Available online at www.sciencedirect.com
ScienceDirect
Journal of Biotechnology xxx (2008) xxx–xxx
Journal of BIOTECHNOLOGY
www.elsevier.com/locate/jbiotec

Human genomic site-specific recombination catalyzed by coliphage HK022 integrase

Gali Harel-Levi [1], Janna Goltsman [1], Chen Nahum Josef Haim Tuby [1], Ezra Yagil, Mikhail Kolot *

Department of Biochemistry, Tel-Aviv University, Tel Aviv 69978, Israel

Received 4 June 2007; received in revised form 6 December 2007; accepted 8 January 2008

Abstract

It has been previously demonstrated that the wild type integrase (Int) protein of coliphage HK022 can catalyze site-specific recombination in human cells between attachment (*att*) sites that were placed on extrachromosomal plasmids. In the present report it is shown that Int can catalyze the site-specific recombination reactions in a human cell culture on the chromosomal level. These include integrative (*attP* × *attB*) as well as excisive (*attL* × *attR*) reactions each in two configurations. In the *cis* configuration both sites are on the same chromosome, in the *trans* configuration one site is on a chromosome and the other on an episome. The reactions in *cis* were observed without any selection force, using the green fluorescent protein (GFP) as a reporter. The reactions in *trans* could be detected only when a selection force was applied, using the hygromycin-resistant (Hyg^R) phenotype as a selective marker. All reactions were catalyzed without the need to supply any of the accessory proteins that are required by Int in its *Escherichia coli* host. The versatility of the *att* sites may be an advantage in the utilization of Int to integrate plasmid DNA into the genome, followed by a partial exclusion of the integrated plasmid.

Keywords: Site-specific recombination; Integrase; Coliphage HK022; Human genome

1. Introduction

Site-specific recombinases have become an important tool in eukaryotic gene manipulations. To date Cre of coliphage P1, Flp of yeast and the integrase of the *Streptomyces* phage ϕC31 (Int-ϕC31) are the leading ones (Sorrell and Kolb, 2005). Cre and Flp are the simpler ones because they recombine between small and identical pairs of DNA targets (*loxP* and *FRT*, respectively), each 34 base pairs (bp) long. Chromosomal integration of plasmid DNA catalyzed by Cre or Flp is unstable because it is flanked by two identical target sites that are substrate of the more favourable reverse excision reaction (Glaser et al., 2005; Nagy, 2000). The integrative target sites of Int-ϕC31 (*attP* and *attB*) that are 39 bp and 34 bp long, respectively, show considerable diversity. An *attP* × *attB* integrative reaction of Int-ϕC31 leads to the recombinant *attL* and *attR* sites that flank the integrated DNA. The reverse excision reaction (*attR* × *attL*), possibly requiring an unknown accessory factor, does not occur and hence integrated DNA remains stable (Calos, 2006; Groth and Calos, 2004).

The integrase protein of coliphage HK022 (Int-HK022) recognizes four different sites of recombination, which makes it another potential tool for site-specific gene manipulations in the eukarya. Similar to its well-studied relative coliphage λ, the phage *attP* target site of Int-HK022 is 225 bp long and the *attB* site on its host DNA (*Escherichia coli*) is only 21 bp long. The integration reaction (*attP* × *attB*) leads to the recombinant *attL* and *attR* sites that flank the integrated prophage. In *E. coli* as well as *in vitro* Int catalysis of the integrative recombination reaction requires the host-encoded integration host factor (IHF) as an accessory DNA-binding protein. In addition to IHF, the reverse excision reaction (*attL* × *attR*) also requires Excisionase (Xis), a phage-encoded binding protein (Azaro and Landy, 2001; Weisberg et al., 1999). However, when the wild type

Abbreviations: Amp, ampicillin; bp, base pairs; Ganc, Ganciclovir; GFP, green fluorescent protein; Hyg^R, hygromycin resistant; IHF, integration host factor; Int, integrase; Neo^R, neomycin resistant; CMV, cytomegalovirus promoter; RMCE, recombination mediated cassette exchange; TK, thymidine kinase; Xis, Excisionase.

* Corresponding author. Tel.: +972 3 6409823; fax: +972 3 6406834.
E-mail address: kolott@post.tau.ac.il (M. Kolot).
[1] These authors contributed equally to this paper.

0168-1656/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.jbiotec.2008.01.002

Int-HK022 is expressed in mammalian cells it can catalyze both integrative and excisive reactions without the need to supply any of the prokaryotic accessory proteins. In the case of λ, only IHF-independent mutants of Int can catalyze the recombination reactions in mammalian cells (Christ et al., 2002). Thus far, the relatively efficient Int-HK022 catalyzed reactions in mammalian cell cultures were only demonstrated using extrachromosomal plasmids (Kolot et al., 2003). In order for Int to be useful for gene manipulations in the mammals it is essential that it catalyzes the recombination reactions when at least one site is located on a chromosome. Here we present experiments showing that the wild type Int-HK022 can catalyze in human cells all possible combinations of the site-specific recombination reactions either when both *att* sites are on the same chromosome, or when one *att* is on a chromosome and the other on an episomal plasmid. By the use of a combined set of wild type *att* sites Int can also catalyze partial removal of an integrated plasmid.

2. Materials and methods

2.1. Cells, growth conditions, transfection, selection

Cells (~4×10^5) of the human embryonic kidney cell lines 293 or 293T were grown and 24 h later transfected with 1–5 μg DNA of each circular plasmid using the calcium phosphate method in a 6-wells culture plate as described before (Kolot et al., 1999). For stable integrations neomycin was added at a final concentration of 1 mg/ml, hygromycin at 150 μg/ml or Ganciclovir at 3 μM, all were added 72 h post-transfection. *E. coli* cells transformed with plasmids were grown in Luria-Bertani medium with the appropriate antibiotics.

2.2. Plasmids and oligomers

Plasmids, and oligomers that were used as primers in PCR reactions, are described in Table 1A and B, respectively.

2.3. DNA techniques

2.3.1. DNA extraction

Plasmid DNA from *E. coli* cultures was prepared using a HiYield™ plasmid mini kit (RBC) or a GenElute™ HP plasmid Maxiprep kit (Sigma). Plasmid DNA from human cells was extracted by the method of Hirt (Hirt, 1967). Genomic DNA was extracted from the human cells using an AccuPrep™ genomic DNA extraction kit (BIONEER).

