

УДК 577.218

ПОДХОДЫ К НАПРАВЛЕННОЙ ИНТЕГРАЦИИ ДНК НА ОСНОВЕ ТРАНСПОЗАЗ И РЕТРОВИРУСНЫХ ИНТЕГРАЗ

© 2011 г. Е. С. Княжанская^{1*}, О. В. Кондрашина², М. Б. Готтих¹

¹Научно-исследовательский институт физико-химической биологии
им. А.Н. Белозерского Московского государственного университета им. М.В. Ломоносова, Москва, 119991

²Химический факультет Московского государственного университета
им. М.В. Ломоносова, Москва, 119991

Поступила в редакцию 04.04.2011 г.

Принята к печати 10.05.2011 г.

Способность ретровирусов и транспозонов встраивать свой геном в геном клетки-хозяина делает их привлекательными объектами для конструирования генно-терапевтических векторов. Однако ферменты, осуществляющие встраивание генетического материала, не проявляют избирательности по отношению к последовательностям-мишеням, что приводит к практически произвольному включению ДНК в геном клетки-реципиента. Произвольное встраивание приводит к мутациям, которые, в свою очередь, могут вызывать нежелательные для клетки последствия и в отдельных случаях неопластическую трансформацию. В связи с этим для успешного функционирования генно-терапевтических систем на основе ретровирусов и транспозонов совершенно необходимо модифицировать их таким образом, чтобы добиться направленного встраивания вектора в заранее заданную последовательность в геноме человека. В данном обзоре рассмотрены известные на настоящий день подходы, разработанные для высокоспецифичной модификации генома, в том числе с использованием гибридных белков на основе ретровирусных интеграз, транспозаз, а также клеточных факторов, взаимодействующих с этими ферментами.

Ключевые слова: интеграна, транспозаза, направленная интеграция, направленная транспозиция.

APPROACHES TOWARDS DIRECTED DNA INTEGRATION BY THE USE OF RETROVIRAL INTEGRASES AND TRANSPOSASES, by E. S. Knyazhanskaya^{1*}, O. V. Kondrashina², M. B. Gottikh¹ (¹Belozersky Institute for Physical and Chemical Biology, Moscow State University, Moscow, 119991 Russia; *e-mail: e.knyazhanskaya@gmail.com; ²Chemical Department, Moscow State University, Moscow, 119991 Russia). The ability of retroviruses and transposases to insert own genome into a host-cell allow us to consider them as a preferable object for constructing gene therapy vectors. However, enzymes that perform the insertion of the genetic material do not display a selectivity towards target nucleotide sequences that results in an almost random DNA introduction into the recipient cell genome. Random insertion leads to mutations which might cause a number of undesirable consequences including neoplastic transformation in the cell. Thereby, in order to achieve a successful functioning of retroviral and trasposonal genetic therapy systems, it is essential to modify them in such a way that directed integration of the vector in a target sequence in the human genome could be achieved. In the review approaches that have been developed for a high specific modification of genome, including the construction of hybrid proteins on the basis of retroviral integrases, transposases, as well as cellular factors interacting with these enzymes, are presented.

Keywords: integrase, transposase, directed integration, directed transposition.

Принятые сокращения: ASLV (avian sarcoma leucosis virus) – вирус саркомы-лейкоза птиц; dbs (DNA binding sequence) – ДНК-связывающая последовательность; LABD (LexA binding domain) – LexA-связывающий домен; LEDGF/p75 (lens-epithelium derived growth factor) – фактор роста эпителия хрусталика; LTR (long terminal repeats) – длинные концевые повторы; Mo-MLV/MLV (Moloney murine leukemia virus) – вирус лейкоза мышей Молони; SB – транспозон *Sleeping Beauty*; UAS (upstream activating sequence) – активирующая последовательность; X1-SCID (X-linked severe combined immunodeficiency) – X-сцепленный с полом тяжелый комбинированный иммунодефицит; ВИЧ-1/HIV-1 – вирус иммунодефицита человека первого типа; ДСД – ДНК-связывающий домен; КСД – “крюк”-связывающий домен; ТСД – транспозаза-связывающий домен.

* Эл. почта: e.knyazhanskaya@gmail.com

ВВЕДЕНИЕ

Генная терапия различных заболеваний – это активно и многопланово развивающаяся область медицины и биотехнологии. Основной прием генной терапии – введение в клетки пациента необходимых генов, называемых “терапевтическими”, при условии не только их успешной доставки, но и стабильной экспрессии в клетках. Добиться этого позволяет использование вирусов в качестве носителей “терапевтического” гена. Не менее 70% клинических протоколов основано на использовании вирусных векторов [1–3]. Самыми удобными, с точки зрения достижения стабильной экспрессии, считаются ретровирусы, так как они обеспечивают долговую и стабильную экспрессию трансгена благодаря способности встраивать свой геном в геном клетки-хозяина. По данным на 2011 г., ретровирусы составляют 21% от всех используемых систем доставки генов, занимая второе место после аденовирусов (24%) [3]. Однако есть объективные факторы, осложняющие эффективное применение ретровирусов для генной терапии. Многие ретровирусы – это возбудители опасных заболеваний человека и животных, поэтому конструирование векторов на их основе требует большой осторожности. Способность ретровирусов встраивать (интегрировать) свой геном в геном зараженной клетки, что и определяет их “привлекательность” как основы для конструирования векторов, также и осложняет их применение. Показано, что интегрированный провирус может активировать экспрессию клеточных генов и, в частности, онкогенов [4–6]. К тому же, такая интеграция может нарушить целостность важных для клетки генов и генетических элементов [7].

Перенос и встраивание ДНК могут осуществляться не только ретровирусными интегразами. Этим свойством обладают и такие мобильные клеточные элементы, как транспозоны [8–10], которые также используют для конструирования генно-терапевтических векторов [11]. К сожалению, встраивание транспозона, называемое транспозицией, также может приводить к необратимым генетическим изменениям в клетке.

На сегодняшний день многие исследователи заняты решением проблем, возникающих при встраивании в клеточную ДНК ретровирусных векторов или векторов на основе транспозонов и, в частности, проблемой повышения специфичности их интеграции. В данном обзоре рассмотрены генно-инженерные системы, которые позволяют добиться повышения специфичности встраивания ДНК, осуществляемого ретровирусами или транспозонами.

СИСТЕМЫ НАПРАВЛЕННОЙ ИНТЕГРАЦИИ НА ОСНОВЕ РЕТРОВИРУСОВ

Этапы интеграции ДНК ретровирусов

Для эффективной экспрессии своих белков ретровирусу необходимо встроить ДНК-копию своего генома в клеточную ДНК. Этот процесс осуществляет входящий в состав вируса фермент, называемый интегразой. Интеграция ДНК начинается в цитоплазме инфицированной клетки и происходит в два этапа [12]. Важный для интеграции элемент в структуре вирусного генома – длинные концевые повторы (LTR – long terminal repeats), фланкирующие ДНК вируса с обоих концов. Каждый LTR содержит три участка: U3, R и U5. На расстоянии двух нуклеотидов от 3'-конца каждой цепи ДНК находится консервативный динуклеотид CA, который встречается в LTR всех ретровирусов [13, 14]. Интеграза в составе прединтеграционного комплекса узнает нуклеотидные последовательности, расположенные на концах участков U3 и U5 вирусной ДНК, связывается с ними, а затем катализирует эндонуклеазное отщепление концевого динуклеотида, что называется 3'-концевым процессингом. В случае ВИЧ-1 это динуклеотид GT [13], для других ретровирусов динуклеотиды могут быть иными.

После отщепления динуклеотидов у сложных ретровирусов (например, лентивирусов) вирусная ДНК в составе прединтеграционного комплекса транспортируется в ядро за счет наличия у интегразы сигнала ядерной локализации. У простых ретровирусов (например, гаммаретровирусов) интеграция может происходить только в делющихся клетках, так как требует разрушения ядерной мембраны [14]. В ядре интеграза катализирует второй этап интеграции, называемый переносом цепи. Он включает нуклеофильную атаку 3'-гидроксильных групп процессированных цепей ДНК по межнуклеотидным фосфатам обеих цепей клеточной ДНК с образованием ковалентного продукта [13]. В случае ВИЧ-1 переетерификации подвергаются межнуклеотидные связи, расположенные в разных цепях ДНК-мишени на расстоянии пяти пар нуклеотидов одна от другой, для других ретровирусов это расстояние иное. Для завершения процесса интеграции требуются: процессинг 5'-концов вирусной ДНК, полимеразная достройка недостающих нуклеотидов и лигирование, которые осуществляются при участии клеточных белков [15].

Интеграционные предпочтения ретровирусов

Необходимо подчеркнуть, что реакция 3'-концевого процессинга специфична к последовательности, т.е. ей подвергаются только последовательности ДНК, первичная структура которых

аналогична U3- или U5-последовательностям ДНК вируса [15]. В то же время интеграция не зависит напрямую от первичной структуры геномной ДНК клетки-хозяина; в геноме можно выделить от 500 до 1000 участков, где интеграция происходит с частотой приблизительно 2.5×10^{-4} на участок [16]. Ранее предпринятые попытки обнаружить какие-либо структурные особенности ДНК-мишени выявили некоторые закономерности интеграции, однако не дали ответа на вопрос: что в целом определяет выбор сайта интеграции в геноме. В ряде экспериментальных работ, проведенных *in vitro*, т.е. с использованием очищенного препарата интегразы (или прединтеграционного комплекса) и молекулы ДНК, служащей моделью клеточной хромосомы, показано, что интеграции благоприятствуют локальные искривления в молекуле ДНК [17–21]. Такие искривления возникают, например, при связывании ДНК с ДНК-связывающими белками или с нуклеосомами [20–22] или в АТ-богатых участках ДНК [20]. Также известны работы, где показана склонность вируса лейкоза мышей Молони (Mo-MLV) встраиваться в транскрипционно-активные участки генома и в районы CpG-островков [23], а для вируса саркомы-лейкоза птиц (ASLV) – в метилированные участки [24]. Компактизация хроматина при добавлении гистона H1 повышает эффективность интеграции ASLV [25]. Отмечалось, что предпочтительными сайтами интеграции MLV, к тому же, могут быть участки генома с повышенной чувствительностью к DNКазе I, т.е. свободные от каких-либо ДНК-связывающих белков [26]. В случае ВИЧ-1 замечено, что его интеграция происходит несколько чаще в районе так называемых Alu-островков [27, 28], а при компактизации хроматина эффективность интеграции подавляется [25].

Однако с появлением полной нуклеотидной последовательности генома человека и ряда других организмов стало возможным выявить некоторые закономерности, определяющие предпочтительные сайты интеграции ретровирусов. Шредер (Shroeder) и соавторы [29] впервые показали, что интеграция ДНК-копии ВИЧ-1 преимущественно происходит в транскрибируемые гены и в то же время “избегает” CpG-островков. Впоследствии эти данные получили подтверждение в других работах [30–36]. Аналогичные закономерности показаны и для других представителей подсемейства лентивирусов: вируса иммунодефицита обезьян (Simian Immunodeficiency Virus – SIV) [37, 38] и вируса иммунодефицита кошек (Feline Immunodeficiency Virus – FIV) [39].

