

# VU Research Portal

## Gene targeting by single-stranded DNA oligonucleotides in mouse embryonic stem cells

Aarts, M.

2010

### **document version**

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

### **citation for published version (APA)**

Aarts, M. (2010). *Gene targeting by single-stranded DNA oligonucleotides in mouse embryonic stem cells*.

### **General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

### **Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

### **E-mail address:**

[vuresearchportal.ub@vu.nl](mailto:vuresearchportal.ub@vu.nl)

Chapter

# 4

## Parameters of oligonucleotide-mediated gene modification in mouse ES cells

Marieke Aarts

Hein te Riele

*J Cell Mol Med* (2009) In press.



*Division of Molecular Biology, The Netherlands Cancer Institute,  
Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands.*



Gene targeting by single-stranded oligodeoxyribonucleotides (ssODNs) is emerging as a powerful tool for the introduction of subtle gene modifications in mouse embryonic stem (ES) cells and the generation of mutant mice. Here, we have studied the role of ssODN composition, transcription and replication of the target locus, and DNA repair pathways to gain more insight into the parameters governing ssODN-mediated gene targeting in mouse ES cells. We demonstrated that unmodified ssODNs of 35-40 nt were most efficient in correcting a chromosomally integrated mutant *neomycin* reporter gene. Addition of chemical modifications did not further enhance the efficacy of these ssODNs. The observed strand bias was not affected by transcriptional activity and may rather be caused by the different accessibility of the DNA strands during DNA replication. Consistently, targeting frequencies were enhanced when cells were treated with hydroxyurea to reduce the rate of replication fork progression. Transient down regulation of various DNA repair genes by RNAi had no effect on the targeting frequency. Taken together, our data suggest that ssODN-mediated gene targeting occurs within the context of a replication fork. This implies that any given genomic sequence, irrespective of transcriptional status, should be amenable to ssODN-mediated gene targeting. The ability of ES cells to differentiate into various cell types after ssODN-mediated gene targeting may offer opportunities for future therapeutic applications.

## INTRODUCTION

Gene modification by synthetic ssODNs (single-stranded oligodeoxyribonucleotides) is a promising alternative to existing strategies for the generation of subtle mutations in mouse ES cells [1,2]. These strategies are based on targeting vectors that upon homologous recombination integrate a selectable marker gene together with the desired genetic alteration into the gene of interest. Construction of these targeting vectors can be laborious and time-consuming. Moreover, although the marker gene is subsequently removed by Cre/lox-mediated recombination, traces of exogenous DNA are inevitably left behind in the target locus [3]. Alternatively, short ssODNs of 40 nt can be used to modify a single or a few bases in a specific location of the genome without altering the genomic organization.

In mouse ES cells, ssODN-mediated gene targeting frequencies range from  $10^{-7}$  to  $10^{-4}$  and are relatively low compared to

those found in other mammalian cell lines [4-7]. Previously, we have demonstrated that the DNA mismatch repair (MMR) system strongly impaired the introduction of genetic alterations by ssODNs in mouse ES cells [2,4]. Transfer of the genetic information requires base pairing between the ssODN and the chromosomal target locus resulting in the formation of DNA mismatches. The MMR system recognizes and removes these mismatches, thereby preventing stable inheritance of the desired genetic alteration. Knockdown of the central MMR gene *Msh2* by RNA interference dramatically improved the targeting frequency in mouse ES cells [1]. This effect has been confirmed in human hepatocytes [8] and in an episomal reporter system in mouse embryonic fibroblasts (MEFs) [9]. Thus, by transiently disabling the MMR system we could develop a simple and rapid procedure for the generation of mutant mouse ES cells. Importantly, the resulting mutant ES

cells retained their pluripotency and mutant alleles could successfully be transmitted through the mouse germ line [1,2]. Further improvement of the targeting frequency in mouse ES cells may greatly enhance the applicability of ssODN-mediated gene targeting, not only as a tool for biomolecular research but also for therapeutic purposes.

Thus far, parameters governing ssODN-mediated gene targeting have been studied in bacteria, yeast and a variety of mammalian cell lines. This has yielded conflicting data which may be ascribed to differences in ssODN composition, type of reporter system and model organism (reviewed in [10]). Many reports have shown that anti-sense ssODNs (*i.e.*, complementary to the non-transcribed strand) are more effective than sense ssODNs [5,11-13]. Protection of ssODNs against nucleolytic degradation by 2'-*O*-methyl RNA residues [5,11], phosphorothioate (PTO) linkages [12,13], or locked nucleic acid (LNA) bases [14,15] seemed to enhance targeting frequencies. Transcriptional activity promoted ssODN-mediated gene targeting in inducible episomal reporter systems in *Escherichia coli* [16,17], *Saccharomyces cerevisiae* [12] and mammalian cells [18]. However, in chromosomal reporter systems in *E. coli* it has been shown that the direction of replication rather than the direction of transcription influenced the targeting frequency [16,17,19,20]. By changing the orientation of the reporter gene towards the replication fork, higher targeting frequencies were observed when ssODNs with the same polarity as the nascent lagging strand were used. DNA replication also seems to play an important role in ssODN-mediated gene targeting in mammalian cells. Synchronization of cells in the S phase of the cell cycle or reducing the rate of replication fork progression improved the frequency of ssODN-mediated gene correction [21-24]. On the other hand, targeting frequencies

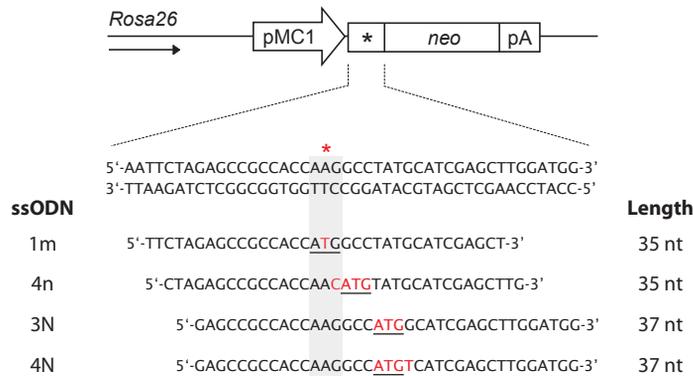
decreased when the replication machinery was blocked, indicating that DNA synthesis is required for effective ssODN-mediated gene correction. These findings suggest that the ssODN anneals to its single-stranded target region within the context of a replication fork. After annealing, the ssODN is extended by the replication machinery resulting in incorporation of the ssODN into the nascent DNA strand [20,25]. Nonetheless, DNA repair proteins may be involved in promoting various steps of the targeting process, since gene targeting frequencies were elevated upon overexpression of homologous recombination proteins [26,27] or endonucleases [28] and upon treatment with DNA damaging agents [29,30].

In this report, we have systematically investigated the role of various parameters in the targeting process in mouse ES cells, such as ssODN composition, transcription and replication of the target locus, and DNA repair pathways.