2.3.2. Plasmid constructions

To construct plasmid pGH786 the *attL* fragment of plasmid pMK25 (Gottfried et al., 2000) was cloned between the BamHI and EcoRI sites of pcDNA3 (Invitrogen). To construct plasmid pGH790 a PCR-generated *attB* fragment was cloned into the BamHI and XhoI sites of plasmid pcDNA3. To construct plasmid pCT909 a PCR-generated CMV-GFP fragment was cloned into the EcoRI and XhoI sites of plasmid pcDNA3.1+. Next, a PCR-generated fragment of *attP* was cloned between the BamHI and EcoRI sites, a PCR-generated *attR* fragment was cloned between the MluI and HindIII sites, a HindIII-BamHI fragment that carries the herpes simplex virus thymidine kinase (TK) gene was extracted from plasmid XpPNT (Tybulewicz et al., 1991) and cloned into the same sites and finally, a PCR-generated fragment that carries the CMV promoter was cloned into the KpnI and BamHI sites.

To construct the plasmid pNS490 the SmaI-BamHI blunted Hyg^R fragment from plasmid p3'SS (Accession number U42371) was cloned into the SmaI site of plasmid Bluescript (SK) obtaining plasmid pNS425. Subsequently the AvaI-NotI Hyg^R fragment of pNS425 was cloned between the AgeI-NotI sites of pMK223 (Kolot et al., 2003) replacing the GFP gene. To construct plasmid pGH625 the Hyg^R fragment from plasmid pNS425 was cloned between the AgeI and NotI sites of pMK242 (Kolot et al., 2003). In plasmids pGH784, pGH785 the *attP*-Stop fragment and the *attR*-Stop fragment extracted from plasmids pMK210 and pMK185 (Kolot et al., 2003), respectively, were cloned between the KpnI and XhoI sites of pGH625.

2.3.3. In vitro site-specific recombination

Reactions were as previously described (Gottfried et al., 2005).

2.3.4. Southern blots

Five to ten micrograms of human cells DNA were cut with the appropriate enzyme(s) and separated on 1–2% agarose gel. Alkaline transfer and hybridization were done according to the "GeneScreen" manual (DuPont). The probes were PCR fragments that were ^{32}P-labeled by the random primer method.

2.3.5. DNA sequencing

DNA sequences were obtained using an ABI PRISM® 3100 Genetic Analyzer (Applied Biosystems and Hitachi).

2.3.6. Fluorescent-activated cell sorting (FACS) analysis

~2×10^6 cells from 1 well of a 6-well plate were collected following trypsin treatment of which 10^4 cells were selected by the FACS sorter (Becton Dickinson Instrument) for fluorescent measurements. Data analysis was performed using the WinMDI2.8 program.

3. Results

Int-HK022 catalysis of integrative (*attP* × *attB*) and the excisive (*attL* × *attR*) recombination, each in two configurations, was tested on the chromosomal level in human cell culture. In the *cis* configurations both recombining sites were located in a tandem orientation on a chromosomally integrated plasmid. In the *trans* configurations one site was located on a chromosome and the other was on a plasmid (episome).

3.1. Recombination reactions in cis

The gene encoding the green fluorescent protein (GFP) was used as the reporter for the *cis*-recombination reactions. For the *attP* × *attB* reaction in *cis* the neomycin-resistant

Please cite this article in press as: Harel-Levi, G., et al., Human genomic site-specific recombination catalyzed by coliphage HK022 integrase, J. Biotechnol. (2008), doi:10.1016/j.jbiotec.2008.01.002

Table 1
List of plasmids and oligomers that were used as primers for the PCR reactions

A. Plasmid	Relevant genotype	Use	Source
pcDNA3	Neo^R oriSV40 vector	Cloning vector	Invitrogen
pcDNA3.1(+)	Neo^R oriSV40 vector	Cloning vector	Invitrogen
pEGFPN1	Neo^R EGFP-N1	Cloning vector	Clontech
pMK52	Int-HK022	Int expression	Kolot et al. (1999)
pMK218	pCMV-*attP*-Stop-*attB*-GFP Neo^R on pEGFP-N1	Chromosomal *cis* reaction	Kolot et al. (2003)
pMK189	pCMV-*attR*-Stop-*attL*-GFP Neo^R on pEGFP-N1	Chromosomal *cis* reaction	Kolot et al. (2003)
pMK221	pCMV-*attP* on pCDNA3	Episome in *trans* reaction	Kolot et al. (2003)
pGH790	pCMV-*attB* on pCDNA3	Episome in *trans* reaction	This work
pAM243	pCMV-*attR* on pCDNA3	Episome in *trans* reaction	Kolot et al. (2003)
pGH786	pCMV-*attL* on pCDNA3	Episome in *trans* reaction	This work
pNS490	Stop-*attB*-Hyg Neo^R on pEGFP-N1	Chromosomal *trans* reaction	This work
pGH784	Stop-*attP*-Hyg Neo^R on pEGFP-N1	Chromosomal *trans* reaction	This work
pGH625	Stop-*attL*-Hyg Neo^R on pEGFP-N1	Chromosomal *trans* reaction	This work
pGH785	Stop-*attR*-Hyg Neo^R on pEGFP-N1	Chromosomal *trans* reaction	This work
pCT909	*attR-TK*-pCMV-*attP*-pCMV-GFP on pcDNA3.1+	Chromosomal *trans* reaction	This work

B. Primers	Sequence	Location
f (forward)	GCCAGGGTTTTCCCAGTCACGA	Bluescript
r (reverse)	GAGCGGATAACAATTTCACACAGG	Bluescript
oEY133	CAAATGATTTTATTTTGACTAATAAGACCCT	P′ arm of *attP*
oEY134	GAAATGATTTTATTTGACCTAATAATGACCT	P′ arm of *attP*
oEY135	AGGTCACTAATACTATCTAAGTAGTTGATTCATAGGACCTGG	P arm of *attP*
oEY204	ATTGACGTCAATGGGAGTTTGTTTTGGC	CMV
oEY206	CGTCGCCGTCCAGCTCGACCAG	GFP
oEY253	GCACGAGATTCTTCGCCCTCCG	Hyg^R
oEY256	TATCCATGGTGAGCAAGGGCGAGG	GFP
oEY257	ATCCATGGTCGTGCTGCTTCATGTGG	GFP
oEY302	CCCAATGTCAAGCACTTCCGGAATC	Hyg^R
oEY331	AATCCGCTTTGTCGACTCAACC	*attB*
oEY389	CGTCATTATTGACGTCAATGGGCGGGGG	CMV-GFP
oEY398	GGGAATAAGGGCGACACGGAAATGTTG	Amp

(Neo^R) plasmid pMK218 that carries the CMV-*attP*-Stop-*attB*-GFP sequence (Kolot et al., 2003) was randomly inserted by neomycin selection into the genome of the human cell line 293 [Fig. 1A(a), bold *att* sites]. CMV is the promoter of the human cytomegalovirus and "Stop" is a transcription terminator that prevents the expression of GFP (Sauer, 1993). For the *attR* × *attL* reaction in *cis* a similar plasmid (pMK189) that carries the CMV-*attR*-Stop-*attL*-GFP sequence (Kolot et al., 2003) was likewise randomly inserted into the human genome [Fig. 1A(a), plain *att* sites]. In both cases, Int-catalyzed recombination is expected to evict the *att*-Stop-*att* fragment [Fig. 1A(c)] leaving a recombined *att* site between the CMV promoter and the GFP open reading frame thereby allowing the expression of GFP [Fig. 1A(b)].

