ОРГАНИЗАЦИЯ И ЭВОЛЮЦИЯ СУБТЕЛОМЕРНЫХ РАЙОНОВ ХРОМОСОМ РЖИ

© *Е. В. Евтушенко*,[1] *Е. А. Елисафенко*,[2] *А. В. Вершинин*[1,*]

[1]*Институт молекулярной и клеточной биологии СО РАН*
и [2]*Институт цитологии и генетики СО РАН, Новосибирск;*
**электронный адрес: avershin@mcb.nsc.ru*

Субтеломерные районы хромосом относятся к наиболее пластичным и быстро эволюционирующим районам эукариотических геномов. Отличительной чертой хромосом ржи (*Secale*) является наличие крупных блоков субтеломерного гетерохроматина на плечах всех 7 пар хромосом. Различия между размером генома культивируемой ржи *Secale cereale* и древним видом *S. silvestre* достигают 15 %, и основный вклад в эти различия вносит увеличение размеров субтеломерных гетерохроматиновых районов. Ранее было показано, что субтеломерный гетерохроматин ржи насыщен множеством копий нескольких семейств тандемно организованных последовательностей ДНК, образующих длинные тяжи мономеров. В настоящей работе мы изучали тонкую молекулярную организацию и взаимное расположение семейств тандемных повторов и окружающей их геномной, нетандемной. Рестрикционный анализ BAC-клонов, содержащих ДНК короткого плеча первой хромосомы ржи, 1RS, показал, что в тяжах основная масса мономеров организована специфическим образом в структуры более высокого порядка — от димеров до пентамеров и даже гексамеров, которые образуются в центральных участках тяжей. Окружающая тяжи тандемных мономеров геномная ДНК представлена копиями семейств различных классов ретротранспозонов, присутствующих также в геномах пшеницы и ячменя. Таким образом, специфическим молекулярным процессом, происходившим при формировании субтеломерного гетерохроматина в процессе дивергенции генома ржи от общего предка злаков, следует рассматривать взрывную амплификацию мономеров семейств тандемных повторов pSc200 и pSc250, приведшую к образованию длинных многочисленных тяжей среди массы различных копий ретротранспозонов. Логично предполагать, что последствием этого процесса стало увеличение размеров субтеломерного гетерохроматина и генома культурной ржи *S. cereale*.

Ключевые слова: тандемно повторяющиеся последовательности ДНК, BAC-библиотека, ретротранспозоны, субтеломерный гетерохроматин, молекулярная организация и эволюция.

Принятые сокращения: т. п. н. — тысячи пар нуклеотидов, BAC — bacterial artificial chromosome, LTR — long terminal repeat.

Субтеломерные районы хромосом, примыкающие непосредственно к теломерам и продолжающиеся до дистально расположенных хромосомоспецифичных последовательностей ДНК, являются одними из наиболее пластичных и быстро эволюционирующих районов эукариотических геномов. Организация субтеломерных районов у разных организмов может иметь свои особенности, однако общей чертой субтеломер большинства эукариотических хромосом является присутствие длинных массивов повторенных последовательностей ДНК, включая микросателлиты, блоки тандемных повторов и производные различных классов мобильных элементов (Levis et al., 1993; Vershinin et al., 1995). Для наиболее изученных субтеломерных районов хромосом человека показано наличие сегментных дупликаций протяженностью от 1 до более чем 200 т. п. н. в зависимости от хромосомы (Riethman et al., 2004). Одним из основных механизмов появления сегментных дупликаций в субтеломерах человека, видимо, являются хромосомные транслокации, возникающие в результате негомологичного обмена на концах хромосом (NHEJ — nonhomologous end joining) в ходе процесса репарации и последующей гомологичной рекомбинации между дуплицированными сегментами различных хромосом (Linardopoulou et al ., 2005). Вследствие интенсивных рекомбинационных событий субтеломерные районы рассматриваются как наиболее эволюционно динамичные (Mefford, Trask, 2002).