## RESULTS

### Reporter system

As a readout for the frequency of ssODN-mediated gene targeting, we used our previously described neomycin (*neo*) reporter system in which the *neo* gene was mutated by a single base substitution in the start codon (ATG to AAG; Figure 1). A single copy of this mutant *neo* gene was stably integrated into the *Rosa26* locus of *Msh2*-deficient and wild-type mouse embryonic stem (ES) cells [4]. ssODNs of 35-37 nt containing one, three or four centrally-located nucleotide substitutions were used to correct the mutant *neo* sequence (*i.e.*, generate a novel ATG start codon) resulting in G418 resistance. Targeting efficiencies were calculated by dividing the number of G418-resistant colonies by the number of cells that were plated after ssODN exposure. Using unmodified ssODNs,



**Figure 1 Mutant neomycin reporter gene and ssODN sequences.**

Sequence of the mutant neomycin (*neo*) reporter gene in which the start codon (ATG) is replaced for AAG (\*). Activity of the *neo* gene can be restored by single-stranded oligodeoxyribonucleotides (ssODNs) that substitute one to four nucleotides (indicated in red) to create a new ATG start codon (underlined). Sequences of sense ssODNs complementary to the transcribed strand are shown. Arrow indicates the direction of the *Rosa26* promoter.

*Msh2*-deficient ES cells showed targeting efficiencies of  $\sim 3 \times 10^{-5}$ , while wild-type ES cells typically showed targeting efficiencies of less than  $\sim 5 \times 10^{-7}$  [2,4].

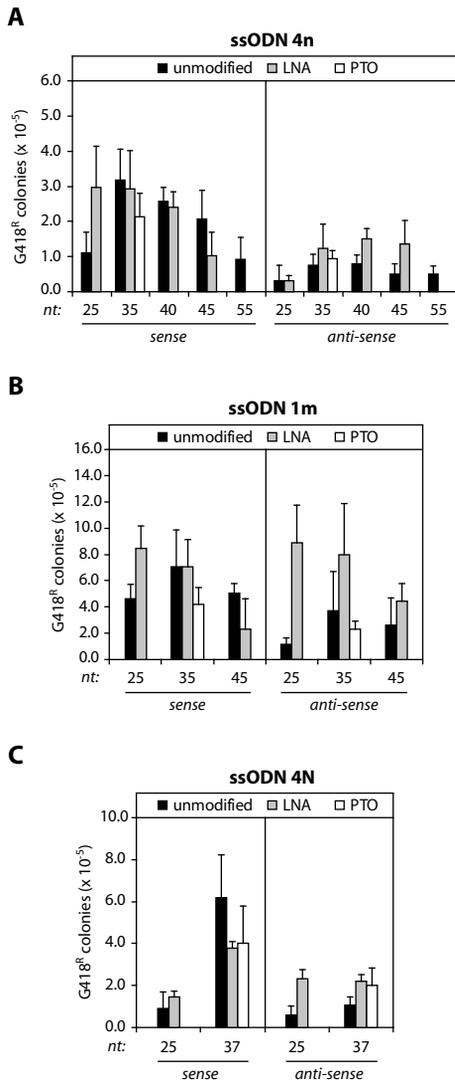
#### Chemical ssODN modifications

The length of the ssODN and its chemical composition may influence the targeting efficiency. To investigate this, we designed three series of ssODNs varying in length and chemical composition: ssODN 4n (Figure 2A), ssODN 1m (Figure 2B), and ssODN 4N (Figure 2C). For all these ssODNs, the optimal length of unmodified ssODNs was 35-40 nt and sense ssODNs were more effective than anti-sense ssODNs (Figure 2, black bars). We previously found a similar optimal length for an ssODN correcting a frameshift *neo* reporter [4].

Many reports have shown that gene targeting was only efficient when ssODNs were protected against nucleolytic degradation by flanking phosphorothioate (PTO) linkages or locked nucleic acid (LNA) bases [6,15,24]. We have tested ssODNs that were modified with a single LNA base at the 5' and 3' end

position (Figure 2, grey bars). Addition of LNA modifications had no significant effect on the efficiency of 35-40 nt sense ssODNs. However, the targeting frequencies of 35-40 nt anti-sense ssODNs were increased by LNA modifications, but they did not exceed the frequencies obtained with sense ssODNs. Short 25 nt ssODNs benefited most strongly from the addition of LNA modifications reaching frequencies similar to those obtained with 35-40 nt unmodified sense ssODNs (Figure 2A, 2B), except for ssODNs 4n anti-sense and 4N sense.

The effect of PTO modifications was tested by adding three PTO linkages to the 5' and 3' ends of ssODNs of 35 and 37 nt (Figure 2, white bars). PTO modifications decreased the targeting efficiency of sense ssODNs, as previously shown using the frameshift *neo* reporter [4], and hardly affected the efficiency of anti-sense ssODNs. Similar results were obtained when LNA- or PTO-modified ssODNs were tested in wild-type ES cells, excluding MMR-dependent effects (data not shown).



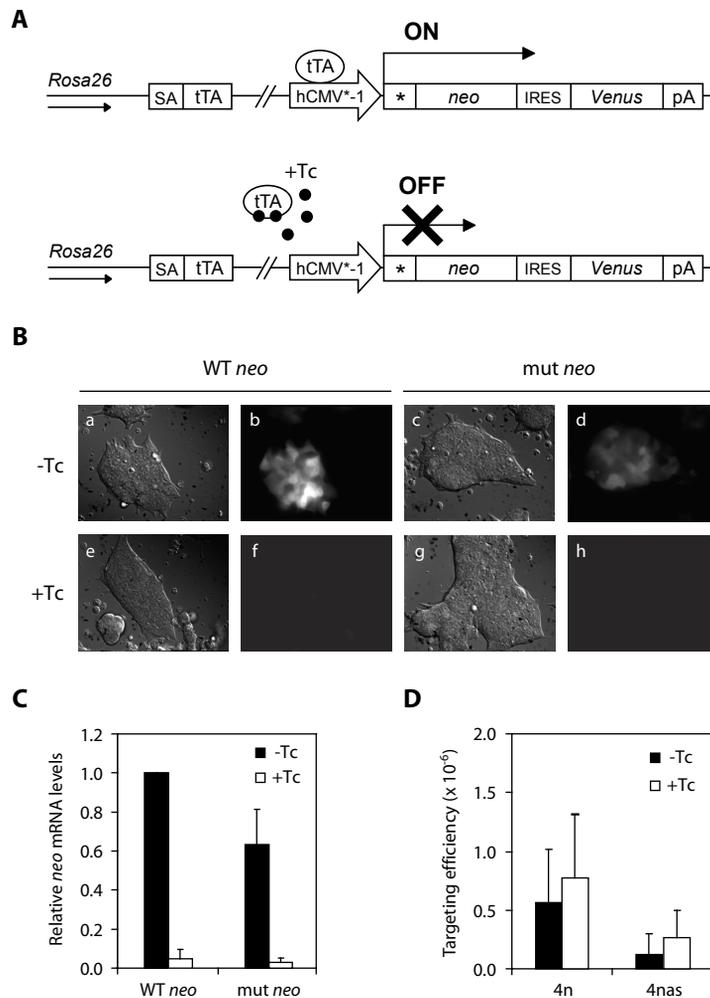
**Figure 2 Gene targeting by different types of ssODN in *Msh2*-deficient ES cells.**

(A) Unmodified (black bars), LNA-modified (grey bars) and PTO-modified (white bars) ssODNs 4n of various length (nt, nucleotides) and polarity (sense or anti-sense) were used to correct the mutant *neo* gene in *Msh2*-deficient ES cells. (B) Correction of the mutant *neo* gene by unmodified (black bars), LNA-modified (grey bars) and PTO-modified (white bars) ssODNs 1m. (C) Correction of the mutant *neo* gene by unmodified (black bars), LNA-modified (grey bars) and PTO-modified (white bars) ssODNs 4N. Targeting efficiency is the number of G418-resistant colonies per  $10^5$  cells that were plated after ssODN exposure. Error bars represent the standard deviation (s.d.) of at least three independent experiments.

