

## **SINE-ЭЛЕМЕНТЫ В ГЕНОМАХ МЛЕКОПИТАЮЩИХ МОГУТ СЛУЖИТЬ ВСПОМОГАТЕЛЬНЫМИ СИГНАЛАМИ ПРИ ФОРМИРОВАНИИ ФАКУЛЬТАТИВНОГО ГЕТЕРОХРОМАТИНА**

© **Н. М. Усманова, В. И. Казаков, Н. В. Томили**

*Институт цитологии РАН, Санкт-Петербург;  
электронный адрес: nmusmanova@gmail.com*

В данной работе методами компьютерного анализа мы установили распределение коротких рассеянных нуклеотидных элементов (*SINEs*) в X-хромосомах человека и мыши. Было показано, что это распределение подобно распределению CpG-островков и генов, но сильно отличается от распределения *LINE1*-элементов. Поскольку некоторые *SINEs* (Alu-повторы человека и B2-повторы мыши) могут содержать сайты связывания для YY1, являющегося белком комплекса Polycomb, мы предположили, что эти повторы могут исполнять роль вспомогательных сигналов («boosters») в Polycomb-зависимом глушении богатых генами сегментов в ходе X-инактивации.

Ключевые слова: инактивация X-хромосомы, *LINE1*, *SINE*, белковый комплекс Polycomb, транскрипционный фактор YY1, Xist-РНК.

Принятые сокращения: H2A-Ub — убиквитинилирование гистона H2A, H3-K27m3 — триметилирование гистона H3 по лизину в 27-м положении, L1 — *LINE1*-повторы, ХИС — центр инактивации X-хромосомы, Xist-РНК — специфический транскрипт, экспрессирующийся в ходе инактивации X-хромосомы, YBS — сайт связывания YY1.

Сайленсинг одной из X-хромосом в клетках млекопитающих у самок инициируется в ходе раннего эмбрионального развития путем экспрессии нетранслируемой РНК, Xist, длиной 17 кб, которая «покрывает» хромосому и необходима для ее инактивации (Heard, 2004). Каким-то образом Xist-РНК инициирует модификации хроматина, в частности триметилирование гистона H3 по лизину в 27-м положении (H3-K27m3) (Rougeulle et al., 2004) и убиквитинилирование гистона H2A (H2A-Ub) (Smith et al., 2004), которые катализируются комплексом Polycomb и подавляют транскрипцию (Cao, Zhang, 2004; Cao et al., 2005). Репрессивные модификации хроматина берут начало в так называемом центре инактивации X-хромосомы (ХИС) и затем распространяются вдоль X-хромосомы с помощью вспомогательных сигналов — «way stations» или «boosters» (Gartler, Riggs, 1983). Была выдвинута гипотеза о том, что ретротранспозоны *LINE1* (L1) могут служить в качестве таких сигналов при распространении факультативного гетерохроматина по X-хромосоме в ходе ее инактивации (Lyon, 1998). Очень высокая плотность L1-элементов характерна для X-хромосом некоторых млекопитающих, а у человека плотность L1 максимальна в ХИС-участке и уменьшается в обе стороны от него (Bailey et al., 2000), однако плотность L1 в X-хромосоме мыши сходна с таковой в аутосомах (Chureau et al., 2002). Основным доказательством участия L1 в инактивации X-хромосомы является то, что эта хромосома обогащена L1, а также то, что они избегают районов, содержащих не инактивированные при гетерохроматинизации гены (Bailey et al., 2000). Однако более детальный анализ