Cells of each transgenic line were transfected with an Int-expressing plasmid (pMK52) and, as control, cells were also transfected with the vector (pcDNA3). Two days later the cells were collected and underwent a FACS analysis. Visually, of four tested *attP*-Stop-*attB* lines two showed fluorescence in the presence of Int. Of three tested *attR*-Stop-*attL* lines all showed fluorescence in the presence of Int. Quantitative FACS results of one line of each reaction (Fig. 1B) show that in the presence of Int, the integrative reaction (*attP* × *attB*) yielded 5.3% fluorescent cells and the excisive reaction (*attR* × *attL*) yielded 11.4% of fluorescent cells. In the absence of Int the fluorescence was negligible. Positive control cells with a GFP-expressing plasmid were always included and showed 80–90% fluorescence (not shown). Fig. 1C shows results of PCRs using genomic DNA of the parental cells and of fluorescent cells as templates, the primers are depicted as bent arrows in Fig. 1A. The PCR products agreed with the expected sizes of the substrates in the parental cells (S in Fig. 1C lanes 1) and appeared also in the fluorescent cells along with the expected products (P) (Fig. 1C, lanes 2). The sequence of the PCR recombination products (P) in each of the two reactions (not shown) has confirmed the expected recombinant *att* site [*attL* in the *attP* × *attB* reaction and *attB* in the *attR* × *attL* reaction [Fig. 1A(b)], indicating that each has resulted from a site-specific recombination reaction. To demonstrate that the reaction was on the chromosomal level we performed Southern blot analyses. Fluorescent cells from each reaction were enriched by FACS sorting and their DNA as well as DNA from the original unrecombined parental cells was cut with EcoRI in the *attP* × *attB* integrative reaction and with EcoRI and NotI in the *attR* × *attL* excisive reaction (Fig. 1A, black triangles). The probe for the integration reaction was a 230 bp PCR product of *attP* and for the excision reaction a 250 bp PCR product that represents an N-terminal sequence of the GFP gene (thin solid lines in Fig. 1A). Fig. 1D, lanes 4 and 5 show the results of the blots of the unrecombined cells and the fluorescent recombinants, respectively, in the *attP* × *attB* and *attR* × *attL* reactions. As positive controls we included in the blots the original plas-

Please cite this article in press as: Harel-Levi, G., et al., Human genomic site-specific recombination catalyzed by coliphage HK022 integrase, J. Biotechnol. (2008), doi:10.1016/j.jbiotec.2008.01.002

