

Роль некодирующих РНК в эпигенетической трансформации генных локусов

Беспалова А.В.*, Куликова Д.А., Фуников С.Ю.

Институт молекулярной биологии им. В.А. Энгельгардта РАН, Москва, Россия

* lina.bespalowa@yandex.ru

Ключевые слова: парамутация; piРНК; эпигенетика; дрозофила; генные piРНК

Мотивация и цель: Феномен парамутации описывает взаимодействие между двумя аллелями, при котором один аллель инициирует наследуемые эпигенетические изменения в другом аллеле, не затрагивая последовательность ДНК [1]. Эпигенетическое преобразование вследствие парамутации заключается в изменении паттерна метилирования гистонов и/или ДНК парамутабельного аллеля и оказывает влияние на экспрессию генов. Механизм эпигенетического преобразования аллелей при парамутации остается малоизученным, однако все ранее описанные случаи объединяет то, что главными эффекторными молекулами, необходимыми для инициации и поддержания эпигенетических модификаций, являются короткие некодирующие РНК [1].

В основе парамутации у дрозофил лежит биохимический путь образования piРНК (короткие некодирующие РНК, ассоциированные с белком Piwi) и механизм ко-транскрипционного сайленсинга с помощью комплекса piРНК/Piwi [2]. Образование piРНК происходит исключительно в репродуктивных органах животных и направлено на подавление активности мобильных элементов. Однако в редких случаях в ходе эпигенетического сайленсинга мобильных элементов белковый комплекс биогенеза piРНК может ошибочно нацеливаться на белок-кодирующие гены и подавлять их экспрессию [3, 4]. У *D. virilis* piРНК-опосредованное подавление аллеля *cdi* является примером подобной “непреднамеренной” регуляции [5]. Кроме того, материнское наследование piРНК к транскриптам *cdi* приводит к эпигенетической трансформации гомологичного локуса вследствие парамутации [6].

В нашем исследовании мы изучили репертуар генных piРНК у линий *D. virilis* из различных географических и лабораторных популяций, исследовали их многообразие и пути возникновения, а также изучили, может ли материнское наследование piРНК комплементарных mРНК белок-кодирующих генов приводить к эпигенетической трансформации генных локусов в результате парамутации.

Методы и алгоритмы: Для анализа генных piРНК мы использовали 11 линий *D. virilis*, представляющих природные и лабораторные популяции. Библиотеки коротких некодирующих РНК приготовлены с помощью набора Truseq small RNA kit (Illumina, США) из тотальной РНК, выделенной из 30 пар яичников самок каждой из линий. Удаление фракции 2S рибосомальной РНК перед созданием библиотеки проведено с помощью выделения 18–30 нт фракции тотальной РНК в полиакриламидном геле. Библиотеки секвенированы на приборах Nextseq-2000. Анализ экспрессии piРНК выполнен с использованием собственного биоинформационно-вычислительного конвейера, написанного на языке R.

Исследование парамутагенных свойств рiРНК-продуцирующих локусов выполнено на модели скрещиваний между самками линии 140 и самцами линии 9. Сначала были отобраны самки линии 140, содержащей потенциальные парамутагенные аллели *Adar/OtopLb*^{140/140} и *RhoGEF3*^{140/140}, и самцы линии 9, аллели которой являются парамутабельными по отношению к аллелям линии 140. Затем были поставлены скрещивания самок линии 140 с самцами линии 9 (♀140 × ♂9) и получены гибриды первого поколения (F1). Далее из гибридов F1 были отобраны девственные самки и проведены возвратные скрещивания с самцами исходной родительской линии 9 для получения гомозиготного потомства по локусам *Adar/OtopLb* и *RhoGEF3* линии 9. Генотипирование выведенных линий, гомозиготных по аллелям *Adar*^{9/9} и *RhoGEF3*^{9/9}, выполнено с помощью аллель-специфичных праймеров. Только потомство от самок с генотипом *Adar*^{9/9} и *RhoGEF3*^{9/9} было использовано для получения библиотек коротких некодирующих РНК.

Результаты: Среди 11 линий *D. virilis* из разных географических и лабораторных популяций мы обнаружили более 100 белок-кодирующих генов, на мРНК которых картируется рiРНК с уровнем экспрессии не менее 50 rpm (reads per million, количество прочтений на миллион выровненных прочтений) хотя бы в одной из исследованных линий. Мы обнаружили, что линии *D. virilis* обладают выраженным полиморфизмом экспрессии генных рiРНК. Из совокупности обнаруженных генов только для 25 последовательностей образование рiРНК наблюдается более чем в 5 линиях, а высокоэкспрессирующиеся генные рiРНК, с покрытием более 500 rpm, обнаружены только в одной из линий – 160.

