

ГЕНОМИКА.  
ТРАНСКРИПТОМИКА

УДК 575.2:582.542.1

ВКЛАД Ty3-gypsy-РЕТРОТРАНСПОЗОНА *Lila* В СПЕЦИФИЧНОСТЬ  
D-ГЕНОМА МЯГКОЙ ПШЕНИЦЫ *Triticum aestivum* L.

© 2012 г. А. Б. Щербань\*, И. Г. Адонина, Е. А. Салина

Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, 630090

Поступила в редакцию 03.08.2011 г.

Принята к печати 07.09.2011 г.

Проведен анализ первичной структуры ВАС-клона 112D20 *Triticum aestivum*, содержащего специфичный для D-генома Ty3-gypsy-ретротранспозон *Lila*. ПЦР-анализ нулли-тетрасомных и делеционных линий *T. aestivum* позволил локализовать этот ВАС-клон в дистальном районе длинного плеча хромосомы 5D. Для ВАС-клона 112D20 характерно высокое содержание Ty3-gypsy-ретротранспозонов (61.7%) и низкое содержание генов (1.2%). Выявлена только одна открытая рамка считывания, гомологичная гену *Aegilops tauschii* с неизвестной функцией. Специфичный для D-генома Ty3-gypsy-ретротранспозон *Lila* в составе ВАС-клона 112D20 имеет длину 14 т.п.н. и ограничен неравными по размеру длинными концевыми повторами. Данные *in situ* гибридизации и ПЦР-анализа различных видов Triticeae свидетельствуют о том, что указанный ретроэлемент подвергся амплификации внутри предкового вида *Ae. tauschii*, донора D-генома. Предполагаемое время амплификации по оценке времени инсерции *Lila* 112D20 составляет 1.7 млн лет, что соответствует периоду до образования первых аллополиплоидных форм пшеницы. Сопоставление с ранее полученными данными позволяет сделать вывод о том, что амплификация специфичных для каждого генома пшеницы ретроэлементов происходила на стадии формирования диплоидных видов — предшественников этих геномов.

**Ключевые слова:** аллополиплоид, геном, Triticeae, ВАС-клон, ретротранспозон.

THE IMPACT OF Ty3-gypsy GROUP RETROTRANSPOSON *Lila* ON D-GENOME SPECIFICITY OF WHEAT *Triticum aestivum* L., by A. B. Shcherban<sup>1</sup>, I. G. Adonina<sup>1</sup>, E. A. Salina<sup>1</sup> (<sup>1</sup>Institute of Cytology and Genetics, Siberian Division, Russian Academy of Sciences, Novosibirsk, 630090 Russia; \*e-mail: atos@bionet.nsc.ru). An analysis of the primary structure of BAC clone 112D20 *T. aestivum*, that contains D-genome specific Ty3-gypsy-retrotransposon *Lila* is presented. PCR analysis of nulli-tetrasomic and deletion lines of *T. aestivum* allowed to localize this BAC clone in the distal region of the long arm of chromosome 5D. Characteristic feature of BAC clone 112D20 is a high concentration of Ty3-gypsy-retrotransposons (61.7%), and low content of the genes (1.2%). Only a single open reading frame was revealed homologous to an unknown gene of *Ae. tauschii*. Specific to the D-genome Ty3-gypsy-retrotransposon *Lila* in the BAC clone 112D20 is 14 kb in length and contains unequal in size long terminal repeats. The data of *in situ* hybridization and PCR analysis of different Triticeae species suggest that this retroelement was amplified within the ancestral species of *Ae. tauschii*, the donor D-genome. The suggested time of amplification based on estimation of insertion time of *Lila* 112D20 is 1.7 million years, which corresponds to the formation of the first allopolyploid forms of wheat. Based on comparison with the previously obtained data, it is concluded that the amplification of retroelements specific to each genome of wheat took place during formation of the diploid progenitors of these genomes.

**Keywords:** allopolyploid, genome, Triticeae, BAC-clone, retrotransposon.

Изучение геномного состава и происхождения аллополиплоидных геномов имеет важное значение для молекулярной генетики растений, так как более 70% растений возникли в результате естественной или искусственной гибридизации. Большинство видов рода *Triticum* — аллополиплоидные, в том числе мягкая, или гексаплоидная пшеница *Triticum aestivum* L. ( $2n = 6x = 42$ , геном ВВААDD), возникшая в результате гибридизации

тетраплоидного предка (ВВАА) с диплоидным донором D-генома [1]. В качестве наиболее вероятных предшественников А- и В-геномов рассматриваются виды *T. monococcum*/*T. urartu* и виды рода *Aegilops* (секция Sitopsis) соответственно, тогда как предполагаемым донором D-генома является *Aegilops tauschii* [2]. Использование молекулярных маркеров позволило значительно повысить эффективность анализа генома на различных уровнях его организации: от отдельных геномов в составе полиплоидных видов до инди-

\* Эл. почта: atos@bionet.nsc.ru

видуальных хромосом и их участков (генетических локусов). Большим достижением стала разработка методов флуоресцентной гибридизации *in situ* (FISH), особенно метода геномной гибридизации (GISH), с помощью которого успешно изучали аллополиплоиды, замещенные линии, а также линии, содержащие транслокации [3–5]. Все это позволило сделать вывод, что наибольший вклад в межгеномные различия вносят повторяющиеся последовательности, в частности, мобильные генетические элементы (МЭ), которые составляют значительную часть генома растений (70–80%).

Разнообразие МЭ определяется большим числом структурных классов и семейств, имеющих различную копийность [6]. Наиболее представительный класс МЭ растений составляют ретротранспозоны. Они кодируют обратную транскриптазу (RT) и интегразу (INT), необходимые для обратной транскрипции РНК-интермедиата и последующего встраивания ДНК-копии МЭ. Благодаря этому механизму ретротранспозоны широко распространены в геномах растений, где занимают от 15% у *Arabidopsis thaliana* до 90% генома у некоторых представителей Liliaceae [7–9]. У мягкой пшеницы *T. aestivum* ретротранспозоны занимают более 60% генома [10]. Класс ретротранспозонов включает две основные группы, выделяемые по наличию–отсутствию длинных концевых повторов (LTR) [6]. У растений преобладают LTR-содержащие ретротранспозоны (LTR-ретротранспозоны), которые делятся на *Ty1-copia*- и *Ty3-gypsy*-подобные в зависимости от расположения RT- и INT-доменов относительно друг друга. LTR-ретротранспозоны вносят существенный вклад в вариабельность размера, а также в структурную дивергенцию генома у цветковых растений [11–13]. Изучение геномов злаков молекулярно-генетическими методами выявило насыщенность межгенных районов LTR-ретротранспозонами [14]. Во многих случаях они формируют так называемые “вложенные инсерции”, т.е. инсерции во встроены ранее элемент [15]. В некоторых районах генома наблюдается кластеризация ретротранспозонов, обусловленная их направленным встраиванием в определенные сайты. Подобную “колонизацию” хромосомных районов специфическими семействами МЭ можно обнаружить при проведении полного или частичного анализа первичной структуры ВАС-клонов, содержащих протяженные фрагменты геномной ДНК, с последующей FISH отдельных субклонированных нуклеотидных последовательностей. Так, например, *Ty3-gypsy*-подобный ретротранспозон *Romani*, выделенный из ВАС-клона *Ae. tauschii*, локализован в дискретных районах пяти хромосом D-генома мягкой пшеницы [16]. Преимущественная локализация САСТА транспозона *Caspar* в субтеломерных районах хромосом