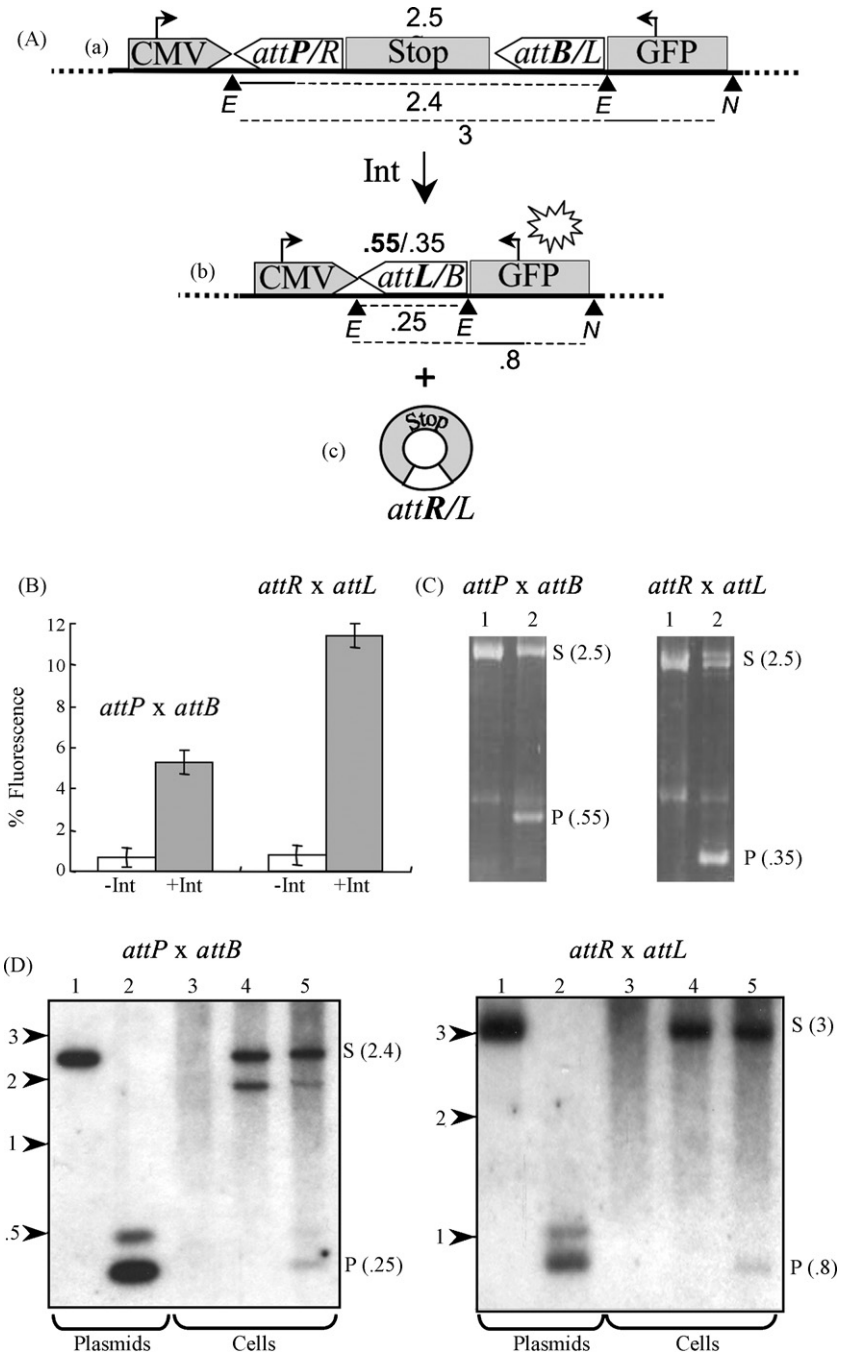

Fig. 1. Recombination reactions in *cis*. (A) Schematic presentation of the recombination reactions. (a) Chromosomal loci of the *attP* × *attB* reaction (in bold) and the *attR* × *attL* reaction. (b) and (c) Expected recombination products. Bent arrows depict the primers used in the PCR reactions. Triangles show the location of the restriction sites (*E* = EcoRI, *N* = NotI) that were used in the Southern blots. Solid thin lines depict the labeled probes and their dotted extensions depict the expected sizes (in kb) of the visible fragments in the Southern blots. (B) FACS analyses, mean of three experiments of the integrative reaction (a) and of the excisive reaction (b). (C) Gels of PCR products of the *attP* × *attB* reaction and of the *attR* × *attL* reaction. Lanes 1 are DNA of the original unrecombined *att* lines and lanes 2 of the fluorescent recombinants. S: substrate, P: product. Primers used were oEY204 + oEY206. (D) Results of Southern blots. (a) *attP* × *attB* reaction, the probe (*attP*) was generated with primers oEY133 + oEY135. (b) *attR* × *attL* reaction. The probe (GFP) was generated with primers oEY256 + oEY257. Lanes 1 are plasmids that were used to create the chromosomal *att* sites (plasmids pMK218 and pMK189, respectively). Lanes 2 are the relevant recombination products of these plasmids generated by *in vitro* site-specific recombination reactions. Lanes 3 are the wild type control cells. Lanes 4 are the relevant chromosomal *att* lines. Lanes 5 are the fluorescent recombinant lines. Arrows indicate molecular weight markers in kb.

mids that were used to create the transgenic lines (lanes 1) and plasmids that carried their recombination products generated by *in vitro* reactions (lanes 2), all cut with the same restriction enzymes. As negative control we used cut genomic DNA of wild type cells (lanes 3). The size of the substrates (S) in the original unrecombined *att* cells (lanes 4) and in the relevant original plasmids (lanes 1) was identical, and so were the products (P) in lanes 5 and 2. Though there are some unexplained additional

Please cite this article in press as: Harel-Levi, G., et al., Human genomic site-specific recombination catalyzed by coliphage HK022 integrase, J. Biotechnol. (2008), doi:10.1016/j.jbiotec.2008.01.002

bands in these blots, the emergence of the expected recombination products in the fluorescent cells along with PCR results confirms that Int has catalyzed both site-specific recombination reactions on both chromosomally located pairs of the *att* sites. The FACS sorting has resulted in about 50% fluorescent cells, nevertheless the signal of the recombination products in lanes 5 was rather weak. We assume that the reason is because the original transgenic lines, that were transformed by the calcium phosphate method, carried more than one copy of the substrate (Wurtele et al., 2003).

3.2. Recombination reactions in trans

For the recombination reactions in *trans* we have tried a similar non-selective approach using the GFP as a reporter. With this reporter, we have earlier demonstrated extrachromosomal Int-promoted integrative and excisive recombination between episomal plasmids (Kolot et al., 2003). However, when one *att* site was first inserted into a chromosome and the other remained as an episome we could not detect any significant Int-promoted fluorescence (not shown). This lead us to conclude that the chromosomal reactions in *trans*, should they occur, are not sufficiently frequent to be detected without a selection force. Therefore, we have constructed a selective system in which only cells that underwent the sought recombination reaction would acquire resistance to hygromycin (Hyg^R). For each reaction (*attP* × *attB* and *attR* × *attL* in *trans*) there are two reciprocal possibilities, one site on the chromosome, the other on the episome and vice versa, all together four reactions as depicted in Table 2 lines (a) and (b). The *att* sites of the four episomal plasmids were placed downstream to the CMV promoter [Fig. 2A(a)]. The four chromosomal plasmids that were inserted into the genome carried each the compatible *att* site flanked upstream with the transcription terminator (Stop) and downstream with the promoter-less open reading frame of the Hyg^R gene [Fig. 2A(b)]. Recombination between the two *att* sites is expected to result in the integration of the episomal plasmid into the chromosome, placing the CMV promoter upstream to the Hyg^R gene [Fig. 2A(c)], thereby facilitating the expression of the Hyg^R gene.

Human 293 cells were stably transfected by neomycin selection with each of the plasmids assigned for the chromosomal insertion [Table 2, line (b)]. Several independent isolates of each of the four types of stable Neo^R *att* cell lines [Table 2, line (c)] were each co-transfected with the Int-expressing plasmid and the episomal plasmid that carried the compatible *att* site [Table 2, line (a)]. Three days later the cultures were supplied with hygromycin. As shown in line (d) of Table 2, only 1–2 out of the 11–16 Neo^R clones of each type of cross generated Hyg^R colonies. One of each of the parental cells that yielded Hyg^R recombinant was cotransfected again with its relevant compatible *att* plasmid and as controls, each line was also transfected with the relevant *att* plasmid without the Int plasmid. No colonies appeared without Int and the number of Hyg^R recombinants obtained in the presence of Int is depicted in Table 2, line (e). Chromosomal DNA extracted from 7 to 10 of these Hyg^R isolates from each of the four