показал, что негативная корреляция между плотностью L1 и числом генов, избегающих X-инактивации, не столь очевидна (Carrel, Willard, 2005). Было показано, что окружение инактивированных генов обогащено олигомерами, большинство из которых является частью L1-повторов. Однако некоторые из олигомеров, являющиеся частью L1, ассоциированы с генами, избегающими инактивации (Carrel et al., 2006). Помимо этого, распределение L1 в неактивной X-хромосоме мышей отличается от распределения убиквитинилированного гистона H2A и Xist-РНК, которые локализуются, как правило, в богатых генами сегментах (Duthie et al., 1999; Smith et al., 2004). В инактивированной X-хромосоме человека гистон H3-K27m3 также сосредоточен в богатых генами цитогенетических дисках (Chadwick, Willard, 2004), которые не совпадают с максимумами плотности L1 (Bailey et al., 2000). Однако ген-богатые регионы в геноме человека и мыши обогащены *SINE*-элементами (Korenberg, Rykowski, 1988; Jurka et al., 2005).

Известно, что многие представители Alu-повторов (распространенного семейства *SINEs*-повторов человека) взаимодействуют с транскрипционным фактором YY1 (Humphrey et al., 1996; Oei et al., 2004). Недавно было обнаружено, что YY1 входит в состав комплекса Polycomb, содержащего метилтрансферазу гистона H3 EZH2, которая катализирует H3-K27-триметилирование. Этот комплекс подавляет экспрессию мышечных «мышечных» генов в недифференцированных миобластах (Caretta et al., 2004), что подтверждает возможность участия YY1 в Polycomb-зависимом глушении генов в клетках млекопитающих.

