функционально значимого метилирования лишь в процессе инактивации X-хромосомы. Исчезновение транскрипта *Tsix* приводит к нарушению метилирования и дерепрессии обоих локусов *Xist*, что приводит к массовой гибели клеток [57, 83]. На основании этих данных было высказано предположение, что *Tsix*-РНК активирует DNMT3а, которая в результате метилирует локус *Xist* активной X-хромосомы. В свою очередь, метилирование обеспечивает репрессию *Xist*. Роль *Tsix* как индуктора метилтрансферазной активности DNMT3а подтверждается наблюдениями, согласно которым фермент присутствует на промоторе гена *Xist* до начала процесса инактивации и даже в клетках, лишенных транскрипта *Tsix* [57].

Альтернативой индуцируемому *Tsix* изменению структуры хроматина промотора *Xist* мог бы быть механизм, связанный с формированием дуплекса *Xist-Tsix* за счет комплементарных участков и затем дегградации РНК *Xist* по механизму, аналогичному РНК-интерференции. Однако, образование дуплекса кажется маловероятным в связи с особенностями сплайсинга РНК *Tsix* [55, 84]. В результате сплайсинга во всех случаях происходит вырезание большей части комплементарного района, сохраняется лишь небольшой фрагмент длиной 1.9 т. н., перекрывающийся с 5'-концом *Xist*, включающим консервативный А-участок. Как было показано, повторы А-участка важны для инактивирующей роли *Xist*, поэтому было высказано предположение, что *Tsix* работает через сплайсированные формы РНК, которые блокируют А-участок *Xist*. Однако подобный механизм также кажется маловероятным, так как было установлено, что присутствие сплайсированных форм РНК *Tsix* само по себе не достаточно для выключения гена *Xist* [84], и, кроме того, большинство молекул РНК *Tsix* являются укороченными в результате преждевременной остановки транскрипции. Так, в случае *Tsix*-РНК человека было показано, что транскрипция останавливается вскоре после пересечения 3'-границы гена *Xist* [85, 86].

Наряду с генами *Xist* и *Tsix* в районе локуса *Xic* относительно недавно был картирован еще один ген, названный *Xite*, в результате транскрипции которого также образуется некодирующая РНК. Направление транскрипции гена *Xite* совпадает с таковым для гена *Tsix*, а множественные сайты начала транскрипции сосредоточены в двух участках, один из которых расположен на 10 т.п.н., а другой на 15 т.п.н. выше старта транскрипции гена *Tsix*, то есть его 5'-конца [55, 87]. Ген экспрессируется на невысоком уровне: количество *Xite*-

РНК в эмбриональных стволовых клетках мышей в 10–60 раз ниже по сравнению с содержанием *Tsix*-РНК. Делеция или искусственное выключение гена *Xite* приводят к нарушению стабильной экспрессии гена *Tsix* в течение процесса дифференцировки и инактивации мутантной X-хромосомы [87]. В то же время, судя по всему, работа *Xite* не зависит от РНК-продукта. Как полагают, ближайший к *Tsix* промоторный участок гена *Xite* содержит энхансер гена *Tsix* [87, 88]. Таким образом, согласно современным данным, *Xite* вместе с *Tsix* отвечает за активное состояние одной из X-хромосом. Благодаря присутствию энхансера *Xite* обеспечивает стабильную экспрессию гена *Tsix* в процессе дифференцировки, а *Tsix*, в свою очередь, инактивирует ген *Xist*.

#### ГИПЕРАКТИВАЦИЯ X-ХРОМОСОМЫ ДРОЗОФИЛЫ. ROX

Компенсация количества генов, связанных с X-хромосомой, происходит также у плодовой мушки *Drosophila melanogaster*, однако механизм этого процесса является прямо противоположным таковому млекопитающих. Обе X-хромосомы самок дрозофилы транскрибируются с одинаковым базовым уровнем, тогда как у самцов, напротив, скорость транскрипции большинства генов X-хромосомы удваивается, чтобы обеспечить эквивалентное образование продукта. Гиперэкспрессия генов зависит от присутствия рибонуклеопротеинового комплекса DCC, состоящего из шести различных белков (рис. 3). DCC связывается с многочисленными сайтами вдоль X-хромосомы самцов, так называемыми «точками входа» [89]. Пять белков комплекса известны как MSL-белки (male-specific lethal): MSL-1, -2 и -3, MLE (maleless) и MOF (males absent on the first). Мутации по этим белкам предотвращают компенсаторную активацию X-хромосомы и приводят к гибели самцов. Шестой компонент DCC – белок JIL-1 – также обнаруживается на X-хромосоме и взаимодействует с MSL-белками [90, 91].

MOF является ацетилтрансферазой гистонов, ацетилирующей остаток лизина в 16-м положении гистона H4 [92, 93]. JIL-1 обладает киназной активностью, способен фосфорилировать гистон H3 по серину в положении 10 *in vitro* [94 – 96]. Функции MOF и JIL-1 указывают на то, что в основе активации транскрипции лежат изменения структуры хроматина X-хромосомы самцов, в частности гистонового кода. Ацетилированный по лизину в 16 положении гистон H4 специфичен для гиперактивированной X-хромосомы.

Вместе с белковыми факторами в компенсаторной регуляции транскрипции генов X-хромосомы *Drosophila* также принимают учас-

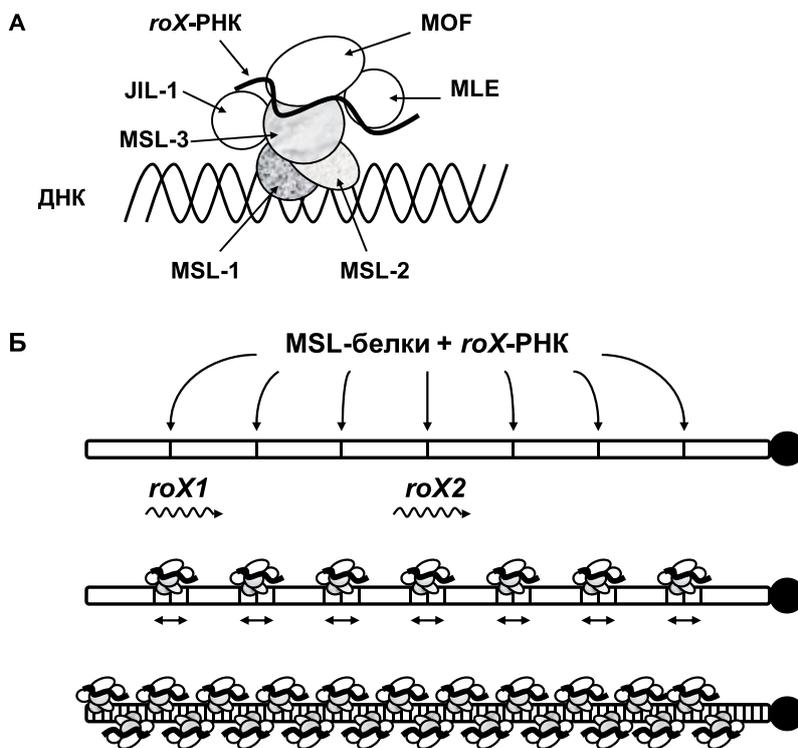


Рис. 3. Дозовая компенсация у *Drosophila melanogaster*

Дозовая компенсация осуществляется нуклеопротеидным комплексом – компенсатором (А). Образование комплекса происходит ступенчато и инициируется белками MSL-1 и MSL-2, которые связываются с 30–35 высокоаффинными сайтами на X-хромосоме – «точками входа» (Б). Для включения в комплекс РНК *roX* необходим белок MLE, который обладает РНК-хеликазной активностью. Белки MSL-3 и MOF также способны взаимодействовать с РНК за счет хромодоменов. JIL-1 – киназа гистонов. От высокоаффинных сайтов компенсаторные комплексы распространяются далее вдоль X-хромосомы, покрывая всю ее длину и вызывая увеличение уровня транскрипции в два раза.

тие молекулы РНК. Было обнаружено, что DCC содержит, по крайней мере, две некодирующие РНК: *roX1* длиной 3.7 т. н. и *roX2* длиной около 0.6 т.н. [97]. Как и в случае с *Xist*, гены *roX* (RNA on the X) картированы на X-хромосоме, а активация транскрипции происходит в результате накопления и распространения РНК-продуктов генов *roX* от места их образования вдоль X-хромосомы вместе с белками DCC. Это было подтверждено в экспериментах, в которых трансген с

локусом *roX* встраивался в одну из аутосом: на расстоянии в несколько тысяч п. н. в обе стороны от места инсерции происходило накопление РНК-продукта и DCC-белков, а также наблюдался повышенный уровень ацетилирования гистона H4 [98, 99]. Данные результаты свидетельствуют также о том, что *roX*-РНК работают независимо от нуклеотидной последовательности мишени и могут вызывать гиперактивацию участков аутосом.

Компенсаторный комплекс собирается исключительно в клетках самцов *Drosophila*, несмотря на то, что белки MSL-1 и -3, MLE и MOF присутствуют также в организме самок и наследуются в составе цитоплазмы ооцитов [100]. Подобная специфичность связана с экспрессией в организме самок, но не самцов, белка SXL – продукта гена *Sex-lethal (Sxl)*, – который препятствует трансляции мРНК MSL-2 [101, 102]. В отсутствие белка MSL-2 DCC не собирается, так как длительная экспрессия и стабильность других MSL белков – MSL-1 и -3 – зависит от его наличия в клетках [103]. Кроме того, показано, что MSL-2 участвует в инициации образования компенсаторного комплекса на X-хромосоме [102]. Так, экспрессия белка MSL-2 с трансгена, мРНК которого лишена сайта связывания белка SXL, в клетках самок приводит к гиперактивации обеих X хромосом и, вследствие этого, к резкому снижению жизнеспособности особей [95, 104, 105]. Белки MSL также требуются для нормальной транскрипции и поддержания стабильности образующихся *roX*-РНК [101, 103].

