

УДК 577.214

ИССЛЕДОВАНИЕ КОНСЕРВАТИВНОСТИ ПОСЛЕДОВАТЕЛЬНОСТЕЙ, ОПРЕДЕЛЯЮЩИХ ТРАНС-СПЛАЙСИНГ В ЛОКУСЕ *MOD(MDG4)* У ВИДОВ ДРОЗОФИЛЫ И ШЕЛКОПРЯДА

© 2025 г. О. Бегинязова, Ю. В. Солдатова, академик РАН П. Г. Георгиев, М. В. Тихонов*

Поступило 15.09.2024 г.

После доработки 05.10.2024 г.

Принято к публикации 05.10.2024 г.

Сплайсинг, являясь ключевым этапом созревания мРНК, играет важную роль в регуляции экспрессии генов эукариот. Образование химерных мРНК из двух транскриптов при сплайсинге является запрещенным процессом, однако у насекомых для нескольких локусов транс-сплайсинг является основным способом увеличения белкового разнообразия. Целью работы является исследование эволюционной консервативности последовательностей, ответственных за транс-сплайсинг в локусе *mod(mdg4)*, у видов из семейства *Drosophilidae* отряда двукрылые и шелкопряда (*Bombyx mori*), который относится к отряду чешуекрылых. С помощью модельных трансгенных линий показано, что последовательности удаленных видов дрозофилы сохраняют способность поддерживать транс-сплайсинг у *D. melanogaster*. Напротив, аналогичные последовательности шелкопряда не поддерживают транс-сплайсинг. Таким образом, РНК-мотивы и связывающие их гипотетические белковые факторы, определяющие транс-сплайсинг, остаются консервативными среди группы *Drosophilidae*, но функционально разошлись у двукрылых и чешуекрылых.

Ключевые слова: альтернативный сплайсинг, транс-сплайсинг, *Drosophila melanogaster*, *Bombyx mori*

DOI: 10.31857/S2686738925010212, **EDN:** tbcvmh

У *Drosophila melanogaster* было показано, что в локусах *mod(mdg4)* и *lola* альтернативные 3'-экзоны присоединяются к общей для всех изоформ пре-мРНК с помощью нового механизма, названного транс-сплайсингом. Этот процесс объединяет 5'- и 3'-концы экзонов, синтезируемых с разных промоторов [5–7]. Транс-сплайсинг был обнаружен у многих высших эукариот, включая млекопитающих [8]. В основном он инициируется спариванием длинных комплементарных последовательностей в интронах, что приводит к сплайсингу между двумя отдельными транскриптами [8, 9]. Однако в локусе *mod(mdg4)* такие комплементарные последовательности не были выявлены [10]. Более того, было показано, что у всех мРНК 3'-экзоны присоединяются исключительно через транс-сплайсинг [11].

Это указывает на существование специального белкового комплекса, который специфически связывается с последовательностями в локусе

mod(mdg4) и инициирует процесс транс-сплайсинга. Основную роль в этом процессе играют последовательности четвертого интрона, который отделяет последний конститутивный экзон от альтернативных экзонов. Каждый из альтернативных экзонов определяет одну из 31 изоформ белка [10, 11] (рис. 1А). В четвертом интроне были обнаружены необходимые для транс-сплайсинга последовательности, удаление которых приводило к нарушению этого процесса [10, 11]. В нижележащей области от них происходит терминация транскрипции без формирования поли(А)-хвоста. Четкой границы, где заканчивается транскрипция, не наблюдается [11].

Ранее было установлено, что структура локуса *mod(mdg4)* у двукрылых насекомых является консервативной [12, 13]. В то же время у насекомых других отрядов структура этого локуса значительно изменилась. Тем не менее транс-сплайсинг в гене *mod(mdg4)* был также обнаружен у шелкопряда (*Bombyx*) из отряда чешуекрылых, который разделен с отрядом двукрылых примерно 290 миллионов лет назад [14–16]. Принцип организации локуса у этих насекомых схож с двукрылыми (рис. 1А). Небольшой участок ДНК кодирует единственный донорный транскрипт, в то время как область,

¹Федеральное государственное бюджетное учреждение науки Институт биологии гена Российской академии наук (ИБГ РАН), Москва, Россия