Далее мы отобрали гены с наибольшим количеством картированных рiРНК; были выбраны три гена – *Adar*, *OtopLb* и *RhoGEF3*, на последовательности которых в 140 линии уникально картируется более 500 rpm рiРНК, а в 9 линии рiРНК отсутствуют. Мы обнаружили, что материнское наследования рiРНК к *Adar*, *OtopLb* и *RhoGEF3* определяет экспрессию рiРНК из данных локусов в следующем поколении и наблюдается только у гибридов ♀140 × ♂9, когда рiРНК наследуются по материнской линии, как и в случае гена *cdi* [6].

Взаимосвязь между образованием генных рiРНК и экспрессией мРНК генов не всегда очевидна. Между уровнем экспрессии рiРНК к последовательности гена *Adar* и уровнем его экспрессии в 140 и 9 линии, а также гибридов от прямого и реципрокного скрещивания наблюдается отрицательная корреляция. Однако для гена *OtopLb*, который находится в одном локусе с геном *Adar*, мы, наоборот, обнаружили положительную корреляцию между экспрессией мРНК и рiРНК в линии 140. Для гена *RhoGEF3* не удалось установить достоверную корреляцию между экспрессией рiРНК и уровнем экспрессии гена. Гены *Adar* и *OtopLb* находятся в одном локусе в области прицентромерного гетерохроматина, а ген *RhoGEF3* на другой хромосоме также в гетерохроматиновой области. Проведя сравнительный анализ строения генных локусов, мы обнаружили, что в локус гена *RhoGEF3* в линии 140, в отличие от 9, произошла вставка ретротранспозона *Nausicaa*. Локусы генов *Adar* и *OtopLb* в линиях 140 и 9, за исключением мелких перестроек, гомологичны последовательности в референсном геноме.

Чтобы проверить, может ли материнское наследование рiРНК, комплементарных транскриптам *Adar*, *OtopLb* и *RhoGEF3*, вызывать трансформацию гомологичных геномных локусов в рiРНК кластеры в результате парамутации, мы провели скрещивания между самками линии 140 и самцами линии 9. В результате эксперимента мы вывели 16 линий с генотипом *Adar/OtopLb*^{9/9} и *RhoGEF3*^{9/9}.

Однако проведенный анализ данных линий показал отсутствие экспрессии piRNA из локусов *Adar/OtopLb* и *RhoGEF3* во всех изученных нами линиях.

Выводы: В данной работе мы провели исследование генных piRNA в 11 линиях *D. virilis* из различных географических и лабораторных популяций. Генные piRNA составляют примерно 5 % от общего пула piRNA в яичниках и обладают выраженным межлинейным полиморфизмом экспрессии. Мы установили, что многообразие генных piRNA у *D. virilis* определяется различными механизмами образования генных piRNA. Используя модель скрещиваний линии *D. virilis* 9 и 140, демонстрирующих асимметричную экспрессию piRNA из локусов генов *Adar*, *Otoplb* и *RhoGEF3*, мы установили, что экспрессии и материнского наследования piRNA к транскриптам генов *Adar*, *OtopLb* и *RhoGEF3* недостаточно для формирования piRNA-продуцирующих кластеров в гомологичных локусах. Мы также обнаружили инсерцию мобильного элемента *Nausicaa* в локусе *RhoGEF3* в линии 140. Учитывая, что инсерции транспозонов отсутствуют в этой же области локуса *RhoGEF3* в линии 9, мы предполагаем, что именно инсерция мобильного элемента является детерминантой образования piRNA кластера в локусе *RhoGEF3* в линии 140.

Финансирование: Работа выполнена при финансовой поддержке гранта Российского научного фонда (проект № 22-74-10050).

The role of non-coding RNAs in the epigenetic conversion of gene loci

Bespalova A.V.*, Kulikova D.A., Funikov S.Y.