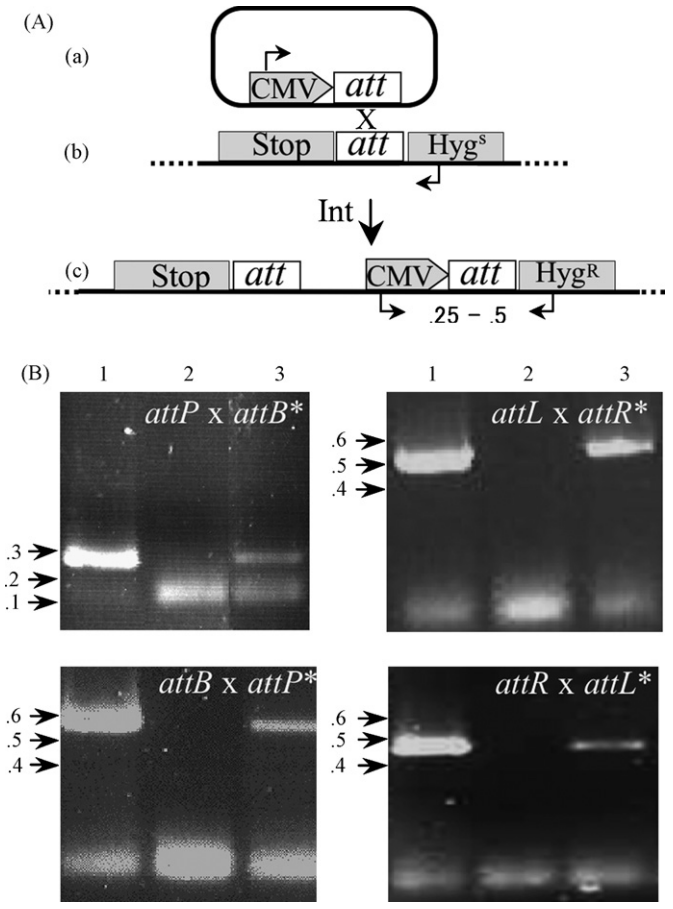

Fig. 2. PCRs of the four *trans* reactions. (A) General schematic presentation of the four *trans* reaction. (a) Loci on plasmid. (b) Loci on chromosome. (c) Expected product. The orientation of the *att* sites in each cross is depicted in Fig. 3 (A–D). Bent arrows depict the primers used in the PCR reactions which were oEY253 + oEY204. (B) Gels showing the PCR products of the four types of Hyg^R recombinants (lanes 3), of the four types of the parental *att* sites as negative controls (lanes 2) and of cells transiently transformed with both *att* plasmids as positive controls (lanes 1). *Indicates chromosomal *att* site and the arrows show the position of size markers in kb.

crosses served as template for PCR reactions, the primer used (Fig. 2A) are excepted to yield PCR products only if these Hyg^R isolates have resulted from a site-specific recombination reaction. Line (f) of Table 2 shows that some of the Hyg^R recombinants yielded positive PCR results. Fig. 2B, lanes 3 show the expected PCR products in each of the four types of Hyg^R recombinants. As negative controls the same PCR reactions were performed (Fig. 2B, lanes 2) using as templates chromosomal DNA from the original chromosomal *att* parents depicted in Fig. 2A(b). As positive controls (Fig. 2B, lanes 1) the DNA templates were extrachromosomal DNA of cells transiently co-transformed with the Int-expressing and both *att* plasmids as previously reported (Kolot et al., 1999). The bottom lines in the PCRs in Fig. 2B were residual free primers. The positive PCR products that were obtained from the four Hyg^R clones (lanes 3) were sequenced and each showed the expected recombinant *att* sites flanked by CMV and Hyg sequences (not shown). This suggests that the PCR-positive Hyg^R clones have resulted from the Int-promoted site-specific recombination reac-

Please cite this article in press as: Harel-Levi, G., et al., Human genomic site-specific recombination catalyzed by coliphage HK022 integrase, J. Biotechnol. (2008), doi:10.1016/j.jbiotec.2008.01.002

Table 2
Number of clones analyzed or obtained in different stages of the reactions in *trans*

	1	2	3	4
(a) Site on episomal plasmid (and plasmid number)	*attP* (pMK221)	*attB* (pGH790)	*attR* (pAM243)	*attL* (pGH786)
(b) Site on chromosome (and plasmid number)	*attB* (pNS490)	*attP* (pGH784)	*attL* (pGH625)	*attR* (pGH785)
(c) Number of Neo^R PCR positive lines tested	14	11	16	13
(d) Number of Neo^R lines that generated Hyg^R recombinants	2/14	1/11	1/16	1/13
(e) Number of Int-dependent Hyg^R colonies in 2nd transformation	16	7	12	10
(f) Number of PCR positive Hyg^R isolates/number tested	10/10	5/7	7/10	2/10

tions in *trans*, between the episomal and the chromosomal *att* sites.

To confirm that all four types of these site-specific recombination reactions in *trans* took place on the chromosomal level, we conducted Southern blot analyses using template DNA from one of each type of the PCR-positive Hyg^R recombinants. Panels A-D in Fig. 3 explain schematically the expected results of the blots. The upper lines (S) represent the relevant loci of the unrecombined parental Hyg^S *att* lines prior to recombination and the lower lines (P) depict the same for the expected Hyg^R recombinants. The diagrams also show the relevant restriction sites (triangles and letters) that were used to cut the DNA and the sizes (in kb) of the fragments expected to be seen in the blots. The solid line in panel A shows the ~300 bp long probe used in all blots, which was a PCR product that covered the *attL* site and 85 bp of the N-terminal region of the Hyg^R gene. Panels a–d in Fig. 3 depict, respectively, the results of the blots. Lanes 1 in panels (a)–(d) represent the blots of the cut plasmids that

