Published online 3 October 2008 *Nucleic Acids Research, 2008, Vol. 36, No. 19* **6333–6342**
doi:10.1093/nar/gkn649

Impact of non-homologous end-joining deficiency on random and targeted DNA integration: implications for gene targeting

Susumu Iiizumi[1,2], Aya Kurosawa[1], Sairei So[2], Yasuyuki Ishii[2], Yuichi Chikaraishi[2], Ayako Ishii[1], Hideki Koyama[2] and Noritaka Adachi[1,2,*]

[1]International Graduate School of Arts and Sciences and [2]Kihara Institute for Biological Research, Yokohama City University, Yokohama, Japan

Received August 12, 2008; Revised and Accepted September 18, 2008

ABSTRACT

In higher animal cells, the principal limitation of gene-targeting technology is the extremely low efficiency of targeted integration, which occurs three to four orders of magnitude less frequently than random integration. Assuming that random integration mechanistically involves non-homologous end-joining (NHEJ), inactivation of this pathway should reduce random integration and may enhance gene targeting. To test this possibility, we examined the frequencies of random and targeted integration in NHEJ-deficient chicken DT40 and human Nalm-6 cell lines. As expected, loss of NHEJ resulted in drastically reduced random integration in DT40 cells. Unexpectedly, however, this was not the case for Nalm-6 cells, indicating that NHEJ is not the sole mechanism of random integration. Nevertheless, we present evidence that NHEJ inactivation can lead to enhanced gene targeting through a reduction of random integration and/or an increase in targeted integration by homologous recombination. Most intriguingly, our results show that, in the absence of functional NHEJ, random integration of targeting vectors occurs more frequently than non-targeting vectors (harboring no or little homology to the host genome), implying that suppression of NHEJ-independent random integration events is needed to greatly enhance gene targeting in animal cells.

INTRODUCTION

Gene targeting by homologous recombination (HR) provides a powerful means for studying gene function by a reverse genetic approach. This technology also offers a potential tool for gene therapeutic applications (1). In higher animal cells, however, it is still quite difficult to generate knockout clones by gene targeting, except for mouse embryonic stem cells (2) and chicken B-lymphocyte DT40 cells (3), even though recent reports using human cell lines such as Nalm-6 and HCT116 indicate the feasibility of human gene targeting (4–9). One reason for the difficulty of gene targeting is that the frequency of targeted integration of targeting vectors through HR is as low as 10^{-6}, and the other reason is that the vectors integrate into random sites of the host genome (random integration) through non-HR at three to four orders of magnitude higher rates (1). It is therefore reasonable to expect that suppressing random integration events may enhance gene targeting by increasing the ratio of targeted to random integration.

Although the precise mechanism of random integration is yet to be elucidated, it has long been postulated that most, if not all, random integration events result from non-homologous end-joining (NHEJ) (10). NHEJ and HR are two major pathways for repairing DNA double-strand breaks (DSBs), which can be caused by a variety of endogenous and exogenous agents (11). In higher animal cells, NHEJ is believed to play a predominant role in DSB repair (12), though how cells choose NHEJ or HR remains obscure. The NHEJ reaction is initiated by the binding of Ku protein (the heterodimer of Ku70 and Ku80) to the ends of a DSB. Subsequently, Ku recruits a complex of DNA-dependent protein kinase catalytic subunit (DNA-PKcs) and Artemis and undergoes end-processing to make ligatable ends, which are finally joined by a complex of XRCC4, XLF (also called Cernunnos) and DNA ligase IV (12–14). This ligase IV complex is absolutely required for the classical (Ku-initiated) NHEJ pathway (4,15,16). Earlier studies reported that Chinese hamster cell lines lacking *Xrcc4* (xrs-1) or *Ku80* (xrs-6) showed decreased random integration frequencies (17,18). Additionally, a mouse cell mutant lacking DNA ligase IV exhibited

*To whom correspondence should be addressed. Tel/Fax: +81 45 787 2228; Email: nadachi@yokohama-cu.ac.jp
Correspondence may also be addressed to Hideki Koyama. Email: koyama@yokohama-cu.ac.jp

significantly (~20-fold) lower random integration frequencies than did wild-type cells (WT) (19). These results support the idea of NHEJ involvement in random integration events.

Despite the expectation that NHEJ deficiency may lead to enhanced gene targeting, there has been no supporting evidence in higher animal cells. For instance, Chinese hamster cells defective in *Ku80* were shown to retain gene-targeting frequencies comparable to those in WT cells (20). In addition, mouse embryonic stem cells deficient in *Ku70*, *Xrcc4* or *Dna-pkcs*, did not exhibit increased gene-targeting efficiencies (21,22). In sharp contrast, however, recent work has demonstrated that NHEJ deficiency leads to highly efficient gene targeting in lower eukaryotes, such as *Saccharomyces cerevisiae* or *Neurospora crassa* (23–30). These observations prompted us to re-examine the impact of NHEJ deficiency on gene targeting using animal cells. In this study, we examine the frequencies of random and targeted integration in NHEJ-deficient chicken DT40 and human Nalm-6 cell lines. We show that most random integration events occur via NHEJ, which, however, is not the sole mechanism of random integration. We also present evidence that NHEJ inactivation does lead to enhanced gene targeting through a reduction of random integration and/or a direct stimulation of targeted integration. Most importantly, our data reveal that random integration events of targeting vectors occur more frequently through a mechanism distinct from that of non-targeting vectors.

MATERIALS AND METHODS

Cells and culture conditions

The chicken B-cell lymphoblastoma DT40 cell line and its *Lig4*$^{-/-}$ derivative were maintained in a 5% CO_2 incubator at 39°C in ES medium (Nissui Seiyaku, Tokyo, Japan) supplemented with 10% FBS (Hyclone, Logan, UT, USA), 1% chicken serum and 10 μM 2-mercaptoethanol, as described (16). For colony formation, cells were grown for 7–11 days in ES medium containing 0.15% agarose (Seakem® LE; Cambrex Bio Science, Rockland, ME, USA), 20% FBS and 2% chicken serum. Note that hygromycin- and puromycin-resistant genes (*Hyg*^{r} and *Puro*^{r}, respectively) were removed from the original *Lig4*$^{-/-}$ cell line (16) by transiently expressed Cre, and one of the resulting marker-free clones was used in this study.

The human pre-B leukemia cell line Nalm-6 and its *LIG4*$^{-/-}$ derivative (6) were maintained in a 5% CO_2 incubator at 37°C in ES medium supplemented with 10% calf serum (Hyclone) and 50 μM 2-mercaptoethanol, as described (31,32). For colony formation, cells were grown for 2–3 weeks in growth medium containing 0.15% agarose (32).