видов Triticeae выявлена в результате анализа соответствующего ВАС-клона [17].

Недавно мы проанализировали первичную структуру ВАС-клона *T. aestivum*, специфичного для В-генома, и с помощью FISH установили, что эта специфичность связана с *Ty3-gypsy*-подобным ретротранспозоном *Fatima*, диспергированным по всем хромосомам указанного генома [18]. Наряду с этим ВАС-клоном мы идентифицировали другой геном-специфичный ВАС-клон (112D20), также показавший дисперсный характер гибридизации с хромосомами одного из геномов мягкой пшеницы [17].

В представленной работе перед нами стояли следующие задачи:

- 1) определить геномную и хромосомную локализацию ВАС-клона 112D20;
- 2) проанализировать первичную структуру этого ВАС-клона и выявить разные типы повторяющейся ДНК, уникальные гены; провести поиск гомологий с известными нуклеотидными последовательностями растений;
- 3) выявить нуклеотидную последовательность, отвечающую за геном-специфичный характер гибридизации ВАС-клона 112D20, а также изучить пути и механизмы ее распространения в ходе эволюции полиплоидного генома пшеницы.

## ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

ВАС-клон 112D20 из геномной ВАС-клонотеки *T. aestivum* сорта “Renan” отбирали согласно [17].

**Субклонирование фрагментов ДНК ВАС-клона 112D20** с целью анализа его первичной структуры проводили с помощью неполного гидролиза ДНК эндонуклеазой рестрикции *Sau3A* и методов случайного фрагментирования [19, 20]. В последнем случае использовали два подхода: обработку ДНК ВАС-клона ДНКазой I и фрагментирование ДНК с помощью небулайзера. Фрагментированную ДНК, очищенную с помощью набора “Qiagen”, лигировали с 50 мкг плазмидной ДНК. Для трансформации методом электропорации использовали высококомпетентные клетки *Escherichia coli* XL-blue MRF+. Рекомбинатные колонии отбирали на агаризованной LB-среде с ампициллином и Xgal/IPTG. Присутствие вставок проверяли с помощью ПЦР колоний. Клоны, содержащие вставки длиной более 200 п.н., отбирали для дальнейшего анализа.

**Анализ нуклеотидной последовательности ДНК.** Субклонированные нуклеотидные последовательности секвенировали в одном направлении, если длина вставки не превышала 1 т.п.н., и в обоих направлениях – в остальных случаях. Использовали универсальные праймеры pUC18/M13 и ABI PRISM Dye Terminator Cycle Sequencing ready reaction kit (“Perkin Elmer”). Продукты реакций

секвенирования анализировали с помощью ABI PRISM 310 Genetic Analyzer (“Perkin Elmer”). Для выравнивания нуклеотидных последовательностей и построения контига ВАС 112D20 использовали программу DNA Dragon Sequence Assembly Software (<http://www.sequentix.de>). Поиск гомологии с известными нуклеотидными последовательностями проводили с использованием программы BLASTN (<http://blast.ncbi.nlm.nih.gov>). С целью выявления повторяющихся последовательностей в составе ВАС-клона 112D20 с помощью алгоритма BLAST проведен скрининг базы данных TREP (<http://wheat.pw.usda.gov>). При поиске генов использовали программу GenScan (<http://argonaute.mit.edu/GENSCAN.html>). Полная нуклеотидная последовательность ВАС-клона 112D20 длиной 76 327 п.н. аннотирована согласно правилам аннотирования геномных последовательностей пшеницы, разработанным Международным Консорциумом по Секвенированию Генома Пшеницы (<http://www.wheatgenome.org/index.php>), и помещена в GenBank под регистрационным номером: JN544193.

**Растительный материал.** В работе использовали виды из коллекции Института цитологии и генетики СО РАН (Новосибирск, Россия): *T. urartu* Tum. (геном A<sup>u</sup>A<sup>u</sup>) TMU06; *Ae. speltoides* Tausch. (геном SS) TS01, TS86, K389; *Ae. umbellulata* Zhuk. (геном UU) TU04; *Ae. longissima* Schweinf. et Muschl. (геном S<sup>S</sup>) TL05; *Ae. tauschii* Coss. (геном DD) TQ27; *T. dicoccoides* Sweinf. (геном BBA<sup>u</sup>A<sup>u</sup>) TTD20; *T. timopheevi* Zhuk. (геном GGA<sup>b</sup>A<sup>b</sup>) K-38555; *T. aestivum* L. (геном BBAADD) сорта “Chinese Spring”, “Renan” и Саратовская 29. Нулли-тетрасомные и делеционные линии мягкой пшеницы *T. aestivum* на основе сорта “Chinese Spring” любезно предоставил В. Gill (University of Kansas, USA).

**ПЦР-анализ.** Для амплификации районов, фланкирующих сайты встраивания ретротранспозона *Lila*, сконструированы специфические праймеры: J1forward/J1reverse (J1F, 5'-gcgtaccagacsaacgagcc-3' и J1R, 5'-tttgcttcttcaagaggg-3') и J2forward/J2reverse (J2F, 5'-cttatcccgttccaggcac-3' и J2R, 5'-cgagcaattgcagtccaac-3'). ПЦР-смесь (25 мкл) содержала 100 нг геномной ДНК, 67 мМ Трис-НСl (рН 8.7), 18 мМ (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 2 мМ MgCl<sub>2</sub>, 0.01% Твин-20, по 0.25 мМ каждого dNTP, по 0.5 мкМ каждого праймера и 1 ед. акт. Таq-полимеразы. Условия ПЦР: денатурация – 1 мин при 94°C; отжиг – 1 мин при 60°C; полимеризация – 30 с при 72°C; число циклов – 35; достраивание ПЦР-фрагментов – 15 мин при 72°C. Продукты ПЦР анализировали с помощью электрофореза в 1–1.5%-м агарозном геле.