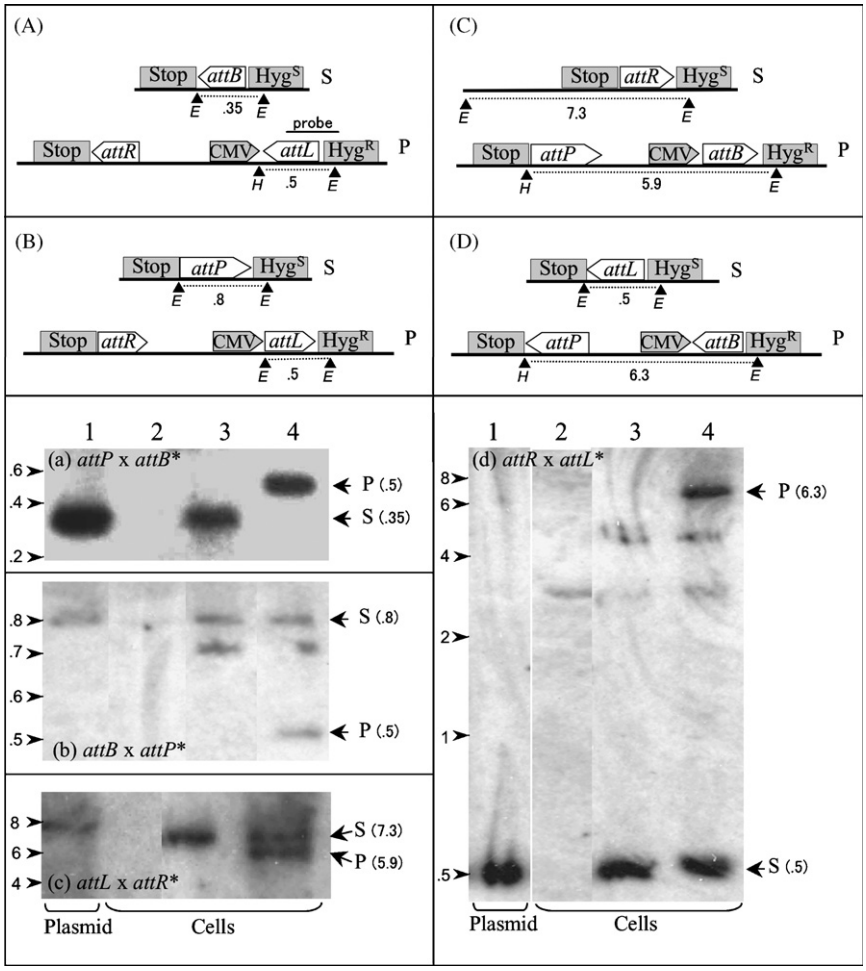

Fig. 3. Southern blots of the four *trans* reactions. (Panels A–D) Loci of the parental *att* sites before recombination (S) and of the Hyg^R recombinants (P). Triangles show the restriction sites (*E* = EcoRI, *H* = HindIII) used to cut the DNA for the blots. Solid thin line (in panel A) shows the probe used in all Southern blots, it was generated with primers oEY133 + oEY253. Dotted lines show the expected sizes (in kb) of the visible fragments in the blots. Panels (a)–(d) are the respective Southern blots. Lanes 1 are DNA of the cut *att* plasmids that were inserted into the genome. Lanes 2 are the non-transformed control cells. Lanes 3 show the parental *att* lines and lanes 4 show the Hyg^R recombinant lines. *Indicates chromosomal *att* site. Numbers with arrows show the fragment sizes in kb. Arrowheads indicate molecular weight markers in kb.

Please cite this article in press as: Harel-Levi, G., et al., Human genomic site-specific recombination catalyzed by coliphage HK022 integrase, J. Biotechnol. (2008), doi:10.1016/j.jbiotec.2008.01.002

were used to construct the relevant parental Hyg^S *att* lines as positive controls. Lanes 2 are the blots of the wild type negative controls. Lanes 3 show the cut DNA extracted from each relevant unrecombined parental Hyg^S *att* line, and lanes 4 show the blots of the Hyg^R recombinants. As expected, the chromosomal *att* lines (lanes 3) show the expected fragments of the unrecombined substrates (S) that are identical to the ones seen from their original plasmid DNA (lanes 1). The Hyg^R lines (lanes 4) show the expected recombination product (P) as well as the substrate [except for panel (a)]. The fact that in three blots of the Hyg^R lines the substrate (S) fragment was evident in addition to the product (P) indicates again that these lines carry more than one copy of the substrate. In panel (a), the Hyg^R line shows the product only and is likely to carry only one copy of the *attB* site. In the blot of the *attL* × *attR** cross [panels C and (c)] we used the single EcoRI site to cut the DNA such that the other site must have resided outside of the recombined cassette. However, since the plasmid (lane 1), the unrecombined *attR* parent (lane 3) and the Hyg^R recombinant (lane 4) all show the same fragment (S) it is likely that the unrecombined *attR* parent carries more than one substrate in tandem which is a known phenomenon in calcium-phosphate transformants (Wurtele et al., 2003). In any case, the appearance of the expected recombinant fragments in all four Hyg^R lines as well as the PCR results, confirms that all of them are the consequence of an Int-catalyzed site-specific recombination between the episomal *att* site and at least one of its chromosomal *att* sites.

3.3. Partial excision of an integrated episome by sequential recombinations

The difference between the Int system and the conventional ones (Cre, Flp and ϕC31) is that Int can utilize two different wild type pairs of compatible recombination sites (*attP* + *attB* and *attL* + *attR*). Assuming that other pairwise combinations are incompatible or much less efficient, this feature may afford the system an advantage in removing undesired sequences of the integrated episome using wild type target (*att*) sites. As a demonstration we have utilized a CMV-*attP* episomal plasmid, to which we added an *attR* site, the gene that encodes the thymidine kinase (TK) of the Herpes simplex virus and the GFP reporter gene with its own promoter [plasmid pCT909, Fig. 4A(a)]. The GFP was used as a reporter for stable transfection and the TK gene was used as a negative selection marker in the presence of Ganciclovir (Chauhan and Gottesman, 1992; Lupton et al., 1991). A sibling of the chromosomal Stop-*attB*-Hyg^S lines described above [Fig. 4A(b)] was co-transfected with plasmid pCT909 along with the Int-expressing plasmid followed by selection of Hyg^R recombinants as before. The four Hyg^R-GFP positive clones that were obtained are expected to have emerged from an *attP* × *attB* reaction that lead to the chromosomal structure shown in Fig. 4A(c). Besides the pre-existing *attR* site (upstream to TK), this reaction has produced the novel recombinant sites *attR* (downstream to Stop) and *attL* (downstream to the CMV promoter). PCR analyses using DNA of the original *attB* line (Fig. 4B, lanes 1–4) and of the recombinant line (lanes 5–8), each with different pair wise oligomers (numbered in circles in Fig. 4A), has confirmed the presence of the expected *attB* site (lane 4) in the original *attB* line and the presence of the expected *att* sites in recombinant left junction (LJ) and right junction (RJ_1) regions in the recombinant line [Fig. 4A(c) and B, lanes 5, 6]. The recombinant also shows a PCR product of *attB* (lane 8) indicating again that more than one copy of the *attB* fragment was present in the unrecombined parent. The sequences (not shown) of the PCR products in lanes 4, 5, 6 have confirmed the predicted *att* sites (*attB* in the unrecombined line, *attR* at the LJ region and *attL* at RJ_1). A Southern blot using the same probe as above (solid line in Fig. 4A) revealed the expected 1 kb fragment (Fig. 4C, lane 1). These results have confirmed the predicted structure of the Hyg^R recombinant depicted in Fig. 4A (c). In a second round of Int-promoted recombination events there are two options: either the reverse *attR* × *attL* reaction that leads back to the original *attB* line [Fig. 4A(b)], or the forward *attR* × *attL* reaction that deletes the fragment that includes the TK gene and the CMV promoter [Fig. 4A(d)]. In both reactions the products are expected to loose the TK gene and become $Ganciclovir^R$. In the reverse reaction the cells are also expected to loose the GFP activity and in the forward reaction they should retain it. The line with the first recombination product [Fig. 4A(c)] was transfected with the Int-expressing plasmid, followed by Ganciclovir treatment. It yielded 27 colonies, 4 of which retained the GFP activity. PCR analysis of a GFP positive line (Fig. 4B, lanes 10–13) has retained the previous LJ region (lane 10), has lost the RJ_1 region (lane 11) and has gained the new RJ_2 region (lane 12). The sequence of the two PCR products has confirmed both predicted *att* sites and the Southern blot has revealed the expected 6.4 kb fragment (Fig. 4C, lane 2). These results have verified the second sequential forward site-specific recombination reaction in which the GFP positive Ganciclovir survivor has deleted a fragment of the previously inserted plasmid DNA. The PCR analysis of a GFP-negative line among the $Ganciclovir^R$ colonies has suggested that it has indeed resulted from a reverse reaction leading to the original *attB* line (Fig. 4B, lanes 14–17). This experiment has demonstrated that it is possible to take advantage of the dual pair wise *att* sites of Int to insert circular DNA into the genome followed by the removal of a desired fragment of the inserted DNA. The implications of this advantage are discussed below.

