

Оригинальные работы

УДК 575.111:575.224:633.71:602.6

Е.А. ФИЛИПЕНКО¹, М.Л. ФИЛИПЕНКО²,
Е.В. ДЕЙНЕКО¹, В.К. ШУМНЫЙ¹

¹ Институт цитологии и генетики
Сибирского отделения Российской академии наук,
пр-т Лаврентьева, 10, Новосибирск, 630090
E-mail: deineko@bionet.nsc.ru, filipenko@bionet.nsc.ru

² Институт химической биологии и фундаментальной медицины
Сибирского отделения Российской академии наук,
пр-т Лаврентьева, 8, Новосибирск, 630090

АНАЛИЗ САЙТОВ ВСТРАИВАНИЯ Т-ДНК ИНСЕРЦИЙ У ТРАНСГЕННЫХ РАСТЕНИЙ ТАБАКА



Проведено клонирование прилежащих к трансгену областей геномной ДНК, выделенной из 17 трансгенных растений табака методом «инвертированной» ПЦР. В результате анализа первичных последовательностей 34 клонированных фрагментов ДНК установлено, что 10 из них имели 100%-ную гомологию с векторными последовательностями, не входящими в Т-ДНК область. Девять полученных нуклеотидных последовательностей имели гомологию с повторяющимися последовательностями в геноме табака. Процент гомологии варьировал от 70 до 90 %, причем идентифицированные повторы принадлежали к разным типам. Для большей части проанализированных фрагментов не выявлено гомологий с последовательностями, депонированными в базах данных. Выравнивание усеченных в процессе встраивания последовательностей левой и правой границ Т-ДНК инсерций выявило значительную кластеризацию (в районе десяти нуклеотидов) сайтов усечения для левой границы. Необходимо отметить, что пять последовательностей имели идентичные сайты усечения +23T, что свидетельствовало о предпочтительном использовании этого нуклеотида. АТ-содержание варьировало от 51 до 72 %, что было близко к суммарному проценту АТ пар в геноме табака.

© Е.А. ФИЛИПЕНКО, М.Л. ФИЛИПЕНКО, Е.В. ДЕЙНЕКО,
В.К. ШУМНЫЙ, 2007

Введение. Для генетической модификации растений, и в особенности двудольных, в настоящее время широко используется метод агробактериальной трансформации, основанный на переносе в растительный геном Т-ДНК области Тi-плазмиды *Agrobacterium tumefaciens*. Тi-плазида содержит гены вирулентности (vir-гены), продукты которых участвуют в вырезании и переносе Т-ДНК в растительную клетку. Вырезание происходит по двум несовершенным прямым повторам, правой и левой бордерным последовательностям, фланкирующим Т-ДНК. Белок VirD2 в присутствии VirD1-белка производит одноцепочечные разрывы ДНК в области бордерных последовательностей и ковалентно присоединяется к 5'-концу правого бордера. Т-ДНК переносится в растительное ядро в виде одноцепочечной молекулы, по всей длине связанной с агробактериальными белками E2 и белком VirD2 на 5'-конце. Интеграция Т-ДНК в растительный геном осуществляется посредством образования разрывов растительной ДНК по механизму негомологичной («незаконной») рекомбинации [1].

Анализ и накопление данных о первичной структуре границ Т-ДНК инсерций и прилежащих к ним областей растительной ДНК у трансгенных растений является необходимым элементом изучения молекулярных механизмов встраивания Т-ДНК в растительный геном. Наличие микрогомологий между Т-ДНК и эндогенной ДНК необходимо для процесса встраивания трансгена [2, 3]. Предполагается, что контекст ДНК, ее АТ-содержание, а также некоторые другие характеристики могут определять предпочтительность встраивания трансгена в растительный геном [4].

Цель настоящей работы – структурный анализ первичных последовательностей прилежащих к трансгену областей геномной ДНК у трансгенных растений табака (*Nicotiana tabacum* L., линия SR1).

