

СРАВНИТЕЛЬНЫЙ СТРУКТУРНЫЙ АНАЛИЗ ТРЕХ КЛАСТЕРОВ piРНК В ГЕНОМЕ *Drosophila melanogaster*

© 2020 г. А. П. Котнова^{а, *}, Ю. В. Ильин^а

^аИнститут молекулярной биологии им. В.А. Энгельгардта Российской академии наук, Москва, 119991 Россия

*e-mail: alina_kotnova@mail.ru

Поступила в редакцию 17.12.2019 г.

После доработки 27.01.2020 г.

Принята к публикации 27.01.2020 г.

В статье предпринята попытка ретроспективного воссоздания последовательности событий, приведших к образованию трех регуляторных piРНК-кластеров у *Drosophila melanogaster*. Исследованы piРНК-кластеры 20A, 38C и *flamenco*. Построены схемы их последовательного формирования и выявлены структурные особенности. В двух из трех исследованных кластеров, 38C и *flamenco*, обнаружены инвертированные последовательности, потенциально способные приводить к образованию двухцепочечных шпилек в РНК-транскрипте. Приведены доводы в пользу известной гипотезы о piРНК-кластерах как о “ловушках транспозонов”. Согласно этой модели, присутствие в геноме только одной копии транспозона свидетельствует о том, что его экспрессия была подавлена механизмами РНК-интерференции сразу после попадания мобильного элемента в piРНК-кластер. Приведены доказательства высокой структурной изменчивости piРНК-кластеров дрозофилы и событий горизонтального переноса мобильных элементов между родственными видами.

Ключевые слова: piРНК-кластеры, мобильные элементы, РНК-интерференция, репрессия транскрипционных позиций, *flamenco*, piРНК-кластер 20A, piРНК-кластер 38C, *Drosophila melanogaster*

DOI: 10.31857/S0026898420030088

ВВЕДЕНИЕ

Геном *Drosophila melanogaster* населяет множество мобильных элементов, составляющих до 22% от всего генома [1]. Интактные, потенциально активные, копии мобильных элементов локализируются как в эухроматине, так и в прицентромерном и теломерном гетерохроматине. Кроме них в геноме дрозофилы присутствуют “молекулярные ископаемые” – дефектные копии, давно утратившие способность к перемещениям. По последним данным, до 77% гетерохроматина дрозофилы представлено мобильными элементами. Большая их часть разрушена инсерциями других транспозонов. Они не способны автономно перемещаться и сконцентрированы в прицентромерных областях хромосом. Эти регионы, первоначально рассматриваемые как кладбища мобильных элементов, вовлечены в систему защиты генома от нестабильности и исполняют роль регуляторов транскрипций [2, 3]. Такие кластеры транскрибируются, после чего транскрипты подвергаются процессингу и образуют короткие РНК, взаимодействующие с белком Piwi (piРНК), и эндогенные малые интерферирующие РНК (endo-siРНК). Короткие РНК в свою очередь приводят к блокировке перемещений мобильных элементов

в гонадах и соматических тканях. Защитные механизмы, основанные на РНК-сайленсинге и направленные против распространения в геноме мобильных элементов, очень консервативны у эукариот [4].

На сегодняшний день крайне мало известно о судьбе деградировавших копий мобильных элементов, несмотря на их важную роль в формировании кластеров, принимающих участие в РНК-интерференции. В данной статье мы попытались воссоздать последовательность образования таких piРНК-кластеров.

ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

В работе использована пятая сборка генома *D. melanogaster*, доступная в базе данных FlyBase (ftp://ftp.flybase.net/genomes/Drosophila_melanogaster/dmel_r5.49_FB2013_01/), и “канонический” набор последовательностей мобильных элементов (https://flybase.org/static_pages/downloads/FB2015_02/transposons/transposon_sequence_set.embl.txt.gz). Для выравнивания нуклеотидных последовательностей использовали программы BLAST+ и MEGA5.

Выравнивание протяженных последовательностей рiРНК-кластеров против всего набора мобильных элементов с помощью BLAST+ позволяет легко найти интактные копии мобильных элементов, а также их крупные фрагменты. В случае сильного повреждения копии мобильного элемента в результате множества делеций и вставок алгоритм BLAST способен найти и выровнять с канонической последовательностью только короткие ее фрагменты, которые получают небольшой скоринг и оказываются беспорядочно разбросанными в огромном массиве данных. С целью найти такие дефектные копии мы сначала сортировали результаты работы BLAST+ не по скорингу, а по положению внутри последовательности рiРНК-кластера, а затем обрабатывали эти данные с помощью нескольких простейших скриптов и вручную.

В каноническом наборе есть мобильные элементы, очень близкие друг другу. Их последовательности схожи, поэтому при сравнении набора с какой-то третьей последовательностью с помощью алгоритма BLAST можно обнаружить в ней участки, схожие не с одним, а сразу с несколькими мобильными элементами. В первую очередь мы отфильтровывали такие перекрывающиеся результаты BLAST+, оставив в наборе данных варианты, имеющие наибольший скоринг и протяженность. Те области рiRNA кластеров, в которых несколько фрагментов одного мобильного элемента были расположены близко друг к другу, анализировали вручную. Если эти фрагменты оказывались однонаправленными и следовали в геноме в точности в той же последовательности, что и в каноническом варианте мобильного элемента, или с небольшими перестройками, этот участок идентифицировали как дефектную копию мобильного элемента.