До сих пор не решён до конца вопрос о роли *roX*-РНК в гиперактивации X-хромосомы. Согласно существующим данным, *roX*-РНК, скорее всего, необходимы для связывания компенсаторных комплексов с X-хромосомой и дальнейшего распространения MSL-белков вдоль хромосомы, чем для формирования собственно DCC. Обнаружено, по крайней мере, 30–35 «точек входа» – высокоафинных участков на X-хромосоме, с которыми происходит связывание компенсаторных комплексов [97]. Установлено, что два из этих сайтов совпадают с локусами генов *roX* [96]. Таким образом, *roX*-РНК, связываясь с MSL-белками вблизи места своего образования, могут направлять DCC к X-хромосоме. Было показано, что в отсутствии *roX*-РНК жизнеспособность самцов резко снижается, и DCC начинают связываться с аутосомами и гетерохроматином [96]. В то же время из экспериментов, в которых мутантов по обоим генам *roX* удавалось спасти с помощью введения дополнительных копий генов *roX* в аутосомы [96], следует, что *roX*-РНК могут также функционировать, находясь в транс-положении по отношению к X-хромосоме, что предполагает участие этих РНК в узнавании X-хромосомы комплексом DCC. Роль

*roX*-РНК в наведении компенсаторных комплексов на X-хромосому подтвердилась также и в случае самок *Drosophila*, экспрессирующих белок MSL-2: при отсутствии транскриптов обоих генов происходило нормальное развитие особей [95, 106]. Обе *roX*-РНК являются взаимозаменяемыми, и функциональные компенсаторные комплексы собираются в отсутствие одного или другого транскрипта [105], однако, согласно некоторым данным *roX2*-РНК имеет большее значение [107]. Тем не менее функциональные элементы *roX*-РНК, которые могли бы служить сайтами связывания MSL-белков или использовались бы для узнавания X-хромосомы, до сих пор остаются неизвестными.

Относительно недавно с помощью делеционного анализа была предпринята попытка определить назначение отдельных участков *roX1*-РНК [106]. Был сконструирован набор делеций, перекрывающих всю длину молекулы РНК, и проверена способность делеций осуществлять дозовую компенсацию *in vivo*. Размер делеции в каждом случае составлял 260 – 400 нуклеотидов. Анализ показал, что за исключением 3'-конца *roX1*-РНК делеции не приводят к нарушению дозовой компенсации. Подобные результаты согласуются с тем, что между самими *roX*-РНК гомология по последовательности нуклеотидов практически отсутствует, и, судя по всему, функционально важной является третичная структура молекул РНК, их пространственная организация. Единственный короткий участок гомологии между *roX*-РНК размером около 30 н. [67] может быть удален без очевидных последствий. Интересно, что в случае *Xist* млекопитающих консервативные последовательности РНК также не имеют функционального значения [96]. В то же время возможно, что функциональные элементы рассеяны вдоль молекулы *roX1*-РНК и находятся в большом избытке, поэтому делеция одного или нескольких таких элементов никак не влияет на дозовую компенсацию. Наряду с тем, что большинство удаленных участков не оказывали никакого эффекта, делеция области на 3'-конце *roX1*-РНК, напротив, приводила к резкому снижению жизнеспособности самцов. В результате анализа нуклеотидной последовательности этого района были обнаружены инвертированные повторы, которые, как было предсказано с помощью компьютерного моделирования, способны образовывать шпильчатую структуру. Как известно, шпильки являются довольно частыми и важными структурными элементами молекул РНК и нередко служат сайтами связывания белков. Тщательный делеционный анализ района повторов показал что шпильчатая структура, по крайней мере частично, ответственна за снижение жизнеспособности самцов. Предполагается, что в 3'-концевом участке *roX1*-РНК также присутствуют дополнительные регуляторные элементы.

С другой стороны, судя по всему, MSL-белки сами по себе, даже в отсутствие *roX*-РНК, способны связываться с сайтами на X-хромосоме и гиперактивировать транскрипцию генов. Так в мутантах, полностью лишенных *roX* [108], DCC хотя и в меньшей степени, но по-прежнему детектируются на X-хромосоме, причем характер их распределения совпадает с таковым в случае особей дикого типа. Кроме того, было показано, что функция локусов *roX* как центров сборки DDC не зависит от транскрипции генов [96]. Наконец, из 30–35 «точек входа» только два совпадают с локусами генов с известными транскриптами. Можно было бы предполагать существование подобных некодирующих РНК и для других сайтов, однако резкое снижение жизнеспособности самцов в случае двойных мутантов по генам *roX* исключает подобную возможность. Таким образом, гены *roX* играют двоякую роль в процессе дозовой компенсации *Drosophila*: с одной стороны, они являются источником некодирующих РНК, которые обеспечивают привлечение MSL-белков преимущественно к X хромосоме, а с другой – сами по себе, наряду с другими 30–35 сайтами, являются участками сборки компенсаторных комплексов, даже в отсутствии транскрипции соответствующих генов. Способность связывать MSL-белки является, скорее всего, свойством последовательностей ДНК сайтов образования компенсаторных комплексов, а *roX*-РНК в этом случае выступают в качестве кофакторов, обеспечивающих предпочтительное связывание DCC с X-хромосомой. В свете этого можно предполагать существование других *roX*-подобных РНК, специфичных для ограниченного набора генов [105]. Пока же *roX*-РНК представляют собой пока единственный известный пример регуляторной системы, опирающейся на некодирующие РНК для привлечения белков-активаторов транскрипции.

#### IV. РНК-ИНТЕРФЕРЕНЦИЯ И СТРУКТУРА ХРОМАТИНА

В последние десять лет был обнаружен и сейчас активно исследуется процесс РНК-интерференции (RNAi) [109–112]. Явление имеет место у эукариот от дрожжей до млекопитающих и описано первоначально как система деградации матричной РНК, направляемая комплементарными малыми РНК (siRNA). Молекулярные механизмы процесса сходны у эукариот. Исходный субстрат – длинная двухцепочечная РНК – возникает несколькими путями (рис. 4). В их числе транскрипция инвертированных повторов, транскрипция участка ДНК в смысловой и антисмысловой ориентации (с двух промоторов), активность РНК-зависимой РНК-полимеразы. Двухцепочечные

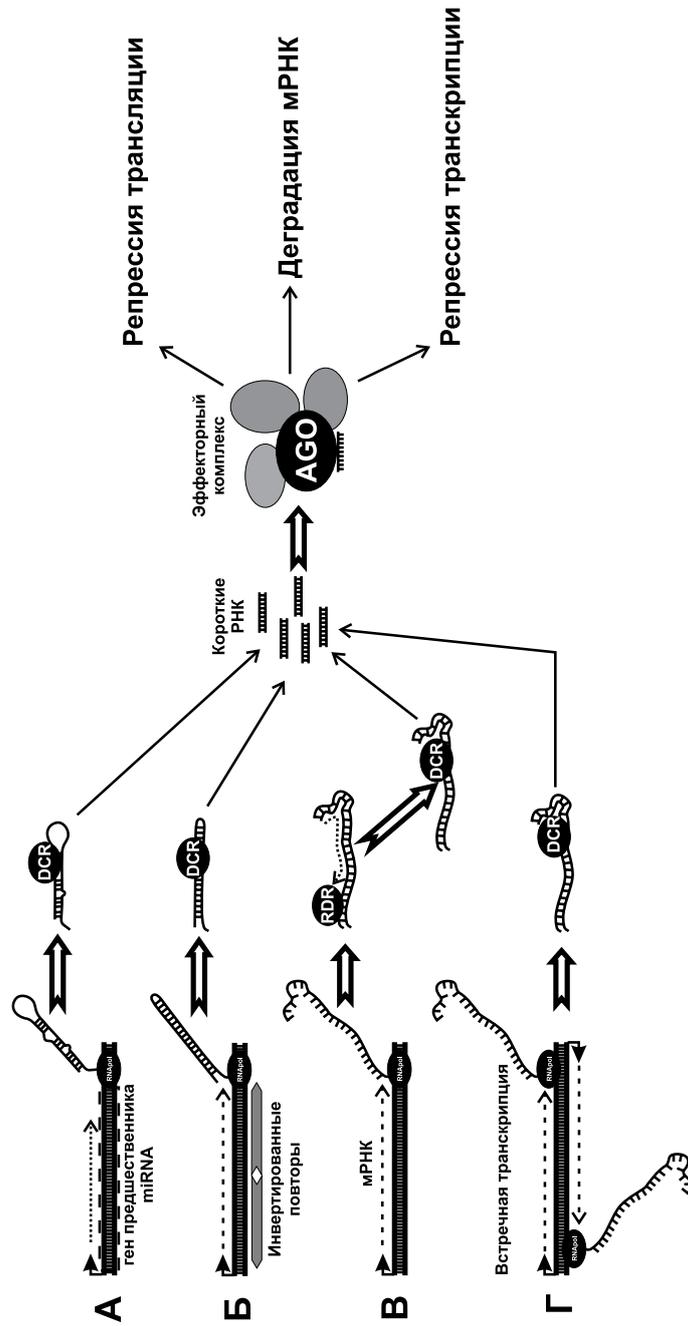


Рис. 4. Схема РНК-интерференции и механизмы образования дцРНК

(А). Транскрипция гена предшественника miRNA (microRNA). В геноме ряда эукариот обнаружены последовательности, транскрибирующиеся в короткие нетранслируемые РНК с развитой вторичной структурой (шпильками). *Dicer* (DCR) расщепляет такие молекулы до miRNA, связывающихся затем с процессорным комплексом аналогично siRNA. (Б). Транскрипция инвертированных повторов (палиндромов). В результате образуется молекула РНК с участками внутренней комплементарности, дающая шпильку. Двухцепочечный участок шпильки расщепляет *Dicer* (DCR), давая siRNA. (В). Транскрипция последовательности с последующей достройкой комплементарной цепи РНК ферментом РНК-зависимой РНК-полимеразой (RDR). (Г). Транскрипция одного участка в смысловой и антисмысловой ориентации с двух промоторов. В результате комплементарного спаривания транскриптов образуется дцРНК.

