



**ИНТЕРФЕРЕНЦИЯ РНК В ФОРМИРОВАНИИ СОМАТИЧЕСКОГО ГЕНОМА У ИНФУЗОРИЙ  
*PARAMECIUM* И *TETRAHYMENA***

© И.В. Некрасова, А.А. Потехин

Санкт-Петербургский государственный университет, Санкт-Петербург

Для цитирования: Некрасова И.В., Потехин А.А. Интерференция РНК в формировании соматического генома у инфузорий *Paramecium* и *Tetrahymena* // Экологическая генетика. — 2018. — Т. 16. — № 4. — С. 5–22. doi: 10.17816/ecogen1645-22.

Поступила: 26.10.2018

Одобрена: 07.12.2018

Принята: 25.12.2018

☼ Инфузории являются отличной моделью для исследования интерференции РНК, играющей важнейшую роль в биологии этих протистов. В обзоре в сравнительном аспекте рассмотрены механизмы геномного сканирования, при каждом половом процессе обеспечивающие формирование соматического генома из хромосом генеративного ядра, у инфузорий *Tetrahymena* и *Paramecium*, относящихся к классу *Oligohymenophorea*. Сравнение нескольких одновременно существующих в клетке геномов осуществляется при посредничестве малых РНК и приводит к точному воспроизведению материнского соматического генома у потомков.

☼ **Ключевые слова:** инфузории; интерференция РНК; малые РНК; макронуклеус; диминуция хроматина; экспансия транспозонов; сайленсинг генов.

**RNA INTERFERENCE IN FORMATION OF THE SOMATIC GENOME OF CILIATES  
*PARAMECIUM* AND *TETRAHYMENA***

© I.V. Nekrasova, A.A. Potekhin

Saint Petersburg State University, Saint Petersburg, Russia

For citation: Nekrasova IV, Potekhin AA. RNA interference in formation of the somatic genome of ciliates *Paramecium* and *Tetrahymena*. *Ecological genetics*. 2018;16(4):5-22. doi: 10.17816/ecogen1645-22.

Received: 26.10.2018

Revised: 07.12.2018

Accepted: 25.12.2018

☼ Ciliates are the model of choice to study RNA interference, the mechanism playing key role in biology of these protists. The genome scanning processes of two ciliates, *Tetrahymena* and *Paramecium* (*Oligohymenophorea*), leading to formation of the somatic genome from the chromosomes of the generative nucleus are compared in the review. Matching of several simultaneously present in one cell genomes is mediated by small RNAs and results in precise reproduction of maternal somatic genome in the sexual progeny.

☼ **Keywords:** ciliates; RNA interference; small RNAs; macronucleus; chromatin diminution; transposon expansion; gene silencing.

**ВВЕДЕНИЕ**

Инфузории (Ciliophora), появившиеся около миллиарда лет назад, представляют собой одну из наиболее изученных и разнообразных групп протистов, включающую примерно 30000 видов [1]. Эволюционное расстояние между инфузориями *Oxytricha* и *Tetrahymena*, относящимися к разным классам, сопоставимо с таковым между человеком и дрожжами, а между представителями одного класса *Tetrahymena* и *Paramecium* — с расстоянием между ланцетником и человеком [2].

Представители родов *Paramecium* и *Tetrahymena* являются признанными модельными объектами клеточной и молекулярной биологии. Именно на этих инфузориях был открыт автосплайсинг интронов [3], впервые найдены отклонения от генетиче-

ского кода [4], обнаружена теломераза [5], найдена ацетилтрансфераза гистонов [6] и начато изучение «гистонового кода», а в последние 15 лет интенсивно изучается РНК-интерференция и ее влияние на функции хроматина и экспрессию генов [7, 8]. Многие из этих открытий были сделаны благодаря уникальной особенности инфузорий — наличию в их клетках двух типов ядер. Мультигеномный макронуклеус (МАК) — это соматическое ядро, экспрессия генов которого обеспечивает практически все вегетативные функции клетки. При вегетативном росте МАК делится уникальным способом, называемым «амитоз»: мини-хромосомы МАК неравномерно и случайно распределяются между дочерними ядрами, а веретено деления отсутствует [9].

При половом процессе МАК разрушается. Основной функцией генеративного микронуклеуса (МИК), содержащего диплоидный или полиплоидный набор хромосом, является хранение генетической информации и передача ее потомству при половом процессе. Во время вегетативного роста МИК делится митотически, а при половом процессе, конъюгации или автогамии, претерпевает мейоз. Новый МАК формируется из продуктов деления синкариона — зиготического ядра, возникающего в результате слияния двух гаплоидных продуктов мейоза МИК. Формирование генома МАК происходит в результате масштабных перестроек зиготического генома, включающих вырезание внутренних элиминирующихся последовательностей, IES (internal eliminated sequences), с последующим сшиванием фланкирующих их участков, фрагментацию хромосом, добавление к полученным фрагментам *de novo* теломерных повторов, и амплификацию образовавшихся в результате всех перестроек мини-хромосом [9]. У некоторых инфузорий (например, у *Euplotes* и *Oxytricha*, класс *Spirotrichea*) их размер оказывается менее 2 т. п. н., число — более 16 000, и на одной мини-хромосоме может находиться лишь один ген [10]. В любом случае мини-хромосомы МАК не являются хромосомами в цитологическом смысле, так как не имеют центромер, выраженных ориджинов репликации, и никогда не конденсируются. Фактически это просто многократно амплифицированные молекулы ДНК разного размера, фланкированные теломерными последовательностями.

Наличие двух геномов в одной клетке, один из которых функционирует в ряду поколений, а второй существует лишь в течение одного жизненного цикла, делает инфузорий исключительно удобной моделью эпигенетики. Действительно, геном МАК не подвержен прямому воздействию естественного отбора; при этом его формирование из хромосом зиготического ядра происходит по схеме, уже заложенной в предшествующем, «старом» МАК клетки. Все молекулярные процессы формирования генома МАК нетривиальны и сложны, но особенно впечатляет процесс удаления из генома тысяч IES — элементов, происходящих, по всей видимости, от транспозонов [11, 12]. Размеры IES могут варьировать, и все IES обладают уникальной последовательностью, то есть способ их удаления должен быть одновременно точным и неспецифичным. Хромосомы МИК подвергаются фрагментации, также сопровождающейся частичной потерей ДНК, причем у *Tetrahymena* из пяти хромосом МИК получается более 180 мини-хромосом МАК [13]. Долгое время после описания масштабной диминуции ДНК при развитии МАК инфузорий [14, 15] молекулярные процессы, приводящие к этому, оставались загадкой, и лишь в последние 15 лет удалось раскрыть ключевые механизмы формирования генома МАК инфузорий. Оказалось, что их главными участниками являются разные группы

некодирующих РНК, а превращение генеративного генома в соматический управляется сложными эпигенетическими механизмами, в основе которых лежит интерференция РНК.

### ОСОБЕННОСТИ ФОРМИРОВАНИЯ ГЕНОМА МАКРОНУКЛЕУСА У *TETRAHYMENA* И *PARAMECIUM*

Процессы формирования генома МАК принципиально различаются у инфузорий разных классов; в этом обзоре мы сфокусируемся на двух модельных представителях класса Oligohymenophorea — *Paramecium tetraurelia* и *Tetrahymena thermophila*. Масштабная диминуция хроматина при развитии генома МАК этих инфузорий захватывает около 25–30 % генома МИК [12, 13]. Основной вклад в этот процесс вносят удаление IES и фрагментация хромосом зиготического ядра.

IES *Tetrahymena* и *Paramecium* существенно различаются. IES *Tetrahymena* в целом значительно длиннее (размер IES варьирует от 136 п. н. до 43,4 т. п. н.), а их общее количество составляет около 12 000 [13]. Многие из них по-прежнему имеют явное сходство с транспозонами, в то время как другие сильно дивергировали. IES тетрахимен зачастую содержат гомологичные участки, указывающие на их общее происхождение [16]. Положение IES в ортологичных локусах сильно различается даже у близких видов, что, по-видимому, связано с продолжавшейся экспансией транспозонов. Сайты вырезания IES сильно различаются и не имеют строгих консенсусов. На концах некоторых IES обнаружены короткие прямые повторы, один из которых остается в геноме МАК после удаления IES. У тетрахимен IES практически никогда не располагаются внутри кодирующих последовательностей, и поэтому часто их эксцизия может происходить по нескольким альтернативным сайтам. При этом обнаруженные при полном секвенировании генома МИК одиннадцать IES, которые расположены внутри генов, вырезаются всегда точно [13, 17]. Эти одиннадцать IES фланкированы консервативными девятинуклеотидными инвертированными повторами ГТААСАСТТ. Недавно было показано, что точное удаление этих IES происходит по другому механизму, чем неточная эксцизия остальных IES тетрахимен [17] (см. ниже). Интересно, что гены, в кодирующих последовательностях которых располагаются эти IES, экспрессируются на поздних этапах конъюгации, совпадающих со временем удаления IES из генома МАК, и кодируют белки, возможно, участвующие в реорганизации хроматина [13] (см. ниже).

IES парамеций представляют собой в основном короткие однокопийные некодирующие элементы размером 26–882 п. н., в геноме *P. tetraurelia* их насчитывается более 45 000 [12]. Последовательности практически всех IES уникальны, и гомология между ними низка [18]. Они часто расположены внутри кодирующих последовательностей, то есть при формировании гено-

ма МАК необходимо их точное вырезание. Большинство IES парамеций являются рудиментами транспозонов семейства Tc1/mariner [12]. Они фланкированы прямыми повторами 5'-ТА-3', один из которых остается в геноме МАК после удаления IES [19]. Других строгих консенсусов на концах IES парамеций не обнаружено, хотя там зачастую присутствует мотив TAYAGYNR [20].

Фрагментация хромосом при формировании генома МАК *Tetrahymena* и *Paramecium* также приводит к удалению из генома некодирующих последовательностей, но при этом концы разрывов не сшиваются, к ним достраиваются теломеры, и так образуются мини-хромосомы МАК. Размер мини-хромосом МАК *Tetrahymena* и *Paramecium* составляет от 50 до 2000 т. п. н. [21–23]. Амплификация мини-хромосом приводит к пloidности МАК около 800п для *P. tetraurelia* и около 50п для *T. thermophila* [7].

У *Tetrahymena* для фрагментации необходимо и достаточно наличие цис-сигнала — высококонсервативной последовательности длиной 15 п. н. (chromosome breakage sequence Cbs), удаление которой из генома захватывает еще и короткие фланкирующие участки [24]. Теломеры присоединяются на расстоянии 5–25 п. н. с каждой стороны от Cbs [25]. Около половины из 225 Cbs, обнаруженных при секвенировании генома МИК *T. thermophila* [13], имеют консенсус 5'-TAAACCAACCTCTTT-3', у остальных Cbs в этой последовательности заменено не более двух нуклеотидов, причем 10 из 15 нуклеотидов строго консервативны, то есть сама Cbs и, соответственно, сайты фрагментации хромосом находятся под действием стабилизирующего отбора. Cbs напоминает сайты-мишени для транспозаз на концах транспозонов [13]. Интересно, что консенсус Cbs содержит в своем составе последовательность C<sub>2</sub>A<sub>2</sub>C<sub>2</sub>, похожую на перестроенный теломерный повтор *T. thermophila* C<sub>4</sub>A<sub>2</sub> [26], а белок Pot2, специфически взаимодействующий с Cbs при фрагментации хромосом [27], относится к тому же семейству, что и белок Pot1, участвующий в поддержании теломер. Являются ли эти совпадения случайными или отражают эволюцию Cbs, пока остается неясным, но предполагают, что сами Cbs и пока неизвестный белок, узнающий Cbs и инициирующий разрыв, могли произойти из мобильного элемента, доместифицированного в геноме именно для обеспечения фрагментации хромосом при формировании генома МАК [13]. Положение Cbs и длина гомологичных мини-хромосом МАК совпадают у разных видов *Tetrahymena*, в отличие от IES, положение которых в геноме сильно варьирует даже у близких видов [13]. Интересно, что не все последовательности, фланкированные в МИК Cbs и приобретающие теломеры в процессе формирования генома, сохраняются в геноме зрелого МАК. Недавно в МАК тетрахимен были обнаружены так называемые «неподдерживаемые» хромосомы (non-maintained chromosomes, NMC), которые образуют-

ся в формирующемся МАК, но исчезают при первых вегетативных делениях эксконъюгантных клеток [13]. В некоторых случаях (см. ниже) такие NMC кодируют белки, гены которых активно экспрессируются именно на поздних этапах конъюгации, то есть пока NMC еще присутствуют в новом соматическом геноме [17].