Vectors

Plasmid vectors used in this study are listed in Table 1. pβActinHis was constructed by subcloning a 3.5-kb BamHI fragment containing the chicken β-actin promoter, hisD coding gene and polyA sequences (11) into BamHI-digested pBluescript SKII(–) (Stratagene, La Jolla, CA, USA). pSV2Puro was constructed by ligating a 0.9-kb HindIII/BglII fragment containing the *Puro*^{r} gene from pPGKPuro and a 3.4-kb HindIII/BamHI fragment containing the SV40 promoter and polyA sequences from pSV2neo (33). pPGKPuro2 was constructed by subcloning a 1.6-kb SalI/ScaI fragment containing the PGK promoter, *Puro*^{r} and polyA sequences from pPGKPuro (34) into SalI/SmaI-digested pUC19 (Takara Bio, Otsu, Japan). Targeting vector pFen1-Puro4 was constructed by shortening the 5′ and 3′ arms of the original vector pFEN1Puro (35). To construct a targeting vector for DNA polymerase β (Polβ), a 6.3-kb genomic fragment containing exons 1–4 was obtained by screening a chicken genomic library (Stratagene). Targeting vector pPolβ-Puro was constructed by replacing a 2.3-kb region containing exon 4 with a floxed *Puro*^{r} gene. Targeting vectors for the human hypoxanthine phosphoribosyltransferase gene (*HPRT*), pHPRT8.9-Puro and pHPRT-Hyg, were constructed by inserting *Puro*^{r} or *Hyg*^{r}, respectively, into the XhoI site of the 8.9-kb *HPRT* fragment (31,36). Targeting vectors for the human *53BP1*, *RAD52*, *RAD54*, *POLβ RAG1* and *ARTEMIS* genes were constructed by a simplified vector construction system (6,37 and Kurosawa *et al.*, unpublished data). All the plasmid vectors were purified with Qiagen Plasmid Maxi Kits (Qiagen K.K., Tokyo, Japan) and linearized with an appropriate restriction enzyme prior to transfection (Table 1).

Table 1. Vectors

Vectors	Linearized with:	Used for:	References
Non-targeting vectors			
pSV2neo	EcoRI	DT40	Southern and Berg (33)
pPGKPuro2	XmnI	DT40	This study
pSV2Puro	EcoRI	DT40	This study
pPGKPuro	ScaI	Nalm-6	Tucker *et al.* (34)
pβActinHis	XmnI	Nalm-6	This study
Targeting vectors			
pTop2α-Puro	SphI	DT40	Adachi *et al.* (38)
pPolβ-Puro	ScaI	DT40	This study
pFen1-Puro4	EcoRI	DT40	This study
pHPRT8.9-Puro	NotI	Nalm-6	This study
p53BP1-Hyg	AhdI	Nalm-6	Iiizumi *et al.* (unpublished)
pRAD52-Puro	AhdI	Nalm-6	Adachi *et al.* (unpublished)
pRAD54-Puro	SwaI	Nalm-6	Kurosawa *et al.* (unpublished)
pPOLβ-Hyg	I-SceI	Nalm-6	Adachi *et al.* (unpublished)
pHPRT-Hyg	NotI	Nalm-6	So *et al.* (36)
pRAG1-Puro	I-SceI	Nalm-6	Kurosawa *et al.* (unpublished)
pRAG1-Hyg	I-SceI	Nalm-6	Kurosawa *et al.* (unpublished)
pARTEMIS-Puro	I-SceI	Nalm-6	Kurosawa *et al.* (37)
pARTEMIS-Hyg	I-SceI	Nalm-6	Kurosawa *et al.* (37)

Transfection and integration assays

DNA transfection was performed in DT40 cells as previously described (16). Briefly, 4×10^6 cells were electroporated with linearized plasmid (3 μg for random integration assays and 4 μg for gene-targeting experiments) per 40-μl chamber of Electro Gene Transfer Equipment (GTE-1; Shimadzu, Kyoto, Japan). After 15 min, cells were transferred into growth medium and cultured for 8 h. The cells

were then collected, counted and replated into agarose medium containing 0.5 μg/ml puromycin (Wako Pure Chemical, Osaka, Japan) or 1.6 mg/ml G418 (Gibco BRL, Gaithersburg, MD, USA). Meanwhile, small aliquots of the transfected cells were replated into drug-free agarose medium to determine the plating efficiency. The resulting colonies after cultivation for 7–11 days were counted, and the integration frequency was calculated by dividing the number of drug-resistant colonies with that of surviving cells. In Nalm-6 cells, random integration assays were carried out in essentially the same manner as in DT40 cells with slight modifications. Briefly, cells were transfected with 4 μg of linearized vector and cultured for 22-h in growth medium. The cells were then replated into agarose medium containing either 0.5 μg/ml puromycin, 0.4 mg/ml hygromycin B (Wako Pure Chemical), or 1.2 mg/ml L-histidinol (Sigma-Aldrich, St Louis, MO, USA), and cultured for 2–3 weeks. For gene-targeting experiments, each targeting vector was transfected into WT or mutant cells, which were subsequently selected for drug-resistant colonies, and correct gene-targeting events were confirmed by Southern blot analysis.

Gene targeting at the human *HPRT* locus

Gene targeting in human Nalm-6 cells was carried out using an *HPRT* targeting system, as described previously (31,36). Briefly, cells were transfected with 4 μg of linearized targeting vector pHPRT8.9-Puro, cultured for 22 h in growth medium and replated into agarose medium with 0.5 μg/ml puromycin. After a 2-week incubation at 37°C, puromycin-resistant colonies were counted to calculate the total integration frequency. Subsequently, single colonies were isolated, expanded and replated into growth medium containing 20 μM 6-thioguanine (6TG; Sigma-Aldrich), a hypoxanthine analog that kills *HPRT*-proficient cells. Genomic DNA was isolated from 6TG-resistant clones and subjected to PCR analysis using primers HPRT-F (5′-TGAGGGCAAAGGATGTGTTACGTG-3′) and HPRT-R (5′-TTGATGTAATCCAGCAGGTCAGCA-3′). The gene-targeting efficiency was calculated by dividing the number of targeted clones with that of drug-resistant clones analyzed. The targeted integration frequency was calculated by multiplying the total integration frequency by the targeting efficiency. The random integration frequency was calculated by subtracting the targeted integration frequency from the total integration frequency.

RESULTS

NHEJ-deficient cells have reduced random integration frequencies

To compare the ability of DT40 WT and NHEJ-deficient cells to integrate transfected DNA into the host genome, we performed quantitative transfection experiments using several plasmid vectors that have no or little homology to the genome (Table 1). (To distinguish from targeting vectors, these vectors are hereafter referred to as non-targeting vectors.) As shown in Figure 1A, the integration frequency of pSV2neo in *Lig4*$^{-/-}$ cells dropped to as low as 7.5% of that in WT cells. Very similar results were obtained using two other non-targeting vectors; the random integration frequencies of pSV2Puro and pPGKPuro2 in *Lig4*$^{-/-}$cells were reduced to 3 and 11%, respectively, relative to WT cells (Figure 1B and C). We note that very similar results were obtained using *Ku70*$^{-/-}$ cells (11) (data not shown). In order to ensure that chromosomal integration of transfected DNA is completely random, we performed Southern blot analysis using individual integrants (at least 20 clones for each cell line) derived from pSV2Puro-transfected WT and *Lig4*$^{-/-}$ cells. We found that in all cell lines, the size and pattern of hybridizing bands were obviously distinct from one another (data not shown). To eliminate the possibility that the difference in random integration frequency between WT and *Lig4*$^{-/-}$ cells was due to a reduced uptake of transfected vectors, we performed transient assays by transfecting a luciferase expression vector, pPGKluc, and confirmed that the *Lig4*$^{-/-}$ mutant exhibited luciferase activity to the level comparable to that of WT cells (data not shown). Collectively, these data indicate that in DT40 cells, NHEJ is responsible for nearly all of random integration events of non-targeting vectors. It is important to emphasize, however, that other mechanism(s) must exist that permit random integration, as the