Достаточно детально изучены особенности интеграции MLV: неоднократно показано, что этот вирус преимущественно интегрируется в интервале ± 2 т.н. от сайтов инициации транскрип-

ции [31, 32, 34, 40], а также в район CpG-островков – а это тоже указание на ассоциацию с транскрипцией. Накопленные к настоящему времени данные о предпочтительных сайтах интеграции разных семейств ретровирусов обобщены в таблице.

Неопределенность местонахождения сайта интеграции ретровирусов является серьезным препятствием для их клинического использования в качестве векторов для генной терапии. Встраиваясь в произвольный участок клеточного генома, вирусная ДНК может оказать мутагенное воздействие на клетку, особенно в том случае, если интеграция затронула онкогены.

Некоторое время тому назад появились данные о том, что при лечении X-сцепленной с полом тяжелой комбинированной иммунной недостаточности (X-linked severe combined immunodeficiency – X1-SCID) у детей с помощью генно-терапевтического вектора на основе MLV [42] у трех из одиннадцати пациентов наблюдалась неконтролируемая пролиферация зрелых Т-клеток, которая привела к развитию лейкоза. Анализ нуклеотидной последовательности сайтов интеграции вектора показал, что ретровирусный вектор интегрировался вблизи промотора протоонкогена LMO2, расположенного в хромосоме 11 [4–6]. Через некоторое время появилось сообщение о том, что еще у двух пациентов интеграция вектора затронула протоонкогены и вызвала развитие лейкоза. В другом клиническом испытании вектора на основе MLV для лечения X1-SCID также был случай развития лейкоза [43, 44].

Особенности интеграции MLV: преимущественное встраивание в сайты инициации транскрипции, а также наличие сильного энхансера транскрипции в U3-фрагменте LTR [45] вполне могут быть причиной такого канцерогенного воздействия векторов на его основе. Анализ распределения интеграции вектора в клетках пациентов подтвердил наличие характерных для MLV особенностей, показанных ранее в экспериментах на культурах клеток [43, 44]. На сегодняшний день с помощью дрожжевой двухгибридной системы удалось выделить ряд клеточных белков, предположительно взаимодействующих с интегразой MLV, – это некоторые транскрипционные факторы (TFIIЕ-β, В-ATF, Znf15, Znf38, Ankrd49, AF9), факторы ремоделирования хроматина (brd2, Enx-1), а также белки, участвующие в репарации ДНК (Ku70, fen1) [46]. Скорее всего, взаимодействие с этими хозяйскими белками определяет особенности распределения сайтов интеграции MLV. Показано, что области интеграции векторов на основе MLV богаты участками связывания транскрипционных факторов [32]. Можно предположить, что в интеграции MLV задействована груп-

Особенности интеграции ретровирусов (по данным [41])

Род ретровирусов	Вирус	Происхождение клетки-хозяина	Количество сайтов ¹	Относительная частота интеграций в различные области генома ²		
				транскрипционные единицы	± 2 т.н. от старта транскрипции	± 2 т.н. от CpG-островков
Лентивирусы	HIV-1	Человек	59869	+ / ++	0	– / 0
		Другие виды	2421			
	HIV-2	Человек	202			
	SIV	Человек	148			
		Обезьяна	501			
	EIAV	Человек	1241			
		Другие виды	70			
FIV	Человек	226				
Альфаретровирусы	ASLV	Человек	695	+	0	+
		Птица	658			
Бетаретровирусы	MMTV	Человек	298	0	0	0
		Мышь	170			
Гаммаретровирусы	MLV	Человек	4005	+	++	++
		Мышь	189			
		Другие виды	953			
	MSCV	Мышь	259			
	PERV	Человек	1962			
	XMRV	Человек	472			
Дельтаретровирусы	HTLV-1	Человек	1235	+	+	+
Спунавирусы	HFV	Человек	3457	0	++	++
		Другие виды	263			
Эндогенные ретровирусы	HERV-K	Человек	1565	+	++	++

¹ – Количество сайтов интеграции, проанализированных в [41].

² – Все частоты приведены на основании баз данных RefSeq.

“0” – Отсутствие статистически значимого отличия от среднего значения по геному; “+ / ++” – положительное статистически значимое отличие от среднего по геному, где “++” для отношения > 2 и “+” для отношения < 2; “–” – отрицательное статистически значимое отличие от среднего по геному.

па белков, так или иначе участвующих в инициации транскрипции, что и определяет итоговое распределение сайтов интеграции этого вируса.

Кроме того, выявлено несколько белков, предположительно взаимодействующих с интегразой ВИЧ-1 в клетке: белок Ini-1 (IN-interactor 1); фактор, препятствующий автоинтеграции (BAF – barrier to autointegration factor); HMGAI (high mobility group A1) и белок теплового шока 60 (Hsp60) [47, 48]. Однако только для одного клеточного белка – LEDGF/p75 (lens epithelium derived growth factor) – экспериментально доказано участие в интеграции ВИЧ-1 [49–52]. Функции этого белка в клетке еще не до конца ясны, но предполагают его участие в активации транскрипции и выживании клеток, а также показана его роль как аутоантигена при atopическом дерматите и раз-

личных воспалениях [41]. Нормальная экспрессия LEDGF/p75 в клетках, инфицированных ВИЧ-1 (или другим лентивирусом), необходима для эффективной интеграции вирусного генома [53–55]. Интересно отметить, что распределение сайтов, с которыми взаимодействует LEDGF/p75 в геноме, сходно с распределением сайтов интеграции для ВИЧ-1, т.е. это в основном внутригенные участки, в то время как промоторы и межгенные области представлены слабо [41]. Интеграция ВИЧ-1 в клетках, нокаутированных по LEDGF/p75, распределяется в соответствии с закономерностями, описанными для MLV [52, 56]. Эти данные позволяют предположить, что в случае интегразы ВИЧ-1 клеточный транскрипционный фактор LEDGF/p75 играет важную роль в “выборе” места интеграции и направляет прединтеграционный комплекс во внутригенные области.

Итак, совершенно ясно, что для обеспечения безопасного использования векторов на основе ретровирусов необходимо изменить их интеграционные особенности таким образом, чтобы перенаправить интеграцию из потенциально опасных для клетки участков в безопасные районы. Далее рассмотрены известные на настоящий день подходы, направленные на изменение и повышение специфичности интеграции ретровирусных векторов.

Подходы к направленной интеграции на основе ретровирусных интеграз или их кофакторов

На сегодняшний день нет большого разнообразия в подходах к изменению интеграционных характеристик ретровирусов. Исходя из ряда экспериментальных данных, можно предположить, что ключевую роль в выборе места интеграции играет вирусный фермент интегразы. В связи с этим практически все варианты направленной интеграции основаны на модификации интегразы с получением гибридных белков за счет присоединения к интегразе некоего ДНК-связывающего домена (ДСД), специфически узнающего определенную последовательность ДНК [57–62].

Хотя этот подход можно считать самым распространенным, он не лишен недостатков. Во-первых, присоединение даже небольшого белкового домена к интегразе может сказаться на ее каталитической активности. Эта проблема отчасти решается включением в конструкцию аминокислотных линкеров между интегразой и ДСД. Во-вторых, для осуществления направленной интеграции модифицированная интегразы должна входить в состав вириона. Включение дополнительного большого белкового домена в молекулу интегразы может повлиять на упаковку вирусной частицы, и, в свою очередь, на инфекционность вируса [57]. Избежать этих проблем можно, если проводить модификацию не самой интегразы, а ее клеточного кофактора, непосредственно участвующего в процессе интеграции. На сегодня такой кофактор описан только для интегразы ВИЧ-1. Это белок LEDGF/p75 (см. выше). Существует одна работа, в которой представлена конструкция, состоящая из ДСД и LEDGF/p75 [63] и изучена ее способность направлять интеграцию. Недавно описан оригинальный метод внесения нуклеотидных последовательностей генов, кодирующих ДСД, в состав генома ретровируса с последующим отбором жизнеспособных мутантов [64].

Еще одна проблема, которую приходится учитывать при конструировании систем направленной интеграции, – выбор сайта интеграции в геноме. В ранних работах такую задачу даже не ставили, так как поначалу исследовали принципиальную возможность изменения интеграционных “предпочте-

ний” интегразы. Модификацию интегразы проводили, используя такие белковые домены, как репрессор фага λ [57, 63], белок LexA *Escherichia coli* [58, 59], домен транскрипционного активатора Sp1 [61] и транскрипционный фактор Zif268 [60]. Для всех этих ДСД не существует специфического сайта узнавания в геноме человека. Однако в более поздних работах использованы конструкции на основе синтетических белков, содержащих структурный мотив “цинковые пальцы”. Белки, связывающие ДНК за счет мотива “цинковые пальцы”, пользуются большой популярностью в различных геном-модифицирующих системах (см. обзоры [65–70]). Такие белки не требуют мультимеризации для эффективного связывания ДНК, к тому же небольшой размер мотива (~30 аминокислотных остатков), а также высокая специфичность взаимодействия “пальца” с конкретным триплетом ДНК [67, 69], позволяют комбинировать различные домены для получения белков с заданной специфичностью. В ряде работ Чоу (Chow) и соавторов [62, 71, 72] описаны конструкции на основе интегразы ВИЧ-1 и синтетического белка E2C, специфически распознающего последовательность *e2c* в 5'-нетранслируемой области гена *erbB-2* в хромосоме 17 человека [73, 74]. В другой работе [64] гены двух различных синтетических белков, содержащих “цинковые пальцы” и способных специфически распознавать конкретные участки в геноме человека, клонировали в разные участки нуклеотидной последовательности *gag-pol* генома вируса MLV.

Прежде чем приступить к подробному описанию подходов к направленной интеграции ДНК, остановимся на методах, используемых для определения специфичности интеграции.

