

ИССЛЕДОВАНИЕ ЭВОЛЮЦИИ НЕКОДИРУЮЩИХ РНК В ГЕНОМЕ *Drosophila melanogaster*

© 2015 г. Е.В. Журавлева, А.А. Миронов

Факультет биоинженерии и биоинформатики Московского государственного университета
им. М.В. Ломоносова, 119234, Москва, ГСП-1, Ленинские горы, 1/73;

Институт проблем передачи информации им. А.А. Харкевича РАН,
127051, Москва, Большой Каретный переулок, 19/1

Email: zhuravlka@mail.ru

Поступила в редакцию 16.07.15 г.

Известно, что некодирующие РНК являются важными регуляторными молекулами клетки и функционируют за счет определенной вторичной структуры. Вторичная структура молекул, определяемая последовательностью некодирующих РНК, может быть представлена такими элементами вторичной структуры, как петли, стебли, псевдоузлы. Каждый локальный участок молекулы РНК, как и вся молекула, стремится к достижению структуры с наименьшей свободной энергией. Новые мутации могут изменить свободную энергию структуры и, таким образом, изменить оптимальную структуру молекулы, что может привести к изменению функциональности. Вероятность того, что замена нуклеотида приведет к изменению структуры молекулы, определяется тем, в каком элементе структуры произошла мутация. В предыдущих исследованиях были показаны различные значения по дивергенции в спаренных и неспаренных участках, что объясняется отбором на поддержание вторичной структуры некодирующих РНК. Выявление более тонких эволюционных различий в петлях и стеблях, а также расширенный анализ, проведенный во всех наиболее значимых классах некодирующих РНК, может представлять интерес для разработки алгоритмов по поиску генов некодирующих РНК в геномах. В этой работе рассмотрен отбор в петлях и стеблях пяти основных классов некодирующих РНК генома *Drosophila melanogaster*.

Ключевые слова: нкРНК, вторичная структура, эволюция нкРНК.

Некодирующие РНК (нкРНК) – это РНК, которые не транслируются в белок. Некодирующие РНК являются важными функциональными молекулами, которые принимают участие во множестве клеточных процессов, таких как регуляция трансляции и транскрипции, химическая модификация молекул в клетке, поддержание теломера [1]. На основе структурных и функциональных особенностей часто подразделяют классы: микроРНК, мяРНК, мякРНК, длинные некодирующие РНК, регуляторные элементы.

Вторичная и четвертичная структура молекул РНК классов имеет большое значение для их функциональности. Структура РНК позволяет ей взаимодействовать с другими молекулами РНК, лигандами и белками, которые связываются с РНК [2]. Механизм регуляции может заключаться в предоставлении нкРНК дополнительных сайтов связывания или же закрытии сайтов связывания белков.

Из длинной линейной последовательности РНК потенциально возможно формирование множества различных вторичных структур. Количество возможных структур растет экспоненциально с увеличением длины последовательности [3]. Часто структура РНК является динамической и претерпевает конформационные изменения в зависимости от условий и характеристик раствора. Молекулы РНК характеризуются высокой чувствительностью и специфичностью к изменениям окружающей среды, изменения в растворе приводят к конкретным изменениям в молекулах РНК с определенной структурой. Эффективность и быстрота подобных механизмов объясняет их широкое использование в клетке, в частности в процессах клеточного ответа на изменения внутренних и внешних условий [4].

Нуклеотиды в составе РНК могут находиться в спаренном и в неспаренном состоянии. На уровне вторичной структуры можно наблюдать различные структурные элементы, такие как стебли и петли различной топологии. Свобод-

Сокращение: нкРНК – некодирующие РНК.

ная энергия структуры складывается из энергий стекинг-взаимодействий пар оснований в стеблях и энтропийного вклада неспаренных оснований в петлях. Эти термодинамические параметры экспериментально определены [5].

В ряде исследований сравнивали скорость нуклеотидных замен в позициях, соответствующих спаренным и неспаренным основаниям в структуре нкРНК. Были попытки придумать аналог оценки отбора для нкРНК, наподобие отношения доли возможных синонимических замен к доле несинонимических (dN/dS) в белках [6]. Однако петли могут быть довольно консервативны, поскольку именно петлевые участки РНК часто участвуют во взаимодействии с другими молекулами. Неспаренные сайты зрелых микроРНК эволюционируют медленнее по сравнению со спаренными [7]. Рассмотрение петель в качестве модели нейтральной эволюции зачастую ведет к ошибочным оценкам силы отбора.