Engelhardt Institute of Molecular Biology, RAS, Moscow, Russia

* lina.bespalowa@yandex.ru

Key words: paramutation; Piwi-interacting RNAs; epigenetics; drosophila; genic piRNAs

Motivation and Aim: The phenomenon of paramutation describes the interaction between two alleles, in which one allele initiates inherited epigenetic changes in the other allele without affecting the DNA sequence [1]. Epigenetic conversion due to paramutation involves changes in methylation pattern of histones and/or DNA of a paramutable allele, and affects gene expression. The mechanism of epigenetic conversion of alleles during paramutation remains poorly understood, however, all previously described cases are united by the fact that the main effector molecules necessary for the initiation and maintenance of epigenetic modifications are small non-coding RNAs [1].

Paramutation in drosophila is based on the biochemical pathway of piRNAs (short non-coding RNAs associated with the Piwi protein) and the mechanism of co-transcriptional silencing by piRNA/Piwi complex [2]. The biogenesis of piRNAs occurs exclusively in the reproductive organs of animals and is aimed at suppressing the activity of transposable elements. However, in rare cases, during epigenetic silencing of transposons, the protein complex of piRNA biogenesis may mistakenly target protein-coding genes and suppress their expression [3, 4]. In *D. virilis*, rRNA-mediated suppression of the *cdi* allele is an example of such “unintended” regulation [5]. In addition, maternal inheritance of piRNA to *cdi* transcripts leads to epigenetic transformation of the homologous locus due to paramutation [6].

Herein, we analyzed the repertoire of genic piRNAs in *D. virilis* strains from various geographical and laboratory populations, investigated their diversity and pathways of origin, and also studied whether maternal inheritance of piRNAs targeted mRNA of protein-coding genes can lead to epigenetic conversion of gene loci as a result of paramutation.

Methods and Algorithms: To analyze genic piRNAs, we used 11 *D. virilis* strains representing natural and laboratory populations. Libraries of small non-coding RNAs were prepared using the Truseq small RNA kit (Illumina, USA) from total RNA isolated from 30 pairs of ovaries of females of each strain. The removal of the 2S ribosomal RNA fraction before the cloning of the library was carried out by isolating the 18–30 nt fraction of total RNA in a polyacrylamide gel. The libraries were sequenced on Nextseq-2000 platform. The analysis of piRNA expression was performed using our own pipeline written in the R language.

The study of the paramutagenic properties of piRNA-producing loci was performed on a model of crosses involving females of strain 140 and males of strain 9. First, females of strain 140 containing potential paramutagenic alleles *Adar/OtopLb*^{140/140} and *RhoGEF3*^{140/140} and males of strain 9, whose alleles are paramutable with respect to alleles of strain 140, were selected. Then crosses of females of strain 140 with males of strain 9 (♀ 140 × ♂ 9) were performed and hybrids of the first generation (F1) were obtained. Further, virgin females were selected from F1 hybrids and recurrent crosses with males of the original parental strain 9 were performed to obtain homozygous offspring at the *Adar/OtopLb* and *RhoGEF3* strain 9 loci. Genotyping of the derived homozygous strains for the *Adar*^{9/9} and *RhoGEF3*^{9/9} alleles was performed using allele-specific primers. Only offspring from females with the *Adar*^{9/9} and *RhoGEF3*^{9/9} genotypes were used to obtain libraries of small non-coding RNAs.

Results: Among 11 *D. virilis* strains from different geographical and laboratory populations, we found more than 100 protein-coding genes, on the mRNA of which piRNA is mapped with an expression level of at least 50 rpm (reads per million, the number of reads per million aligned reads) in at least one of the studied lines. We found that *D. virilis* strain have a pronounced polymorphism of genic piRNA expression. Of the total of the detected genes, only for 25 sequences, the formation of piRNAs is observed in more than 5 lines, and highly expressed genic piRNAs (more than 500 rpm) were found in only one of the strains – 160.

Next, we selected the genes with the largest number of mapped piRNAs. Three genes were selected – *Adar*, *OtopLb* and *RhoGEF3*, on the sequence of which more than 500 rpm piRNAs are uniquely mapped in strain 140, and there are no piRNAs in strain 9. We found that maternal inheritance of piRNA to *Adar*, *OtopLb* and *RhoGEF3* determines the expression of piRNA from these loci in the next generation and is observed only in hybrids 140×9 when piRNAs are inherited maternally, as in the case of the *cdi* gene [6].