Taken together, we have shown that unmodified sense ssODNs of 35-40 nt were most effective in correcting the mutant *neo* sequence in *Msh2*-deficient ES cells. Chemical ssODN modifications did not further improve the targeting efficiency and were not required for optimal ssODN-mediated gene targeting.

### Transcription

We showed that ssODNs in the sense orientation were more effective than ssODNs in the anti-sense orientation (Figure 2). In contrast, others have shown that anti-sense ssODNs, *i.e.*, complementary to the non-transcribed strand, were more effective than sense ssODNs [5,11-13]. This strand bias has been attributed to the increased accessibility of the non-transcribed strand for ssODN annealing during transcription, suggesting that transcription has a beneficial effect on ssODN-mediated gene targeting [18]. To gain insight into the role of transcription in ssODN-mediated gene targeting in ES cells, we used the ROSA-TET system described by Masui *et al.* [31]. In this Tet-off system, the *neo* gene is under the control of the hCMV\*-1 promoter containing a tetracycline response element (TRE) and the transcriptional transactivator tTA, stably integrated into the *Rosa26* locus of wild-type ES cells (Figure 3A). In culture medium without tetracycline (Tc), tTA binds to the hCMV\*-1 promoter and induces the expression of the *neo* reporter gene as well as *Venus*, whose fluorescence is used as an internal control for Tc-responsiveness.



**Figure 3** Effect of transcription on ssODN-mediated gene targeting in wild-type ES cells.

(A) A single copy of a tetracycline (Tc)-controllable mutant neo reporter gene was stably integrated into the *Rosa26* locus of wild-type ES cells. In the resulting ROSA-TET cell line, binding of tTA to the hCMV\*-1 promoter induces expression of a neo-IRES-Venus reporter gene in culture medium without Tc. (B) Tc-induced inhibition of Venus expression in ROSA-TET ES cells carrying a wild-type (WT) or mutant (*mut*) neo reporter gene. Cells were cultured for 24 hrs in medium without tetracycline (-Tc; panels a, b, c and d) or medium plus tetracycline (+Tc; panels e, f, g and h) before visualization of Venus fluorescence through a YFP filter (panels b, d, f and h). Corresponding bright field images are shown in panels a, c, e and g, respectively. (C) Quantitative RT-PCR showing Tc-induced inhibition of *neo* expression in ROSA-TET ES cells carrying a wild-type (WT) or mutant (*mut*) neo reporter gene. Mouse  $\beta$ -actin mRNA levels were used for normalization. (D) Correction of the mutant *neo* gene in ROSA-TET ES cells by ssODN 4n (sense orientation) or ssODN 4nas (anti-sense orientation) in -Tc medium (transcription "on") or +Tc medium (transcription "off"). Targeting efficiency is the number of G418-resistant colonies per  $10^6$  cells that were plated after ssODN exposure. Error bars represent the s.d. of nine independent experiments.

*Msh2*-proficient ROSA-TET ES cell lines were constructed containing either a wild-type (ATG) or a mutant (AAG) *neo* reporter gene. To test the Tc-responsiveness, ROSA-TET ES cells were cultured in -Tc (transcription “on”) or +Tc (transcription “off”) culture medium for 24 hrs. Addition of Tc to the culture medium resulted in a 20-fold reduced expression of *neo* in both the wild-type and the mutant *neo* ROSA-TET cell lines, as determined by Venus fluorescence (Figure 3B) and quantitative RT-PCR analysis of *neo* mRNA levels (Figure 3C).

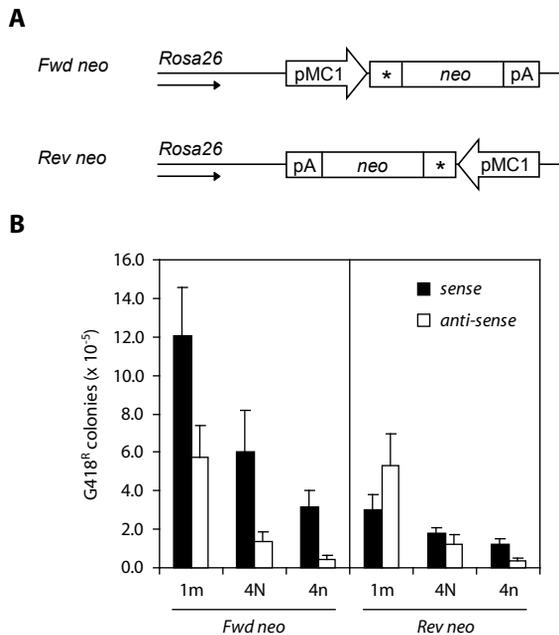
To investigate whether the correction frequency of the mutant *neo* reporter gene was influenced by the level of transcription, ROSA-TET ES cells were cultured in -Tc or +Tc culture medium from 24 hrs before to 24 hrs after ssODN transfection. Then, all cells were cultured in -Tc medium to induce expression of the *neo* gene before G418 selection was started (48 hrs after transfection). Although Tc addition resulted in a 20-fold decrease in *neo* expression levels, the targeting efficiencies of sense ssODN 4n and anti-sense ssODN 4nas were not significantly altered (Figure 3D). Independent of the level of transcriptional activity, sense ssODNs were still more effective than anti-sense ssODNs. Our findings demonstrate that transcription does not play a role in ssODN-mediated gene targeting in ES cells, suggesting that poorly or even non-transcribed genes are accessible to ssODN-mediated gene targeting.

### **Replication**

As the observed strand bias was not affected by differences in transcriptional activity, we investigated whether the different accessibility of the leading and lagging strand during DNA replication plays a role in dictating the strand bias. In *E. coli*, consistently higher targeting frequencies were obtained with ssODNs corresponding in sequence to the nascent lagging strand [16,17,19,20]. During discontinuous lagging strand synthe-

sis there are more single-stranded regions available than during the more continuous leading strand synthesis, thereby possibly facilitating annealing of the ssODN to its target site. To examine whether the strand bias correlated with the direction of replication, we constructed *Msh2*-deficient ES cell lines that contain a single copy of the *neo* gene in the opposite orientation at the same chromosomal position in the *Rosa26* locus (Figure 4A). In the resulting *Rev neo* and original *Fwd neo* reporter cell lines, the transcribed and non-transcribed strands are the same, while leading and lagging strands have (supposedly) interchanged. If the strand bias is directed by preferred incorporation of sense ssODNs into the nascent lagging strand in the *Fwd neo* cell line, then anti-sense ssODNs should be most effective in the *Rev neo* cell line.

As shown in Figure 4B, the efficiency of all sense ssODNs was 3- to 4-fold decreased in the *Rev neo* cell line compared to the *Fwd neo* cell line. While this may indicate that the strand bias was indeed dictated by the direction of replication, the decreased efficiency of the sense ssODNs was not accompanied by an increase in efficiency of the anti-sense ssODNs in the *Rev neo* cell line. In fact, the targeting efficiencies of the anti-sense ssODNs were not affected by reversal of the *neo* reporter. Consequently, the strand bias observed in the *Fwd neo* cell line was no longer present in the *Rev neo* cell line: sense and anti-sense ssODNs were equally effective. These findings were confirmed using cell lines carrying a frameshift *neo* reporter in opposite orientations into the *Rosa26* locus (Supplementary Figure 1). Our results obtained with the sense ssODNs indicate that the direction of replication is an important factor in dictating the strand bias in ssODN-mediated gene targeting; however, the results obtained with the anti-sense ssODNs remain enigmatic.