Спаренные основания в стеблях должны отличаться меньшей дивергенцией с точки зрения их большей роли в поддержании консервативной вторичной структуры. Были попытки классификации позиций в стеблях на основе их термодинамических особенностей на терминальные, субтерминальные и позиции в теле стебля [6]. В этом исследовании было показано, что в рРНК, мяРНК и некоторых других классах нкРНК консервативность определенных позиций коррелирует с величиной ожидаемого термодинамического эффекта, т.е. структура поддерживается действием отрицательного отбора.

Сила действия отбора в различных элементах вторичной структуры зависит как от важности поддержания функциональной стабильной структуры в растворе, так и от поддержания сайтов связывания на нкРНК. Необходим сравнительный анализ эволюционных закономерностей для элементов структуры различных классов нкРНК, поскольку каждый класс характеризуется собственными структурными и функциональными особенностями. Улучшение понимания закономерностей действия отбора в различных элементах структуры нкРНК может значительно улучшить понимание механизмов функционирования молекул нкРНК и их идентификации в геномах, основанной в том числе на диффузионных процессах [8].

Хотя многие классы нкРНК – микроРНК, мякРНК, мяРНК, регуляторные элементы и некоторые длинные некодирующие РНК – имеют высококонсервативные вторичные структуры, другие классы – пиРНК (piwi-белок взаимодействующих РНК) – на настоящий момент счи-

таются слабоструктурированными [9]. Однако в силу небольшого количества данных об этом недавно определенном классе нкРНК исследование их эволюции представляется важной задачей современной биологии.

МАТЕРИАЛЫ И МЕТОДЫ

В данном разделе описаны процедуры подготовки данных и методология анализа. Для изучения отбора, действующего на последовательность, были посчитаны отношения dN и dS – частоты встречаемости межвидовых замен по последовательностям нкРНК различных классов и по нейтрально эволюционирующим сайтам соответственно. В качестве нейтрально эволюционирующих сайтов были рассмотрены четырежды вырожденные позиции кодонов аминокислот. Межвидовая вариация была посчитана между видами *D. melanogaster* и *D. simulans*.

Были рассмотрены нкРНК из генома *Drosophila melanogaster*. Координаты нкРНК в геноме *D. melanogaster* взяты из базы данных Rfam [10], которая включает в себя семейства нкРНК с консервативной вторичной структурой. Предсказание вторичных структур по ковариационным моделям Rfam можно считать наиболее точным, поскольку оно основано на эволюционных данных, а также на экспериментальных данных о структуре. Мы использовали средства пакета Infernal [11] для предсказания структуры нкРНК по ковариационным моделям Rfam. На основе предсказанной структуры мы разделили позиции в нкРНК на относящиеся к стеблю или петлевым участкам. В совокупности нами было рассмотрено 298 нкРНК из 140 семейств в пяти классах нкРНК: микроРНК, мяРНК, мяшРНК, длинная некодирующая РНК и регуляторные элементы. Координаты пиРНК в геноме *Drosophila melanogaster* были взяты из базы данных piRNABank [12].

Для оценки скорости эволюции последовательности нкРНК были посчитаны отношения по дивергенции и полиморфизму между петлями и стеблями к нейтрально эволюционирующим сайтам – четырежды вырожденным позициям кодонов аминокислот. Для оценки параметра α – доли замен нуклеотидов, зафиксированных в результате положительного отбора, был использован эволюционный тест Макдональда–Крейтмана [13].

Значение α определено от $-\infty$ до 1. Отрицательные значения α могут свидетельствовать о недооценке этого параметра в силу присутствия слабобредных мутаций. Для того чтобы улучшить оценивание этого параметра, были

Таблица 1. Дивергенция (*D. melanogaster* vs. *D. simulans*)

	$D_{\text{нкРНК}}/D_s$, петля	$D_{\text{нкРНК}}/D_s$, стебель	Отношение (петля/стебель)
Длинные некодирующие	0,05	0,05	1,15
Регуляторные элементы	0,08	0,11	0,76
мяРНК	0,05	0,04	1,3
мякРНК	0,1	0,14	0,74
микроРНК	0,08	0,03	2,97
Все классы	0,08	0,06	1,24

Примечание. Показано отношение по дивергенции в петлях и стеблях классов нкРНК.