Методы определения специфичности интеграции в модельной системе in vitro

Методы определения направленности интеграции *in vitro* сводятся к воспроизведению реакции интеграции на модельной системе, имитирующей интеграцию в клетке. Чаще всего такая система включает:

1) выделенный из культуры *E. coli* препарат рекомбинантной интегразы либо гибрида на основе интегразы и ДСД;

2) короткий ДНК-дуплекс, последовательность которого соответствует либо U3-, либо U5-участку вирусной ДНК (обычно обозначается как субстрат интегразы), где в одной из цепей отсутствует концевой динуклеотид, отщепляемый интегразой в реакции 3'-концевого процессинга [75]. Такой дуплекс имитирует процессированный субстрат интегразы, которая распознает его и встраивает либо в молекулу другой ДНК (обозна-

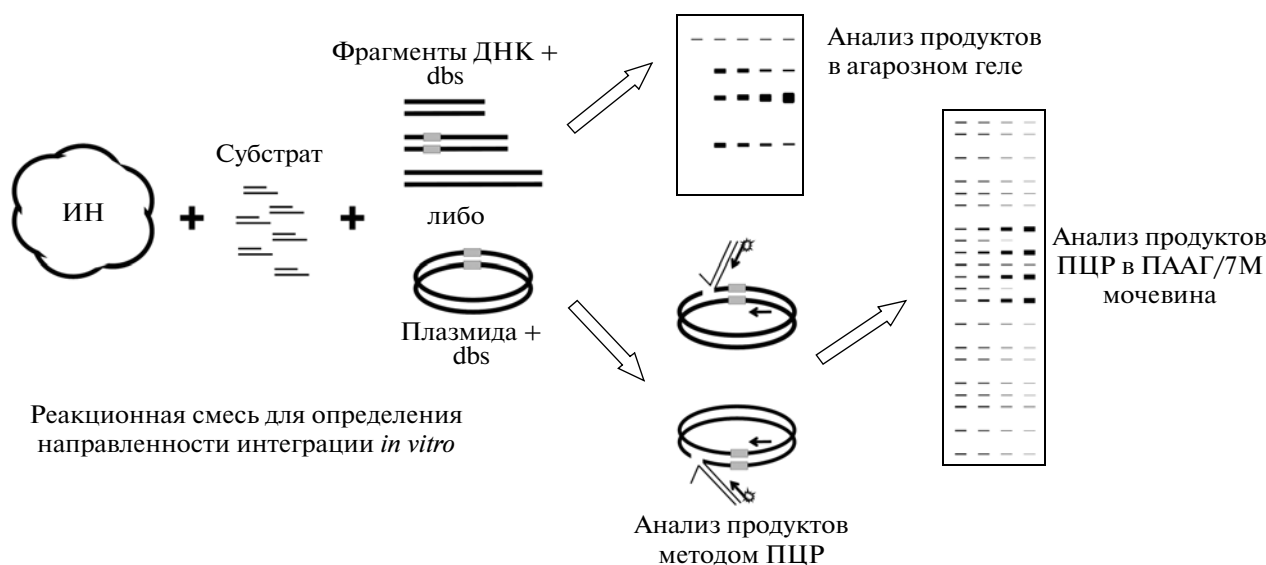


Рис. 1. Схематическое изображение экспериментальной системы для оценки направленности интеграции *in vitro* (по-яснения в тексте). ИН – препарат рекомбинантной интегразы или гибрида ИН-ДСД; во всех фрагментах ДНК серым цветом выделен участок *dbs* (участок связывания ДСД); стрелка – праймер; стрелка со звездой – праймер с радиоактивной меткой на 5'-конце.

чаемую как мишень), либо в другую молекулу этого же субстрата [76];

3) мишень интеграции, в качестве которой используется либо несколько длинных фрагментов ДНК, содержащих последовательность для узнавания ДСД, либо плазмида, также содержащая специфический сайт (далее обозначен как *dbs* – DNA binding sequence) (рис. 1).

В том случае, если мишенью служат фрагменты ДНК, лишь один или несколько из них содержат *dbs*. И тогда используется субстрат, содержащий радиоактивную метку на 5'-конце интегрируемой цепи. Направленность интеграции можно оценить по процентному содержанию субстрата, встроенного во фрагмент (фрагменты), содержащий(е) *dbs*, относительно количества субстрата, встроенного во все фрагменты.

Если мишенью служит плазмида, распределение интеграции оценивают методом ПЦР (рис. 1). В этом случае используют такую систему, где матрицей служит продукт интеграции субстратной ДНК в плазмиду с *dbs*, один из праймеров комплементарен участку плазмиды на небольшом расстоянии от *dbs* (обычно 100–200 п.о.), а второй – субстрату и несет радиоактивную метку на 5'-конце. Можно использовать два типа матриц в качестве контрольных: либо плазмиду, не содержащую *dbs* после проведения в нее интеграции, либо плазмиду, содержащую *dbs*, но интеграция в которую проводилась интегразой дикого типа. Анализ продуктов ПЦР в денатурирующем ПААГ позволяет определить, есть ли в данной системе какие-

то изменения в распределении интеграции по сравнению с контрольной реакцией. Обычно отмечается снижение частоты интеграции непосредственно в сам сайт *dbs* (защита сайта за счет связывания с ним ДСД), в то время как частота интеграции вблизи него возрастает.

Методы определения специфичности интеграции в клеточный геном

Включение гибрида интегразы с ДСД в состав вирионов обычно проводят методом *in trans* [71, 72, 77, 78]. Этот подход основан на одновременной трансфекции клеток тремя плазмидами. Одна из них кодирует интегразу или гибридный белок на ее основе и дополнительно несет ген белка упаковки *vpr* на 5'-конце гена интегразы. Вторая кодирует геном вируса, в котором ген интегразы содержит мутацию, приводящую к замене D64V в каталитической триаде кодируемого белка и, как следствие, к полной потере ее каталитической активности. Также в геноме отсутствует последовательность, кодирующая белок оболочки Env, она заменена геном устойчивости к антибиотику. Белок оболочки закодирован в третьей плазмиде. Таким образом, если произойдет включение гибридной интегразы в вирионы, то будет восстановлена их жизнеспособность. Полученными вирусными частицами заражают культуру клеток и проводят селекцию антибиотиком, ген устойчивости к которому закодирован во второй плазмиде. Оценка количества колоний после селекции позволяет сравнить инфекционность вирионов,

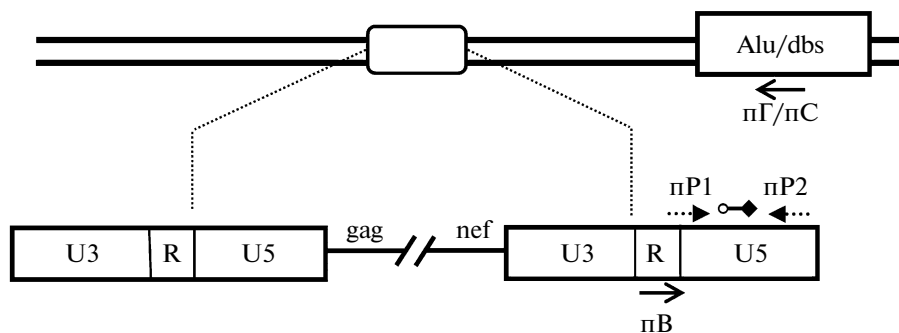


Рис. 2. Схематическое изображение экспериментальной системы для определения частоты интеграции в целом по геному. В ПЦР П1 используют пару праймеров пВ/пГ, что позволяет нормировать частоту интеграции в целом по геному (например, относительно Alu-островков). В ПЦР П2 используют пару праймеров пВ/пС — для определения частоты интеграции вблизи целевой последовательности *dfs*. В гнездовой ПЦР, Пг, используют пару праймеров пР1/пР2, а также флуоресцентный зонд, комплементарный амплифицируемому участку, что позволяет количественно оценить продукты реакций П1 и П2. Стрелка с ромбом — флуоресцентно-меченный зонд для наблюдения за накоплением продукта в Пг.

содержащих гибридную интегразу, с вирионами, содержащими интегразу дикого типа. Использование подхода *in trans* связано с тем, что прямое включение гибридной интегразы в геном приводит к потере инфекционной способности вируса [57]. Добавление дополнительного домена на С-конец интегразы может повлиять на фолдинг, упаковку или на процессинг полипептида Gag-Pol. Еще одна проблема заключается в том, что 3'-конец гена интегразы перекрывается с геном белка Vif [79].

Анализ интеграции провирусной ДНК в геном чаще всего проводят с помощью последовательной серии ПЦР (рис. 2). Очищенную геномную ДНК из инфицированных клеток обрабатывают ферментами рестрикции, после чего используют в качестве матрицы для первой ПЦР (П1), где один из праймеров (праймер пВ) комплементарен участку, находящемуся внутри последовательности провируса (обычно в районе LTR), а второй — комплементарен некоторому повторяющемуся генетическому элементу, например, Alu-островкам (праймер пГ). Этот эксперимент позволяет нормировать частоту интеграции в целом по геному. Далее с тем же препаратом геномной ДНК проводят другую ПЦР (П2), с парой праймеров пВ и пС, где пС — праймер, комплементарный участку последовательности, в район которой предполагается направить интеграцию. Выход специфического продукта как в случае П1, так и в случае П2 крайне низок, поэтому для количественного определения частоты интеграции проводят гнездовую ПЦР в реальном времени (Пг) с использованием пары праймеров, комплементарных участку внутри LTR ниже по цепи от сайта отжига праймера пВ. Статистическая обработка данных Пг позволяет определить частоту интеграции в район целевой последовательности по

сравнению с общей частотой интеграции. Подробное описание методики дано в работах [71] и [72].

Гибриды на основе интеграз и природных ДНК-связывающих белков

Как уже отмечалось выше, в пионерских работах по модификации ретровирусных интеграз с целью изменения их ДНК-связывающих свойств к интегразе присоединяли хорошо изученные ДСД из природных источников. Первая работа в этой области проведена Ф. Бушманом (Bushman) в 1994 г. на основе интегразы ВИЧ-1. В этой работе в качестве ДСД на N-конец интегразы был присоединен N-концевой домен репрессора фага λ (λR , аминокислотные остатки 1–133) [57]. λR состоит из двух независимых доменов, соединенных подвижным линкером, один из которых отвечает за связывание ДНК (N-концевой, остатки 1–92), а второй — за димеризацию репрессора (С-концевой, остатки 132–236) [80]. По данным [57], полученный гибридный белок сохранял каталитическую активность в реакциях *in vitro* на уровне интегразы дикого типа и прочно связывался с оператором фага λ . Анализ направленности интеграции с использованием набора фрагментов фаговой ДНК в качестве мишени показал, что гибридный белок чаще всего направлял интеграцию во фрагменты, содержащие оператор фага λ . Так была доказана принципиальная возможность изменения интеграционных характеристик интегразы путем присоединения к ней дополнительного ДНК-связывающего домена. Однако следует заметить, что включение модифицированной интегразы в состав вирусных частиц приводило к полной потере их инфекционности.

Впоследствии тот же ДСД (λR) был использован для получения гибридов с LEDGF/p75 [63]. Как уже отмечалось выше, модификация клеточно-

го партнера интегразы помогает избежать трудностей, возникающих при модификации самого фермента. Присоединение λR к N-концу LEDGF/p75 изменило распределение интеграции в системе, состоящей из ДНК фага λ , гибридного белка, интегразы ВИЧ-1 и ее ДНК-субстрата, и привело к тому, что частота интеграции в районе сайта оператора фага значительно повысилась. Таким образом, показано, что изменить направление интеграции можно также путем модификации клеточного партнера интегразы.

Некоторые исследователи при создании гибридов для направленной интеграции использовали ДНК-связывающий белок LexA *E. coli*. LexA состоит из двух доменов: N-концевого ДНК-связывающего (LexA binding domain, LABD, остатки 1–87) и C-концевого, отвечающего за димеризацию [81, 82]. Специфическим сайтом для LexA считается *recA* оператор. В серии работ [58, 77, 78] описаны гибриды на основе интегразы ВИЧ-1 и LexA. С целью проверить возможность изменения направления интеграции только с помощью ДНК-связывающего домена LexA создали конструкторы, в которых к C-концу интегразы присоединен либо полноразмерный белок LexA, либо LABD [58]. Так же, как и в работе [57], присоединение белковых доменов к интегразе не влияло на ее каталитическую активность *in vitro*. Оба гибридных белка сохраняли способность связывать последовательность *recA*, хотя LABD-содержащий гибрид связывал мишень слабее, чем гибрид с полноразмерным LexA. Специфичность интеграции, катализируемой этими гибридами, анализировали как на фрагментах ДНК, так и на плазмиде, и показали, что гибриды предпочтительно направляют интеграцию в район сайта *recA*, причем гибрид на основе полноразмерного белка LexA примерно в два раза эффективнее LABD-содержащего.