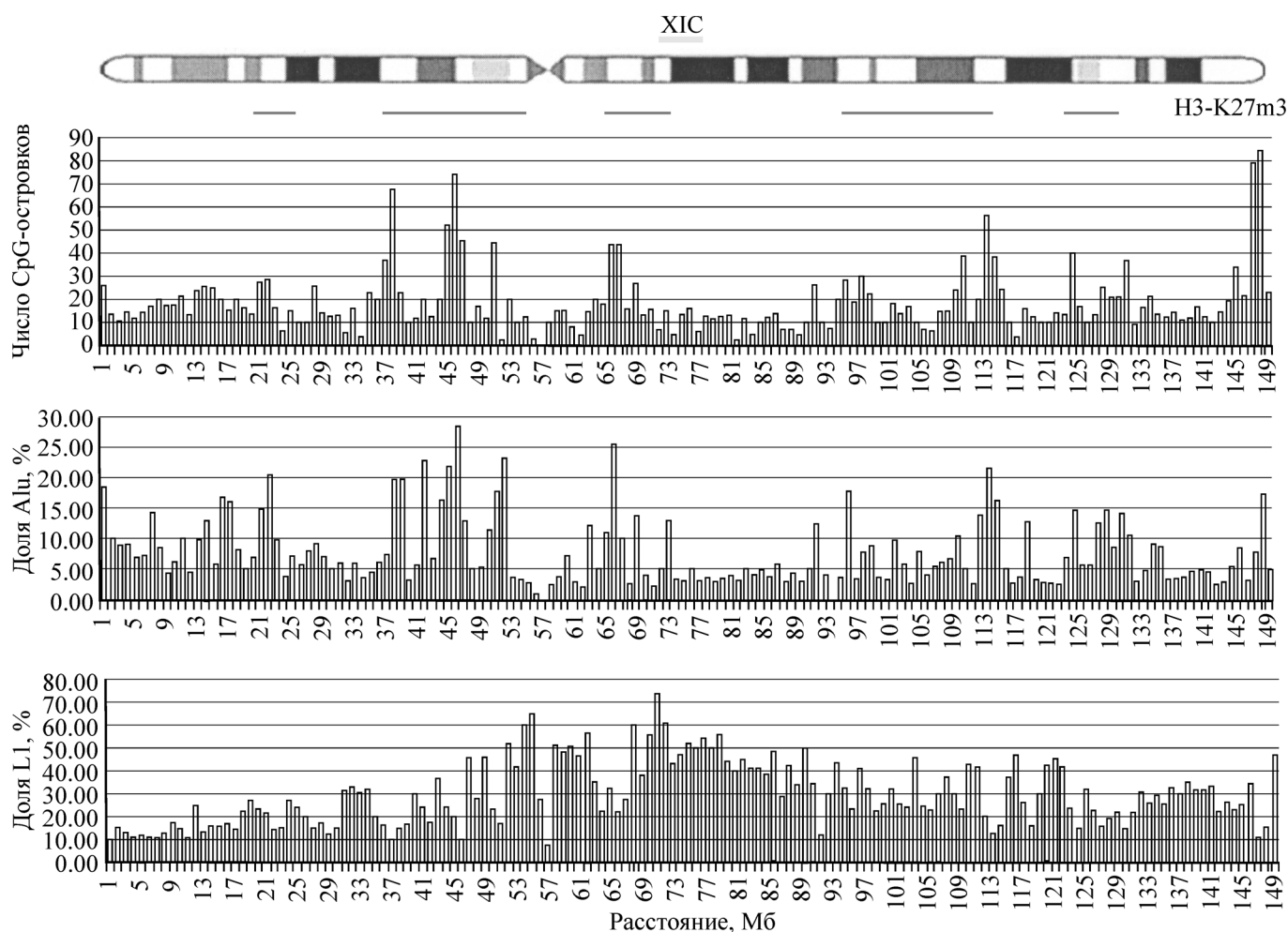


Рис. 1. Распределение Alu-повторов, CpG-островков и *LINE1*-повторов (L1) в X-хромосоме человека. Распределение H3-K27m3 в инактивированной X-хромосоме человека взято из: Chadwick, Willard, 2004.

CpG-островки также могут влиять на статус X-хромосомы, поскольку метилирование ДНК — один из механизмов поддержания ее в инактивированном состоянии. Анализ X-хромосомы показал, что распределение генов, ассоциированных с островками, не различается между генами, избегающими инактивации, и полностью инактивированными генами (Carrel, Willard, 2005).

Задачей данной работы явилось установление распределений *SINE*-повторов, L1-повторов, CpG-островков и потенциальных сайтов связывания YY1 (YBS) в X-хромосомах человека и мыши. Нуклеотидная последовательность X-хромосомы взята из базы данных Ensembl v.44 ([www.ensembl.org](http://www.ensembl.org)). CpG-островки были определены в составе X-хромосомы при помощи программы CPGPLOT (<http://www.ebi.ac.uk/emboss/cpgplot/>). Для их поиска были использованы стандартные критерии: длина островка должна составлять как минимум 200 н. п., состав GC должен быть не менее 50 %, отношение наблюдаемого числа CpG-динуклеотидов к ожидаемому — более 0.6. Alu- и L1-повторы были идентифицированы при помощи программы RepeatMasker (<http://www.repeatmasker.org/>). Поиск сайтов связывания транскрипционного фактора YY1 (последовательность CCATSTTGNC) проводили при помощи программы RSA-tools-patser (<http://rsat.ulb.ac.be/rsat/>) с использованием следующих опций командной строки: bin/patser-Aa:t 0.3 c:g 0.2-m tmp/patser.

2006\_03\_27.153217.matrix -b 1 -c -d1 -ls 11 -f. Эти опции позволяют идентифицировать только точно совпадающие с указанным мотивом последовательности.

После этого мы сравнили распределение повторов L1, *SINE* и CpG-островков в X-хромосомах человека и мыши. Как видно на рис. 1, эти распределения сходны для Alu-повторов и CpG-островков в X-хромосоме человека, в то время как распределение L1 отличается от распределения Alu-повторов и CpG-островков. L1-повторы преимущественно группируются в районах с низкой плотностью Alu и CpG. Следует также заметить, что кластеры Alu-повторов и CpG-островков перекрываются с кластерами известных и предсказанных генов, представленных в базе данных Ensembl v.44, и районами, обогащенными гистонотом H3-K27m3, в инактивированной X-хромосоме (Chadwick, Willard, 2004).