В геноме МИК *Paramecium* [28] также присутствует несколько сотен сайтов фрагментации. Однако у видов комплекса *P. aurelia* сигнальных последовательностей для фрагментации не обнаружено. Расположение сайта фрагментации при формировании нового МАК определяется у них не только последовательностью ДНК, но и эпигенетическими механизмами — клеточные линии с идентичными геномами МИК могут воспроизводимо реализовывать разные паттерны фрагментации при каждом половом процессе, то есть наследование паттернов фрагментации является «цитоплазматическим» [7] (см. ниже). По-видимому, в этом случае процесс фрагментации связан с неточной элиминацией сравнительно длинных микронуклеарных последовательностей вокруг сайтов, по которым происходит фрагментация [28]. Эти последовательности представляют собой в основном повторы, мини-сателлиты и транспозоны [28] и отличаются от IES большим размером и вариабельностью точек вырезания [28]. Иногда фрагментация у парамеций является факультативной, и элиминация ДНК сопровождается или разрывом, или сшиванием свободных концов [8]. В итоге при формировании МАК *P. tetraurelia* 50–60 хромосом МИК (точное количество хромосом МИК *P. tetraurelia* до сих пор неизвестно) фрагментируются примерно на 200 мини-хромосом МАК, к которым *de novo* добавляются теломеры [22]. Неточная фрагментация приводит в свою очередь к отсутствию точной позиции для добавления теломерных повторов, которое происходит в участках длиной от 800 п. н. до 2 т. п. н. [19]. Кроме того, у некоторых мини-хромосом существуют альтернативные участки присоединения теломер, находящиеся на расстоянии 2–20 т. п. н. друг от друга [7]. Таким образом, даже внутри одного генома МАК можно обнаружить по-разному процессированные гомологичные мини-хромосомы [29].

Долгое время оставалось совершенно непонятным, как регулируется быстрое, одновременное и точное протекание нескольких молекулярных процессов реорганизации генома: основные события формирования генома МАК разворачиваются в течение 3–4 часов и включают удаление тысяч IES, фрагментацию хромосом по сотням сайтов, синтез теломер и амплификацию полученных молекул ДНК. Другой загадкой были гомологозависимые материнские эффекты, наблюдаемые при трансформации МАК. Так, например, введение в МАК парамеций векторов, содержащих некоторые IES, вызывает нарушение эксцизии этих IES из генома нового МАК после прохождения транс-

формантами полового процесса [30–32]. Этот эффект наблюдается только для 7 % IES, получивших название «материнских» IES (matIES) [33]. При этом они не имеют никаких признаков, отличающих их от других, не-matIES [31]. Короткие (менее 50 п. н.) IES, как правило, не являются matIES [31, 33, Некрасова, Потехин, неопубл. данные]. Также для некоторых matIES характерно наличие на обоих концах редкого варианта консенсуса 5'-TAYAGYNR... -3', где в четвертой позиции находится T (5'-TATT... -3') [31], но это не является универсальной чертой всех matIES. Если же в геноме старого МАК происходит делеция фрагмента ДНК, то она также будет наследоваться в ряду поколений, то есть все половые потомки клетки, где произошла делеция в МАК, при развитии генома нового МАК будут удалять из него соответствующую делеции последовательность. Таким образом, любая отсутствующая в геноме старого МАК последовательность будет удалена из формирующегося генома нового МАК, а добавленная в старый МАК МИК-специфичная последовательность будет с высокой вероятностью сохранена в геноме нового МАК [34]. Вопрос о посредниках в передаче информации между старым МАК и новым, развивающимся соматическим ядром долгое время оставался открытым, на эту роль выдвигали молекулы РНК [35]. Эта гипотеза обрела основания с открытием механизма интерференции РНК. Как оказалось, формирование нового МАК у инфузорий управляется в первую очередь эпигенетическими механизмами с участием малых РНК и, действительно, происходит по трафарету генома родительского МАК. Для объяснения эпигенетических эффектов, наблюдаемых при формировании нового МАК у инфузорий *Paramecium* и *Tetrahymena*, была предложена модель сканирования геномов [36].

### МОДЕЛЬ ГЕНОМНОГО СКАНИРОВАНИЯ

Процесс геномного сканирования в целом сходен у тетрахимен и парамеций. По всей видимости, он эволюционировал в первую очередь как способ защиты генома от экспансии мобильных элементов и других чужеродных последовательностей. Вероятно, именно развитие и специализация механизмов интерференции РНК привели к возникновению ядерного дуализма у инфузорий, когда все «лишние» последовательности очень эффективно удаляются из соматического генома, в то же время свободно, практически незаметно для естественного отбора накапливаясь в молчащем геноме генеративного ядра. В геномное сканирование вовлечены как традиционные участники путей интерференции РНК, так и целый ряд белков, функции которых специфичны именно для инфузорий. На сегодняшний день эпигенетические механизмы формирования генома МАК, с одной стороны, в общих чертах ясны, с другой же стороны, по мере добавления деталей картина продолжает усложняться. Разберем процесс геномного

сканирования поэтапно, параллельно сравнивая события у представителей двух модельных родов инфузорий (рис. 1).

#### 1. Двухнаправленная транскрипция генома и образование сканРНК

В самом начале полового процесса как у тетрахимен, так и у парамеций происходит двухнаправленная транскрипция генома МИК с образованием двуцепочечных РНК (дцРНК). Считается, что для парамеций характерна полная транскрипция всего генома МИК, в то время как у тетрахимен преимущественно транскрибируются прителомерные и прицентромерные районы [37]. Именно из этих РНК-дуплексов затем будут образовываться малые сканирующие РНК (сканРНК), главные молекулы-посредники в сканировании геномов. Тотальная транскрипция генома требует участия особых белков-регуляторов. У *P. tetraurelia* было обнаружено два гена, кодирующих факторы транскрипции Spt5. Наличие в геноме двух генов *SPT5* редко встречается у эукариот [38]. Как оказалось, ген *SPT5v* экспрессируется при вегетативном росте, а экспрессия гена *SPT5m* достигает максимума на ранних стадиях полового процесса скорее всего во время мейоза. В начале автогамии белок Spt5m локализуется в мейотическом МИК, затем — в восьми продуктах мейоза; он еще присутствует в зиготическом ядре и продуктах его митотического деления, но по мере созревания новых МИК и МАК исчезает. Нокаут *SPT5m* приводит к высочайшей смертности — после полового процесса выживает только несколько процентов клеток, которые не способны делиться и умирают перед или сразу после первого деления. Искусственное снижение количества белка Spt5m приводит к 40-кратному уменьшению количества сканРНК, то есть, по сути, они не образуются в клетке.

Кроме того, нарушается элиминация почти всех МИК-специфичных последовательностей из генома развивающегося МАК, в том числе тех, вырезание которых не зависит от сканРНК. При этом отсутствие белка Spt5m не препятствует протеканию мейоза, формированию новых МАК и амплификации ДНК. Не исключено, что Spt5m котранскрипционно вносит некие модификации, необходимые в дальнейшем для эксцизии IES [38]. У *Tetrahymena* пока не обнаружено факторов Spt5 или белков, выполняющих аналогичную функцию.

Образовавшиеся в МИК дцРНК разрезаются эндонуклеазами семейства Dicer на короткие фрагменты размером 25 н. у *Paramecium* и 28 н. у *Tetrahymena* — собственно сканРНК. Вообще белки семейства Dicer, РНКазы III типа, разрезают дцРНК на короткие (21–28 п. н.) дуплексные молекулы с характерным монофосфатом на 5'-конце и концевыми двухнуклеотидными выступами [39]. Эти белки могут иметь до шести доменов; у разных групп организмов те или иные

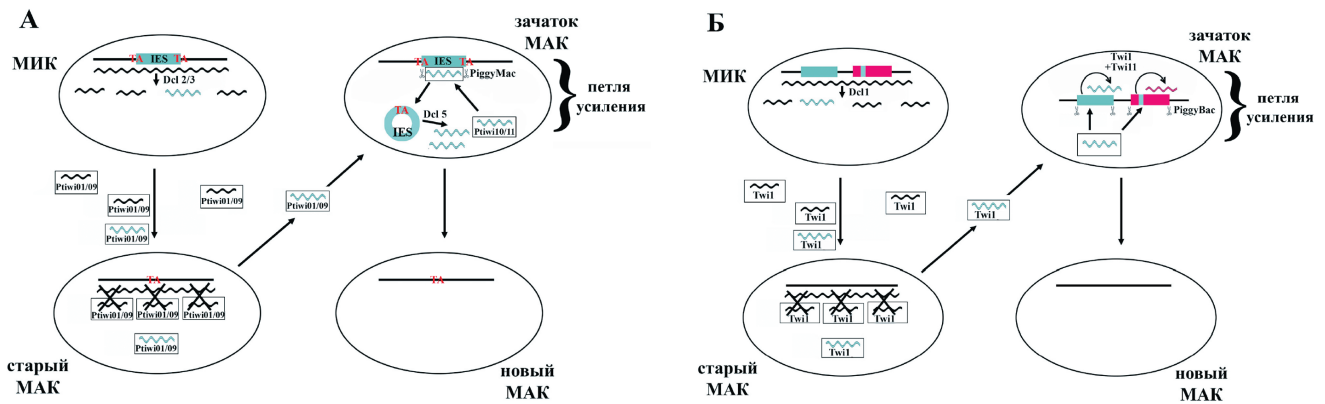


Рис. 1. Процесс геномного сканирования у *Paramecium* (А) и *Tetrahymena* (Б).

А. Геномное сканирование у *P. tetraurelia*. На ранних этапах полового процесса происходит полная двунаправленная транскрипция генома МИК. Полученные транскрипты МИК (показаны волнистой линией) с помощью белков Dcl2/3 процессируются в сканРНК (короткие волнистые линии). СканРНК образуют комплексы с белками Ptiwi01/09 (белые прямоугольники) и переносятся в старый МАК. ДНК старого МАК также полностью транскрибируется; сканРНК, комплементарные синтезированным нкРНК, деградируют. Оставшиеся сканРНК, гомологичные IES и прочим последовательностям, отсутствовавшим в старом МАК (показаны голубыми короткими волнистыми линиями) переносятся в формирующийся МАК, где узнают комплементарные им IES (голубые прямоугольники), в результате чего эти IES удаляются белком PiggyMac. На матрице вырезанных IES (голубые кольца) с помощью белка Dcl5 синтезируются iesРНК (голубые волнистые линии), которые взаимодействуют с белками Ptiwi10/11 и усиливают мечение и удаление гомологичных IES. В сайтах удаления IES сохраняются динуклеотиды ТА.

Б. Геномное сканирование у *T. thermophila*. На ранних этапах конъюгации происходит двунаправленная транскрипция генома МИК, затрагивающая как кодирующие последовательности (прямая черная линия), так и А-IES (голубой прямоугольник). В-IES (малиновый прямоугольник) не транскрибируются. Полученные транскрипты МИК (показаны волнистой линией) с помощью белка Dcl1 процессируются в ранние сканРНК (короткие волнистые линии). Ранние сканРНК образуют комплексы с белком Twi1 (белые прямоугольники) и переносятся в старый МАК. ДНК старого МАК также транскрибируется; ранние сканРНК, комплементарные синтезированным нкРНК, деградируют. Оставшиеся ранние сканРНК, гомологичные А-IES (показаны голубыми короткими волнистыми линиями), переносятся в формирующийся МАК. Там ранние сканРНК узнают не только А-IES, но и В-IES, имеющие короткие участки гомологии с А-IES (голубая полоска внутри малинового прямоугольника). В результате такого распознавания индуцируется синтез поздних сканРНК (малиновые волнистые линии) с А-IES и В-IES. Поздние сканРНК взаимодействуют с белками Twi1 и Twi11. Вместе ранние и поздние сканРНК обеспечивают полную элиминацию IES из формирующегося генома МАК

Fig. 1. Genome scanning process in *Paramecium* (A) and *Tetrahymena* (B).