*e-mail: me@mtih.me

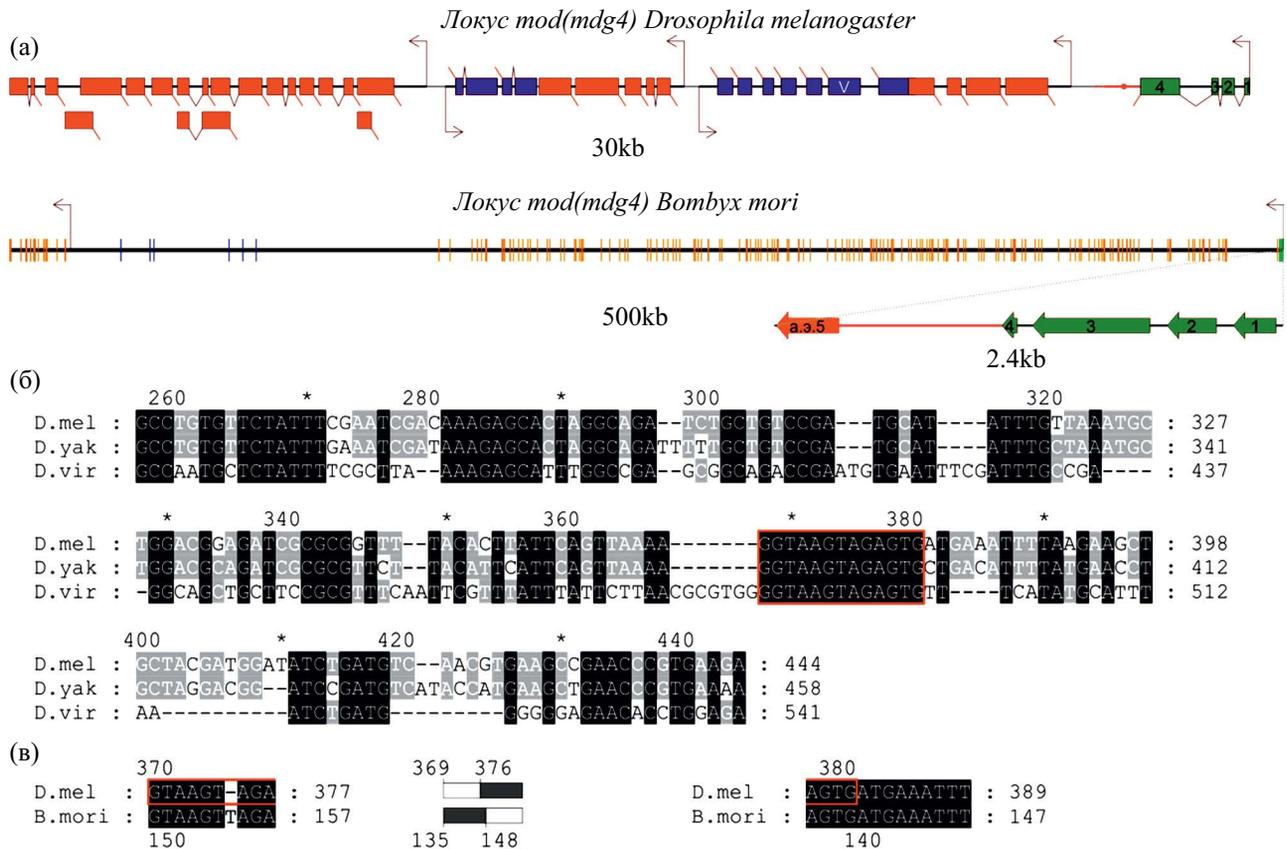


Рис. 1 (А) Структура локуса *mod(mdg4)* у *Drosophila melanogaster* и *Bombyx mori*. Аннотированные промоторы обозначены стрелками. Общие экзоны, присутствующие во всех изоформах, показаны зеленым цветом. Четвертый интрон *Drosophila melanogaster*, содержащий мотивы, необходимые для транс-сплайсинга, и его гомолог у *Bombyx mori* выделены красным. Альтернативные экзоны, уникальные для конкретной изоформы, отображены оранжевым цветом (на одной и той же цепи ДНК) и синим цветом (на противоположной цепи). **(Б)** Выравнивание последовательности, необходимой для транс-сплайсинга у *Drosophila melanogaster*, с гомологичными участками интронов *Drosophila yakuba* и *Drosophila virilis*. Консервативный мотив 13 п.о., с которым, вероятно, связывается U1-snRNP, выделен красным прямоугольником. В верхней строке приведены координаты относительно начала интрона-4 *D. melanogaster*, каждая десятая позиция отмечена числом или звездочкой (*). **(В)** Сравнение схожих последовательностей в интронах *Drosophila melanogaster* и *Bombyx mori*. Схематично показана перестановка порядка следования этих последовательностей в интронах.

отвечающая за акцепторные транскрипты, расширена до 500 kb и содержит более 200 альтернативных экзонов. Как и у двукрылых, были обнаружены изоформы, расположенные на противоположной цепи ДНК, однако их количество может быть недооценено из-за ограничений существующих алгоритмов аннотации транскриптов.

Целью настоящего исследования стало выяснение, насколько консервативными являются последовательности, которые определяют транс-сплайсинг в локусе *mod(mdg4)* *Drosophila melanogaster*.

Ранее нами была выявлена область в интроне 4, делеция которой нарушала транс-сплайсинг. Мы провели сравнительный анализ последовательности *D. melanogaster*, важной для транс-сплайсинга, и интронов далекого вида *D. virilis*