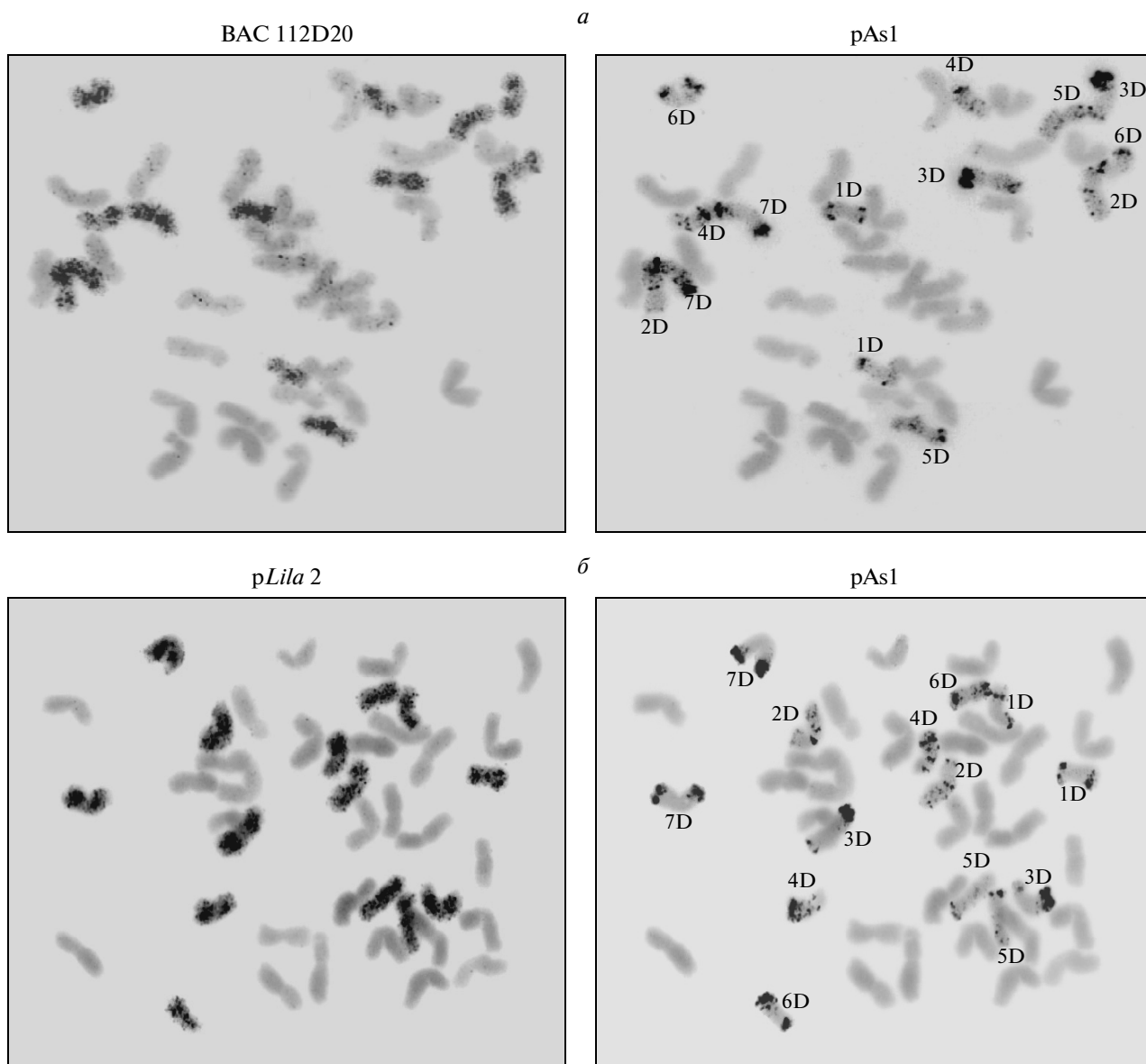
**Флуоресцентная гибридизация *in situ* (FISH).** В качестве зонда для FISH использовали ДНК ВАС-клона 112D20, а также различные субклони-

рованные фрагменты данного клона, меченные дигоксигенин-11-dUTP (“Roche Applied Science”). Одновременно в гибридизационный раствор добавляли клонированную субтеломерную последовательность rAs1, меченную биотин-11-dUTP (“Roche Applied Science”). Эта последовательность специфична для D-генома мягкой пшеницы и позволяет идентифицировать хромосомы этого генома [21]. Мечение обоих зондов проводили с помощью метода ник-трансляции согласно инструкции производителя. Получение препаратов метафазных хромосом, гибридизацию и детекцию зондов на хромосомах осуществляли согласно [22]. Меченный дигоксигенином зонд выявляли с помощью антител к дигоксигенину, связанных с родамином (Anti-digoxigenin-rhodamine Fab fragments, “Roche Applied Science”). Биотинилированный зонд выявляли с использованием флуоресцеин-авидина (Fluorescein Avidin D, “Vector Laboratories”, США). Препараты анализировали микроскопически (“Axioskop” 2 Plus, “Carl Zeiss”). Изображение регистрировалось CCD-камерой VC-44 (PCO) и обрабатывалось компьютерной программой Adobe Photoshop® (“Adobe Systems”, Великобритания).

## РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

### *Анализ геномной локализации ВАС-клона 112D20*

Геномную локализацию ВАС-клона 112D20 определяли методом FISH ВАС-клона и клонированной геном-специфичной ДНК с хромосомами *T. aestivum* сорта “Renan” (см. “Экспериментальную часть”). На рис. 1а представлен результат гибридизации одновременно с двумя зондами – ВАС-клоном 112D20 и rAs 1 в качестве маркера D-генома мягкой пшеницы. Наиболее интенсивный сигнал гибридизации ВАС-клона выявлен на 14 хромосомах, относящихся к D-геному. Аналогичный результат получен и для других сортов мягкой пшеницы: сорта “Chinese Spring” и Саратовская 29 (результат не представлен). Наряду с этим, на одном препарате проведена FISH ВАС-клона 112D20 с диплоидными видами: *T. monococcum* (AA), *Ae. speltoides* (SS) и *Ae. tauschii* (DD). В последнем случае выявлен равномерный сигнал гибридизации на всех хромосомах, тогда как у двух других видов сигнал полностью отсутствовал (результат не представлен). Таким образом, использование в качестве зонда ВАС-клона 112D20 позволяет эффективно идентифицировать хромосомы D-генома. Дисперсный характер гибридизации указывает на присутствие в составе ВАС-клона повторяющейся ДНК (МЭ), получившей широкое распространение в D-геноме.