Материалы и методы. Для анализа районов встраивания Т-ДНК инсерций использовали 17 независимо полученных нами ранее методом агробактериального переноса трансгенных растений табака (*N. tabacum* L., линия SR1). Клонирование прилежащих к трансгену областей геномной ДНК проводили методом «инвертированной» ПЦР [5]. Так как генетические конструкции для получения трансгенных растений были созданы на основе бинарного векто-

Описание нуклеотидных последовательностей клонированных фрагментов ДНК

№	Трансгенное растение	Размер секвенированного района растительной ДНК, п.н.		A/T, %	Наличие гомологии с последовательностями из GENE BANK
		левый бордер (RB)	правый бордер (RB)		
1	10,48	139		65	*
2			300		Вектор
3	16,70		171	52	D17456
4		224		58	*
5	121,41	130		74	*
6			133	51	*
7	121,57	163		66	*
8			212	61	Y08010
9	121,76	190		53	Z73149
10		211		71	*
11	121,83	223		70	*
12			103	65	Y12534
13	121,86	350		67	*
14			166	66	X80829, PAR-ген
15	Res60		300		Вектор
16			165	69	*
17		230			Вектор
18	Res66	550			Вектор
19		480			Вектор
20			169	56	*
21			279	61	*
22	Res36	320			Вектор
23			164	62	*
24	Res47	313		64	*
25			256	55	X59606
26	Res79	320			Вектор
27	Her	434		54	*
28			270		Вектор
29	Res91	538		68	X79006
30	10,18	430			Вектор
31			318	70	M77826 X14059 AB058958 AJ416574 X14593 AJ131837
32	121, 89		320	56	*
33	Res73	92		70	Z73149
34			200		Вектор

* — гомологии не выявлено.

ра рVI121, то концевые последовательности T-ДНК этого вектора были использованы для синтеза олигонуклеотидных праймеров.

Геномную ДНК трансгенных растений табака выделяли, как описано ранее [6]. ДНК (100–200 нг) гидролизвали эндонуклеазами *TaqI* или *MspI*, смесь экстрагировали фенолом, ДНК из смеси осаждали этанолом. ДНК растворяли в лигазном буфере (1–2 нг/мкл) и подвергали самолигированию в течение ночи при 8 °С. ДНК осаждали этанолом, растворяли в воде и использовали в качестве матрицы для ПЦР [7]. ПЦР проводили в 50 мкл буфера, содержащего 67 мМ трис-НСl (рН 8,9), 16 мМ (NH₄)₂SO₄, 1,5 мМ MgCl₂, 0,01% Tween 20, 10 мМ β-меркаптоэтанола, 100 мкМ dNTP, 1 мкМ праймеры, 2 ед. акт. Taq-ДНК-полимеразы. Амплификацию осуществляли в течение 32 циклов в следующем режиме: денатурация — 1 мин при 94 °С, отжиг праймеров — 1 мин при 52 °С для первого цикла и 0,5 мин при 62 °С для всех последующих циклов, элонгация — 3 мин при 72 °С. Далее 1/10 амплификационных смесей использовали для второго раунда амплификации с «внутренними» праймерами в аналогичных условиях ПЦР. Для клонирования амплифицированные фрагменты ДНК подвергали гидролизу эндонуклеазой PstI (для праймеров 1 и 3) или BamHI (для праймеров 5 и 7) и лигировали с плазмидой pBluscriptKSII, линейризованной соответствующей эндонуклеазой рестрикции. Нуклеотидные последовательности клонированных фрагментов ДНК определяли с использованием ABI PRISM BigDye Terminator v3.0 Ready Reaction Cycle Sequencing Kit («Amersham», Англия) в соответствии с инструкциями фирмы-производителя.