Эффекторный комплекс связывает siRNA. Затем возможны репрессия транскрипции (RITS), расщепление мРНК (RISC) или репрессия трансляции. В состав эффекторных комплексов обычно входят белки семейства *Argonaute*. На рисунке не указан еще один источник дцРНК – вирусы, так как это источник экзогенного происхождения.

РНК расщепляются до фрагментов размером 22–26 п.н. (siRNA) рибонуклеазами семейства *Dicer* и затем связываются с белковыми комплексами (RISC), в составе которых вызывают деградацию или нарушение трансляции гомологичных мРНК. Белковые комплексы RISC включают в себя белки семейства *Argonaute*, функция siRNA в таких комплексах – распознавание мишени, молекулы мРНК, которая должна подвергнуться расщеплению. Продукты деградации, в свою очередь, могут повторно участвовать в процессе, тем самым обеспечивая эффективную инактивацию генов. В последнее время получены данные о том что siRNA могут приводить к репрессии не только на посттранскрипционном уровне, но и на уровне структуры хроматина [111]. Наиболее изученные системы такого рода – зависимое от РНК формирование гетерохроматина при участии комплекса RITS у дрожжей и опосредованное РНК метилирование ДНК (RdDM) у растений, однако есть указания на участие siRNA в регуляции структуры хроматина у других организмов. В таблице представлены гены, участвующие в процессах формирования гетерохроматина и метилирования ДНК у дрожжей и растений, соответственно.

#### RITS-КОМПЛЕКС ДРОЖЖЕЙ

Почкующиеся дрожжи *Schizosaccharomyces pombe* являются, по-видимому, наиболее простыми эукариотами, у которых имеется *гетерохроматин*. Гетерохроматиновую структуру имеют прицентромерные области, теломеры и участок расположения локуса определения типа спаривания (MAT-локус).

Таблица.  
**Белки, участвующие в процессах РНК-зависимого метилирования ДНК у растений и РНК-зависимой сборки гетерохроматина у дрожжей**

Функция	Дрожжи	Арабидопсис	
Семейство <i>Dicer</i>	DCR1 [120]	DCL3 [147]	Компоненты системы RNAi
Семейство <i>Argonaute</i>	AGO1 [119]	AGO4 [147, 148]	
РНК-зависимая РНК-полимераза	RDP1 [120]	RDR2 [147]	
Компоненты РНК-полимеразы IV	нет	NRPD1a [147]	
	нет	NRPD1b [149]	
	нет	NRPD2a [147]	
PolyA-полимераза	CID12 [120]	?	
РНК-хеликаза	HRR1 [120]	?	
Рибонуклеаза III	ERI1 [124]	?	
Метилтрансфераза гистонов	CLR4 [115]	KYP [155]	Хроматиновые белки
ДНК-метилтрансферазы	нет	MET1 [42]	
	нет	DRM2 [42]	
	нет	CMT3 [42]	
Деацетилаза гистонов	CLR3 [115]	HDA6 [156]	
	CLR6 [114]	?	
	SIR2 [113]	?	
Моделирование хроматина	?	DRD1 [157]	
	?	DDM1 [143]	
Белки хроматина, содержащие хромодомен	SWI6 [158]	?	
	CHP1 [159]	?	
	нет	CMT3 [42]	
Деметилирование ДНК	нет	ROS1 [152]	
	нет	DME [154]	
Неизвестно	TAS3 [123]	?	

Белки сгруппированы по биологической функции. У дрожжей нет РНК-полимеразы IV и белков, связанных с метилированием ДНК, поэтому в соответствующих строчках стоит пометка «нет». Если белок, выполняющий сходную функцию в другом организме, неизвестен – стоит знак «?». В квадратных скобках даны литературные ссылки.

Прицентромерные области дрожжей имеют размер 35–110 т. п. н. и состоят из центральной части, окруженной районами повторов – *imr* (innermost) и *otr* (outermost). Район *otr* содержит tandemные повторы двух типов – *dh* и *dg*. MAT-локус содержит участок размером 3 т. п. н., на 97% гомологичный повторам *dh/dg*. Участок называется *cenH* и область размером примерно 10 т. п. н. вокруг него гетерохроматинизирована. В области теломерных повторов также имеются фрагменты с высокой степенью гомологии с *dh*.

В молекулярной организации гетерохроматина у растений, животных и дрожжей имеются общие черты. У дрожжей, в отличие от растений и млекопитающих, нет метилирования ДНК. При формировании структуры гетерохроматина происходит деацетилирование гистонов деацетилазами CLR3, CLR6, SIR2. Затем гистон H3 метилируется по лизину в девятом положении метилтрансферазой гистонов CLR4 [113–116]. Белок SWI6 (эквивалент HP1 у дрозофилы и млекопитающих) узнает метилированный гистон и связывается с хроматином. Комплекс, включающий SWI6, CHP1, CLR4 способен к олигомеризации и распространяется вдоль хроматина, начиная от точки сборки. Таким образом происходит гетерохроматинизация района размером несколько тысяч пар нуклеотидов [117]. Гены в этой области репрессируются.

Было обнаружено, что формирование гетерохроматина в прицентромерных районах и MAT-локусе происходит с участием компонентов системы РНК-интерференции и siRNA, возникающих при процессинге более длинных транскриптов прицентромерных повторов [118].

У дрожжей имеется по одному гену семейства *Dicer* (*Dcr1*), *Argonaute* (*Ago1*) и РНК-зависимой РНК-полимеразы (*Rdp1*). Делеция любого из этих генов вела к нарушению системы RNAi, а также к нарушению формирования прицентромерного гетерохроматина. В районе центромеры не происходило накопления метилированного по лизину-9 гистона H3, связывания SWI6, а также наблюдалась дерепрессия искусственно введенных трансгенных конструкций. Обнаружено было также накопление полноразмерных транскриптов tandemных повторов *dh/dg* (в норме расщепляемых до siRNA) [118]. Аналогичные наблюдения были сделаны в отношении MAT-локуса, содержащего участок *cenH*, гомологичный прицентромерным повторам.

Биохимическими методами был очищен белковый комплекс, названный RITS [119]. В состав комплекса входят CHP1 (белок, содержащий хромодомен), AGO1, TAS1 (функция неизвестна), а также собственно siRNA.

Компоненты комплекса RITS связаны с прицентромерным хроматином дрожжей, причем это связывание зависит от DICER – а следовательно, от siRNA. Таким образом, простейшая модель предполагает, что siRNA, образующиеся при расщеплении рибонуклеазой DICER транскриптов прицентромерных повторов, входят в состав RITS и обеспечивают его связывание с гомологичными участками генома – такими как сами повторы или *cenH* в локусе типа спаривания. После этого за счет белок – белковых взаимодействий привлекаются SWI6 и метилтрансфераза гистонов CLR4 и инициируется сборка гетерохроматина.

Однако реальная ситуация гораздо сложнее. Помимо RITS, был идентифицирован комплекс RDRC, в состав которого входят RDP1 (РНК-зависимая РНК полимеразы), HRR1 (РНК-хеликаза) и CID12, polyA-полимераза. Комплекс обладает РНК-зависимой РНК-полимеразной активностью и, как и RITS, ассоциирован с хроматином прицентромерных областей. При этом связывание с гетерохроматином зависит от DICER, от компонентов RITS и от CLR4/SWI6. И напротив, мутации по компонентам RDRC нарушают связывание RITS-комплекса и формирование гетерохроматина [120, 121]. Таким образом, для обусловленного RNAi образования гетерохроматина необходим процесс синтеза РНК на РНК.

Далее было обнаружено, что связывание RITS-комплекса с хроматином зависит от метилтрансферазы гистонов CLR4. У мутантов по этому гену очищенный RITS – комплекс не содержал siRNA, и общее количество siRNA из прицентромерных районов уменьшалось [120]. Нарушалось также связывание RDP1 с хроматином. Таким образом, насколько верно что siRNA в комплексе с RITS инициирует сборку гетерохроматина, настолько верно и то, что структурные компоненты гетерохроматина необходимы для сохранения интактного комплекса RITS-RDRC и эффективного расщепления прицентромерных транскриптов до siRNA.

Две работы содержат прямое указание на то, что комплекс RITS связывается не с ДНК, а с новосинтезированной РНК. Было показано, что вызванные RNAi изменения хроматина происходят только если последовательность-мишень, гомологичная siRNA, транскрибируется. При этом компонент RITS белок AGO1 связывается с новосинтезированной РНК, а мутация в RPB2, одной из субъединиц РНК-полимеразы II, приводит к нарушению образования siRNA, метилирования гистона H3 и формирования гетерохроматина. Нарушений собственно транскрипции мутация не вызывает. Это прямое указание на существование физической связи между системой

образованием siRNA и базовой транскрипционной машиной [122].