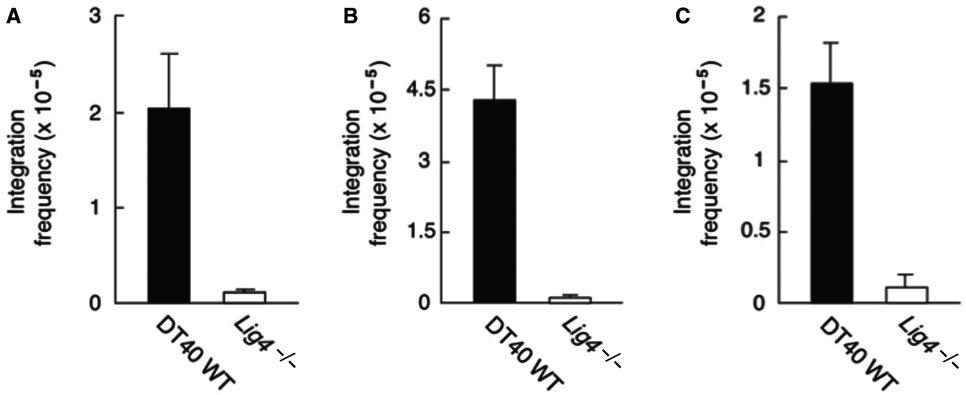

Figure 1. *Lig4*$^{-/-}$ DT40 cells have reduced random integration frequencies. Shown are the mean ± SD of three to five independent experiments with pSV2neo (**A**), pSV2Puro (**B**) and pPGKPuro2 (**C**).

frequency of random integration does not drop to zero in the absence of functional NHEJ.

We next employed the human pre-B cell line Nalm-6 to examine the impact of NHEJ deficiency on random integration. As shown in Figure 2, the random integration frequencies of pβActinHis and pPGKPuro in *LIG4*$^{-/-}$ cells were decreased to ~50% of that in WT cells, contrasting with the significant drop observed in DT40 cells. The vector uptake after transfection was unaffected in NHEJ-deficient human cells, as judged by transient assays with a luciferase expression vector, pCMVluc (data not shown). These data indicate that the NHEJ pathway is responsible for only approximately half of random integration events for non-targeting vectors in the human Nalm-6 cell line, confirming the notion that some random integration events occur through an NHEJ-independent mechanism(s), whose contribution to random integration may differ between cell lines or species.

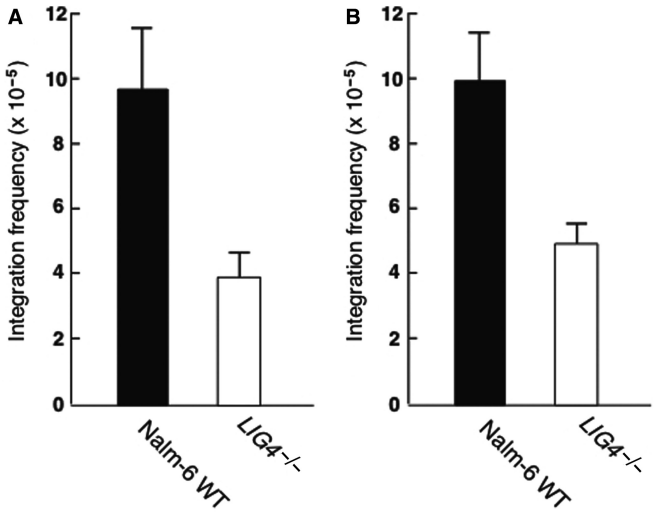

Figure 2. *LIG4*$^{-/-}$ Nalm-6 cells have reduced random integration frequencies. Shown are the mean ± SD of six independent experiments with pβActinHis (**A**) and pPGKpuro (**B**).

Enhanced gene targeting in NHEJ-deficient DT40 cells

Because random integration was substantially suppressed by *Lig4* deficiency in DT40 cells, we expected that the *Lig4* deficiency could enhance gene targeting. To test this, we first performed gene targeting at the *Fen1* locus (Figure 3A). Specifically, pFen1-Puro4 vector was transfected into WT and *Lig4*$^{-/-}$ cells and puromycin-resistant colonies were counted to calculate the total integration frequency. Additionally, these colonies were subjected to Southern blot analysis (Figure 3B) and classified into random and targeted integrants to calculate the frequencies of random and targeted integration. As shown in Figure 4A, the frequency of random integration in *Lig4*$^{-/-}$ cells was decreased to ~27% of that in WT cells, while the frequency of targeted integration was marginally affected by *Lig4* deficiency. As a consequence, the gene-targeting efficiency was elevated ~2.6-fold in the mutant (Figure 4B). We next performed gene targeting at the *Polβ* locus by using pPolβ-Puro vector (Figure 3C and D). The random integration frequency in *Lig4*$^{-/-}$ cells was decreased to ~25% (Figure 4C), resulting in an ~1.7-fold increased gene-targeting efficiency (Figure 4D). We further employed pTop2α-Puro vector (38) to target the *Top2α* locus, and found that the random integration frequency in *Lig4*$^{-/-}$ cells was decreased to ~36% of that in WT cells (Figure 4E), resulting in an ~1.8-fold increased gene-targeting efficiency (Figure 4F). Taken together, these results indicate that *Lig4* deficiency leads to

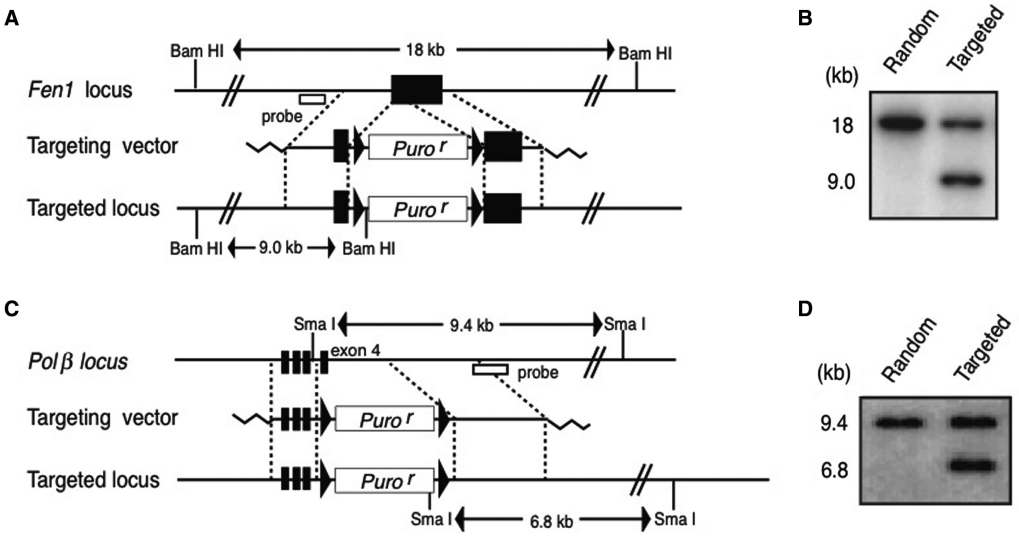

Figure 3. Schematic representation of gene targeting at the chicken *Fen1* and *Polβ* loci. (**A**) The *Fen1* locus, the targeting vector, and the targeted locus are shown. Closed boxes indicate exons, and triangles designate *lox*P sequences. (**B**) Southern blot analysis of the *Fen1* locus. BamHI-digested genomic DNA from representative random and targeted clones was hybridized with the probe shown in (A). (**C**) The *Polβ* locus, the targeting vector, and the disrupted loci are shown. (**D**) Southern blot analysis of the *Polβ* locus. SmaI-digested genomic DNA from representative random and targeted clones was hybridized with the probe shown in (C).

enhanced gene targeting in DT40 cells, by virtue of a reduction of random integration.