Предположения об очередности встраивания мобильных элементов в геном строили на основании анализа последовательностей, фланкирующих вставку мобильного элемента. Если элемент А был фланкирован в геноме справа и слева последовательностями, похожими на элемент Б, причем точка разрыва элемента Б совпадала слева и справа (с учетом дубликации сайта-мишени), то мы предполагали, что элемент А встроился в элемент Б.

Предположения о времени встраивания строили только для некоторых ретротранспозонов, имеющих длинные концевые повторы (LTR). Механизм их перемещения таков, что LTR у новой копии ретротранспозона в геноме всегда будут идентичными. Наличие внутри копии перестроек, к сожалению, не позволяет судить о времени ее перемещения, так как известны случаи перемещения дефектных копий ретротранспозонов с нарушенной кодирующей частью за счет активности белков, синтезированных на других, полноценных, копиях этого элемента.

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

В 2007 году Brenneke с соавт. [2] продемонстрировали, что последовательности, с которых транскрибируются рiРНК в геноме дрозофилы, не разбросаны по геному, а кластеризуются в небольшом числе дискретных локусов. Они определили 142 места в геноме, активно генерирующих короткие РНК. Эти сайты продуцируют 81% всех уникальных рiРНК. Среди указанных 142 кластеров рiРНК только 7 расположены в потенциально эухроматических областях, а остальные локализуются в прицентромерном и теломерном гетерохроматине. Размеры этих кластеров у *D. melanogaster* существенно различаются. Наибольшую известность среди них приобрел локус *flamenco*, охарактеризованный как регулятор транспозиций мобильных элементов задолго до открытия механизма РНК-интерференции.

В конце 1980-х—начале 1990-х было получено несколько генетически нестабильных линий *D. melanogaster*, в яйцниках которых активно перемещался ретротранспозон *gypsy*. Оказалось, что нестабильность в этих линиях была вызвана комбинацией двух факторов: 1) присутствием транспозиционно активной копии *gypsy* и 2) мутацией, регулирующей перемещения этого мобильного элемента. Генетическими методами эта мутация была картирована в позиции 65,9 X-хромосомы (секция 20A1–A3) — в непосредственной близости от β-гетерохроматина, накапливающего множество деградировавших мобильных элементов [5]. Локус получил название *flamenco*, поскольку “заставлял цыгана (*gypsy*) танцевать”. Примерно в этой же области X-хромосомы в 1997 году картировали мутацию, контролирующую перемещения двух других ретротранспозонов: *ZAM* и *Idefix*. В связи с тем, что в мутантной линии с активно перемещающимися *ZAM* и *Idefix* не было обнаружено активных перемещений *gypsy*, *Gorgiaux* и др. [5] считали, что имеют дело с двумя различными генными локусами, расположенными по соседству: *flamenco* и *COM*. В последующие годы в лабораториях разных стран проводились интенсивные, но не увенчавшиеся успехом исследования, направленные на клонирование гена *flamenco*. Только после открытия явления РНК-интерференции появились предположения, что *flamenco* может оказаться не классическим геном-репрессором, а источником коротких регуляторных РНК. Но настоящий прорыв в понимании структуры локуса *flamenco/COM* произошел только в 2007 году, когда Brenneke с соавт. [2] охарактеризовали его как один из значимых рiРНК-кластеров.

Нами предпринята попытка исследовать формирование трех таких кластеров и определить в них последовательности инсерций различных мобильных элементов.