Кроме того, присоединение одного из компонентов комплекса RITS, белка TAS3, к новосинтезированному транскрипту гена *ura4* при помощи введенного в последовательность TAS3 РНК-связывающего домена, инициирует репрессию гена-мишени, формирование гетерохроматина и появление соответствующих гену-мишени siRNA [123]. Инициация транскрипции *ura4* при этом не нарушается (количество связанной с геном РНК-полимеразы не изменяется).

В работе Бухлера и др. [123] было также продемонстрировано, что появление соответствующих *ura4* siRNA при присоединении TAS3 не приводит к репрессии второй копии *ura4* в геноме. Это подтверждает модель, согласно которой образование siRNA у дрожжей процесс локальный и происходит котранскрипционно. Был обнаружен ген, *Eri1*, мутация в котором приводила к репрессии второй копии *ura4*. *Eri1* (enhancer of RNAi) это высококонсервативная рибонуклеаза, обнаруженная также у нематоды *C. elegance*. У дрожжей мутация по *Eri1* усиливает вызванную RNAi репрессию [124], видимо, нарушая деградацию siRNA.

Таким образом, представляется что у дрожжей образование siRNA (по крайней мере в случае прицентромерных повторов) сопряжено с транскрипцией и осуществляется сложным комплексом белков, выполняющим несколько функций, а именно – расщепление РНК до siRNA, амплификацию РНК с помощью РНК-зависимой РНК-полимеразы, сборку гетерохроматина, и, возможно, расщепление siRNA. Присутствие DICER в составе комплекса не показано, и, возможно, он требуется только для инициации процесса. В то же время, DICER в дрожжах локализован в ядре. Предполагаемая схема РНК-зависимого формирования гетерохроматина представлена на рис. 5.

#### RDDM У РАСТЕНИЙ. ОПОСРЕДОВАННОЕ РНК МЕТИЛИРОВАНИЕ ДНК

Метилирование ДНК – это обратимая ковалентная модификация некоторых оснований (обычно цитозина в симметричных последовательностях CpG, CpNpG или ассиметричных CpHpH, где Н – А, С, Т), широко распространенная у растений и млекопитающих. В то же время у насекомых, дрожжей и нематоды *Caenorabditis elegans* процесс метилирования ДНК либо вообще отсутствует, либо не играет особой роли в регуляции генов.

У растений и млекопитающих метилирование ДНК считается одним из основных механизмов эпигенетического наследования. Обычно метилирование ассоциируется с транскрипционно-неак-

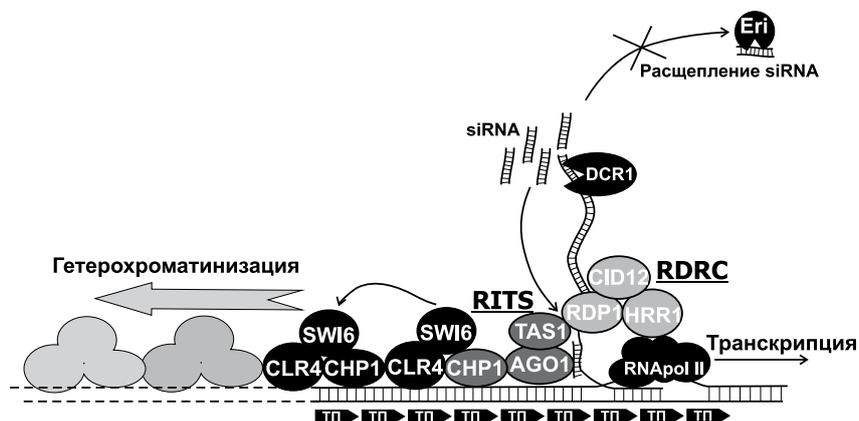


Рис. 5. РНК-зависимое формирование гетерохроматина у дрожжей.

Предполагаемая схема взаимодействий при формировании гетерохроматина прицентромерных районов и репрессии локуса типа спаривания у дрожжей. Транскрипция с прицентромерных прямых tandemных повторов (ТП) приводит к синтезу РНК, которая непосредственно на месте образования конвертируется в дцРНК с участием комплекса RDRC (в составе HRR1, SID12 и РНК-зависимой РНК-полимеразы RDP1). дцРНК немедленно расщепляется рибонуклеазой DCR1 (DICER) до siRNA. DICER у дрожжей имеет ядерную локализацию. siRNA входит в состав комплекса RITS (включает TAS1, CHP1 и AGO1, представитель семейства *Argonaute*). RITS, содержащий siRNA, связывается с транскриптом tandemных повторов за счет гомологичного взаимодействия siRNA и запускает процесс формирования гетерохроматина (возможно, через CHP1, общий компонент для гетерохроматинового комплекса SWI6-CLR4 и RITS). Гетерохроматин затем распространяется от точки инициации за счет самосборки (CLR4 метилирует гистон H3 по лизину в 9-м положении, с модифицированным гистонном связывается SWI6, с которым, в свою очередь, связывается CLR4).

Представляется, что компоненты RITS, RDRC и гетерохроматиновые белки взаимодействуют между собой и формируют единый комплекс. siRNA образуется на месте и здесь же вступает в реакцию. Свободной диффузии (и выключению последовательностей, гомологичных tandemным повторам в других районах генома) препятствует рибонуклеаза *Eri1*. См. также пояснения в тексте.

тивным состоянием хроматина – метилированными оказываются прицентромерный гетерохроматин, инактивированная X-хромосома млекопитающих, различные повторы и мобильные элементы. Метилируются также регуляторные области генов при их инактивации, причем в ряде случаев у растений активное/инактивированное состояние стабильно наследуется в ряду поколений, давая так называемые *эпигенетические аллели*. С метилированием связаны случаи *геномного импринтинга* – различий в уровнях экспрессии аллелей материнского и отцовского происхождения.

Эффективность метилирования как механизма эпигенеза связано с системой поддержания метилированных симметричных последовательностей (CpG), аналогичной механизму полуконсервативной репликации ДНК. После репликации ДНК такие сайты оказываются в «полуметилированном» состоянии, метилазы поддержания (MET1 у арабидопсиса, DNMT1 у млекопитающих) дometилирует их, при этом у растений в процессе участвуют белки, модифицирующие гистоны (HDA6 – деацетилаза гистонов) и модификатор структуры хроматина DDM1. Механизм поддержания несимметричных сайтов более сложен и включает для сайтов вида CpNpG хромометилазу CMT3 (белок одновременно обладает ДНК-метиلاзной активностью и содержит хромодомен), KYP (метилтрансферазу гистонов, метилирует гистон H3 по лизину-9), модификатор структуры хроматина DRD1 (семейство swi/snf) и ДНК-метилазу *de novo* DRM2 [125, 126].

Если за поддержание метилированного состояния отвечает комплекс ДНК-метилаз и хроматиновых белков, то установление репрессии, по современным представлениям, связано с siRNA – продуктами функционирования системы РНК-интерференции. Собственно, у растений впервые был описан случай вызванных siRNA эпигенетических модификаций генома.

В геноме табака вводились трансгенные конструкции, содержащие ДНК, соответствующую РНК вириоида (PSTVd, potato spindle tuber viroid). Было обнаружено, что последовательности вириоида в составе генома метилируются в том случае, когда растение заражается вириодом и вириод проходит цикл автономной репликации РНК ([127]. Таким образом, последовательность ДНК в составе генома метилировалась при наличии в клетке комплементарной двухцепочечной РНК.

Транскрипция конструкции, содержащей участок промотора гена NOS (nopaline synthase) в составе инвертированного повтора приводит к репрессии гена и его метилированию. РНК, получающаяся при транскрипции такой конструкции, содержит шпильку – двухцепочечный участок. Этот участок расщеплялся до siRNA длиной примерно 23 нуклеотида, что соответствует деградации дцРНК по механизму РНК-интерференции [128].

В ходе исследований были обнаружены многочисленные примеры РНК-зависимого метилирования регуляторных последовательностей генов, мобильных элементов, прямых и инвертированных повторов и прицентромерных сателлитных ДНК [42, 111, 128–140].

Система РНК-интерференции у растений оказалась очень сложной и избыточной. У арабидопсиса было обнаружено 4 гена – члена

семейства *Dicer*, 10 членов семейства *Argonaute*. Кроме того, в систему входят РНК-зависимые РНК-полимеразы (6 генов) и специфичная для растений ДНК-зависимая РНК-полимераза IV [141]. У растений имеет место как связанная с системой РНК-интерференции транскрипционная репрессия, так и классическая посттранскрипционная (PTGS – posttranscriptional gene silencing). Более того, разные типы генетических элементов (гены, прицентромерные сателлиты, прямые и инвертированные повторы) имеют разные механизмы транскрипционной репрессии с участием разных белков.

Лучше всего в настоящее время охарактеризована система регуляции метилирования и транскрипционной репрессии гена *FWA* у *Arabidopsis thaliana*. Продукт гена представляет собой транскрипционный фактор, отвечающий за время цветения арабидопсиса. Ген экспрессируется в эндосперме, в прочих тканях неактивен. С неактивным состоянием ассоциировано метилирование участка ДНК, соответствующего области в районе 5'-конца транскрипта гена [142]. Эта область образована ретротранспозоном *SINE3* [143] и содержит две пары прямых повторов, наличие которых необходимо для репрессии. В случае *FWA* имеет место метилирование как симметричных последовательностей CpG, так и несимметричных CpHpH. Деметилирование повторов в тканях взрослого растения приводит к экспрессии гена и, соответственно, «фенотипу позднего цветения», при этом состояние стабильно наследуется в ряду поколений (так называемая эпимутация *fwa*) [142].

