

УДК 576.316:577.113+577.315.42

Множественные места локализации нуклеосом ассоциированы с участками связывания ДНК с ядерным матриксом

Федосеева В.Б.*

Институт молекулярной генетики РАН

Аннотация. Нами получены карты нуклеосомной локализации для последовательности более 20 фрагментов ДНК, связывающихся с ядерным матриксом. Проведено сравнение карты с локализацией контекстных характеристик (пурин-пиримидиновые тракты, АТG-, АТC-последовательности, повторы мотивов T₁₋₅G₁₋₅, C₁₋₅A₁₋₅, повторы динуклеотидов TG, CA, TA, мотивы теломерных повторов, АТ-богатые последовательности, сайты ТороII-изомеразы). Выяснено, что для участков связывания с ядерным матриксом длиной более 800 п.н. встречаются множественные сильные сайты нуклеосомной локализации, часто последовательно расположенные друг за другом. С ними соседствуют, как правило, зоны пониженного потенциала нуклеосомной локализации. С этими зонами ассоциируется повышенная частота встречаемости контекстных характеристик (пурин-пиримидиновых трактов, T₁₋₅G₁₋₅/C₁₋₅A₁₋₅-мотивов, АТG-, АТC-последовательностей, повторов пуринов-пиримидиновых динуклеотидов, отдельных теломерных повторов, сайтов ТороII-изомеразы). Исключение в представленной выборке составил пример гена бета-глобина человека, интрон которого содержит сайт связывания с ядерным матриксом, некоторые контекстные характеристики, но не имеет близко расположенных сильных сайтов нуклеосомной локализации.

Ключевые слова: *ядерный матрикс, локализация нуклеосом.*

ВВЕДЕНИЕ

В основаниях петельных структур хроматина при их связывании с ядерным матриксом расположены специфические фрагменты хромосомной ДНК [1]. Связь сохраняется и при удалении гистонов. Такая организация обеспечивает сохранение топологической обособленности одного хромосомного домена от другого. Эти последовательности ДНК предпочтительно связываются с изолированным ядерным матриксом в присутствии значительного избытка конкурентной ДНК [2]. Связывание MAR-элементов с ядерным матриксом не являлось специфичным в отношении типа ткани и вида организма, из клеток которого выделен ядерный матрикс [2, 3]. Эти регионы прикрепления к матриксу (в дальнейшем MARs/SARs) были изолированы и клонированы для определения нуклеотидной последовательности. Существует общепринятый набор характеристик последовательности фрагментов закоривания [4]: полиА-, полиТ-тракты, различные T_nG_m-мотивы, пуринов-пиримидиновые тракты, (А,Т)-богатые фрагменты, участки гибкости ДНК (пуринов-пиримидиновые динуклеотидные тракты). Однако локализация этих характеристик не упорядочена. Для последовательностей, ранее опубликованных в литературе, нами осуществлен комплексный анализ, основанный на возможности позиционирования нуклеосом на фрагментах ДНК [5], локализации сайтов топоизомеразы II (известно, что MARs

*fvb@img.ras.ru

обогатены участками расщепления этой изомеразы [2, 6, 7], и сам фермент изомеразы ассоциирован с ядерным матриксом) и других вышеупомянутых контекстных характеристик. Добавлены также участки узнавания белка, присутствующего в том или ином количестве в матриксе разных тканей и связывающегося с MARs из разных видов [8]. Для фрагментов генома человека в схему анализа включены также теломерные повторы млекопитающих, так как известно, что теломерные повторы связываются с ядерным матриксом [9, 10].