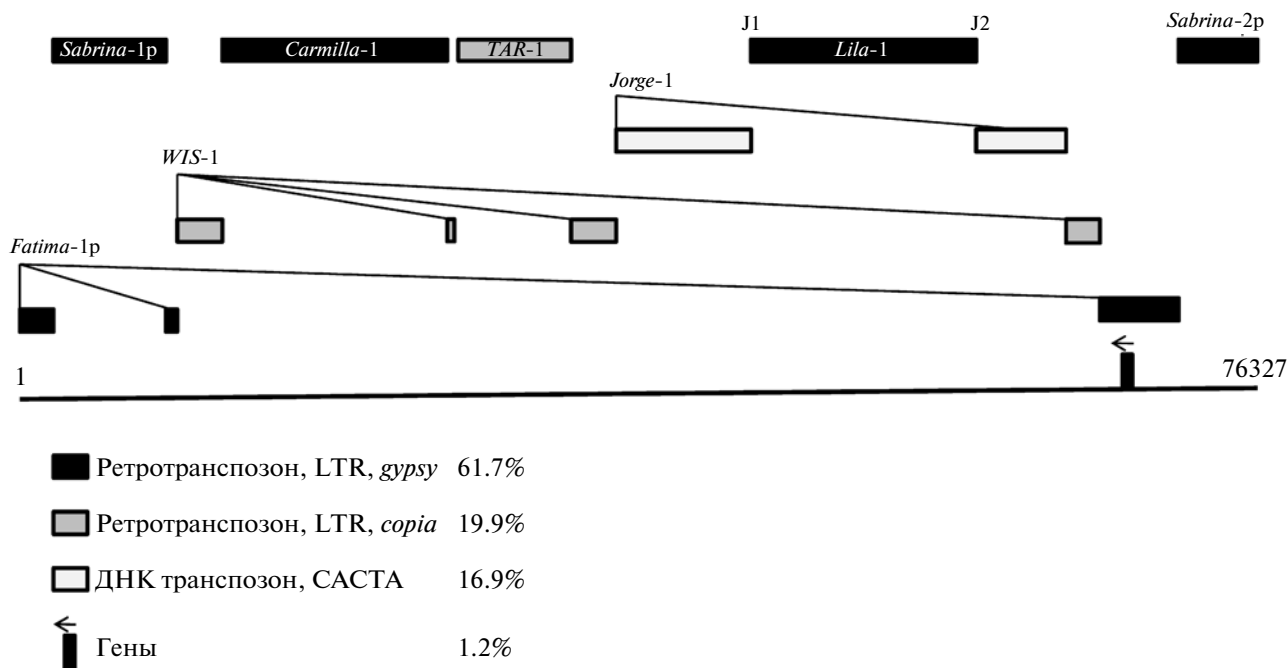
**Рис. 1.** FISH метафазных хромосом *T. aestivum* (сорт “Renan”) с зондами. *a* – BAC-клон 112D20 и pAs1 (отдельно); *б* – pLila2 и pAs1.

### Субклонирование и анализ первичной структуры BAC-клона 112D20

Нами проведен детальный анализ первичной структуры BAC-клона 112D20 (см. “Экспериментальную часть”). Предварительно мы получили библиотеку субклонированных фрагментов ДНК, содержащую 793 клон со вставками размером 0.2–3 т.п.н. С использованием программы BLAST выявлены и исключены из анализа последовательности вектора pIndigoBAC-5 (7.5 т.п.н.) и геномной ДНК *E. coli* (около 7.5% от общего числа субклонов). Оставшиеся субклоны обеспечили 4–6-кратное покрытие вставки BAC-клона 112D20 размером 76327 п.н.

Выявлена высокая насыщенность BAC-клона 112D20 МЭ, причем преобладали среди них LTR-ретротранспозоны (81.6% общей длины BAC-клона, таблица). Наиболее распространенное суперсемейство Ty3-gypsy-подобных элементов занимает 61.7% BAC-клона. Обнаружены “вложенные инсерции” LTR-ретротранспозонов *Sabrina*, *Carmilla*, *Lila*, *Tar* в другие элементы того же класса: *Fatima*, *WIS*, а также в ДНК-транспозон семейства САСТА (рис. 2).

Класс ДНК-транспозонов представлен единственной копией САСТА-элемента *Jorge* 112D20-1, составляющей 16.9% BAC-клона 112D20 (таблица). Границы этого транспозона точно определе-



**Рис. 2.** Структурная организация геномного фрагмента *T. aestivum* длиной 76327 п.н. в составе ВАС-клона 112D20. “р” в названиях мобильных элементов означает укороченный элемент.

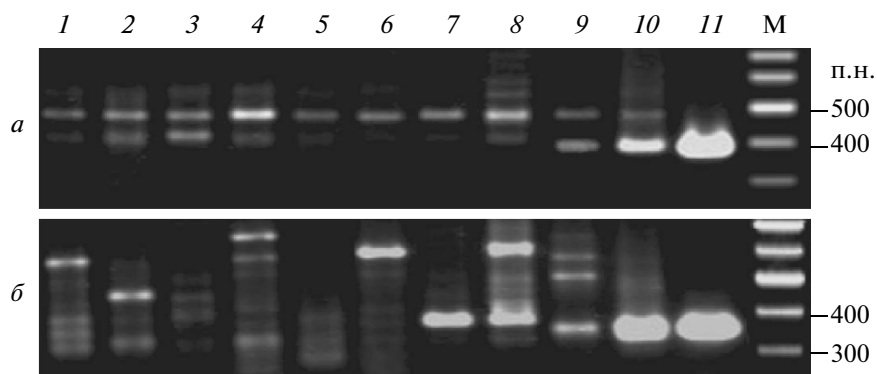
ны по наличию повтора САСТА в инвертированной ориентации, ограниченного дупликацией сайта-мишени САА (рис. 2). Концевые участки *Jorge* 112D20-1 длиной около 400 п.н. имеют 96% гомологии с соответствующими участками элемента *Jorge* AF326781-1 из генома *T. monocosmum*. Внутренняя часть *Jorge* 112D20-1 имеет лишь частичную гомологию с другими САСТА-транспозонами и включает общий для этих элементов участок, кодирующий транспозазу. В централь-

ную часть элемента *Jorge* 112D20-1 встроен *Ty3-gypsy*-ретротранспозон *Lila* 112D20-1 длиной 14 т.п.н. Он ограничен дупликацией сайта-мишени ССТГ и содержит несимметричные LTR. 5'-LTR на 0.4 т.п.н. короче 3'-LTR (1.1 т.п.н.) из-за делеции в центральной части.