У растений субтеломерные районы подробно исследованы у видов, имеющих небольшие, практически полностью секвенированные геномы. К ним относятся арабидопсис и рис. Субтеломерные районы хромосом 2 и 4 *Arabidopsis thaliana* содержат тандемно организованные гены рибосомальной ДНК (*NOR2* и *NOR4*), район *NOR4* непосредственно примыкает к теломерному повтору (Copenhaver, Pikaard, 1996). Остальные восемь плеч хромосом *A. thaliana* имеют небольшие субтеломерные районы протяженностью до 5 т. п. н., в которых отсутствуют повторяющаяся ДНК и транспозоны (Copenhaver Pikaard, 1996; Heacock et al., 2004). В отличие от арабидопсиса субтеломерные районы хромосом риса содержат несколько семейств повторяющейся ДНК. Из 24 концов хромосом вида риса *Oruza sativa ssp. japonica* 5 плеч имеют

массивы тандемно организованной ДНК, в которых повтор TrsA составляет кластеры из 3—106 копий в зависимости от хромосомы, а между массивами TrsA могут находиться районы, содержащие гены (Mizuno et al., 2008). Семейства тандемно организованных последовательностей в субтеломерных районах описаны у табака, ячменя, кукурузы, томата, пшеницы и картофеля (Sharma, Raina, 2005; Torres et al., 2011), но детальная структура субтеломер этих видов изучена значительно меньше.

Геном ржи *Secale cereale* относится к одним из самых больших геномов растений — его размер составляет $8.3 \cdot 10^9$ п. н., тогда как средний размер генома цветковых растений составляет $5.6 \cdot 10^9$ п. н. (Rabinowicz, Bennetzen, 2006). Отличительной чертой хромосом ржи является наличие крупных блоков субтеломерного гетерохроматина на плечах всех семи пар хромосом, что нехарактерно для хромосом пшеницы и ячменя — видов, так же как и рожь, относящихся к трибе Triticeae. Размеры геномов отдельных видов в роде *Secale* колеблются: различия между размером генома культивируемой ржи *S. cereale* и древним видом *S. silvestre* достигают 15 %, и основный вклад в эти различия вносит увеличение размеров субтеломерных гетерохроматиновых районов (Bennett et al., 1977). Доля их у *S. cereale* составляет около 12 % всего генома и в 2 раза превышает размер этих районов у *S. silvestre* (рис. 1).

Ранее было показано, что субтеломерный гетерохроматин ржи насыщен множеством копий тандемно организованных последовательностей ДНК, принадлежащих к нескольким семействам — pSc119.2, pSc200 и pSc250 (McIntyre et al., 1990; Vershinin et al., 1995). Мономеры этих семейств различаются по размеру и составляют 118, 379 и 571 п. н. соответственно. Создание BAC-библиотеки, содержащей ДНК, выделенную из короткого плеча первой хромосомы, 1RS (Simkova et al., 2008), дало нам возможность изучить тонкую молекулярную организацию, взаимное расположение семейств тандемных повторов и окружающей их геномной, нетандемной ДНК и представить молекулярные процессы, происходившие в субтеломерных районах хромосом ржи в процессе эволюции.

Гидролиз редкощепящей рестриктазой *Bst*X1 высокомолекулярной ДНК транслокационной линии пшеницы 1BL. 1RS и последующий пульс-электрофорез и гибридизация на меченую пробу pSc250 выявили несколько фрагментов гибридизации с мол. массой от 38 до 300 т. п. н. Это указывает на наличие нескольких протяженных тяжей тандемно организованных мономеров pSc250 в 1RS. Рестрикционный анализ BAC-клонов, содержащих тяжи мономеров pSc250, показал, что основная масса мономеров организована в структуры более высокого порядка — от димеров до пентамеров и даже гексамеров (рис. 2, клоны 17C17 и 19H3). Такие структуры характерны не для всех тяжей: например, они отсутствуют в ДНК BAC-клона 122F3 (рис. 2, *а*, *б*). Последующее секвенирование BAC-клонов показало, что клон 122F3 содержит протяженный тяж мономеров pSc250 длиной 57 т. п. н., фланкированный с обеих сторон участками нетандемной ДНК с сайтами для рестриктазы *Bst*X1. Таким образом, наличие отдельных тяжей с различной внутренней организацией мономеров pSc250 было подтверждено как пульс-электрофорезом высокомолекулярной ДНК, так и данными секвенирования. Подобная высокая гетерогенность характерна и для внутренней организации тяжей мономеров pSc200 (рис. 2, *в*, *г*).