A. Genome scanning in *P. tetraurelia*. At early stage of the sexual process the micronuclear genome is totally transcribed. The transcripts (curly lines) are processed into scanRNAs (short curly lines) by Dcl2/3p. ScanRNAs in complexes with Ptiwi01/09 (white rectangles) are transferred to the old macronucleus (MAC). The genome of the old MAC is also totally transcribed, and scanRNAs homologous to these transcripts get degraded. Remaining scanRNAs homologous to IESs and other sequences absent in the old MAC (blue short curly lines) are transferred to the new MAC, where they label complementary IESs (blue rectangles) for the excision by PiggyMac. The excised IESs (blue circles) serve as templates for synthesis of iesRNAs by Dcl5p (amplification loop, marked with a brace); iesRNAs (blue curly lines) interact with Ptiwi10/11 and enhance excision of homologous IESs. TA dinucleotides remain at the excision sites.

B. Genome scanning in *T. thermophila*. At early stage of the conjugation the coding sequences and A-IESs (blue rectangle) are bidirectionally transcribed in the micronuclear genome. B-IESs (purple rectangle) are not transcribed. The transcripts (curly lines) are processed by Dcl1p into early scanRNAs (short curly lines), which make complexes with Twi1p (white rectangles) and get transferred to the old MAC. The genome of the old MAC is also transcribed, and early scanRNAs which find homology among the transcripts get degraded. Remaining early scanRNAs, complementary to A-IESs (blue short curly lines), are transferred to the new MAC, where they recognize not only A-IESs but also short homologous regions in B-IESs. As a result of this, the late scanRNAs (purple curly lines) synthesis from A-IES and B-IES templates is induced. The late scanRNAs interact with Twi1p and Twi11p. Both early scanRNAs and late scanRNAs ensure complete elimination of IESs from the new MAC genome (amplification loop, marked with a brace)

домены могут отсутствовать, но наличие двух доменов с активностью РНКазы III является консервативным практически для всех эукариот [40]. У *T. thermophila* обнаружено три белка Dicer. Один из них, Dcr1, присутствует в клетке постоянно, но, по-видимому, каталитически неактивен, так как нарушение функции Dcr1 никак не проявляется фенотипически [41, 42]. Белок Dcr2 необходим для синтеза siРНК, участвующих в регуляции экспрессии различных последовательностей в вегетативных клетках [42]. Третий белок, Dcl1 (Dicer-like), имеет домен РНКазы и РНК-связывающий домен [41, 42], в небольшом количестве присутствует в вегетативных клетках и при голодании, но максимально его ген экспрессируется в конъюгирующих клетках [41, 42]. Главная функция белка Dcl1 тетрахимен заключается именно в образовании сканРНК [42, 43]. При скрещивании клеток с нокаутированным DCL1 с клетками дикого типа в клетках dcl1 — не формируются сканРНК. При этом синтез исходных длинных диРНК в клетках с нокаутированным DCL1 никак не нарушается. Белок Dcl1 также необходим для митотической и мейотической сегрегации хромосом МИК и нормального выстраивания хромосом в профазе мейоза, а у клеток с нокаутированным DCL1 МИК имеют меньшие размеры, чем в норме, и количество ДНК в них значительно ниже, чем у клеток дикого типа [42].

В геноме *P. tetraurelia* обнаружено восемь генов, кодирующих белки с доменами РНКазы III [44, 45]. Здесь необходимо отметить, что в эволюции рода *Paramecium* имели место как минимум три полные дубликации генома, две последние из которых произошли непосредственно перед дивергенцией видов-двойников комплекса *P. aurelia* [22]. В итоге в геномах этих видов гены в основном представлены группами онологов, насчитывающих максимально до восьми членов.

По-видимому, все гены семейства Dicer у *P. tetraurelia* как раз произошли от одного предкового гена [22]. Три из восьми генов (*DCR1*, 2, 3) кодируют типичные белки Dicer, содержащие по два хеликазных субдомена и по два домена РНКазы III [46], но из них лишь белок Dcr1 является каталитически активным, подобно белку Dcr2 у тетрахимен. Нокаут *DCR1* во время автогамии не препятствует развитию функционального нового МАК [46]. Еще пять генов *P. tetraurelia* кодируют белки, которые содержат только домен РНКазы и последовательность из 350 аминокислотных остатков на N-конце, функции которой неясны [44, 46]. Эти белки структурно близки к Dcl1 *T. thermophila* и были названы Dcl1-Dcl5 [46]. Как и у тетрахимен, именно Dicer-подобные белки Dcl2 и Dcl3 необходимы для образования сканРНК. Гены *DCL2* и *DCL3* активируются в начале автогамии [46]. Сайленсинг любого из этих двух генов по отдельности во время полового процесса приводит к частичным нарушениям эксцизии матIES, но не снижает выживаемости полового потомства. Результатом их совместного сайленсинга является полное выключение синтеза сканРНК, нарушение вырезания практически всех матIES и в итоге нежизнеспособность потомства [44, 46]. Оказалось, что белок Dcl3 отвечает за продукцию малых РНК размером от 24 до 30 н. с последовательностью UNG на 5'-конце [44], в то время как белок Dcl2 продуцирует малые РНК размером 25 н. предпочтительно с 5'-UNN или 5'-AGA, но затем Piwi-белки (см. ниже) отбирают те малые РНК, на 5'-конце которых находится U [47]. Вероятно, белки Dcl2 и Dcl3 работают в комплексе, так как при их совместной работе образуются сканРНК размером ровно 25 н. и с UNG на 5'-конце (рис. 2).

В клетках с инактивированными *DCL2* и *DCL3* гистоновые метки H3K27me3 и H3K9me3, маркирующие

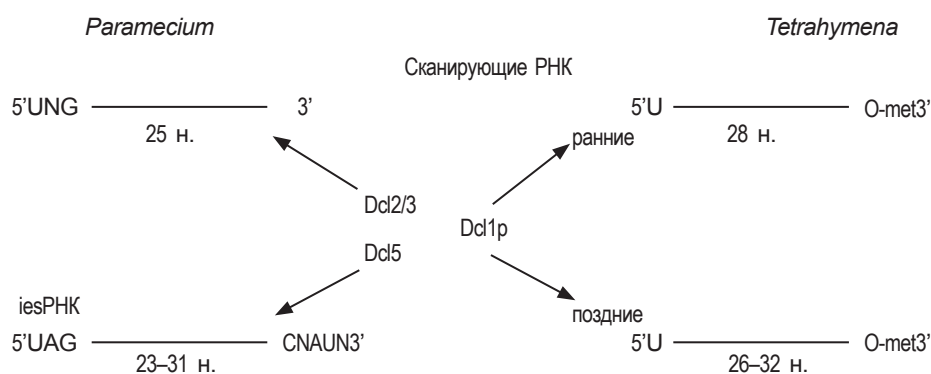


Рис. 2. Малые РНК, участвующие в геномном сканировании у парамеций и тетрахимен. У парамеций они представлены сканРНК, образующимися в результате работы белков Dcl2/3, и iesРНК, образующимися с помощью белка Dcl5 в петле усиления. У тетрахимен в геномном сканировании участвуют ранние и поздние сканРНК, формирующиеся с помощью белка Dcl1 и имеющие метилированную рибозу на 3'-конце

Fig. 2. Small RNAs acting in genome scanning in *Paramecium* and *Tetrahymena*. In *Paramecium* these are scanRNAs produced by Dcl2/3p, and iesRNAs produced by Dcl5p in the amplification loop. In *Tetrahymena* early and late scanRNAs are formed by Dcl1p

последовательности, подлежащие элиминации, выявляются в формирующемся МАК на более поздних сроках его развития и выражены слабее, чем в норме [33], что, вероятно, также связано с отсутствием сканРНК. Гены *DCL1* и *DCL4* на низком уровне экспрессируются на протяжении всего вегетативного роста клетки [44], функции этих белков не установлены, хотя показано, что белок Dcl4 также может производить малые РНК размером 25–30 н., отличающиеся по 3'-концу от тех малых РНК, за формирование которых отвечает наиболее родственный ему белок Dcl3 [47]. Наконец, белок Dcl5 отвечает за синтез iesРНК [44], обеспечивающих петлю усиления в процессе геномного сканирования (рис. 1; см. ниже).

СканРНК *Tetrahymena* размером 28 н. претерпевают 2'-О-метилирование рибозы на 3'-конце, что необходимо для обеспечения их функции и стабильности [48]. СканРНК парameций размером 25 н. имеют характерный тринуклеотид UNG на 5'-конце. СканРНК образуются из предшествующих дцРНК в виде дуплексов с 2-н. выступами на 3'-концах обеих цепей, что характерно для обычных продуктов реакций, катализируемых РНКазой III, а затем цепь, содержащая 5'-UNG, предпочтительно стабилизируется [46]. Такая стабилизация, вероятно, достигается благодаря ее специфичному взаимодействию с белками семейства Piwi, также участвующими в геномном сканировании [46]. Комплементарная цепь, по-видимому, быстро деградирует [46]. Общие характеристики сканРНК тетрахимен и парameций показаны на рис. 2.

**2. Взаимодействие сканРНК с белками Piwi**

Синтезированные сканРНК перемещаются в цитоплазму, где формируют комплекс с белками Piwi. Белки семейства Argonaute-Piwi являются наиболее консервативными из всех белков, задействованных в интерференции РНК. Они распадаются на две, по-видимому, паралогичные группы — Argonaute-подобные (Ago), взаимодействующие с различными классами малых РНК у большинства эукариот, и Piwi-подобные, функционирующие на определенных стадиях развития у животных. Животные в основном содержат белки обеих подгрупп; грибы (*Opisthokonta*), зеленые водоросли и растения (*Archaeplastida*) содержат преимущественно белки группы Ago, а инфузории [49, 50] — только Piwi-белки. Для этих белков характерно наличие двух главных доменов — PAZ и PIWI [51]. PAZ-домен связывает 3'-конец малых РНК, в то время как их 5'-конец закрывается в консервативном сайте PIWI-домена, обладающего активностью РНКазы [49]. После формирования комплекса белки Ago-Piwi разрезают одну из цепей, что приводит к ее диссоциации из комплекса [52].

В геноме *T. thermophila* закодировано 12 белков подсемейства Piwi, которые у этой инфузории называются Twi [40]. Большинство белков Twi *Tetrahymena* имеет установленные функции и вовлечено в интерференцию РНК в вегетативных клетках. Периоды экспрессии кодирующих белки Twi генов в жизненном цикле *T. thermophila* представлены на рис. 3.

Из всех членов семейства только белок Twi1 способен катализировать расщепление цепи в дуплексах

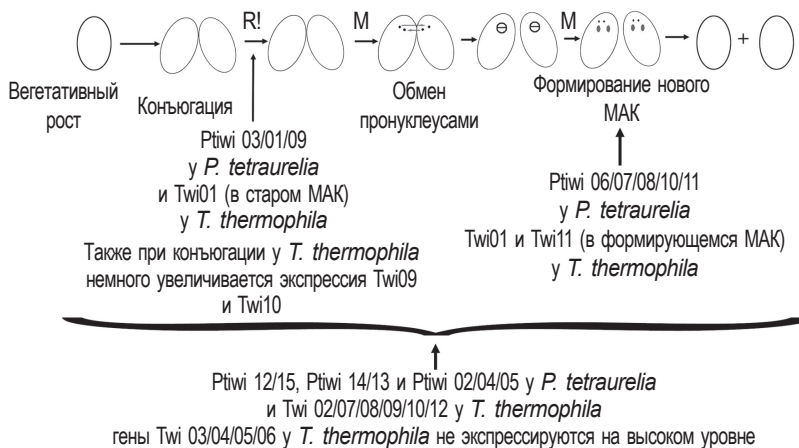


Рис. 3. Периоды экспрессии генов белков Piwi в жизненном цикле парameций и тетрахимен. Редукционное деление микронуклеусов конъюгирующих клеток обозначено символом «R!». В каждой из конъюгирующих клеток из гаплоидных ядер, образовавшихся в результате мейоза микронуклеуса(ов), остается только одно, остальные разрушаются. Это единственное гаплоидное ядро делится митотически (M), в результате чего формируются два генетически идентичных гаплоидных пронуклеуса, один из которых (стационарный пронуклеус) остается в исходной клетке, а второй мигрирует в клетку партнера по половому процессу, где сливается со стационарным пронуклеусом второй клетки. Образовавшееся зиготическое ядро, синкарион, далее делится митотически (также обозначено M)

Fig. 3. Expression periods of the PIWI proteins in life cycles of *Paramecium* and *Tetrahymena*. MIC(s) of each conjugating cell passes meiosis (R!), and all haploid nuclei except one get degraded. Remaining nucleus passes mitosis (M) leading to formation of two haploid pronuclei. One of pronuclei of each cell migrates to the partner cell and fuses with the stationary pronucleus of that cell. Zygotic nucleus, synkaryon, then divides mitotically (also labeled M)

малых РНК. Белок *Tw1* отсутствует в вегетативных клетках, а его ген экспрессируется только при конъюгации, через 3 часа после ее начала [36]. Клетки с нокаутом гена *TW11* оказываются не в состоянии завершить половой процесс и погибают [36]. Фрагментация хромосом при формировании МАК в клетках с нокаутированным *TW11* ингибируется, хотя не блокируется полностью [36]. Белок *Tw1* образует комплексы со сканРНК и необходим для их накопления [53]. При конъюгации *Tw1* сначала локализуется в цитоплазме, затем перемещается в старый, родительский МАК [53]. Мутации в гене *TW11*, приводящие к нарушению аккумуляции сканРНК, в дальнейшем сказываются и на установлении метки H3K9 на подлежащем элиминации хроматине [54] (см. ниже). Недавно было выяснено, что ген *TW11* экспрессируется не только в старом МАК, но и в формирующемся зиготическом МАК. Также в новом МАК сразу начинается экспрессия гена белка *Tw11*; экспрессия *TW11* и *TW111* из зиготического ядра связана с петлей усиления в сканировании генома (см. ниже).

