

TRF4 И TRF5 - ГЕНЫ, КОДИРУЮЩИЕ НОВЫЕ TBP-ПОДОБНЫЕ БЕЛКИ D.MELANOGASTER

Куршакова Мария Михайловна

канд. биол. наук, научный сотрудник, Институт биологии гена РАН, Россия, Москва

Копытова Дарья Владимировна

канд. биол. наук, старший научный сотрудник, Институт биологии гена РАН, Россия, Москва

trf4 and trf5 - genes coding for new TBP-related proteins of D.melanogaster

Maria Kurshakova

Cand. Biol. in Science, Researcher, Institute of Gene Biology Russian Academy of Sciences, Russia, Moscow

Darya Kopytova

Cand. Biol. Sci., Senior Researcher, Institute of Gene Biology, Russian Academy of Sciences, Moscow, Russia

Аннотация. В геноме *D. melanogaster* были найдены гены, кодирующие новые TBP-подобные белки, названные *trf4* и *trf5*. Гомологи *trf4* и *trf5* найдены в геномах других видов *Drosophila*, но не обнаружены в геномах других организмов. *trf4* и *trf5* были клонированы, определена их экзон-интронная структура. Предсказанные белки TRF4 и TRF5 имеют вырожденную гомологию с TBP и TBP-подобным доменом TRF2. В аминокислотной последовательности TRF4 и TRF5 мутированы последовательность, связывающаяся с промотором, и последовательность сигнала ядерной локализации. Можно предположить, что TRF4 и TRF5 выполняют функции отличные от функций известных TBP-подобных белков TBP и TRF2.

Ключевые слова: TBP; TRF2; TRF1; TBP-подобные белки.

Keywords: TBP; TRF2; TRF1; TBP-related factors.

Введение. Белок TBP, входящий в состав фактора TFIID вместе с ассоциированными с ним белками (TAFs), распознает ТАТА-последовательность промоторов и таким образом регулирует место старта транскрипции. С-концевой домен TBP высоко консервативен у эукариот и содержит два симметричных повтора, которые сворачиваются в седловидную структуру, взаимодействующую с ДНК. Ранее считалось, что TBP является универсальным и единственным транскрипционным фактором, необходимым всем трем РНК-полимеразам для распознавания промоторов и инициации транскрипции у эукариот. В настоящее время известно, что у многоклеточных существуют белки с высокой гомологией к коровому домену

ТВР (ТВР-подобные факторы), которые могут выполнять функции ТВР в процессе транскрипции определенного набора генов, не регулируемых ТВР. ТВР-подобные белки участвуют в связывании транскрипционного аппарата с регуляторными элементами промоторов и общими факторами транскрипции, направляя транскрипцию с ТАТА-независимых промоторов, ген-специфичных и тканеспецифичных промоторов. На сегодняшний день описаны следующие ТВР-подобные белки: TRF1, TRF2 (TLF), TRF3 (TBP2) [1].

TRF1, также называемый TRF, TBP-related factor, был открыт у *D.melanogaster* и специфичен для насекомых. Он имеет высокий уровень экспрессии в центральной нервной системе эмбрионов и клетках половой системы самцов. TRF1 имеет 63% идентичных аминокислотных остатка с ТВР в коровом домене, способен связываться с ТАТА боксом, ассоциирован с Brf1 и присутствует на большинстве промоторов РНК-полимеразы III, стимулирует базальную транскрипцию *in vivo* и *in vitro* [2].

TRF2, также известный как TBP-like factor, TLF, TBPL1 — более дальний паралог ТВР, который найден во всех многоклеточных. Он проявляет около 40% идентичности с коровым доменом ТВР. Это единственный представитель ТВР-семейства, у которого отсутствует способность связываться с ТАТА боксом. Он, однако, взаимодействует с TFIIA и TFIIB [3, 4]. TRF2 принимает участие в регуляции образования хроматина: частичная инактивация TRF2 в мухах приводит к ослабленной премейотической конденсации хромосом [5], а уменьшение уровня TRF2 в слюнных железах *D.melanogaster* вызывает серьезные нарушения структуры хроматина в политенных хромосомах [2]. Показано, что TRF2 участвует в регуляции транскрипции генов в ходе эмбриогенеза у мухи, лягушки и рыбы [6, 7]. У *D.melanogaster* и мышей TRF2 необходим для правильного развития половых клеток, так как у самцов мышей, с делецией *trf2* в геноме, наблюдается нарушение сперматогенеза, самцы стерильны [8], а у *D. melanogaster* при частичной инактивации TRF2 нарушается развитие мужских и женских половых клеток [5]. В целом, имеющиеся данные указывают на то, что TRF2, по-видимому, поддерживает альтернативный механизм инициации транскрипции, который может включать ремоделирование хроматина.