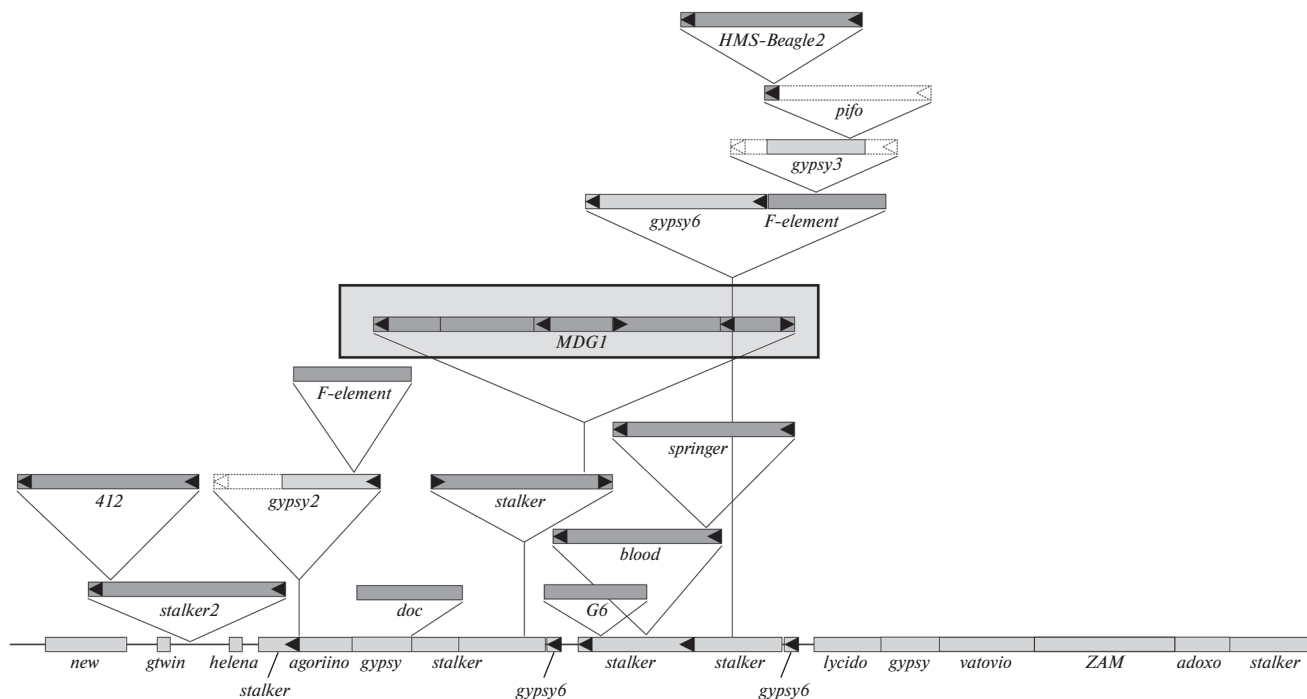


Рис. 1. Схема строения локуса *flamenco*. Светло-серым цветом обозначены деградировавшие фрагменты мобильных элементов (“молекулярные ископаемые”), темно-серым — копии, имеющие значительное сходство с каноническими. Не закрашены делетированные области. Треугольниками обозначены длинные концевые повторы (LTR). Рамкой выделена область, потенциально способная формировать участки двухцепочечной РНК.

Формирование piРНК-кластера *flamenco*

Локус *flamenco*, также известный как *COM*-локус, — главный регулятор перемещений *gypsy*, а также некоторых других ретротранспозонов, в частности *ZAM* и *Idefix*. *Flamenco* имеет длину примерно 180 т.п.н. и расположен в прицентромерном гетерохроматине в области 20A1–A3 X-хромосомы. Его последовательность большей частью составлена из фрагментов мобильных элементов, принадлежащих главным образом к семейству LTR-ретротранспозонов. В яичниках дрозофилы локус *flamenco* служит источником piРНК, а в соматических тканях — источником endo-siРНК. Роль *flamenco* заключается в предотвращении экспрессии *gypsy* и других мобильных элементов, фрагменты которых расположены в этом локусе. По этой причине он считается компонентом “адаптивного иммунитета” против ретротранспозонов [4].

Изучение локуса *flamenco* сыграло важную роль в понимании того, как малые некодирующие РНК репрессируют транспозиции мобильных элементов. От других piРНК-локусов *flamenco* отличается по некоторым параметрам. Во-первых, 94% всех уникальных РНК, кодируемых этим локусом, оказались партнерами белка Piwi. Во-вторых, большинство копий ретротранспозонов встроились во *flamenco* в одной и той же ори-

ентации, поэтому локус продуцирует piРНК с заметной асимметрией цепей [5].

Следует сказать, что *flamenco* относится не только к самым изученным, но и к самым интригующим piРНК-кластерам, поэтому мы проанализировали его в первую очередь. На рис. 1 приведена схема, иллюстрирующая последовательность встраивания мобильных элементов при формировании локуса *flamenco*.

Мы предполагаем, что основой для образования локуса *flamenco* стала дублированная несколько раз последовательность, содержащая дивергировавшую копию мобильного элемента *stalker* или его предка. В данном случае речь идет не о нескольких независимых инсерциях этого мобильного элемента в близлежащие участки хромосомы, а о рекомбинационных событиях в геноме. Это подтверждается тем фактом, что внутри дублированного фрагмента обнаружены одни и те же перестройки и мутации.

На начальном этапе формирования кластера в эту последовательность встроились ретротранспозоны *ZAM*, *gypsy*, *gypsy2*, *gypsy3*, *gypsy6*, а также недавно описанные *agoriino*, *lycido*, *vatovio* и *adoxo*. Все эти элементы представлены в составе *flamenco* сильно деградировавшими фрагментами, что свидетельствует о давности произошедших событий.

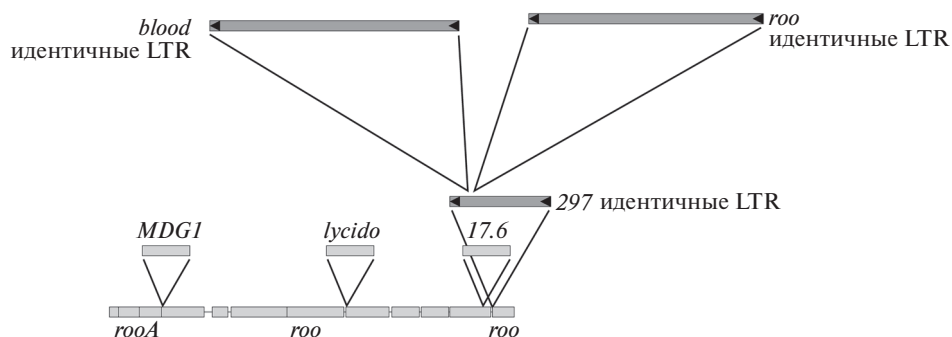


Рис. 2. Схема строения piРНК-кластера 20A. Светло-серым цветом обозначены деградировавшие фрагменты мобильных элементов (“молекулярные ископаемые”), темно-серым — копии, имеющие значительное сходство с каноническими. Треугольниками обозначены длинные концевые повторы (LTR).