В составе ВАС-клона 112D20 выявлена только одна предполагаемая открытая рамка считывания (ОРС), не кодирующая МЭ (таблица, рис. 2). Эта

Элементы, идентифицированные в составе ВАС-клона 112D20 (76327 п.н.)

Класс, суперсемейство, семейство	Число копий	Длина, п.н.	Относительное содержание, %
<b>Класс 1 (LTR-ретротранспозоны)</b>	<b>7</b>	<b>62299</b>	<b>81.6</b>
<i>Gypsy</i>	<b>5</b>	<b>47118</b>	<b>61.7</b>
RLG_ <i>Fatima</i> _112D20-1p	1	7808	
RLG_ <i>Sabrina</i> _112D20-1p, 2p	2	7247, 3647	
RLG_ <i>Carmilla</i> _112D20-1	1	14381	
RLG_ <i>Lila</i> _112D20-1	1	14035	
<i>Copia</i>	<b>2</b>	<b>15181</b>	<b>19.9</b>
RLC_ <i>WIS</i> _112D20-1	1	8509	
RLC_ <i>TAR</i> _112D20-1	1	6672	
<b>Класс 2 (ДНК-транспозоны)</b>	<b>1</b>	<b>12924</b>	<b>16.9</b>
САСТА, DTC_ <i>Jorge</i> _112D20-1			
<b>Гены</b>	<b>1</b>	<b>921</b>	<b>1.2</b>
<b>Неидентифицированные последовательности</b>			<b>0.3</b>



**Рис. 3.** ПЦР-анализ геномной ДНК различных видов Triticeae с использованием специфических праймеров. *a* – к J1-району; *б* – к J2-району. 1 – *T. urartu*; 2 – *Ae. speltooides* TS01; 3 – *Ae. speltooides* K389; 4 – *Ae. speltooides* TS86; 5 – *Ae. umbellulata*; 6 – *T. dicoccoides*; 7 – *Ae. longissima*; 8 – *T. timopheevii*; 9 – *T. aestivum* (“Chinese Spring”); 10 – *Ae. tauschii*; 11 – ВАС 112D20. В качестве маркера длины фрагментов ДНК использован “100 bp DNA ladder”.

ОРС кодирует последовательность из 307 аминокислотных остатков, которая имеет 49% гомологии с продуктом неизвестного гена *Ae. tauschii*, расположенного вблизи гена устойчивости к листовая ржавчине *Lr21* (AF532104.1). Внутри изучаемого ВАС-клона эта ОРС с обеих сторон ограничена *Ty3-gypsy*-элементом *Fatima* 112D20-1р.

#### Поиск последовательности ВАС-клона 112D20, специфичной для D-генома

С целью идентификации нуклеотидной последовательности, специфичной для D-генома, нами проведена сравнительная оценка полиморфизма всех МЭ, входящих в состав ВАС-клона 112D20. Анализ гомологичных последовательностей в банках данных с помощью программы BLASTN показал, что большинство семейств МЭ этого ВАС-клона широко представлены последовательностями из различных геномов мягкой пшеницы и соответствующих видов-предшественников. Анализ полиморфизма этих семейств указывает на отсутствие выраженной дивергенции между геномами и на значительную перекрестную гомологию (82–95%) между представителями того или иного семейства, выделенными из разных геномов. Единственное исключение представляет *Ty3-gypsy*-ретротранспозон *Lila* 112D20-1. С помощью программы BLASTN найдена только одна последовательность (AF488415.1) из генома *T. monosocum*, гомологичная данному элементу (75–80% гомологии).

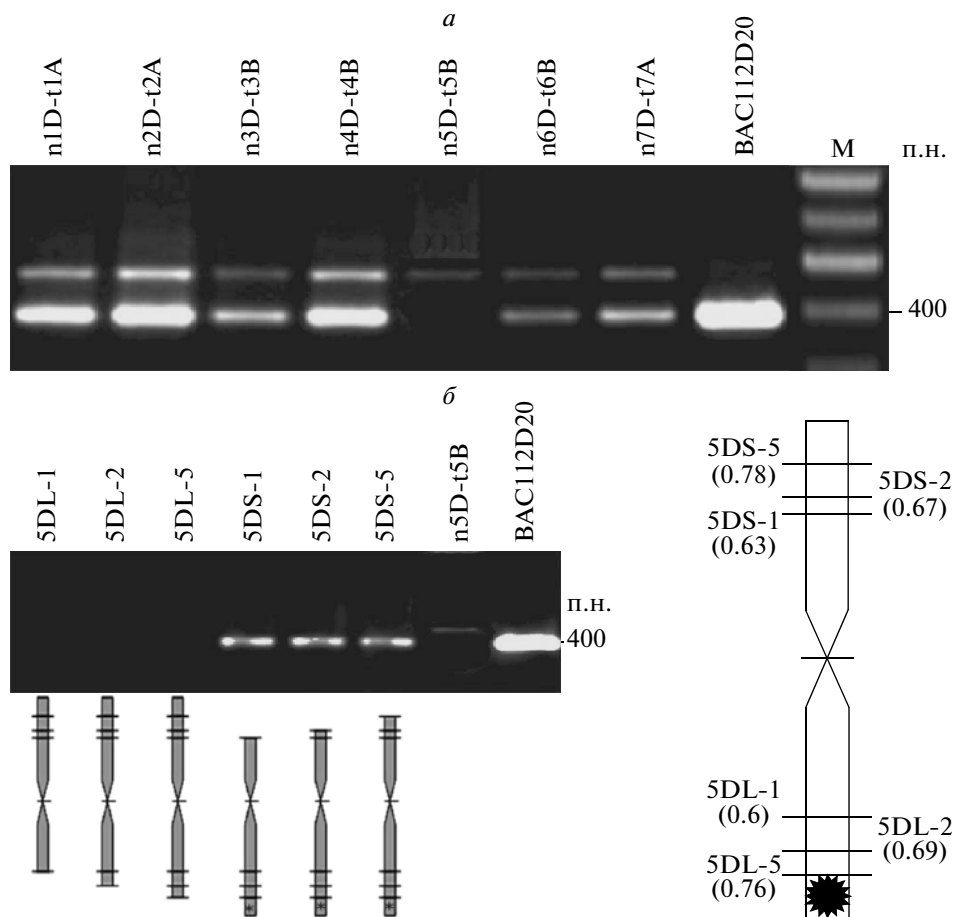
Для дальнейшего поиска маркера D-генома мы сконструировали специфические праймеры к районам, фланкирующим сайты встраивания ретро-транспозона *Lila* в составе ВАС-клона 112D20 (районы J1 и J2 длиной 402 и 355 п.н. соответственно, рис. 2). С помощью этих праймеров проведен ПЦР-анализ геномной ДНК различных диплоидных и полиплоидных видов Triticeae (рис. 3).