Для ПЦР использовали следующие структуры олигонуклеотидных праймеров («Биосан», Новосибирск):

pbi 1: GACCTGCAGTCTCATATTCACTCTCAATCC
 pbi 2: GATGGATCCATAAAATCCCCTCGGTATCCA
 pbi 3: GACCTGCAGTCGTTTCCCGCCTTCAGT
 pbi 4: GATGGATCCTTAATTCTCCGCTCATGATC
 pbi 5: GACGGATCCACTACGTGAACCATCAC
 pbi 6: CACGGATCCGTCTATCAGGGCCGATGG
 pbi 7: GACGGATCCAACGTCCGCAATGTGTTATTAAG
 pbi 8: CACAAGCTTGCCCCTCTCACTGGTGA

Результаты исследований и их обсуждение. Результаты анализа первичных последовательностей клонированных нами фрагментов ДНК у 17 трансгенных растений табака приве-

■ *Анализ сайтов встраивания Т-ДНК инсерций у трансгенных растений табака* ■

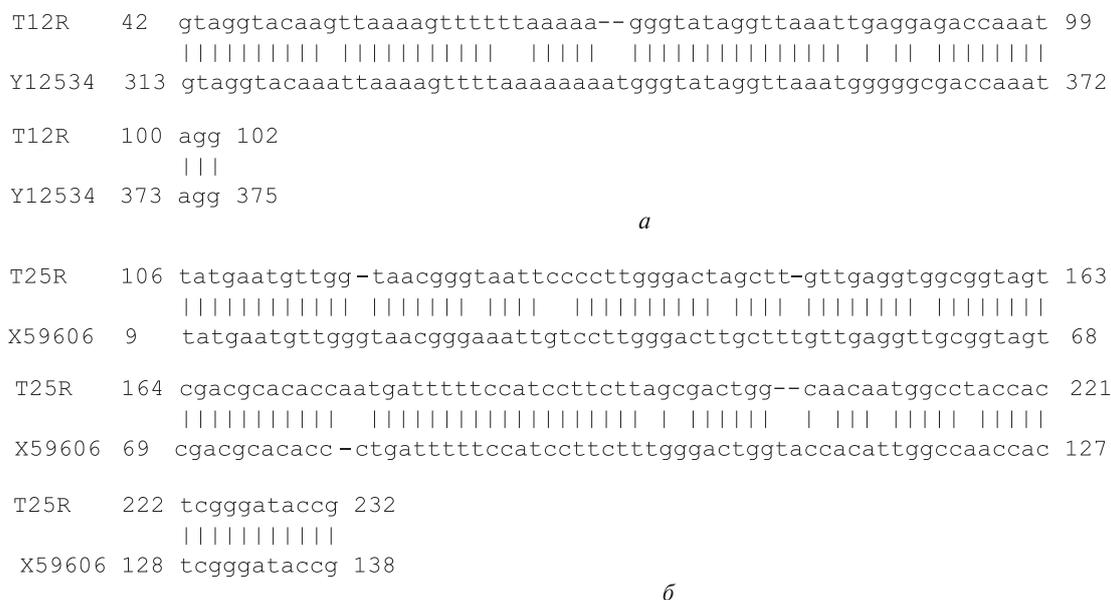


Рис. 1. Пример выравнивания нуклеотидных клонированных последовательностей прилежащих к трансгену областей растительной ДНК с нуклеотидными последовательностями, которые найдены в GENE BANK: а – фрагмент Т12R с Y12534 (*N. tabacum* unknown DNA fragment, clone H59) 87 % гомологии; б – фрагмент Т25R

дены в таблице. Как видно из представленных данных, для 10 фрагментов ДНК, принадлежащих восьми растениям, выявлена 100%-ная гомология с векторными последовательностями, не входящими в Т-ДНК область. Доля растений, несущих в геноме фрагменты векторных ДНК, составила 47,1 % от общего числа проанализированных трансформантов табака. Полученные результаты хорошо согласуются с опубликованными нами ранее данными о частоте встраивания векторных последовательностей в геном трансгенных растений [8]. По данным других авторов, при агробактериальной трансформации частота встраивания векторных последовательностей в геном двудольных растений варьирует в широких пределах – от 1,3 [9] до 75,0 % [10].

Причины и механизмы образования фрагментов, состоящих из Т-ДНК и векторной ДНК, к настоящему времени еще до конца не ясны. Известно, что мутации по VirD2-белку могут приводить к неэффективному узнаванию и разрезанию данным белком пограничных повторов и их «проскакиванию», что влечет встраивание в геном растения длинных конкатимеров плазмидной ДНК [11, 12].