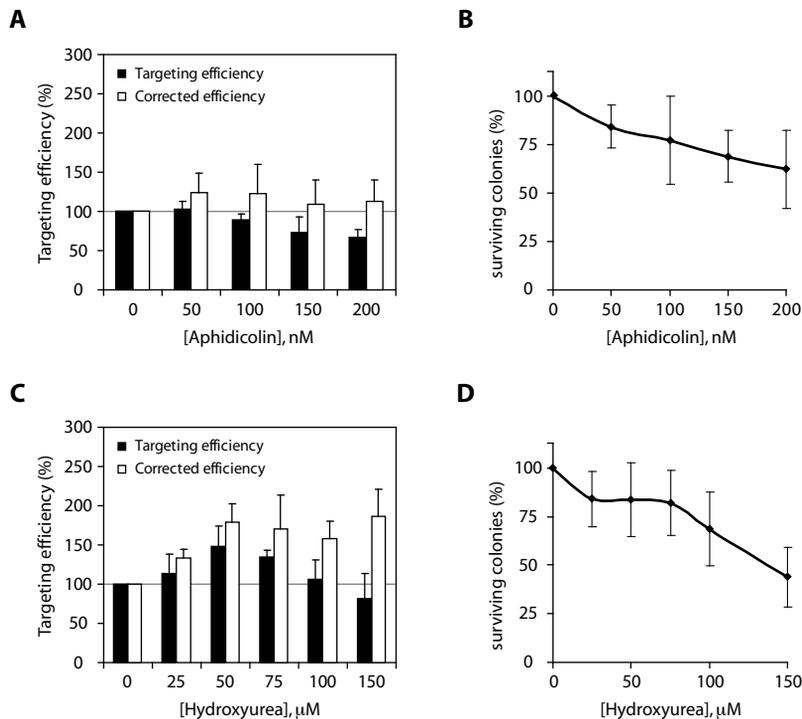
**Figure 4 Effect of the orientation of the neo reporter gene on ssODN-mediated gene targeting.** (A) Fwd and Rev neo cell lines contain a single copy of the mutant neo gene integrated in opposite directions at the same chromosomal position in the Rosa26 locus of Msh2-deficient ES cells. The direction of the Rosa26 promoter is indicated by an arrow. (B) Correction of the neo gene by sense ssODNs (black bars) or anti-sense ssODNs (white bars) in the Fwd and Rev neo cell lines. Targeting efficiency is the number of G418-resistant colonies per 10<sup>5</sup> cells that were plated after ssODN exposure. Error bars represent the s.d. of at least three independent experiments.

#### Slowing replication fork progression

More evidence that replication plays an important role in ssODN-mediated gene targeting comes from the observation that S phase appeared to be the preferred cell cycle phase for gene correction [23,24]. Synchronization of cells at the G1/S phase or slowing down replication fork progression has led to increased targeting frequencies in a variety of mammalian cell lines [21-24]. Since approximately 50% of asynchronously growing ES cells is already in S phase (unpublished observation), we attempted to reduce the rate of replication fork progression rather than increasing the number of cells progressing through S phase. We used low doses of two different replication in-

hibitors: aphidicolin, which directly inhibits the replicative polymerases  $\alpha$ ,  $\delta$  and  $\epsilon$  by competing with dCTP binding [32], and hydroxyurea, which inactivates ribonucleotide reductase thereby depleting cellular dNTP pools [33].

Msh2-deficient ES cells were exposed to increasing concentrations of aphidicolin or hydroxyurea between 6 hrs before and 22 hrs after transfection of ssODN 3N. Then, part of the cells was plated to assess the colony-forming ability while the remainder of the cells was plated in G418-containing medium to assess the neo correction frequency. In each independent experiment, the results were normalized to the targeting efficiency of untreated control cells. As



**Figure 5 Effect of replication inhibition on ssODN-mediated gene targeting.**

(A) *Msh2*-deficient ES cells carrying the mutant *neo* gene were incubated with increasing concentrations of aphidicolin or hydroxyurea (C) between 6 hrs before and 22 hrs after exposure to ssODN 3N. Relative targeting efficiencies are the number of G418-resistant colonies per  $10^5$  cells that were plated after ssODN exposure, normalized to the efficiency found in untreated control cells (black bars). Corrected targeting efficiencies are adjusted for the decreased colony survival after treatment with aphidicolin or hydroxyurea (white bars). Error bars represent the s.d. of three independent experiments. (B) Colony-forming ability of *Msh2*-deficient ES cells after exposure to increasing doses of aphidicolin or hydroxyurea (D) between 6 hrs before and 22 hrs after ssODN exposure. Error bars represent the s.d. of three independent experiments.

shown in Figure 5A, aphidicolin treatment resulted in a dose-dependent decrease in targeting efficiency compared to untreated control cells (Figure 5A, black bars). However, aphidicolin treatment also reduced the colony-forming ability of ES cells (Figure 5B). When targeting efficiencies were corrected for this decreased survival, aphidicolin treatment had no significant effect on the targeting efficiency (Figure 5A, white bars). In contrast, when cells were treated with hydroxyurea, targeting frequencies increased

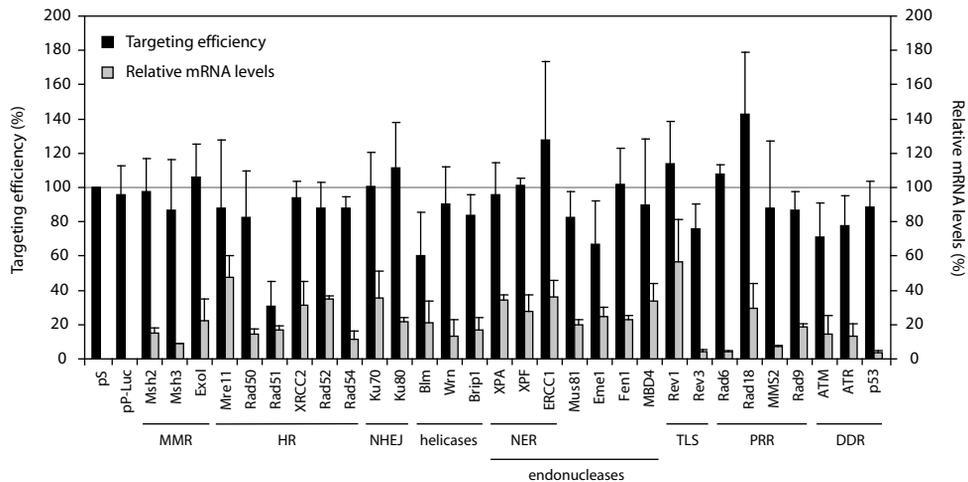
to 148% at 50  $\mu$ M hydroxyurea (Figure 5C, black bars). Higher concentrations of hydroxyurea resulted in a gradual decrease in targeting efficiency, reflecting the reduced colony survival at these doses (Figure 5D). After correction for this decreased survival, targeting efficiencies reached levels of 179% at 50  $\mu$ M hydroxyurea to 186% at 150  $\mu$ M hydroxyurea (Figure 5C, white bars). These findings suggest that the ssODN anneals to its homologous target region within the context of a replication fork.

