



УДК 577.214.(337+622)

## ЭКЗОН-ИНТРОННАЯ СТРУКТУРА ГЕНА *fet5<sup>+</sup>* *Schizosaccharomyces pombe* И ФИЗИЧЕСКОЕ КАРТИРОВАНИЕ ОКАЙМЛЯЮЩИХ ЕГО УЧАСТКОВ ГЕНОМА\*

© 1998 г. Г. В. Шпаковский<sup>#</sup>, Е. Н. ЛебедекоИнститут биоорганической химии им. М.М. Шемякина и Ю.А. Овчинникова РАН,  
117871, Москва, ГСП-7, ул. Миклаухо-Маклая, 16/10

Поступила в редакцию 04.07.97 г. Принята к печати 28.08.97 г.

Из геномной клонотеки делящихся дрожжей *Schizosaccharomyces pombe* изолирована плаزمида рYUK3, несущая ген *fet5<sup>+</sup>*, и построена детальная физическая карта всей геномной вставки (~9,6 т. п. о.). Установлена первичная структура гена *fet5<sup>+</sup>* и окружающих его областей. Ген *fet5<sup>+</sup>* содержит единственный интрон длиной 45 п. о., расположенный в дистальной части; в 5'-некодирующей области гена находится классический ТАТА-участок (ТАТААГ), расположенный на расстоянии ~50 п. о. от предполагаемого старта транскрипции; 3'-некодирующая область содержит АТ-богатые палиндромы, участвующие, по-видимому, в терминии синтеза *fet5<sup>+</sup>*-транскрипта. Недалеко от *fet5<sup>+</sup>* обнаружен ранее не описанный ген *Sz. pombe*, кодирующий белок с некоторыми чертами сходства с одним из активаторов инициации транскрипции из ТВР-группы (ТАТА-binding protein group) SPT-факторов транскрипции. Поскольку кДНК гена *fet5<sup>+</sup>* была обнаружена как функциональный супрессор генетического дефекта ядерных РНК-полимераз I–III (Биоорганическая химия, 1997, т. 23, № 3, с. 234–237), такое соседство является первым свидетельством в пользу возможной сопряженности в геноме делящихся дрожжей генов, осуществляющих регуляцию процесса транскрипции.

**Ключевые слова:** делящиеся дрожжи, ген *fet5<sup>+</sup>*, интрон, факторы транскрипции эукариот, глутаминбогатый активаторный домен.

Транскрипция у эукариот, осуществляемая тремя ядерными ДНК-зависимыми РНК-полимеразами I–III, регулируется большим числом разнообразных транс-факторов белковой природы. Такая регуляция может быть как положительной, так и отрицательной и осуществляется, как правило, на стадии инициации процесса. Различные факторы, регулирующие транскрипцию, были обнаружены с помощью биохимических и генетических методов. Выяснилось, что детально охарактеризованные биохимически общие факторы транскрипции, такие, как ТАТА-связывающий белок (ТВР), консервативны у всех эукариот, от дрожжей до человека [2, 3]. В то же время детальные механизмы активации промоторов у одно- и многоклеточных эукариот различны и осуществляются через медиаторный комплекс холо-

фермента у дрожжей [4] или с использованием различных белковых факторов, ассоциированных с ТВР (TBP-associated factors или TAFs), у дрожзофилы и человека [5].

Многие факторы, участвующие в транскрипционном контроле, были обнаружены с помощью генетических подходов, в подавляющем большинстве случаев на примере дрожжей *Saccharomyces cerevisiae*. Среди них можно отметить многочисленные факторы, влияющие на фенотипическую экспрессию репортерных генов с таких промоторов, как рTy, рM $\beta$ Gal, рM $\alpha$ 1 и др., и кодируемые генами ADA [6], SNF/SWI [7], SPT [8, 9], SRB [4].