и более близкого вида *D. yakuba*. В анализ также был включен интрон тутового шелкопряда. Последовательности *Drosophilidae* оказались схожими (рис. 1Б). В центре этих последовательностей расположен консервативный мотив длиной 13 п.о., с которым может связываться U1 snRNP [10] (рис. 1Б, отмечен красным). Хотя делеция этого участка не приводит к нарушению транс-сплайсинга [11], вероятно, он является одним из сайтов, на которые происходит посадка РНК-связывающих белков. Консервативные мотивы также были найдены в окружающих последовательностях. Напротив, интроны *D. melanogaster* и *B. mori* значительно отличаются. Можно отметить только две сходные последовательности, одна из которых пересекается с консервативным мотивом 13 п.о. (рис. 1В).

Ранее в нашей лаборатории для экспериментального тестирования ДНК последовательностей, участвующих в транс-сплайсинге, была разработана модельная система (рис. 2А) [11]. Одна конструкция (названная “Донор”) содержит регуляторную область локуса *mod(mdg4)* и все общие экзоны и интроны. Последний общий экзон 4 был модифицирован вставкой последовательности сайта внутренней инициации трансляции (IRES) из гена *reaper* [17], который используется для инициации трансляции из внутренних частей мРНК по кэп-независимому механизму. Вторая конструкция,

названная “Акцептор”, содержит ген, кодирующий 3'-часть изоформы V. Этот ген был разделен последовательностью гена люциферазы светлячка (Fluc). Перед 3'-сайтом сплайсинга расположен ATG кодон, инициация трансляции с которого приводила к формированию коротких белковых продуктов, препятствуя трансляции люциферазы с несплайсированных акцепторных транскриптов. Только в результате транс-сплайсинга люцифераза могла транслироваться – IRES из донорного транскрипта оказывался перед кодирующей областью Fluc акцептора.