отфильтрованы низкочастотные полиморфизмы (с частотой < 5%), как это было предложено в работе [14].

Четырежды вырожденные позиции в геноме *Drosophila melanogaster* были получены из множественного выравнивания ортологичных кодирующих последовательностей базы данных FlyBase [15] для филогенетической группы *melanogaster*. Координаты экзонов в транскриптах, а также координаты кодирующей части последовательности в экзонах были получены при помощи средств BioMart Ensembl [16].

Координаты внутривидовых замен и их аллельные частоты по последовательностям нкРНК и нейтрально эволюционирующим позициям были посчитаны по 623 геномам *Drosophila melanogaster* [17–19]. Полное выравнивание этих геномов было взято из проекта *Drosophila Genome Nexus 1.0* [20]. Мы использовали разработанные нами скрипты для идентификации и подсчета частот по SNP.

Мы посчитали межвидовую вариацию из полногеномных выравниваний, полученных с помощью MultiZ [21] из UCSC Genome Browser [22]. Все расчеты проводили для пяти версий генома *Drosophila melanogaster*. В работе использовали конвертер координат, представленный на FlyBase [15]. Мы посчитали частоты нуклеотидных замен в спаренных и неспаренных позициях нкРНК, а также в четырежды вырожденных позициях кодонов аминокислот.

Для каждой структуры нкРНК была измерена ее свободная энергия средствами gnaeval пакета ViennaRNA [23]. Далее мы разделили нкРНК в каждом классе на две равные группы по величине удельной свободной энергии, нормализованной на длину. Данные по вариации были проанализированы в каждой из этих групп.

Мы оценивали эффект однонуклеотидных полиморфизмов на изменение энергии нкРНК. Для этого использовали два подхода. В первом сравнивали значения свободной энергии

для нкРНК с учетом и без учета замены. Во втором использовали предсказание эффекта SNP на локальный участок структуры с помощью RnaSNP [24].

В работе мы использовали разработанные нами скрипты на python и bash. Для данных по нкРНК, вариации и свободной энергии была создана база данных MySQL на кластере нашей лаборатории.

РЕЗУЛЬТАТЫ

Мы разделили позиции в последовательности нкРНК, как принадлежащие петлевым или стеблевым участкам. В каждом классе позиций – в петлях, стеблях и в четырежды вырожденных позициях кодонов – были посчитаны доли замен нуклеотидов по отношению к общему числу соответствующих сайтов.

Отношения по дивергенции и по полиморфизму для петель и стеблей к нейтрально эволюционирующим сайтам ($D_{\text{нкРНК}}/D_s$ и $P_{\text{нкРНК}}/P_s$) приведены в табл. 1 и 2. В табл. 2 приведены отношения по полиморфизмам для SNP с низкими (< 5%) и высокими (> 5%) аллельными частотами в последовательностях нкРНК.

Значение отношения по дивергенции больше единицы соответствует более высокой скорости эволюции рассматриваемой последовательности по сравнению с нейтральными сайтами, меньше – указывает на консервативность последовательности. Как петли, так и стебли в рассматриваемых классах нкРНК более консервативны по сравнению с нейтрально эволюционирующими последовательностями. При этом наиболее консервативны петли и стебли длинных некодирующих РНК, мяРНК, стебли микроРНК.

Более высокое значение отношения по дивергенции является свидетельством действия положительного отбора в соответствующих сайтах или же фиксации большого числа замен в результате нейтральной эволюции, что соответ-

Таблица 2. Полиморфизм

	$P_{\text{нкРНК}}/P_s$, петля		$P_{\text{нкРНК}}/P_s$, стебель		Отношение (петля/стебель)	
	Низкие частоты	Высокие частоты	Низкие частоты	Высокие частоты	Низкие частоты	Высокие частоты
Длинные некодирующие	0,57	0,07	0,52	0,07	1,09	0,94
Регуляторные элементы	0,16	0,03	0,33	0,05	0,47	0,53
мяРНК	1,07	0,07	1,34	0,07	0,80	0,99
мякРНК	0,90	0,09	1,23	0,13	0,73	0,68
микроРНК	0,49	0,05	0,21	0,02	2,29	2,46
пиРНК	0,07	0,02	0,07	0,02	1,00	1,04
Все классы (кроме пиРНК)	0,68	0,06	0,65	0,06	1,05	1,07

Примечание. Показано отношение по полиморфизму в петлях и стеблях классов нкРНК, для полиморфизмов на низких частотах ($< 0,05$) и для полиморфизмов на высоких частотах ($> 0,05$).

