

УДК 575.22:595.773.4

СРАВНЕНИЕ ФУНКЦИОНАЛЬНОЙ РОЛИ КОРОТКОЙ И ДЛИННОЙ ИЗОФОРМ БЕЛКА TRF2 У *Drosophila melanogaster*

И. С. Осадчий*, академик РАН П. Г. Георгиев, О. Г. Максименко

Поступило 05.03.2019 г.

Белок TRF2 (TBP-related factor 2) способен замещать TBP, формируя альтернативные ТАТА-независимые комплексы на промоторах генов, в том числе гистона H1 и репрессирующих транспозоны *piwi*-кластеров. Ген *trf2* дрозофилы кодирует две изоформы белка — короткую и длинную. В данной работе мы показали, что длинная изоформа белка TRF2 обладает большей функциональной активностью по сравнению с короткой изоформой при их сниженной экспрессии под собственными промоторами гена. При экспрессии только длинной изоформы белка самцы и самки имели нормальную выживаемость. Экспрессия только короткой изоформы приводила к значительному снижению количества самцов и нарушению сегрегации хромосом в мейозе — фенотипу, описанному при инактивации белка TRF2 и дерепрессии транспозонов.

Ключевые слова: TBP-related factor 2 (TRF2), дерепрессия транспозонов, нарушение сегрегации хромосом.

DOI: <https://doi.org/10.31857/S0869-56524864514-518>

Инициация транскрипции у высших эукариот наиболее хорошо изучена на примере образования комплекса TFIIID при связывании белка TBP (TATA-box-binding protein) с последовательностью ТАТА-бокса в коровой части промотора. Однако только около 30% РНК-полимераза II-зависимых генов *D. melanogaster* содержат ТАТА-бокс [1]. Достаточно рано в процессе эволюции произошла дупликация гена, кодирующего TBP, в результате которой возник паралог TRF2 (TBP-related factor 2), потерявший ДНК-связывающую активность, но способный замещать TBP, формируя альтернативные ТАТА-независимые комплексы [2]. У дрозофилы TRF2 является ключевым белком в активности промоторов генов, регулирующих развитие, генов “домашнего хозяйства”, в частности связанных с репликацией, клеточным циклом, белками рибосом, гистонами H1 [3–5]; в инициации транскрипции в гетерохроматине *piRNA*-кластеров, необходимых для репрессии транспозонов [6].

Ген *trf2* кодирует две изоформы белка — короткую и длинную. Длинная изоформа p175 (1715 а.о.) отличается от короткой p75 (632 а.о.) наличием дополнительного протяжённого N-концевого домена, слабо консервативного среди *Drosophilidae* и состо-

ящего из повторов [7]. Предполагается, что протяжённый N-концевой домен может дополнительно стабилизировать TRF2 на промоторах и/или усиливать активацию транскрипции.

Целью настоящего исследования является выяснение функциональности короткой и длинной изоформ белка TRF2. Для этого была создана трансгенная линия *D. melanogaster*, содержащая *attP*-сайт в интроне гена *trf2* и позволяющая с помощью фаговой рекомбиназы ϕ C31 заменить кодирующую часть гена на последовательность, кодирующую длинную или короткую изоформу. В трансгенной линии, экспрессирующей только длинную изоформу белка, самцы и самки имели нормальную выживаемость и встречались в равных пропорциях. При экспрессии только короткой изоформы происходило описанное ранее в [8] нарушение расхождения хромосом в мейозе, что приводило к снижению количества развивающихся эмбрионов и наблюдаемому низкому количеству самцов относительно самок (около 5%).