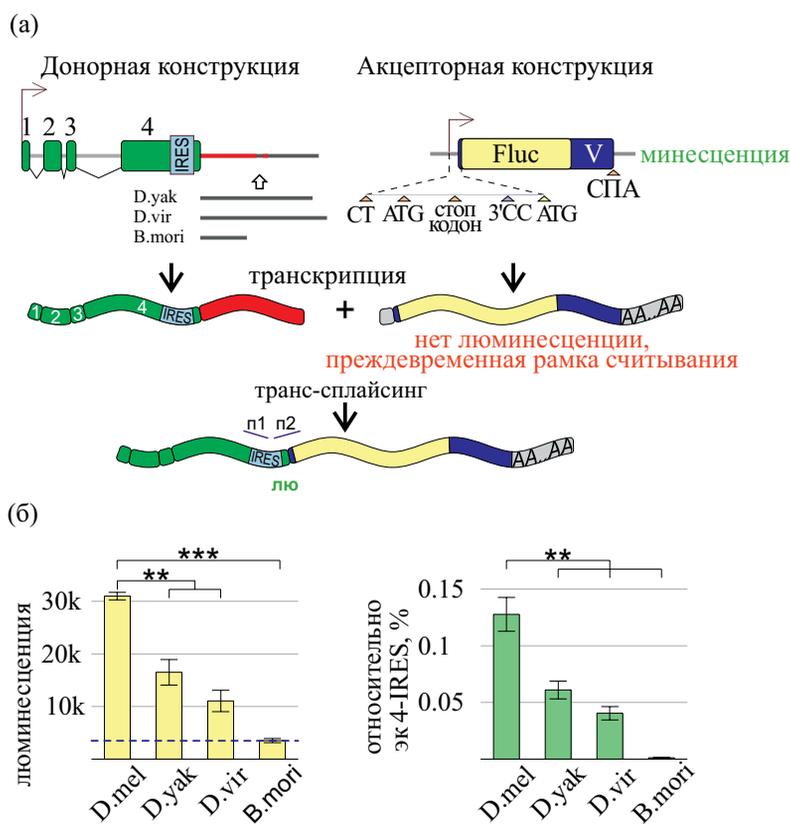


Рис. 2. А) Схема модельной системы, предназначенной для исследования функциональной роли последовательностей четвертого интрона гена *mod(mdg4)* из различных видов насекомых в поддержании транс-сплайсинга. Конструкция “Донор” (слева) включает промотор *mod(mdg4)* (обозначен стрелкой), кодирующую область, состоящую из четырех общих экзонов (зеленые прямоугольники), IRES и интрон 4 (темно-серый цвет; красным выделена транскрибируемая часть, штриховой линией обозначена зона терминации транскрипции). Конструкция “Акцептор” (справа) содержит промотор (стрелка, СТ – старт транскрипции) и альтернативный экзон V (синий), разделенный репортерным геном Fluc (желтый прямоугольник). В результате транс-сплайсинга формируется мРНК, состоящая из 5'-части донорного транскрипта и 3'-части акцепторного. Места отжига праймеров (п1 и п2) для анализа эффективности транс-сплайсинга обозначены синими линиями. Б) Оценка эффективности транс-сплайсинга у гетерозигот донор/акцептор была проведена с использованием двух методов: анализа активности люциферазы (слева) и ОТ-кПЦР (справа). На графике (слева) показаны абсолютные уровни люминесценции, измеренные в лизатах с равным содержанием тотального белка. Прерывистая синяя линия обозначает уровень фоновой автолюминесценции. При использовании ОТ-кПЦР эффективность транс-сплайсинга оценивалась по соотношению количеств ампликонов из области соединения четвертого экзона и люциферазы (п2) к количеству ампликонов из области пересечения четвертого экзона и IRES (п1). В обоих методах каждая комбинация трансгетерозигот была проанализирована не менее чем в трех повторностях. Планки погрешности отображают стандартные отклонения (n = 3). Звездочки обозначают уровни значимости: **P < 0.01, ***P < 0.001.

Обе конструкции интегрировались по attP сайту в район 22A с помощью фС31 интегразы [18]. При скрещивании двух линий получали потомство, одновременно несущее две конструкции в одном локусе. В качестве положительного контроля использовали полный четвертый интрон *Drosophila melanogaster*. По такой же схеме, как и контрольный, в донорные конструкции были интегрированы гомологичные интроны других видов: *Drosophila yakuba*, *Drosophila virilis* и *Bombyx mori*.

Лизат для определения активности люциферазы и тотальную РНК выделяли из 2–3 дневных самцов, которые содержали один из вариантов донорной конструкции и акцепторную конструкцию. Лизат использовали для измерения активности люциферазы: к ним добавляли субстрат, разведенный в буфере для измерения (Biotium), и оценивали активность фермента с помощью люминометра (Promega). Тотальную РНК использовали для синтеза кДНК. Количественная ПЦР позволяла оценивать два участка. Первый, расположенный на стыке экзона 4 и IRES, определяет общее количество донорного транскрипта (рис. 2А). Второй участок амплифицируется только с продуктов транс-сплайсинга: прямой праймер находится на донорном транскрипте, а обратный – на люциферазе акцептора. Соотношение амплифицированных продуктов второго и первого участков показывает долю донорного транскрипта, который претерпел транс-сплайсинг с акцептором.