4. Discussion

The wild type Int of phage HK022 can promote in the human cells integrative and excisive recombination in the *cis* configurations as well as all four possible combinations in the *trans* configuration. Hence, it can become a potential tool for site-specific gene manipulations in mammalian cells. The recombination reactions in the *cis* configurations are sufficiently frequent such that no selective force is required to detect them. However, the *trans* configurations are much less frequent, such that a selective force was needed to detect them. In the case of Cre and Flp the frequency of the site-specific *trans* reactions between a single pair of recombination sites is lower than the frequency of random insertions such that they also required

Please cite this article in press as: Harel-Levi, G., et al., Human genomic site-specific recombination catalyzed by coliphage HK022 integrase, J. Biotechnol. (2008), doi:10.1016/j.jbiotec.2008.01.002

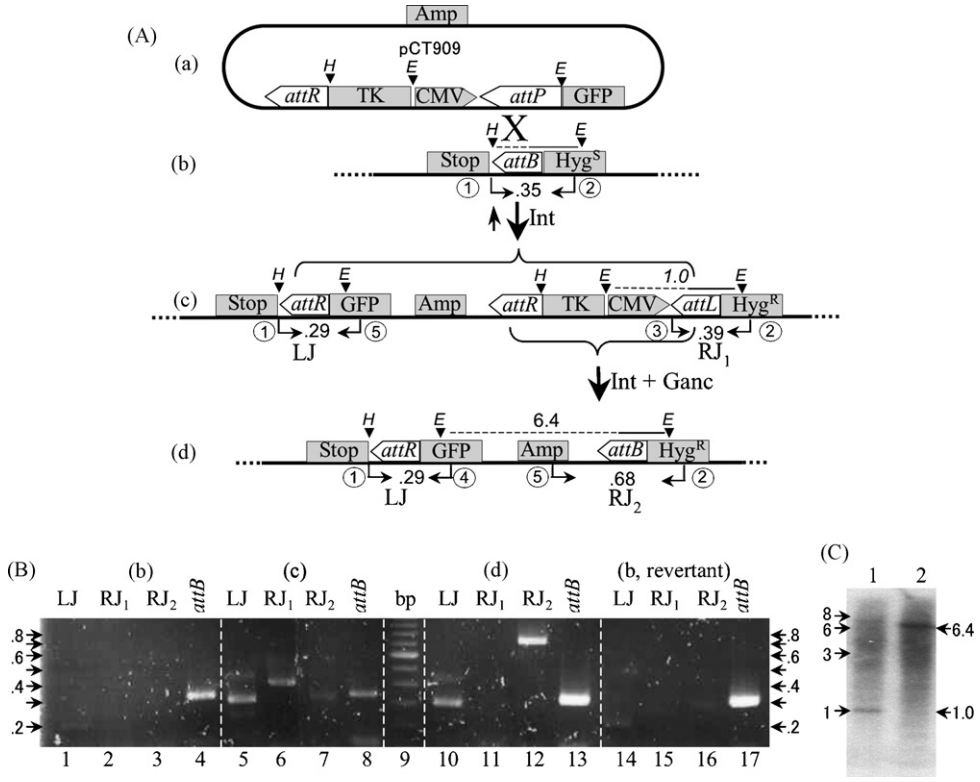

Fig. 4. Sequential site-specific recombinations. (A) Schematic presentation of the episomal plasmid designed for two subsequent recombination reactions (a), of the chromosomal loci in the *attB* line (b), of the expected structure of the first recombination product (c) and of the expected structure of the second forward recombination product (d). Bent arrows and circled numbers depict the primers used in the PCR reactions (1. oEY331, 2. oEY302, 3. oEY134, 4. oEY389, 5. oEY398). Triangles show the location of the restriction sites (*E* = EcoRI, *H* = HindIII) used in the Southern blots. Solid thin lines depict the labeled probes and their dotted extensions depict the expected sizes (in kb) of the visible fragments in the Southern blots. (B) Results of PCR reactions of respective recombinants using the pair wise oligomers depicted in panel A. Horizontal arrows indicate molecular weight markers in kb. Lanes 1–4 are the *attB* parental line (b). Lanes 5–8 are the first recombinant (c). Lane 9 represents a 100 bp ladder. Lanes 10–13 are the second round GFP positive recombinant (d). Lanes 14–17 are the second round GFP-negative revertant. (C) Southern blots. Lane 1: the first round recombinant (c). Lane 2: the second round GFP-positive recombinant (d). Arrows indicate molecular weight markers in kb.

a selection force to be detected (Baer and Bode, 2001; Sauer, 1994). In the mammalian cells as well as in the plants (Gottfried et al., 2005) the Int-catalyzed recombination reactions were evident without the need to supply the accessory proteins IHF and Xis that are required in the bacterial host and the reasons were previously discussed (Kolot et al., 2003). In phage λ only an IHF-dependent Int can promote site-specific recombination reactions in mammalian cells and the supply of IHF alleviates the reaction (Corona et al., 2003). In the experiments where the recombination reactions could be quantified (Fig. 1B, and others, not shown) the excisive *attR* × *attL* reaction were more efficient than integrative *attP* × *attB* reactions, confirming previous results on the episomal level (Kolot et al., 2003). The reason may be that, as in λ, excision is less dependent on IHF than is integration (Weisberg and Landy, 1983). Moreover, in the excisive reaction both *att* sites carry arms with high-affinity binding sites for Int that are capable to form presynapse complexes (Mumm et al., 2006), whereas in the integrative reaction only *attP* carries high-affinity binding sites for Int.

In both the *cis* reactions and in particular the *trans* reactions, only some of the transgenic cells that carried the *att* sites allowed recombination in the presence of Int [Table 2, line (d)]. The reasons might be due to the random integration of the substrate that has either resulted in a position effect that prevented the access of Int, or that the integrity of sequences that are essential to monitor the recombination reactions was distorted owing to the random integration of the substrate. Only about 25% of the DNA in the *att* plasmids were available for random chromosomal insertion without interrupting sequences that are essential to repost the recombination.

In the last experiment reported above we have demonstrated how, with the use of Int, one can insert a plasmid into a single site, followed by the removal of a part of the inserted DNA. The important feature in this approach is our assumption that the *attR* and *attP* sites on the episomal plasmid pCT909 are incompatible thereby avoiding an undesired *cis* recombination between them. This approach may prove useful in designing future insertion of desired genes, followed by the removal of markers that carry antibiotic resistance genes that are indispensable for the selective insertion reaction but whose presence in the final transgene is no more desired. However, compatibility tests between "illegal" pairs of recombination (e.g. *attP* × *attR*, *attP* × *attL*, etc.) must still be determined.