Ген *trf2* *D. melanogaster*, размером около 25 т.п.н., находится на X-хромосоме в районе 7E7 и имеет длинную регуляторную область, которая включает два промотора (рис. 1а). Изначально планировалось получение делеции гена *trf2* с одновременной интеграцией репортерного гена *mCherry* и *attP*-сайта (рис. 1б) с помощью метода CRISPR/Cas9 и гомо-

Институт биологии гена Российской Академии наук,
Москва

*E-mail: untie@mail.ru

логичной рекомбинации [9]. Для этого с помощью программы Fly CRISPR Optimal Target Finder (Университет Висконсин) были подобраны гидовые РНК, направляющие Cas9 в область интрона с 5'-стороны от кодирующих экзонов (tctctgtcactactcttagc) и в область 3'-НТО (tgcttttcgcttcggtgtcc), внося при этом разрывы на расстоянии 6,7 т.п.н. ДНК-последовательности, кодирующие гидовые РНК, были клонированы в экспрессирующий вектор на основе pCFD4-U6:1_U6:3 tandemgRNAs (Addgene # 49411) с помощью рестриктазы *BbsI* [10]. ДНК-последовательности плечей для гомологичной рекомбинации (рис. 1б) были клонированы в плазмидный вектор с двух сторон от интегрируемой кассеты, которая состояла из сайта *attP* для фаговой интегразы фС31 и находящегося между двумя *loxP*-сайтами репортерного гена *mCherry* под контролем актинового промотора [10]. Смесь плазмид (10 : 1 — плаزمид для гомологичной рекомбинации : гидовые РНК,

конечная концентрация 500 нг/мкл) была инъецирована в эмбрионы линии # 54591 (Bloomington *Drosophila* Stock Center). Эта линия содержит на X-хромосоме ген, экспрессирующий белок Cas9 под контролем промотора *nanos* [11].

Полученные после инъекции мухи F0 были скрещены индивидуально с мухами, несущими балансир по X-хромосоме, FM7с. Идентификация потомков F1, в которых потенциально могла произойти замена гена *trf2* на *attP*-сайт и репортер *mCherry*, осуществлялась по флуоресценции *mCherry*. В результате анализа потомков 140 мух F0 было получено только одно событие. При этом наблюдались пониженная выживаемость эмбрионов и личинок и слабая фертильность мух F0, в которых одновременно детектировалось свечение *mCherry*. Вероятно, низкая эффективность получения событий замены в гене *trf2* является следствием высокой эффективности работы Cas9, в результате которой происходят од-