Относительно недавно в локус *flamenco* попали мобильные элементы *doc*, *G6*, *F-element*, ретротранспозоны *412*, *blood*, *springer*, *HMS-Beagle2* и *MDG1*. Кроме того, нами обнаружены довольно свежие копии *stalker* и *stalker2*. По-видимому, эти события произошли в более поздний период, так как указанные мобильные элементы накопили мало замен по сравнению с образцом из канонического набора, а в LTR-ретротранспозонах присутствуют идентичные LTR. Известно, что в процессе обратной транскрипции таких элементов у новообразовавшейся копии формируются полностью идентичные друг другу концевые повторы [6, 7], поэтому по числу накопленных в LTR мутаций можно очень приблизительно оценить давность вставки.

Мобильный элемент *pifo* представлен одиночным LTR, что свидетельствует о вставке этого элемента и последующем его вырезании в результате гомологичной рекомбинации по LTR. Позднее внутри оставшегося во *flamenco* LTR *pifo* встроилась копия ретротранспозона *HMS-Beagle2*.

Довольно интересна, с нашей точки зрения, вставка в локус *flamenco* мобильного элемента *MDG1*, содержащая сложную перестройку, представляющую сочетание инверсий и дупликаций (выделена рамкой на рис. 1). При транскрипции такой последовательности в РНК могут образовываться двухцепочечные участки, потенциально способные запускать механизм РНК-интерференции.

В 2013 году Zanni и др. [8] уже исследовали локус *flamenco*, однако их внимание было сосредоточено не на происхождении этого локуса, а на его функции как регулятора экспрессии мобильных элементов. Zanni с соавт. сравнивали “работу” локуса *flamenco* в трех различных линиях дрозофилы, характеризующихся разным уровнем экспрессии ретротранспозонов *ZAM* и *Idefix* (линии w^{IR6} , *Iso1A* и *Rev*). Авторы определили, что структурные отличия локуса между линиями до-

вольно велики: многочисленные случаи гомологичной и негомологичной рекомбинации, делеции, сегментальные дупликации, утеря и новые инсерции мобильных элементов. Построенная нами карта во многом сходна с опубликованной в статье Zanni и др. [8]. Так, по мнению авторов, основной образования локуса *flamenco* послужила дублированная несколько раз последовательность элемента *stalker4*, в которую встраивались в разное время мобильные элементы *gypsy*, *MDG1*, *412*, *F*, *doc*, *ZAM* и другие. Но есть и немало отличий между предложенной нами картой и опубликованной Zanni и др. Например, в области *flamenco* нами выявлены ретротранспозоны *gtwin*, *helena*, *agorino*, *springer*, *lycido*, *vatovio* и другие, не представленные на схеме Zanni с соавт. Кроме того, нами обнаружен еще один, ранее неизвестный мобильный элемент, пока не получивший собственного названия и обозначенный на схеме как “new”. Мы не выявили эухроматических копий этого элемента в геноме *D. melanogaster*, но обнаружили их у родственных видов: *D. erecta* и *D. yakuba*. По всей видимости, мы имеем дело со случаем горизонтального переноса ретроэлемента *new* в геном *D. melanogaster*.

Формирование piРНК-кластера 20A

В силу того, что исследовать структуру первого, самого протяженного piРНК-кластера, 42AB, вышеописанным способом оказалось невозможным именно по причине его протяженности, после *flamenco* мы изучили строение кластера 20A (аннотирован под номером 2 у Brennecke с соавт. [2]). Он расположен в секции 20A X-хромосомы, его длина составляет примерно 40 т.п.н. Как и в случае *flamenco*, кластер 20A следует отнести к однонаправленным: почти все детектируемые внутри него piРНК считываются с плюс-цепи (78.4% против 0.2% по минус цепи) [2]. Схема строения кластера 20A представлена на рис. 2.

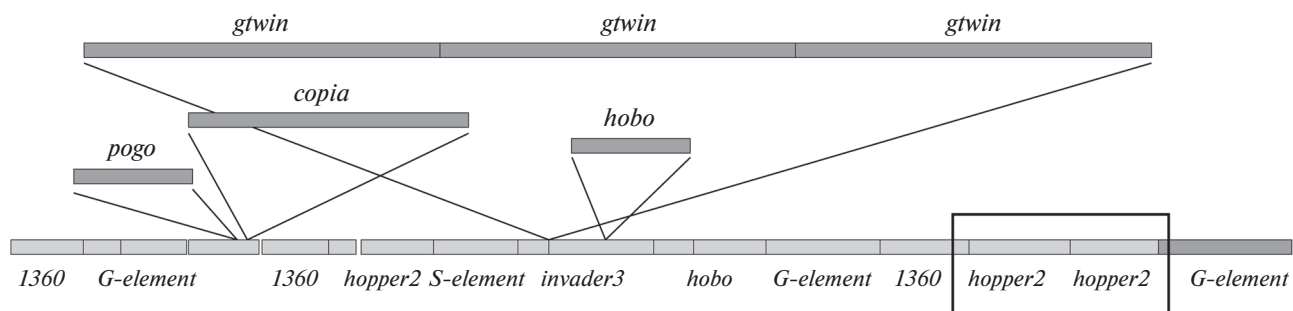


Рис. 3. Схема строения рiРНК-кластера 38С. Светло-серым цветом обозначены деградировавшие фрагменты мобильных элементов (“молекулярные ископаемые”), темно-серым — копии, имеющие значительное сходство с каноническими. Рамкой выделена область, потенциально способная формировать участки двухцепочечной РНК.