МАТЕРИАЛЫ И МЕТОДЫ

Для анализа мы использовали последовательности около 20 нуклеотидных фрагментов (MARs/SARs) в диапазоне длин от ~300 п.н. до ~11 т.п.н. Ранее нами был предложен комбинированный метод определения положения нуклеосом относительно последовательности ДНК [5], в этом методе используются две матрицы иерархически выстроенных тринуклеотидных вкладов в изгиб ДНК [11, 12] в комбинациях, зависящих от АТ-состава текущего нуклеосомного фрагмента (145 п.н.), что позволяет вычисленным сайтам локализации наилучшим образом соответствовать различным экспериментальным данным. Представлен потенциал $P(x)$ позиционирования нуклеосом для x -координаты вдоль фрагмента ДНК (5'-конец нуклеосомы). Длина нуклеосомного фрагмента выбрана равной 145 н.п. В работе на картах нуклеосомного позиционирования также представлены сайты топоизомеразы II – предсказания на основе консенсуса для ТороII плодовой мухи *Drosophila m.* [13] или для ТороII позвоночных [14].

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

Были исследованы последовательности из GenBank, отнесенные к группе MARs/SARs: фрагменты ДНК золотистого хомяка *C. aureus*, связывающиеся с ядерным матриксом (X96546 – X96549 [15]), фрагменты ДНК клеток табака (AF065877-AF065888) [16], фрагменты 19-ой хромосомы человека (Z35288, Z35290, Z35291, Z35279, Z35280, Z35282) [17], фрагменты хромосомы человеческого бета-глобинового комплекса [18], область между генами гистонами H1 и H3 плодовой мухи *D. melanogaster* [6], MARs последовательности для региона гена бета-интерферона человека [19], MARs последовательности, отстоящие от 5'-конца гена лизоцима *Gallus Gallus* [20] на 11.1 т.п.н. Для приведенных последовательностей построены карты зависимости потенциала нуклеосомного позиционирования вдоль длины фрагмента. Карта нуклеосомного позиционирования получена на основе опубликованного ранее комбинированного метода их локализации [5]. При этом каждой точке вдоль последовательности ставится в соответствие потенциал нуклеосомного позиционирования (в нашем случае 5'-конца нуклеосомного фрагмента) после усреднения в окне 30 п.н. Сильные пики соответствуют более стабильно локализованным нуклеосомам, чем слабо выраженные, иерархия пиков может соответствовать последовательности заполнения гистонами сайтов нуклеосомной локализации. Одновременно представлены количественные данные таких контекстных характеристик, как длина пиримидиновых и/или пуриновых трактов в виде гистограмм вдоль последовательности фрагментов. Интерес представляют достаточно протяженные пуриновые и пиримидиновые тракты, выявленные ранее как контекстная характеристика MARs/SARs. Они тракуются, как участки H-форм, представляющие собой горячие точки рекомбинационных событий. Обычно рассматриваются тракты, начиная с 10 н.п., но с учетом того, что трактов разных длин в рассматриваемых областях очень много, и, с другой стороны, H-форма остается достаточно стабильной и для протяженных трактов [21], содержащих одно и двух буквенные замены, мы учитываем также тракты начиная с 12 н.п. с однобуквенной заменой и 16 н.п. с двухбуквенными заменами. Кроме того, на рисунках представлены сплошные повторы типа T₁₋₅G₁₋₅ (или/и C₁₋₅A₁₋₅) (ими также обогатены

протяженные MARs/SARs-последовательности [22]) длиной, начиная с 10 п.н. при отсутствии буквенных замен, 12 н.п. с одной допустимой заменой или 16 н.п. при двух нуклеотидных заменах. На этих же рисунках для сравнения представлены места расщепления ДНК белком топоизомеразы II, вычисленными по консенсусу для *D. melanogaster* [13] или для позвоночных [14]. В отдельных случаях на общих схемах представлены повторы альтернирующих пиримидин-пуриновых динуклеотидов, если они достаточно длинные. Они также могут служить горячими точками рекомбинации. Так как известно, что теломерные повторы связываются с ядерным матриксом, их присутствие также отмечено на общей схеме, в частности, для млекопитающих отмечен повтор TTAGGG.