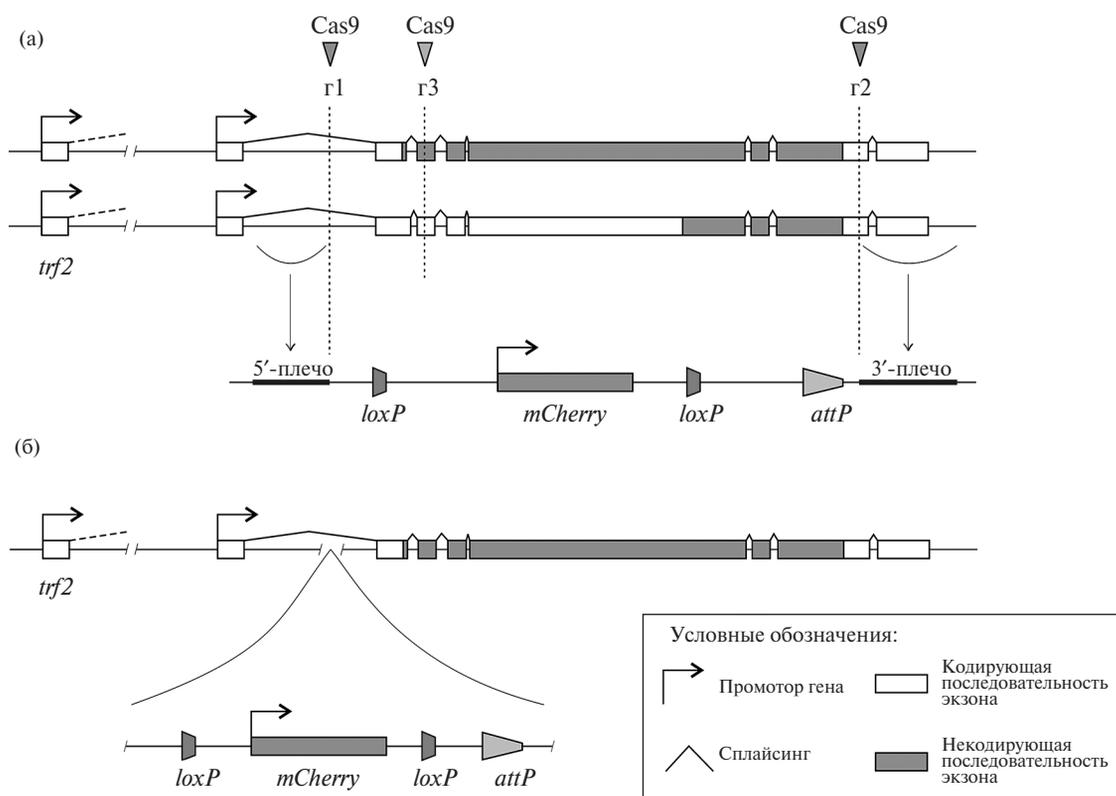


Рис. 1. а — схема гена *trf2* и матрицы для гомологичной рекомбинации при замене геномной последовательности методом CRISPR/Cas9. Показано два альтернативных промотора гена *trf2* и возможность формирования двух альтернативных мРНК; трансляция короткой изоформы белка TRF2^{p75} осуществляется с внутреннего сайта инициации трансляции мРНК, не имеющей первый ATG-кодон (с которого запускается трансляция длинной изоформы белка). 5'-плечо — 500 п.н. гомологичной ДНК-последовательности; 3'-плечо — 700 п.н. гомологичной ДНК-последовательности; Cas9 — направляемая гидовыми РНК эндонуклеаза; g1, g2, g3 — гидовые РНК; *loxP* — сайты, узнаваемые Cre-рекомбиназой бактериофага P1; *attP* — последовательность для сайт-специфической интеграции конструкций, узнаваемая интегразой бактериофага фС31; *mCherry* — репортерный ген под контролем актинового промотора, кодирующий красный флуоресцентный белок. б — схема модификации гена и результат интеграции конструкции. Гомологичная рекомбинация произошла только по 5'-разрыву. Репортерный ген *mCherry* и *attP*-сайт встроились в интрон *trf2*.

новременно события во всех аллелях гена *trf2*. Анализ полученной линии с помощью ПЦР показал наличие встройки репортера и *attP*-сайта в интрон гена, по 5'-разрыву, но с отсутствием делеции *trf2*.

Для повторной инъекции были изменены 3'-гиговая РНК (ассааgtagctagagactta) и соответствующее вносимому разрыву 3'плечо в матрице для рекомбинации так, чтобы разрывы вносились на расстоянии 1,1 т.п.н. друг от друга, приводя к инактивации гена *trf2*. Полученные в F0 эмбрионы также имели низкую выживаемость, при этом у развивающихся личинок наблюдалась флуоресценция репортера mCherry в близких к месту инъекции тканях. Личинки с наиболее интенсивным уровнем флуоресценции погибали на стадии куколки или в первые дни после выхода из куколки. Выжившие мухи были индивидуально скрещены с мухами линии FM7c. Среди потомков F1 не было обнаружено событий встройки. Эти наблюдения подтверждают высокую эффективность внесения разрывов и гомологичной рекомбинации одновременно в двух аллелях гена, что приводит к гибели мух F0.

Для дальнейшей работы были использованы мухи, имеющие встройку репортера и *attP*-сайта в интрон *trf2*, расположенный перед кодирующими белок экзонами. Полученные мухи $y^1 cas9 w^1 trf2[mCherry, attP]/FM7c$ были рекомбинированы с мухами $y^+ w^1$ для удаления из линии гена *Cas9*. В результате была получена гомозиготная линия $y + w^1$

trf2[mcherry, attP], которая не имела морфологических изменений или снижения жизнеспособности по сравнению с исходной линией. Одновременно было обнаружено присутствие одной дополнительной копии конструкции, которая встроилась на близком генетическом расстоянии от *trf2* (что не позволяет удалить её при помощи мейотической рекомбинации). При индукции рекомбинации между *loxP*-сайтами возникал летальный фенотип, связанный с рекомбинацией между *loxP*-сайтами из разных копий конструкции и делецией области генома, находящейся между ними. Наличие нескольких копий репортерного гена *mCherry* было подтверждено с помощью количественной ПЦР. Следствием этого стали, с одной стороны, невозможность вырезать репортерный ген с использованием Cre/*loxP*-рекомбинации, а с другой стороны, высокая вероятность двух встроек при интеграции тестируемых конструкций по *attP*-сайту.