Эпимутация *fwa* может возникнуть в силу случайных причин либо быть индуцирована мутациями в генах *DDM1* и *MET1* [142]. *DDM1* представляет собой белок, относящийся к семейству *swi/snf* и осуществляющий АТФ-зависимое изменение структуры хроматина [144]. *MET1* – это ДНК-метилаза, осуществляющая поддерживающее метилирование симметричных последовательностей вида CpG. В мутантах по *DDM1* и *MET1* происходит постепенное деметилирование ряда элементов генома на протяжении нескольких поколений и это приводит к нарушениям регуляции генов и разнообразным фенотипическим эффектам [145].

В геном арабидопсиса при помощи бактериальной трансформации можно ввести конструкцию, содержащую дополнительную копию гена *FWA*. Трансген интегрируется в геном и в норме практически со 100% вероятностью оказывается метилирован и репрессирован. Таким образом, трансформация растений дикого типа геном *FWA* не приводит к «фенотипу позднего цветения» – и эндогенная и трансгенная копии неактивны.

Очевидно, что попадающий в геном трансген для репрессии должен быть метилирован заново. Если продукты каких-либо генов арабидопсиса необходимы для метилирования и репрессии *FWA*, то при введении трансгенной конструкции в мутантов по этим генам инактивации не произойдет. В ходе основанных на таком подходе скринингов был обнаружен ряд необходимых для метилирования *de novo* генов. В их число входят:

*DRM2* – ДНК – метилтрансфераза, необходимая для метилирования *de novo* симметричных (CpG) и несимметричных последовательностей.

*DRD1* – Белок относится к семейству swi/snf белков, моделирующей структуру хроматина (chromatin remodeling protein).

*AGO4* – Один из представителей семейства *Argonaute*. Белки этого семейства представлены у всех эукариот, способны связывать короткие РНК (например, siRNA) и входят в состав эффекторных комплексов (RITS, RISC).

*DCL3* – Один из представителей семейства *Dicer*. Рибонуклеаза, осуществляющая разрезание длинных дцРНК до siRNA.

*RDR2* – РНК-зависимая РНК-полимераза. Синтез второй цепи РНК по РНК-матрице приводит к образованию дцРНК, которая затем разрезается до siRNA.

*NRPD1a*, *NRPD1b* – субъединицы специфичной для растений ДНК-зависимой РНК-полимеразы IV (RNAPol IV). RNAPol IV у растений транскрибирует повторы и необходима для репрессии мобильных элементов и прочих типов повторенных последовательностей. Существует в двух вариантах – комплекс *NRPD1a*–*NRPD2* (RNAPol IVa) и комплекс *NRPD1b*–*NRPD2* (RNAPol IVb), при этом *NRPD2* – общая субъединица для обоих типов комплексов. Комплексы различаются по их функциям – RNAPol IVa необходима для синтеза siRNA, а RNAPol IVb – для метилирования и транскрипционной репрессии.

Помимо этого, было обнаружено что с прямых повторов 5'-области гена *FWA* образуется siRNA [143, 146].

Таким образом, для метилирования и инактивации трансгена *FWA de novo*, помимо модификатора структуры хроматина *DRD1* и метилтрансферазы *DRM2*, также требуются компоненты системы РНК-интерференции растений. Это прямо указывает на связь между RNAi и транскрипционной репрессией [147]. Случай с *FWA* не единственный, *AGO4* участвует в инактивации гена *SUPERMAN* [148].

Дальнейшие генетические и молекулярно-биологические исследования *FWA* позволили понять последовательность событий, приводящих к репрессии. В работе Чана и др. [149] было показано,

что DCL3, NRPD1a и RDR2 необходимы для образования siRNA с повторов *FWA*. NRPD1b, AGO4, DRD1 и DRM2 действуют на следующей стадии и отвечают за метилирование и репрессию трансгена.

Исследования методом иммуноокрашивания ядер позволили определить внутриядерную локализацию компонентов, отвечающих за RdDM. Оказалось, что DCL3, RDR2, NRPD1b и AGO4 локализуются в тельце Кохаля – расположенном недалеко от ядрышка образовании, отвечающем за процессинг ряда форм РНК. DRD1, DRM2, NRPD1a обнаруживаются на хроматине в районах расположения повторов. NRPD1b и AGO4, помимо тельца Кохаля, также обнаруживаются на хроматине.

С учетом имеющихся данных возможна следующая схема RdDM для *FWA* (рис. 6). В норме *FWA* инактивирован и его ген метилирован в районе прямых повторов. Метилированы как симметричные (CpG), так и несимметричные (CpHpH) сайты. После цикла репликации ДНК симметричные сайты воспроизводятся по полуконсервативному механизму с участием комплекса MET1–DDM1–HDA6, где MET1 – метилаза поддержания, DDM1 – белок, моделирующий структуру хроматина, а HDA6 – деацетилаза гистонов. Этот механизм независим от RNAi. Несимметричные сайты необходимо метилировать после каждого цикла удвоения ДНК, и процесс схож с метилированием *de novo* введенной в геном трансгенной конструкции.

Для метилирования *de novo* требуется система, состоящая из метилазы (DRM2), siRNA к регулируемой области и компонентов системы RNAi. Представляется, что RNAPol IVa транскрибирует прямые повторы в 5'-области *FWA*, после чего образующиеся молекулы РНК транспортируются в тельце Кохаля, где подвергаются обратной транскрипции с помощью RDR2 и нарезаются до siRNA посредством DCL3. siRNA связываются с комплексом, содержащим AGO4 и NRPD1b. Комплекс затем транспортируется к регулируемому гену и индуцирует метилирование и репрессию – с участием метилазы DRM2 и хроматинового белка DRD1. siRNA в составе комплекса нужна для правильной локализации репрессирующей активности.

Предполагается, что siRNA нацеливает комплекс за счет узнавания комплементарных последовательностей, в случае *FWA* – прямых повторов в 5'-области гена. Процесс комплементарного спаривания требует присутствия мишени (РНК или ДНК) в одноцепочечной форме. Соответственно, роль RNAPol IVb может быть связана либо с локальным расплетением ДНК, либо с синтезом транскрипта, с которым и связывается siRNA-содержащий комплекс. При введении трансгена в растения его ДНК присутствует в одноцепочечной форме, возможно, это упрощает связывание репрессирующего комплекса и обеспечивает полную инактивацию.

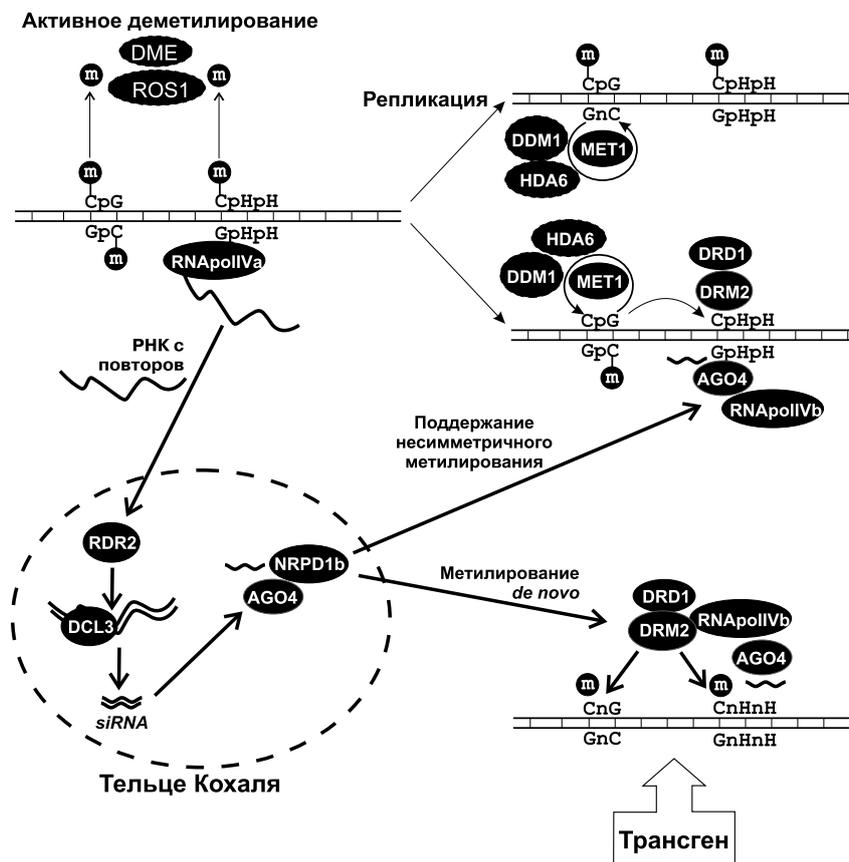


Рис. 6. РНК-зависимое метилирование ДНК у арабидопсиса на примере гена *FWA*. В метилировании несимметричных сайтов (CpHpH) после цикла репликации ДНК или при введении трансгенной конструкции участвуют белки системы РНК-интерференции RDR2, DCL3 и AGO4, а также siRNA прямых повторов из 5'-области гена *FWA*. Подробное описание в тексте.

Репрессия *FWA* – наиболее детально исследованный пример RdDM у растений: идентифицированы белковые факторы, вовлеченные в процесс и имеется информация о их биохимических функциях. Вместе с тем, детали сборки репрессирующего комплекса на гени-мишени неизвестны. Кроме того, исследования растений, несущих стабильную эпимутацию *fwa*, показали, что у таких растений производится siRNA к повторам из 5'-области гена. [149]. *fwa*, как указывалось выше, представляет собой неметилированный и активно транскрибирующийся ген *FWA*, такое состояние стабильно

наследуется в растениях дикого типа, содержащих все компоненты системы репрессии. Следовательно, только присутствия siRNA к повторам недостаточно для инактивации, ген должен обладать «компетентностью» к сигналам RNAi. В случае трансгенного *FWA* «компетентность» приобретается в ходе процесса трансформации и интеграции в геном, а поддерживающее метилирование несимметричных сайтов по механизму RNAi зависит от метилирования симметричных сайтов [149].