В геноме *P. tetraurelia* найдено 17 генов *Piwi*-белков, у этой инфузории называемых *Ptiwi*, 12 из которых являются парными онологами последней полной дубликации генома [45, 49] (рис. 3). В процессе сканирования генома у *P. tetraurelia* участвуют два из них — *Ptiwi01* и *Ptiwi09*, имеющие гомологию друг с другом 99 % и по-видимому, родственные белку *Tw1* тетрахимен [45, 49]. Еще два белка, *Ptiwi10* и *Ptiwi11*, тоже гомологичные между собой на 98 %, связаны с петлей усиления [55] (см. ниже). Было показано, что гены белков *Ptiwi01* и *Ptiwi09* специфично активируются во время мейоза. В начале полового процесса эти белки локализируются в цитоплазме, а далее перемещаются сначала в старый МАК, а затем в новый, формирующийся МАК [49, 55]. Уменьшение количества любого из этих двух белков по отдельности не приводит к критическим нарушениям в формировании генома МАК, хотя вызывает нарушения материнского наследования делеций в функциональных генах [49]. Результатом же совместного нокаута *Ptiwi01* и *Ptiwi09* является блокировка вырезания IES, причем не только «материнских» [49], нарушение эксцизии транспозонов из формирующегося соматического генома и, как следствие, нежизнеспособность потомства.

### 3. Двухнаправленная транскрипция генома старого МАК и взаимодействие нкРНК со сканРНК

На следующем этапе сканРНК из МИК должны переместиться в старый МАК, где они «выяснят», какие именно последовательности ДНК из генома МИК там присутствуют — ведь именно эти последовательности должны сохраниться и в новом соматическом геноме. Белки *Piwi* доставляют сканРНК в старый МАК, где к тому времени также происходит полная двухнаправленная транскрипция генома. Как именно происходит

взаимодействие между сканРНК и полногеномными транскриптами старого МАК, неизвестно. Важно, что все сканРНК, которые найдут себе гомологию среди транскриптов старого МАК, выбывают из процесса геномного сканирования, видимо, деградируя в старом МАК [46].

Важную роль в перемещении комплексов *Piwi*-сканРНК играют GW-белки (белки, содержащие повторы из глицина и триптофана), имеющие сигнал ядерной локализации на С-конце, в то время как их N-концевой домен отвечает за способность связывать РНК, дцРНК, одноцепочечную ДНК и РНК/ДНК-дуплексы. Сначала такие белки, *Nowa1* и *Nowa2*, активность которых необходима для выживания полового потомства, были обнаружены у *P. tetraurelia* [56]. Гены этих белков фоновно экспрессируются у клеток в вегетативной фазе, но максимальная экспрессия наблюдается непосредственно перед началом мейоза и далее при половом процессе. С использованием GFP было показано, что белок *Nowa1* вскоре после начала мейоза обнаруживается в старом МАК, а затем перемещается в развивающийся МАК, так же как белок *Ptiwi09*. Белки *Nowa1* и *Nowa2*, вероятно, выступают посредниками во взаимодействии между нкРНК старого МАК и сканРНК, доставленными в старый МАК белками *Ptiwi01/09* [57]. Белки *Nowa* необходимы для нормальной элиминации транспозонов и матIES [56]. Предполагается, что эти белки играют важную роль в фильтрации сканРНК в старом МАК и затем участвуют в переносе не связавшихся в старом ядре молекул в новый МАК [56, 58].

У *Tetrahymena* также были обнаружены два GW-белка, *Wag1* и *CnjB*, участвующие в процессе формирования генома нового МАК [59]. Эти белки не являются гомологами *Nowa1* и *Nowa2*, но имеют сходную структуру. Как и *Nowa*-белки парамеций, белки *Wag1* и *CnjB* присутствуют в вегетативных клетках, но в небольшом количестве, и экспрессия их генов сильно возрастает при половом процессе. При вегетативном росте мутанты по гену *WAG1* не обладают выраженным фенотипом, в то время как мутанты по гену *CnjB* имеют мелкие, «дефектные» МИК, что говорит о возможной роли белка *CnjB* в поддержании хромосом МИК [59]. При половом процессе оба белка сначала локализируются в старом МАК, а потом — в развивающемся МАК, а их гены транскрибируются как в старом, так и в зиготическом МАК. У клеток, мутантных одновременно по обоим генам, *WAG1* и *CnjB*, нарушается взаимодействие комплексов *Tw1p*/сканРНК с нкРНК в старом МАК и, соответственно, весь дальнейший процесс сканирования. Интересно, что у таких двойных мутантов сохраняется нормальная локализация белка *Tw1* на всех стадиях полового процесса, в то время как одной из предполагаемых функций белков *Nowa* парамеций является как раз транспорт связанных с *Ptiwi*



сканРНК. Методом коиммунопреципитации было показано, что белки Wag1 и CnjB непосредственно или через другие белки взаимодействуют как друг с другом, так и с белком Twi1. У двойных мутантов по генам *WAG1* и *CnjB* серьезно нарушается развитие генома МАК, и в развивающемся МАК не появляются нуклеопротеиновые комплексы, формируемые элиминирующей ДНК и белком Pdd1 (см. ниже), в то время как у мутантов по каждому из двух этих генов в отдельности таких нарушений выявлено не было [59]. Отсутствие только белка CnjB не сказывается ни на жизнеспособности клеток, ни на их способности элиминировать IES из нового генома. Мутанты же по гену *WAG1* не могут преодолеть стадию конъюгации, на которой происходит диминуция хроматина и сшивание фрагментов ДНК. При этом даже выживающие 1–2 % клеток очень слабы. Таким образом, экспрессия *WAG1* при конъюгации необходима для получения нормального потомства, а CnjB, вероятно, оптимизирует работу *WAG1*.

#### 4. Двухнаправленная транскрипция генома формирующегося МАК, взаимодействие сканРНК с нкРНК в новом МАК

Геном развивающегося зачатка МАК также подвергается полной транскрипции. Очевидно, что в тотальной транскрипции геномов старого и нового МАК при половом процессе у инфузорий также должны быть задействованы особые факторы транскрипции. Действительно, у *P. tetraurelia* обнаружили шесть генов, кодирующих гомологи фактора TFIS, из которых три экспрессируются постоянно, а три индуцируются при автогамии, причем особенно сильно — ген *TFIS4*. Этот ген кодирует фактор элонгации, взаимодействующий с РНК-полимеразой II и необходимый для синтеза нкРНК в новом МАК. Его нокаут приводит к нарушению вырезания значительной части (около 48 %) IES, а также многих других МИК-специфических последовательностей при формировании нового МАК в результате, к высокой смертности у потомства [60]. Выживает около 15 % потомков, большая часть которых также погибает после нескольких делений. Белок TFIS4 аккумулируется в новом МАК до эксцизии IES и исчезает сразу после их удаления. Все IES, элиминация которых зависит от белков Nowa1 и Nowa2, нуждаются и в TFIS4, а все IES, удаление которых связано с функцией белка TFIS4, в свою очередь, входят в число IES, зависимых от гистоновой метилтрансферазы Ezl1 (см. ниже). Предполагают, что взаимодействие нкРНК со сканРНК в новом МАК приводит затем к установлению метки на хроматин, подлежащий элиминации из формирующегося генома. У *T. thermophila* был обнаружен гомолог TFIS, экспрессия гена которого также возрастает при половом процессе [61].

Еще одним важным участником событий этого этапа является хеликаза. Белок Ema1, появляющийся в МАК конъюгирующих клеток в профазе мейоза I и исчезаю-

щий через 14 часов после начала конъюгации, когда начинается элиминация ДНК из зачатка МАК, был обнаружен у тетрахимен. Эта ассоциированная с белком Twi1 РНК-хеликаза перемещается из старого МАК в развивающийся МАК [62]. По-видимому, ее основная роль состоит в обеспечении связывания комплекса Twi1-сканРНК с нкРНК в составе хроматина в старом и формирующемся МАК [62], она не нужна для созревания сканРНК и для их связывания с белком Twi1. Хеликаза Ema1 взаимодействует также с GW-белками. Нокаут гена *EMA1* у конъюгирующих клеток приводит к потере меток H3K9 и H3K27 на ДНК, подлежащей элиминации, нарушению вырезания многих IES в формирующемся МАК и в итоге к нежизнеспособности потомства. Недавно у *P. tetraurelia* также была обнаружена хеликаза Ptmb.220, гомологичная Ema1 тетрахимен и, видимо, выполняющая аналогичные функции [63]. Этот белок функционирует при половом процессе, когда, так же как и хеликаза Ema1 тетрахимен, сначала появляется в старом МАК, а затем — в формирующемся новом МАК, демонстрируя тот же паттерн локализации, что и белок Nowa1 парамеций [63]. Так же ведет себя еще один белок *P. tetraurelia* PtCAF-1 (*P. tetraurelia* Chromatin assembly factor-1-like protein), имеющий домен для связывания с гистонами H3 и H4 [57]. Нокаут гена *PtCAF-1* препятствует вырезанию матIES и транспозонов из формирующегося генома и, как и в других аналогичных случаях, приводит к исключительно высокой смертности клеток после полового процесса. Было показано, что белок PtCAF-1 необходим для элиминации МАК-специфичных сканРНК из общего пула сканРНК в процессе сканирования генома [57].

#### 5. Мечение подлежащих деградации последовательностей ДНК в развивающемся МАК

У тетрахимен нуклеосомы, связанные с помеченными с помощью сканРНК и подлежащими элиминации из развивающегося генома последовательностями, приобретают *de novo* гистоновые метки H3K27me3 [64] и H3K9me3 [65]. Модификация гистона H3K27me3 встречается и в МИК, и в МАК вегетативных клеток и, возможно, является промежуточным этапом для сканРНК-зависимого H3K9-метилирования [64]. Модификация H3K9, которая у большинства эукариот ассоциирована с гетерохроматином, у тетрахимен встречается только в подлежащем элиминации хроматине зачатка МАК [65]. Большинство IES *Paramecium* значительно короче, чем IES *Tetrahymena*, и их длины недостаточно, чтобы связать хотя бы одну нуклеосому. Это долгое время заставляло думать о том, что у *Paramecium* мечение ДНК на деградацию должно происходить не через модификации гистонов, а по-иному, например метилированием нуклеотидов [58]. Тем не менее, как оказалось, гистоновые метилтрансферазы необходимы для правильного удаления последовательностей при формировании генома МАК у обеих инфузорий.

У *Tetrahymena* было выявлено три гистоновые метилтрансферазы Ezl1, Ezl2 и Ezl3, специфически модифицирующие H3K27 [64]. Как оказалось, ген *EZL1* экспрессируется исключительно на ранних этапах конъюгации, в то время как *EZL2* и *EZL3* работают и в вегетативных клетках, хотя их экспрессия при конъюгации возрастает. Эксперименты по нокауту этих генов показали, что метилтрансфераза Ezl1 отвечает за установление метки H3K27me3 в родительском МАК и в зачатке нового МАК во время полового процесса, а Ezl2 и Ezl3 отвечают за другие, не перекрывающиеся события метилирования H3K27, в том числе в вегетативных клетках [64]. Клетки с нокаутированным геном *EZL1* после конъюгации не могут оставить живое потомство.