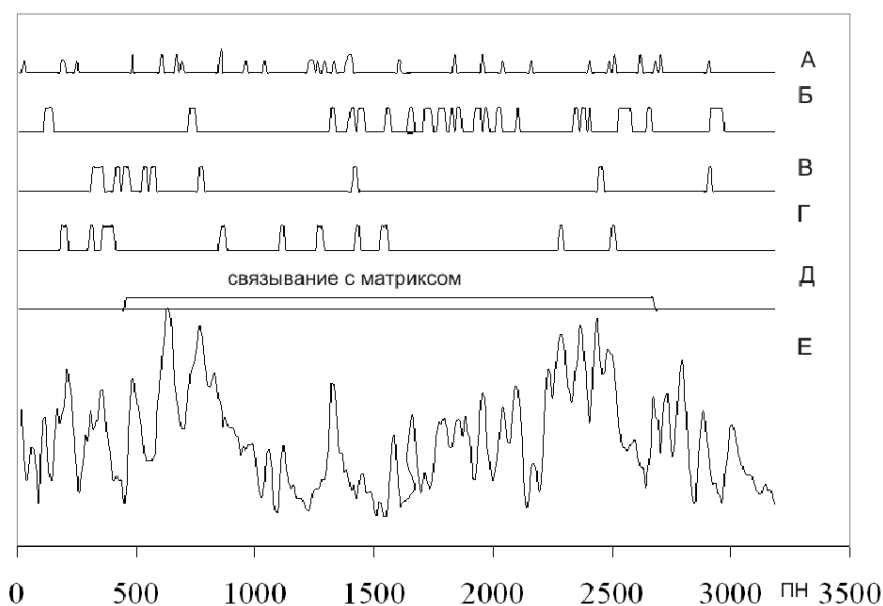


Рис. 1. Схема локализации контекстных характеристик MAR, отстоящего на 11.9 т.н.п. от 5'-конца гена лизоцима *Gallus Gallus*. **А.** Локализация сайтов топоизомеразы II, полученных в соответствии с консенсусом для позвоночных (максимум соответствует 90% совпадению). **Б.** Локализация ATG-, ATC-содержащих фрагментов. **В.** Локализация пуриновых и пиримидиновых трактов. **Г.** Локализация повторов T₁₋₅G₁₋₅-мотивов или/и C₁₋₅A₁₋₅. **Д.** Объединение фрагментов рестрикции, имеющих преимущественное связывание с ядерным матриксом. **Е.** Локализация мест нуклеосомного позиционирования (максимум соответствует 180 условным единицам, минимум – 100).

В дополнение к перечисленным контекстным характеристикам, следует воспользоваться информацией о белках, которые связываются с последовательностями, входящими в MARs/SARs. Один из них [8] связывается с AT-богатыми мотивами с асимметричным расположением гуаниновых оснований между двумя нитями (в одну нить двойной спирали входит гуанин, а в комплементарную - цитозин). Это так называемые ATC- или ATG-распознающие белки. Выявленная длина участков связывания составляет от 8 до 32 п.н. Нами также проанализированы исследованные последовательности с целью выявления ATC- и ATG-мотивов, что также отражено в виде гистограмм на общей схеме. Показано, что короткие MARs длиной около 300 н.п. (*Homo sapiens*) из списка встретившихся последовательностей MARs/SARs практически не содержат средние по силе сайты нуклеосомного позиционирования, но включают ТороII-сайты по консенсусу для позвоночных [14], также пуриновые и пиримидиновые тракты от 8 до 70 п.н. длиной. Особенно сильные сайты нуклеосомного позиционирования представлены в последовательностях MARs, начиная с 800 н.п. Анализ всех представленных последовательностей и их контекстных характеристик и сравнение с картами локализации нуклеосом свидетельствует о

наличие ряда общих закономерностей. Приведем несколько примеров анализа общих схем.