Наконец, в эндосперме имеет место процесс активации *FWA*, связанный с деметилированием [150]. Деметилирование ДНК происходит в результате активности ДНК-гликозилаз, из которых у арабидопсиса охарактеризованы DEMETER и ROS1 (Repressor of silencing 1) [150–154]. После деметилирования *FWA* экспрессируется в эндосперме, но так как эндосперм не становится частью взрослого растения, активированное состояние не передается потомкам.

## V. ЗАКЛЮЧЕНИЕ

За последние два десятилетия существенно изменились представления о механизмах регуляции транскрипции у эукариот. Классическая модель предполагала, что уровень экспрессии определяется транскрипционными факторами, связывающимися с регуляторными областями гена. Факторы напрямую взаимодействуют с компонентами аппарата транскрипции, инициируя синтез матричной РНК. Хроматину (совокупности связанных с ДНК гистонов и структурных негистоновых белков) отводилась пассивная роль – в основном обеспечения компактной укладки нити ДНК в ядре.

В дальнейшем была установлена активная роль хроматина в регуляции транскрипции. Был обнаружен так называемый гистоновый код – совокупность ковалентных модификаций гистонов, неодинаковая в разных районах генома (к примеру, в гетерохроматине и активно работающих генах). Гистоновый код интерпретируется белковыми комплексами, способными менять структуру хроматина и делать участки генома доступными для транскрипционных факторов либо, напротив, переводить их в компактное неактивное состояние. В качестве примера можно привести комплекс гетерохроматиновых белков, сборка которого приводит к репрессии генов и компактизации хроматина затронутой области.

В последние годы было обнаружено, что в ряде случаев сборка и поддержание меняющих структуру хроматина белковых комплексов происходит с участием некодирующих РНК. Гетерохроматинизация X-хромосомы млекопитающих для достижения дозовой компенсации

требует *Xist*, увеличение уровня транскрипции в два раза с X-хромосомы самцов дрозофилы – *roX*. Формирование гетерохроматина дрожжей происходит при участии siRNA, возникающих при расщеплении транскриптов прицентромерных прямых повторов, репрессия некоторых генов растений также запускается и поддерживается siRNA.

Во всех случаях функция некодирующих РНК, по-видимому – целеуказание, определение районов генома, с которыми должен связаться комплекс модифицирующих структуру хроматина белков. Специфичность узнавания достигается за счет гомологичных взаимодействий между РНК в составе комплекса и последовательностью мишени – ДНК или свежееобразованной РНК. Далее происходит модификация гистонов, метилирование ДНК, компактизация хроматина – изменения, определяемые свойствами белкового комплекса, ассоциированного с РНК. В случае инактивации X-хромосомы млекопитающих, RdDM и сборки гетерохроматина у дрожжей имеет место репрессия; при гиперактивации X-хромосомы дрозофилы – активация транскрипции.

Если РНК (например, siRNA у растений) участвует в репрессии последовательности, с которой она образовалась, то возникает противоречие – для эффективной репрессии оказывается необходима активная транскрипция. По-видимому, существует несколько путей разрешения ситуации. У растений имеется дополнительная (к трем у прочих эукариот) РНК-полимераза IV, которая транскрибирует повторы и, возможно, гетерохроматинизированное состояние не является препятствием для ее активности. Далее, опять же у растений, репрессированное состояние гена может сохраняться за счет поддерживающего метилирования симметричных повторов. И, наконец, участие в siRNA-опосредованной репрессии РНК-зависимой РНК-полимеразы и у растений, и у дрожжей указывает на возможность существования «контура амплификации» – системы, воспроизводящей siRNA без транскрипции с ДНК.

В обзоре подробно рассмотрены лишь четыре примера влияния РНК на структуру хроматина, однако, к примеру, дозовая компенсация затрагивает сотни и тысячи генов. Реальные масштабы связанных с некодирующими РНК модификаций структуры хроматина пока не поддаются оценке, но, несомненно, весьма велики. К примеру, у дрозофилы сотни мишеней действия белков группы *Polycomb* и компонентов гетерохроматина, а оба типа модифицирующих структуру хроматина комплексов, по-видимому, зависят от некодирующих РНК. Дальнейшие исследования позволят получить новую информацию о чрезвычайно сложных механизмах регуляции экспрессии генов эукариот.

## ЛИТЕРАТУРА

1. Brown, C.J., Ballabio, A., Rupert, J.L., Lafreniere, R.G., Grompe, M., Tonlorenzi, R., Willard, H.F. (1991) *Nature*, **349**, 38–44.
2. Yao, M.C., Fuller, P., Xi, X. (2003) *Science*, **300**, 1581–1584.
3. Selker, E.U. (2003) *Science*, **300**, 1517–1518.
4. Morris, K.V., Chan, S.W., Jacobsen, S.E., Looney, D.J. (2004) *Science*, **305**, 1289–1292.
5. Kawasaki, H., Taira, K. (2004) *Nature*, **431**, 211–217.
6. Fukagawa, T., Nogami, M., Yoshikawa, M., Ikeno, M., Okazaki, T., Takami, Y., Nakayama, T., Oshimura, M. (2004) *Nat. Cell Biol.*, **6**, 784–791.
7. Pal-Bhadra, M., Leibovitch, B.A., Gandhi, S.G., Rao, M., Bhadra, U., Birchler, J.A., Elgin, S.C. (2004) *Science*, **303**, 669–672.
8. Wolffe, A.P. (1994) *Trends Biochem. Sci.*, **19**, 240–244.
9. Gregory, P.D., Horz, W. (1998) *Eur. J. Biochem.*, **251**, 9–18.
10. Marendza, D.R., Zraly, C.B., Dingwall, A.K. (2004) *Dev. Biol.*, **267**, 279–293.
11. Peterson, C.L. (1996) *Curr. Opin. Genet. Dev.*, **6**, 171–175.
12. Wade, P.A., Geggion, A., Jones, P.L., Ballestar, E., Aubry, F., Wolffe, A.P. (1999) *Nat. Genet.*, **23**, 62–66.
13. Katsani, K.R., Mahmoudi, T., Verrijzer, C.P. (2003) *Curr. Top Microbiol. Immunol.*, **274**, 113–141.
14. Jenuwein, T., Allis, C.D. (2001) *Science*, **293**, 1074–1080.
15. Thiagalingam, S., Cheng, K.H., Lee, H.J., Mineva, N., Thiagalingam, A., Ponte, J.F. (2003) *Ann. N. Y. Acad. Sci.*, **983**, 84–100.
16. Turner, B.M. (2002) *Cell*, **111**, 285–291.
17. Braunstein, M., Sobel, R.E., Allis, C.D., Turner, B.M., Broach, J.R. (1996) *Mol. Cell Biol.*, **16**, 4349–4356.
18. Tsukada, Y., Fang, J., Erdjument-Bromage, H., Warren, M.E., Borchers, C.H., Tempst, P., Zhang, Y. (2006) *Nature*, **439**, 811–816. Epub. 2005 Dec 2018.
19. Ebert, A., Schotta, G., Lein, S., Kubicek, S., Krauss, V., Jenuwein, T., Reuter, G. (2004) *Genes Dev.*, **18**, 2973–2983.
20. Schotta, G., Ebert, A., Dorn, R., Reuter, G. (2003) *Semin. Cell Dev. Biol.*, **14**, 67–75.
21. Weiler, K.S., Wakimoto, B.T. (2002) *Mol. Genet. Genomics*, **266**, 922–932.
22. Czermin, B., Schotta, G., Hulsman, B.B., Brehm, A., Becker, P.B., Reuter, G., Imhof, A. (2001) *EMBO Rep.*, **2**, 915–919.
23. Schotta, G., Ebert, A., Krauss, V., Fischer, A., Hoffmann, J., Rea, S., Jenuwein, T., Dorn, R., Reuter, G. (2002) *EMBO J.*, **21**, 1121–1131.
24. Min, J., Zhang, Y., Xu, R.M. (2003) *Genes Dev.*, **17**, 1823–1828.
25. Fischle, W., Wang, Y., Jacobs, S.A., Kim, Y., Allis, C.D., Khorasanizadeh, S. (2003) *Genes Dev.*, **17**, 1870–1881.
26. Brehm, A., Tufteland, K.R., Aasland, R., Becker, P.B. (2004) *Bioessays*, **26**, 133–140.
27. Akhtar, A., Zink, D., Becker, P.B. (2000) *Nature*, **407**, 405–409.
28. Bannister, A.J., Zegerman, P., Partidge, J.F., Miska, E.A., Thomas, J.O., Allshire, R.C., Kouzarides, T. (2001) *Nature*, **410**, 120–124.
29. Shareef, M.M., Badugu, R., Kellum, R. (2003) *Genetica*, **117**, 127–134.
30. Seum, C., Pauli, D., Delattre, M., Jaquet, Y., Spierer, A., Spierer, P. (2002) *Genetics*, **161**, 1125–1136.
31. Pirrotta, V., Poux, S., Melfi, R., Pilyugin, M. (2003) *Genetica*, **117**, 191–197.
32. Simon, J.A. (2003) *Curr. Biol.*, **13**, R79–80.