стует низкой силе отрицательного отбора в этих областях. Петли отличаются большей дивергенцией по сравнению со стеблями в классах длинных некодирующих РНК, мяРНК и микроРНК. Наибольшее различие для отношения по дивергенции в петлях и стеблях наблюдается в классе микроРНК.

Большая часть наблюдаемых полиморфизмов являются нейтральными либо почти нейтральными. Высокие значения отношения по полиморфизмам соответствуют большему количеству внутривидовых замен на рассматриваемых сайтах и являются свидетельством слабого действия отрицательного отбора. В целях разделения предположительно нейтральных и слабовредных полиморфизмов в табл. 2 приведены отдельно значения по полиморфизмам с высокими аллельными частотами – больше 5% и с низкими аллельными частотами – меньше 5%. В обоих случаях отношения по полиморфизму по петлям и стеблям практически сохраняются.

Большая полиморфность петель микроРНК в совокупности с данными по дивергенции указывают на высокую консервативность стеблей микроРНК и большое давление отрицательного отбора на этот структурный класс микроРНК в сравнении с другими классами нкРНК. Низкие значения по полиморфизмам в петлях регуляторных элементов и мякРНК по сравнению со стеблями указывают на большую консервативность петель этих классов нкРНК. Вероятно, это объясняется расположением функциональных сайтов связывания в петлевых участках этих нкРНК.

Для пиРНК отношения по полиморфизмам идентично для петель и стеблей, более того сохраняется как на высоких частотах SNP, так

Таблица 3. Доля замен, зафиксированных в результате положительного отбора (α)

	α , петли	α , стебли
Длинные некодирующие	-0,23	-0,52
Регуляторные элементы	0,67	0,52
мяРНК	-0,58	-1,08
мякРНК	0,18	0,10
микроРНК	0,30	0,16
Все классы	0,19	0,06

Примечание. Представлена доля замен, зафиксированных в результате положительного отбора, полученная в тесте Макдональда–Крейтмана для петель и стеблей нкРНК различных классов.

и на низких. Это подтверждает данные о низкой структурированности этих молекул.

Результаты теста Макдональда–Крейтмана для нкРНК приведены в табл. 3. Значения доли замен, зафиксированных в результате положительного отбора по всем рассмотренным классам нкРНК выше в петлях. Наиболее сильный положительный отбор наблюдается в классе регуляторных элементов, как по спаренным, так и по неспаренным позициям. Отрицательные значения α для некоторых классов получены в результате недооценивания доли слабовредных замен и неэффективности фильтрации низкочастотных замен в этих классах.

нкРНК во всех классах были разделены на группы по удельному значению энергии на нуклеотид, измеренному с помощью gnaeval. Значения по дивергенции и полиморфизму для нкРНК приведены в табл. 4 и 5. Несколько большей дивергенцией характеризуются петли нкРНК с более стабильной структурой и соответственно меньшими значениями свободной

Таблица 4. Дивергенция по структурам с высокой (в.э.) и низкой (н.э.) удельной энергией

	$D_{\text{нкРНК}}/D_s$, петля		$D_{\text{нкРНК}}/D_s$, стебель		Отношение (петля/стебель), н.э.	Отношение (петля/стебель), в.э.
	н.э.	в.э.	н.э.	в.э.		
Длинные некодирующие	0,12	0,11	0,06	0,13	2,13	0,87
Регуляторные элементы	0,09	0,05	0,13	0,06	0,66	0,80
мяРНК	0,10	0,08	0,12	0,06	0,88	1,30
мякРНК	0,27	0,22	0,23	0,59	1,18	0,37
микроРНК	0,17	0,14	0,04	0,02	4,17	5,78

Примечание. Некодирующие РНК были разделены в каждом классе нкРНК на две группы по величине удельной энергии структуры, отношению по дивергенции в каждом типе структуры и классе нкРНК.