Этот кластер сформировался на базе деградировавшей копии ретротранспозона *roo*, содержащей вставки фрагментов других мобильных элементов: *MDGI*, *lycido*, *rooA* и *17.6*, — возраст которых определить не представляется возможным. Относительно недавно в этот кластер встроился мобильный элемент *297*, в который, в свою очередь, внедрились ретротранспозоны *blood* и *roo*. Хотя внутри тела этих ретротранспозонов присутствуют делеции и вставки, мы полагаем, что инсерции произошли сравнительно недавно, так как LTR этих ретротранспозонов полностью идентичны друг другу. Кроме того, ретротранспозоны с внутренними делециями в некодирующих областях иногда способны перемещаться посредством ретротранспозиции за счет активности белков, синтезированных на других, полноценных копиях этого элемента.

Формирование рiРНК-кластера 38С

Кластер 38С рiРНК (аннотирован под номером 5 у Vrenneske с соавт. [2]) заинтересовал нас как потенциальный регулятор ретротранспозиций элемента *gtwin*, который был впервые клонирован и много лет изучался в нашей лаборатории [9–14].

Этот локус протяженностью свыше 79 т.п.н. локализован в секции 38С в левом плече хромосомы 3. Свыше 63% рiРНК считываются по минус-цепи ДНК, в то время как по плюс-цепи — 23.4% [2]. Схема строения рiРНК-кластера 38С приведена на рис. 3.

Несколько деградировавших копий мобильных элементов, подвергшихся неоднократным перестройкам, стали основой для формирования рiРНК-кластера 38С. Одна из перестроек привела к дубликации значительного участка кластера, содержащего мобильные элементы *G-element*, *1360* и *hopper2*. Другая перестройка, по всей видимости, сопровождалась дубликацией и инверсией одной из копий элемента *hopper2*, что привело к

образованию структуры, на которой, предположительно, может формироваться двухцепочечная РНК (выделена рамкой на рис. 3). О рекомбинационном происхождении этих фрагментов генома свидетельствуют идентичные перестройки в одних и тех же местах мобильных элементов.

Этот локус содержит несколько относительно недавно появившихся вставок мобильных элементов, принадлежащих к различным группам (процент их гомологии с последовательностями из канонического набора очень велик, а LTR идентичны). Вероятно, последней из них была вставка трех копий ретротранспозона *gtwin*, соединенных друг с другом общими LTR.

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Не вызывает сомнения тот факт, что рiРНК-кластеры могут быть устроены по-разному, что наводит на мысль о существовании различных механизмов их образования. Еще до открытия механизмов РНК-интерференции, при изучении случаев гибридного дисгенеза, исследователи наблюдали целый спектр негативных последствий всплеск транспозиционной активности для организма-хозяина, но нестабильность всегда оказывалась временным событием. Организм хозяина успешно брал “захватчика” под свой контроль [15–18]. Для некоторых мобильных элементов — например, Р-элемента — удалось показать существование молекулы-репрессора, кодируемой самим транспозоном. Для других транспозонов — в частности *gypsy* — были картированы генные локусы, контролирующие их перемещения, однако механизм такого репрессирования долгое время оставался неизвестным. Даже для Р-элемента, авторегуляция которого за счет репрессора уже была детально изучена, обнаружили, что инсерция копии или ее фрагмента в определенные места генома может приводить к стабилизации дисгенной линии [19, 20]. С открытием РНК-интерференции стало очевидным, что любая из описанных выше моде-

лей устойчивости к транспозициям может быть напрямую связана с воздействием разных видов коротких РНК.

В представленной работе мы попытались воссоздать ретроспективную последовательность образования трех piРНК-кластеров, контролирующих перемещения мобильных элементов в геноме *D. melanogaster*, изучить их детальное строение и сравнить между собой их внутреннее устройство.

Ранее показано, что при определенных условиях даже единичные копии мобильных элементов могут стать основой для формирования piРНК-кластеров [21]. Так, Оловников с соавт. [22] показали, что попадание в геном конструкции, содержащей транскрибируемый фрагмент I-элемента, приводит к образованию нового piРНК- и endo-siРНК-продуцирующего локуса. Короткие РНК обеих полярностей генерировались как с введенного в геном участка, так и с фланкирующих геномных последовательностей; при этом транскрипты распознавались белком Piwi наравне с эндогенными I-специфичными piРНК. Подобное образование кластера piРНК, зависимое от сайта встраивания, наблюдается для всех классов мобильных элементов и добавляет новый уровень к piРНК-обеспечиваемой защите от экспансии мобильных элементов. Однако наши коллеги (Шпиз и др. [21] и Оловников и др. [22]) обнаружили, что не все копии мобильных элементов или инсерции транскриптов индуктируют образование мощных piРНК-кластеров. По всей видимости, формирование мастер-локуса *de novo* зависит от транскрипционного и хроматинного статуса мишени [23]. Можно предположить, что исследованные нами piРНК-кластеры формировались именно по такому пути: древняя инсерция мобильного элемента в определенный участок гетерохроматина по какой-то причине вызвала образование коротких РНК, после чего все последующие встроившиеся туда мобильные элементы оказывались “в ловушке”.

На сегодняшний день не удалось получить четкого представления о том, какой именно механизм играет роль пускового крючка для запуска экспрессии коротких РНК на определенных участках генома, что служит сигналом к их образованию и как регулируются эти процессы. Так, в 2016 году Guida и соавт. [4] показали, что способность локуса *flamenco* контролировать транспозиции мобильных элементов довольно изменчива и, по всей видимости, тесно связана с его хромосомным окружением, а также с “потерями” (делециями) и “приобретениями” (инсерциями) мобильных элементов внутри самого локуса. Они описали механизм косупрессии, вызванной инсерцией в эухроматин (не в локус *flamenco*!) новой копии *gypsy*, в линии *D. melanogaster*, до этого не подавляющей перемещения этого ретротранспозона.