### DNA repair pathways

Treatment with DNA damaging agents has been shown to stimulate ssODN-mediated gene targeting, which has been attributed to the induction of DNA double-strand breaks (DSBs) [29,30,34]. In addition, (over)expression of the homologous recombination (HR) proteins Rad51 and Rad54 may promote annealing of the ssODN with its target region [26,27], while nucleotide excision repair (NER) endonucleases ERCC1/XPF and XPG may be involved in resection of the target region to facilitate ssODN incorporation [28].

We examined the involvement of different DNA repair pathways in the ssODN-mediated gene targeting process in ES cells using transient down-regulation by RNA interference. The role of proteins involved

in MMR (Msh3, Exo1), HR (Mre11, Rad50, Rad51, XRCC2, Rad52, Rad54), non-homologous end-joining (Ku70, Ku80), NER (XPA, XPF, ERCC1), translesion synthesis (Rev1, Rev3), post-replicative repair (Rad6, Rad18, MMS2, Rad9), and DNA damage signaling (ATM, ATR, p53) was investigated as well as that of various DNA endonucleases (Mus81, Eme1, Fen1, MBD4) and DNA helicases (Blm, Wrn, Brip1). *Msh2*-deficient and wild-type ES cells containing the mutant *neo* reporter gene were transfected with a pSUPER vector expressing a short RNA hairpin, followed by puromycin selection for 2 days to eliminate untransfected cells. Then, cells were transfected with ssODN 3N to correct the *neo* reporter gene. In each experiment, empty pSUPER (pS) transfected ES cells were used for normalization of



**Figure 6 Effect of RNAi-mediated down-regulation of DNA repair genes on ssODN-mediated gene targeting.**

Correction of the mutant *neo* gene by ssODN 3N in *Msh2*-deficient ES cells after transient down-regulation of various DNA repair genes by vector-mediated RNAi. Relative targeting efficiencies are the number of G418-resistant colonies per  $10^5$  cells that were plated after ssODN exposure, normalized to the efficiency found in empty vector (pS)-transfected control cells (black bars). Error bars represent the s.d. of at least three independent experiments. Expression levels of the various DNA repair genes were quantified by RT-PCR and normalized using mouse  $\beta$ -actin mRNA levels (grey bars). MMR, mismatch repair; HR, homologous recombination; NHEJ, non-homologous end-joining; NER, nucleotide excision repair; TLS, translesion synthesis; PRR, post-replicative repair; DDR, DNA damage response.

the targeting efficiencies. Luciferase (pP-Luc) and *Msh2* knockdown vectors served as additional controls. As shown in Figure 6, none of the tested knockdown vectors resulted in a significant change in targeting efficiency in *Msh2*-deficient ES cells (Figure 6, black bars) or in wild-type ES cells (data not shown), while the median knockdown level was 79% (ranging from 96% to 43%) as determined by quantitative RT-PCR (Figure 6, grey bars). Down-regulation of Rad51 and Blm appeared to reduce the targeting efficiency in *Msh2*-deficient ES cells to 30 and 60%, respectively, however, this could be fully attributed to a reduction in the percentage of surviving colonies (29% and 53% compared to pS-transfected control cells, respectively).

In conclusion, knockdown of the investigated DNA repair proteins had no significant effect on the targeting efficiency in either *Msh2*-deficient or wild-type ES cells, indicating that these proteins are neither essential for nor detrimental to ssODN-mediated gene targeting.

## DISCUSSION

Numerous reports have demonstrated that ssODNs can be used for site-specific modification of a genomic target sequence. This technique would be particularly useful in ES cells for a number of reasons. First, ES cells are diploid, requiring only two gene copies to be modified for phenotypic analyses *in vitro*; second, mouse ES cells can be used to generate mice, allowing the modification to be studied in the context of a living animal; third, ES cells or ES-cell-like cells may become the preferred target cell for gene correction for therapeutic purposes. These applications will strongly benefit from in-depth insight into the parameters governing ssODN-directed gene modification in ES cells. We have previously shown that DNA

mismatch repair raises a strong barrier to ssODN-mediated gene targeting in ES cells and that this problem can be overcome by transient down-regulation of the central MMR gene *Msh2*. In the present study, we have systematically investigated the role of various parameters in the targeting process, such as transcription, replication and DNA repair pathways.

Using our *neo* reporter system in mouse ES cells, we have shown that ssODNs in the sense orientation (with respect to the direction of transcription of the *neo* gene) are more effective than ssODNs in the anti-sense orientation, independent of ssODN length (Figure 2). This preference for sense ssODNs has been confirmed in some reports [7,17,34,35], while others have shown a preference for targeting by anti-sense ssODNs [5,11-13]. A possible explanation for this difference in reported strand biases could be the type of readout used to determine the targeting frequency. The chromosomally integrated mutant *EGFP* [13,22] or *LacZ* [5,11] reporter genes that were used in these studies, allow detection of targeted cells by flow cytometry or (fluorescence) microscopy already 24 hrs after ssODN transfection. Since anti-sense ssODNs are incorporated into the transcribed strand, gene correction by these ssODNs will lead to instant EGFP/LacZ expression without the need for an extra round of DNA replication. In contrast, correction by sense ssODNs does require an extra round of replication to copy the base alteration(s) to the transcribed strand before EGFP/LacZ expression can be detected. Consequently, reporter genes with an early readout may skew the results towards anti-sense ssODNs being the most effective. In our assay, correction of the *neo* reporter gene is monitored by the appearance of G418-resistant colonies after 8-10 days, which eliminates this “early” transcriptional strand bias.

Besides polarity, we have also investigated the effect of ssODN length and chemical composition on the targeting frequency (Figure 2). Extending our previous results, we found that the optimal length of unmodified ssODNs was around 35-40 nt. LNA modifications did not affect the targeting efficiency of sense ssODNs of this size, but improved the efficiency of most anti-sense ssODNs. Addition of LNA modifications particularly improved the targeting efficiencies of short ssODNs of 25 nt. Shorter ssODNs may be more susceptible to nucleolytic degradation, which is prevented by the addition of chemical modifications. Additionally, LNA modifications may also increase the stability of the heteroduplex that is formed between the ssODN and its target sequence. Since 25 nt ssODNs can form fewer base pairs with the target sequence than their longer counterparts, LNA may be more beneficial for short ssODNs.

To investigate the role of transcription on the targeting frequency and the observed strand bias, we have developed a ROSA-TET cell line in which expression of the *neo* gene could be regulated in a reversible manner in the same clonal cell line. In this system, 20-fold increased transcription did clearly not promote ssODN-mediated gene targeting (Figure 3D), thereby challenging previous findings. When targeting frequencies of two different CHO-K1 clones carrying a randomly integrated mutant *LacZ* reporter gene were compared, higher targeting frequencies were observed in the clone that showed 20-fold higher expression levels [18]. Yet, this difference in targeting frequency may reflect locus-specific differences in chromatin state or clonal variations rather than an actual effect of transcription. Although transcriptional activity also promoted ssODN-mediated gene targeting in inducible episomal reporter systems [12,16-18], these findings may not apply to ssODN-mediated

targeting of chromosomally located genes [5].