При скрининге BAC-библиотеки мы идентифицировали пять BAC-клонов, содержащих тяжи двух семейств

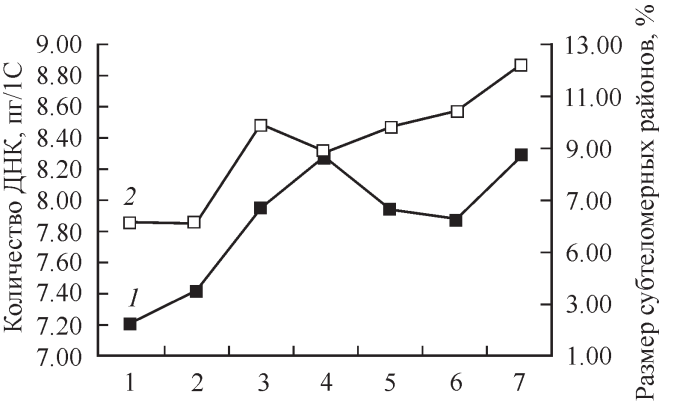

Рис. 1. Динамика изменений размеров генома (*1*) и районов субтеломерного гетерохроматина (*2*) в процессе эволюции рода *Secale*.

По оси абсцисс — виды ржи (*Secale*): 1 — *S. silvestris*, 2 — *S. africanum*, 3 — *S. kuprijanovii*, 4 — *S. montanum*, 5 — *S. ancestrale*, 6 — *S. dighoricum*, 7 — *S. cereale*, сорт Petkus; *по оси ординат*: *слева* — размер гаплоидного генома в пикограммах ДНЕ (*1*), *справа* — размер районов субтеломерного гетерохроматина (*2*), % от общей длины хромосом гаплоидного набора.

тандемных повторов — pSc200 и pSc119.2. Рестрикционное картирование и секвенирование этих клонов указывают на их происхождение из одного района генома, который наиболее полно представлен в BAC 84/C 15 и простирается на 125—130 т. п. н. Остальные 4 клона содержат более короткие вставки, обрывающиеся в различных участках тяжей мономеров pSc200 и pSc119.2. Обработка ферментом рестрикции *Apa*I ДНК клонов, содержащих тяжи pSc200, выявила наличие мультимерных единиц высокого порядка — тетрамеров — наряду с димерами и мономерами в 4 из 5 клонов. Однако тетрамеры отсутствовали в BAC-клона 130/H7 с самым коротким участком тяжа pSc200, не более 5 т. п. н. от левой границы. Аналогичная тенденция характерна и для тяжей мономеров семейства pSc119.2, а именно: структуры высокого порядка присутствуют только в протяженных тяжах. Эти данные указывают на то, что мультимерные структуры высокого порядка образуются в центральной части тандемных тяжей, тогда как вблизи границ тяж представлен мономерами и димерами.

Поиск гомологий в различных базах данных выявил разнообразные пути возникновения мономеров семейств, локализованных в субтеломерном гетерохроматине. Наиболее очевидным является происхождение *Xba*I — семейства, описанного нами ранее (Евтушенко и др., 2010). Мономер этого семейства имеет длину 576 п. н., и 475 из них показывает 82 % гомологии к известному LTR-ретротранспозону ячменя Cereba. В остальной части *Xba*I-мономера присутствует участок длиной 37 п. н., имеющий 79 % гомологии к пшеничному Gypsy-подобному ретротранспозону Quinta. С другой стороны, мономер семейства pSc250, имеющий почти такой же размер (571 п. н.), состоит из множества коротких сегментов с гомологией к различным классам ретротранспозонов. Четыре наиболее длинных участка гомологии имеют размер 51 п. н. и максимум (85 %) гомологии к последовательностям мобильных элементов. Подобную структуру имеет и мономер pSc200. Принципиальным отличием является момент возникновения этих мономеров в процессе эволюции. Если семейство *Xba*I обнаружено в геномах большинства видов злаков, то pSc250 появилось только в процессе дивер-