Недавно с помощью генетического подхода, гетерологичной межвидовой супрессии, нами был идентифицирован у делящихся дрожжей *Schizosaccharomyces pombe* новый белковый фактор транскрипции Fet5, способный подавлять условный дефект по функции общей субъединицы ABC10 $\beta$ , незаменимого компонента РНК-полимераз I–III *S. cerevisiae* [10]. В качестве начала, супрессирующего этот дефект, нами была выделена полноразмерная кДНК гена *fet5<sup>+</sup>*, кодирующая белок из 276 а. о. с вычисленной молекулярной массой 31,5 кДа и изоэлектрической точкой рI 4,22. В продолжение этого исследования мы изолировали из геномной клонотеки делящихся

\* Результаты данной работы были частично доложены на Европейском конгрессе по молекулярной биологии клетки ЕСВО 97, г. Брайтон (Англия), 22–25 марта 1997 г. (см. [1]). Сокращения: SPT – генетические супрессоры мутаций, вызываемых дрожжевым ретротранспозоном Ty (suppressors of Ty); ТВР – белок, связывающийся с ТАТА-участком (ТАТА-binding protein), в *Saccharomyces cerevisiae* кодируется геном SPT15; ORF – открытая рамка считывания.

<sup>#</sup> Автор для переписки (тел.: (095)330-65-83; e-mail: gvs@ibch.siocb.ras.ru; факс: (095)335-71-03).

дрожжей *Sz. pombe* соответствующий плазмидный клон, составили его подробную физическую карту и выяснили полную первичную структуру гена *fet5<sup>+</sup>* и некоторых близлежащих участков генома.

Просеивание геномной клонотеки *Sz. pombe* [11] провели методом последовательных разведений с применением ПЦР [12, 13]. В качестве праймеров использовали олигонуклеотиды oGVS308, (5')GGAAGTCCCGCATATTAAC (соответствует нуклеотидам 492–510 *fet5<sup>+</sup>* кДНК [10]), и oGVS309, (5')ССТТСТАТАСГТГТАТСТСС (комплементарен нуклеотидам 851–832, лежащим непосредственно за стоп-кодоном *fet5<sup>+</sup>* [10]). В результате положительный сигнал был обнаружен в первичном разведении А1 геномной клонотеки. При этом полученный в ПЦР фрагмент ДНК оказался длиннее фрагмента, образующегося при амплификации *fet5<sup>+</sup>* кДНК с теми же праймерами, что указывает на наличие интрона в данной области гена (рис. 1).

После четырех циклов последовательных субразведений был изолирован индивидуальный клон (плазмида рYUK3 = А1-4-4-3), несущий ген *fet5<sup>+</sup>*. Длина вставки геномной ДНК *Sz. pombe* (9.6 т. п. о.) и положение гена *fet5<sup>+</sup>* в ней были установлены путем физического картирования рYUK3 с помощью редкощепящих эндонуклеаз рестрикции: *Bam*HI, *Bgl*II, *Eco*RI, *Eco*RV, *Hind*III, *Kpn*I, *Not*I, *Pst*I, *Pvu*I, *Sac*I, *Sal*I, *Sma*I, *Sph*I, *Stu*I, *Xba*I и *Xho*I (рис. 2). Параллельно была построена подробная физическая карта челночного вектора рНL10 геномной клонотеки [11], способного реплицироваться в клетках *E. coli* и *Sz. pombe* (рис. 2). В качестве фенотипических маркеров этот вектор несет бактериальный ген β-лактамазы и ген делящийся дрожжей *ura4<sup>+</sup>* [14], гомологичный гену *URA3 S. cerevisiae*.

Секвенирование плазмиды рYUK3 с использованием специфических праймеров oGVS304 (5' GTTTGTGGTGTGGCCAGCAGTG), oGVS307 (5' ССАТТГАТГАТГТТАТГГА) и уже упомянутых oGVS308 и oGVS309 показало, что ген состоит из двух экзонов (691 и 186 п. о.), разделенных интроном длиной 45 п. о. В первичной структуре интрона представлены все необходимые для сплайсинга нуклеотидные последовательности, характерные для этих элементов у *Sz. pombe* [15] (рис. 3). Интрон расположен в дистальной части гена *fet5<sup>+</sup>*, при этом все участки гена, кодирующие пуриноклеотидсвязывающие домены (Р-петля, мотивы С и G; см. [10]), находятся в первом экзоне. Не исключена возможность того, что и несплайсированная копия мРНК *fet5<sup>+</sup>* может приводить к образованию белкового продукта укороченной длины (236 а. о., см. рис. 3). Вопрос, использует ли природа у одноклеточных эукариот