Таблица 5. Полиморфизм

	$P_{\text{нкРНК}}/P_s$, петля		$P_{\text{нкРНК}}/P_s$, стебель		Отношение (петля/стебель), н.э.	Отношение (петля/стебель), в.э.
	н.э.	в.э.	н.э.	в.э.		
Длинные некодирующие	0,07	0,07	0,06	0,09	1,11	0,74
Регуляторные элементы	0,01	0,04	0,03	0,07	0,40	0,58
мяРНК	0,06	0,09	0,07	0,09	0,93	0,98
мякРНК	0,11	0,06	0,11	0,22	0,95	0,29
микроРНК	0,06	0,05	0,02	0,02	2,72	2,23

Примечание. Некодирующие РНК были разделены в каждом классе нкРНК на две группы по величине удельной энергии структуры, отношению по полиморфизму в каждом типе структуры и классе нкРНК.

Таблица 6. Доля замен, зафиксировавшихся в результате положительного отбора (α)

	α , петли		α , стебли	
	н.э.	в.э.	н.э.	в.э.
Длинные некодирующие	0,41	0,42	-0,12	0,33
Регуляторные элементы	0,87	0,14	0,78	-0,20
мяРНК	0,40	-0,08	0,44	-0,44
мякРНК	0,60	0,70	0,51	0,62
микроРНК	0,67	0,63	0,50	0,04

Примечание. Некодирующие РНК были разделены в каждом классе нкРНК на две группы по величине удельной энергии структуры (α) теста Макдональда–Крейтмана в каждом типе структуры и классе нкРНК.

энергии. В стеблях в различных классах нкРНК наблюдаются различные соотношения по дивергенции. Отношение по дивергенции в стеблях выше для более стабильных по энергии представителей классов регуляторных элементов, мяРНК, микроРНК. Обратное наблюдается для длинных нкРНК и мякРНК. Это свидетельствует о возможно более низкой роли структуры и отбора на структуру в классах длинных нкРНК и мякРНК. Большие значения по дивергенции для более стабильных нкРНК по

сравнению с менее стабильными (регуляторные элементы, мяРНК, микроРНК) говорит о либо большем положительном отборе на структуру, либо о более слабом отрицательном отборе для более стабильных по энергии структур.

Значения по полиморфизму в табл. 5 указывают на количество нейтральных замен в элементах структуры нкРНК с высокой и низкой удельной энергией. С их помощью мы посчитали значения параметра α теста Макдональда–Крейтмана по типам структуры и по энергии (табл. 6).

По результатам теста Макдональда–Крейтмана мы видим несколько ситуаций для классов нкРНК. Для нкРНК классов регуляторных элементов, мяРНК, микроРНК характерен сильный положительный отбор на структуру. Некодирующие РНК регуляторных элементов и мяРНК с низкой удельной энергией показывают более высокие значения α по петлям и стеблям по сравнению с менее стабильными структурами. Для микроРНК это утверждение верно по отношению к стеблям, но не к петлям. Доля замен нуклеотидов, зафиксировавшихся в результате положительного отбора, не отличается между петлями микроРНК с высокими и низкими значениями удельной энергии. Это под-