На рис. 1 представлена общая схема локализации перечисленных характеристик для MAR, ассоциированного с 5'-концом гена лизоцима *Gallus Gallus*. Из рис. 1Е видно, что потенциал нуклеосомного позиционирования соответствует, по крайней мере, двум зонам сильных пиков, соседствующими с участками низкого потенциала нуклеосомной локализации. Вид первой зоны (490 – 895 п.н.) дает представление об эквидистантной стабильной локализации трех нуклеосом. Как следует из сравнения рис. 1А, Б, Г с рис. 1Е ATG-, АТС- содержащие фрагменты, T_nG_m-тракты, пурин/пиримидиновые последовательности и сайты ТороII в значительной степени приходятся на участки низкого потенциала нуклеосомной локализации.

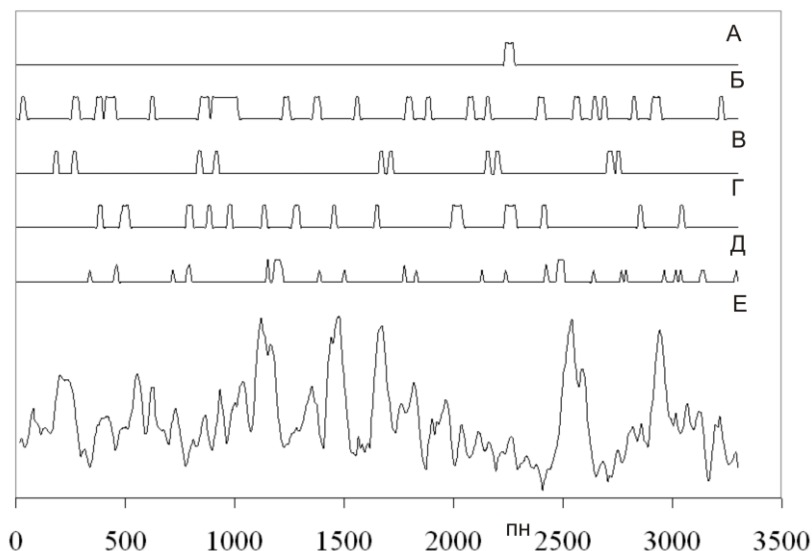


Рис. 2. Схема локализации контекстных характеристик MAR золотистого хомячка, нуклеотидная последовательность с идентификатором X96548 в GenBank. А Динуклеотидные повторы TG. Б. Локализация ATG-, АТС-содержащих фрагментов. В. Локализация пуриновых и пиримидиновых трактов. Г. Локализация повторов T₁₋₅G₁₋₅- или C₁₋₅A₁₋₅ мотивов. Д. Локализация сайтов топоизомеразы II, полученных в соответствии с консенсусом для позвоночных. Е. Локализация мест нуклеосомного позиционирования (максимум соответствует 180 условным единицам, минимум – 100).

Как видно из рис. 2Е, картина мест локализации нуклеосом представлена 5-ю сильными пиками, 3 из которых представляют собой компактную группу (координаты 1075 – 1675 п.н.), окруженную зонами низкого потенциала нуклеосомной локализации (0-1075 и 1746-2439). Сравнение рис. 2Е с рисунками 2А, 2Б, 2Г, 2Д создает представление о том, что короткие ATG-, АТС-содержащие фрагменты, T_nG_m-тракты и пурин/пиримидиновые последовательности относительно равномерно распределены по всему представленному участку, однако более длинные варианты этих последовательностей соответствуют зонам низкого потенциала нуклеосомной локализации.

На рис. 3 представлена схема контекстных характеристик для MAR между генами гистонов H1 и H3 *D. melanogaster*. Как следует из рис. 3Г и 3Д, область связывания с ядерным матриксом (720-1209 п.н.) ограничена, по крайней мере, с 3'-конца цепочкой пиков нуклеосомной локализации с координатами 1209, 1350, 1490, 1650, 1801 п.н. Также видно, что на область связывания с ядерным матриксом (рис. 4А, 4Б, 4В) приходится большее количество протяженных трактов (ATG-, АТС-последовательностей, T_nG_m-повторов) и сайтов ТороII. Районы генов гистонов H1 и H3 обеднены пиками потенциала нуклеосомной локализации и прочими контекстными характеристиками по сравнению с межгенным пространством.