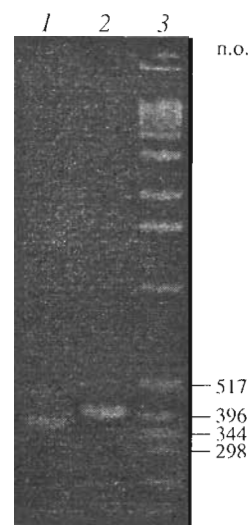


Рис. 1. Электрофорез в 1.5% агарозном геле продуктов ПЦР, полученных со специфическими праймерами oGVS308 и oGVS309 на матрице *fet5<sup>+</sup>* кДНК (плазмида рGVS510 [10]) (дорожка 1) и ДНК первичного разведения А1 геномной клонотеки (дорожка 2). Дорожка 3 – маркеры молекулярной массы (1Kb DNA Ladder, BRL, США).

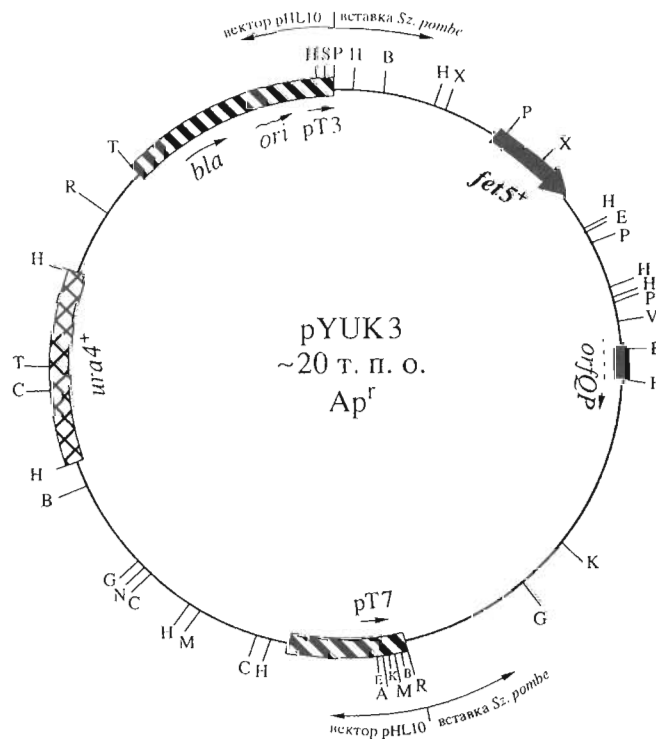


Рис. 2. Физическая карта плазмиды рYUK3, несущей фрагмент генома *Sz. pombe* с геном *fet5<sup>+</sup>*. Эндонуклеазы рестрикции: А – *Sac*I, В – *Bam*HI, С – *Eco*RV, Е – *Eco*RI, G – *Bgl*II, H – *Hind*III, К – *Kpn*I, М – *Sma*I, N – *Not*I, P – *Pst*I, R – *Xba*I, S – *Sph*I, Т – *Stu*I, V – *Pvu*I, X – *Xho*I. В составе вектора клонотеки рНL10 не обнаружены участки узнавания эндонуклеаз рестрикции *Xho*I и *Sal*I, причем последний отсутствует и во вставке.



такой “альтернативный несплайсинг”, требует дальнейшего изучения.

Из приведенной первичной структуры *DraI*-фрагмента видно, что *fet5<sup>+</sup>* представляет собой ген с четко выраженной ТАТА-последовательностью (ТАТААГ) в промоторной области и пиримидин-богатыми последовательностями в двух возможных участках старта транскрипции, удовлетворяющих консенсусу YYANWYY для эукариотических генов [16] и расположенных на расстоянии 50 и 60 п. о. от участка ТАТА (рис. 3). Начало *fet5<sup>+</sup>* мРНК [10] находится в первой из этих последовательностей и соответствует структуре инициаторных участков у генов дрожжей: Y $\Delta$ AR [17]. 3'-Некодирующая область гена содержит два палиндромных участка, состоящих только из АТ-пар (AAAATTTT и ТАТТТАААТА) и участву-

ющих, по-видимому, в терминации синтеза *fet5<sup>+</sup>*-транскрипта (рис. 3).