TRF3, также известный как TBP2, TBPL2 был найден в геномах позвоночных и является наиболее близким паралогом ТВР, имея 95% идентичности в коровом домене. Он способен связываться с ТАТА боксом, взаимодействовать с TFIIA и TFIIB и способствовать транскрипции *in vitro* [9, 10].

В работе был проведен поиск новых генов, кодирующих ТВР-подобные белки, в геноме *D. melanogaster*. Были найдены два гена, названные *trf4* и *trf5*. Обнаружено, что гомологи *trf4* и *trf5* присутствуют в геномах других видов *Drosophila*, но не найдены в геномах других организмов. Проведен анализ гомологии предсказанных белковых продуктов TRF4 и TRF5 с другими ТВР-подобными белками *D.melanogaster*: исследуемые белки имеют вырожденную гомологию с ТВР и ТВР-подобным доменом TRF2. Опубликованные экзон-интронные структуры *trf4* и *trf5* были подтверждены. Проведено сравнение экзон-интронной структуры генов *trf4*, *trf5*, *trf2* и *tbp*. Проанализирована консервативность белковых мотивов у гомологов TRF4 и TRF5 из разных видов *Drosophila*: TRF4 и TRF5 сильно изменены по сравнению с ТВР и TRF2, и могут выполнять функции, отличные от функций ранее описанных ТВР-подобных белков.

Материалы и методы.

Поиск и анализ нуклеотидных последовательностей.

Поиск генов проводили при помощи программы Blast. Ген *trf4* (CG9879) and ген *trf5* (CG15398) были найдены в базе данных GeneBank: NM_134855 and NM_134856, соответственно. Нуклеотидные последовательности гомологов *trf4* и *trf5* из других видов *Drosophila* были также найдены в базе данных GeneBank. Сравнение нуклеотидных и аминокислотных последовательностей проводили с использованием программы Clustal W 1.83 (<http://www.genebee.msu.su/clustal/>).

Клонирование.

Чтобы картировать экзон-интронную структуру генов *trf4* и *trf5* был проведен анализ методом быстрой амплификации 5' и 3' концов кДНК генов с использованием Circ ligase (Epicentre) по протоколу производителя. Использовали мРНК, выделенную из имаго *D.melanogaster* при помощи Trizol (Invitrogen). Первая цепь кДНК была синтезирована с использованием RevertAid (Thermo Fisher Scientific) обратной транскриптазы. Для амплификации ПЦР использовали праймеры к кодирующей части генов. На основе последовательностей, полученных в этих экспериментах, были синтезированы ген-специфические праймеры для амплификации полноразмерных кДНК генов. ПЦР продукты были клонированы в вектор pJet1.2 и секвенированы. Экзон-интронная структура клонов была определена путем сравнения кДНК и геномных последовательностей.

Результаты.

Поиск генов *trf4* и *trf5*.

С целью найти новые TBP-подобные белки *D.melanogaster* был проведен поиск при помощи программы Blast в базе белков RefSeq и в базе белковых продуктов генов GeneBank. Поиск проводили по аминокислотной последовательности TBP-подобного домена TBP-подобного белка TRF2. Ген, аналогичный гену белка TRF3, у дрозофилы отсутствует. Белок TRF2 имеет большую длину, чем остальные проанализированные белки, а именно 1715 а.о., но содержит вблизи С-конца область гомологии с TBP. В результате были найдены ранее не охарактеризованные TBP-подобные белки - продукты генов CG9879 и CG15398, которым мы дали названия *trf4* и *trf5* (ген, аналогичный гену белка TRF3, у дрозофилы отсутствует). Ген *trf4* по данным базы GenBank кодирует один полипептид размером 331 а.о., которому мы дали название TRF4, а ген *trf5* кодирует полипептид размером 305 а.о., названный TRF5.