У *P. tetraurelia* недавно также была обнаружена метилтрансфераза гистонов Ezl1, необходимая для триметилирования гистона H3 в развивающемся МАК, и было показано, что она необходима для нормальной реаранжировки генома [33]. Всего в геноме *P. tetraurelia* было установлено 4 гена, кодирующих H3K27-специфичные метилтрансферазы, но ни одного — для H3K9-специфичных метилтрансфераз. В норме метки H3K27me3 и H3K9me3 никогда не выявляются у парамеций в транскрипционно активном МАК или в МИК при вегетативном росте. Все гены *EZL* экспрессируются исключительно при половом процессе, но *EZL4* экспрессируется сразу после мейоза, а *EZL2* — в самом начале развития нового МАК. Экспрессия *EZL3* и в первую очередь *EZL1* резко повышается в ходе мейоза и продолжается несколько часов, предшествуя появлению в зачатке МАК белка Dcl5 и началу эксцизии IES [33]. У клеток с инактивированной Ezl1 метки H3K27me3 и H3K9me3 практически не выявляются. В норме же метка H3K27me3 выявляется в МИК уже при первом делении мейоза и обнаруживается в старом МАК ближе к концу мейоза. Метка H3K9me3 на этих стадиях не определяется нигде. Зато обе метки выявляются с самого начала формирования нового МАК на всех стадиях его развития [33]. Нокаут гена метилтрансферазы гистонов *EZL1* во время автогамии летален, поставогазные клетки не могут совершить даже первые вегетативные деления. В геноме МАК таких клеток сохраняются все протяженные МИК-специфичные последовательности, как длинные IES, повторы и транспозоны, так и участки генов, отсутствующих в геноме родительской клетки. Полногеномный анализ показал, что МИК-специфичные последовательности отличаются по чувствительности к нокауту гена *EZL1* [33]. Этот фермент необходим для удаления 70 % из 45000 IES. При этом белок Ezl1 нужен и для точного вырезания 40 % коротких (длиной до 52 п. н.) IES; IES-мишени Ezl1 более GC-богаты и часто локализуются внутри кодирующих последовательностей. Таким

образом, гистоновые метки все же как-то участвуют и в детерминации на удаление последовательностей, длины которых недостаточно для самостоятельного связывания хотя бы одной нуклеосомы.

Другой модификацией гистонов в подлежащей элиминации ДНК в геноме формирующегося МАК является сумоилирование. Было показано, что экспрессия двух генов, продукты которых, SUMO и Uba2, участвуют в сумоилировании нуклеосом, усиливается при половом процессе у *Paramecium* и *Tetrahymena* [66]. Нокаут каждого из этих генов у *Paramecium* приводит к нарушению эксцизии IES, а потомки полового процесса при этом выживают за счет регенерации старого МАК. Функция этих генов в вегетативных клетках не очевидна [66].

#### 6. Распознавание гистоновых модификаций на ДНК, подлежащей элиминации, в старом МАК и формирующемся МАК

Несмотря на в целом высокое сходство систем геномного сканирования у тетрахимен и парамеций, существует и ряд белков, обнаруженных только у той или другой инфузории. Так, белки, вовлеченные в распознавание гистоновых модификаций на удаляемой ДНК, пока известны только для тетрахимен. Белки Pdd1, Pdd2 и Pdd3 (Programmed DNA Degradation) не имеют выраженной гомологии с какими-либо белками *Paramecium*. Белки Pdd1 и Pdd2 выявляются совместно и только в период 5–15 часов после начала полового процесса в старом МАК и затем в зачатке нового МАК [61, 67, 68]. Сначала белки Pdd1 и Pdd2 расположены в зачатке МАК диффузно, затем они формируют крупные структуры — комплексы с элиминируемой ДНК, концентрирующиеся у периферии ядра [67]. Pdd1 имеет в своем составе хромодомены и способен *in vitro* связывать метки H3K9me3 и H3K27me3 [67, 68], являющиеся у тетрахимен сигналом для элиминации ДНК в зачатке нового МАК и для деградации ДНК в старом МАК [69].

Подавление экспрессии гена *PDD1* при конъюгации приводит к нарушению аккумуляции сканРНК в новом МАК [36] и уменьшению метилирования H3K9 [69]. Клетки с нокаутированным геном *PDD1* обычно умирают после конъюгации, а у выживающих инфузорий регенерирует старый МАК [68]. Клетки с нокаутированным геном *PDD2* также не способны завершить половой процесс [70], так как по крайней мере часть МИК-специфичной ДНК, как IES, так и содержащие Cbs-последовательности, у них не может быть элиминирована. Таким образом, белок Pdd2 необходим и для вырезания IES, и для фрагментации хромосом [70]. Однако снижение количества белка Pdd2 не мешает накоплению сканРНК [36]. Белок Pdd3 также содержит один хромодомен и, как и Pdd1, способен связывать метку H3K9. В конъюгирующих клетках он тоже колокализуется с ДНК, подлежащей элиминации [69].

Неродственные друг другу белки Lia (localize in macronuclear anlagen), как и белки Pdd, в основном не имеют гомологов у парамеций [71]. Lia1 и Lia3 являются небольшими положительно заряженными белками, что говорит об их вероятной способности связываться с нуклеиновыми кислотами. Lia2 содержит домены, характерные для РНК-хеликаз. В начале полового процесса белки Lia локализируются в основном в зачатке МАК, а также в цитоплазме и в конъюсоме, образующейся в формирующемся МАК тетрахимен особой структуре с неизвестными функциями, главным компонентом которой является белок Pdd1 [72]. После расхождения конъюгирующих клеток белки Lia вместе с Pdd1 обнаруживаются в дискретных зонах внутри нового МАК, что связывает белки Lia с процессами реаранжировки генома МАК [71]. Отдельного внимания заслуживает белок Lia5, который мы рассмотрим ниже.

### 7. Удаление МИК-специфичных последовательностей из генома МАК

Транспозазы, отвечающие собственно за вырезание помеченных на удаление МИК-специфичных последовательностей в развивающемся МАК, были впервые обнаружены у *P. tetraurelia* [73].

Как сказано выше, IES парамеций являются рудиментами транспозонов Tc/mariner, но их эксцизия происходит не так, как у предковых мобильных элементов. С обеих сторон IES возникают двунитевые разрывы по сайтам ТА, при этом на обоих 5'-концах образуются 4 н. — одонитевые выступы; на месте удаленной IES сохраняется один из сайтов ТА [74]. Двунитевые разрывы вносятся белком PiggyMac (PGM), представляющим собой «одомашненную» парамециями транспозазу piggyBac [73]. Транспозоны PiggyBac, исходно обнаруженные в геноме мотылька *Trichoplusia ni*, кодируют транспозазу, которая осуществляет их эксцизию и обеспечивает встраивание транспозонов в сайт-мишень ТТАА [75, 76]. PGM структурно сходна с транспозазами PiggyBac, но, в отличие от них, ген PGM не является частью транспозона и, соответственно, всегда сохраняется в геноме МАК [22, 77]. Таким образом, ген PGM стал частью генома *Paramecium*, а кодируемый им белок выполняет важнейшие функции. В настоящий момент одомашненные гены транспозаз PiggyBac найдены у разных эукариот [78], причем один из таких генов у человека кодирует белок PGBD5, участвующий в перестройках ДНК в раковых клетках [79]. Экспрессия гена PGM *P. tetraurelia* приурочена к завершению полового процесса, когда белок PGM обнаруживается в развивающемся МАК [73]. Нокаут гена PGM во время полового процесса приводит к нарушению эксцизии практически всех IES и других МИК-специфичных последовательностей в формирующемся МАК и к невозможности фрагментации хромосом, что фатально для потомков. Для эффективного удаления IES необходимо формирование комплекса, состоящего минимум из двух

белков PGM [80]. Недавно было обнаружено еще пять групп генов-паралогов PGM, кодирующих белки, получившие общее название PgmL (Pgm-like). В эти пять групп входят гены девяти белков PgmL, различающихся по последовательности и организации доменов. Гены, кодирующие белки PgmL, экспрессируются при половом процессе [81]. Белки каждой из пяти групп локализируются в развивающемся МАК, являются необходимыми компонентами общего ассоциированного с PGM комплекса и важны для правильной локализации PGM в ядре и точного вырезания IES [81].

Открытие транспозазы из группы PiggyBac у парамеций заставило искать ее и у тетрахимен. Основанием для этого в том числе послужило то, что у *Tetrahymena* при вырезании IES также образуются 4 н. 5'-одонитевые выступы на обоих концах, хотя консервативный сайт ТА у них отсутствует [82]. Действительно, в геноме *Tetrahymena* было обнаружено несколько генов транспозаз PiggyBac, TPB (Tetrahymena piggyBac transposase) [17, 83]. Таким образом, ключевой фермент пути геномного сканирования, осуществляющий собственно эксцизию ненужной в МАК ДНК, по-видимому, появился в геноме инфузорий в составе транспозона и одомашнивание транспозаз PiggyBac произошло единожды на раннем этапе эволюции инфузорий класса Oligohymenophorea [73]. Ген TPB2 кодирует белок, отвечающий за эксцизию МИК-специфичных последовательностей при формировании нового МАК [83]. Его экспрессия сильно возрастает на поздних этапах конъюгации, через 12–14 часов после смешивания клеток противоположных типов спаривания. Белок Trb2 локализуется в развивающемся МАК с метками H3K27 и H3K9 и белком Pdd1. Нокаут гена TPB2 приводит к нарушению вырезания IES и фрагментации хромосом и, соответственно, к гибели потомства [83]. Как и все транспозазы PiggyBac, Trb2 распознает сайты 5'-ТТАА-3' и вносит двунитевой разрыв перед первым Т, но в качестве субстрата могут использоваться и другие последовательности — например, 5'-GTAG-3' и 5'-AGTGAT-3'. Недавно было показано [17, 84], что две другие транспозазы PiggyBac, Trb1 и Trb6, вместе отвечают за эксцизию точно вырезаемых IES, расположенных внутри экзон и имеющих на концах девяти-нуклеотидные консенсусы, типичные для транспозонов piggyBac. Было показано [17], что активно экспрессирующийся в течение полового процесса ген TPB6 расположен на одной из NMC, исчезающей из генома зрелого МАК тетрахимен. Интересно, что белок Lia5, имеющий домен «цинковый палец» для связи с ДНК, по-видимому, также исходно произошел от транспозаз PiggyBac, но утратил домены, необходимые для выполнения функции транспозазы [85]. Этот белок требуется для зависимого от Trb2 удаления IES из формирующегося соматического генома при половом процессе и может являться функциональным гомологом белков

PgmL *P. tetraurelia*, хотя на данный момент неизвестно, взаимодействует ли он напрямую с Trb2 [81, 85].

#### 8. Лигирование МАК-специфичных последовательностей ДНК

После эксцизии IES сшивание концов обеспечивают лигазы. Такие лигазы пока что обнаружены только для парамеций [86], хотя и у тетрахимен они, очевидно, должны существовать. Образующиеся концы ДНК соединяются белками системы NHEJ (non-homologous ends joining), у других эукариот обеспечивающей сшивание двунитевых разрывов хромосом. Ключевыми компонентами этой системы являются АТФ-зависимые лигазы и их кофакторы. В геноме *P. tetraurelia* найдено два гена (*LIG4a* и *LIG4b*) лигаз типа IV, репарирующих двунитевые разрывы в пути NHEJ. Также был обнаружен один ген, кодирующий гомолог XRCC4, необходимый компонент комплекса лигаз, и два гена *XLFCernunnos*, кодирующих кофакторы комплекса Lig4-XRCC4 [86]. Экспрессия генов *LIG4a* и *XRCC4* индуцируется при половом процессе, а белок Lig4a локализуется в развивающемся МАК. При этом экспрессия *LIG4a* и *XRCC4* начинается раньше, чем PGM вносит двунитевые разрывы в ДНК нового МАК, то есть активация лигирующего комплекса не индуцируется разрывами, а выступает частью запрограммированного процесса. Инактивация генов *LIG4a* и *XRCC4* приводит к нарушению формирования хромосом развивающегося МАК и регенерации старого МАК у выживающих после полового процесса клеток [86].