Авторы обнаружили, что “умолкание” *gypsy* сопровождалось “умолканием” других мобильных элементов, перемещения которых также регулирует *flamenco*. Этот механизм косупрессии зависел от посттранскрипционной регуляции с вовлечением путей как endo-siРНК, так и piРНК. Предполагается, что новая инсерция *gypsy* вызвала посттранскрипционный сайленсинг как смысловых, так и бессмысловых последовательностей *gypsy*, что в итоге привело к изменениям в активности локуса *flamenco* [4].

В структуре piРНК-кластеров 20A и *flamenco* нами обнаружены инвертированные последовательности, которые могут облегчать формирование двухцепочечных участков РНК-транскрипта. Подобные структуры довольно часто можно встретить в гетерохроматине. Иногда инвертируются сравнительно небольшие участки ДНК, но в некоторых случаях речь идет о тысячах пар нуклеотидов [24]. Выполняемая ими функция на данный момент остается неясной, однако их консервативность рассматривается как косвенное свидетельство важности выполняемых ими функций. Они могут участвовать в пока неизвестных механизмах репрессии транспозиций мобильных элементов или/и принимать участие в формировании устойчивых эпигенетических связей.

В 2006 году Bergman с соавт. [25] предположили, что “гнезда” подвергшихся перестройкам транспозонов, расположенные в β -гетерохроматине, служат ловушкой для новых инсерций мобильных элементов, таким образом играя роль компонента “адаптивного иммунитета” хозяинского генома. Разные линии дрозофилы заметно отличаются друг от друга по составу мобильных элементов, попавших в кластеры piРНК, что может быть причиной наблюдаемых различий в их способности подавлять разные семейства мобильных элементов. Присутствие в piРНК-кластерах, в том числе и в локусе *flamenco*, только одной копии транспозона свидетельствует о том, что это семейство мобильных элементов было подавлено механизмами РНК-интерференции сразу после попадания в локус, что предотвратило образование новых копий [25]. В 2013 году коллектив авторов из Франции проверил гипотезу Bergman и др. на разных линиях дрозофилы. Zanni и др. [8] обнаружили, что среди 52 различных мобильных элементов, присутствующих в локусе *flamenco*, подавляющее большинство (49) представлено в виде уникальной копии [8]. Это наблюдение согласуется с выдвинутой Bergman и др. [25] гипотезой о функционировании piРНК-кластеров как “ловушек транспозонов”. Результаты проведенного нами исследования также подтверждают эту гипотезу. В случае кластера 20A, как и в случае *flamenco*, только один элемент — *roo* — встречается более одного раза. Остальные

инсерции уникальны. Возможно, древняя копия ретротранспозона *roo*, послужившая родоначальником кластера 20A, стала слишком сильно отличаться от современного варианта, поэтому в него смогла встроиться еще одна, “свежая”, копия этого мобильного элемента.

Основа кластера 38C была сформирована за счет дубликаций и инверсий участков генома, содержащих *G-element*, *1360* и *hopper2* (здесь речь идет о рекомбинационных процессах, а не о многократном встраивании в локус одного и того же транспозона). А три копии *gtwin* могли попасть в локус вместе, в виде своеобразного тримера, либо дублицироваться во время репликации. Ранее нами показано, что формирование таких структур характерно для *gtwin* [13].

Во многих исследованиях локуса *flamenco* выявлена его высокая структурная изменчивость. Так, между разными линиями *D. melanogaster* выявлено множество различий, возникающих в результате делеций, инсерций или дубликаций. Кроме того, при анализе последовательностей мобильных элементов, содержащихся в локусе *flamenco*, обнаружено, что многие из них, по всей видимости, только недавно встроились в этот локус. В работе Zanni и др. [8] в составе *flamenco* выявлено 12 новых мобильных элементов. Интересно, что 8 из них оказались близки к мобильным элементам *D. simulans*, *D. sechellia*, *D. yakuba* или *D. erecta*, что, по всей видимости, свидетельствует о случаях недавнего горизонтального переноса между видами *Drosophila*, принадлежащими к подгруппе *melanogaster* (см. [8, 12, 26] и элемент *new* в этой статье). Полученные нами результаты хорошо согласуются с этими данными: свежие инсерции мобильных элементов обнаружены и в локусе 20A (*297*, *blood*, *roo*), и в локусе 38C (*pogo*, *copia*, *hobo*, *gtwin*), и во *flamenco* (*doc*, *G6*, *F-element*, ретротранспозоны *412*, *blood*, *springer*, *HMS-Beagle2*, *MDG1* и др.). Следовательно, можно говорить о том, что рiРНК-кластеры представляют собой довольно динамичные структуры.