Как и ожидалось, большинство проанализированных последовательностей не имели большой гомологии с последовательностями, депонированными в GENE BANK. Найденные нами гомологии приведены в таблице. В основном они связаны с повторяющимися последовательностями в геноме табака. Процент гомологии варьирует от 70 до 90 %, причем идентифицированные повторы принадлежат к разным типам повторов (рис. 1).

Выравнивание усеченных в процессе встраивания последовательностей левого и правого бордеров приведено на рис. 2. Для левого бордера наблюдается значительная кластеризация (в районе десяти нуклеотидов) сайтов усечения, причем пять последовательностей имели идентичные сайты усечения +23 T, что говорит о предпочтительном использовании этого нуклеотида. Сайт усечения правого бордера не имеет выраженного предпочтения.

В настоящее время в литературе существует мнение о том, что встраивание Т-ДНК происходит в АТ-богатые области [13]. Данные об АТ-содержании секвенированных нами прилежащих к трансгену областей приведены в таб-

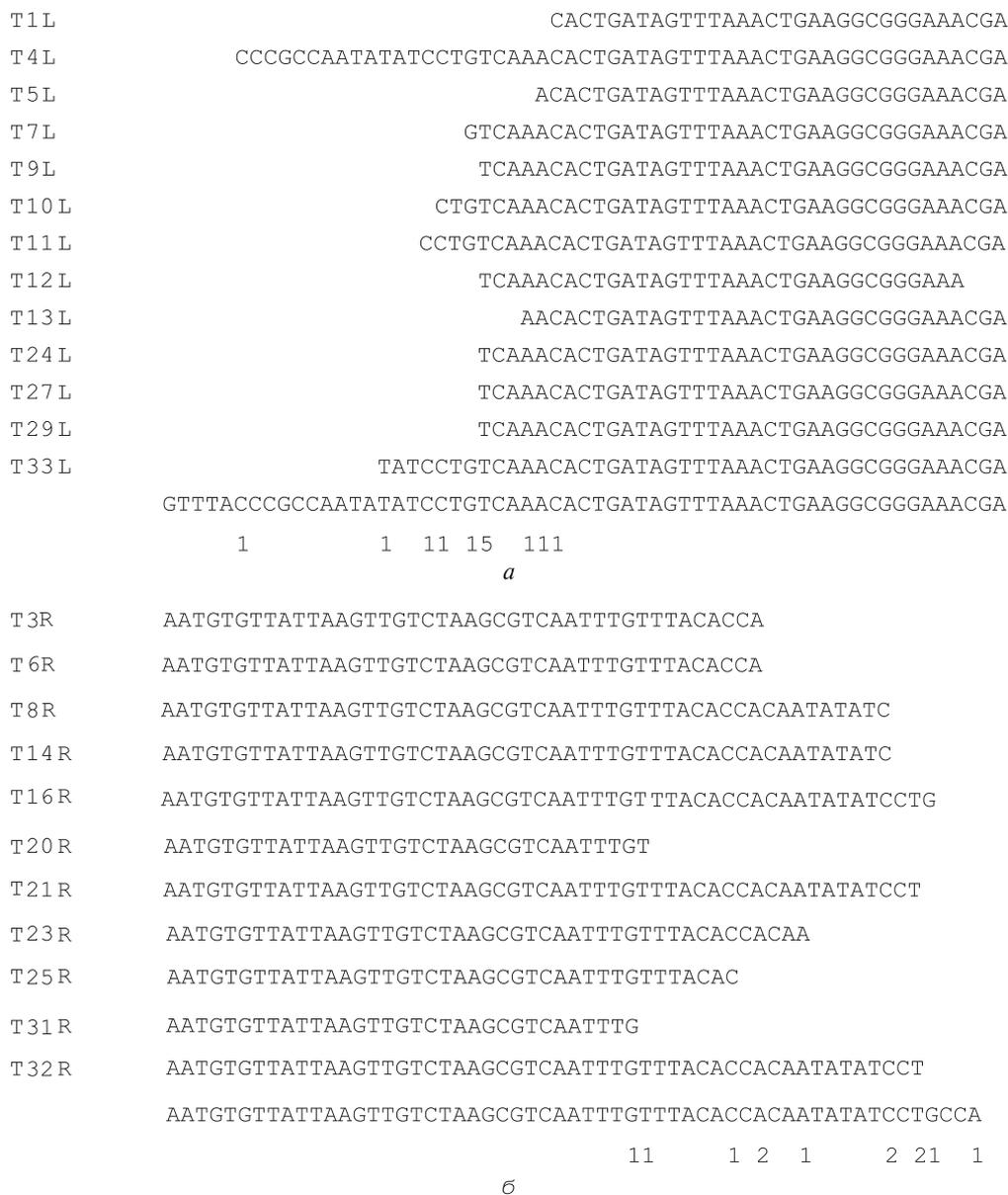


Рис. 2. Выравнивание «усеченных» последовательностей «левого» (а) и «правого» (б) бордеров в исследуемых прилежащих к трансгену областях ДНК

лице. АТ-содержание варьирует от 51 до 72 % при общем АТ-содержании в геноме табака 64 % [14]. Таким образом, в нашем исследовании такая тенденция не наблюдалась. Более того, Бруно с соавт. [15] в процессе анализа большого количества Т-ДНК инсерций предположили, что сайту интеграции в районе левого бордера часто предшествует поли-Т-тракт, чего мы не наблюдаем при анализе наших последователь-

ностей. Скорее опубликование данных о секвенировании генома табака должно добавить информативности нашему анализу.

Роль последовательностей ДНК так же, как и топологическая структура генома вокруг сайта встраивания Т-ДНК, еще до конца не изучена. Сравнительная характеристика сайтов встраивания в *Arabidopsis thaliana* не показала преимущественной интеграции в специ-

фические генные структуры хромосом [16]. Тем не менее у *Arabidopsis thaliana* показан повышенный процент Т в позиции, расположенной выше сайта встраивания, что наводит на мысль о влиянии нуклеотидного состава на место встраивания. Этим можно объяснить наблюдаемое повышенное Т-содержание и предпочтительность интеграции в район гена выше стартового кодона, а также большее количество встроок в интронах по сравнению с экзонами [15].