На рис. 2 представлены соответствующие распределения *SINE*-повторов (B1 и B2—B4), L1-повторов и CpG-островков в X-хромосоме мыши. Как видно на рис. 2, повторы B1 и B2—B4 имеют распределения, подобные распределению CpG-островков, но отличные от распределения L1. Помимо этого, оказалось, что распределение *SINE*-повторов и CpG-островков также перекрываются с распределением генов и с указанным в литературных источниках распределением Xist-РНК в хромосомах во время поздних стадий митоза (Smith et al., 2004).

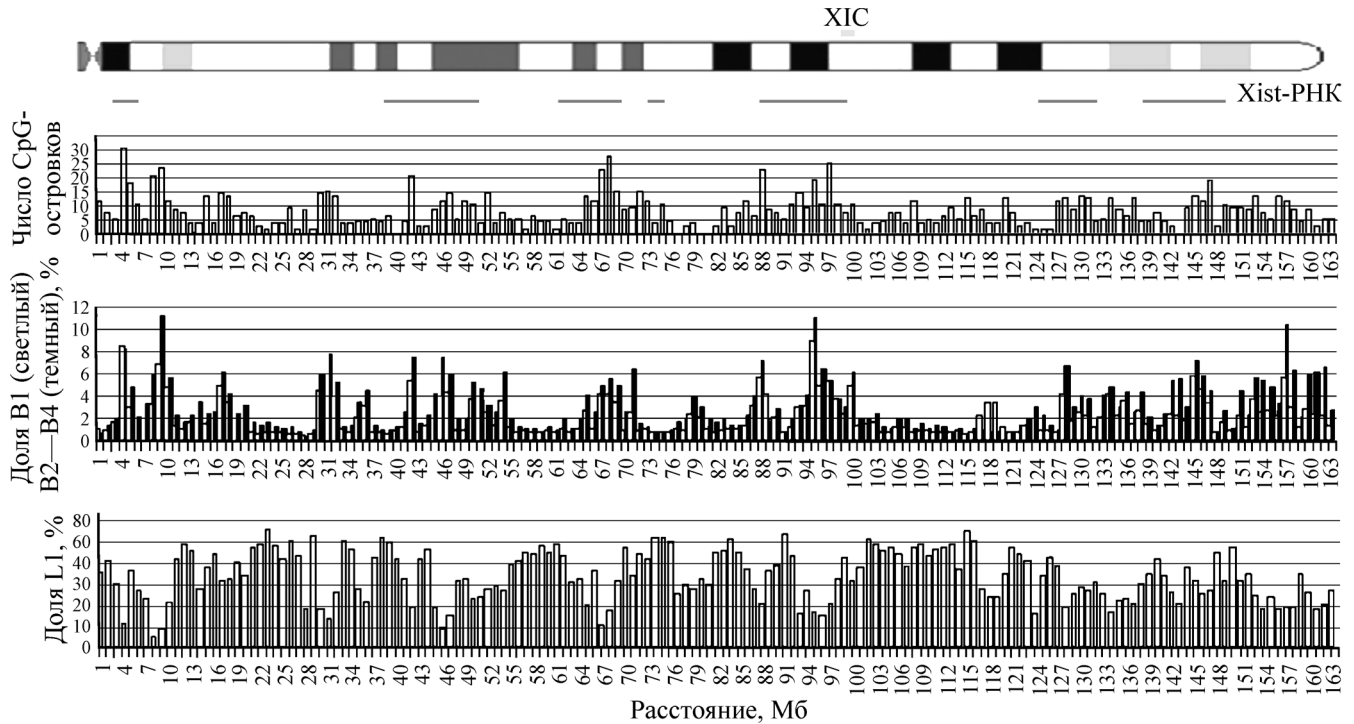
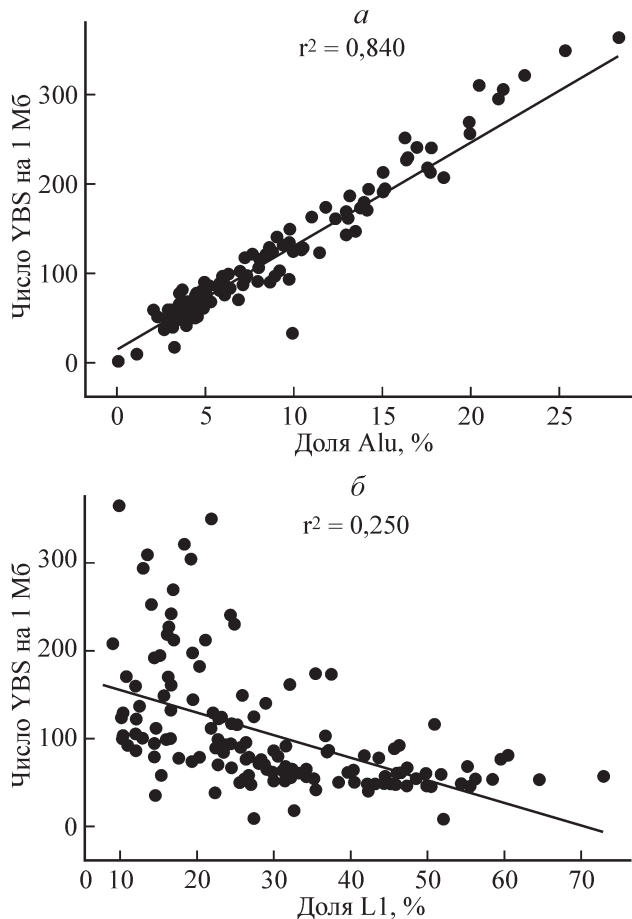