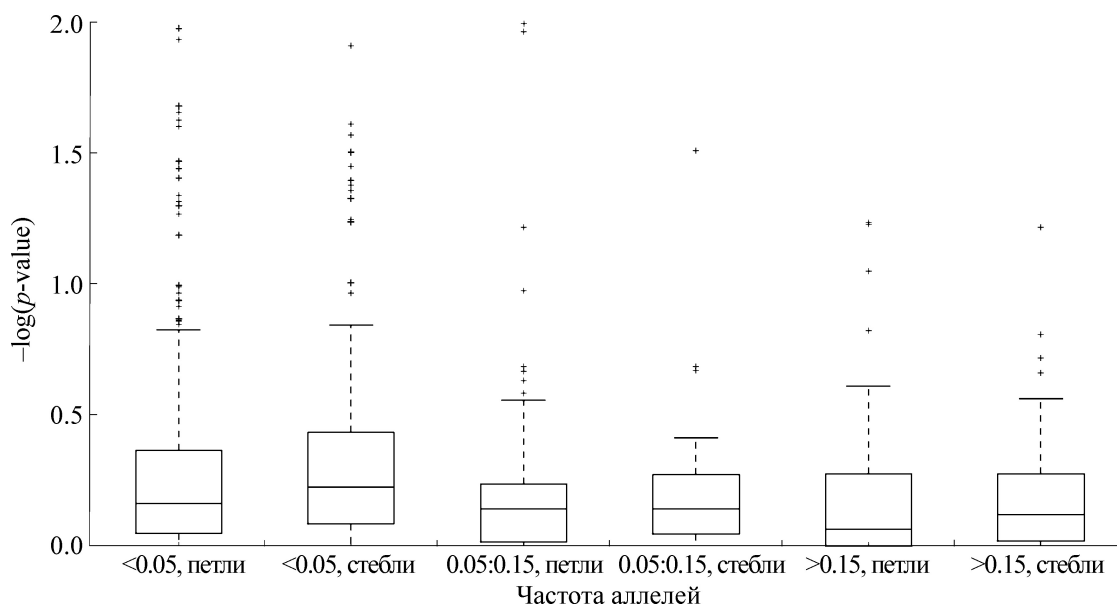


Рис. 1. Распределение p -value RnaSNP по частотам полиморфизмов и по классам структуры.

черкивает силу положительного отбора на стабильность структуры в этом классе нкРНК, которая обеспечивается за счет энергетической стабильности стеблей. При этом сила положительного отбора в петлях микроРНК сравнима для стабильных и менее стабильных молекул, что говорит о низкой силе отрицательного отбора на структуру. Для мяРНК и регуляторных элементов получены заниженные значения параметра α , наблюдается тенденция действия положительного отбора в нкРНК со стабильной структурой. Низкие значения дивергенции по петлям нкРНК с высокой удельной энергией структуры говорят о еще большем, чем для микроРНК, значении структуры в этих классах нкРНК, поскольку могут быть свидетельством не только положительного отбора, но и отрицательного отбора на структуру.

Некодирующие РНК класса малых ядрышковых очень слабо структурированы, их структура представлена большей частью петлевыми участками. Из оценки параметра α для этого класса нкРНК также следует низкая роль поддержания низкой энергии структуры этих молекул. Значения α практически не отличаются в петлях и стеблях нкРНК с различной удельной энергией. Об отсутствии отбора на структуру говорит и значение α по классу длинных нкРНК. Значение α недооценено, α выглядит существенно более низким для позиций в стеблях более стабильных по энергии нкРНК, что соответствует большой доли нейтральных межвидовых замен в этих участках структуры. Так

же как и в классе мякРНК, для длинных нкРНК структура имеет менее определяющее значение.

В табл. 7 представлено разбиение по частотам полиморфизмов в зависимости от эффекта, производимого на структуру. Эффект был оценен по изменению энергии, измеренной с g_{naeval} . Полиморфизмы, которые увеличивают энергию структуры, преобладают на низких частотах по сравнению со стабилизирующими (p -value в тесте Фишера = 0,01). В то же время на более высоких частотах преобладают полиморфизмы с положительным эффектом на структуру. Это свидетельствует о более сильном отрицательном отборе на замены нуклеотидов, уменьшающих стабильность структуры молекулы.

Результаты измерения эффекта полиморфизма на структуру, полученные с помощью RnaSNP (рис. 1), показывают, что в среднем более значимый эффект на структуру молекулы оказывают наименее частотные варианты. Это объясняется тем, что среди полиморфизмов с

Таблица 7. Стабилизирующий и дестабилизирующий эффекты SNP на структуру

Эффект	SNP, частота > 5%	SNP, частота < 5%
Дестабилизирующий	17	454
Стабилизирующий	34	441

Примечание. Показано распределение SNP по частотам в группах полиморфизмов, ухудшающих и улучшающих энергию структуры молекулы.