Известно, что АТ-последовательности являются районами с низкой стабильностью ДНК-дуплекса [17] и стерическим изгибом хромосомной структуры [18]. Вероятно, узнавание пространственной структуры ДНК является важным фактором при интеграции чужеродной ДНК в геном [19]. Повышенное АТ-содержание в сайтах встраивания трансгенной ДНК показано также для табака [20] и риса [18]. Такой нуклеотидный состав характерен для районов ядерного матрикса, посредством которых хроматиновая структура крепится к оболочке ядра, образуя петлевые функциональные домены [21, 22]. Предполагают, что указанные районы играют важную роль в регулировании экспрессии интегрированных генов [23]. Кроме того, топоизомераза II, фермент, делающий двухцепочечные разрывы в ДНК, имеет высокую аффинность к сайтам ядерного матрикса [24]. Закономерно предположение о том, что такие сайты могут быть предпочтительнее для проявления механизма незаконной рекомбинации.

Бруно с соавт. [15] показали, что у *A. thaliana* трансгены в равной мере распределены во всех пяти хромосомах, причем встраивание происходило реже в район центромер. Около 40 % трансгенов встраивалось в районы генов (определенных по базе данных AGI), что включает 54 % генома. Было показано, что у *A. thaliana* встраивание происходит преимущественно в 5'UTR, 3'UTR и промоторные области. Возможно, раскрученная спираль ДНК во время инициации транскрипции и терминации транскрипции на этих участках способствует большей доступности для встраивания Т-ДНК.

Не было выявлено определенной категории генов, обладающих более или менее других преимуществами для встраивания чужерод-

ных генов. При анализе наших последовательностей у восьми мы наблюдали гомологию с последовательностями повторов табака. В настоящий момент мы не можем делать выводов о предпочтительности встраивания Т-ДНК в область повторов, так как известно, что геном табака содержит большое количество этих последовательностей. Но и отвергать эту гипотезу не имеет смысла до момента появления в открытом доступе результатов секвенирования генома табака, которые позволят подсчитать встречаемость повторяющихся элементов.