Рис. 2. Распределение *SINE*-повторов (B1, B2-B4), CpG-островков и L1 в X-хромосоме мыши.

Распределение Xist-PHC в инактивированной X-хромосоме мыши взято из: Smith et al., 2004.



Известно, что привлечение белкового комплекса Polycomb, который индуцирует триметилирование H3-K27, зависит от Xist-PHC (Heard, 2004). Таким образом, наши результаты согласуются с идеей о том, что именно *SINE*-повторы млекопитающих (а не L1) служат дополнительными сигналами глушения транскрипции в генсодержащих районах X-хромосомы в процессе ее инактивации по пути, зависящему от белкового комплекса Polycomb. Однако мы не исключаем участия *LINE1*-повторов в гетерохроматинизации районов генома, не содержащих генов, по пути, не зависящему от Polycomb.

Поскольку многие Alu-повторы специфически связываются с YY1, мы решили проанализировать распределение потенциальных Alu-ассоциированных YBS в X-хромосоме и сопоставить его с распределением Alu-повторов. Мы обнаружили (рис. 3), что распределение Alu-повторов (но не L1-повторов) строго коррелирует с распределением YBS, т. е. большое число Alu-повторов, локализованных в X-хромосоме, может связывать YY1. Основная фракция потенциальных YBS в промоторах «мышечных» генов человека также ассоциирована с Alu-повторами. Однако мотив CCATSTTGNC не обнаруживает корреляции с *SINE*-повторами мыши. Возможно, что YBS у мыши слегка отличается от такового у человека. Консенсусная последовательность B2-элементов у мыши содержит мотив GCCATCTCACC, который, вероятно, и является сайтом связывания YY1 у мышей, что, однако, требует экспериментального подтверждения.

ДНК-связывающие домены YY1 и его гомологов весьма консервативны в ряду дрозофила—человек

Рис. 3. Корреляция между потенциальными сайтами связывания YY1 (YBS) и плотностью Alu-повторов (a) и L1 (b) в X-хромосоме человека.

YBS = CCATSTTGNC.



factor results in peri-implantation lethality. *Mol. Cell Biol.* 19 : 7237—7244.