Описанные конструкторы на основе интегразы ВИЧ-1 и LexA включали в состав вирионов ВИЧ-1 методом *in trans* [77, 78]. Вирионы, содержащие гибриды интегразы, оказались жизнеспособными, но их инфекционность составляла примерно 50% от инфекционности вирионов на основе интегразы дикого типа. Каких-либо данных о направленности интеграции, катализируемой такими вирионами, в литературе не встречалось.

Также описаны гибриды белка LexA и интегразы ASLV [59]. В этом случае ДНК-связывающий домен присоединяли как к N-, так и к C-концу интегразы. Полученные конструкторы также сохраняли каталитическую активность и способность к интеграции в условиях *in vitro* в линейаризованную плазмиду, содержащую два сайта *recA*. Анализ продуктов интеграции методом ПЦР показал, что, также как и в случае гибридов LexA с интегразой ВИЧ-1 [58], гибриды на основе интегразы

ASLV направляют интеграцию вблизи сайта узнавания LexA, но не в сам сайт; при этом эффективность направленной интеграции не зависит от того, по N- или C-концу интегразы присоединен LexA.

В ряде работ предприняты попытки создания гибридов интеграз с природными белками, содержащими “цинковые пальцы”, например, Zif268 и Sp1. Так, гибрид на основе ДНК-связывающего домена мышинового транскрипционного фактора Zif268, присоединенного к C-концу интегразы ВИЧ-1, описан еще в 1997 году [60]. Предполагалось, что маленький размер домена (93 аминокислотных остатка [83]) не должен влиять на упаковку вирусных частиц. Линию клеток НЕК 293Т заражали полученными *in trans* вирусными частицами. Из инфицированных клеток выделяли предынтеграционный комплекс и проверяли его способность направлять интеграцию во фрагмент ДНК, содержащий сайт связывания Zif268. Повышенную частоту интеграции в районе 20 п.н. от сайта узнавания Zif268 действительно зафиксировали.

Наконец, известна работа по модификации интегразы MLV путем присоединения к ней природного белкового домена Sp1 [61], взаимодействующего с последовательностью (-GGGCGG-) [84, 85]. Белок Sp1 присоединяли к C-концу интегразы MLV в составе вириона, что незначительно снизило инфекционность получившихся мутантных вирусных частиц. Это важный результат, означающий, что присоединение небольшого белкового домена (60 аминокислотных остатков) к C-концу интегразы MLV, в отличие от интегразы ВИЧ-1 [57], не приводит к потере инфекционности вирусных частиц. К сожалению, в этой работе не проведен полногеномный анализ точек интеграции. Тем не менее, картирование сайтов интеграции пяти клонов, положительных на наличие сайта узнавания Sp1 вблизи места интеграции, выявило некоторую тенденцию в способности описанного конструктора специфически направлять интеграцию.

Следует отметить, что в работах с использованием доменов Zif268 и Sp1 достигнуть строгой специфичности интеграции так и не удалось. Одно из возможных объяснений следующее: для интеграции использованы гибриды, состоящие из двух белков, каждый из которых обладает сродством к определенным последовательностям ДНК; интеграза же в значительной степени связывает ДНК неспецифически, и “пересилить” эту неспецифичность может только ДСД с высоким сродством к своей мишени. Сродство белков, построенных из нескольких доменов “цинковые пальцы”, к ДНК возрастает с увеличением числа таких доменов в белке. Белок Sp1 включает всего два домена “цинковые пальцы”, а белок Zif268 — три. По-ви-

димому, двух–трех таких доменов недостаточно для формирования прочного ДНК-белкового комплекса, который будет существенно устойчивее неспецифических комплексов интегразы с ДНК.

Итак, в ряде научных работ показано, что присоединение к интегразе (или к ее клеточному партнеру) другого ДНК-связывающего белка приводит к изменению направления интеграции. Однако выбор белков, использованных в этих работах в качестве ДСД, не оптимален: ни у одного из них нет специфического сайта узнавания в геноме человека и, следовательно, невозможно их использование для направленной интеграции ДНК в геном человека. Таким образом, полученные конструкторы не пригодны для создания генно-терапевтических векторов.

Гибриды на основе синтетических белков, содержащих мотив “цинковые пальцы”

Как уже отмечалось, белки, связывающие ДНК за счет доменов “цинковые пальцы”, — очень привлекательные кандидаты на роль ДСД в гибридах с интегразой, поскольку они образуют специфичные комплексы с ДНК, но при этом не требуют мультимеризации. С точки зрения прочности и специфичности образуемых с ДНК комплексов особенно интересны синтетические белки, содержащие 6 доменов “цинковые пальцы”. Комбинируя домены, взаимодействующие с разными триплетами в ДНК, теоретически можно сконструировать белки, способные узнавать различные уникальные 18-нуклеотидные последовательности в геноме человека.

Авторы ряда работ [62, 71, 72] использовали синтетический шестидоменный белок E2C, специфически распознающий 18-нуклеотидную последовательность в 5'-нетранслируемой области гена *erbB-2*. Белок E2C был присоединен либо к N-, либо к C-концу интегразы ВИЧ-1 [62]. Как и в случае гибридов этой интегразы с другими ДСД [57, 58], присоединение E2C не отразилось на каталитической активности фермента *in vitro*. Полученные гибриды сохраняли способность специфично связывать сайт узнавания белка E2C и направлять интеграцию в плазмиду, содержащую этот сайт. Интересно, что более специфичным оказался гибрид, содержащий белок E2C на C-конце интегразы. Напомним, что в случае присоединения белка LexA к интегразе ASLV [59] не было замечено различий в специфичности направленной интеграции N-концевого и C-концевого гибридов.

Полученные методом *in trans* вирусные частицы с интегразой, содержащей E2C на C-конце [71], проявляли способность инфицировать культуру клеток человека. Полногеномный анализ

распределения интеграции в клетках, зараженных мутантными вирионами, показал, что частота интеграции вблизи последовательности *erbB-2* возрастает в 7–10 раз по сравнению с интеграцией, осуществляемой вирусом дикого типа.

Эти работы отличает ряд особенностей. Во-первых, роль ДСД в них выполняет белок, специфически связывающий определенную последовательность в геноме человека. Во-вторых, белок E2C содержит 6 доменов “цинковые пальцы” и связывает 18-звенную нуклеотидную последовательность, поэтому его комплекс с ДНК прочнее, чем в случае Zif268 (трехдоменный белок) или Sp1 (двухдоменный белок). Соответственно и специфичность интеграции, направляемой этим гибридом, выше. Таким образом, гибриды интегразы ВИЧ-1 с E2C можно рассматривать как гораздо более перспективные, чем те, которые описаны нами выше. Однако и в этих работах не решена такая важная проблема, как выбор безопасного участка интеграции, поскольку последовательность *erbB-2* — это часть транскрипционно-активного геномного локуса.

Недавно появилась работа, в которой представлен достаточно оригинальный подход к направленной интеграции [64]. Ген, кодирующий белок ZFD1, содержащий 6 доменов “цинковые пальцы”, встроен в большой набор случайных участков нуклеотидной последовательности *gag-pol* вируса MLV с использованием транспозонной системы [86]. Белок ZFD1, как и E2C, сконструирован искусственно для того, чтобы специфически распознавать 18-звенную последовательность в человеческом геноме — неподалеку от γ -глобинового локуса. Селекция полученной библиотеки клонов в культуре клеток позволила отобрать жизнеспособные и вирулентные штаммы, содержащие ген белка ZFD1 в составе *gag-pol*. Авторам удалось выделить 7 штаммов, которые содержали этот ген в двух разных положениях участка p12 гена *gag*, в двух положениях в начале гена обратной транскриптазы и, наконец, в трех положениях гена интегразы, а именно: в начальном участке N-концевого домена, в каталитическом домене и вблизи C-концевого домена. Анализ распределения интеграции, осуществленной отобранными гибридными вирионами, показал, что при встраивании гена белка ZFD1 в ген интегразы вируса MLV характерная для MLV частота интеграции в участки начала транскрипции (см. выше и таблица) значительно снизилась. Также, хоть и не столь значительно, снизилась интеграция в CpG-островки. Таким образом, удалось установить точные положения в последовательности гена *gag-pol*, в которые можно встраивать гены чужеродных белков и при этом не нарушать инфекционность вируса. Затем в три положения, определенные в гене интегразы, клонировали ген другого шестидоменного белка — ZFD2

[87], узнающего 18-звенную нуклеотидную последовательность в гене фактора пигментированного эпителия (pigment epithelium-derived factor, PEDF). При этом: а) вирусы сохраняли жизнеспособность и вирулентность; б) такие мутанты “уводили” интеграцию от сайтов инициации транскрипции – частота интеграции в районе ± 5 т.н. от сайтов инициации транскрипции падала в 4 раза; в) частота интеграции в CpG-островки была сильно снижена – в 30 раз в районе 1 т.н. от CpG островков. Кроме этих уникальных по своей значимости результатов, в этой работе точно указано местоположение тех участков гена *gag-pol* MLV, при встраивании в которые дополнительной последовательности не происходит потери жизнеспособности вируса.

Подводя итог, можно сказать, что на сегодняшний день активно развивается конструирование систем направленной интеграции на основе гибридов ретровирусных интеграз (преимущественно интегразы ВИЧ-1) и различных ДНК-связывающих белков, в частности содержащих “цинковые пальцы”. Разработаны и описаны подходы к включению гибридных белков в вирион, например метод *in trans* или метод множественного встраивания и последующей селекции. Таким образом, принципиальная возможность направленной интеграции уже показана, но, чтобы ее успешно применить, необходимо решить еще много задач. Основными из них, на наш взгляд, являются: 1) поиск или искусственное конструирование еще более специфичных ДНК-связывающих доменов; 2) выбор сайтов интеграции, достаточно удаленных от критически важных для жизнедеятельности клетки участков генома; 3) оптимизация методов включения гибридных интеграз в состав вирионов. Перспективным направлением остается также поиск и модификация клеточных кофакторов вирусных интеграз, как это сделано для кофактора интегразы ВИЧ-1 LEDGF/p75. Известно, что LEDGF может связываться с другими лентивирусными интегразами, такими как интегразы вирусов иммунодефицитов обезьян, кошек [88], что делает возможным построение систем направленной интеграции на их основе.

ПОХОДЫ К НАПРАВЛЕННОЙ ИНТЕГРАЦИИ НА ОСНОВЕ ТРАНСПОЗАЗ

Как уже отмечалось, транспозоны тоже могут быть использованы как векторы для доставки “терапевтических” генов. Транспозоны – это мобильные генетические элементы, которые могут перемещаться внутри хромосом или между вектором и хромосомой (см. обзоры [8–10]). Фермент транспозаза, кодируемый геном транспозона, относится к тому же классу ферментов, что и ретрови-

русная интеграза, – классу трансфераз. Транспозаза состоит из двух доменов: N-концевого, отвечающего за мультимеризацию, транспортировку в ядро и связывание транспозона, и C-концевого – каталитического. Она узнает специфические инвертированные повторы на концах транспозона и осуществляет его транспозицию – процесс вырезания и встраивания на новое место. Таким образом транспозоны, содержащие нужный ген между концевыми повторами, теоретически, могут быть использованы для генной терапии. Векторы на основе транспозонов состоят из двух компонентов: один содержит “терапевтический” ген, а другой кодирует транспозазу. Несмотря на низкую эффективность трансформации и низкий уровень экспрессии, некоторые из них находятся на стадии предклинических испытаний [89].