In *E. coli*, a clear correlation was observed between the direction of replication and preferential targeting by ssODNs with the same sequence as the nascent lagging strand [16,17,19,20]. In contrast, the strand bias did not change in two CHO cell lines carrying an *mEGFP* reporter gene in opposite directions at the same chromosomal location [13]. In the *mEGFP*-rev CHO cell line, the correction frequency of anti-sense ssODNs was 2-fold lower than in the *mEGFP*-fwd CHO cell line, while correction by sense ssODNs was undetectable in both cell lines. Similarly, in our *Rev neo* ES cell line, sense ssODNs were less effective than in the *Fwd neo* ES cell line, while the frequency of anti-sense ssODNs remained unchanged (Figure 4B). Moreover, the strand bias that was evident in the *Fwd neo* cell line was absent in the *Rev neo* cell line. In mammalian cells, DNA replication is initiated at numerous replication origins. However, in a single cell cycle only a fraction of all potential replication origins in the genome are activated (reviewed in [36]). Epigenetic modifications play an important role in origin choice and timing of DNA replication. Possibly, reversal of the *neo* gene has led to changes in chromatin structure that may interfere with replication fork movement or choice of origin activation [37], thereby obscuring our results. Although the impact on the strand bias remained somewhat unclear, changing the orientation of the *neo* gene clearly had an impact on the targeting frequency. This suggests that the direction of replication does play a role in ssODN-mediated gene targeting.

More evidence for the involvement of replication in the targeting process comes from the observation that hydroxyurea treatment improved the targeting efficiency in *Msh2*-deficient ES cells in a dose-dependent

manner (Figure 5C). Yet, slowing down replication fork movement by aphidicolin had no effect on the targeting efficiency (Figure 5A). Replication fork stalling can induce uncoupling of the replicative polymerase and helicase activities, resulting in the formation of ssDNA [38,39]. Perhaps, higher doses of aphidicolin were required to effectively inhibit the replicative polymerases resulting in this uncoupling and the formation of ssDNA. Unfortunately, ES cells are highly sensitive to inhibition of replication fork progression and readily undergo apoptosis, since higher doses of replication inhibitors clearly reduced the viability of the cells (Figure 5B, 5D). On the other hand, even without addition of replication inhibitors or DNA damaging agents, already a relatively large proportion of ES cells progresses through S phase during ssODN exposure.

ES cells have robust mechanisms to repair DNA damage in order to prevent accumulation of mutations in their progeny. Although DNA repair proteins may facilitate ssODN incorporation by stabilizing the heteroduplex or by increasing the accessibility of the target locus [26-28], our results suggest that the DNA repair proteins tested are no major determinants in the targeting process in mouse ES cells. RNAi-mediated down-regulation of various DNA repair proteins did not significantly improve or decrease the targeting efficiency in either *Msh2*-deficient or wild-type ES cells (Figure 6).

In conclusion, we have demonstrated that unmodified ssODNs of 35-40 nt are most efficient in correcting a chromosomal *neo* reporter gene in mouse ES cells. Addition of chemical modifications does not further enhance the efficacy of these ssODNs. The observed strand bias is likely caused by the different accessibility of the DNA strands during DNA replication rather than during transcription of the target locus. This im-

plies that any given genomic sequence, irrespective of transcriptional status, should be amenable to ssODN-mediated gene targeting. The optimal orientation of the ssODN seems to be dictated by the direction of the replication fork through the target locus. Since this will be different for each locus, we recommend testing ssODNs in both orientations. Although treatment with hydroxyurea improved the targeting efficiency, we argue against standard application in mouse ES cells since these cells are highly sensitive to hydroxyurea. Taken together, our data support a model in which the ssODN is physically incorporated into the genomic target locus during replication.

## ACKNOWLEDGEMENTS

We thank Dr. Shinji Masui (Research Institute, International Research Center of Japan) for providing the EBRTcH3 ES cells and the pPthC-Oct3/4 and pCAGGS-Cre plasmids and also for helpful suggestions. We are grateful to Edith van de Vijver for her technical assistance in construction of the ROSA-TET cell lines. The authors would like to thank Sietske Bakker, Rob Dekker and Eva Wielders for critical reading of the manuscript.

This work was supported by grants from the Netherlands Genomics Initiative (Horizon Breakthrough project 050-71-007 and Horizon project 050-71-051).

## MATERIALS AND METHODS

### **Cell lines and culture conditions**

129/Ola-derived E14-IB10 ES cells [40] were cultured on MEF feeders in Glasgow minimal essential medium (GMEM) supplemented with 10% fetal calf serum (FCS), 1mM sodium pyruvate, 1x non-essential amino acids, 1 mM 2-mercaptoethanol and 1000 U per ml of leukemia inhibitory factor. For transfections and antibiotic selections, ES cells were cultured onto gelatin-coated plates in BRL (Buffalo rat liver cells)-conditioned medium [40].

### **Neo reporter cell lines**

We developed a selectable mutant neomycin (*neo*) reporter cell line in which the start codon of the *neo* resistance gene was mutated from ATG to AAG (Figure 1). A single copy of the mutant *neo* reporter gene was stably integrated into the *Rosa26* locus of *Msh2*-deficient and wild-type ES cells as described previously [4]. To create a *Rev neo* cell line that carries the *neo* gene in the opposite direction relative to the *Rosa26* promoter, we cloned the *neo* gene in the opposite orientation into the *Rosa26-His* targeting vector. Single copy integration was confirmed by Southern blot analysis.

### **ROSA-TET cell lines**

To establish ES cell lines that express the *neo* gene under the control of the Tc-controllable promoter, we used the ROSA-TET system described by Masui et al. [31]. E14tg2a-derived EBRTcH3 ES cells (carrying the tetracycline-controllable transactivator tTA and IRES-Venus ORFs in the *Rosa26* locus) and plasmids pPthC-Oct3/4 and pCAGGS-Cre were a kind gift from Dr. Masui. To construct an exchange vector with the mutant *neo* cDNA, the *neo* gene was amplified by PCR using forward primer 5'-CCGCTCGAGCGACCCAATTCTA-3' and reverse primer 5'-ATATGCGGCCGCTCAGAAGAAGCTGCAAG-3' and the *Rosa26-His-neo* targeting vector as template to introduce *Xho*I and *Not*I sites (underlined). The *Xho*I/*Not*I digested *neo* fragment was inserted into exchange vector pPthC-Oct-3/4 to obtain pPthC-(AAG)neo. Similarly, a pPthC-(ATG)neo exchange vector with the wild-type *neo* start codon was constructed for control purposes (PCR amplification with forward primer 5'-CCGCTCGAGATGGCCTATGCATCG-3' to introduce a correct **ATG** start codon). We established ROSA-TET ES cells by co-transfection of 5 µg exchange vector pPthC-neo and 2 µg pCAGGS-Cre with 10 µl of Lipofectamine 2000 (Invitrogen, Breda, The Netherlands) following previously described procedures [31]. Single copy integration was confirmed by Southern blot analysis. The recombinant cell lines were maintained in 1.5 mg/ml puromycin (Sigma-Aldrich, Zwijndrecht, The Netherlands). For targeting experiments,  $7 \times 10^5$  ROSA-TET ES cells were seeded onto a gelatin-coated 6-well in BRL-conditioned medium supplemented with 1 µg/ml tetracycline (+Tc; Sigma-Aldrich) or without tetracycline (-Tc) the day before ssODN transfection. Cells were exposed to ssODNs in +Tc or -Tc medium for 24 hrs as described elsewhere. Following transfection, cells were washed twice with phosphate buffered saline (PBS), counted and replated onto three gelatin-coated 100 mm dishes in medium without Tc.