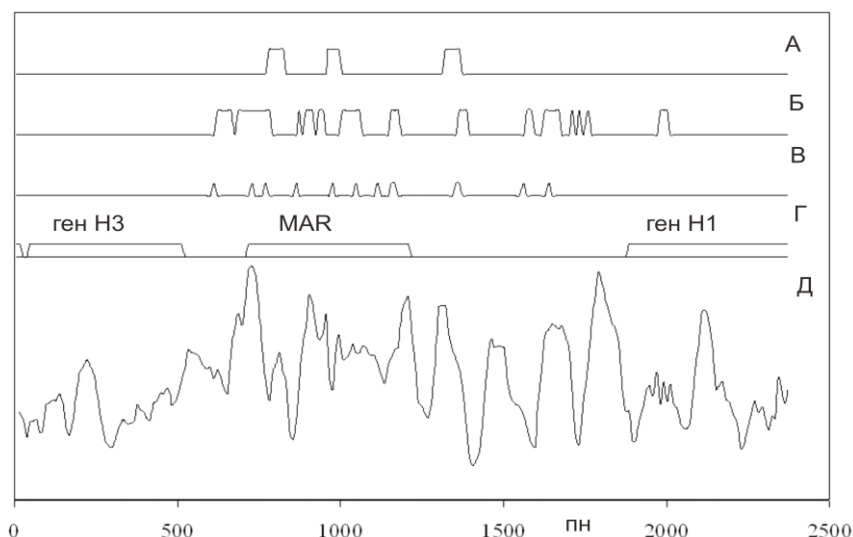


Рис.3. Схема локализации контекстных характеристик MAR, локализованного между генами H1 и H3 генами *D. melanogaster*. А. (АТ)-тракты. Б. Локализация АТG-, АТC-содержащих фрагментов. В. Локализация сайтов топоизомеразы II, полученных в соответствии с консенсусом для *D. melanogaster*. Г. Схема расположения генов гистонов H1 и H3 *D. melanogaster* и непосредственно MAR-последовательности.