LIG4 deficiency in Nalm-6 cells leads to enhanced gene targeting without reduced random integration of -targeting vectors

To determine the frequency of random and targeted integration in Nalm-6 cells, we designed an assay system utilizing the X-chromosome-linked *HPRT* gene (illustrated in Figure 5A; Nalm-6 is of male origin and has a single *HPRT* locus), which enabled us to select for targeted integrants without the initial requirement of genomic analysis. Specifically, puromycin-resistant colonies were transferred into 6TG-containing medium, and those colonies that survived the 6TG selection were subjected to PCR analysis using a set of primers that flank exon 3 of the *HPRT* gene.

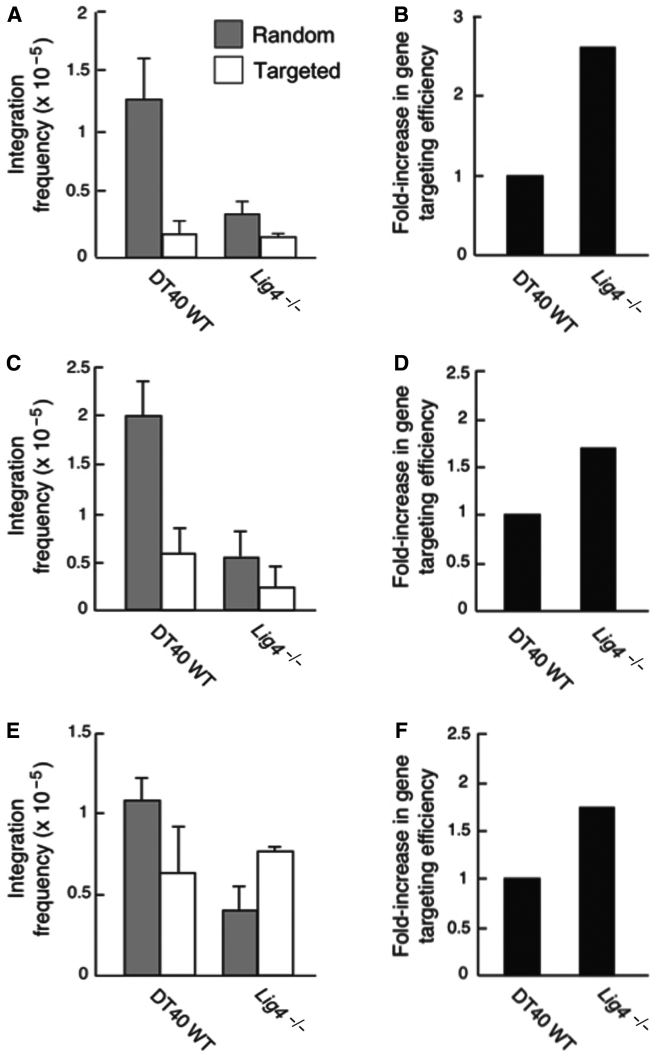

Figure 4. Reduced random integration of targeting vectors and enhanced gene-targeting efficiencies in $Lig4^{-/-}$ DT40 cells. (**A**, **C**, **E**) Random and targeted integration frequencies of targeting vectors for the chicken *Fen1* (A), *Polβ* (C) and *Top2α* (E) genes. Shown are the mean ± SD of three to four independent experiments. (**B**, **D** and **F**) Fold-increase in gene-targeting efficiency in $Lig4^{-/-}$ DT40 cells at the chicken *Fen1* (B), *Polβ* (D) and *Top2α* (F) loci.

In this analysis, random integrants, like untransfected cells, display a 0.5-kb band, whereas targeted integrants display a 2.2-kb band due to $Puro^r$ insertion (Figure 5B). Thus, 6TG-resistant colonies that gave no 0.5-kb band in the PCR analysis were counted as targeted integrants (correct gene targeting was finally verified by Southern blot), and the rest were counted as random integrants.

Surprisingly, the random integration frequency of the targeting vector pHPRT8.9-Puro was not reduced but rather slightly increased in human $LIG4^{-/-}$ cells (Figure 5C). Intriguingly, however, the targeted integration frequency was elevated >2-fold in the mutant, thus resulting in an increase in gene-targeting efficiency (Figure 5D). To further examine the impact of *LIG4* deficiency in human Nalm-6 cells, we employed *53BP1* targeting vector to perform gene targeting. As shown in Table 3, WT cells gave rise to no targeted clones, whereas two targeted clones were obtained from $LIG4^{-/-}$ cells with the efficiency of 3.3%. Further, we performed gene-targeting experiments at four additional autosomal loci, *RAD52*, *RAD54*, *POLβ* and *ARTEMIS*. As summarized in Table 3, relative to WT cells, $LIG4^{-/-}$cells showed increased gene-targeting efficiencies in all cases examined, which was statistically significant ($P < 0.05$; data from the four loci). From these results, we conclude that *LIG4* deficiency leads to enhanced gene targeting in Nalm-6 cells. It should be emphasized, however, that this enhancement is likely caused by direct stimulation of targeted integration, as our data reveal that loss of DNA ligase IV does not reduce random integration of targeting vectors.

Table 2. Summary of gene targeting at the human *HPRT* locus

Experiment	Cell line	No. of PCR-positive clones (a)	No. of 6TG-resistant clones (b)	No. of puromycin-resistant clones analyzed (c)	Gene-targeting efficiency (%) (a/c × 100)
Exp. 1	WT	9	9	144	6.3
	$LIG4^{-/-}$	23	25	144	16
Exp. 2	WT	18	18	176	10
	$LIG4^{-/-}$	27	30	192	14
Exp. 3	WT	21	21	192	11
	$LIG4^{-/-}$	33	40	192	17
Exp. 4	WT	10	10	192	5.2
	$LIG4^{-/-}$	33	33	192	17

Table 3. Summary of gene targeting experiments in Nalm-6 cells at autosomal loci

Locus	Selection marker	Gene-targeting efficiency	
		WT (%)	$LIG4^{-/-}$ (%)
53BP1	Hyg	0/120 (<0.8)	2/60 (3.3)
RAD52	Puro	1/185 (0.5)	3/193 (1.6)
RAD54	Puro	4/103 (3.9)	4/87 (4.6)
POLB	Hyg	3/72 (4.2)	5/108 (4.6)
ARTEMIS	Puro	1/142 (0.7)	2/140 (1.4)
ARTEMIS	Hyg	1/240 (0.4)	2/238 (0.8)

Table 4. Summary of total integration frequency of various vectors in Nalm-6 cells

Vectors	Fold-increase in total integration frequency (*LIG4*$^{-/-}$ versus WT)
Non-targeting vectors	
pβActinHis	0.43
pPGKPuro	0.50
Targeting vectors	
pHPRT8.9-Puro	1.27
pHPRT-Hyg	1.11
pRAG1-Puro	1.10
pRAG1-Hyg	1.13

Indeed, slight increases in random integration frequency in human *LIG4*$^{-/-}$ cells were similarly observed with other targeting vectors (Table 4).