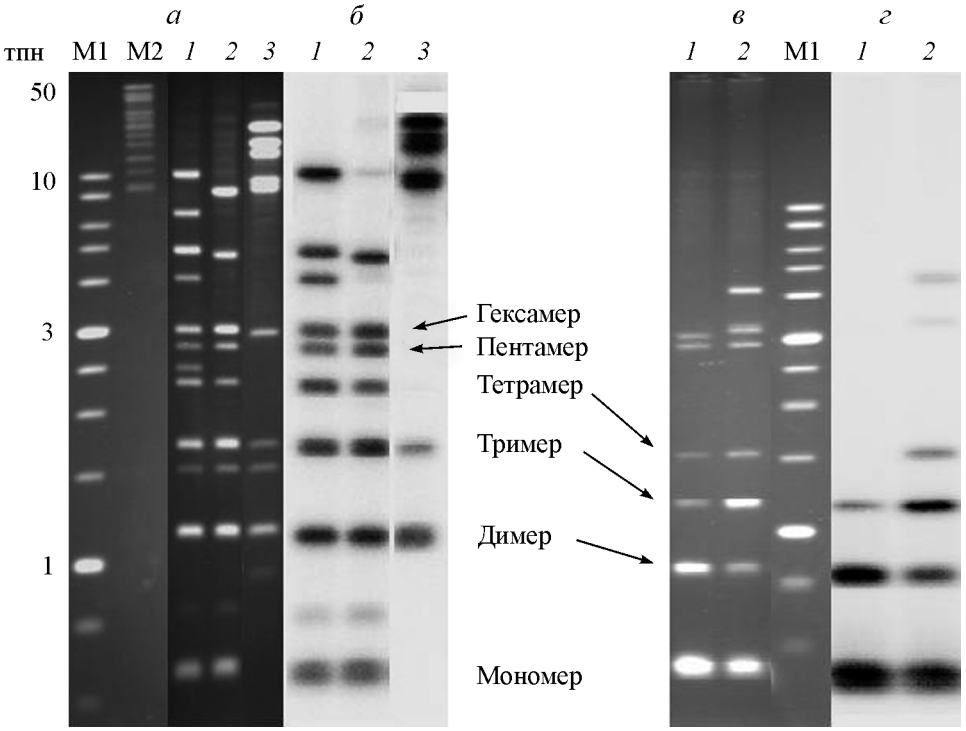

Рис. 2. Гетерогенность внутренней организации тяжей тандемно организованных повторов ржи.

Пульс-электрофорез ДНК BAC-клонов после гидролиза рестриктазой *Pst*1: *а* — окраска бромистым этидием, BAC-клоны 17C17 (1), 19H13 (2) и 122F3 (3); *б* — фрагменты того же геля после переноса на нитроцеллюлозный фильтр и гибридизации на пробу мономера pSc250; *в* — окраска бромистым этидием, BAC-клоны 119C15 (*дорожка 1*) и 119M22 (*дорожка 2*); *г* — фрагменты того же геля после гибридизации на пробу мономера pSc200, M1 и M2 — маркеры размеров ДНК (1kb — от фирмы «СибЭнзим», 8—48 kb — от «BioRad»).

генции генома ржи от общего предка злаков (Vershinin et al., 1996; Kishii et al., 2001).

Анализ первичной структуры, окружающей тандемные тяжи мономеров геномной, нетандемной ДНК, показал, что в исследованных BAC-клонах она полностью представлена различными классами ретротранспозонов. Последовательностей другой молекулярной природы не выявлено. Как правило, участки гомологии к ретротранспозонам имеют большую протяженность. Гомология к исходным элементам, обнаруженным в геномах ближайших родственников ржи — пшениц *Triticum monococcum* и *T. aestivum*, а также ячменя *H. vulgare*, варьирует от 57 до 90 %. Всего в 10 BAC-клонах выявлены 4 элемента класса Gypsy, 5 элементов класса Copia, 1 Solo-LTR и 1 LINE-элемент.