Транспозиция происходит в определенную последовательность ДНК, однако ее маленькая длина (всего 2–4 нуклеотида) не решает проблемы неспецифического встраивания и связанных с этим проблем [90]. Например, транспозоны из семейства *Tc1/mariner* встраиваются в динуклеотид TA [91, 92], а транспозон *piggyBac* – в четырехнуклеотидную последовательность TТАА [93, 94]. В отличие от интеграции ретровирусов каких либо значимых предпочтений при транспозиции не замечено. В ряде работ, однако, показано, что транспозон *Sleeping Beauty* склонен к транспозиции в экспрессируемые гены [90, 95], а транспозон *Mos1* – в гены, кодирующие рРНК в *C. elegans* [96], и в ген *cat* в *E. coli* [97]. Кроме того, отмечено, что транспозиция может увеличиваться в области метилированных CpG-островков [98]. Таким образом, создание системы для направленного встраивания ДНК на основе транспозаз столь же актуально, как и в случае ретровирусных интеграз.

На сегодняшний день в литературе описаны три основных подхода к направленной транспозиции (рис. 3). Первый из них аналогичен описанным выше моделям на основе ретровирусных интеграз и заключается в прямой модификации транспозазы путем присоединения к ней ДСД (рис. 3а). Вторым подходом предполагается модификацию самого транспозона (рис. 3б). В состав транспозона вносится специальная последовательность – “крюк”, узнаваемая специфическим белком КСД (“крюк”-связывающий домен), в то время как сам белок КСД входит в состав гибрида КСД-ДСД (подробно см. в обзоре [99]). Наконец, третий подход напоминает попытки модификации LEDGF/p75 как природного кофактора интегразы ВИЧ-1 [63] и предполагает присоединение ДСД к некоему белку, способному специфически связываться с транспозазой (ТСД – транспозазасвязывающий домен) и таким образом привлекать комплекс транспозазы с транспозоном в район запланированного места встраивания “терапевтиче-

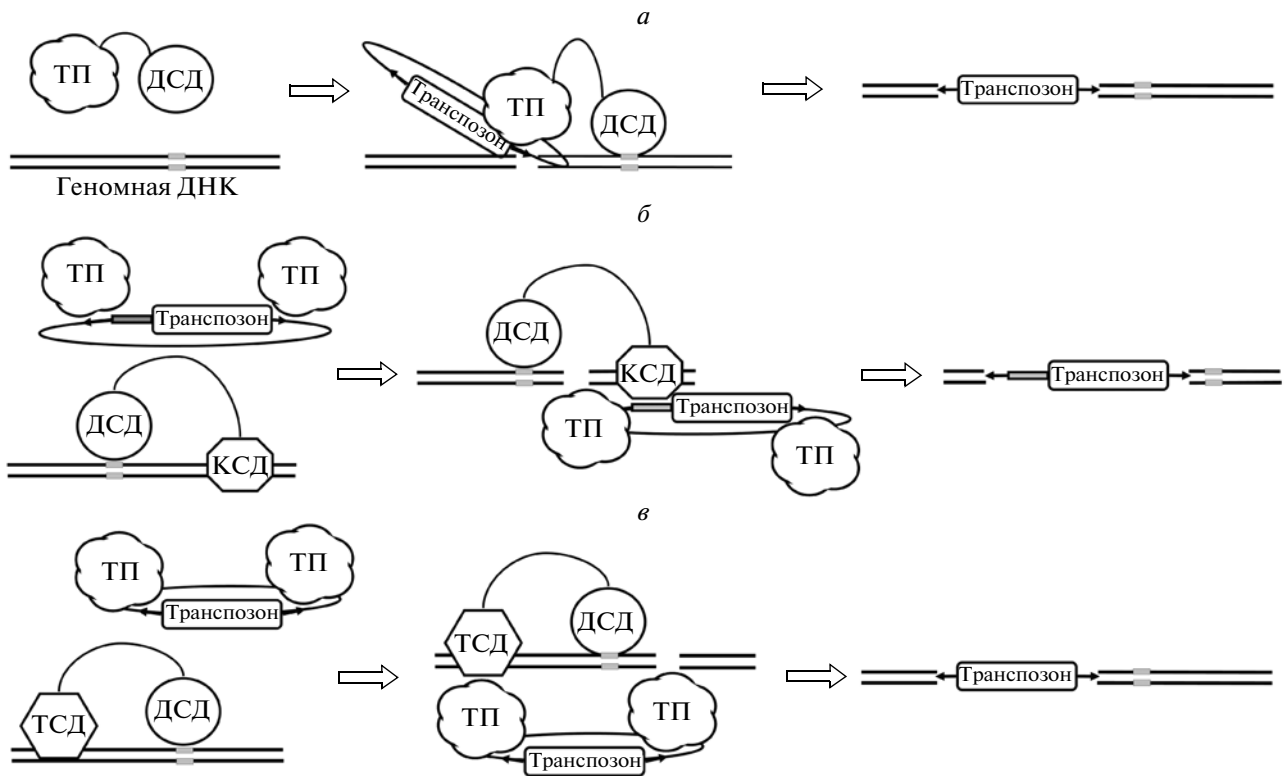


Рис. 3. Три подхода к направленной транспозиции. *а* – Модификация транспозазы; *б* – модификация транспозона; *в* – модификация белка, связывающего транспозазу (пояснения в тексте). В составе геномной ДНК серым цветом выделена (=) последовательность *dbS*; ТП – транспозаза; ДСД – ДНК-связывающий домен; КСД – “крюк”-связывающий домен; ТСД – транспозаза-связывающий домен.

ского” гена (рис. 3*в*). Далее мы подробно рассмотрим преимущества и недостатки каждого из этих подходов, а также способы изменения направления транспозиции. Однако стоит сразу отметить, что подход, основанный на прямой модификации транспозазы, хотя и встречается в литературе наиболее часто, представляется наименее перспективным, так как во многих случаях присоединение ДНК-связывающего белка к транспозазе значительно снижает или приводит к полной потере ее каталитической активности. Степень снижения каталитической активности транспозазы зависит, во-первых, от природы присоединяемого к ней ДСД; во-вторых, от транспозазы, используемой в эксперименте; и, в-третьих, от длины и состава аминокислотного линкера, который соединяет транспозазу с ДСД.

Метод определения специфичности транспозиции

Для анализа направленности транспозиции применяется метод интерплазмидного анализа [91], включающий использование системы из трех плазмидных векторов. Один из них, называ-

емый донорным, содержит последовательность транспозона. Вторым, акцепторным, содержит *dbS*. Третий вектор кодирует транспозазу или гибридные белки и называется хелперным. При этом в состав последовательности транспозона включен ген устойчивости к одному антибиотику (*aX*), а акцепторная плаزمид содержит ген устойчивости к другому антибиотику (*aY*). Все три вектора совместно вводят в клетки, на которых проводится исследование направленности транспозиции, после чего проводят селекцию сначала антибиотиком *aY*, а затем *aX*. Таким способом удастся отобрать только те клетки, в которых прошел полноценный перенос транспозона в акцепторную плазмиду. Затем акцепторную плазмиду выделяют, и распределение транспозиции в ней оценивают методом ПЦР, аналогично тому, как это показано на рис. 1.

Модификация транспозазы для повышения специфичности транспозиции

В ряде работ описано присоединение к различным транспозазам ДНК-связывающих белков с целью придания транспозазам новых ДНК-свя-

зывающих характеристик. В большинстве случаев в качестве объекта модификации использована транспозаза *Sleeping Beauty* (SB) из семейства *Tc1/mariner*, так как транспозон SB — один из наиболее изученных транспозонов, и механизм его транспозиции хорошо описан [100].

На основе транспозазы SB созданы гибридные белки с доменами “цинковые пальцы” (E2C, Jazz, Sp1), а также другими ДНК-связывающими белками (TetR, Gal4). В ряде работ, описывающих прямую модификацию транспозазы SB [101–103], показано, что “посадка” белкового домена на С-конец транспозазы полностью лишает ее активности в реакции транспозиции, а на N-конец — значительно снижает активность (от 25 до 90%). Такие выводы сделаны для белков Sp1 [101], E2C [102, 103] и Jazz [102]. Вместе с тем, присоединение синтетического белка ZNF202, содержащего “цинковые пальцы”, не влияло на активность транспозазы SB — она осталась на уровне природного фермента [101]. В данном случае для повышения каталитической активности модифицированного фермента использовали так называемую гиперактивную форму транспозона (pT3) и активную форму транспозазы (SB12).

Данные о влиянии длины и структуры линкера, соединяющего транспозазу и ДСД, на активность гибридного белка очень противоречивы. С одной стороны, не замечено влияния длины линкера на активность гибридов транспозазы SB12 и белка ZNF202 [101]. С другой стороны, при детальном исследовании свойств гибрида состава E2C-linker-SB выявлено, что оптимальная активность комплекса достигается при длине 15 аминокислотных остатков в линкере (Gly-Gly-Ser)_n, т.е. при $n = 5$ [102]. Наконец, согласно данным работы [103], присоединение к SB-транспозазе белка E2C через линкер G10 (10 остатков глицина) приводит к полной потере ферментативной активности в реакции транспозиции, а присоединение через такой же линкер белка Jazz снижает активность гибридного белка на 90% относительно природной транспозазы.

Результаты, полученные интерплазмидным методом оценки специфичности транспозиции, свидетельствуют о том, что конструкция E2C-(GGS)₅-SB направляет транспозицию во фрагменты, содержащие соответствующий сайт узнавания, в 11 раз, а конструкция Gal4-(GGS)₅-SB в 8 раз эффективнее, чем немодифицированная SB-транспозаза. Однако при исследовании направленности транспозиции гибридом E2C-(GGS)₅-SB в геноме человека, где домен E2C узнает последовательность 5'-нетранслируемой области гена *erbB-2* в хромосоме 17 (см. выше), повышенного встраивания в район *erbB-2* не зафиксировано [102]. Гибрид E2C-(G)₁₀-SB оказался вообще не

способным к направленной транспозиции — даже в интерплазмидной системе [103].

Транспозон Mos1, как и SB, относится к семейству *Tc1/mariner*, и сайтом интеграции для него также служит динуклеотид TA. Гибрид на основе транспозазы Mos1 и белка Gal4 анализировали в интерплазмидной системе, инъекционной в эмбрионы *Aedes aegypti*. Домен Gal4 специфически узнает последовательность UAS (upstream activating sequence, 5'-CGGAGTACTGTCCCTCCG-3'). Транспозиция гибридом Mos1-Gal4 в плазмиду, содержащую UAS-последовательность, происходила почти в 13 раз эффективнее, чем в плазмиду, не содержащую этого сайта узнавания для белка Gal4 [104]. Более того, авторы особо отмечают, что 96% случаев интеграции, катализируемой белком Mos1-Gal4, происходило в динуклеотид TA, расположенный на расстоянии 954 п.н. от последовательности-мишени UAS. Эти результаты позволяют выявить определенные изменения в распределении участков транспозиции в плазмиде, содержащей UAS, относительно картины транспозиции в плазмиде, не содержащей сайта узнавания Gal4.