ПЦР-продукты нужной длины обнаружены только у мягкой пшеницы *T. aestivum* и диплоидного вида *Ae. tauschii*. По-видимому, ретро-транспозон *Lila* встроился в ВАС-клон 112D20 после дивергенции диплоидных видов внутри предка *Ae. tauschii*, и в дальнейшем этот участок вошел в состав гексаплоидной пшеницы в результате скрещивания тетраплоидной пшеницы (ВВАА) с предковым видом *Ae. tauschii* – предполагаемым донором D-генома.

С целью подтверждения специфичности ретро-транспозона *Lila* для D-генома нами проведена FISH с отдельными короткими клонированными фрагментами, представляющими различные районы данного элемента: LTR, а также внутренний кодирующий район. Наиболее специфичный сигнал обнаружен при гибридизации последовательности *pLila 2* длиной 1.2 т.п.н. из участка, кодирующего обратную транскриптазу *Lila* (рис. 1б). Таким образом показано, что входящий в состав ВАС-клона 112D20 *Ty3-gypsy*-ретротранспозон *Lila* 112D20-1 обуславливает преимущественную гибридизацию этого ВАС-клона с хромосомами D-генома мягкой пшеницы.

#### Хромосомная локализация ВАС-клона 112D20

Хромосомную локализацию ВАС-клона 112D20 мы определяли с помощью ПЦР-анализа нулли-тетрасомных и делеционных линий *T. aestivum* с использованием специфических праймеров к районам J1 и J2 (см. “Экспериментальную часть”). В линии n5D-t5B отсутствовал ПЦР-продукт, размер которого соответствовал району J1 в составе ВАС-клона 112D20, что позволило предположить локализацию данного ВАС-клона в хромосоме 5D мягкой пшеницы (рис. 4а). Образование у этой линии ПЦР-продукта, соответствующего J2-району (результат не представлен), можно объяснить недостаточной специфично-



**Рис. 4.** ПЦР-анализ нулли-тетрасомных (а) и делеционных (б) линий *T. aestivum* с использованием праймеров к J1-району. Внизу и справа — схема хромосомы 5D с указанием мест разрыва у соответствующих делеционных линий (цифры в скобках обозначают относительную длину плеча за исключением участка делеции). Звездочкой обозначен предполагаемый район локализации клона BAC 112D20. Размер фрагментов ДНК (п.н.) указан в соответствии с положением маркера “100 bp DNA ladder”.

стью праймеров, или же тем, что указанный район не уникален. Более точно локализацию BAC-клона в составе хромосомы 5D определили с использованием линии пшеницы с перекрывающимися делециями разной длины на обоих плечах этой хромосомы (рис. 4б). Проведение ПЦР с использованием праймеров к району J1 выявило отсутствие искомого продукта в трех линиях с делециями на длинном плече 5D, тогда как все остальные линии такой продукт содержали. Этот результат указывает на локализацию BAC-клона 112D20 в дистальном районе длинного плеча хромосомы 5D *T. aestivum*.

#### ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Наиболее хозяйственно ценными видами рода *Triticum* являются аллополиплоидные, тетра- и гексаплоидные виды ( $2n = 28, 42$ ). Изучение геномного состава и происхождения этих видов, и, в частности гексаплоидной пшеницы *T. aestivum*,

до сих пор весьма актуально для генетики и селекции. Геномная *in situ* гибридизация показала свою эффективность в случае сложных аллополиплоидных геномов, при поиске хромосомных перестроек (транслокаций), идентификации чужеродных хромосом и их сегментов на однородном генетическом фоне и т.д. [3–5]. Основное ограничение, которое необходимо учитывать при использовании этого метода заключается в том, что в ходе эволюции аллополиплоидного вида происходит “выравнивание” отдельных геномов, входящих в его состав, за счет процессов гомогенизации, обмена и перераспределения повторяющихся ДНК, в частности, МЭ. Так, при изучении аллополиплоидов *Nicotiana* показано, что эффективность метода GISH в идентификации отдельных геномов уменьшается при увеличении времени формирования аллополиплоидов, что предполагает непрерывный обмен повторяющейся ДНК между родительскими геномами [5]. Недостаточная эффективность GISH для дискриминации А-,

В- и D-геномов гексаплоидной пшеницы показана с использованием в качестве зондов суммарной ДНК соответствующих диплоидных видов-доноров – *T. urartu*, *Ae. speltoides* и *Ae. tauschii* [23]. В связи с этим возникает необходимость в поиске маркеров межгеномных различий. Решение данной задачи стало возможным с созданием ВАС-клонотек протяженных фрагментов геномной ДНК и их детального анализа с последующим выявлением отдельных последовательностей в качестве молекулярных маркеров генетического материала различных уровней организации: геномов, индивидуальных хромосом, а также генетических локусов.

Большинство ВАС-клонов геномной ДНК пшеницы демонстрирует дисперсный сигнал *in situ* гибридизации с хромосомами, что свидетельствует о высоком содержании повторяющихся МЭ в составе этих ВАС-клонов [24]. МЭ, учитывая их широкое распространение и скорость эволюции, претендуют на роль наиболее эффективных маркеров отдельных геномов. В частности, у пшеницы выделен целый ряд геном-специфичных МЭ, использованных для анализа полиплоидных видов *Triticum* [10, 16, 25–27]. Наиболее существенный вклад в вариацию размера генома и его дивергенцию у растений вносят LTR-ретротранспозоны [11–13].

Ранее мы провели скрининг ВАС-клонотеки *T. aestivum* сорта “Репан”, чтобы отобрать ВАС-клоны, содержащие субтеломерные повторы [17]. Два из отобранных клонов имели дисперсный сигнал гибридизации с хромосомами *T. aestivum*. При этом в обоих случаях гибридизация затрагивала только один из трех геномов. Анализ первичной структуры ВАС-клона 2383A24, специфичного для В-генома, показал, что подобная специфичность обусловлена присутствием *Ty3-gypsy*-ретротранспозона *Fatima* [18].