TRF4 и TRF5 *D.melanogaster* имеют около 26% идентичных и 40% похожих аминокислот. При помощи программы Blast был проведен поиск в базе белков RefSeq гомологичных TRF4 и TRF5 белков у других организмов. Гомологи *trf4* и *trf5* были обнаружены у других видов *Drosophila* и не найдены у других организмов. Совпадение между аминокислотной последовательностью TBP-подобных доменов предсказанных белков TRF4 *D.melanogaster* (а.о. 117 - 331) и более близкородственных [11] *D.yakuba* (а.о. 115 - 329), *D.erecta* (а.о. 120 - 334) около 81%, а между TBP-подобными доменами белков TRF4 *D.melanogaster* (а.о. 117 - 331) и более эволюционно далекими видами, такими как *D.virilis* (а.о. 298 - 512) and *D.obscura* (а.о. 186 - 400) – около 41%. TRF5 также проявляет эволюционную консервативность, но в меньшей степени. Выравнивание аминокислотных последовательностей предсказанных гомологичных белков TRF4 и TRF5 различных видов *Drosophila* привело к построению двух отдельных ветвей филогенетического дерева: одна соответствует гомологам TRF4, а другая – гомологам TRF5 (Рисунок 1). Интересно, что гомологи TRF4 и TRF5 обнаружены в геномах различных видов *Drosophila*, но отсутствуют в геномах других организмов, включая других насекомых.

Степень гомологии TRF4 и TRF5 *D.melanogaster* с TRF2 низкая: идентичность аминокислотной последовательности между TBP-подобным доменом TRF2 и TRF4 составляет 8,8% на 331 выравниваемых а.о., а для TRF5 - 8,6% на 305 выравниваемых а.о. Сравнение аминокислотных последовательностей TRF4 и TRF5 с аминокислотной последовательностью TBP показывает 13,1% и 13,5% идентичных аминокислот соответственно. Область гомологии сравниваемых белков находится в С-концевой области.

Экзон-интронная структура *trf4* и *trf5*.

Сравнение геномной последовательности гена *trf4* *D.melanogaster* и предсказанной кДНК показало, что *trf4* содержит два интрона. Аналогичное сравнение показало, что *trf5* также содержит два интрона. Опубликованная экзон-интронная структура генов была исследована методом быстрой амплификации 5' и 3' концов кДНК генов и последующим клонированием и секвенированием. кДНК *trf4* и *trf5* также были клонированы и секвенированы.