33. Pirrotta, V. (1995) *Curr. Opin. Genet. Dev.*, **5**, 466–472.
34. Chanas, G., Lavrov, S., Iral, F., Cavalli, G., Maschat, F. (2004) *Dev. Biol.*, **272**, 522–535.
35. Ringrose, L., Rehmsmeier, M., Dura, J.M., Paro, R. (2003) *Dev. Cell*, **5**, 759–771.
36. Negre, N., Henmetin, J., Sun, L.V., Lavrov, S., Bellis, M., White, K.P., Cavalli, G. (2006) *PLoS Biol.*, **4**, e170.
37. Sewalt, R.G., Kwaks, T.H., Hamer, K., Otte, A.P. (2004) *Methods Enzymol.*, **377**, 282–296.
38. Pasini, D., Bracken, A.P., Helin, K. (2004) *Cell Cycle*, **3**, 396–400.
39. Zhang, H., Azevedo, R.B., Lints, R., Doyle, C., Teng, Y., Haber, D., Emmons, S.W. (2003) *Dev. Cell*, **4**, 903–915.
40. Reyes, J.C., Grossniklaus, U. (2003) *Semin. Cell Dev. Biol.*, **14**, 77–84.
41. Weissmann, F., Muyrers-Chen, I., Musch, T., Stach, D., Wiessler, M., Paro, R., Lyko, F. (2003) *Mol. Cell Biol.*, **23**, 2577–2586.
42. Chan, S.W., Henderson, I.R., Jacobsen, S.E. (2005) *Nat. Rev. Genet.*, **6**, 351–360.
43. Liu, K., Wang, Y.F., Cantemir, C., Muller, M.T. (2003) *Mol. Cell Biol.*, **23**, 2709–2719.
44. Zemach, A., Grafi, G. (2007) *Trends Plant. Sci.*, **12**, 80–85.
45. Fatemi, M., Wade, P.A. (2006) *J Cell Sci*, **119**, 3033–3037.
46. Wade, P.A. (2001) *Bioessays*, **23**, 1131–1137.
47. Huyen, K.D., Lee, J.T. (2003) *Nature*, **426**, 857–862.
48. Okamoto, I., Otte, A.P., Allis, C.D., Reinberg, D., Heard, E. (2004) *Science*, **303**, 644–649.
49. Avner, P., Heard, E. (2001) *Nat. Rev. Genet.*, **2**, 59–67.
50. Penny, G.D., Kay, G.F., Sheardown, S.A., Rastan, S., Brockdorff, N. (1996) *Nature*, **379**, 131–137.
51. Marahrens, Y., Panning, B., Dausman, J., Strauss, W., Jaenisch, R. (1997) *Genes Dev.*, **11**, 156–166.
52. Lee, J.T., Jaenisch, R. (1997) *Nature*, **386**, 275–279.
53. Brockdorff, N., Ashworth, A., Kay, G.F., Cooper, P., Smith, S., McCabe, V.M., Norris, D.P., Penny, G.D., Patel, D., Rastan, S. (1991) *Nature*, **351**, 329–331.
54. Lee, J.T., Davidow, L.S., Warshawsky, D. (1999) *Nat. Genet.*, **21**, 400–404.
55. Sado, T., Wang, Z., Sasaki, H., Li, E. (2001) *Development*, **128**, 1275–1286.
56. Herzing, L.B., Romer, J.T., Horn, J.M., Ashworth, A. (1997) *Nature*, **386**, 272–275.
57. Sun, B.K., Deaton, A.M., Lee, J.T. (2006) *Mol. Cell*, **21**, 617–628.
58. Marahrens, Y., Loring, J., Jaenisch, R. (1998) *Cell*, **92**, 657–664.
59. Boggs, B.A., Cheung, P., Heard, E., Spector, D.L., Chinault, A.C., Allis, C.D. (2002) *Nat. Genet.*, **30**, 73–76.
60. Silva, J., Mak, W., Zvetkova, I., Appanah, R., Nesterova, T.B., Webster, Z., Peters, A.H., Jenuwein, T., Otte, A.P., Brockdorff, N. (2003) *Dev. Cell*, **4**, 481–495.
61. Chadwick, B.P., Willard, H.F. (2003) *Hum. Mol. Genet.*, **12**, 2167–2178.
62. Rasmussen, T.P., Mastrangelo, M.A., Eden, A., Pehrson, J.R., Jaenisch, R. (2000) *J. Cell Biol.*, **150**, 1189–1198.
63. Csankovszki, G., Nagy, A., Jaenisch, R. (2001) *J. Cell Biol.*, **153**, 773–784.
64. Mak, W., Baxter, J., Silva, J., Newall, A.E., Otte, A.P., Brockdorff, N. (2002) *Curr. Biol.*, **12**, 1016–1020.
65. Plath, K., Fang, J., Mlynarczyk-Evans, S.K., Cao, R., Worringer, K.A., Wang, H., De La Cruz, C.C., Otte, A.P., Panning, B., Zhang, Y. (2003) *Science*, **300**, 131–135.
66. Csankovszki, G., Panning, B., Bates, B., Pehrson, J.R., Jaenisch, R. (1999) *Nat. Genet.*, **22**, 323–324.
67. Wutz, A., Rasmussen, T.P., Jaenisch, R. (2002) *Nat. Genet.*, **30**, 167–174.

68. Wutz, A., Jaenisch, R. (2000) *Mol. Cell*, **5**, 695–705.
69. Hendrich, B.D., Brown, C.J., Willard, H.F. (1993) *Hum. Mol. Genet.*, **2**, 663–672.
70. Nesterova, T.B., Slobodyanyuk, S.Y., Elisaphenko, E.A., Shevchenko, A.I., Johnston, C., Pavlova, M.E., Rogozin, I.B., Kolesnikov, N.N., Brockdorff, N., Zakian, S.M. (2001) *Genome Res.*, **11**, 833–849.
71. Mak, W., Nesterova, T.B., De Napolles, M., Appanah, R., Yamanaka, S., Otte, A.P., Brockdorff, N. (2004) *Science*, **303**, 666–669.
72. Luikenhuis, S., Wutz, A., Jaenisch, R. (2001) *Mol. Cell Biol.*, **21**, 8512–8520.
73. Stavropoulos, N., Lu, N., Lee, J.T. (2001) *Proc. Natl. Acad. Sci. USA*, **98**, 10232–10237.
74. Tada, T., Obata, Y., Tada, M., Goto, Y., Nakatsuji, N., Tan, S., Kono, T., Takagi, N. (2000) *Development*, **127**, 3101–3105.
75. Norris, D.P., Patel, D., Kay, G.F., Penny, G.D., Brockdorff, N., Sheardown, S.A., Rastan, S. (1994) *Cell*, **77**, 41–51.
76. Ariel, M., Robinson, E., Mccarrey, J.R., Cedar, H. (1995) *Nat. Genet.*, **9**, 312–315.
77. Zuccotti, M., Monk, M. (1995) *Nat. Genet.*, **9**, 316–320.
78. Mcdonald, L.E., Kay, G.F. (1997) *Biotechniques*, **22**, 272–274.
79. Sado, T., Okano, M., Li, E., Sasaki, H. (2004) *Development*, **131**, 975–982.
80. Chen, T., Ueda, Y., Dodge, J.E., Wang, Z., Li, E. (2003) *Mol. Cell Biol.*, **23**, 5594–5605.
81. Mcdonald, L.E., Paterson, C.A., Kay, G.F. (1998) *Genomics*, **54**, 379–386.
82. Sado, T., Tada, T., Takagi, N. (1996) *Dev. Dyn.*, **205**, 421–434.
83. Lee, J.T. (2002) *Nat. Genet.*, **32**, 195–200.
84. Shibata, S., Lee, J.T. (2003) *Hum. Mol. Genet.*, **12**, 125–136.
85. Migeon, B.R., Chowdhury, A.K., Dunston, J.A., Mcintosh, I. (2001) *Am. J. Hum. Genet.*, **69**, 951–960.
86. Chow, J.C., Hall, L.L., Clemson, C.M., Lawrence, J.B., Brown, C.J. (2003) *Genomics*, **82**, 309–322.
87. Ogawa, Y., Lee, J.T. (2003) *Mol. Cell*, **11**, 731–743.
88. Stavropoulos, N., Rowntree, R.K., Lee, J.T. (2005) *Mol. Cell Biol.*, **25**, 2757–2769.
89. Jin, Y., Wang, Y., Johansen, J., Johansen, K.M. (2000) *J. Cell Biol.*, **149**, 1005–1010.
90. Smith, E.R., Pannuti, A., Gu, W., Steurnagel, A., Cook, R.G., Allis, C.D., Lucchesi, J.C. (2000) *Mol. Cell Biol.*, **20**, 312–318.
91. Akhtar, A., Becker, P.B. (2000) *Mol. Cell*, **5**, 367–375.
92. Jin, Y., Wang, Y., Walker, D.L., Dong, H., Conley, C., Johansen, J., Johansen, K.M. (1999) *Mol. Cell*, **4**, 129–135.
93. Wang, Y., Zhang, W., Jin, Y., Johansen, J., Johansen, K.M. (2001) *Cell*, **105**, 433–443.
94. Kelley, R.L., Kuroda, M.I. (2000) *Cell*, **103**, 9–12.
95. Meller, V.H., Wu, K.H., Roman, G., Kuroda, M.I., Davis, R.L. (1997) *Cell*, **88**, 445–457.
96. Meller, V.H., Rattner, B.P. (2002) *EMBO J.*, **21**, 1084–1091.
97. Kelley, R.L., Meller, V.H., Gordadze, P.R., Roman, G., Davis, R.L., Kuroda, M.I. (1999) *Cell*, **98**, 513–522.
98. Franke, A., Dernburg, A., Bashaw, G.J., Baker, B.S. (1996) *Development*, **122**, 2751–2760.
99. Rastelli, L., Richman, R., Kuroda, M.I. (1995) *Mech. Dev.*, **53**, 223–233.
100. Gebauer, F., Grskovic, M., Hentze, M.W. (2003) *Mol. Cell*, **11**, 1397–1404.
101. Palmer, M.J., Richman, R., Richter, L., Kuroda, M.I. (1994) *Genes Dev.*, **8**, 698–706.
102. Kelley, R.L., Solovyeva, I., Lyman, L.M., Richman, R., Solovyev, I.