### **Transfection**

ssODNs and pSUPER plasmids were transfected following the TransFast-mediated transfection method described earlier [1]. Generally,  $7 \times 10^5$  ES cells were seeded onto a gelatin-coated 6-well in BRL-conditioned medium the day before transfection. For one well, 3 µg of ssODNs and 27 µl of TransFast transfection agent (Promega, Leiden, The Netherlands) were diluted in 1.4 ml serum-free medium and incubated for 15 min at room temperature. After 75 min of exposure to the transfection mixture at 37 °C, 4 ml of BRL-conditioned medium with serum was added to the cells. The day after ssODN transfection, cells were counted and replated onto three gelatin-coated 100 mm dishes. G418 selection (800 µg/ml for *Fwd neo* and *Rev neo* cell lines, 2.5 mg/ml for ROSA-TET cell line; GIBCO-Invitrogen) was started 48 hrs after transfection. After 10 days, G418-resistant colonies were stained with Leishman's eosin methylene blue solution (Merck, Darmstadt, Germany) and counted. Transfection efficiencies were calculated by dividing the number of G418-resistant colonies by the number of cells that were

plated 24 hrs after ssODN transfection.

### ***Oligonucleotides***

ssODNs were designed to correct the mutant start region of the *neo* reporter gene by introducing an in-frame ATG start codon (sequences of sense ssODNs are listed in Figure 1). Anti-sense ssODNs introduce identical base changes but have opposite polarity (*i.e.*, are complementary to the non-transcribed strand). When indicated, ssODNs were modified with either three PTO linkages or a single LNA base on each end to protect them against nucleolytic degradation. Unmodified and PTO-modified ssODNs, deprotected and desalted, were obtained from Sigma-Aldrich, Inc. LNA-modified ssODNs were obtained from Sigma-Prologo and Eurogentec.

### ***Fluorescence microscopy***

ROSA-TET ES cells carrying the mutant or wild-type *neo* reporter gene were plated at a density of  $1.5 \times 10^5$  cells per well onto 35 mm glass-bottom Willco dishes in both +Tc and -Tc medium. After 24 hrs, dishes were transferred to a Zeiss Axiovert 200M microscope, equipped with a 0.55 numeric aperture condenser and a 40X Plan-Neo DIC objective (N.A. 1.3) using a Photometrics Coolsnap HQ CCD camera (Roper Scientific, Vianen, The Netherlands) with a CFP/YFP dual band pass filter set to select specific fluorescence. Images were processed using MetaMorph software (Universal Imaging, Downington, PA, USA).

### ***Drug treatment and colony survival***

The sensitivity of ES cells to increasing doses of replication inhibitors was determined by measuring their colony-forming ability. ES cells were plated at  $7 \times 10^5$  cells per well onto gelatin-coated 6-well plates 24 hrs prior to drug treatment. Cells were exposed to increasing doses of aphidicolin (100  $\mu$ M stock in DMSO; Sigma-Aldrich) or hydroxyurea (100 mM stock in PBS; Sigma-Aldrich) for 6 hrs, transfected with ssODNs as described above and exposed to drugs for another 22 hrs. Then, cells were counted and plated at 500 cells/100 mm dish to determine colony survival, whereas the remainder of the cells was plated for G418 selection. After 8 days, surviving colonies were stained and counted.

### ***RNA interference***

We used the pSUPER vector containing a puromycin resistance gene for expression of shRNAs [41]. For each target gene, three different pSUPER vectors were constructed (sequences available upon request). An empty vector without gene-specific 19 nt sequence served as a non-silencing control (pS), while luciferase (pP-Luc) and pS-MSH2 [1] vectors served as additional controls. For transient down regulation of the various DNA repair proteins, ES cells were first transfected with 3  $\mu$ g of pSUPER vector followed by selection with 20  $\mu$ g/ml of puromycin (Sigma-Aldrich) for 2 days. Cells were washed twice with PBS, trypsinized and plated at a density of  $7 \times 10^5$  cells per well for ssODN transfection as described above.

### ***RNA isolation and quantitative real-time PCR***

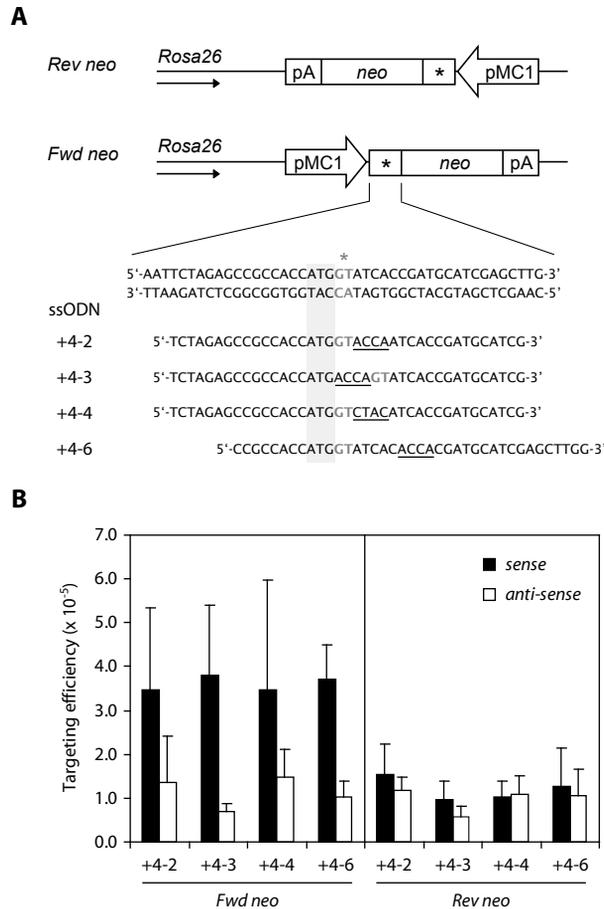
ES cells were transfected with the different pSUPER constructs as described above. After puromycin selection for 2 days, total RNA was extracted using RNA-Bee Total RNA Isolation Reagent (Campro Scientific, Veenendaal, The Netherlands). For each reaction, 1  $\mu$ g of RNA was used for first-strand cDNA synthesis using SuperScript II Reverse Transcriptase (Invitrogen) and random primers. Quantitative real-time PCR was performed in an ABI Prism 7000 Sequence Detection System (Applied Biosystems, Nieuwerkerk a/d IJssel, The Netherlands) using the standard amplification program. Amplification was performed in a final volume of 25  $\mu$ l, containing 15-30 ng of cDNA from the reversed transcribed reaction, 0.4  $\mu$ M each of forward and reverse primers, and 12.5  $\mu$ l of 2x SYBR Green PCR Master Mix (Applied Biosystems). Where possible, primers were designed to overlap adjacent exon boundaries to exclude detection of genomic DNA (Primer Express 1.0 software, Applied Biosystems). The final mRNA levels of the studied genes were normalized to mouse  $\beta$ -actin mRNA levels.