Результаты анализа, представленные на рис. 2Б, показывают, что интроны из более отдаленного вида *D. virilis* и из более близкого вида *D. yakuba* способны поддерживать транс-сплайсинг у *D. melanogaster*. Однако эффективность этого процесса у интронов обоих видов снижена. По данным ОТ-кПЦР, для интрона *D. virilis* снижение эффективности составило в 2.5 раза, а для интрона из *D. yakuba* – в 1,9 раза. Интрон шелкопряда полностью исключает возможность транс-сплайсинга. Результаты обоих методов согласуются между собой.

Ранее нами было показано [11], что успешное прохождение транс-сплайсинга требует наличия пока неопределенных мотивов (кроме уже известного 13 п.о. [10]) в небольшом участке (73 п.о.) четвертого интрона. Эти мотивы, вероятно, распознаются еще неидентифицированными РНК-связывающими белками, которые участвуют в транс-сплайсинге. В нативном интроне такие сайты представлены в избыточном количестве. У представителей семейства *Drosophilidae*, участвовавших в исследовании, регуляторные области консервативны и демонстрируют высокий уровень сходства. Белковые факторы *Drosophila melanogaster*, участвующие в транс-сплайсинге, вероятно, могут взаимодействовать с гомологичными

последовательностями других видов *Drosophilidae*. Однако интрон шелкопряда, несмотря на наличие наиболее консервативного мотива, не способен поддерживать транс-сплайсинг у *Drosophila melanogaster*. Это, вероятно, связано с высокой степенью расхождения РНК-связывающих белков, которые участвуют в транс-сплайсинге, и их мотивов у дрозофилы и шелкопряда.

Можно заключить, что РНК-мотивы и распознающие их белковые факторы остаются консервативными в пределах таксона *Drosophilidae*. Однако, несмотря на наличие схожих последовательностей в регуляторной области интрона 4 у *D. melanogaster* и *B. mori*, эволюционная дивергенция между этими видами привела к тому, что белковые факторы плодовой мухи не способны в достаточной степени взаимодействовать с интроном шелкопряда.

БЛАГОДАРНОСТИ

В работе использовали оборудование Центра высокоточного редактирования и генетических технологий для биомедицины ИБГ РАН.

ИСТОЧНИКИ ФИНАНСИРОВАНИЯ

Исследование выполнено за счет гранта РФ (проект № 23-44-00038).

КОНФЛИКТ ИНТЕРЕСОВ

Авторы заявляют об отсутствии конфликта интересов.

СОБЛЮДЕНИЕ ЭТИЧЕСКИХ НОРМ И СТАНДАРТОВ

План работы одобрен биоэтической комиссией ФГБУН Института биологии гена РАН (заключение № 17 от 15 декабря 2022).

СПИСОК ЛИТЕРАТУРЫ

1. Gehring, N.H. and J.Y. Roignant. Anything but Ordinary – Emerging Splicing Mechanisms in Eukaryotic Gene Regulation // Trends Genet. 2021. 37(4). P. 355-372.
2. McManus, C.J., et al. Global analysis of trans-splicing in *Drosophila* // Proc Natl Acad Sci U S A. 2010. 107(29). P. 12975-9.
3. Baralle, F.E. and J. Giudice. Alternative splicing as a regulator of development and tissue identity // Nat Rev Mol Cell Biol. 2017. 18(7). P. 437-451.