Таким образом, ранее описанная гипотеза о “приручении чужака” [12, 27] подтверждена и результатами проведенного нами исследования. Для мобильного элемента выгоднее всего попасть в “девственный” геном, не содержащий гомологичных ему последовательностей. Наилучший способ найти такой геном – это заразить животных близкого вида посредством горизонтального переноса. Попав в новый для себя организм, мобильный элемент не подвергается никаким репрессирующим механизмам и его копии успешно “расселяются” по геному (период нестабильности). В этом случае вставки новых копий в какой-либо мастер-локус регуляции довольно вероятны, так как, по всей видимости, в этот процесс вовлечены неизвестные пока механизмы, делающие эти ло-

кусы ловушками для новых инсерций. Попадание нового элемента в рiРНК-кластер приводит к образованию соответствующих коротких РНК, репрессирующих его дальнейшую экспрессию. Наступает период стабильности. В пользу этой гипотезы свидетельствуют данные, что в большинстве случаев в мастер-локусе регуляции присутствует только одна копия подавляемого локусом элемента. Последующие делеционные перестройки локуса могут привести к потере одной или нескольких копий мобильного элемента; при этом присутствующие в геноме копии элемента снова приобретают способность размножаться и начинается новый виток нестабильности. Таким образом, транспозиционные всплески, периоды стабильности/нестабильности напрямую связаны с мутационными событиями, затрагивающими рiРНК-кластеры, подобные *flamenco*.

Раунды высокой транспозиционной активности могут вызывать генетическую нестабильность и сопряженные с нею вредные мутации, но нет никаких сомнений, что они также играют важную роль в эволюции видов. Геном дрозофилы можно рассматривать как некий отпечаток, или даже исторический срез, таких транспозиционных всплесков для большинства семейств мобильных элементов, так как в нем присутствуют и древние, и недавно встроившиеся копии. К ним относятся и идентифицированные нами *blood*, *stalker*, *roo*; тогда как древних копий *pifo* в геноме *D. melanogaster* не обнаружено, из чего можно сделать вывод о недавно произошедшем случае его горизонтального переноса.

Таким образом, вопросы о начальных этапах формирования рiРНК-кластеров, последующих перестройках и детальных механизмах их функционирования до сих пор остаются открытыми, и для ответа на них необходимы дальнейшие исследования.

Работа поддержана грантом Программы фундаментальных исследований Президиума РАН “Постгеномные технологии и перспективные решения в биомедицине” по теме: “Сборка РНП-частиц некодирующих транскриптов и транскриптов с отложенной трансляцией”.

Настоящая статья не содержит каких-либо исследований с участием людей или животных в качестве объектов исследований.

Авторы заявляют об отсутствии конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. Kapitonov V.V., Jurka J. (2003) Molecular paleontology of transposable elements in the *Drosophila melanogaster* genome. *Proc. Natl. Acad. Sci. USA*. **100**, 6569–6574.
2. Brennecke J., Aravin A.A., Stark A., Dus M., Kellis M., Sachidanandam R., Hannon G.J. (2007) Discrete