Аналогичная конструкция создана на основе транспозазы piggyBac: piggybac-Gal4. В этом случае интерплазмидный анализ в эмбрионах *Aedes aegypti* выявил почти 12-кратное увеличение частоты интеграции в UAS-содержащую плазмиду, причем 67% транспозиций происходило в сайт, расположенный на расстоянии 912 п.н. от UAS [104]. В работе [105] эта же система проверена на культуре клеток *Bombyx mori* (тутовый шелкопряд) и *Drosophila*. В клетках *B. mori* транспозиция в плазмиду, содержащую последовательность UAS, возрастала в 4 раза, а в клетках плодовой мушки — в 7.5 раз. Таким образом, система piggyBac-Gal4/UAS может считаться пригодной для направленной транспозиции в клетках насекомых, но ее эффективность определяется организмом — источником клеток. Предполагается, что на специфичность транспозиции влияют клеточные факторы, взаимодействующие с транспозазой, но экспериментально такая гипотеза не подтверждена.

Транспозон ISY100 также относится к семейству *Tc1/mariner* [92]. Авторы статьи [106] проанализировали 101 сайт интеграции, катализируемой транспозазой ISY100, и выявили консенсусную нуклеотидную последовательность, наиболее предпочтительную для встраивания транспозона: 5'-ADWTAWHT-3' (где W = A или T, D ≠ C и H ≠ G). Кроме того авторы попытались оценить влияние линкера между транспозазой ISY100 и белком Zif268 на каталитическую активность гибридных белков. Получено 3 гибрида с длиной линкера 9 (L1), 12 (L2) и 45 (L3) аминокислот. Линкеры L1

и L2 снижали каталитическую активность транспозазы в 3 раза по сравнению с диким типом, а линкер L3 – в 9 раз. Оценку эффективности всех гибридов проводили интерплазмидным методом на искусственной мишени, моделирующей оптимальную последовательность для транспозиции, – [TANN]₉, которая была клонирована в плазмиду рядом с сайтом узнавания Zif268. В этой мишени сайт TA повторяется девять раз, но в различном окружении, подобранном с учетом указанной выше консенсусной последовательности. При оценке направленности встраивания транспозона показано, что только гибрид с длинным линкером, L3, направляет транспозицию в область, расположенную вблизи мишени белка Zif268, – в область [TANN]₉.

В отличие от других описанных здесь гибридов конструктор на основе транспозазы бактериального транспозона IS30 и репрессора фага λ [107], присоединенного к ее С-концу, оказался каталитически активным и осуществлял транспозицию на уровне транспозазы дикого типа. 55% транспозиций, катализируемых таким гибридом в интерплазмидной системе, происходит на расстоянии, не превышающем 400 п.н. от сайта OR (оператор фага λ). Исследование проводили на эмбрионах рыбок данио-рерио (*Danio rerio*) и на культуре клеток человека HeLa, в которых эта транспозаза прокариот также активна. Аналогично для гибрида IS30-Gli1 (ДСД транскрипционного фактора Gli1) показано, что 50% транспозиций происходит в пределах 100 п.н. от сайта узнавания Gli1 [107].

Повышение специфичности транспозиции путем модификации транспозона

Как показано выше, прямая модификация транспозазы SB сильно снижает ее каталитическую активность. В работе [103] предложен альтернативный подход к направленной транспозиции – модификация клеточного ДСД путем присоединения к нему белкового домена, узнающего определенную нуклеотидную последовательность внутри транспозона (последовательность-“крюк”) (рис. 3б). Этот белковый домен назвали “крюк”-связывающим (КСД). При разработке указанного подхода создали гибриды на основе консервативного домена (SAF-box) белка Scaffold attachment factor A (SAF-A), участвующего в прикреплении ДНК к ядерному матриксу в так называемых MAR-областях (matrix attachment regions), и белка LexA из *E. coli*, выступающего в качестве КСД. Аналогичным образом сконструирован гибрид состава TetR/NLS/LexA, узнающий последовательность тетрациклинового оператора [103]. В транспозон ввели специфический сайт узнавания для белка LexA и показали, что его присутствие не влияет на эффективность транспозиции. Анализ

транспозиции проводился *in vivo*, в культуре клеток HeLa. При анализе сайтов транспозиции гибрида SAF/NLS/LexA с использованием методов статистической обработки показано, что гибрид может изменять распределение транспозиции для SB и, более того, повышать частоту транспозиции вблизи MAR-областей. Последовательность тетрациклинового оператора вводили в геном клеток HeLa, после чего анализировали распределение транспозиции в таких клетках гибридом TetR/NLS/LexA. Вблизи области тетрациклинового оператора обнаружили два сайта, по которым прошла транспозиция. Таким образом, подход, основанный на модификации транспозона, можно рассматривать как потенциально эффективный для направленной транспозиции.

Повышение специфичности транспозиции за счет модификации белка, связывающего транспозазу

На модели транспозазы SB разрабатывали еще один подход к направленной транспозиции – путем модификации белка, специфически связывающегося с транспозазой (рис. 3в). Известно, что N-концевой участок транспозазы SB (N-57) отвечает за ее межсубъединичные взаимодействия, а совместная экспрессия полной и укороченной формы транспозазы не влияет на эффективность транспозиции [108]. Прежде всего, сконструировали плазмиду, кодирующую гибридный белок (TetR/NLS/linker/N-57), состоящий из N-концевого участка транспозазы N-57 (в качестве ТСД) и белка TetR (в качестве ДСД) [103]. Проанализировав сайты встраивания транспозона в интерплазмидной системе, включающей четыре плазмиды, идентифицировали 12 “горячих точек” транспозиции, из них наибольшее количество транспозиций пришлось на сайт, расположенный на расстоянии 44 п.н. от последовательности узнавания тетрациклинового оператора TetR, а остальные оказались разбросанными в пределах 2600 п.н. от нее. Таким образом, модификация белков, взаимодействующих с транспозазой, может занять достойное место в развитии методов направленной транспозиции.

Итак, на сегодняшний день цель большинства исследований, проводившихся по направленной транспозиции, – получение гибридных белков на основе транспозазы с целью модификации ее ДНК-связывающих свойств. Однако этот подход кажется малоперспективным, поскольку присоединение к большинству транспозаз дополнительного белкового домена чаще всего приводит к частичной или полной потере каталитической активности фермента. Присоединение ДСД на С-конец транспозазы оказалось невозможным во всех случаях, кроме транспозазы IS30. Присоединение ДСД на N-конец транспозазы в ряде случа-

ев давало активный гибридный белок, особенно, если авторам удавалось подобрать подходящий линкер.

Повышение специфичности транспозиции путем создания конструкторов на основе ДСД и белка, специфически связывающего транспозазу, — задача не из простых, так как в этом случае требуется поиск белков, образующих прочные комплексы с транспозазой. Тем не менее, этот подход — неплохая альтернатива методам непосредственной модификации транспозазы.

Модификация транспозона путем введения в его состав последовательности-“крюка”, связываемой специфическим белком, который, в свою очередь, присоединен к ДСД, представляется наиболее перспективным направлением, так как, во-первых, не затрагивает каталитическую активность транспозазы, а во-вторых, не ограничивает свободу в выборе сайта встраивания ДНК. К сожалению, на сегодняшний день этот подход опробован лишь на одной системе.

ЗАКЛЮЧЕНИЕ

Из проведенного нами анализа существующих на сегодняшний день подходов для направленной интеграции ДНК можно сделать следующий вывод: все разрабатываемые методологии основаны на создании гибридных белков, один из которых — фермент, осуществляющий интеграцию, а другой — белок, способный специфично связывать ДНК. Наиболее изучены системы на основе ферментов из семейства трансфераз: транспозаз и ретровирусных интеграз. Результаты описанных здесь работ показывают, что положение ДНК-связывающего белка в случае интеграз не имеет значения для каталитической активности гибридного белка, но может влиять на специфичность интеграции. В случае же транспозаз большинство гибридных белков, содержащих ДНК-связывающий белок на С-конце транспозазы, теряют каталитическую активность.

Активность гибридов на основе транспозаз также зависит от длины линкера, соединяющего два компонента гибридного белка. Но и тут не всегда наблюдается корреляция. В одном из исследований оптимальным оказался 15-аминокислотный линкер, в других — хорошо “работали” линкеры и иных размеров. В случае же гибридов на основе интеграз специальных исследований по оптимизации длины и последовательности линкера не проводили. Почти во всех случаях линкер состоял из двух аминокислот, кодируемых сайтом рестрикции, необходимым для генно-инженерных манипуляций.

Модификация клеточного кофактора интегразы ВИЧ-1 — белка LEDGF/p75 — позволила избе-

жать проблемы падения инфекционности вириона. В связи с этим создание гибридных белков на основе не самих ферментов, осуществляющих интеграцию, а их белковых кофакторов — очень перспективное направление. Однако развитие этой методологии ограничено тем, что на сегодняшний день белок LEDGF/p75 — единственный описанный белковый кофактор ретровирусных интеграз, чья роль в процессе интеграции экспериментально доказана. А для транспозаз неизвестно ни одного природного кофактора.

Оригинальный подход к направленной интеграции транспозонов основан на модификации последовательности самого транспозона путем внесения в нее уникальной нуклеотидной последовательности-“крюка”. По-видимому, аналогичную модификацию можно применить и для направленной интеграции ретровирусов, однако таких работ пока нет.

Стоит отметить, что, независимо от используемого подхода, принципиальный момент при создании конструкций для направленной интеграции — это выбор сайта интеграции в геноме. И это отнюдь не простая задача, так как сайт должен удовлетворять двум очевидным критериям: он должен располагаться в активно транскрибируемом участке генома, однако интеграция в него не должна нарушать генетических систем, важных для жизнедеятельности клетки и организма в целом. Ни в одной из опубликованных на настоящий момент работ не предложен сайт интеграции, соответствующий перечисленным критериям. Тем не менее, работы в этой области активно ведутся, что позволяет надеяться на появление в ближайшее время “терапевтических” конструкций, созданных для направленного и безопасного встраивания в геном человека.