DISCUSSION

Despite little supporting evidence, NHEJ has long been postulated to be responsible for random integration of transfected DNA (10), leading to the idea that NHEJ deficiency should reduce random integration and may enhance gene targeting. Our present study has revealed that this is indeed the case in chicken DT40 cells. Intriguingly, however, our data show that in human Nalm-6 cells, NHEJ deficiency can enhance targeted integration but does not reduce random integration of targeting vectors. Recent evidence suggests the existence of a competition, upon DSBs, between the NHEJ and HR pathways (39,40). Given this, eliminating the NHEJ pathway would shunt DSBs toward HR. Indeed, in Chinese hamster cells *Dna-pkcs* or *Xrcc4* deficiency enhanced spontaneous as well as DSB-induced HR (41,42). In the present study, a direct stimulation of targeted integration by NHEJ deficiency was observed in Nalm-6 cells (Figure 5) but was not evident in DT40 cells (Figure 4). Possibly, this discrepancy may reflect an intrinsic difference in DSB repair mechanism between human and other animal cells. Further work is required to clarify the competitive relationship between NHEJ and HR in gene-targeting events.

In *Lig4*$^{-/-}$ DT40 cells, random integration frequencies of non-targeting vectors were reduced to 3–11% of WT levels (Figure 1). In addition, *LIG4*$^{-/-}$ Nalm-6 cells showed an ~50% reduction in random integration frequency (Figure 2). We previously reported that a *Lig4*-deficient mouse cell mutant showed a markedly reduced level of random integration (19). Others reported earlier that *Xrcc4* or *Ku80* deficient Chinese hamster cell lines showed reduced random integration frequencies (17,18). Taken together, these observations strongly support the view that random integration events occur through the NHEJ pathway (10). We would like to emphasize, however, that a small number of random integrants always arise in *Lig4*$^{-/-}$ DT40 cells (Figure 1), and a considerably large fraction of random integrants do appear in *LIG4*$^{-/-}$ Nalm-6 cells (Figure 2). Thus, random integration events cannot be completely suppressed by ablating the classical NHEJ pathway, pointing to a contribution of an NHEJ-independent pathway(s) to the residual random integration events. Indeed, recent studies by several groups have revealed that a backup pathway for NHEJ (termed 'alternative end-joining') functions in V(D)J and class switch recombination at unexpectedly robust levels (43–45). This alternative pathway appears to be operating in Nalm-6 cells, as, for example, I-SceI-induced chromosomal DSBs are efficiently rejoined even when DNA ligase IV is absent (31). It is therefore highly likely that this alternative end-joining pathway is responsible for random integration events in the human *LIG4*$^{-/-}$ mutant, and possibly contributes to the residual random integrants observed in the NHEJ-deficient chicken cells.

Given that animal cells have three genetically distinct genes for DNA ligase (*LIG1*, *LIG3* and *LIG4*), it is conceivable that a *LIG1* and/or *LIG3* product(s) are responsible for NHEJ-independent random integration events. Recent evidence indicates that DNA ligase IIIα, which is one of the *LIG3* products, is involved in the alternative end-joining pathway (46–48), suggesting that this ligase has a role in random integration. Yet, since our preliminary experiments showed that *LIG1* and *LIG3* knockdown both reduced random integration frequencies (Iiizumi *et al.*, unpublished observations), we speculate that multiple DNA ligases are responsible for the residual random integrants observed with our mutant cell lines described herein. Apparently, however, a more comprehensive analysis is required to elucidate the nature of, and the relative contribution of multiple DNA ligases to, NHEJ-independent random integration.

Intriguingly, when targeting vectors were transfected into NHEJ-deficient DT40 cells, the frequency of random integration was only decreased to 25–36% of that in WT cells (Figure 4), contrasting with the case of non-targeting vectors (Figure 1). More surprisingly, despite an ~50% decrease in random integration with non-targeting vectors (Figure 2), *LIG4*$^{-/-}$ Nalm-6 cells showed no decrease rather a slight increase in random integration with targeting vectors (Figure 5 and Table 4). These results clearly show that in addition to the mechanism for random integration of non-targeting vectors, integration of targeting vectors would involve some other mechanism(s) that do not rely on DNA ligase IV. Why do targeting vectors integrate more efficiently than do non-targeting vectors in the absence of NHEJ? One possibility is that targeting vectors may be able to integrate into the genome in a microhomology-dependent manner. In this regard, Merrihew *et al.* (49) proposed a model for microhomology-dependent random integration, where a short homologous sequence of transfected vector primes DNA synthesis within a single-stranded chromosomal region, and resolution by endonuclease cleavage leaves the vector linked to the chromosome. This idea may be supported by the fact that targeting vectors, in general, contain repetitive DNA sequences, involving short interspersed nuclear elements (SINEs) such as Alu sequences. It was reported earlier that such repetitive sequences could participate in genetic recombination (50–52). A more recent study using murine HC11 cells showed that the integration frequency of a SINE-flanked vector was 3.5-fold higher than that of