- small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell*. **128**, 1089–1103.
3. Saito K., Nishida K.M., Mori T., Kawamura Y., Miyoshi K., Nagami T., Siomi H., Siomi M.C. (2006) Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the *Drosophila* genome. *Genes Dev.* **20**, 2214–2222.
 4. Guida V., Cernilogar F.M., Filograna A., De Gregorio R., Ishizu H., Siomi M.C., Schotta G., Belenchi G.C., Andrenacci D. (2016) Production of small noncoding RNAs from the *flamenco* locus is regulated by the *gypsy* retrotransposon of *Drosophila melanogaster*. *Genetics*. **204**(2), 631–644.
 5. Goriaux C., Théron E., Brassat E., Vaury C. (2014) History of the discovery of a master locus producing piRNAs: the *flamenco/COM* locus in *Drosophila melanogaster*. *Front. Genet.* **5**, 257. <https://doi.org/10.3389/fgene.2014.00257>
 6. Gilboa E., Mitra S., Goff S., Baltimore D. (1979) A detailed model of reverse transcription and a test of crucial aspects. *Cell*. **18**(1), 93–100.
 7. Luciw P., Leung N. (1992) Mechanisms of retroviral replication. In: *The Retroviridae*. Ed. Levy J.A. New York: Plenum Press, 159–298.
 8. Zanni V., Eymery A., Coiffet M., Zytnicki M., Luyten I., Quesneville H., Vaury C., Jensen S. (2013) Distribution, evolution, and diversity of retrotransposons at the *flamenco* locus reflect the regulatory properties of piRNA clusters. *Proc. Natl. Acad. Sci. USA*. **110**(49), 19842–19847.
 9. Котнова А.П., Карпова Н.Н., Феоктистова М.А., Любомирская Н.В., Ким А.И., Ильин Ю.В. (2005) Ретротранспозон *gtwin*: структурный анализ и распределение в линиях дрозофилы. *Генетика*. **41**(1), 23–29.
 10. Глухов И.А., Карпова Н.Н., Котнова А.П., Любомирская Н.В., Ильин Ю.В. (2004) Структурные особенности третьей открытой рамки считывания ретротранспозона *gtwin* в различных линиях *Drosophila melanogaster*. *ДАН*. **399**(2), 257–259.
 11. Котнова А.П., Феоктистова М.А., Глухов И.А., Саленко В.Б., Любомирская Н.В., Ким А.И., Ильин Ю.В. (2006) *GTWIN* – ретротранспозон дрозофилы, специфичный для подгруппы *Melanogaster*. *ДАН*. **409**(5), 691–693.
 12. Kotnova A.P., Glukhov I.A., Karpova N.N., Salenko V.B., Lyubomirskaya N.V., Ilyin Yu.V. (2007) Evidence for recent horizontal transfer of *gypsy*-homologous LTR-retrotransposon *gtwin* into *Drosophila erecta* followed by its amplification with multiple aberrations. *Gene*. **396**(1), 39–45.
 13. Стефанов Ю.Э., Котнова А.П., Глухов И.А., Любомирская Н.В., Ильин Ю.В. (2009) Особенности ретротранспозиционной активности эндогенного ретровируса *gtwin* в линии Г32 *Drosophila melanogaster*. *ДАН*. **424**(4), 555–558.
 14. Саленко В.Б., Котнова А.П., Глухов И.А., Стефанов Ю.Э., Сурков С.А., Любомирская Н.В., Ильин Ю.В. (2011) Редкое подсемейство ретротранспозона *gtwin*, характеризующееся мутацией в сайте связывания тРНК-затравки для синтеза минус-цепи кДНК, амплифицировано в линии Г-32 *Drosophila melanogaster*. *ДАН*. **436**(4), 549–552.
 15. Bingham P.M., Kidwell M.G., Rubin G.M. (1982) The molecular basis of P-M hybrid dysgenesis: the role of the P element, a P-strain-specific transposone family. *Cell*. **29**, 995–1004.
 16. Engels W.R. (1989) P elements in *Drosophila melanogaster*. In: *Mobile DNA*. Eds Berg D.E., Howe M.M. Washington, DC: American Society for Microbiology, 437–484.
 17. Finnegan D.J. (1989) The I factor and I-R hybrid dysgenesis in *Drosophila melanogaster*. In: *Mobile DNA*. Eds Berg D.E., Howe M.M. Washington, DC: American Society for Microbiology, 503–517.
 18. Bucheton A. (1990) I transposable elements and I-R hybrid dysgenesis. *Trends Genet.* **6**, 16–19.
 19. Marin L., Lehmann M., Nouaud D., Izaabel H., Anxolabehere D., Ronsseay S. (2000). P-element repression in *Drosophila melanogaster* by a naturally occurring defective telomeric P copy. *Genetics*. **155**(4), 1841–1854.
 20. Stuart J.R., Haley K.J., Swedzinski D., Lockner S., Kocian P.E., Merriman P.J., Simmons M.J. (2002) Telomeric P elements associated with cytotype regulation of the P transposon family in *Drosophila melanogaster*. *Genetics*. **162**, 1641–1654.
 21. Shpiz S., Ryazansky S., Olovnikov I, Abramov Yu., Kalmykova A. (2014) Euchromatic transposon insertions trigger production of novel Pi- and endo-siRNAs at the target sites in the *Drosophila* germline. *PLoS Genet.* **10**(2), e1004138. <https://doi.org/10.1371/journal.pgen.1004138>
 22. Olovnikov I., Ryazansky S., Shpiz S., Lavrov S., Abramov Y., Vaury C., Jensen S., Kalmykova A. (2013) *De novo* piRNA cluster formation in the *Drosophila* germ line triggered by transgenes containing a transcribed transposon fragment. *Nucleic Acids Res.* **41**, 5757–5768.
 23. Akulenko N., Ryazansky S., Morgunova V., Komarov P., Olovnikov I., Vaury C., Jensen S., Kalmykova A. (2018) Transcriptional and chromatin changes accompanying *de novo* formation of transgenic piRNA clusters. *RNA*. **24**(4), 574–584.
 24. Котнова А.П., Саленко В.Б., Любомирская Н.В., Ильин Ю.В. (2009) Структурная организация гетерохроматина *D. melanogaster*: инвертированные повторы кластеров мобильных элементов. *ДАН*. **429**(1), 120–122.
 25. Bergman C.M., Quesneville H., Anxolabéhère D., Ashburner M. (2006) Recurrent insertion and duplication generate networks of transposable element sequences in the *Drosophila melanogaster* genome. *Genome Biol.* **7**(11):R112.
 26. Bartolomé C., Bello X., Maside X. (2009) Widespread evidence for horizontal transfer of transposable elements across *Drosophila* genomes. *Genome Biol.* **10**(2), R22.
 27. Salenko V.B., Kotnova A.P., Lyubomirskaya N.V., Ilyin Yu.V. (2008) *Russian-European Workshop on DNA Repair and Epigenetic Regulation of Genome Stability*, St. Petersburg, Russia. Abstract book, Abstract 23, p. 81.

COMPARATIVE ANALYSIS OF THE STRUCTURE OF THREE piRNA CLUSTERS IN THE *Drosophila melanogaster* GENOME

A. P. Kotnova^{1,*} and Yu. V. Ilyin¹

¹*Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, 119991 Russia*

**e-mail: alina_kotnova@mail.ru*

Here we attempt to reconstruct the sequence of events that led to the formation of three regulatory piRNA clusters, namely, 20A, 38C and *flamenco*, in *Drosophila melanogaster* genome. Both 38C and *flamenco* clusters include inverted sequences, which may potentially form double-stranded RNA hairpins. We present evidence in favor of the well-known hypothesis of piRNA clusters as “transposon traps”. According to this model, the presence of the only copy of the transposon in the genome indicates that its expression was suppressed by an RNA-interference mechanism immediately after the mobile element entered the piRNA cluster. We also discuss high structural variability of piRNAs in *Drosophila* clusters and the cases of horizontal transmobile elements between related species.

Keywords: piRNA clusters, mobile elements, RNA-interference, repression of transpositions, *flamenco*, 20A piRNA cluster, 38C piRNA cluster, *Drosophila melanogaster*