СПИСОК ЛИТЕРАТУРЫ

1. Young L.S., Searle P.F., Onion D., Mautner V. 2006. Viral gene therapy strategies: from basic science to clinical application. *J. Pathol.* **208**(2), 299–318.
2. Pluta K., Kasprzak M.M. 2009. Use of HIV as a gene transfer vector. *Acta Biochim. Pol.* **56**, 531–595.
3. www.wiley.co.uk/genmed/clinical/.
4. Hacein-Bey-Abina S., von Kalle C., Schmidt M., et al. 2003. LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science.* **302**, 415–419.
5. Deichmann A., Hacein-Bey-Albina S., Schmidt M., et al. 2007. Vector integration is nonrandom and clustered and influences the fate of lymphopoiesis in SCID-X1 gene therapy. *J. Clin. Invest.* **117**, 2225–2232.
6. Hacein-Bey-Abina S., Garrigue A., Wang G.P., et al. 2008. Insertional oncogenesis in 4 patients after retro-

- virus-mediated gene therapy of SCID-X1. *J. Clin. Invest.* **118**, 3132–3142.
7. Holmes-Son M.L., Appa R.S., Chow S.A. 2001. Molecular genetics and target site specificity of retroviral integration. *Adv. Genet.* **43**, 33–69.
 8. Delaurière L., Chénais B., Hardivillier Y., Gauvry L., Casse N. 2009. *Mariner* transposons as genetic tools in vertebrate cells. *Genetica.* **137**, 9–17.
 9. Izsvák Z., Chuah M.K., Vandendriessche T., Ivics Z. 2009. Efficient stable gene transfer into human cells by the *Sleeping Beauty* transposon vectors. *Methods.* **49**(3), 287–297.
 10. Hackett P.B., Largaespada D.A., Cooper L.J. 2010. A transposon and transposase system for human application. *Mol. Ther.* **18**, 674–683.
 11. Ivics Z., Izsvák Z. 2006. Transposons for gene therapy! *Curr. Gene Ther.* **6**, 593–607.
 12. Engelman A., Cherepanov P. 2008. The lentiviral integrase binding protein LEDGF/p75 and HIV-1 replication. *PLoS Pathog.* **4**, e1000046.
 13. Brown P.O. 1990. Integration of retroviral DNA. *Curr. Top. Microbiol. Immunol.* **157**, 19–48.
 14. Hindmarsh P., Leis J. 1999. Retroviral DNA integration. *Microbiol. Mol. Biol. Rev.* **63**, 836–843.
 15. Агапкина Ю.Ю., Приказчикова Т.А., Смолов М.А., Готтих М.Б. 2005. Структура и функции интегразы ВИЧ-1. *Успехи биол. химии.* **45**, 87–122.
 16. Спиринов П.В., Вильгельм А.Э., Прасолов В.С. 2008. Лентивирусные векторы. *Молекуляр. биология.* **42**, 913–926.
 17. Pruss D., Bushman F.D., Wolffe A.P. 1994. Human immunodeficiency virus integrase directs integration to sites of severe DNA distortion within the nucleosome core. *Proc. Natl. Acad. Sci. USA.* **91**, 5913–5917.
 18. Pruss D., Reeves R., Busman F.D., Wolffe A.P. 1994. The influence of DNA and nucleosome structure on integration events directed by HIV-1 integrase. *J. Biol. Chem.* **269**, 25031–25041.
 19. Bor Y.C., Bushman F.D., Orgel L.E. 1995. *In vitro* integration of human immunodeficiency virus type 1 cDNA into targets containing protein-induced bends. *Proc. Natl. Acad. Sci. USA.* **92**, 10334–10338.
 20. Muller H.-P., Varmus H.E. 1994. DNA bending creates favored sites for retroviral integration: an explanation for preferred insertion sites in nucleosomes. *EMBO J.* **13**, 4704–4714.
 21. Pryciak P.M., Varmus H.E. 1992. Nucleosomes, DNA-binding proteins and DNA sequence modulate retroviral integration target site selection. *Cell.* **69**, 769–780.
 22. Pryciak P.M., Sil A., Varmus H.E. 1992. Retroviral integration into minichromosomes *in vitro*. *EMBO J.* **11**, 291–303.
 23. Scherdin U., Rhodes K., Breindl M. 1990. Transcriptionally Active Genome Regions Are Preferred Targets for Retrovirus Integration. *J. Virol.* **64**, 907–991.
 24. Kitamura Y., Lee Y.M., Coffin J.M. 1992. Nonrandom integration of retroviral DNA *in vitro*: effect of CpG methylation. *Proc. Natl. Acad. Sci. USA.* **89**, 5532–5536.
 25. Taganov K.D., Cuesta I., Daniel R., Cirillo L.A., Katz R.A., Zaret K.S., Skalka A.M. 2004. Integrase-specific enhancement and suppression of retroviral DNA integration by compacted chromatin structure *in vitro*. *J. Virol.* **78**, 5848–5855.
 26. Rohdewohld H., Weiher H., Reik W., Jaenisch R., Breindl M. 1987. Retrovirus integration and chromatin structure: Moloney murine leukemia proviral integration sites map near DNaseI-hypersensitive sites. *J. Virol.* **61**, 336–343.
 27. Stevens S.W., Griffith J.D. 1996. Sequence analysis of the human DNA flanking sites of human immunodeficiency virus type 1 integration. *J. Virol.* **70**, 6459–6462.
 28. Lewinski M.K., Yamashita M., Emerman M., et al. 2006. Retroviral DNA integration: viral and cellular determinants of target-site selection. *PLoS Pathogens.* **2**, e60.
 29. Schroeder A.R.W., Shinn P., Chen H., Berry C., Ecker J.R., Bushman F. 2002. HIV-1 integration in the human genome favors active genes and local hotspots. *Cell.* **110**, 521–529.
 30. Bushman F., Lewinski M., Ciuffi A., Barr S., Leipzig J., Hannenhalli S., Hoffmann C. 2005. Genome wide analysis of retroviral DNA integration. *Nat. Rev. Microbiol.* **3**, 848–856.
 31. Wu X., Li Y., Crise B., Burgess S.M. 2003. Transcription start regions in the human genome are favored targets for MLV integration. *Science.* **300**, 1749–1751.
 32. Felice B., Cattoglio C., Cittaro D., et al. 2009. Transcription factor binding sites are genetic determinants of retroviral integration in the human genome. *PLoS One.* **4**, e4571.
 33. Barr S.D., Leipzig J., Shinn P., Ecker J.R., Bushman F.D. 2005. Integration targeting by avian sarcoma leucosis virus and human immunodeficiency virus in the chicken genome. *J. Virol.* **79**, 12035–12044.
 34. Mitchell R.S., Beitzel B.F., Schroeder A.R., et al. 2004. Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. *PLoS Biol.* **2**(8), e234.
 35. Han Y., Lassen K., Monie D., et al. 2004. Resting CD4+ T cells from human immunodeficiency virus type 1 (HIV-1)-infected individuals carry integrated HIV-1 genomes within actively transcribed host genes. *J. Virol.* **78**, 6122–6133.
 36. Elleder D., Pavlíček A., Paces J., Hejnar J. 2002. Preferential integration of human immunodeficiency virus type 1 into genes, cytogenetic R bands and GC-rich DNA regions: insight from the human genome sequence. *FEBS Lett.* **517**, 285–286.
 37. Crise B., Li Y., Yuan C., Morcock D.R., et al. 2005. Simian immunodeficiency virus integration preference is similar to that of human immunodeficiency virus type 1. *J. Virol.* **79**, 12199–12204.

38. Hematti P., Hong B.K., Ferguson C., et al. 2004. Distinct genomic integration of MLV and SIV vectors in primate hematopoietic stem and progenitor cells. *PLoS Biol.* **2**, 2183–2190.
39. Kang Y., Moressi C.J., Scheetz T.E., et al. 2006. Integration site choice of a feline immunodeficiency virus vector. *J. Virol.* **80**, 8820–8823.
40. de Palma M., Montini E., Santoni de Sio F.R., Benedicenti F., Gentile A., Medico E., Naldini L. 2005. Promoter trapping reveals significant differences in integration site selection between MLV and HIV vectors in primary hematopoietic cells. *Gen. Ther.* **105**, 2307–2315.
41. Desfarges S., Ciuffi A. 2010. Retroviral integration site selection. *Viruses.* **2**, 111–130.
42. Hacein-Bey H., Cavazzana-Calvo M., Le Deist F., et al. 1996. Gamma-c gene transfer into SCID-X1 patients' B-cell lines restores normal high-affinity interleukin-2 receptor expression and function. *Blood.* **87**, 3108–3116.
43. Thrasher A.J., Gaspar H.B., Baum C., et al. 2006. Gene therapy: X-SCID transgene leukaemogenicity. *Nature.* **443**, E5–6; discussion E6–7.
44. Schwarzwaelder K., Howe S.J., Schmidt M., et al. 2007. Gammaretrovirus-mediated correction of SCID-X1 is associated with skewed vector integration site distribution *in vivo*. *J. Clin. Invest.* **117**, 2241–2249.
45. Wahlers A., Kustikova O., Zipfel P.F., et al. 2002. Upstream conserved sequences of mouse leukemia viruses are important for high transgene expression in lymphoid and hematopoietic cells. *Mol. Ther.* **6**, 313–320.
46. Studamire B., Goff S.P. 2008. Host proteins interacting with the Moloney murine leukemia virus integrase: multiple transcriptional regulators and chromatin binding factors. *Retrovirology.* **5**, 48.
47. van Maele B., Debyzer Z. 2005. HIV-1 integration: an interplay between HIV-1 integrase, cellular and viral proteins. *AIDS Rev.* **7**, 26–43.
48. van Maele B., Busschots K., Vandekerckhove L., Christ F., Debyzer Z. 2006. Cellular co-factors of HIV-1 integration. *Trends. Biochem. Sci.* **31**, 98–105.
49. Cherepanov P., Maertens G., Proost P., et al. 2003. HIV-1 integrase forms stable tetramers and associates with LEDGF/p75 protein in human cells. *J. Biol. Chem.* **278**, 372–381.
50. Turlure F., Maertens G., Rahman S., Cherepanov P., Engelman A. 2006. A tripartite DNA-binding element, comprised of the nuclear localization signal and two AT-hook motifs, mediates the association of LEDGF/p75 with chromatin *in vivo*. *Nucl. Acids Res.* **34**, 1653–1665.
51. Maertens G., Cherepanov P., Plumers W., Busschots K., de Clercq E., Debyzer Z., Engelborghs Y. 2003. LEDGF/p75 is essential for nuclear and chromosomal targeting of HIV-1 integrase in human cells. *J. Biol. Chem.* **278**, 33528–33539.
52. Ciuffi A., Llano M., Poeschla E., et al. 2005. A role for LEDGF/p75 in targeting HIV DNA integration. *Nat. Med.* **11**, 1287–1289.
53. Shun M.C., Raghavendra N.K., Vandegraaf N., et al. 2007. LEDGF/p75 functions downstream from preintegration complex formation to effect gene-specific HIV-1 integration. *Genes Dev.* **21**, 1767–1778.
54. Llano M., Saenz D.T., Meehan A., et al. 2006. An essential role for LEDGF/p75 in HIV-1 integration. *Science.* **314**, 461–464.
55. Vandekerckhove L., Christ F., van Maele B., et al. 2006. Transient and stable knockdown of the integrase cofactor LEDGF/p75 reveals its role in the replication cycle of human immunodeficiency virus. *J. Virol.* **80**, 1886–1896.
56. Marshall H.M., Ronen K., Berry C., et al. 2007. Role of PSIP1/LEDGF/p75 in lentiviral infectivity and integration targeting. *PLoS ONE.* **2**, e1340.
57. Bushman F.D. 1994. Tethering human immunodeficiency virus 1 integrase to a DNA site directs integration to nearby sequences. *Proc. Natl. Acad. Sci. USA.* **91**(20), 9233–9237.
58. Goulaouic H., Chow S.A. 1996. Directed integration of viral DNA mediated by fusion proteins consisting of human immunodeficiency virus type 1 integrase and *Escherichia coli* LexA protein. *J. Virol.* **70**, 37–46.
59. Katz R.A., Merkel G., Skalka A.M. 1996. Targeting of retroviral integrase by fusion to a heterologous DNA binding domain: *in vitro* activities and incorporation of a fusion protein into viral particles. *Virology.* **217**, 178–190.
60. Bushman F.D., Miller M.D. 1997. Tethering human immunodeficiency virus type 1 preintegration complexes to target DNA promotes integration at nearby sites. *J. Virol.* **71**, 458–464.
61. Peng W.-J., Chang C.-M., Lin T.-H. 2002. Target integration by a chimeric Sp1 zinc finger domain-Moloney murine leukemia virus integrase *in vivo*. *J. Biomed. Sci.* **9**, 171–184.
62. Tan W., Zhu K., Segal D.J., Barbas C.F., 3rd, Chow S.A. 2004. Fusion proteins consisting of human immunodeficiency virus type 1 integrase and the designed polydactyl zinc finger protein E2C direct integration of viral DNA into specific sites. *J. Virol.* **78**, 1301–1313.
63. Ciuffi A., Diamond T.L., Hwang Y., Marshall H.M., Bushman F.D. 2006. Modulating target site selection during human immunodeficiency virus DNA integration *in vitro* with an engineered tethering factor. *Hum. Gene Ther.* **17**, 960–967.
64. Lim K., Klimczak R., Yu J.H., Schaffer D.V. 2010. Specific insertions of zinc finger domains into Gag-Pol yield engineered retroviral vectors with selective integration properties. *Proc. Natl. Acad. Sci. USA.* **107**, 12475–12480.
65. Urnov F.D., Rebar E.J., Holmes M.C., Zhang H.S., Gregory P.D. 2010. Genome editing with engineered zinc finger nucleases. *Nat. Rev. Genet.* **11**, 636–646.
66. Klug A. 2010. The discovery of zinc fingers and their development for practical applications in gene regula-