- V., Kuroda, M.I. (1995) *Cell*, **81**, 867–877.
103. Lyman, L.M., Copps, K., Rastelli, L., Kelley, R.L., Kuroda, M.I. (1997) *Genetics*, **147**, 1743–1753.
104. Meller, V.H. (2003) *Mech. Dev.*, **120**, 759–767.
105. Meller, V.H., Gordadze, P.R., Park, Y., Chu, X., Stuckenholtz, C., Kelley, R.L., Kuroda, M.I. (2000) *Curr. Biol.*, **10**, 136–143.
106. Franke, A., Baker, B.S. (1999) *Mol. Cell*, **4**, 117–122.
107. Stuckenholtz, C., Meller, V.H., Kuroda, M.I. (2003) *Genetics*, **164**, 1003–1014.
108. Kageyama, Y., Mengus, G., Gilfillan, G., Kennedy, H.G., Stuckenholtz, C., Kelley, R.L., Becker, P.B., Kuroda, M.I. (2001) *EMBO J.*, **20**, 2236–2245.
109. Carthew, R.W. (2001) *Curr Opin Cell Biol*, **13**, 244–248.
110. Tijsterman, M., Plasterk, R.H. (2004) *Cell*, **117**, 1–3.
111. Meister, G., Tuschl, T. (2004) *Nature*, **431**, 343–349.
112. Matzke, M.A., Birchler, J.A. (2005) *Nat Rev Genet*, **6**, 24–35.
113. Shankaranarayana, G.D., Motamedi, M.R., Moazed, D., Grewal, S.I. (2003) *Curr. Biol.*, **13**, 1240–1246.
114. Grewal, S.I., Bonaduce, M.J., Klar, A.J. (1998) *Genetics*, **150**, 563–576.
115. Nakayama, J., Rice, J.C., Strahl, B.D., Allis, C.D., Grewal, S.I. (2001) *Science*, **292**, 110–113.
116. Partridge, J.F., Borgstrom, B., Allshire, R.C. (2000) *Genes Dev.*, **14**, 783–791.
117. Grewal, S.I., Moazed, D. (2003) *Science*, **301**, 798–802.
118. Volpe, T.A., Kidner, C., Hall, I.M., Teng, G., Grewal, S.I., Martienssen, R.A. (2002) *Science*, **297**, 1833–1837.
119. Verdel, A., Jia, S., Gerber, S., Sugiyama, T., Gygi, S., Grewal, S.I., Moazed, D. (2004) *Science*, **303**, 672–676.
120. Motamedi, M.R., Verdel, A., Colmenares, S.U., Gerber, S.A., Gygi, S.P., Moazed, D. (2004) *Cell*, **119**, 789–802.
121. Sugiyama, T., Cam, H., Verdel, A., Moazed, D., Grewal, S.I. (2005) *Proc. Natl. Acad. Sci. USA*, **102**, 152–157.
122. Schramke, V., Sheedy, D.M., Denli, A.M., Bonila, C., Ekwall, K., Hannon, G.J., Allshire, R.C. (2005) *Nature*, **435**, 1275–1279.
123. Buhler, M., Verdel, A., Moazed, D. (2006) *Cell*, **125**, 873–886.
124. Iida, T., Kawaguchi, R., Nakayama, J. (2006) *Curr. Biol.*, **16**, 1459–1464.
125. Lindroth, A.M., Shultis, D., Jasencakova, Z., Fuchs, J., Johnson, L., Schubert, D., Patnaik, D., Pradhan, S., Goodrich, J., Schubert, I., Jenuwein, T., Khorasanizadeh, S., Jacobsen, S.E. (2004) *EMBO J.*, **23**, 4286–4296.
126. Tran, R.K., Zilberman, D., De Bustos, C., Ditt, R.F., Henikoff, J.G., Lindroth, A.M., Delrow, J., Boyle, T., Kwong, S., Bryson, T.D., Jacobsen, S.E., Henikoff, S. (2005) *Genome Biol*, **6**, R90.
127. Wassenegger, M., Heimes, S., Riedel, L., Sanger, H.L. (1994) *Cell*, **76**, 567–576.
128. Mette, M.F., Aufsatz, W., Van Der Winden, J., Matzke, M.A., Matzke, A.J. (2000) *EMBO J.*, **19**, 5194–5201.
129. Aufsatz, W., Mette, M.F., Van Der Winden, J., Matzke, A.J., Matzke, M. (2002) *Proc. Natl. Acad. Sci. USA*, **99 Suppl 4**, 16499–16506.
130. Beclin, C., Boutet, S., Waterhouse, P., Vaucheret, H. (2002) *Curr. Biol.*, **12**, 684–688.
131. Jones, L., Ratcliff, F., Baulcombe, D.C. (2001) *Curr. Biol.*, **11**, 747–757.
132. Kawasaki, H., Taira, K. (2005) *Curr. Opin. Mol. Ther.*, **7**, 125–131.

133. Lippman, Z., May, B., Jordan, C., Singer, T., Martienssen, R. (2003) PLoS Biology, **1**, E67.
134. May, B.P., Lippman, Z.B., Fang, Y., Spector, D.L., Martienssen, R.A. (2005) PLoS Genetics, **1**, e79.
135. Melquist, S., Bender, J. (2003) Genes Dev., **17**, 2036–2047.
136. Pelissier, T., Wassenegger, M. (2000) RNA, **6**, 55–65.
137. Qin, H., Dong, Y., Von Arnim, A.G. (2003) Plant Mol. Biol., **52**, 217–231.
138. Sijen, T., Vijn, I., Rebocho, A., Van Blokland, R., Roelofs, D., Mol, J.N., Kooter, J.M. (2001) Curr. Biol., **11**, 436–440.
139. Steimer, A., Amedeo, P., Afsar, K., Fransz, P., Mittelsten Scheid, O., Paszkowski, J. (2000) Plant Cell, **12**, 1165–1178.
140. Vaucheret, H., Fagard, M. (2001) Trends Genet., **17**, 29–35.
141. Herr, A.J., Jensen, M.B., Dalmay, T., Baulcombe, D.C. (2005) Science, **308**, 118–120.
142. Soppe, W.J., Jacobsen, S.E., Alonso-Blanco, C., Jackson, J.P., Kakutani, T., Koornneef, M., Peeters, A.J. (2000) Mol. Cell, **6**, 791–802.
143. Lippman, Z., Gendrel, A.V., Black, M., Vaughn, M.W., Dedhia, N., McCombie, W.R., Lavine, K., Mital, V., May, B., Kasschau, K.D., Carrington, J.C., Doerge, R.W., Colot, V., Martienssen, R. (2004) Nature, **430**, 471–476.
144. Jeddelloh, J.A., Stokes, T.L., Richards, E.J. (1999) Nat. Genet., **22**, 94–97.
145. Kakutani, T., Jeddelloh, J.A., Flowers, S.K., Munakata, K., Richards, E.J. (1996) Proc. Natl. Acad. Sci. USA, **93**, 12406–12411.
146. Lu, C., Tej, S.S., Luo, S., Haudenschild, C.D., Meyers, B.C., Green, P.J. (2005) Science, **309**, 1567–1569.
147. Chan, S.W., Zilberman, D., Xie, Z., Johansen, L.K., Carrington, J.C., Jacobsen, S.E. (2004) Science, **303**, 1336.
148. Zilberman, D., Cao, X., Jacobsen, S.E. (2003) Science, **299**, 716–719.
149. Chan, S.W.-L., Zhang, X., Bernatavichute, Y.V., Jacobsen, S.E. (2006) PLoS Biology, **4**, e363.
150. Kinoshita, T., Miura, A., Choi, Y., Kinoshita, Y., Cao, X., Jacobsen, S.E., Fischer, R.L., Kakutani, T. (2004) Science, **303**, 521–523.
151. Kapoor, A., Agarwal, M., Andreucci, A., Zheng, X., Gong, Z., Hasegawa, P.M., Bressan, R.A., Zhu, J.K. (2005) Curr. Biol., **15**, 1912–1918.
152. Gong, Z., Morales-Ruiz, T., Ariza, R.R., Roldan-Arjona, T., David, L., Zhu, J.K. (2002) Cell, **111**, 803–814.
153. Choi, Y., Gehring, M., Johnson, L., Hannon, M., Harada, J.J., Goldberg, R.B., Jacobsen, S.E., Fischer, R.L. (2002) Cell, **110**, 33–42.
154. Choi, Y., Harada, J.J., Goldberg, R.B., Fischer, R.L. (2004) Proc. Natl. Acad. Sci. USA, **101**, 7481–7486.
155. Jackson, J.P., Lindroth, A.M., Cao, X., Jacobsen, S.E. (2002) Nature, **416**, 556–560.
156. Aufsatz, W., Mette, M.F., Van Der Winden, J., Matzke, M., Matzke, A.J. (2002) EMBO J., **21**, 6832–6841.
157. Kanno, T., Mette, M.F., Kreil, D.P., Aufsatz, W., Matzke, M., Matzke, A.J. (2004) Curr. Biol., **14**, 801–805.
158. Verdel, A., Moazed, D. (2005) FEBS Lett., **579**, 5872–5878.
159. Petrie, V.J., Wuitschick, J.D., Givens, C.D., Kosinski, A.M., Partridge, J.F. (2005) Mol. Cell Biol., **25**, 2331–2346.