Работа выполнена при поддержке гранта РФФИ № 05-04-48925.

SUMMARY. From the total DNA of 17 transgenic tobacco plants the DNA fragments containing T-DNA/plant DNA junctions were amplified using inverse polymerase chain reaction. Comparison of the nucleotide sequences of 34 fragments with the GENE BANK sequences revealed homology with vector sequences outside T-DNA in 10 cases and no homology with the known nucleotide sequences in most clones. The AT-content varied from 51 up to 72 % that is close to the total percentage of AT pairs in tobacco genome. Alignment of the sequences truncated during embedding of the left and the right borders has shown that for the left border significant clusterization (10 bp region) of truncation sites was observed, and five sequences had identical sites of truncation (+23 T) that showed the preferable use of this nucleotide. Nine created nucleotide sequences were homologous to the repeating sequences in tobacco genome. The percentage of homology varied from 70 up to 90 %. The identified repeats belong to different types.

РЕЗЮМЕ. Проведено клонування прилеглих до трансгена регіонів геномної ДНК, виділеної з 17 трансгенних рослин тютюну методом «інвертованої» ПЛР. У результаті аналізу первинних послідовностей 34 клонованих фрагментів ДНК встановлено, що 10 з них мали 100%-ну гомологію з векторними послідовностями, які не входять у Т-ДНК область. Дев'ять отриманих нуклеотидних послідовностей мали гомологію з повторювальними послідовностями в геномі тютюну. Відсоток гомології варіював від 70 до 90 %, причому ідентифіковані повтори належали до різних типів. Для більшої частини проаналізованих фрагментів не виявлено гомологій з послідовностями, депонованими в базах даних. Вирівнювання усічених у процесі вбудовування послідовностей лівої та правої границь Т-ДНК інсерцій виявило значну кластеризацію (у районі десяти нуклеотидів) сайтів усікання для лівої границі. Необхідно зазначити, що п'ять

последовательностей мали ідентичні сайти усікання +23T, що свідчило про переважне використання цього нуклеотиду. АТ-вміст варіював від 51 до 72 %, що було близько до сумарного відсотку АТ пар в геномі тютюну.