#### 9. «Петля усиления»

Как было недавно обнаружено, процесс сканирования генома как у тетрахимен, так и у парамеций имеет «петлю усиления» — дополнительный раунд синтеза малых РНК, обеспечивающих максимально полное мечение IES в формирующемся геноме. В новом МАК появляются малые РНК особого класса, гомологичные исключительно IES [16, 44]. Эти малые РНК, особенности их образования и механизм действия, по-видимому, различаются у двух родов инфузорий. Неожиданно оказалось, что у тетрахимен не для всех IES в МИК образуются обычные, «ранние» сканРНК. Те IES, для которых сканРНК образуются сразу, получили название А-IES; однако В-IES, для которых ранние сканРНК отсутствуют, также элиминируются из генома нового формирующегося МАК. Оказалось, что за удаление В-IES отвечают «поздние» сканРНК размером 26–32 н. (см. рис. 1), которые формируются в зачатке нового МАК до эксцизии IES и взаимодействуют с белками Twi11 и Twi1, синтезируемыми за счет экспрессии генов в новом МАК. При этом Twi1, кодируемый в старом МАК, связывает только ранние сканРНК, распознающие А-IES [16]. Однако если нокаут гена *TWI1* в старом МАК приводит к нарушению элиминации большей части IES, то нокаут генов *TWI1* и *TWII1* в новом МАК вызывает лишь к нарушению удаления части В-IES, то есть ран-

ние сканРНК все равно необходимы для удаления всех IES. Внутри А-IES и В-IES присутствуют одинаковые повторы, что позволяет ранним сканРНК, продуцируемым с А-IES, и распознавать В-IES (то есть ранние сканРНК обладают транс-функцией), и запускать синтез вторичных, поздних сканРНК для надежной эксцизии всех IES. Как и ранние сканРНК, поздние сканРНК образуются в результате работы белка Dcl1 и не отличаются от ранних сканРНК ничем, кроме размера (см. рис. 2). Для синтеза поздних сканРНК нужны белки Ezl1 и Pdd1, то есть, вероятно, сайты синтеза дополнительных сканРНК определяются гистоновыми метками. В свою очередь, поздние сканРНК тоже способны к транс-распознаванию многих IES, поэтому наличие даже небольшого количества ранних сканРНК в результате «петли усиления» оказывается достаточно для удаления практически всех IES [16].

У парамеций не выделяют ранних и поздних сканРНК, но помимо сканРНК у парамеций был обнаружен новый класс малых РНК, участвующих в реорганизации генома МАК при половом процессе [44]. Эти малые РНК гомологичны только последовательностям IES и были названы iesРНК (см. рис. 1). Размер iesРНК варьирует от 23 до 31 н., их концевые последовательности отличаются от таковых сканРНК (см. рис. 2). В отличие от сканРНК, за продукцию которых отвечают белки Dcl2 и Dcl3, iesРНК синтезируются при работе Dcl5. Ген *DCL5* экспрессируется на поздних стадиях развития МАК, как раз когда происходит эксцизия IES. В отличие от многих других белков, участвующих в процессе сканирования, Dcl5p, по-видимому, не участвует в фрагментации хромосом [33] и важен только для эксцизии IES, зависимых от iesРНК. Сайленсинг гена *DCL5* приводит к сохранению в геноме примерно 12 % IES, но не является летальным для клетки [44]. Недавно было показано, что Dcl5p может распознавать в дцРНК именно сайты, соответствующие концам IES [47], разрезая транскрипты искусственных конкатемеров из двух IES точно по их границе. Замена любого из нуклеотидов в сайте сшивки IES-IES вызывала нарушение распознавания этого сайта белком Dcl5. По ходу формирования генома МАК количество сканРНК падает, а количество iesРНК, наоборот, нарастает. Если сканРНК взаимодействуют с белками Ptiwi01/09, синтезируемыми за счет экспрессии генов в старом МАК, то iesРНК образуют комплексы с белками Ptiwi10/11, синтезируемыми в зиготическом МАК [55]. Предполагается, что iesРНК формируются при транскрипции уже вырезанных из ДНК IES, обеспечивая положительную петлю обратной связи — количественно усиливая мечение и удаление IES, составляющих примерно 7–15 % от общего числа IES [44]. Поскольку, в отличие от IES тетрахимен, IES парамеций уникальны, их транс-распознавание, по-видимому, невозможно.

## 10. Белки пути сканирования с неизвестными функциями

У парамеций и тетрахимен есть также белки, появляющиеся при половом процессе и необходимые для завершения реаранжировки генома, точные функции которых остаются невыясненными. Так, у *P. tetraurelia* был обнаружен экспрессирующийся при половом процессе ген, продукт которого обнаруживается исключительно в развивающемся новом МАК в период, когда происходит вырезание IES. Этот белок, получивший название Die5 (Defective IES Excision), необходим для удаления из генома МИК-специфичных последовательностей [87]. Гомологичный ген со сходным профилем экспрессии обнаружен и у *T. thermophila*. Его продукт детектируется в новом МАК только на ранних этапах развития и необходим для удаления IES и фрагментации хромосом [87], но зато на всех этапах конъюгации присутствует в МИК [87]. Die5p не имеет белковых доменов с известной функцией, кроме двух классических сигналов ядерной локализации. *Tetrahymena* с инактивированным *DIE5* не могут завершить амплификацию генома, в том время как мутанты по *DIE5* *Paramecium* геном амплифицируют, но не могут правильно его реорганизовать [87].

Для правильной перестройки генома тетрахимен необходимо также наличие центромерспецифичного варианта гистона H3, который встречается у всех эукариот [88, 89]. У *T. thermophila* этот белок называется Spa1. Ген *CNA1* на низком уровне экспрессируется в вегетативных клетках, выключается при голодании и вновь экспрессируется при конъюгации. Такой паттерн экспрессии соответствует роли Spa1 в формировании центромер: в вегетативных клетках постоянно реплицируются хромосомы МИК, при голодании МИК замирает в фазе G2, а во время конъюгации проходит несколько делений пре- и постзиготических митозов. По мере формирования нового МАК Spa1p перестает детектироваться в этом ядре, что коррелирует с исчезновением из него центромерных участков по мере элиминации МИК-специфичной ДНК. Как оказалось, экспрессия *CNA1* при конъюгации необходима для выживания потомства. Если нокаутировать в конъюгирующих клетках ген *CNA1*, это препятствует формированию комплексов белков Pdd с элиминируемой ДНК. Возможно, именно центромерные участки хромосом МИК образуют кор для формирования этих структур в развивающемся зачатке МАК [90].

У парамеций при половом процессе индуцируется экспрессия 28 не имеющих паралогов генов [91], 18 из которых являются уникальными и не имеют гомологов в базах данных. При этом для двух из этих генов оказалось, что их экспрессия необходима для выживания полового потомства (при нокауте летальность превышала 90 %). Они были названы *PDSG1* и *PDSG2* (*Paramecium* protein involved in Development of the Somatic

Genome). Однако свойства белков, кодируемых этими генами, сильно различались. Белок Pdsg1 присутствует как в старом, так и в развивающемся МАК, в то время как Pdsg2 — только в развивающемся МАК. При этом Pdsg1 важен для элиминации только некоторых матIES, а Pdsg2 оказался необходим для элиминации всех проверенных IES [91]. Кроме того, нокаут *PDSG1* блокирует процесс сканирования — нарушается элиминация сканРНК, гомологичных последовательностям, присутствующим в старом МАК. Нокаут *PDSG2* никак не влияет на сканРНК, но затрагивает iesРНК. Интересно, что из всех белков *Paramecium*, участвующих в элиминации IES, только Pdsg1 не задействован и в удалении транспозонов; уменьшение количества Pdsg2 приводит к значительному удержанию транспозонов двух семейств, то есть белок Pdsg2 вовлечен и в процесс фрагментации хромосом.

Как уже было сказано выше, удаление не всех IES парамеций обеспечивается процессом сканирования генома. Значительная часть IES не зависит от сканРНК, особенно это касается относительно коротких IES, давно существующих в геномах парамеций и сильно редуцированных по сравнению с транспозонами, от которых они произошли. Каким способом обеспечивается их вырезание? На данный момент однозначного ответа на этот вопрос нет, но существует ряд предположений, что в качестве сигналов могут выступать и модификации ДНК [8]. Метилирование цитозина, стандартная модификация ДНК бактерий, у инфузорий практически отсутствует. В геноме МАК *Paramecium* и *Tetrahymena* зато обнаруживается другая характерная для бактерий, архей и некоторых грибов модификация ДНК — N<sub>6</sub>-метилирование аденина [92, 93], происходящее по сайтам GATC в развивающемся МАК как раз в то время, когда происходит элиминация ДНК из формирующегося генома этого ядра. Возможно, хотя пока никак не доказано, что это может иметь значение для распознавания удаляемых последовательностей [8].

## ЗАКЛЮЧЕНИЕ

Многообразие раскрытых за последние двадцать лет эпигенетических механизмов прекрасно иллюстрирует роль интерференции РНК в базовых процессах жизнедеятельности клетки. Именно эпигенетические механизмы делают различными клетки многоклеточного организма, обладающие идентичными геномами, они вовлечены в процессы определения пола и видообразования [94]. У инфузорий *Tetrahymena* и *Paramecium* малые РНК лежат в основе сложнейшего процесса геномного сканирования, в результате которого в одной клетке инфузорий всегда существует два типа ядер, содержащих физически и функционально различающиеся геномы. При половом процессе все геномы полностью транскрибируются, формирующиеся в МИК малые

сканирующие РНК, перемещаясь между ядрами, обеспечивают прямое «сравнение» геномов МИК и старого МАК, по гомологии взаимодействуя с транскриптами старого, а затем нового МАК. В результате этого сравнения появляются соответствующие гистоновые метки в хроматине развивающегося МАК, что у других организмов приводит к гетерохроматинизации. Инфузории заходят дальше в этом процессе, элиминируя ДНК, связанную с мечеными нуклеосомами, из соматического генома, который, по сути, можно назвать эпигеномом. Так как соматический геном формируется заново из хромосом генеративного ядра в каждом жизненном цикле, инфузории являются уникальной моделью для изучения механизмов эпигенетического наследования в ряду поколений. В результате геномного сканирования не только быстро и точно, по образцу родительского генома, формируется сложнейший геном МАК в клетках-потомках, но и обеспечивается материнское наследование признаков [37]. Так, наследование типов спаривания у инфузорий *P. tetraurelia* [95], один из первых описанных случаев цитоплазматического наследования признаков, более 70 лет остававшееся загадкой эпигенетики, недавно [96] было объяснено именно тем, что при геномном сканировании точно воспроизводится материнское состояние аллелей генов типов спаривания и делеции в промоторе или кодирующей последовательности родительской клетки наследуются у клетки дочерней. При этом отсутствующие части генов не являются IES или рудиментами транспозонов, они ошибочно, по наличию цис-сигналов, могли быть распознаны как IES-подобные последовательности и затем в ряду поколений стали стабильно удаляться из генома МАК [97]. Очевидно, что аналогичным образом могут регулироваться и наследоваться и другие признаки у инфузорий. То есть система, эволюционировавшая для защиты генома от экспансии транспозонов и удаления из соматического генома некодирующей ДНК, может быть адаптирована и для других целей. Особенности организации ядерного аппарата инфузорий иногда мешают проводить прямые параллели между эпигенетическими феноменами, обнаруженными у инфузорий и выявленными у других эукариот. Однако нет сомнений, что в будущем сведения о механизмах и роли интерференции РНК, полученные при исследованиях инфузорий — казалось бы, очень специфичной группы эукариот, — постоянно будут находить аналогии и параллели в биологии других организмов, как это уже происходит сейчас.

Авторы признательны Эрику Майеру (Институт биологии, Высшая нормальная школа, Париж, Франция) за длительные дискуссии по эпигенетике парамециев, способствовавшие написанию этого обзора. Исследование выполнено в рамках работ по гранту РНФ 16-14-10157.