As with Cre and Flp, the reaction of Int in *cis* was more favourable. To prevent the reversibility of the *trans* reaction and to improve gene integration technology, the recombination-mediated cassette exchange (RMCE) system was developed. In this system a DNA segment flanked by two different and incom-

Please cite this article in press as: Harel-Levi, G., et al., Human genomic site-specific recombination catalyzed by coliphage HK022 integrase, J. Biotechnol. (2008), doi:10.1016/j.jbiotec.2008.01.002

patible pairs of recombination sites is exchanged by site-specific recombination with a chromosomal fragment that is flanked by the same heterologous sites (Baer and Bode, 2001). This approach has significantly improved the frequency of the desired gene insertion (Bethke and Sauer, 1997). However, in the case of the Cre-*lox* and Flp-*FRT* systems one of the recombining pair of sites must be mutated and be incompatible with the wild type site (Feng et al., 1999; Langer et al., 2002). If recombination between several "illegal" *att* sites will prove incompetent, it may render the Int-HK022 and the λ systems advantageous in an RMCE reaction avoiding the need to use mutated *att* sites.

Acknowledgments

We thank Ms. Nava Silberstein for devoted technical help. This work was supported by grants from the US-Israel Binational Science Foundation (grant 2003304), from the Israel Cancer Research Fund and from the Israel Science Foundation (grant no. 637/02).

References

Azaro, M.A., Landy, A., 2001. Integrase and the λ int family. In: Craig, N.L., Craigie, R., Gellert, M., Lambowitz, A. (Eds.), Mobile DNAII. ASM Press, Washington, DC, pp. 118–148.

Baer, A., Bode, J., 2001. Coping with kinetic and thermodynamic barriers: RMCE, an efficient strategy for the targeted integration of transgenes. Curr. Opin. Biotechnol. 12, 473–480.

Bethke, B., Sauer, B., 1997. Segmental genomic replacement by Cre-mediated recombination: genotoxic stress activation of the p53 promoter in single-copy transformants. Nucleic Acids Res. 25, 2828–2834.

Calos, M.P., 2006. The phiC31 integrase system for gene therapy. Curr. Gene Ther. 6, 633–645.

Chauhan, S.S., Gottesman, M.M., 1992. Construction of a new universal vector for insertional mutagenesis by homologous recombination. Gene 120, 281–285.

Christ, N., Corona, T., Dröge, P., 2002. Site-specific recombination in eukaryotic cells mediated by mutant lambda integrases: implications for synaptic complex formation and the reactivity of episomal DNA segments. J. Mol. Biol. 319, 305–314.

Corona, T., Bao, Q.Y., Christ, N., Schwartz, T., Li, J.M., Dröge, P., 2003. Activation of site-specific DNA integration in human cells by a single chain integration host factor. Nucleic Acids Res. 31, 5140–5148.

Feng, Y.Q., et al., 1999. Site-specific chromosomal integration in mammalian cells: highly efficient CRE recombinase-mediated cassette exchange. J. Mol. Biol. 292, 779–785.

Glaser, S., Anastassiadis, K., Stewart, A.F., 2005. Current issues in mouse genome engineering. Nat. Genet. 37, 1187–1193.

Gottfried, P., Yagil, E., Kolot, M., 2000. Core-binding specificity of bacteriophage integrase. Mol. Gen. Genet. 263, 619–624.

Gottfried, P., et al., 2005. Site-specific recombination in *Arabidopsis* plants promoted by Integrase protein of coliphage HK022. Plant Mol. Biol. 57, 435–444.

Groth, A.C., Calos, M.P., 2004. Phage integrases: biology and applications. J. Mol. Biol. 335, 667–678.

Hirt, B., 1967. Selective extraction of polyoma DNA from infected mouse cell cultures. J. Mol. Biol. 26, 365–369.

Kolot, M., Silberstein, N., Yagil, E., 1999. Site-specific recombination in mammalian cells expressing the Int recombinase of bacteriophage HK022. Mol. Biol. Rep. 26, 207–213.

Kolot, M., Meroz, A., Yagil, E., 2003. Site-specific recombination in human cells catalyzed by the wild-type integrase protein of coliphage HK022. Biotechnol. Bioeng. 84, 56–60.

Langer, S.J., Ghafoori, A.P., Byrd, M., Leinwand, L., 2002. A genetic screen identifies novel non-compatible IoxP sites. Nucleic Acids Res. 30, 3067–3077.

Lupton, S.D., Brunton, L.L., Kalberg, V.A., Overell, R.W., 1991. Dominant positive and negative selection using a hygromycin phosphotransferase-thymidine kinase fusion gene. Mol. Cell Biol. 11, 3374–3378.

Mumm, J.P., Landy, A., Gelles, J., 2006. Viewing single lambda site-specific recombination events from start to finish. EMBO J. 25, 4586–4595.

Nagy, A., 2000. Cre recombinase: the universal reagent for genome tailoring. Genesis 26, 99–109.

Sauer, B., 1993. Manipulation of transgenes by site-specific recombination: use of Cre recombinase. Methods Enzymol. 225, 890–900.

Sauer, B., 1994. Site-specific recombination: developments and applications. Curr. Opin. Biotechnol. 5, 521–527.

Sorrell, D.A., Kolb, A.F., 2005. Targeted modification of mammalian genomes. Biotechnol. Adv. 23, 431–469.

Tybulewicz, V.L.J., Crawford, C.E., Jackson, P.K., Bronson, R.T., Mulligan, R.C., 1991. Neonatal lethality and lymphopenia in mice with a homozygous disruption of the C-Abl protooncogene. Cell 65, 1153–1163.

Weisberg, R.A., Landy, A., 1983. Site-specific recombination in phage lambda. In: Hendrix, R.W., Roberts, J.W., Stahl, F.W., Weisberg, R.A. (Eds.), Lambda II. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, pp. 211–250.

Weisberg, R.A., Gottesmann, M.E., Hendrix, R.W., Little, J.W., 1999. Family values in the age of genomics: comparative analyses of temperate bacteriophage HK022. Annu. Rev. Genet. 33, 565–602.

Wurtele, H., Little, K.C.E., Chartrand, P., 2003. Illegitimate DNA integration in mammalian cells. Gene Ther. 10, 1791–1799.

Please cite this article in press as: Harel-Levi, G., et al., Human genomic site-specific recombination catalyzed by coliphage HK022 integrase, J. Biotechnol. (2008), doi:10.1016/j.jbiotec.2008.01.002