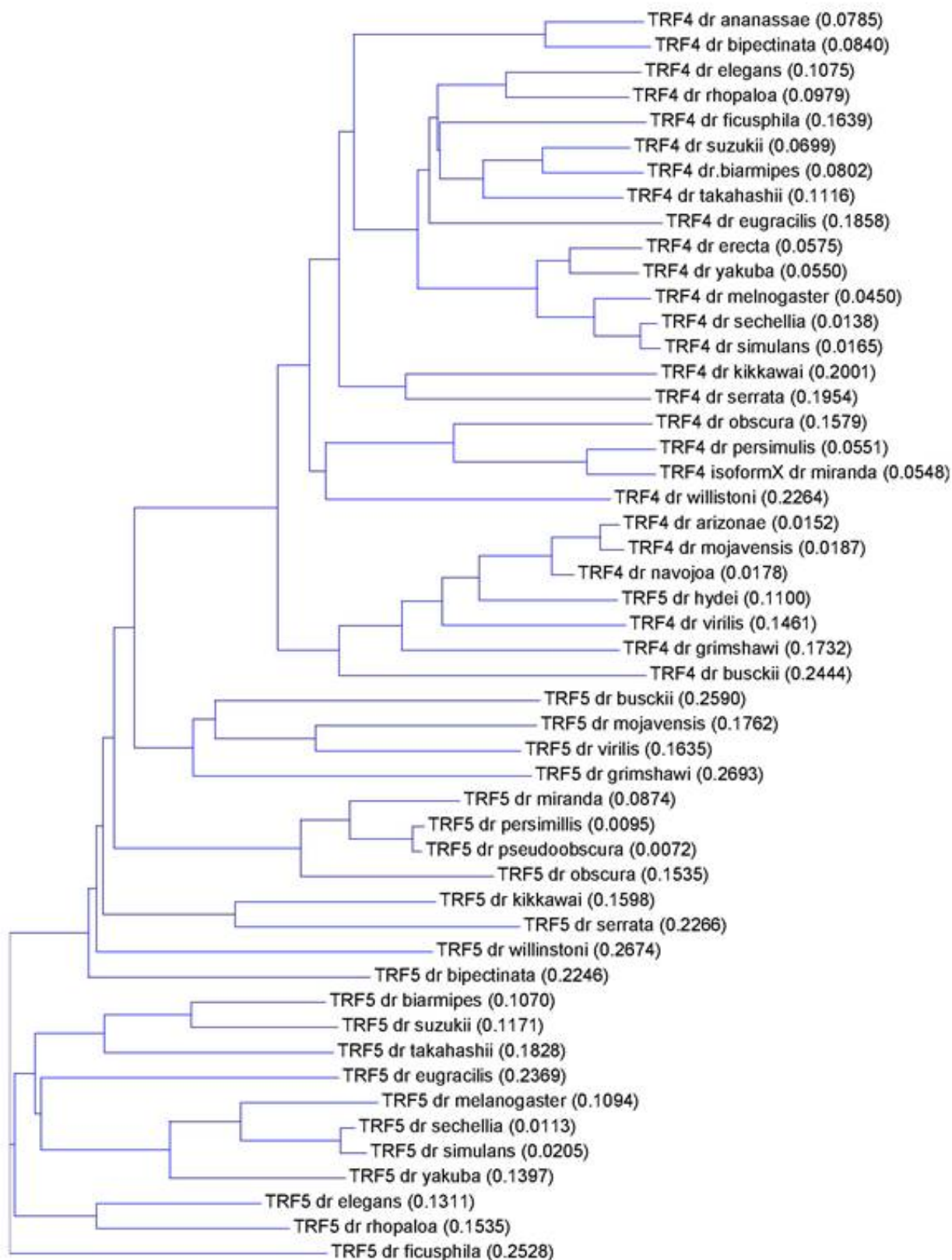


Рисунок 1. Филогенетическое дерево гомологов TRF4 и TRF5 у разных видов *Drosophila*

Результаты подтвердили предсказанную структуру генов *trf4* и *trf5*. Основываясь на них, каждый из генов кодирует один транскрипт. Для того, чтобы сравнить экзон-интронные

структуры генов *D.melanogaster*, кодирующих TBP-подобные белки, мы провели выравнивание геномных последовательностей и кДНК генов *trf4*, *trf5*, *trf2* и *tbp*. Все сравниваемые гены имеют интроны и экзон-интронная структура каждого гена уникальна, однако, в ДНК-последовательности гена *trf5* отсутствует участок, соответствующий последовательности одного из интронов *trf4*. Этот факт а также то, что уровень гомологии между предсказанными продуктами генов *trf4* и *trf5* - белками TRF4 и TRF5 - выше, чем между ними и другими TBP-подобными белками *D.melanogaster* позволяет предположить, что *trf4* и *trf5* являются родственными генами и произошли от какого-то гена-предшественника, кодирующего TBP-подобный белок. Можно предположить, что *trf4* сформировался в результате дупликации гена в ходе эволюции *Drosophila*, а *trf5* является вырожденной ретрокопией *trf4*, которая в ходе эволюции значительно изменилась, в том числе приобрела свои интроны.

Анализ белковых мотивов TRF4 и TRF5.

Анализ белковых доменов и мотивов TRF4 и TRF5 *D.melanogaster*, проведенный при помощи программ на сайте <http://elm.eu.org/> показал, что аминокислотные последовательности TRF4 и TRF5 содержат мутации в TATA-связывающей области TBP-подобного домена. Это также наблюдается в случае TRF2, у которого также изменена правильная последовательность TATA-связывающей области и который не способен связываться с ДНК. Также у TRF4 изменена последовательность сигнала ядерной локализации PILKKFKK на PILAKLKN, а у TRF5 - на PTLYRLKD. Наблюдается высокая консервативность некоторых участков аминокислотных последовательностей гомологичных TRF4 белков и гомологичных TRF5 белков у разных видов *Drosophila*, причем у TRF4 и TRF5 белков консервативными являются различные мотивы. Такая сохранность указывает на важность этих мотивов для работы белков, а также дает основания полагать, что TRF4 и TRF5 функционально диверсифицировались в ходе эволюции. Интересно, что у TBP и TRF2 эти мотивы отсутствуют. Такие изменения в аминокислотной последовательности, которые являются консервативными у разных видов *Drosophila*, могут указывать на отличие функций TRF4 и TRF5 от функций TBP, который присутствует в ядре и регулирует транскрипцию.