## ЛИТЕРАТУРА

1. Adl SM, Simpson AG, Lane CE, et al. The revised classification of eukaryotes. *J Eukaryot Microbiol.* 2012;59(5):429-493. doi: 10.1111/j.1550-7408.2012.00644.x.
2. Parfrey LW, Lahr DJ, Knoll AH, Katz LA. Estimating the timing of early eukaryotic diversification with multigene molecular clocks. *Proc Natl Acad Sci USA.* 2011;108(33):13624-13629. doi: 10.1073/pnas.1110633108.
3. Cech TR, Zaug AJ, Grabowski PJ. *In vitro* splicing of the ribosomal RNA precursor of *Tetrahymena*: involvement of a guanosine nucleotide in the excision of the intervening sequence. *Cell.* 1981;27(3 Pt 2):487-496. doi: 10.1016/0092-8674(81)90390-1.
4. Caron F, Meyer E. Does *Paramecium primaurelia* use a different genetic code in its macronucleus? *Nature.* 1985;314(6007):185-188. doi: 10.1038/314185a0.
5. Greider CW, Blackburn EH. Identification of a specific telomere terminal transferase activity in *Tetrahymena* extracts. *Cell.* 1985; 43:405-413. doi: 10.1016/0092-8674(85)90170-9.
6. Brownell JE, Zhou J, Ranalli T, et al. *Tetrahymena* histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell.* 1996;84(6):843-851. doi: 10.1016/S0092-8674(00)81063-6.
7. Meyer E, Garnier O. Non-Mendelian inheritance and homology-dependent effects in ciliates. *Adv Genet.* 2002;46:305-337. doi: 10.1016/S0065-2660(02)46011-7.
8. Coyne RS, Lhuillier-Akakpo M, Duharcourt S. RNA-guided DNA rearrangements in ciliates: is the best genome defence a good offence? *Biol Cell.* 2012;104(6):309-325. doi: 10.1111/boc.201100057.
9. Coyne RS, Chalker DL, Yao MC. Genome downsizing during ciliate development: nuclear division of labor through chromosome restructuring. *Annu Rev Genet.* 1996;30:557-578. doi: 10.1146/annurev.genet.30.1.557.
10. Swart EC, Bracht JR, Magrini V, et al. The *Oxytricha trifallax* macronuclear genome: a complex eukaryotic genome with 16,000 tiny chromosomes. *PLoS Biol.* 2013;11(1): e1001473. doi: 10.1371/journal.pbio.1001473.
11. Chalker DL, Yao MC. DNA elimination in ciliates: transposon domestication and genome surveillance. *Annu Rev Genet.* 2011;45:227-246. doi: 10.1146/annurev-genet-110410-132432.
12. Arnaiz O, Mathy N, Baudry C, et al. The *Paramecium* germline genome provides a niche for intragenic parasitic DNA: evolutionary dynamics of internal eliminated sequences. *PLoS Genet.* 2012;8(10):e1002984. doi: 10.1371/journal.pgen.1002984.

13. Hamilton EP, Kapusta A, Huvos PE, et al. Structure of the germline genome of *Tetrahymena thermophila* and relationship to the massively rearranged somatic genome. *Elife*. 2016;5. doi: 10.7554/eLife.19090.
14. Ammermann D, Steinbrück G, von Berger L, Hennig W. The development of the macronucleus in the ciliated protozoan *Stylonychia mytilus*. *Chromosoma*. 1974;45(4):401-429. doi: 10.1007/BF00283386.
15. Yao MC, Gorovsky MA. Comparison of the sequences of macro- and micronuclear DNA of *Tetrahymena pyriformis*. *Chromosoma*. 1974;48(1):1-18. doi: 10.1007/BF00284863.
16. Noto T, Kataoka K, Suhren JH, et al. Small-RNA-Mediated Genome-wide trans-Recognition Network in *Tetrahymena* DNA Elimination. *Mol Cell*. 2015;59(2):229-242. doi: 10.1016/j.molcel.2015.05.024.
17. Feng L, Wang G, Hamilton EP, et al. A germline-limited piggyBac transposase gene is required for precise excision in *Tetrahymena* genome rearrangement. *Nucleic Acids Res*. 2017;45(16):9481-9502. doi: 10.1093/nar/gkx652.
18. Duret L, Cohen J, Jubin C, et al. Analysis of sequence variability in the macronuclear DNA of *Paramecium tetraurelia*: a somatic view of the germline. *Genome Res*. 2008;18(4):585-596. doi: 10.1101/gr.074534.107.
19. Meyer E, Butler A, Dubrana K, et al. Sequence-specific epigenetic effects of the maternal somatic genome on developmental rearrangements of the zygotic genome in *Paramecium primaurelia*. *Mol Cell Biol*. 1996;17(7):3589-3599. doi: 10.1128/MCB.17.7.3589.
20. Kloutcher LA, Herrick G. Consensus inverted terminal repeat sequence of *Paramecium* IESs: resemblance to termini of Tc1-related and *Euplotes* Tec transposons. *Nucleic Acids Res*. 1995;23(11):2006-13. doi: 10.1093/nar/23.11.2006.
21. Eisen JA, Coyne RS, Wu M, et al. Macronuclear genome sequence of the ciliate *Tetrahymena thermophila*, a model eukaryote. *PLoS Biol*. 2006;4(9): e286. doi: 10.1371/journal.pbio.0040286.
22. Aury JM, Jaillon O, Duret L, et al. Global trends of whole-genome duplications revealed by the ciliate *Paramecium tetraurelia*. *Nature*. 2006;444(7116):171-8. doi: 10.1038/nature05230.
23. Nekrasova IV, Przybos E, Rautian MS, Potekhin AA. Electrophoretic karyotype polymorphism of sibling species of the *Paramecium aurelia* complex. *J Eukaryot Microbiol*. 2010;57(6):494-507. doi: 10.1111/j.1550-7408.2010.00507.x.
24. Fan Q, Yao MC. A long stringent sequence signal for programmed chromosome breakage in *Tetrahymena thermophila*. *Nucleic Acids Res*. 2000;28(4):895-900. doi: 10.1093/nar/28.4.895.
25. Fan Q, Yao MC. New telomere formation coupled with site-specific chromosome breakage in *Tetrahymena thermophila*. *Mol Cell Biol*. 1996;16(3):1267-1274. doi: 10.1128/MCB.16.3.1267.
26. Yao MC, Zheng K, Yao CH. A conserved nucleotide sequence at the sites of developmentally regulated chromosomal breakage in *Tetrahymena*. *Cell*. 1987;48(5):779-788. doi: 10.1016/0092-8674(87)90075-4.
27. Cranert S, Heyse S, Linger BR, et al. *Tetrahymena* Pot2 is a developmentally regulated paralog of Pot1 that localizes to chromosome breakage sites but not to telomeres. *Eukaryot Cell*. 2014;13(12):1519-1529. doi: 10.1128/EC.00204-14.
28. Le Mouël A, Butler A, Caron F, Meyer E. Developmentally regulated chromosome fragmentation linked to imprecise elimination of repeated sequences in *paramecia*. *Eukaryotic Cell*. 2003;2(5):1076-1090. doi: 10.1128/EC.2.5.1076-1090.2003.
29. Caron F. A high degree of macronuclear chromosome polymorphism is generated by variable DNA rearrangements in *Paramecium primaurelia* during macronuclear differentiation. *J Mol Biol*. 1992;225(3):661-78. doi: 10.1016/0022-2836(92)90393-X.
30. Duhaucourt S, Butler A, Meyer E. Epigenetic self-regulation of developmental excision of an internal eliminated sequence in *Paramecium tetraurelia*. *Genes Dev*. 1995;9(16):2065-2077. doi: 10.1101/gad.9.16.2065.
31. Duhaucourt S, Keller A-M, Meyer E. Homology-dependent maternal inhibition of developmental excision of internal eliminated sequences in *Paramecium tetraurelia*. *Mol Cell Biol*. 1998;18(12):7075-7085. doi: 10.1128/MCB.18.12.7075.
32. Chalker DL, Yao MC. Non-Mendelian, heritable blocks to DNA rearrangement are induced by loading the somatic nucleus of *Tetrahymena thermophila* with germline-limited DNA. *Mol Cell Biol*. 1996;16(7):3658-67. doi: 10.1128/MCB.16.7.3658.
33. Lhuillier-Akakpo M, Frapporti A, Wilkes CD, et al. Local effect of enhancer of Zeste-like reveals cooperation of epigenetic and cis-acting determinants for zygotic genome rearrangements. *PLoS Genet*. 2014;10(9):e1004665. doi: 10.1371/journal.pgen.1004665.
34. Garnier O, Serrano V, Duhaucourt S, Meyer E. RNA-mediated programming of developmental genome rearrangements in *Paramecium tetraurelia*. *Mol Cell Biol*. 2004; 24(17):7370-7379. doi: 10.1128/MCB.24.17.7370-7379.2004.
35. Forney JD, Yantiri F, Mikami K. Developmentally controlled rearrangement of surface protein genes in *Paramecium tetraurelia*. *J Eukaryot Microbiol*. 1996;43(6):462-467. doi: 10.1111/j.1550-7408.1996.tb04505.x.
36. Mochizuki K, Fine NA, Fujisawa T, Gorovsky MA. Analysis of a piwi-related gene implicates small RNAs in genome rearrangement in *Tetrahymena*.

- Cell*. 2002;110(6):689-699. doi: 10.1016/S0092-8674(02)00909-1.
37. Noto T, Mochizuki K. Small RNA-Mediated trans-Nuclear and trans-Element Communications in *Tetrahymena* DNA Elimination. *Curr Biol*. 2018;28(12):1938-49 e1935. doi: 10.1016/j.cub.2018.04.071.
  38. Gruchota J, Denby Wilkes C, Arnaiz O, et al. A meiosis-specific Spt5 homolog involved in non-coding transcription. *Nucleic Acids Res*. 2017;45(8):4722-4732. doi: 10.1093/nar/gkw1318.
  39. Elbashir SM, Harborth J, Lendeckel W, et al. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature*. 2001;411(6836):494-498. doi: 10.1038/35078107.
  40. Cerutti H, Casas-Mollano JA. On the origin and functions of RNA-mediated silencing: from protists to man. *Curr Genet*. 2006;50(2):81-99. doi: 10.1007/s00294-006-0078-x.
  41. Lee SR, Collins K. Two classes of endogenous small RNAs in *Tetrahymena thermophila*. *Genes Dev*. 2006;20(1):28-33. doi: 10.1101/gad.1377006.
  42. Mochizuki K, Gorovsky MA. A Dicer-like protein in *Tetrahymena* has distinct functions in genome rearrangement, chromosome segregation, and meiotic prophase. *Genes Dev*. 2005;19(1):77-89. doi: 10.1101/gad.1265105.
  43. Malone CD, Anderson AM, Motl JA, et al. Germ line transcripts are processed by a Dicer-like protein that is essential for developmentally programmed genome rearrangements of *Tetrahymena thermophila*. *Mol Cell Biol*. 2005;25(20):9151-9164. doi: 10.1128/MCB.25.20.9151-9164.2005.
  44. Sandoval PY, Swart EC, Arambasic M, Nowacki M. Functional diversification of Dicer-like proteins and small RNAs required for genome sculpting. *Dev Cell*. 2014;28(2):174-188. doi: 10.1016/j.devcel.2013.12.010.
  45. Marker S, Carradec Q, Tanty V, et al. A forward genetic screen reveals essential and non-essential RNAi factors in *Paramecium tetraurelia*. *Nucleic Acids Res*. 2014;42(11):7268-7280. doi: 10.1093/nar/gku223.
  46. Lepere G, Nowacki M, Serrano V, et al. Silencing-associated and meiosis-specific small RNA pathways in *Paramecium tetraurelia*. *Nucleic Acids Res*. 2009;37(3):903-915. doi: 10.1093/nar/gkn1018.
  47. Hoehener C, Hug I, Nowacki M. Dicer-like Enzymes with Sequence Cleavage Preferences. *Cell*. 2018;173(1):234-247 e237. doi: 10.1016/j.cell.2018.02.029.
  48. Kurth HM, Mochizuki K. 2'-O-methylation stabilizes Piwi-associated small RNAs and ensures DNA elimination in *Tetrahymena*. *RNA*. 2009;15(4):675-685. doi: 10.1261/rna.1455509.
  49. Bouhouche K, Gout JF, Kapusta A, et al. Functional specialization of Piwi proteins in *Paramecium tetraurelia* from post-transcriptional gene silencing to genome remodelling. *Nucleic Acids Res*. 2011;39(10):4249-64. doi: 10.1093/nar/gkq1283.
  50. Couvillion MT, Lee SR, Hogstad B, et al. Sequence, biogenesis, and function of diverse small RNA classes bound to the Piwi family proteins of *Tetrahymena thermophila*. *Genes Dev*. 2009;23(17):2016-2032. doi: 10.1101/gad.1821209.
  51. Cerutti L, Mian N, Bateman A. Domains in gene silencing and cell differentiation proteins: the novel PAZ domain and redefinition of the Piwi domain. *Trends Biochem Sci*. 2000;25(10):481-482. doi: 10.1016/S0968-0004(00)01641-8.
  52. Matranga C, Tomari Y, Shin C, et al. Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes. *Cell*. 2005;123(4):607-620. doi: 10.1016/j.cell.2005.08.044.
  53. Mochizuki K, Gorovsky MA. Conjugation-specific small RNAs in *Tetrahymena* have predicted properties of scan (scn) RNAs involved in genome rearrangement. *Genes Dev*. 2004;18(17):2068-2073. doi: 10.1101/gad.1219904.
  54. Liu Y, Taverna SD, Muratore TL, et al. RNAi-dependent H3K27 methylation is required for heterochromatin formation and DNA elimination in *Tetrahymena*. *Genes Dev*. 2007;21(12):1530-1545. doi: 10.1101/gad.1544207.
  55. Furrer DI, Swart EC, Kraft MF, et al. Two Sets of Piwi Proteins Are Involved in Distinct sRNA Pathways Leading to Elimination of Germline-Specific DNA. *Cell Rep*. 2017;20(2):505-520. doi: 10.1016/j.celrep.2017.06.050.
  56. Nowacki M, Zagorski-Ostojka W, Meyer E. Nowa1p and Nowa2p: novel putative RNA binding proteins involved in trans-nuclear crosstalk in *Paramecium tetraurelia*. *Curr Biol*. 2005;15(18):1616-1628. doi: 10.1016/j.cub.2005.07.033.
  57. Ignarski M, Singh A, Swart EC, et al. *Paramecium tetraurelia* chromatin assembly factor-1-like protein PtCAF-1 is involved in RNA-mediated control of DNA elimination. *Nucleic Acids Res*. 2014;42(19):11952-11964. doi: 10.1093/nar/gku874.
  58. Lepere G, Betermier M, Meyer E, Duharcourt S. Maternal noncoding transcripts antagonize the targeting of DNA elimination by scanRNAs in *Paramecium tetraurelia*. *Genes Dev*. 2008;22(11):1501-1512. doi: 10.1101/gad.473008.
  59. Bednenko J, Noto T, DeSouza LV, et al. Two GW repeat proteins interact with *Tetrahymena thermophila* argonaute and promote genome rearrangement. *Mol Cell Biol*. 2009;29(18):5020-5030. doi: 10.1128/MCB.00076-09.
  60. Maliszewska-Olejniczak K, Gruchota J, Gromadka R, et al. TFIIS-Dependent Non-coding Transcription Regulates Developmental Genome Rearrangements.