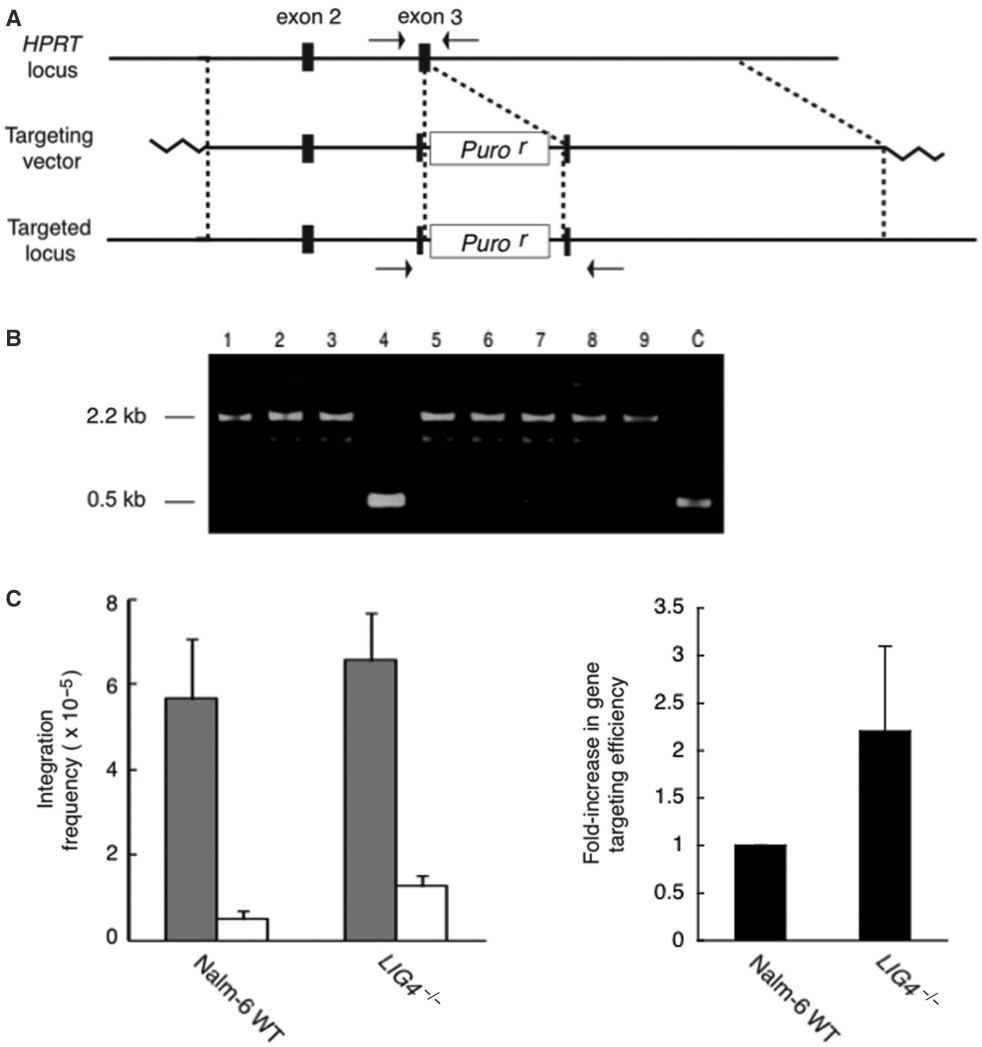

Figure 5. Enhanced gene targeting at the HPRT locus in *LIG4*$^{-/-}$ Nalm-6 cells. (**A**) Scheme of gene targeting at the human *HPRT* locus. The targeting vector was designed to insert into exon 3. Black boxes represent exons. Arrows indicate PCR primers. (**B**) PCR analysis to confirm gene targeting. Shown are the results with nine 6TG-resistant clones from *LIG4*$^{-/-}$ cells (1–9) and untransfected cells (marked by C). (**C**) Random and targeted integration frequencies of the *HPRT* targeting vector in WT and *LIG4*$^{-/-}$ Nalm-6 cells. Symbols are as in Figure 4. Shown are the mean ± SD of four independent experiments (Table 2). (**D**) Fold-increase in gene-targeting efficiency in *LIG4*$^{-/-}$ cells at the *HPRT* locus. Shown are the mean ± SD of four independent experiments.

SINE-less controls (53), and this increase was attributed to HR between the exogenous and endogenous SINEs. Considering that Alu sequence-mediated HR has been suggested in several studies (54–56), it is possible that NHEJ-independent random integration of targeting vectors may be related to HR between SINE sequences present in the arms of targeting vector and the host genome. The existence of such repetitive sequence-dependent random integration may well explain why *LIG4*$^{-/-}$ Nalm-6 cells, unlike NHEJ-deficient DT40 cells, failed to show reduced random integration frequencies with targeting vectors (Figures 4 and 5), as the human genome has a >10-fold greater number of interspersed repeats than the chicken genome (3×10^6 versus 0.25×10^6 copies) (57,58). Do, then, targeting vectors devoid of any repetitive sequences allow for more efficient gene targeting? To gain insights into this issue, we compared the structures of the targeting vectors used for human cells, as summarized in Supplementary Table 1. Thus far, however, the relationship between the gene-targeting efficiency and the amount of repetitive sequences appears unlikely. For instance, the *LIG4* and *RAD54* targeting vectors gave similarly high gene-targeting efficiencies, despite the apparent difference in the amount of SINE/LINE sequences between these vectors. This may imply, however, that the frequency of gene targeting may be more strongly governed by other possible factors, such as chromosomal localization or transcriptional status of the target site (59,60). Thus, a more detailed analysis will clarify those issues, for example, by creating and comparing multiple targeting vectors for the same gene with similar sizes but with different lengths of repetitive sequences. It will also be interesting to make a series of 'pseudo-targeting vectors', in which either arm of the targeting vector is missing or both arms are reversed; these vectors may help clarify the contribution of repetitive

sequences to the DNA ligase IV-independent integration events.

In species with a far smaller genome size, NHEJ deficiency successfully leads to efficient gene targeting with significantly reduced random integration (23–30). For instance, in *N. crassa*, whose genome size is ~100-times smaller than that of humans (61), NHEJ deficiency does completely abolish random integration (25,26). This may support the aforementioned idea that DNA ligase IV-independent random integration of targeting vector depends upon the presence of repetitive sequences. Unfortunately, our results presented here indicate that such a great increase in gene targeting cannot be achieved by loss of DNA ligase IV in higher animal cells, even though it remains possible that a more dramatic effect on gene targeting may be observed in cells with higher rates of random integration (and thus we will continue to monitor this issue as we knock out the *LIG4* gene in an increasing number of other human cell lines). It should be mentioned, however, that Hendrickson and colleagues (62) very recently showed using the HCT116 cell line that, unlike in rodent or chicken cells, even a heterozygous disruption of the *KU80* gene resulted in an increased gene-targeting frequency. Although only AAV-based targeting vectors were employed in that study, this finding may suggest that suppressing expression of upstream NHEJ factors such as Ku or DNA-PKcs, rather than DNA ligase IV (the most downstream factor), is a plausible way of enhancing gene targeting, at least in human cells. It is possible that absence or decrease of Ku or DNA-PKcs (or even Artemis) can shunt DSBs toward other repair pathways more efficiently than the absence of DNA ligase IV, consistent with previous work (15,16).

In summary, we have shown that most random integration events occur via NHEJ, which, however, is clearly not the sole mechanism of random integration in animal cells, unlike in certain lower eukaryotes. Thus, in animal cells, NHEJ inactivation alone does not necessarily lead to a significant reduction of random integration. Nevertheless, NHEJ inactivation does enhance gene targeting in both DT40 and Nalm-6 cells, though this enhancement is far less prominent than that observed in lower eukaryotes. Finally, our data presented here have implications for gene-targeting experiments. We have demonstrated for the first time that random integration of targeting vectors occurs more efficiently, in the absence of DNA ligase IV, than that of non-targeting vectors. We therefore suggest that, in addition to ablating the classical NHEJ pathway, suppressing the additional mechanism for random integration may greatly enhance the efficiency of gene targeting in animal cells.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Chie Nishigaki, Kaoru Okamura, Kyoko Murai and Sachiko Yamamoto for excellent technical assistance.

FUNDING

Yokohama City University (Strategic Research Project, No. W18006, K19009, and K20005); Grant-in-Aids from the Ministry of Education, Culture, Sports, Science and Technology (MEXT). Funding for open access charge: Yokohama City University.

Conflict of interest statement. None declared.