Обсуждение. Обнаруженные в геноме *D.melanogaster* гены *trf4* и *trf5* кодируют белки TRF4 и TRF5, которые имеют вырожденную гомологию с TBP и TBP-подобным доменом TRF2. Однако, в геномах различных видов *Drosophila* найдены гены, кодирующие белки, гомологичные TRF4 и TRF5. По-видимому, гены *trf4* и *trf5* образовались в ходе эволюции *Drosophila* и являются специфичными для этого семейства. Сравнение экзон-интронной структуры *trf4* и *trf5* показало, что они являются родственными генами, но значительно диверсифицировались в ходе эволюции. Гомологи TRF4 и гомологи TRF5 у разных видов *Drosophila* проявляют эволюционную консервативность, но эта сохранность наблюдается в различных участках аминокислотных последовательностей. Вероятно, белки TRF4 и TRF5 выполняют различные функции в клетке. Основываясь на том, что у TRF4 и TRF5 белков отсутствуют правильная последовательность связывания с TATA-элементом промоторов и сигнал ядерной локализации, а также высоко консервативны мотивы, отсутствующие у TBP и TRF2 белков, можно предположить, что TRF4 и TRF5 приобрели в ходе эволюции *Drosophila* некоторые другие функции, не связанные с работой в ядерных процессах, которые были поддержаны естественным отбором.

Благодарности.

Работа была поддержана грантом РФФИ № 16-04-00823А "Исследование нового TBP-подобного белка TRF4 *D.melanogaster*".

Список литературы:

1. Akhtar W. and Veenstra G.J. (2011) TBP-related factors: a paradigm of diversity in transcription initiation. *Cell&Bioscience*, 1: 23.

2. Isogai Y., Takada S., Tijan R., Keles S. (2007) Novel TRF1/BRF target genes revealed by genome-wide analysis of *Drosophila* PolIII transcription. *EMBO J.*, 26(1):79-89.
3. Moore P. A., Ozer J., Salunek M., Jan G., Zerby D., Campbell S., Lieberman P. M. (1999) A human TATA binding protein-related protein with altered DNA binding specificity inhibits transcription from multiple promoters and activators. *Mol Cell Biol*, 19(11): 7610-7620.
4. Teichmann M., Wang Z., Martinez E., Tjernberg A., Zhang D., Vollmer F., Chait B. T., Roeder R. G. (1999) Human TATA-binding protein-related factor-2 (hTRF2) stably associates with hTFIIA in HeLa cells. *Proc Natl Acad Sci USA*, 96(24): 13720-13725.
5. 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. (2006) Two isoforms of *Drosophila* TRF2 are involved in embryonic development, premeiotic chromatin condensation, and proper differentiation of germ cells of both sexes. *Mol Cell Biol*, 26(20): 7492-7505.
6. Jacobi U. G., Akkers R. C., Pierson E. S., Weeks D. L., Dagle J. M., Veenstra G. J. (2007) TBP paralogs accommodate metazoan- and vertebrate-specific developmental gene regulation. *EMBO J*, 26(17): 3900-3909.
7. Dantonel J. C., Quintin S., Lakatos L., Labouesse M., Tora L. (2000) TBP-like factor is required for embryonic RNA polymerase II transcription in *C.elegans*. *Mol Cell*, 6(3): 715-722.
8. Martianov I., Fimia G. M., Dierich A., Parvinen M., Sassone-Corsi P., Davidson I. (2001) Late arrest of spermiogenesis and germ cell apoptosis in mice lacking the TBP-like TLF/TRF2 gene. *Mol Cell*, 7(3): 509-515.
9. Bartfai R., Balduf C., Hilton T., Rathmann Y., Hadzhiev Y., Tora L., Orban L., Muller F. (2004) TBP2, a vertebrate-specific member of the TBP family, is required in embryonic development of zebrafish. *Curr Biol*, 14(7): 593-598.
10. Jallow Z., Jacobi U. G., Weeks D. L., Dawid I. B., Veenstra G.J. (2004) Specialized and redundant roles of TBP and vertebrate-specific TBP paralog in embryonic gene regulation in *Xenopus*. *Proc Natl Acad Sci USA*, 101(37): 13525-13530.
11. Seetharam A. S., Stuart G. W. (2013) Whole genome phylogeny for 21 *Drosophila* species using predicted 2b-RAD fragments. *PeerJ* 1: e226; DOI 10.7717/peerj.226.