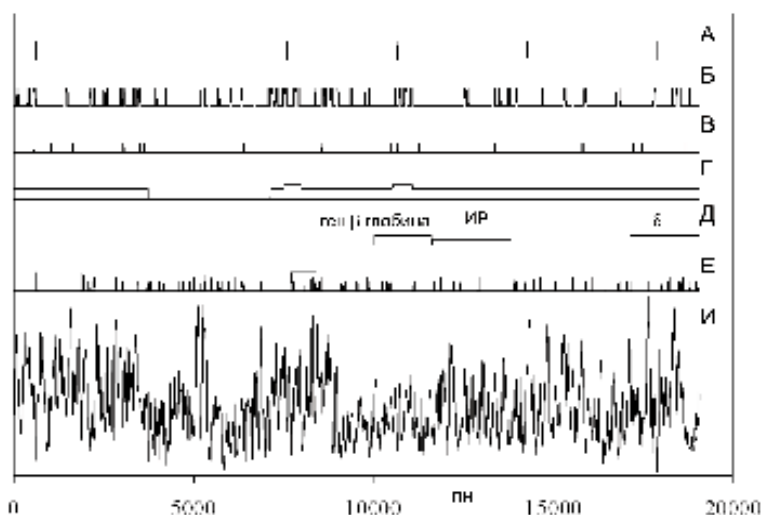


Рис.4. Схема локализации контекстных характеристик для части бета-глобинового локуса *Homo sapiens* по последовательности NT_009237 GeneBank. А. Гистограмма локализации фрагментов альтернирующих пурин-пиримидиновых динуклеотидов. Б. Локализация АТG-, АТC-содержащих фрагментов. В. Локализация ТTAGGG-последовательностей, составляющих теломерные повторы млекопитающих и человека. Г. Схема локализации различных фрагментов рестрикции (SARs), амплитуда соответствует силе связывания каждого фрагмента с ядерным матриксом. Д. Схема расположения бета- и дельта-глобиновых генов, а также множественных мест инициации репликации ДНК (ИР). Е. Локализация сайтов для топоизомеразы II, полученных в соответствии с консенсусом для позвоночных (максимум соответствует 90% совпадению), сплошная линия – экспериментальные сайты для области бета-глобинового гена. И. Локализация мест нуклеосомного позиционирования, максимум соответствует 180 условным единицам, минимум – 100.

На рис. 4 представлены карты локализации описанных выше характеристик для MAR-последовательности для части бета-глобинового локуса *H. sapiens*. Представленный регион содержит обширную область связывания с ядерным матриксом (рис. 4Г), в частности, участки (7627-8159 и 10770-11257 п.н.) соответствуют наиболее

сильной связи с ядерным матриксом. На рис. 4И выявляются зоны, включающие пики повышенного потенциала нуклеосомного позиционирования: 0-3645, 5106-5262, 7053-8500, 14344-16308 п.н. Эти зоны прерываются участками потенциала низкого уровня, например, около точки 1920 п.н., 3656-5095, 5275-7053, в районе точки 8190, в зоне 9272-14344 п.н. Как следует из рисунков 4А-4В, плотность АТG-, АТC-содержащих последовательностей, повторов альтернирующих пурипиримидиновых динуклеотидов повышена для участков, свободных от сильных сигналов нуклеосомного позиционирования. Обращает на себя внимание и тот факт, что участку сильного связывания с ядерным матриксом 7627 – 8159 п.н. сопутствует с 3'-конца ряд последовательных пиков потенциала нуклеосомной локализации, начиная с координаты 8298 п.н., а в окрестности участка 10770-11257 п.н. заметные по силе пики нуклеосомного позиционирования не встречаются. Заметим, что этот участок, в отличие от первого, приходится на интрон бета-глобинового гена.

Аналогичные наблюдения о присутствии областей с последовательно расположенными сильными пиками нуклеосомного позиционирования и примыкании к ним областей с низким потенциалом без сильных пиков можно отметить и для большинства из 20-ти последовательностей MARs/SARs, перечисленных выше. С областями низкого потенциала в значительной степени ассоциируются контекстные характеристики (различные T_nG_m -мотивы, пурипиримидиновые тракты, АТG-, АТC-последовательности, пурипиримидиновые динуклеотидные тракты, сайты ТороII-изомеразы) или, в случае трактов, их большая протяженность. В этом ряду несколько обособленное место занимают MARs/SARs растений, так как для них именно характерно присутствие АТ-богатых последовательностей в области низкого потенциала.

Пики из областей высокого потенциала, как правило, обогащены $A_{3-5}T_{3-5}$ -трактами для всех проанализированных последовательностей.

ОБСУЖДЕНИЕ

В дополнение к известным ранее характеристикам нуклеотидного контекста, соответствующих MARs/SARs, удалось выявить сильные пики нуклеосомного позиционирования. В одних случаях – это изолированные пики; в случае сравнительно длинных фрагментов – это зоны повышенного потенциала позиционирования, включающие цепочки пиков, последовательно отстоящих на расстояние, превышающее или сравнимое с 145 п.н. Этим пикам, как правило, соответствуют А- и Т-тракты длиной 3-5 п.н. Вероятно, сильные пики определяют более устойчивое состояние нуклеосомной локализации, чем слабо выраженные, в которых нуклеосомы могут и вовсе отсутствовать на определенных фазах клеточного цикла, либо иерархия пиков определяет последовательность образования нуклеосом. Заметная корреляция наблюдается в том, что рядом с сильными пиками наблюдаются зоны выраженного минимума потенциала позиционирования, иногда достаточно протяженного. В этих зонах наблюдается заметная обогащенность пуриновыми и пиримидиновыми трактами. Как принято считать эти тракты создают горячие точки гомологичной рекомбинации. Для этих зон потенциального минимума характерны и тракты, обозначенные здесь как T_nG_m , они представляют собой разнообразные мотивы, в которых заметны ассоциации с теломерными последовательностями дрожжей (TGTGTGGGTGTGG), позвоночных (TTAGGG), теломерные повторы *Oxytricha* (TTTTGGGG), а также цепочки пурипиримидиновых динуклеотидов типа (TG/CA), которые также, как и вышеупомянутые тракты, могут инициировать рекомбинацию. Нами также проанализированы исследованные последовательности с целью выявления АТC- и АТG-мотивов. Определенное их количество выявлено в области зон с низким потенциалом позиционирования нуклеосом, значительно большее, чем в районах пиков потенциала нуклеосомной локализации.