В настоящей работе проанализирован другой ВАС-клон – 112D20. С помощью FISH установлена его преимущественная локализация на хромосомах D-генома (рис. 1а). Результаты ВАС-FISH с диплоидными видами – предшественниками А-, В- и D-геномов – позволяют отчетливо идентифицировать хромосомы *Ae. tauschii* (DD) на фоне двух других видов. У *T. aestivum* межгеномные различия не столь ярко выражены из-за незначительной перекрестной гибридизации. Аналогичная картина наблюдается и в случае ВАС-клона 2383A24 [18], что подтверждает существование обмена между гомеологичными хромосомами в ходе эволюции аллополиплоидного генома *T. aestivum*.

Анализ первичной структуры ВАС-клона 112D20 позволил предположить, что его специфичность для D-генома определяется *Ty3-gypsy*-ретротранспозоном, входящим в малоизученное

семейство *Lila*. Наряду с гомологичной последовательностью AF488415.1 из генома *T. monococcum*, в базе данных TREP найдено несколько элементов из генома *T. aestivum*, сходных по структуре с описанным нами *Lila* 112D20-1. Максимальную гомологию проявляет LTR, а также 5'-кодирующая область (80–85%). Однако малое количество гомологичных *Lila* последовательностей не позволяет провести достоверный анализ их дивергенции, как это проделано в случае элемента *Fatima* [18]. FISH с использованием в качестве зонда различных участков ретротранспозона *Lila* 112D20 подтвердила его специфичность для D-генома и показала, что эти участки дивергировали с разной скоростью. Наибольшей специфичностью для D-генома обладает последовательность, расположенная в 3'-кодирующей области (RT-домен) (рис. 1б). Это же наблюдали и при изучении *Ty1-copia*-ретротранспозонов *Ae. speltoides* [26]. Использование в качестве зонда RT-домена этих элементов позволило более эффективно идентифицировать хромосомы В-генома.

Разработан метод поиска маркеров генетических локусов, основанный на подборе праймеров, специфичных к сайтам встраивания МЭ, и последующем анализе растительного материала с помощью ПЦР [25]. Мы использовали этот подход для более точной хромосомной локализации ВАС-клона 112D20 и определения приблизительного времени инсерции ретротранспозона *Lila* 112D20-1. С использованием праймеров к районам J1 и J2, фланкирующим этот ретротранспозон, проведен ПЦР-анализ различных видов *Triticaceae* (рис. 3). Показано, что инсерция ретротранспозона произошла внутри предкового вида *Ae. tauschii* до его скрещивания с тетраплоидным видом (ВВАА) и образования гексаплоидной пшеницы. Ранее была выведена формула расчета времени инсерции ретротранспозона по степени гомологии его LTR, основанная на линейном характере их взаимной дивергенции во времени:  $T = D/2t$ , где  $T$  – время инсерции;  $D$  – оценка дивергенции LTR;  $t$  – количество замен на нуклеотидный сайт в год [14]. Степень дивергенции LTR оценивали по методу Кимуры [28], в качестве  $t$  использовали значение  $1.3 \times 10^{-8}$  [29]. С помощью этой формулы мы получили ориентировочное время встраивания – 1.7 млн лет. Этот срок соответствует периоду между формированием диплоидных видов и появлением дикой тетраплоидной пшеницы *T. dicoccoides* (2.5–0.5 млн лет) [1, 30, 31]. На основании филогенетического анализа последовательностей *Fatima* ранее был сделан вывод, что амплификация В-специфичного варианта этого ретротранспозона также произошла до формирования тетраплоидной пшеницы (ВВАА), в составе донора В-генома – предкового вида *Ae. speltoides* [18]. Интересно, что инсерция В-специфичного ретротранспозона *Fatima* в ВАС-кло-



не 2383A24 произошла примерно в то же время (1.6 млн лет), что и инсерция *Lila* 112D20-1. По-видимому, геном-специфичные элементы у диплоидных видов Triticeae формировались в один и тот же период времени после их дивергенции от общего предка. Тот факт, что 80% МЭ в составе А- и В-геномов размножилось до формирования аллополиплоидных пшениц, также подтверждает эту гипотезу [10].

В изученном нами ВАС-клоне 112D20 МЭ занимают 98.5% нуклеотидной последовательности, тогда как на долю генов и неидентифицированных участков приходится 1.5% (таблица). Среди них преобладают МЭ класса LTR-ретротранспозонов, а именно, *Ty3-gypsy*-подобные элементы (61.7%). Найдена только одна ОРС, не относящаяся к кодирующим последовательностям МЭ (рис. 2). Кодируемая этой ОРС аминокислотная последовательность на 49% гомологична продукту неизвестного гена, который находится на расстоянии 6 т.п.н. от гена *Lr21* (AF532104). Нельзя исключить, что указанная ОРС имеет отношение к функционированию ретроэлемента *Fatima*, поскольку в составе ВАС-клона 112D20 она ограничена непосредственно этим элементом. Отсутствие значимых генов не позволяет установить точную локализацию ВАС-клона 112D20 по синтении с другими геномами. С помощью ПЦР мы определили его приблизительную локализацию в пределах дистальной четверти длинного плеча хромосомы 5D (рис. 4). В составе ВАС-клона 112D20 не найден субтеломерный повтор Spelt 52, несмотря на то что ранее его выявили с помощью ПЦР [17]. Вероятно, положительный результат ПЦР был обусловлен случайной реассоциацией праймеров с короткими последовательностями, и принадлежность ВАС-клона 112D20 к субтеломерному району остается невыясненной.

В целом, изученный ВАС-клон – это генетически инертный участок генома с высокой концентрацией МЭ, образующих структуры типа “вложенных инсерций” (nested structure). Сопоставив данные FISH, первичную структуру ВАС-клона, а также результаты ПЦР-анализа различных видов Triticeae, мы сделали вывод, что наблюдаемая картина гибридизации ВАС-клона 112D20 обусловлена присутствием *Ty3-gypsy*-ретротранспозона *Lila*, специфичного для D-генома. Этот ретроэлемент подвергся значительной амплификации внутри диплоидного вида – предка D-генома. Вместе с ранее полученными данными эти результаты свидетельствуют о том, что амплификация А-, В- и D-геном-специфичных ретроэлементов происходила до образования аллополиплоидных видов пшеницы на стадии формирования диплоидных видов-предшественников.