С целью поиска генов, соседних с *fet5<sup>+</sup>*, была определена первичная структура ряда других участков геномной вставки плазмиды рYUK3, перенесенных в удобные для секвенирования векторы рUC18 и рUC19 [18]. Интересным результатом этой работы явилось обнаружение в *EcoRI/HindIII*-фрагменте длиной 250 п. о. (рис. 2) неполной открытой рамки считывания (ORF), обладающей статистически достоверным подобием гену активатора транскрипции *SPT20/ADA5* [19]. Действительно, этот *EcoRI/HindIII*-фрагмент способен кодировать полипептид длиной 84 а. о., два участка которого явно сходны с фрагментами фактора транскрипции *Spt20p*:

```

NSDVKFYDGC LTVR I IDH      KPNTNSPPVYHTVLRPTPETLWQDL
+ + + + * * + * * + + + * *   * + * * * + * + * * * + * +
Spt20p 201 DCN I QFYEGNL I LQVYDH 218 272 KASFKRPRVYRTLKPNDL TTYUDM 296

```

(\* – идентичные аминокислотные остатки, + – консервативные замены).

Между этими сегментами в обоих случаях предполагаются участки, обогащенные остатками аспарагина и глутамина (соответственно 43 и 28% N + Q), а также пролина (соответственно 14 и 10%). Поскольку ПЦР на суммарной кДНК *Sz. pombe* со специфическими для обнаруженной нами ORF (*orfQP*) праймерами дает четкий сигнал (данные не приведены), не вызывает сомнения, что обсуждаемая *orfQP* является частью ранее не описанного гена делящихся дрожжей *Sz. pombe*.

К настоящему времени в первичной структуре активаторов транскрипции, к которым относится *Spt20p/Ada5p* [19, 20], выявляют три типа функциональных доменов, ответственных за активацию транскрипции: **кислотный**, т.е. обогащенный остатками глутаминовой и аспарагиновой кислот (acidic domain), **глутаминбогатый** и **пролинбогатый домены** [21].

Вовлеченность в регуляцию процесса транскрипции у дрожжей гена *SPT20/ADA5* хорошо документирована недавними статьями двух групп исследователей, изолировавших этот ген в результате независимых, различающихся по используемым подходам, поисков мутантов, устраняющих дефекты транскрипции маркерных генов с ряда сложно регулируемых промоторов [19, 20]. При этом *ADA5/SPT20* явился первым связующим звеном между, казалось бы, совершенно различными наборами *ADA*- и *SPT*-генов, участвующих в регуляции транскрипции у дрожжей [19, 20]. Представляется весьма интересным и тот факт, что ген *SPT20/ADA5* отнесен к группе генов, сход-

ных по своей функции с *SPT15*, который кодирует ТАТА-связывающий белок (ТВР-группа *SPT*-мутантов) [22, 23]. *SPT20*, по-видимому, способен влиять на регуляцию процессов транскрипции, катализируемых всеми тремя типами эукариотических РНК-полимераз (I, II и III). Последнее обстоятельство сближает его функциональные свойства не только с белком ТВР [24], но и с действием *Fet5<sup>+</sup>*, способного супрессировать *ts*-мутацию по субъединице ABC10 $\beta$ , являющейся незаменимым компонентом всех трех ядерных РНК-полимераз [10].

### ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

ДНК геномной клонотекы делящихся дрожжей *Sz. pombe* [10] была трансформирована в штамм *E. coli* MH5 (*trpC-9830 pyrF::Tn5 galU galK hsr rpsL lacZ $\Delta$ X74*), и трансформанты рассеяны на 24 первичных разведения (A1 – A24) по 150–200 колоний в каждом. Вектор клонотекы рHL10 представляет собой челночную плазмиду на основе ампициллинустойчивой фагамиды рBS M13 (+) = рВЕ1 (фирма Stratagene; номер депонирования нуклеотидной последовательности в GenBank L08783\*, в которую по устраненному в процессе клонирования *DraII*-сайту встроен составной *EcoRI*-фрагмент длиной 6.7 т. п. о., содер-

\* В процессе секвенирования вставок гена *fet5<sup>+</sup>* и других генетических детерминант *Sz. pombe* нами обнаружено, что первичная структура участка вектора клонотекы, ведущего свое происхождение из фагамиды рВЕ1, содержит дополнительное звено Т в позиции 961 последовательности L08783.