4. *Shenasa, H. and D.L. Bentley.* Pre-mRNA splicing and its cotranscriptional connections // *Trends Genet.* 2023. 39(9). P. 672–685.
5. *Büchner, K., et al.* Genetic and molecular complexity of the position effect variegation modifier *mod(mdg4)* in *Drosophila* // *Genetics.* 2000. 155(1). P. 141–57.
6. *Dorn, R., G. Reuter, and A. Loewendorf,* Transgene analysis proves mRNA trans-splicing at the complex *mod(mdg4)* locus in *Drosophila* // *Proc Natl Acad Sci U S A.* 2001. 98(17). P. 9724–9.
7. *Horiuchi, T., E. Giniger, and T. Aigaki.* Alternative trans-splicing of constant and variable exons of a *Drosophila* axon guidance gene, *lola* // *Genes Dev.* 2003. 17(20). P. 2496–501.
8. *Lei, Q., et al.* Evolutionary Insights into RNA trans-splicing in Vertebrates // *Genome Biol Evol.* 2016. 8(3). P. 562–77.
9. *Kong, Y., et al.* The evolutionary landscape of intergenic trans-splicing events in insects // *Nat Commun.* 2015. 6. P. 8734.
10. *Gao, J.L., et al.* A conserved intronic U1 snRNP-binding sequence promotes trans-splicing in *Drosophila* // *Genes Dev.* 2015. 29(7). P. 760–71.
11. *Tikhonov, M., et al.* Conserved sequences in the *Drosophila mod(mdg4)* intron promote poly(A)-independent transcription termination and trans-splicing // *Nucleic Acids Res.* 2018.
12. *Krauss, V. and R. Dorn.* Evolution of the trans-splicing *Drosophila* locus *mod(mdg4)* in several species of *Diptera* and *Lepidoptera* // *Gene.* 2004. 331. P. 165–76.
13. *Labrador, M. and V.G. Corces.* Extensive exon reshuffling over evolutionary time coupled to trans-splicing in *Drosophila* // *Genome Res.* 2003. 13(10). P. 2220–8.
14. *Shao, W., et al.* Alternative splicing and trans-splicing events revealed by analysis of the *Bombyx mori* transcriptome // *RNA.* 2012. 18(7). P. 1395–407.
15. *Tong, K.J., et al.* INSECT PHYLOGENOMICS. Comment on “Phylogenomics resolves the timing and pattern of insect evolution” // *Science.* 2015. 349(6247). P. 487.
16. *Misof, B., et al.* Phylogenomics resolves the timing and pattern of insect evolution // *Science.* 2014. 346(6210). P. 763–7.
17. *Hernández, G., et al.* Internal ribosome entry site drives cap-independent translation of reaper and heat shock protein 70 mRNAs in *Drosophila* embryos // *RNA.* 2004. 10(11). P. 1783–97.
18. *Bischof, J., et al.* An optimized transgenesis system for *Drosophila* using germ-line-specific *phiC31* integrases // *Proc Natl Acad Sci USA.* 2007. 104(9). P. 3312–7.

AN INVESTIGATION OF THE CONSERVATISM OF SEQUENCES DEFINING TRANS-SPLICING IN THE *MOD(MDG4)* LOCUS ACROSS *DROSOPHILA* AND SILKWORM SPECIES

O. Beginyazova, Iu. V. Soldatova, Academician of the RAS P. G. Georgiev, M. V. Tikhonov*

Institute of Gene Biology Russian Academy of Sciences, Moscow, Russian Federation

*e-mail: me@mtih.me

Splicing, a key step in mRNA maturation, plays a crucial role in the regulation of eukaryotic gene expression. The formation of chimeric mRNAs from two transcripts during splicing is typically a prohibited process; however, in insects, trans-splicing is a primary mechanism for increasing protein diversity for several loci. The aim of this study is to investigate the evolutionary conservativeness of sequences responsible for trans-splicing in the *mod(mdg4)* locus among species from the *Drosophilidae* family (order *Diptera*) and the silkworm (*Bombyx mori*), which belongs to the order *Lepidoptera*. Using model transgenic lines, it was shown that sequences from distant *Drosophila* species retain the ability to support trans-splicing in *D. melanogaster*. In contrast, analogous sequences in the silkworm do not support trans-splicing. Thus, the RNA motifs and their binding hypothetical protein factors, defining trans-splicing, remain conserved among the *Drosophilidae* group, but have functionally diverged between *Diptera* and *Lepidoptera*.

Keywords: alternative splicing, trans-splicing, *Drosophila melanogaster*, *Bombyx mori*