- tion and genome manipulation. *Q. Rev. Biophys.* **43**, 1–21.
67. Sera T. 2009. Zinc-finger-based artificial transcription factors and their applications. *Adv. Drug. Deliv. Rev.* **61**, 513–526.
 68. Segal D.J., Barbas C.F. 3rd. 2001. Custom DNA-binding proteins come of age: polydactyl zinc-finger proteins. *Curr. Opin. Biotechnol.* **12**, 632–637.
 69. Beerli R.R., Barbas C.F. 3rd. 2002. Engineering polydactyl zinc-finger transcription factors. *Nat. Biotechnol.* **20**, 135–141.
 70. Wu H., Yang W.P., Barbas C.F. 3rd. 1995. Building zinc fingers by selection: toward a therapeutic application. *Proc. Natl. Acad. Sci. USA.* **92**, 344–348.
 71. Tan W., Dong Z., Wilkinson T.A., Barbas C.F. 3rd, Chow S.A. 2006. Human immunodeficiency virus type 1 incorporated with fusion proteins consisting of integrase and the designed polydactyl zinc finger protein E2C can bias integration of viral DNA into a predetermined chromosomal region in human cells. *J. Virol.* **80**, 1939–1948.
 72. Su K., Wang D., Ye J., Kim Y.C., Chow S.A. 2009. Site-specific integration of retroviral DNA in human cells using fusion proteins consisting of human immunodeficiency virus type 1 integrase and the designed polydactyl zinc-finger protein E2C. *Methods.* **47**, 269–276.
 73. Beerli R.R., Segal D.J., Dreier B., Barbas C.F. 3rd. 1998. Toward controlling gene expression at will: specific regulation of the *erbB-2/HER-2* promoter by using polydactyl zinc finger proteins constructed from modular building blocks. *Proc. Natl. Acad. Sci. USA.* **95**, 14628–14633.
 74. Beerli R.R., Dreier B., Barbas C.F. 3rd. 2000. Positive and negative regulation of endogenous genes by designed transcription factors. *Proc. Natl. Acad. Sci. USA.* **97**, 1495–1500.
 75. Chow S.A. 1997. *In vitro* assays for activities of retroviral integrase. *Methods.* **12**, 306–317.
 76. Bushman F.D., Fujiwara T., Craigie R. 1990. Retroviral DNA integration directed by HIV integration protein *in vitro*. *Science.* **249**, 1555–1558.
 77. Holmes-Son M.L., Chow S.A. 2000. Integrase-lexA fusion proteins incorporated into human immunodeficiency virus type 1 that contains a catalytically inactive integrase gene are functional to mediate integration. *J. Virol.* **74**, 11548–11556.
 78. Holmes-Son M.L., Chow S.A. 2002. Correct integration mediated by integrase-LexA fusion proteins incorporated into HIV-1. *Mol. Ther.* **5**, 360–370.
 79. Purcell D.F., Martin M.A. 1993. Alternative splicing of human immunodeficiency virus type 1 mRNA modulates viral protein expression, replication and infectivity. *J. Virol.* **67**, 6365–6378.
 80. Pabo C.O., Sauer R.T., Sturtevant J.M., Ptashne M. 1979. The lambda repressor contains two domains. *Proc. Natl. Acad. Sci. USA.* **76**, 1608–1612.
 81. Little J.W., Mount D.W. 1982. The SOS regulatory system of *Escherichia coli*. *Cell.* **29**, 11–22.
 82. Schnarr M., Oertel-Buchheit P., Kazmaier M., Granger-Schnarr M. 1991. DNA binding properties of the Lex-A repressor. *Biochemie.* **73**, 423–431.
 83. Pavletich N.P., Pabo C.O. 1991. Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Å. *Science.* **252**, 809–817.
 84. Mitchell P.J., Tjia R. 1989. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science.* **245**, 371–378.
 85. Gill G., Pascal E., Tseng Z.H., Tjian R. 1994. A glutamine-rich hydrophobic patch in transcription factor Sp1 contacts the dTAFII110 component of the Drosophila TFIID complex and mediates transcriptional activation. *Proc. Natl. Acad. Sci. USA.* **91**, 192–196.
 86. Yu J.H., Schaffer D.V. 2006. High-throughput, library-based selection of a murine leukemia virus variant to infect nondividing cells. *J. Virol.* **80**, 8981–8988.
 87. Tan S., Guschin D., Davalos A., et al. 2003. Zinc-finger protein-targeted gene regulation: genome-wide single-gene specificity. *Proc. Natl. Acad. Sci. USA.* **100**, 11997–12002.
 88. Busschots K., Vercammen J., Emiliani S., Benarous R., Engelborghs Y., Christ F., Debyser Z. 2005. The interaction of LEDGF/p75 with integrase is lentivirus-specific and promotes DNA binding. *J. Biol. Chem.* **280**, 17841–17847.
 89. Nakanishi H., Higuchi Y., Kawakami S., Yamashita F., Hashida M. 2009. Development and therapeutic application of transposon-based vectors. *Yakugaku Zasshi.* **129**, 1433–1443.
 90. Vigdal T.J., Kaufman C.D., Izsvák Z., Voytas D.F., Ivics Z. 2002. Common physical properties of DNA affecting target site selection of *Sleeping Beauty* and other Tc1/*mariner* transposable elements. *J. Mol. Biol.* **323**, 441–452.
 91. Ivics Z., Hackett P.B., Plasterk R.H., Izsvák Z. 1997. Molecular reconstruction of *Sleeping Beauty*, a Tc1-like transposon from fish, and its transposition in human cells. *Cell.* **91**, 501–510.
 92. Plasterk R.H., Izsvák Z., Ivics Z. 1999. Resident aliens: the Tc1/*mariner* superfamily of transposable elements. *Trends Genet.* **15**, 326–332.
 93. Fraser M.J., Cary L., Boonvisudi K., Wang H.G. 1995. Assay for movement of Lepidopteran transposon IFP2 in insect cells using a baculovirus genome as a target DNA. *Virology.* **211**, 397–407.
 94. Fraser M.J., Ciszczon T., Elick T., Bauser C. 1996. Precise excision of TTAA-specific lepidopteran transposons piggyBac (IFP2) and tagalong (TFP3) from the baculovirus genome in cell lines from two species of Lepidoptera. *Insect. Mol. Biol.* **5**, 141–151.
 95. Walisko O., Schorn A., Rolfs F., Devaraj A., Miskey C., Izsvák Z., Ivics Z. 2008. Transcriptional activities of the *Sleeping Beauty* transposon and shielding its genetic cargo with insulators. *Mol. Ther.* **16**, 359–369.

96. Granger L., Martin E., Ségalat L. 2004. Mos as a tool for genome-wide insertional mutagenesis in *Caenorhabditis elegans*: results of a pilot study. *Nucl. Acids Res.* **32**, e117.
97. Crénès G., Ivo D., Hérisson J., Dion S., Renault S., Bigot Y., Petit A. 2009. The bacterial Tn9 chloramphenicol resistance gene: an attractive DNA segment for Mos1 *mariner* insertions. *Mol. Genet. Genomics.* **281**, 315–328.
98. Yusa K., Takeda J., Horie K. 2004. Enhancement of *Sleeping Beauty* transposition by CpG methylation: possible role of heterochromatin formation. *Mol. Cell. Biol.* **24**, 4004–4018.
99. Demattei M.V., Thomas X., Carnus E., Augé-Gouillou C., Renault S. 2010. Site-directed integration of transgenes: transposons revisited using DNA-binding-domain technologies. *Genetica.* **138**, 531–540.
100. Izsvák Z., Ivics Z. 2004. *Sleeping beauty* transposition: biology and application for molecular therapy. *Mol. Ther.* **9**, 147–156.
101. Wilson M.H., Kaminski J.M., George A.L. Jr. 2005. Functional zinc finger/*sleeping beauty* transposase chimeras exhibit attenuated overproduction inhibition. *FEBS Lett.* **579**, 6205–6209.
102. Yant S.R., Huang Y., Akache B., Kay M.A. 2007. Site-directed transposon integration in human cells. *Nucleic Acids Res.* **35**, e50.
103. Ivics Z., Katzer A., Stuwe E.E., Fielder D., Knespel S., Izsvák Z. 2007. Targeted *Sleeping Beauty* Transposition in Human Cells. *Mol. Ther.* **15**, 1137–1144.
104. Maragathavally K.J., Kaminski J.M., Coates C.J. 2006. Chimeric Mos1 and piggyBac transposases result in site-directed integration. *FASEB J.* **20**, 1188–1195.
105. Wang N., Jiang C.Y., Jiang M.X., Zhang C.X., Cheng J.A. 2010. Using chimeric piggyBac transposase to achieve directed interplasmid transposition in silkworm *Bombyx mori* and fruit fly *Drosophila* cells. *J. Zhejiang Univ. Sci. B.* **11**, 728–734.
106. Feng X., Bednarz A.L., Colloms S.D. 2009. Precise targeted integration by a chimaeric transposase zinc-finger fusion protein. *Nucl. Acids Res.* **38**, 1204–1216.
107. Szabó M., Müller F., Kiss J., Balduf C., Strähle U., Olsz F. 2003. Transposition and targeting of the prokaryotic mobile element IS30 in zebrafish. *FEBS Lett.* **550**, 46–50.
108. Izsvák Z., Khare D., Behlke J., Heinemann U., Plasterk R.H., Ivics Z. 2002. Involvement of a bifunctional, paired-like DNA-binding domain and a transpositional enhancer in *Sleeping Beauty* transposition. *J. Biol. Chem.* **277**, 34581–34588.