Для выяснения функциональности короткой и длинной изоформ белка TFR2 были созданы и клонированы в плазмидный вектор для интеграции по *attP*-сайту три конструкции (рис. 2). Две из них содержали акцепторный сайт сплайсинга интрона *trf2*, кодирующие части короткой (TRF2^{p75}) или длинной (TRF2^{p175}) изоформ *trf2* и сигнал полиаденилирования из вируса SV40. В качестве контроля использована конструкция, содержащая только акцепторный сайт сплайсинга и сигнал полиаденилирования из

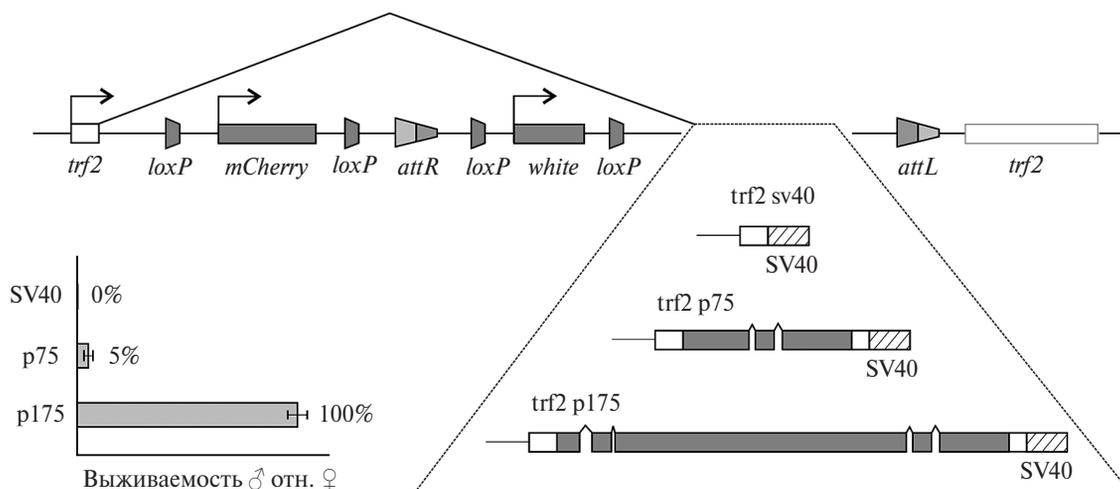


Рис. 2. Схема модифицированного гена *trf2* и введённых конструкций с его мутационными производными. Каждая конструкция содержит репортерный ген *white*, акцепторный сайт сплайсинга из интрона *trf2* и сигнал полиаденилирования вируса SV40 для предотвращения транскрипции эндогенного *trf2*. Производная *trf2 sv40* приводит к летальности самцов и гомозиготных самок; *trf2 p75* — не влияет на жизнеспособность гомозиготных самок, но приводит к резкому снижению количества самцов (около 5% от количества самок); *trf2 p175* самцы и гомозиготные самки имеют нормальную жизнеспособность (гистограмма слева — выживаемость самцов относительно самок ($N = 200$); для линии *trf2_sv40* представлено соотношение самцов *trf2^{sv40}/Y* к самкам *trf2^{sv40}/FM7c*). *attL/attR* — последовательности, возникающие в результате интеграции конструкции по *attP*-сайту, *white* — репортерный ген, отвечающий за красную пигментацию глаз.

вируса SV40. Вектор для интеграции содержал *attB*-сайт и окруженный *loxP*-сайтами репортерный ген *white*, отвечающий за красную пигментацию глаз.

Смесь плазмид в соотношении 10:1 (тестируемая конструкция: вспомогательная плазида, экспрессирующая фаговую интегразу ϕ C31 [12]; конечная концентрация 500 нг/мкл) была инъецирована в эмбрионы полученной линии y^+w^1 *trf2[mCherry, attP]*. В результате был получен ряд независимых *white*⁺-линий с одним или несколькими событиями интеграции. Отбор линий с целевой встройкой конструкции осуществлялся ПЦП с праймерами к участку генома с 5'-стороны от встройки (aaatctacatccggcagcaggcat) и к сигналу полиаденилирования SV40 (ctctacaatgtggtatggct) с помощью полимеразы NEB LongAmp Taq. Зависимость степени красной окраски глаз от дозы гена *white* позволила отобрать линии с единичной встройкой. Линии *white*⁺, не имеющие встройки по целевому *attP*-сайту, имели нормальный фенотип и жизнеспособность, что предполагает нейтральность внесённых Cas9 мутаций и нецелевых интеграций конструкций.