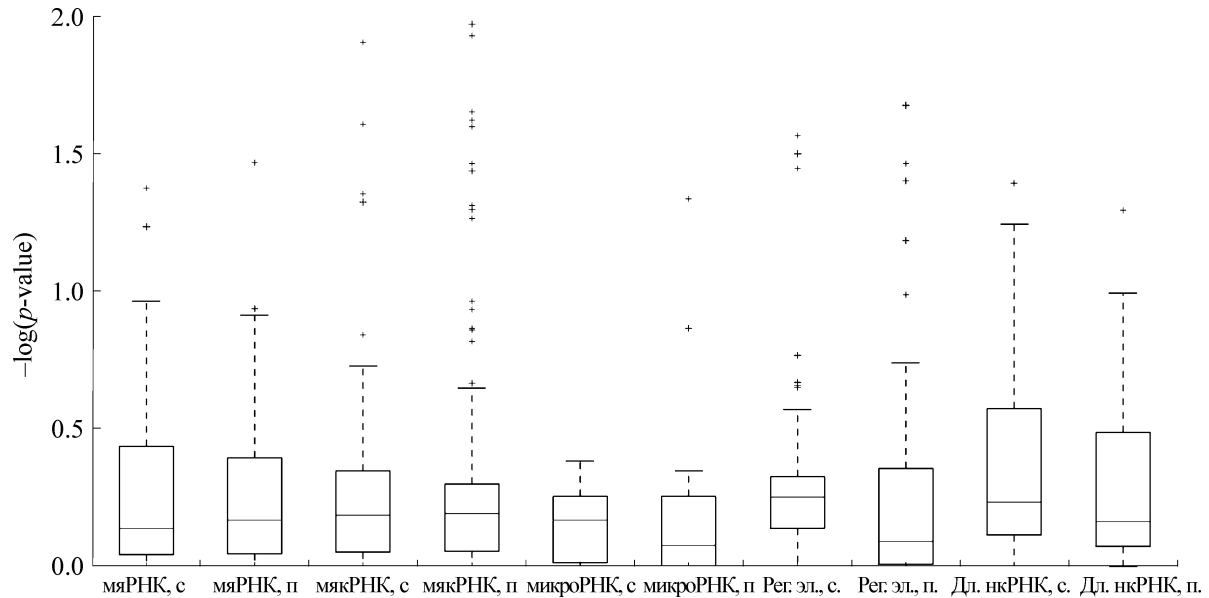


Рис. 2. Распределение p -value RnaSNP по классам нкРНК и классам структуры. Частоты полиморфизмов < 5%.

низкой частотой наибольшая доля слабавредных мутаций. А также свидетельствует о поддержке оптимальной энергичной структуры отрицательным отбором.

Распределение результатов измерения эффекта SNP на структуру по классам нкРНК представлено на рис. 2. Использованы только полиморфизмы с низкими аллельными частотами. По классам нкРНК значимость эффекта между петлями и стеблями выше по медиане для полиморфизмов в стеблях нкРНК различных классов.

ЗАКЛЮЧЕНИЕ

Нами были проанализированы основные эволюционные особенности нкРНК пяти основных классов: микроРНК, мяРНК, мякРНК, длинные некодирующие РНК, регуляторные элементы. Мы проанализировали действие положительного и отрицательного отборов на структуру нкРНК с помощью теста Макдональда–Крейтмана и исследовали эффект индивидуальных замен нуклеотидов в нкРНК. Мы показали, что нкРНК различных классов с точки зрения отбора на минимизацию энергии структуры отличаются друг от друга. В классах регуляторных элементов, мяРНК и микроРНК минимизация энергии структуры поддерживается действием отбора, в отличие от мякРНК и длинных некодирующих РНК. Анализ полиморфизмов в пнРНК не показал структурированности этих молекул.

Работа выполнена при финансовой поддержке Российского научного фонда (грант № 14-14-00088).