The relationship between the formation of genic piRNAs and the expression of mRNA is not always obvious. There is a negative correlation between the level of expression of piRNA to the sequence of the *Adar* gene and the level of its mRNA in strains 140 and 9, as well as hybrids from direct and reciprocal crossing. However, for the *OtopLb* gene, which is located at the same locus with the *Adar* gene, on the contrary, we found a positive correlation between the expression of mRNA and piRNA in strain 140. For the *RhoGEF3* gene, it was not possible to establish a reliable correlation between the expression of piRNA and the level of gene expression. The *Adar* and *OtopLb* genes are

located at the same locus in the region of the centromeric heterochromatin, and the *RhoGEF3* gene on the other chromosome is also in the heterochromatin region. After conducting a comparative analysis of the structure of the gene loci, we found that the retrotransposon *Nausicaa* was inserted into the locus of the *RhoGEF3* gene in strain 140, in contrast to strain 9. The loci of the *Adar* and *OtopLb* genes in 140 and 9 strains, with the exception of minor rearrangements, are homologous to the sequence in the reference genome.

To test whether maternal inheritance of piRNAs complementary to the *Adar*, *OtopLb* and *RhoGEF3* transcripts can cause the conversion of homologous genomic loci into piRNA clusters as a result of paramutation, we performed crosses between females of strain 140 and males of strain 9. As a result of the experiment, we obtained 16 strain with the genotype *Adar/OtopLb*^{9/9} and *RhoGEF3*^{9/9}. However, the analysis of these strains showed the absence of expression of piRNA from the *Adar/OtopLb* and *RhoGEF3* loci in all the obtained so far.

Conclusion: In this work, we performed a study of genic piRNAs in 11 *D. virilis* strains from various geographical and laboratory populations. Genic piRNAs make up approximately 5 % of the total pool of piRNAs in the ovaries and have a pronounced intrastrain polymorphism of expression. We have established that the diversity of genic piRNAs in *D. virilis* is determined by various mechanisms of their formation. Using the crosses involving strains 9 and 140, demonstrating asymmetric expression of piRNA from the loci of the *Adar*, *OtopLb* and *RhoGEF3* genes, we found that the expression and maternal inheritance of piRNA is not sufficient for the formation of piRNA clusters in homologous loci. We also found the insertion of the transposable element *Nausicaa* at the *RhoGEF3* locus in strain 140. Considering that there are no transposon insertions in the same region of the *RhoGEF3* locus in strain 9, we assume that transposon insertion is determining the formation of a piRNA cluster at the *RhoGEF3* locus in strain 140.

Funding: The study is supported by Russian Science Foundation (grant No. 22-74-10050).

Список литературы/References

1. Hollick J.B. Paramutation and related phenomena in diverse species. *Nat Rev Genet.* 2017(1):5-23. doi 10.1038/nrg.2016.115
2. Czech B., Munafò M., Ciabrelli F., Eastwood E.L., Fabry M.H., Kneuss E., Hannon G.J. piRNA-Guided Genome Defense: From Biogenesis to Silencing. *Annu Rev Genet.* 2018;52:131-157. doi 10.1146/annurev-genet-120417-031441
3. Blumenstiel J.P., Erwin A.A., Hemmer L.W. What Drives Positive Selection in the *Drosophila* piRNA Machinery? The Genomic Autoimmunity Hypothesis. *Yale J Biol Med.* 2016;89(4):499-512
4. Rozhkov N.V., Aravin A.A., Zelentsova E.S., Schostak N.G., Sachidanandam R., McCombie W.R., Hannon G.J., Evgen'ev M.B. Small RNA-based silencing strategies for transposons in the process of invading *Drosophila* species. *RNA.* 2010(8):1634-1645. doi 10.1261/rna.2217810
5. Erwin A.A., Galdos M.A., Wickersheim M.L., Harrison C.C., Marr K.D., Colicchio J.M., Blumenstiel J.P. piRNAs are associated with diverse transgenerational effects on gene and transposon expression in a hybrid dysgenic syndrome of *D. virilis*. *PLoS Genet.* 2015;11(8):e1005332. doi 10.1371/journal.pgen.1005332
6. Dorador A.P., Dalikova M., Cerbin S., Stillman C.M., Zych M.G., Hawley R.S., Miller D.E., Ray D.A., Funikov S.Y., Evgen'ev M.B., Blumenstiel J.P. Paramutation-like epigenetic conversion by pima at the telomere of *Drosophila virilis*. *Biology (Basel).* 2022;11(10):1480. doi 10.3390/biology11101480