Суммируя представленные результаты, можно сделать вывод, что в составе субтеломерного гетерохроматина 1RS присутствует по несколько тяжей тандемно организованных мономеров семейств pSc200 и pSc250. Эти тяжи характеризуются гетерогенностью внутренней организации, со специфической для каждого тяжа организацией мономеров в структуры более высокого порядка. Мономеры разных семейств имеют различные пути происхождения: это могут быть амплифицированная часть LTR-ретротранспозона (мономер семейства *Xba*I) и амплифицированная последовательность, состоящая из коротких сегментов, относящихся к различным мобильным элементам (мономеры семейств pSc200 и pSc250). Окружающая тяжи тандемных мономеров геномная ДНК представлена копиями семейств различных классов ретротранспозонов, присутствующих также в геномах пшеницы и ячменя. Таким образом, специфическим молекулярным процессом, происходившим при формировании субтеломерного гетерохроматина в процессе дивергенции генома ржи от общего предка злаков, следует считать взрывную амплификацию мономеров pSc200 и pSc250, приведшую к образованию длинных многочисленных тяжей этих семейств. Логично предполагать, что последствием этого процесса стало увеличение размера субтеломерного гетерохроматина у культурной ржи *S. cereale*, и соответствующее увеличение размера генома этого вида.

Авторы выражают глубокую благодарность докторам J. Dolezel и J. Safar (Институт экспериментальной ботаники, г. Оломоуц, Чехия) за предоставленную BAC-библиотеку 1RS, С. А. Мироновой и А. И. Белоусову за участие в части экспериментов.

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (проект 12-04-00512) и программы интеграционных проектов фундаментальных исследований СО РАН (проект 51).

Список литературы

Евтушенко Е. В., *Елисафенко Е. А.*, *Вершинин А. В. 2010.* Взаимное расположение двух семейств тандемных повторов в гетерохроматине ржи. Молекуляр. биол. 44 (1) : 3—10.

Bennett M. D., *Gustafson J. P.*, *Smith J. B. 1977*. Variation in nuclear DNA in the genus *Secale*. Chromosoma. 62 : 149—176.

Copenhaver G. P., *Picaard C. S. 1996.* RFLP and physical mapping with an rDNA-specific endonuclease reveals that nucleolus organizer regions of *Arabidopsis thaliana* adjoin the telomeres on chromosomes 2 and 4. Plant J. 9 : 259—272.

Heacock M., *Spangler K.*, *Puizina R. J.*, *Shippen D. E. 2004*. Molecular analysis of telomere fusions in *Arabidopsis*: multiple pathways for chromosome end-joining. EMBO J. 23 : 2304—2313.

Kishii M., *Nagaki K.*, *Tsujimoto H. 2001.* A tandem repetitive sequence located in the centromeric region of common wheat (*Triticum aestivum*) chromosomes. Chromosome Res. 9 : 417—428.

Levis R. V., *Ganesan R.*, *Houtchens K.*, *Tolar A.*, *Sheen F. M. 1993.* Transposons in place of telomeric repeats at a *Drosophila* telomere. Cell. 75 : 1083—1093.

Linardopoulou E. V., *Williams E. M.*, *Fan Y.*, *Friedman C.*, *Young G. M.*, *Trask B. J. 2005.* Human subtelomeres are hot spots of interchromosomal recombination and segmental duplication. Nature. 437 : 94—100.

McIntyre C. L., *Pereira S.*, *Moran L. B.*, *Appels R. 1990.* New *Secale cereale* (rye) DNA derivatives for the detection of rye chromosome segments in wheat. Genome. 33 :317—323.

Mefford H. C., *Trask B. J. 2002*. The complex structure and dynamic evolution of human subtelomeres. Nat. Rev. Genet. 3 : 91—102.

Mizuno H., *Wu J.*, *Katayose Y.*, *Kanamori H.*, *Sasaki T.*, *Matsumoto T. 2008.* Characterization of chromosome ends on the basis of the structure of TrsA subtelomeric repeats in rice (*Oryza sativa* L.). Mol. Genet. Genomics. 280 : 19—24.