REFERENCES

1. Vasquez,K.M., Marburger,K., Intody,Z. and Wilson,J.H. (2001) Manipulating the mammalian genome by homologous recombination. *Proc. Natl Acad. Sci. USA*, **98**, 8403–8410.
2. Capecchi,M.R. (1989) Altering the genome by homologous recombination. *Science*, **244**, 1288–1292.
3. Yamazoe,M., Sonoda,E., Hochegger,H. and Takeda,S. (2004) Reverse genetic studies of the DNA damage response in the chicken B lymphocyte line DT40. *DNA Repair*, **3**, 1175–1185.
4. Grawunder,U., Zimmer,D., Fugmann,S., Schwarz,K. and Lieber,M.R. (1998) DNA ligase IV is essential for V(D)J recombination and DNA double-strand break repair in human precursor lymphocytes. *Mol. Cell*, **2**, 477–484.
5. Adachi,N., So,S., Iiizumi,S., Nomura,Y., Murai,K., Yamakawa,C., Miyagawa,K. and Koyama,H. (2006) The human pre-B cell line Nalm-6 is highly proficient in gene targeting by homologous recombination. *DNA Cell Biol.*, **25**, 19–24.
6. Iiizumi,S., Nomura,Y., So,S., Uegaki,K., Aoki,K., Shibahara,K., Adachi,N. and Koyama,H. (2006) Simple one-week method to construct gene-targeting vectors: application to production of human knockout cell lines. *Biotechniques*, **41**, 311–316.
7. Li,G., Nelsen,C. and Hendrickson,E.A. (2002) Ku86 is essential in human somatic cells. *Proc. Natl Acad. Sci. USA*, **99**, 832–837.
8. Hiyama,T., Katsura,M., Yoshihara,T., Ishida,M., Kinomura,A., Tonda,T., Asahara,T. and Miyagawa,K. (2006) Haploinsufficiency of the Mus81-Eme1 endonuclease activates the intra-S-phase and G2/M checkpoints and promotes rereplication in human cells. *Nucleic Acids Res.*, **34**, 880–892.
9. Shiomi,N., Mori,M., Tsuji,H., Imai,T., Inoue,H., Tateishi,S., Yamaizumi,M. and Shiomi,T. (2007) Human RAD18 is involved in S phase-specific single-strand break repair without PCNA monoubiquitination. *Nucleic Acids Res.*, **35**, e9.
10. Roth,D. and Wilson,J.H. (1988) Illegitimate recombination in mammalian cells. In Kucherlapati,R. and Smith,G. (eds), *Genetic Recombination*. American Society for Microbiology, Washington, pp. 621–653.
11. Takata,M., Sasaki,M.S., Sonoda,E., Morrison,C., Hashimoto,M., Utsumi,H., Yamaguchi-Iwai,Y., Shinohara,A. and Takeda,S. (1998) Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells. *EMBO J.*, **17**, 5497–5508.
12. Lieber,M.R., Yu,K. and Raghavan,S.C. (2006) Roles of nonhomologous DNA end joining, V(D)J recombination, and class switch recombination in chromosomal translocations. *DNA Repair*, **5**, 1234–1245.
13. Ahnesorg,P., Smith,P. and Jackson,S.P. (2006) XLF interacts with the XRCC4-DNA ligase IV complex to promote DNA nonhomologous end-joining. *Cell*, **124**, 301–313.
14. Buck,D., Malivert,L., de Chasseval,R., Barraud,A., Fondaneche,M.C., Sanal,O., Plebani,A., Stephan,J.L., Hufnagel,M., le Deist,F. *et al.* (2006) Cernunnos, a novel nonhomologous end-joining factor, is mutated in human immunodeficiency with microcephaly. *Cell*, **124**, 287–299.
15. Karanjawala,Z.E., Adachi,N., Irvine,R.A., Oh,E.K., Shibata,D., Schwarz,K., Hsieh,C.L. and Lieber,M.R. (2002) The embryonic lethality in DNA ligase IV-deficient mice is rescued by deletion of Ku: implications for unifying the heterogeneous phenotypes of NHEJ mutants. *DNA Repair*, **1**, 1017–1026.
16. Adachi,N., Ishino,T., Ishii,Y., Takeda,S. and Koyama,H. (2001) DNA ligase IV-deficient cells are more resistant to ionizing