Как и ожидалось, интеграция сигнала полиаденилирования SV40 (линия *trf2_sv40*) приводила к эмбриональной летальности гемизиготных самцов *trf2^{sv40}/Y* и гомозиготных самок *trf2^{sv40}/trf2^{sv40}*. Однако гетерозиготные самки *trf2^{sv40}/FM7c* имели нормальный фенотип и жизнеспособность, что идентично фенотипу ранее описанных нуль-мутаций по гену *trf2* [13] и подтверждает, что в линии *trf2_sv40* происходит полная инактивация гена *trf2*.

В трансгенной линии *trf2(p175)_sv40* происходило полное восстановление активности гена *trf2*: гомозиготные самки *trf2^{p175}/trf2^{p175}* и гемизиготные самцы *trf2^{p175}/Y* имели нормальный фенотип и жизнеспособность. Таким образом, можно предположить, что наличие репортерных генов *mCherry* и *white* в интроне перед кодирующей частью тестируемой изоформы *trf2* слабо влияет на экспрессию самой конструкции; при этом длинная изоформа TRF2^{p175} может полностью компенсировать инактивацию гена *trf2*.

В трансгенной линии *trf2(p75)_sv40* не происходило полное восстановление функции TRF2. Описанный ранее феномен нерасхождения хромосом при слабых мутациях в регуляторной области гена *trf2* [14] или снижения его экспрессии с помощью РНК-интерференции [8] приводил к повышенной эмбриональной летальности, появлению самок с генотипом XX/Y, и, как следствие, наблюдаемому низкому количеству самцов относительно самок

(около 5%), а в случае расхождения хромосом приводил к спорадическому появлению в потомстве мух, имеющих фенотип линий, с которыми линия *trf2(p75)_sv40* скрещивалась ранее (*FM7c* и y^+w^1). При этом наблюдаемые гомозиготные самки *trf2^{p75}/trf2^{p75}* и гемизиготные самцы *trf2^{p75}/Y* имели нормальный фенотип.

Ранее в [7] было показано, что при оверэкспрессии TRF2 под промотором гена *Su(Hw)* обе изоформы способны восстанавливать нуль-мутации по *trf2*, а в [4] продемонстрировали, что обе изоформы TRF2 могут привлекаться на одни и те же модельные промоторы. Проблема функциональной необходимости длинной изоформы оставалась невыясненной. В полученных нами линиях *trf2(p75)_sv40* и *trf2(p175)_sv40* экспрессия изоформ частично снижена из-за присутствия двух репортерных генов в интроне *trf2*. Однако сниженная экспрессия трансгенов под собственными промоторами и в нативном геномном окружении показала различие двух изоформ TRF2, p75 и p175, к способности восстановления нуль-мутации. Наиболее чувствительной системой к варианту изоформы и количеству белка TRF2 оказалась система регуляции экспрессии транспозонов, роль длинных некодирующих РНК которых в организации перичентромерных и теломерных регионов и вовлечённость в митотическую организацию хромосом становится всё более установленной [14, 8].

Способность сниженного количества TRF2^{p175} так же компенсировать нуль-мутацию, как и оверэкспрессия короткой изоформы TRF2^{p75}, свидетельствует о том, что, вероятно, неконсервативная N-концевая часть участвует во взаимодействии с ДНК-связывающими белками, образуя более стабильный преинициаторный комплекс на промоторах, и/или обладает дополнительной активирующей способностью.

Источники финансирования. Работа выполнена при поддержке Российского фонда фундаментальных исследований (17–00–00285).