жащий *HindIII*-вставку (1.8 т. п. о.) гена *ura4<sup>+</sup>* из хромосомы III делящихся дрожжей, участок, ответственный за репликацию плазмиды в *Sz. pombe*, а также синтетический полилинкер с уникальными участками узнавания рестриктаз *SfiI* и *NotI*. Ген *ura4<sup>+</sup>* кодирует один из ферментов биосинтеза урацила, оротидин-5'-фосфатдекарбоксилазу, является гомологом гена *URA3 S. cerevisiae* и используется как фенотипический маркер в делящихся дрожжах [14]. Размер геномных вставок *Sz. pombe* в плаزمидах клонотекки варьирует от 7 до 15 т. п. о. со средним значением 8.5 т. п. о. (по результатам анализа ~50 клонов). Просеивание геномной клонотекки *Sz. pombe* проводили методом последовательных разведений, описанным нами ранее [12, 13].

**Полимеразную цепную реакцию** со специфическими праймерами оGVS308 и оGVS309 на матрице кДНК или геномной ДНК *Sz. pombe* проводили в течение 35 циклов: денатурация – 1 мин при 94°C, отжиг – 1 мин при 55°C и элонгация – 1.5 мин при 72°C. Использовали рекомбинантную *Taq*-полимеразу производства лаборатории биотехнологии ИБХ РАН. Реакцию проводили в буфере состава: 67 мМ трис-HCl (pH 8.8), 6.7 мМ MgCl<sub>2</sub>, 1 мМ дитиотреит, 16.6 мМ (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> и 6.4 мкМ EDTA на амплификаторе DNA Thermal Cycler фирмы Perkin-Elmer Cetus. Продукты ПЦР анализировали электрофорезом в 1.5% агарозном геле в 0.5-кратном трис-боратном буфере [25].

**Физическое картирование и клонирование** осуществляли с помощью ферментных препаратов следующих фирм: New England Biolabs (эндонуклеазы рестрикции *BglII*, *EcoRI*, *EcoRV*, *HindIII*, *KpnI*, *PstI*, *PvuI*, *SacI*, *SalI*, *SphI* и *StuI*), Bethesda Research Laboratory (эндонуклеазы рестрикции *BamHI*, *SmaI*, *XbaI*, *XhoI* и *DraI*; ДНК-лигаза фага T4) и United States Biochemicals (эндонуклеазы рестрикции *NotI* и *SmaI*). Все ферментативные реакции проводили в условиях, рекомендованных производителями.

**ДНК секвенировали** с использованием набора Sequenase (version 2.0) фирмы USB (США) на приборе Macrophor фирмы LKB (Швеция). В качестве метки применяли [ $\alpha$ -<sup>32</sup>P]dATP (ИБХ РАН) с удельной активностью 74 ПБк/моль.

**Анализ первичных последовательностей ДНК** и их сравнение с базами данных GenBank/EMBL проводили с помощью пакета программ BLAST сервера <http://www.ncbi.nlm.nih.gov> [26].

Установленные в настоящей работе первичные структуры гена *fet5<sup>+</sup>* и его кДНК депонированы в GenBank соответственно под номерами AF020907 и U80218.

Настоящая работа частично поддержана грантом № 96-04-49867 Российского фонда фундаментальных исследований (РФФИ) и грантом Государственной научно-технической программы

России “Новейшие направления биоинженерии” (Генная и клеточная инженерия).

Авторы благодарны А.Л. Каюшину за синтез олигонуклеотидных праймеров, Ю.А. Берлину за ценные замечания при подготовке рукописи статьи.