radiation in the absence of Ku70: implications for DNA double-strand break repair. *Proc. Natl Acad. Sci. USA*, **98**, 12109–12113.
17. Hamilton,A.A. and Thacker,J. (1987) Gene recombination in X-ray-sensitive hamster cells. *Mol. Cell. Biol.*, **7**, 1409–1414.
18. Jeggo,P.A. and Smith-Ravin,J. (1989) Decreased stable transfection frequencies of six X-ray-sensitive CHO strains, all members of the xrs complementation group. *Mutat. Res.*, **218**, 75–86.
19. Sado,K., Ayusawa,D., Enomoto,A., Suganuma,T., Oshimura,M., Sato,K. and Koyama,H. (2001) Identification of a mutated DNA ligase IV gene in the X-ray-hypersensitive mutant SX10 of mouse FM3A cells. *J. Biol. Chem.*, **276**, 9742–9748.
20. Liang,F., Romanienko,P.J., Weaver,D.T., Jeggo,P.A. and Jasin,M. (1996) Chromosomal double-strand break repair in Ku80-deficient cells. *Proc. Natl Acad. Sci. USA*, **93**, 8929–8933.
21. Pierce,A.J., Hu,P., Han,M., Ellis,N. and Jasin,M. (2001) Ku DNA end-binding protein modulates homologous repair of double-strand breaks in mammalian cells. *Genes Dev.*, **15**, 3237–3242.
22. Dominguez-Bendala,J., Masutani,M. and McWhir,J. (2006) Down-regulation of PARP-1, but not of Ku80 or DNA-PKcs, results in higher gene targeting efficiency. *Cell Biol. Int.*, **30**, 389–393.
23. Yamana,Y., Maeda,T., Ohba,H., Usui,T., Ogawa,H.I. and Kusano,K. (2005) Regulation of homologous integration in yeast by the DNA repair proteins Ku70 and RecQ. *Mol. Genet. Genomics*, **273**, 167–176.
24. Kooistra,R., Hooykaas,P.J. and Steensma,H.Y. (2004) Efficient gene targeting in *Kluyveromyces lactis*. *Yeast*, **21**, 781–792.
25. Ninomiya,Y., Suzuki,K., Ishii,C. and Inoue,H. (2004) Highly efficient gene replacements in *Neurospora* strains deficient for nonhomologous end-joining. *Proc. Natl Acad. Sci. USA*, **101**, 12248–12253.
26. Ishibashi,K., Suzuki,K., Ando,Y., Takakura,C. and Inoue,H. (2006) Nonhomologous chromosomal integration of foreign DNA is completely dependent on MUS-53 (human Lig4 homolog) in *Neurospora*. *Proc. Natl Acad. Sci. USA*, **103**, 14871–14876.
27. Krappmann,S., Sasse,C. and Braus,G.H. (2006) Gene targeting in *Aspergillus fumigatus* by homologous recombination is facilitated in a nonhomologous end-joining-deficient genetic background. *Eukaryot. Cell*, **5**, 212–215.
28. Takahashi,T., Masuda,T. and Koyama,Y. (2006) Enhanced gene targeting frequency in ku70 and ku80 disruption mutants of *Aspergillus sojae* and Aspergillus oryzae. *Mol. Genet. Genomics*, **275**, 460–470.
29. Poggeler,S. and Kuck,U. (2006) Highly efficient generation of signal transduction knockout mutants using a fungal strain deficient in the mammalian ku70 ortholog. *Gene*, **378**, 1–10.
30. Haarmann,T., Lorenz,N. and Tudzynski,P. (2008) Use of a non-homologous end joining deficient strain (Deltaku70) of the ergot fungus *Claviceps purpurea* for identification of a nonribosomal peptide synthetase gene involved in ergotamine biosynthesis. *Fungal Genet. Biol.*, **45**, 35–44.
31. So,S., Adachi,N., Lieber,M.R. and Koyama,H. (2004) Genetic interactions between BLM and DNA ligase IV in human cells. *J. Biol. Chem.*, **279**, 55433–55442.
32. Adachi,N., Kurosawa,A. and Koyama,H. (2008) Highly proficient gene targeting by homologous recombination in the human pre-B cell line Nalm-6. *Methods Mol. Biol.*, **435**, 17–29.
33. Southern,P.J. and Berg,P. (1982) Transformation of mammalian cells to antibiotic resistance with a bacterial gene under control of the SV40 early region promoter. *J. Mol. Appl. Genet.*, **1**, 327–341.
34. Tucker,K.L., Beard,C., Dausmann,J., Jackson-Grusby,L., Laird,P.W., Lei,H., Li,E. and Jaenisch,R. (1996) Germ-line passage is required for establishment of methylation and expression patterns of imprinted but not of nonimprinted genes. *Genes Dev.*, **10**, 1008–1020.
35. Matsuzaki,Y., Adachi,N. and Koyama,H. (2002) Vertebrate cells lacking FEN-1 endonuclease are viable but hypersensitive to methylating agents and H_2O_2. *Nucleic Acids Res.*, **30**, 3273–3277.
36. So,S., Nomura,Y., Adachi,N., Kobayashi,Y., Hori,T., Kurihara,Y. and Koyama,H. (2006) Enhanced gene targeting efficiency by siRNA that silences the expression of the Bloom syndrome gene in human cells. *Genes Cells*, **11**, 363–371.
37. Kurosawa,A., Koyama,H., Takayama,S., Miki,K., Ayusawa,D., Fujii,M., Iiizumi,S. and Adachi,N. (2008) The requirement of Artemis in double-strand break repair depends on the type of DNA damage. *DNA Cell Biol.*, **27**, 55–61.
38. Adachi,N., Suzuki,H., Iiizumi,S. and Koyama,H. (2003) Hypersensitivity of nonhomologous DNA end-joining mutants to VP-16 and ICRF-193: implications for the repair of topoisomerase II-mediated DNA damage. *J. Biol. Chem.*, **278**, 35897–35902.
39. Sonoda,E., Hochegger,H., Saberi,A., Taniguchi,Y. and Takeda,S. (2006) Differential usage of non-homologous end-joining and homologous recombination in double strand break repair. *DNA Repair*, **5**, 1021–1029.
40. Roth,D.B. and Wilson,J.H. (1985) Relative rates of homologous and nonhomologous recombination in transfected DNA. *Proc. Natl Acad. Sci. USA*, **82**, 3355–3359.
41. Allen,C., Kurimasa,A., Brenneman,M.A., Chen,D.J. and Nickoloff,J.A. (2002) DNA-dependent protein kinase suppresses double-strand break-induced and spontaneous homologous recombination. *Proc. Natl Acad. Sci. USA*, **99**, 3758–3763.
42. Delacote,F., Han,M., Stamato,T.D., Jasin,M. and Lopez,B.S. (2002) An xrcc4 defect or Wortmannin stimulates homologous recombination specifically induced by double-strand breaks in mammalian cells. *Nucleic Acids Res.*, **30**, 3454–3463.
43. Yan,C.T., Boboila,C., Souza,E.K., Franco,S., Hickernell,T.R., Murphy,M., Gumaste,S., Geyer,M., Zarrin,A.A., Manis,J.P. *et al.* (2007) IgH class switching and translocations use a robust non-classical end-joining pathway. *Nature*, **449**, 478–482.
44. Soulas-Sprauel,P., Le Guyader,G., Rivera-Munoz,P., Abramowski,V., Olivier-Martin,C., Goujet-Zalc,C., Charneau,P. and de Villartay,J.P. (2007) Role for DNA repair factor XRCC4 in immunoglobulin class switch recombination. *J. Exp. Med.*, **204**, 1717–1727.
45. Corneo,B., Wendland,R.L., Deriano,L., Cui,X., Klein,I.A., Wong,S.Y., Arnal,S., Holub,A.J., Weller,G.R., Pancake,B.A. *et al.* (2007) Rag mutations reveal robust alternative end joining. *Nature*, **449**, 483–486.
46. Sallmyr,A., Tomkinson,A.E. and Rassool,F.V. (2008) Up-regulation of WRN and DNA ligase IIIα in chronic myeloid leukemia: consequences for the repair of DNA double strand breaks. *Blood*, **112**, 1413–1423.
47. Wang,H., Rosidi,B., Perrault,R., Wang,M., Zhang,L., Windhofer,F. and Iliakis,G. (2005) DNA ligase III as a candidate component of backup pathways of nonhomologous end joining. *Cancer Res.*, **65**, 4020–4030.
48. Audebert,M., Salles,B. and Calsou,P. (2004) Involvement of poly(ADP-ribose) polymerase-1 and XRCC1/DNA ligase III in an alternative route for DNA double-strand breaks rejoining. *J. Biol. Chem.*, **279**, 55117–55126.
49. Merrihew,R.V., Marburger,K., Pennington,S.L., Roth,D.B. and Wilson,J.H. (1996) High-frequency illegitimate integration of transfected DNA at preintegrated target sites in a mammalian genome. *Mol. Cell. Biol.*, **16**, 10–18.
50. Barsh,G.S., Seeburg,P.H. and Gelinas,R.E. (1983) The human growth hormone gene family: structure and evolution of the chromosomal locus. *Nucleic Acids Res.*, **11**, 3939–3958.
51. Cross,M. and Renkawitz,R. (1990) Repetitive sequence involvement in the duplication and divergence of mouse lysozyme genes. *EMBO J.*, **9**, 1283–1288.
52. Lehrman,M.A., Goldstein,J.L., Russell,D.W. and Brown,M.S. (1987) Duplication of seven exons in LDL receptor gene caused by Alu-Alu recombination in a subject with familial hypercholesterolemia. *Cell*, **48**, 827–835.
53. Kang,Y.K., Park,J.S., Lee,C.S., Yeom,Y.I., Chung,A.S. and Lee,K.K. (1999) Efficient integration of short interspersed element-flanked foreign DNA via homologous recombination. *J. Biol. Chem.*, **274**, 36585–36591.
54. Wallenburg,J.C., Nepveu,A. and Chartrand,P. (1987) Integration of a vector containing rodent repetitive elements in the rat genome. *Nucleic Acids Res.*, **15**, 7849–7863.

55. Shen,M.R. and Deininger,P.L. (1992) An in vivo assay for measuring the recombination potential between DNA sequences in mammalian cells. *Anal. Biochem.*, **205**, 83–89.
56. Watson,J.E., Slorach,E.M., Maule,J., Lawson,D., Porteous,D.J. and Brookes,A.J. (1995) Human repeat-mediated integration of selectable markers into somatic cell hybrids. *Genome Res.*, **5**, 444–452.
57. Lander,E.S., Linton,L.M., Birren,B., Nusbaum,C., Zody,M.C., Baldwin,J., Devon,K., Dewar,K., Doyle,M., FitzHugh,W. *et al.* (2001) Initial sequencing and analysis of the human genome. *Nature*, **409**, 860–921.
58. Hillier,L.W., Miller,W., Birney,E., Warren,W., Hardison,R.C., Ponting,C.P., Bork,P., Burt,D.W., Groenen,M.A., Delany,M.E. *et al.* (2004) Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature*, **432**, 695–716.
59. Yáñez,R.J. and Porter,A.C.G. (2002) A chromosomal position effect on gene targeting in human cells. *Nucleic Acids Res.*, **30**, 4892–4901.
60. Thyagarajan,B., Johnson,B.L. and Campbell,C. (1995) The effect of target site transcription on gene targeting in human cells *in vitro*. *Nucleic Acids Res.*, **23**, 2784–2790.
61. Galagan,J.E., Calvo,S.E., Borkovich,K.A., Selker,E.U., Read,N.D., Jaffe,D., FitzHugh,W., Ma,L.J., Smirnov,S., Purcell,S. *et al.* (2003) The genome sequence of the filamentous fungus *Neurospora crassa*. *Nature*, **422**, 859–868.
62. Fattah,F.J., Lichter,N.F., Fattah,K.R., Oh,S. and Hendrickson,E.A. (2008) Ku70, an essential gene, modulates the frequency of rAAV-mediated gene targeting in human somatic cells. *Proc. Natl Acad. Sci. USA*, **105**, 8703–8708.