СПИСОК ЛИТЕРАТУРЫ

1. Ohler U., Liao G., Niemann H., Rubin G.M. // Genome Biol. 2002. V. 3. № 12. P. RESEARCH0087.
2. Vo ngoc L., Wang Y.-L., Kassavetis G.A., Kadonaga J.T. // Genes Dev. 2017. V. 31. № 13. P. 1289–1301.
3. Isogai Y., Keles S., Prestel M., Hochheimer A., Tjian R. // Genes Dev. 2007. V. 21. № 22. P. 2936–2949.

4. Kedmi A., Zehavi Y., Glick Y., Orenstein Y., Ideses D., Wachtel C., Doniger T., Waldman Ben-Asher H., Muster N., Thompson J., et al. // *Genes Dev.* 2014. V. 28. № 19. P. 2163–2174.
5. Wang Y.-L., Duttke S.H.C., Chen K., Johnston J., Kassavetis G.A., Zeitlinger J., Kadonaga J.T. // *Genes Dev.* 2014. V. 28. № 14. P. 1550–1555.
6. Andersen P.R., Tirian L., Vunjak M., Brennecke J. // *Nature.* 2017. V. 549. № 7670. P. 54–59.
7. Kopytova D. V., Krasnov A.N., Kopantceva M.R., Nabirochkina E.N., Nikolenko J. V., Maksimenko O., Kurshakova M.M., Lebedeva L.A., Yerokhin M.M., Simonova O.B., Korochkin L.I., Tora L., Georgiev P.G., Georgieva S.G. // *Mol. Cell. Biol.* 2006. V. 26. № 20. P. 7492–7505.
8. Vorontsova Y.E., Cherezov R.O., Simonova O.B. // *Genetika.* 2013. V. 49. № 6. P. 669–80.
9. Gratz S.J., Ukken F.P., Rubinstein C.D., Thiede G., Donohue L.K., Cummings A.M., O'Connor-Giles K.M. // *Genetics.* 2014. V. 196. № 4. P. 961–71.
10. Zolotarev N., Maksimenko O., Kyrchanova O., Sokolinskaya E., Osadchiy I., Girardot C., Bonchuk A., Ciglar L., Furlong E.E.M., Georgiev P. // *Nucleic Acids Res.* 2017. V. 45. № 21. P. 12285–12300.
11. Port F., Chen H.-M., Lee T., Bullock S.L. // *Proc. Nat. Acad. Sci. U. S. A.* 2014. V. 111. № 29. P. E2967–76.
12. Bischof J., Maeda R.K., Hediger M., Karch F., Basler K. // *Proc. Nat. Acad. Sci.* 2007. V. 104. № 9. P. 3312–3317.
13. Peter A., Schöttler P., Werner M., Beinert N., Dowe G., Burkert P., Mourkioti F., Dentzer L., He Y., Deak P., et al. // *EMBO Rept.* 2002. V. 3. № 1. P. 34–8.
14. Morgunova V., Akulenko N., Radion E., Olovnikov I., Abramov Y., Olenina L. V., Shpiz S., Kopytova D. V., Georgieva S.G., Kalmykova A. // *Nucleic Acids Res.* 2015. V. 43. № 18. P. 8762–8773.

FUNCTIONAL COMPARISON OF SHORT AND LONG ISOFORMS OF THE TRF2 PROTEIN IN *Drosophila melanogaster*

I. S. Osadchiy, Academician of the RAS P. G. Georgiev, O. G. Maksimenko

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

Received March 5, 2019

The TRF2 protein (TBP-related factor 2) can substitute for TBP in the TFIID complex forming alternative transcription initiation complexes for TATA-less promoters, including promoters of histone H1 and required for transposon repression piRNA-clusters. The *Drosophila trf2* gene codes for two isoforms: a “short” and a “long” one, in which the same short TRF2 sequence is preceded by a long N-terminal domain. Here, we demonstrated that the long TRF2 isoform has greater functional activity than the short isoform by expressing each of them in reduced rate under the endogenous promoters. Expression of the long isoform alone doesn't affect neither the flies' viability nor the sex ratio. Expression of the short isoform alone leads to the phenotype described for the *trf2* gene insufficiency and derepression of transposable elements, that is, decreased viability, violation of homologous chromosome pairing and segregation, and apparent female-biased sex ratio.

Keywords: TBP-related factor 2 (TRF2), derepression of transposons, abnormal chromosome segregation.