Rabinowicz P. D., *Bennetzen J. L. 2006.* The maize genome as a model for efficient sequence analysis of large plant genomes. Curr. Opin. Plant Biol. 9 : 149—156.

Riethman H., *Ambrosini A.*, *Castaneda C.*, *Finklestein J.*, *Hu X.-L.*, *Mudunuri U.*, *Wei J. 2004.* Mapping and initial analysis of human subtelomeric sequence assemblies. Genome Res. 14 : 18—28.

Sharma S., *Raina S. N. 2005*. Organization and evolution of highly repeated satellite DNA sequences in plant chromosomes. Cytogenet. Genome Res. 109 :15—26.

Simkova H., *Safar J.*, *Suchankova P.*, *Kovarova P.*, *Bartos J.*, *Kubalakova M.*, *Janda J.*, *Cihalikova J.*, *Mago R.*, *Lelley T.*, *Dolezel J. 2008*. A novel resource for genomics of Triticeae: BAC library specific for the short arm of rye (*Secale cereale* L.) chromosome 1R(1RS). BMC Genomics. 9 : 237.

Torres G. A., *Gong Z.*, *Iovene M.*, *Hirsch C. D.*, *Buell C. R.*, *Bryan G. J.*, *Novák P.*, *Macas J.*, *Jiang J. 2011.* Organization and evolution of subtelomeric satellite repeats in the potato genome. Genes Genomes Genet. 1 : 85—92.

Vershinin A. V., *Alkhimova E. G.*, *Heslop-Harrison J. S. 1996*. Molecular organization of tandemly organized DNA sequences and heterochromatic chromosome regions in some Triticeae species. Chromosome Res. 4 : 517—525.

Vershinin A. V., *Schwarzacher T.*, *Heslop-Harrison J. S. 1995.* The large-scale organization of repetitive DNA families at the telomeres of rye chromosomes. Plant Cell. 7 : 1823—1833.

Поступила 26 XI 2012

ORGANIZATION AND EVOLUTION OF THE SUBTELOMERIC REGIONS OF THE RYE CHROMOSOMES

E. V. Evtushenko,[1] *E. A. Elisafenko*,[2] *A. V. Vershinin*[1, *]

[1] Institute of Molecular and Cellular Biology SB RAS and [2] Institute of Cytology and Genetics SB RAS, Novosibirsk;
* e-mail: avershin@mcb.nsc.ru

Subtelomeric regions of chromosomes are particularly dynamic and variable parts in the evolution of the eukaryotic genomes. A specific feature of the rye (*Secale*) chromosomes is large heterochromatin blocks at the arms of all seven pairs of chromosomes. Within the genus *Secale*, an interspecific variation in the genome size reaches nearly 15 % between *Secale cereale* (cultivated rye) and the ancient species *S. silvestre* and an increase in the size of subtelomeric heterochromatin regions is the main contributor to these differences. Earlier works have demonstrated that the subtelomeric hetechromatin of rye is enriched for a few multicopy tandemly organized DNA families which form the long arrays of monomers. Here we aimed to clarify the fine large-scale organization and mutual arrangement within the tandem arrays of these families and the flanking genomic nonarray DNA. Restriction analysis of the BAC-clones containing genetic material of the short arm of the first rye chromosome (1RS) showed that within arrays the bulk of monomers is organized in the specific higher-order repeat units (up to hexamers) which are generated in the central part of tandem arrays, while only monomers and dimers are present near the boundaries. Sequencing of the genomic nonarray DNA flanking the tandem arrays has demonstrated that this DNA in the studied BAC-clones consists completely of retrotransposons of various classes which are also present in the wheat and barley genomes. Thus, only an explosive amplificcation of pSc200 and pSc250 monomers, on the background of a saturated mixture of various retrotransposons, can be regarded as a specific molecular process in formation of subtelomeric hetechromatin during the divergence of rye genome from the common ancestor of cereals. Evidently, this process resulted in the enlargement of subtelomeric hetechromatin of *S. cereale* and an increasing of its genome size.

Key words: tandemly organized repetitive DNA sequences, BAC library, retrotransposons, subtelomeric hetechromatin, molecular organization and evolution.