Работа выполнена при финансовой поддержке Министерства образования и науки РФ (Госконтракт 16.512.11.2223).

## СПИСОК ЛИТЕРАТУРЫ

1. Feldman M., Lupton F.G.H., Miller T.E. 1995. *Evolution of crop plants*. London: Longman group, 184–192.
2. Гончаров Н.П. 2002. *Сравнительная генетика пшениц и их сородичей*. Новосибирск: Сиб. университетское изд-во, 251 с.
3. Fuchs J., Houben A., Brandes A., Schubert I. 1996. Chromosome “painting” in plants a feasible technique? *Chromosoma*. **104**, 315–320.
4. Heslop-Harrison J.S., Schwarzacher T. 1996. Genomic Southern and *in situ* hybridization for plant genome analysis. In: *Methods of genome analysis in plants*. Ed. Jauhar P.P. Boca Raton, N.Y.: CRC Press, 163–179.
5. Lim K.Y., Kovarik A., Matyasek R., et al. 2007. Sequence of events leading to near-complete genome turnover in allopolyploid *Nicotiana* within five million years. *New Phytol.* **175**, 756–763.
6. Wicker T., Sabot F., Hua-Van A., et al. 2007. A unified classification system for eukaryotic transposable elements. *Nat. Rev. Genet.* **8**, 973–982.
7. Kumar A., Bennetzen J. 1999. Plant retrotransposons. *Annu. Rev. Genet.* **33**, 479–532.
8. Vitte C., Panaud O. 2005. LTR retrotransposons and flowering plant genome size: emergence of the increase/decrease model. *Cytogenet. Genome Res.* **110**, 91–107.
9. Sabot F., Schulman A.H. 2006. Parasitism and the retrotransposon life cycle in plants: a hitch-hiker’s guide to the genome. *Heredity*. **97**, 381–388.
10. Charles M., Belcram H., Just J., et al. 2008. Dynamics and differential proliferation of transposable elements during the evolution of the B and A genomes of wheat. *Genetics*. **180**, 1071–1086.
11. Bennett M.D., Leitch I.J. 1995. Nuclear DNA amounts in *Angiosperms*. *Ann. Bot.* **76**, 113176.
12. Bennetzen J.L. 1998. The evolution of grass genome organization and function. *Symp. Soc. Exp. Biol.* **51**, 123–126.
13. Piegu B., Guyot R., Picault N., et al. 2006. Doubling genome size without polyploidization: dynamics of retrotransposition driven genomic expansions in *Oryza australiensis*, a wild relative of rice. *Genome Res.* **16**, 1262–1269.
14. SanMiguel P., Gaut B.S., Tikhoniv A., et al. The paleontology of intergene retrotransposons in maize. *Nat. Genet.* **20**, 43–45.
15. Choulet F., Wicker T., Rustenholz C., et al. 2010. Megabase level sequencing reveals contrasted organization and evolution patterns of the wheat gene and transposable element spaces. *Plant Cell*. **22**, 1686–1701.
16. Zhang P., Li W., Friebe B., Gill B.S. 2004. Simultaneous painting of three genomes in hexaploid wheat by ВАС-FISH. *Genome*. **47**, 979–987.
17. Salina E.A., Sergeeva E.M., Adonina I.G., et al. 2009. Isolation and sequence analysis of the wheat B genome subtelomeric DNA. *BMC Genomics*. **10**, 414–428.

18. Salina E.A., Sergeeva E.M., Adonina I.G., et al. 2011. The impact of Ty3-gypsy group LTR retrotransposons *Fatima* on B-genome specificity of polyploidy wheats. *BMC Plant Biol.* DOI: 10.1186/1471-2229-11-99.
19. Sambrook J., Fritsch E.F., Maniatis T. 1989. *Molecular cloning: a laboratory manual*, 2<sup>nd</sup> edn. N.Y., Cold Spring Harbor: Cold Spring Harbor Lab. Press.
20. Surzycki S. 2000. *Basic methods in molecular biology*. N.Y.: Springer-Verlag.
21. Nagaki K., Tsujimoto H., Isono K., Sasakuma T. 1995. Molecular characterization of a tandem repeat, *Afa* family, and its distribution among Triticeae. *Genome*. **38**, 479–486.
22. Salina E.A., Lim K.Y., Badaeva E.D., et al. 2006. Phylogenetic reconstruction of *Aegilops* section *Sitopsis* and the evolution of tandem repeats in the diploids and derived wheat polyploids. *Genome*. **49**, 1023–1035.
23. Mukai Y., Nakahara Y., Yamamoto M. 1993. Simultaneous discrimination of the three genomes in hexaploid wheat by multicolor fluorescence in situ hybridization using total genomic and highly repeated DNA probes. *Genome*. **36**, 489–494.
24. Zhang P., Li W., Fellers J., et al. 2004. BAC-FISH in wheat identifies chromosome landmarks consisting of different types of transposable elements. *Chromosoma*. **112**, 288–299.
25. Paux E., Roger D., Badaeva E., et al. 2006. Characterizing the composition and evolution of homoeologous genomes in hexaploid wheat through BAC-end sequencing on chromosome 3B. *Plant J.* **48**, 463–474.
26. Raskina O., Belyaev A., Nevo E. 2002. Repetitive DNAs of wild emmer wheat (*Triticum dicoccoides*) and their relation to S-genome species: molecular cytogenetic analysis. *Genome*. **45**, 391–401.
27. Zhang P., Friebe B., Gill B.S. 2002. Variation in the distribution of a genome-specific DNA sequences on chromosomes reveals evolutionary relations in the *Triticum* and *Aegilops* complex. *Plant Syst. Evol.* **235**, 169–179.
28. Kimura M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16**, 111–120.
29. Ma J., Bennetzen J.L. 2004. Rapid recent growth and divergence of rice nuclear genomes. *Proc. Natl. Acad. Sci. USA*. **101**, 12404–12410.
30. Blake N.K., Leffler B.R., Lavin M., Talbert L.E. 1999. Phylogenetic reconstruction based on low copy DNA sequence data in an allopolyploid: the B genome of wheat. *Genome*. **42**, 351–360.
31. Huang S., Sirikhachornkit A., Su X., et al. 2002. Genes encoding plastid acetyl-CoA carboxylase and 3-phosphoglycerate kinase of the *Triticum/Aegilops* complex and the evolutionary history of polyploid wheat. *Proc. Natl. Acad. Sci. USA*. **99**, 8133–8138.