## СПИСОК ЛИТЕРАТУРЫ

1. Shpakovski G.V., Lebedenko E.N. // Abstracts of Papers Presented at the European Congress for Molecular Cell Biology “ECBO 97” (Brighton, England), 1997. P. 85 (Abstract E-5128).
2. Conaway R.C., Conaway J.W. // Annu. Rev. Biochem. 1993. V. 62. P. 161–190.
3. Roeder R.G. // Trends Biochem. Sci. 1996. V. 21. P. 327–335.
4. Koleske A.J., Buratowski S., Nonet M., Young R.A. // Cell. 1992. V. 69. P. 883–894.
5. Verrijzer C.P., Tjian R. // Trends Biochem. Sci. 1996. V. 21. P. 338–342.
6. Berger S.L., Pina B., Silverman N., Marcus G.A., Agapite J., Regier J.L., Triezenberg S.J., Guarente L. // Cell. 1992. V. 70. P. 251–265.
7. Carlson M., Laurent B.C. // Curr. Opin. Cell Biol. 1994. V. 6. P. 396–401.
8. Simchen G., Winston F., Styles C.A., Fink G.R. // Proc. Natl. Acad. Sci. USA. 1984. V. 81. P. 2431–2434.
9. Winston F. // Transcriptional Regulation / Eds S.L. McKnight, K.R. Yamamoto. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press, 1992. P. 1271–1293.
10. Шпаковский Г.В., Лебедеко Е.Н. // Биоорганическая химия. 1997. Т. 23. С. 234–237.
11. Weaver D.C., Shpakovski G.V., Caputo E., Levin H.L., Boeke J.D. // Gene. 1993. V. 131. P. 135–139.
12. Shpakovski G.V. // Gene. 1994. V. 147. P. 63–69.
13. Шпаковский Г.В., Лебедеко Е.Н., Тюрью П. // Биоорганическая химия. 1997. Т. 23. С. 110–117.
14. Bach M.L. // Curr. Genet. 1987. V. 12. P. 527–534.
15. Prabhala G., Rosenberg G.H., Kaufer N.F. // Yeast. 1992. V. 8. P. 171–182.
16. Javahery R., Khachi A., Lo K., Zenzie-Gregory B., Smale S.T. // Mol. Cell. Biol. 1994. V. 14. P. 116–127.
17. Furter-Graves E.M., Hall B.D. // Mol. Gen. Genet. 1990. V. 223. P. 407–416.
18. Yanisch-Perron C., Vieira J., Messing J. // Gene. 1985. V. 33. P. 103–119.
19. Roberts S.M., Winston F. // Mol. Cell. Biol. 1996. V. 16. P. 3206–3213.
20. Marcus G.A., Horiuchi J., Silverman N., Guarente L. // Mol. Cell. Biol. 1996. V. 16. P. 3197–3205.
21. Latchman D.S. Eukaryotic Transcription Factors. London; San Diego; Sydney; Tokyo; Toronto: Acad. Press, 2nd Ed., 1995. P. 239–278.

22. Eisenmann D.M., Dollard C., Winston F. // Cell. 1989. V. 58. P. 1183–1191.
23. Hahn S., Buratowski S., Sharp A., Guarente L. // Cell. 1989. V. 58. P. 1173–1181.
24. Hernandez N. // Genes Dev. 1993. V. 7. P. 1291–1308.
25. Sambrook J., Fritsch E.F., Maniatis T. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press, 1989.
26. Altschul S.F., Gish W., Miller W., Myers E.W., Lipman D.J. // J. Mol. Biol. 1990. V. 215. P. 403–410.

## Exon-Intron Structure of the *fet5<sup>+</sup>* Gene of *Schizosaccharomyces pombe* and the Physical Mapping of the Encompassing Regions of the Genome

G. V. Shpakovski and E. N. Lebedenko

*Shemyakin–Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, ul. Miklukho-Maklaya 16/10, GSP-7 Moscow, 117871 Russia*

**Abstract**—Plasmid pYUK3 bearing the *fet5<sup>+</sup>* gene of *Schizosaccharomyces pombe* was isolated from a genomic library of the fission yeast, and a detailed physical map of the whole genomic insert (ca. 9.6 Kbp) was constructed. The primary structure of the *fet5<sup>+</sup>* gene and its flanking regions is established. The gene contains a single 45-bp intron in its distal part. A typical TATA-box (TATAAG) was found in the 5'-noncoding region ca. 50 bp upstream of the putative start of transcription, and the 3'-noncoding region contains AT-rich palindromes, which are probably involved in termination of the *fet5<sup>+</sup>* transcription. A previously unidentified gene of *Sz. pombe* encoding a protein with some similarity to one of the transcriptional activators from the TBP (TATA-binding protein) group of SPT factors of transcription was found in the vicinity of the *fet5<sup>+</sup>* gene. Taking into account that cDNA of the *fet5<sup>+</sup>* gene was isolated as a suppressor of the genetic defect of nuclear RNA polymerases I – III (*Bioorg. Khim.*, 1997, vol. 23, № 3, pp. 234–237), this vicinity may be the first evidence of possible clustering, in the genome of the fission yeast, of genes participating in transcription regulation.

*Key words:* fission yeast, *fet5<sup>+</sup>* gene, intron, eukaryotic transcription factors, glutamine-rich activatory domain.