СПИСОК ЛИТЕРАТУРЫ

1. Gasser S.M., Laemmli U.K. *Trends Genet.* 1987. **3**. 16-22.
2. Cockerill P.N., Garrard W.T. *Cell.* 1986. **44**. 273-282.
3. Amati B., Gasser S.M. *Mol. Cell Biol.* 1990.**10**. 5442-5454
4. Gasser S.M., Amati B.B., Cardenas M.E. Hofmann J.F.Z. *Internat. Rev. Cytology.* 1989. **119**. 57-96.
5. Fedoseyeva V.B., Alexandrov A.A. *J. Biomol. Struct. Dynam.* 2007. **24**. 481-488.
6. Gasser S.M., Laemmli U.K. *EMBO J.* 1986. **5**. 511-518.
7. Gasser S.M., Laroche T., Falquel J., Boy de la Tour E., Laemmli U.K. *J. Mol. Biol.* 1986. **188**. 613-629.
8. Dickinson L.A., Tadashi J., Kohri Y., Kohwi-Shigematsu T. *Cell.* 1992. **70**. 631-645.
9. De Lange T. *EMBO J.* 1992. **11**. 717-724.
10. Luderus M.E., van Steensel B., Chong L., Sibon O.C. et al. *J. Cell. Biol.* 1996. **135**. 867-881.
11. Bruckner I., Sanchez R., Suck D., Pondor S. *EMBO J.* 1995. **14**. 1812-1818.
12. Satchwell S.C., Drew H.R., Travers A. *J.Mol.Biol.* 1986. **191**. 659-675.
13. Sander M., Tao-Shih H. *Nucl.Acids Res.* 1985. **13**. 1057-1071.
14. Spitzner J.R., Chung I.K., Muller M.T. *Nucl.Acids Res.* 1989. **18**. 1-11.
15. Fernandez M.A., Barav B., Prigent M., Toledo F. et al. *J.Cell.Biochem.* 1997. **67**. 541-551.
16. Michalowski S.M., Allen G.C., Hall G.E.Jr., Thompson W.F., Spiker S. *Biochem.* 1999. **38**. 12795-12804.
17. Nikolaev L.G., Tsevegiyn T., Akopov S.B., Ashworth L.K., Sverdlov E.D. *Nucl. Acids Res.* 1996. **24**. 1330-1336.
18. Jarnan A.P., Higgs D.R. *EMBO J.* 1988. **7**. 3337-3344.
19. Bode J., Maass K. *Biochem.* 1988. **27**. 4706-4711.
20. Loc Phi-Van, Stratling W.H. *EMBO J.* 1988. **7**. 655-664.
21. Belotserkovskii B.P., Veselkov A.G., Filippov S.A., Dobrynin V.N. et al. *Nucl. Acids Res.* 1990. **18**. 6621-6624.
22. Глазко Г.В., Рогозин И.Б., Глазков М.В. *Мол. биол.* 2000. **34**. 5-10.

Материал поступил в редакцию 27.04.2007, опубликован 07.06.2007