СПИСОК ЛИТЕРАТУРЫ

1. K. V. Morris and J. S. Mattick, *Nat. Rev. Genet.* **15**, 423 (2014). doi:10.1038/nrg3722.
2. J. C. Nacher and N. Araki, *BioSystems* **101** (1), 10 (2010). doi:10.1016/j.biosystems.2010.02.005.
3. I. M. Meyer, *Brief Bioinform.* **8** (6), 396 (2007). doi:10.1093/bib/bbm011.
4. A. Manuscript, *NIH Public Access Changes* **29** (6), 997 (2012). doi:10.1016/j.biotechadv.2011.08.021.Secreted.
5. J. A. Henderson and J. S. Lucas, *Nature* **232** (5313), 655 (1971).
6. N. K. Mimouni, R. B. Lyngsø, S. Griffiths-Jones, and J. Hein, *Mol. Biol. Evol.* **26** (1), 209 (2009). doi:10.1093/molbev/msn240.
7. A. G. Clark, M. B. Eisen, D. R. Smith, et al., *Nature* **450** (7167), 203 (2007). doi:10.1038/nature06341.
8. R. A. Soldatov and A. A. Mironov, *Biophysics (Oxf.)* **58** (2), 142 (2013). doi:10.1134/S0006350913020206.
9. A. Girard, R. Sachidanandam, G. J. Hannon, and M. A. Carmell, *Nature* **442** (7099), 199 (2006). doi:10.1038/nature04917.
10. S. Griffiths-Jones, A. Bateman, M. Marshall, et al., *Nucl. Acids Res.* **31** (1), 439 (2003). doi:10.1093/nar/gkg006.
11. E. P. Nawrocki and S. R. Eddy, *Bioinformatics* **29** (22), 2933 (2013). doi:10.1093/bioinformatics/btt509.
12. S. Sai Lakshmi and S. Agrawal, *Nucl. Acids Res.* **36** (Suppl. 1), 173 (2008). doi:10.1093/nar/gkm696.

13. J. H. McDonald and M. Kreitman, *Nature* **351** (6328), 652 (1991). doi:10.1038/351652a0.
14. J. C. Fay, G. J. Wyckoff, and C. I. Wu, *Genetics* **158** (3), 1227 (2001). <http://www.ncbi.nlm.nih.gov/pubmed/11454770>.
15. G. Dos Santos, A. J. Schroeder, J. L. Goodman, et al., *Nucl. Acids Res.* **43** (D1), D690 (2014). doi:10.1093/nar/gku1099.
16. R. J. Kinsella, A. Kähäri, S. Haider, et al., *Database* **2011**, 1 (2011). doi:10.1093/database/bar030.
17. C. H. Langley, K. Stevens, C. Cardeno, et al., *Genetics* **192** (2), 533 (2012). doi:10.1534/genetics.112.142018.
18. T. F. C. Mackay, S. Richards, E. A. Stone, et al., *Nature* **482** (7384), 173 (2012). doi:10.1038/nature10811.
19. J. E. Pool, R. B. Corbett-Detig, R. P. Sugino, et al., *PLoS Genet.* **8** (12), e1003080 (2012). doi:10.1371/journal.pgen.1003080.
20. Online E. *Drosophila* melanogaster* (2015).
21. M. Blanchette, W. J. Kent, C. Riemer, et al., *Genome Res.* **14** (4), 708 (2004). doi:10.1101/gr.1933104.
22. K. R. Rosenbloom, J. Armstrong, G. P. Barber, et al., *Nucl. Acids Res.* **43** (D1), D670 (2014). doi:10.1093/nar/gku1177.
23. R. Lorenz, S. H. Bernhart, C. Höner zu Siederdisen, et al., *Algorithms Mol. Biol.* **6** (1), 26 (2011). doi:10.1186/1748-7188-6-26.
24. R. Sabarinathan, H. Tafer, S. E. Seemann, et al., *Hum. Mutat.* **34** (2009), 546 (2013). doi:10.1002/humu.22273.

Evolution of non-Coding RNAs in *Drosophila melanogaster* Genome

E.V. Juravleva* and A.A. Mironov* **

*Department of Bioengineering and Bioinformatics, Lomonosov Moscow State University,
Leninskie Gory 1/73, Moscow, 119234 Russia

**Kharkevich Institute of Information Transmission Problems, Russian Academy of Sciences,
Bolshoj Karetnyj per. 19/1, Moscow, 127051 Russia

Non-coding RNAs are important regulatory molecules that function due to their specific RNA secondary structure. The secondary RNA structure depends on RNA sequence. It can be represented as a combination of loops, stems and pseudoknot regions. By the structure formation, each separate local part of a RNA structure as well as the whole molecule tend to achieve the minimum of free energy. New mutations may change the free energy of the current RNA structure and thus change the optimal one and change the function of the molecule. The chance to change the structure for a mutation depends on where it is located in the secondary RNA structure. In this research we studied natural selection acting on loop and stem regions of ncRNA of five classes. We tried to associate their selection characteristics to the changes of free energy value. We considered *Drosophila melanogaster* genome.

Key words: non-coding RNA, secondary RNA structure, evolution of non-coding RNA