- PLoS Genet.* 2015;11(7):e1005383. doi: 10.1371/journal.pgen.1005383.
61. Miao W, Xiong J, Bowen J, et al. Microarray analyses of gene expression during the *Tetrahymena thermophila* life cycle. *PLoS One.* 2009;4(2):e4429. doi: 10.1371/journal.pone.0004429.
  62. Aronica L, Bednenko J, Noto T, et al. Study of an RNA helicase implicates small RNA-noncoding RNA interactions in programmed DNA elimination in *Tetrahymena*. *Genes Dev.* 2008;22(16):2228-2241. doi: 10.1101/gad.481908.
  63. Nowak JK, Gromadka R, Juszczuk M, et al. Functional study of genes essential for autogamy and nuclear reorganization in *Paramecium*. *Eukaryot Cell.* 2011;10(3):363-372. doi: 10.1128/EC.00258-10.
  64. Liu Y, Taverna SD, Muratore TL, et al. RNAi-dependent H3K27 methylation is required for heterochromatin formation and DNA elimination in *Tetrahymena*. *Genes Dev.* 2007;21(12):1530-1545. doi: 10.1101/gad.1544207.
  65. Liu Y, Mochizuki K, Gorovsky MA. Histone H3 lysine 9 methylation is required for DNA elimination in developing macronuclei in *Tetrahymena*. *Proc Natl Acad Sci USA.* 2004;101(6):1679-1684. doi: 10.1073/pnas.0305421101.
  66. Matsuda A, Forney JD. The SUMO pathway is developmentally regulated and required for programmed DNA elimination in *Paramecium tetraurelia*. *Eukaryot Cell.* 2006;5(5):806-815. doi: 10.1128/EC.5.5.806-815.2006.
  67. Madireddi MT, Coyne RS, Smothers JF, et al. Pdd1p, a novel chromodomain-containing protein, links heterochromatin assembly and DNA elimination in *Tetrahymena*. *Cell.* 1996;87(1):75-84. doi: 10.1016/S0092-8674(00)81324-0.
  68. Coyne RS, Nikiforov MA, Smothers JF, et al. Parental expression of the chromodomain protein Pdd1p is required for completion of programmed DNA elimination and nuclear differentiation. *Cell.* 1999;4(5):865-872. doi: 10.1016/S1097-2765(00)80396-2.
  69. Taverna SD, Coyne RS, Allis CD. Methylation of histone H3 at lysine 9 targets programmed DNA elimination in *Tetrahymena*. *Cell.* 2002;110(6):701-711. doi: 10.1016/S0092-8674(02)00941-8.
  70. Nikiforov MA, Smothers JF, Gorovsky MA, Allis CD. Excision of micronuclear-specific DNA requires parental expression of pdd2p and occurs independently from DNA replication in *Tetrahymena thermophila*. *Genes Dev.* 1999;13(21):2852-2862.
  71. Yao MC, Yao CH, Halasz LM, et al. Identification of novel chromatin-associated proteins involved in programmed genome rearrangements in *Tetrahymena*. *J Cell Sci.* 2007;120(Pt 12):1978-1989. doi: 10.1242/jcs.006502.
  72. Janetopoulos C, Cole E, Smothers JF, et al. The conjusome: a novel structure in *Tetrahymena* found only during sexual reorganization. *J Cell Sci.* 1999;112:1003-1011.
  73. Baudry C, Malinsky S, Restituto M, et al. PiggyMac, a domesticated piggyBac transposase involved in programmed genome rearrangements in the ciliate *Paramecium tetraurelia*. *Genes Dev.* 2009;23(21):2478-83. doi: 10.1101/gad.547309.
  74. Gratiás A, Bétermier M. Processing of double-strand breaks is involved in the precise excision of *Paramecium* IESs. *Mol Cell Biol.* 2003;23(20):7152-7162. doi: 10.1128/MCB.23.20.7152-7162.2003.
  75. Elick TA, Bauser CA, Fraser MJ. Excision of the piggyback transposable element *in vitro* is a precise event that is enhanced by the expression of its encoded transposase. *Genetica.* 1996;98(1):33-41. doi: 10.1007/BF00120216.
  76. Fraser MJ, Ciszczon T, Elick T, Bauser C. Precise excision of TTAA-specific lepidopteran transposons piggyBac (IFP2) and tagalong (TFP3) from the baculovirus genome in cell lines from two species of *Lepidoptera*. *Insect Mol Biol.* 1996;5(2):141-151. doi: 10.1111/j.1365-2583.1996.tb00048.x.
  77. Arnaiz O, Cain S, Cohen J, Sperling L. Paramecium-DB: a community resource that integrates the *Paramecium tetraurelia* genome sequence with genetic data. *Nucleic Acids Res.* 2007;35(Database issue):D439-444. doi: 10.1093/nar/gkl777.
  78. Bouallegue M, Rouault JD, Hua-Van A, et al. Molecular Evolution of piggyBac Superfamily: From Selfishness to Domestication. *Genome Biol Evol.* 2017;9(2):323-339. doi: 10.1093/gbe/evw292.
  79. Henssen AG, Koche R, Zhuang J, et al. PGBD5 promotes site-specific oncogenic mutations in human tumors. *Nat Genet.* 2017;49(7):1005-1014. doi: 10.1038/ng.3866.
  80. Dubois E, Mathy N, Regnier V, et al. Multimerization properties of PiggyMac, a domesticated piggyBac transposase involved in programmed genome rearrangements. *Nucleic Acids Res.* 2017;45(6):3204-16. doi: 10.1093/nar/gkw1359.
  81. Bischerour J, Bhullar S, Denby Wilkes C, et al. Six domesticated PiggyBac transposases together carry out programmed DNA elimination in *Paramecium*. *Elife.* 2018;7. doi: 10.7554/eLife.37927.
  82. Saveliev SV, Cox MM. Developmentally programmed DNA deletion in *Tetrahymena thermophila* by a transposition-like reaction pathway. *EMBO J.* 1996;15(11):2858-2869. doi: 10.1002/j.1460-2075.1996.tb00647.x.
  83. Cheng CY, Vogt A, Mochizuki K, Yao MC. A domesticated piggyBac transposase plays key roles in heterochromatin dynamics and DNA cleavage during programmed DNA deletion in *Tetrahymena thermophila*. *Mol Biol Cell.* 2010;21(10):1753-1762. doi: 10.1091/mbc.E09-12-1079.

84. Cheng CY, Young JM, Lin CG, et al. The piggyBac transposon-derived genes TPB1 and TPB6 mediate essential transposon-like excision during the developmental rearrangement of key genes in *Tetrahymena thermophila*. *Genes Dev.* 2016;30(24):2724-2736. doi: 10.1101/gad.290460.116.
85. Shieh AW, Chalker DL. LIA5 is required for nuclear reorganization and programmed DNA rearrangements occurring during tetrahymena macronuclear differentiation. *PLoS One.* 2013;8(9):e75337. doi: 10.1371/journal.pone.0075337.
86. Kapusta A, Matsuda A, Marmignon A, et al. Highly precise and developmentally programmed genome assembly in *Paramecium* requires ligase IV-dependent end joining. *PLoS Genet.* 2011;7(4):e1002049. doi: 10.1371/journal.pgen.1002049.
87. Matsuda A, Shieh AW, Chalker DL, Forney JD. The conjugation-specific Die5 protein is required for development of the somatic nucleus in both *Paramecium* and *Tetrahymena*. *Eukaryot Cell.* 2010;9(7):1087-99. doi: 10.1128/EC.00379-09.
88. Henikoff S, Dalal Y. Centromeric chromatin: what makes it unique? *Curr Opin Genet Dev.* 2005;15(2):177-184. doi: 10.1016/j.gde.2005.01.004.
89. Smith MM. Centromeres and variant histones: what, where, when and why? *Curr Opin Cell Biol.* 2002;14(3):279-285. doi: 10.1016/S0955-0674(02)00331-9.
90. Cui B, Gorovsky MA. Centromeric histone H3 is essential for vegetative cell division and for DNA elimination during conjugation in *Tetrahymena thermophila*. *Mol Cell Biol.* 2006;26(12):4499-4510. doi: 10.1128/MCB.00079-06.
91. Arambasic M, Sandoval PY, Hoehener C, et al. Pdsg1 and Pdsg2, novel proteins involved in developmental genome remodelling in *Paramecium*. *PLoS One.* 2014;9(11):e112899. doi: 10.1371/journal.pone.0112899.
92. Cummings DJ, Tait A, Goddard JM. Methylated bases in DNA from *Paramecium aurelia*. *Biochim Biophys Acta.* 1974;374(1):1-11. doi: 10.1016/0005-2787(74)90194-4.
93. Harrison GS, Karrer KM. DNA synthesis, methylation and degradation during conjugation in *Tetrahymena thermophila*. *Nucleic Acids Res.* 1985;13(1):73-87. doi: 10.1093/nar/13.1.73.
94. Epigenetics. Ed. by C.D. Allis, M.L. Caparros, T. Jenkinson, D. Reinberg. Cold Spring Harbor: Cold Spring Harbor Laboratory Press; 2015.
95. Sonneborn TM. Sex, sex inheritance and sex determination in *Paramecium aurelia*. *Proc Natl Acad Sci USA.* 1937;23(7):378-385. doi: 10.1073/pnas.23.7.378.
96. Singh DP, Saudemont B, Guglielmi G, et al. Genome-defence small RNAs exapted for epigenetic mating-type inheritance. *Nature.* 2014;509(7501):447-452. doi: 10.1038/nature13318.
97. Orias E, Singh DP, Meyer E. Genetics and Epigenetics of Mating Type Determination in *Paramecium* and *Tetrahymena*. *Annu Rev Microbiol.* 2017;71:133-156. doi: 10.1146/annurev-micro-090816-093342.

✉ Информация об авторах

**Ирина Владимировна Некрасова** — доцент, биологический факультет. Санкт-Петербургский государственный университет, Санкт-Петербург. E-mail: alexey.potekhin@spbu.ru.

**Алексей Анатольевич Потехин** — профессор, биологический факультет. Санкт-Петербургский государственный университет, Санкт-Петербург. E-mail: alexey.potekhin@spbu.ru.

✉ Information about the authors

**Irina V. Nekrasova** — Associate Professor, Faculty of Biology. Saint Petersburg State University, Saint Petersburg, Russia. E-mail: alexey.potekhin@spbu.ru.

**Alexey A. Potekhin** — Professor, Faculty of Biology. Saint Petersburg State University, Saint Petersburg, Russia. E-mail: alexey.potekhin@spbu.ru.