СПИСОК ЛИТЕРАТУРЫ

1. Tinland B. The integration of T-DNA into plant genomes // Trends Plant Sci. USA. — 1996. — **1**, № 6. — P. 178–184.
2. Gheysen G., Van Montagu M.V., Zambryski P. Integration of *Agrobacterium tumefaciens* transfer DNA (T-DNA) involves rearrangements of target plant DNA // Proc. Nat. Acad. Sci. USA. — 1987. — **84**. — P. 6169–6173.
3. Matsumoto S., Ito Y., Hosoi T., Takahashi Y., Mashida Y. Integration of *Agrobacterium* T-DNA into a tobacco chromosome: possible involvement of DNA homology between T-DNA and plant DNA // Mol. Gen. Genet. — 1990. — **224**. — P. 309–316.
4. Shimizu K., Takahashi M., Goshima N. et al. Presence of SAR-like sequence in junction regions between an introduced transgene and genomic DNA of cultured tobacco cells: its effect on transformation frequency // Plant J. — 2001. — **26**. — P. 375–384.
5. Ochman H., Ajioka J.W., Garza D., Hartl D.L. PCR technology: principles and applications for DNA amplification // Biotechnology. — 1990. — **8**. — P. 759–760.
6. Dellaporta S.L., Wood J., Hicks J.B. A plant DNA miniprep: version 2 // Plant Mol. Biol. Rep. — 1983. — **1**. — P. 19–22.
7. Филипенко Е.А., Филипенко М.Л., Мурашева С.В., Дейнеко Е.В., Загорская А.А., Сидорчук Ю.В., Шумный В.К. Анализ сайтов встраивания Т-ДНК у трансгенных растений ДНК с мутантным фенотипом // Докл. АН РАН. — 2000. — **370**, № 2. — P. 273–276.
8. Пухначева Н.В., Новоселя Т.В., Зоткевич Е.А., Дейнеко Е.В. Встраивание векторных последовательностей в геном трансгенных растений // Генетика. — 2005. — **41**, № 9. — P. 1203–1209.
9. Ramanathan V., Veluthambi K. Transfer of non-T-DNA portions of the *Agrobacterium tumefaciens* Ti plasmid pTiA6 from the left terminus of TL-DNA // Plant. Mol. Biol. — 1995. — **28**. — P. 1149–1154.
10. Kononov M., Bassuner B., Gelvin S. Integration of T-DNA binary vector «backbone» sequences into the tobacco genome: evidence for multiple complex patterns of integration // Plant J. — 1997. — **11**. — P. 945–957.
11. Tinland B., Schoumacher F., Gloeckler V., Bravo-Angel A.M., Hohn B. The *Agrobacterium tumefaciens* virulence D2 protein is responsible for precise integration of T-DNA into the plant genome // EMBO J. — 1995. — **14**, № 14. — P. 3585–3595.
12. Wolters A.-M.A., Trindade L.M., Jacobsen E., Visser R.G.F. Fluorescence *in situ* hybridization on extended DNA fibres as a tool to analyse complex T-DNA loci in potato // Plant J. — 1998. — **13**. — P. 837–847.
13. Daniels G.R., Deininger P.L. Integration sites preferences of the Alu family and similar repetitive DNA sequences // Nucl. Acids Res. — 1985. — **13**. — P. 8939–8954.
14. Karlin S., Mrazek J. Compositional differences within and between eukariotic genomes // Proc. Nat. Acad. Sci. USA. — 1997. — **94**. — P. 10227–10232.
15. Brunaud V., Balzergue S., Dubreucq B., Aubourg S., Samson F. et al. T-DNA integration into the *Arabidopsis* genome depends on sequences of pre-insertion sites // EMBO Rep. — 2002. — **3**, № 12. — P. 1152–1157.
16. Azpiroz-Leehan R., Feldmann K.A. T-DNA integration mutagenesis in *Arabidopsis*: going back and forth // Trends Genet. — 1997. — **13**. — P. 152–156.
17. Muller A., Kamisugi Y., Grunberg R., Niedenhof R., Huxford R.J., Meyer P. Palindromic sequences and A+T-rich DNA elements promote illegitimate recombination in *Nicotiana tabacum* // J. Mol. Biol. — 1999. — **291**. — P. 28–46.
18. Takano M., Egawa H., Ikeda J.-E., Wakasa K. The structures of integration sites in transgenic rice // Plant J. — 1997. — **11**, № 3. — P. 353–361.
19. Sawasaki T., Takahashi M., Goshima N., Morikawa H. Structures of transgene loci in transgenic *Arabidopsis* plants obtained by particle bombardment: Junction regions can bind to nuclear matrices // Gene. — 1998. — **218**. — P. 27–35.
20. Gheysen G., Herman L., Breyne P., Gielen J., Van Montagu M., Depicker A. Cloning and sequence analysis of truncated T-DNA inserts from *Nicotiana tabacum* // Gene. — 1990. — **94**. — P. 155–163.
21. Amati B., Pick L., Laroche T., Gasser S. M. Nuclear scaffold attachment stimulates, but is not essential for ARS activity in *Saccharomyces cerevisiae*: analysis of the *Drosophila* ftz SAR // EMBO J. — 1990. — **9**. — P. 4007–4016.
22. Hall G.Jr., Allen G.C., Loser D.S., Thompson W.F., Spiker S. Nuclear scaffolds and scaffold-attachment regions in higher plants // Proc. Nat. Acad. Sci. USA. — 1991. — **88**. — P. 9320–9324.
23. Thompson E.M., Adenot P., Tsuji F.L., Renard J.-P. Real time imaging of transcriptional activity in live mouse preimplantation embryos using a secreted luciferase // Proc. Nat. Acad. Sci. USA. — 1995. — **92**. — P. 1317–1321.
24. Sperry A.O., Blasquez V.C., Garrard W.T. Disfunction of chromosomal loop attachment sites: Illegitimate recombination linked to matrix association regions and topoisomerase II // Proc. Nat. Acad. Sci. USA. — 1989. — **86**. — P. 5497–5501.

Поступила 10.10.06