Duthie S. M., Nesterova T. B., Formstone E. J., Keohane A. M., Turner B. M., Zakian S. M., Brockdorff N. 1999. Xist RNA exhibits a banded localization on the inactive X chromosome and is excluded from autosomal material in cis. *Hum. Mol. Genet.* 8 : 195—204.

Gartler S. M., Riggs A. D. 1983. Mammalian X-chromosome inactivation. *Annu. Rev. Genet.* 17 : 155—190.

Heard E. 2004. Recent advances in X-chromosome inactivation. *Curr. Opin. Cell Biol.* 16 : 247—255.

Humphrey G. W., Englander E. W., Howard B. H. 1996. Specific binding sites for a pol III transcriptional repressor and pol II transcription factor YY1 within the inter-nucleosomal spacer region in primate Alu repetitive elements. *Gene Expr.* 6 : 151—168.

Jin Q., van Eynde A., Beullens M., Roy N., Thiel G., Stalmans W., Bollen M. 2003. The protein phosphatase-1 (PP1) regulator, nuclear inhibitor of PP1 (NIP1), interacts with the polycomb group protein, embryonic ectoderm development (EED), and functions as a transcriptional repressor. *J. Biol. Chem.* 278 : 30 677—30 685.

Jurka J., Kohany O., Pavlicek A., Kapitonov V. V., Jurka M. V. 2005. Clustering, duplication and chromosomal distribution of mouse SINE retrotransposons. *Cytogenet. Genome Res.* 110 : 117—123.

Korenberg J. R., Rykowski M. C. 1988. Human genome organization: Alu lines, and the molecular structure of metaphase chromosome bands. *Cell.* 53 : 391—400.

Kuzmichev A., Jenuwein T., Tempst P., Reinberg D. 2004. Different EZH2-containing complexes target methylation of histone H1 or nucleosomal histone H3. *Mol. Cell.* 14 : 183—193.

Lyon M. F. 1998. X-chromosome inactivation: a repeat hypothesis. *Cytogenet. Cell Genet.* 80 : 133—137.

Oei S. L., Babich V. S., Kazakov V. I., Usmanova N. M., Kropotov A. V., Tomilin N. V. 2004. Clusters of regulatory signals for RNA polymerase II transcription associated with Alu family repeats and CpG islands in human promoters. *Genomics.* 83 : 873—882.

Rougeulle C., Chaumeil J., Sarma K., Allis C. D., Reinberg D., Avner P., Heard E. 2004. Differential histone H3 Lys-9 and Lys-27 methylation profiles on the X chromosome. *Mol. Cell Biol.* 24 : 5475—5484.

Satijn D. P., Hamer K. M., den Blaauwen J., Otte A. P. 2001. The polycomb group protein EED interacts with YY1, and both proteins induce neural tissue in *Xenopus* embryos. *Mol. Cell Biol.* 21 : 1360—1369.

Smith K. P., Byron M., Clemson C. M., Lawrence J. B. 2004. Ubiquitinated proteins including uH2A on the human and mouse inactive X chromosome: enrichment in gene rich bands. *Chromosome.* 113 : 324—335.

Поступила 20 IX 2007

#### SINEs IN MAMMALIAN GENOMES CAN SERVE AS ADDITIONAL SIGNALS INFORMATION OF FACULTATIVE HETEROCHROMATIN

N. M. Usmanova, V. I. Kazakov, N. V. Tomilin

Institute of Cytology RAS, St. Petersburg;  
e-mail: nmusmanova@gmail.com

Using computer-based methods we determined the global distribution of short interspersed nuclear elements (SINEs) in the human and mouse X chromosomes. It has been shown that this distributions is similar to the distributions of CpG islands and genes but is different from the distribution of *LINE1* elements. Since SINEs (human Alu and mouse B2) may have binding sites for Polycomb protein YY1, we suggest that these repeats can serve as additional signals («boosters») in Polycomb-dependent silencing of gene rich segments during X inactivation.