## REFERENCES

1. **Aarts M, Dekker M, de Vries S, van der Wal A, te Riele H.** (2006) Generation of a mouse mutant by oligonucleotide-mediated gene modification in ES cells. *Nucleic Acids Res.* 34, e147.
2. **Dekker M, Brouwers C, Aarts M, van der Torre J, de Vries S, van de Vrugt H, te Riele H.** (2006) Effective oligonucleotide-mediated gene disruption in ES cells lacking the mismatch repair protein MSH3. *Gene Ther.* 13, 686-94.
3. **Hasty P, Ramirez-Solis R, Krumlauf R, Bradley A.** (1991) Introduction of a subtle mutation into the Hox-2.6 locus in embryonic stem cells. *Nature.* 350, 243-6.
4. **Dekker M, Brouwers C, te Riele H.** (2003) Targeted gene modification in mismatch-repair-deficient embryonic stem cells by single-stranded DNA oligonucleotides. *Nucleic Acids Res.* 31, e27.
5. **Nickerson HD, Colledge WH.** (2003) A comparison of gene repair strategies in cell culture using a lacZ reporter system. *Gene Ther.* 10, 1584-91.
6. **Pierce EA, Liu Q, Igoucheva O, Omarrudin R, Ma H, Diamond SL, Yoon K.** (2003) Oligonucleotide-directed single-base DNA alterations in mouse embryonic stem cells. *Gene Ther.* 10, 24-33.
7. **Murphy BR, Moayedpardazi HS, Gewirtz AM, Diamond SL, Pierce EA.** (2007) Delivery and mechanistic considerations for the production of knock-in mice by single-stranded oligonucleotide gene targeting. *Gene Ther.* 14, 304-15.
8. **Igoucheva O, Alexeev V, Anni H, Rubin E.** (2008) Oligonucleotide-mediated gene targeting in human hepatocytes: implications of mismatch repair. *Oligonucleotides.* 18, 111-22.
9. **Maguire KK, Kmiec EB.** (2007) Multiple roles for MSH2 in the repair of a deletion mutation directed by modified single-stranded oligonucleotides. *Gene.* 386, 107-14.
10. **Igoucheva O, Alexeev V, Yoon K.** (2004) Oligonucleotide-directed mutagenesis and targeted gene correction: a mechanistic point of view. *Curr Mol Med.* 4, 445-63.
11. **Igoucheva O, Alexeev V, Yoon K.** (2001) Targeted gene correction by small single-stranded oligonucleotides in mammalian cells. *Gene Ther.* 8, 391-9.
12. **Liu L, Rice MC, Drury M, Cheng S, Gamper H, Kmiec EB.** (2002) Strand bias in targeted gene repair is influenced by transcriptional activity. *Mol Cell Biol.* 22, 3852-63.
13. **Olsen PA, Randol M, Luna L, Brown T, Krauss S.** (2005) Genomic sequence correction by single-stranded DNA oligonucleotides: role of DNA synthesis and chemical modifications of the oligonucleotide ends. *J Gene Med.* 7, 1534-44.
14. **Parekh-Orlmedo H, Drury M, Kmiec EB.** (2002) Targeted nucleotide exchange in *Saccharomyces cerevisiae* directed by short oligonucleotides containing locked nucleic acids. *Chem Biol.* 9, 1073-84.
15. **Andrieu-Soler C, Casas M, Faussat AM, Gandolphe C, Doat M, Tempe D, Giovannangeli C, Behar-Cohen F, Concordet JP.** (2005) Stable transmission of targeted gene modification using single-stranded oligonucleotides with flanking LNAs. *Nucleic Acids Res.* 33, 3733-42.
16. **Huen MS, Lu LY, Liu DP, Huang JD.** (2007) Active transcription promotes single-stranded oligonucleotide mediated gene repair. *Biochem Biophys Res Commun.* 353, 33-9.
17. **Li XT, Costantino N, Lu LY, Liu DP, Watt RM, Cheah KS, Court DL, Huang JD.** (2003) Identification of factors influencing strand bias in oligonucleotide-mediated recombination in *Escherichia coli*. *Nucleic Acids Res.* 31, 6674-87.
18. **Igoucheva O, Alexeev V, Pryce M, Yoon K.** (2003) Transcription affects formation and processing of intermediates in oligonucleotide-mediated gene alteration. *Nucleic Acids Res.* 31, 2659-70.
19. **Ellis HM, Yu D, DiTizio T, Court DL.** (2001) High efficiency mutagenesis, repair, and engineering of chromosomal DNA using single-stranded oligonucleotides. *Proc Natl Acad Sci U S A.* 98, 6742-6.
20. **Huen MS, Li XT, Lu LY, Watt RM, Liu DP, Huang JD.** (2006) The involvement of replication in single stranded oligonucleotide-mediated gene repair. *Nucleic Acids Res.* 34, 6183-94.
21. **Brachman EE, Kmiec EB.** (2005) Gene repair in mammalian cells is stimulated by the elongation of S phase and transient stalling of replication forks. *DNA Repair (Amst).* 4, 445-57.
22. **Engstrom JU, Kmiec EB.** (2008) DNA replication, cell cycle progression and the targeted gene repair reaction. *Cell Cycle.* 7, 1402-14.
23. **Wu XS, Xin L, Yin WX, Shang XY, Lu L, Watt RM, Cheah KS, Huang JD, Liu DP, Liang CC.** (2005) Increased efficiency of oligonucleotide-mediated gene repair through slowing replication fork progression. *Proc Natl Acad Sci U S A.* 102, 2508-13.
24. **Olsen PA, Randol M, Krauss S.** (2005) Implications of cell cycle progression on functional sequence correction by short single-stranded DNA oligonucleotides. *Gene Ther.* 12, 546-51.

25. **Radecke S, Radecke F, Peter I, Schwarz K.** (2006) Physical incorporation of a single-stranded oligodeoxynucleotide during targeted repair of a human chromosomal locus. *J Gene Med.* 8, 217-28.
26. **Liu L, Cheng S, van Brabant AJ, Kmiec EB.** (2002) Rad51p and Rad54p, but not Rad52p, elevate gene repair in *Saccharomyces cerevisiae* directed by modified single-stranded oligonucleotide vectors. *Nucleic Acids Res.* 30, 2742-50.
27. **Liu L, Maguire KK, Kmiec EB.** (2004) Genetic re-engineering of *Saccharomyces cerevisiae* RAD51 leads to a significant increase in the frequency of gene repair in vivo. *Nucleic Acids Res.* 32, 2093-101.
28. **Igoucheva O, Alexeev V, Scharer O, Yoon K.** (2006) Involvement of ERCC1/XPF and XPG in oligodeoxynucleotide-directed gene modification. *Oligonucleotides.* 16, 94-104.
29. **Ferrara L, Parekh-Olmedo H, Kmiec EB.** (2004) Enhanced oligonucleotide-directed gene targeting in mammalian cells following treatment with DNA damaging agents. *Exp Cell Res.* 300, 170-9.
30. **Ferrara L, Kmiec EB.** (2004) Camptothecin enhances the frequency of oligonucleotide-directed gene repair in mammalian cells by inducing DNA damage and activating homologous recombination. *Nucleic Acids Res.* 32, 5239-48.
31. **Masui S, Shimosato D, Toyooka Y, Yagi R, Takahashi K, Niwa H.** (2005) An efficient system to establish multiple embryonic stem cell lines carrying an inducible expression unit. *Nucleic Acids Res.* 33, e43.
32. **Oguro M, Suzuki-Hori C, Nagano H, Mano Y, Ikegami S.** (1979) The mode of inhibitory action by aphidicolin on eukaryotic DNA polymerase alpha. *Eur J Biochem.* 97, 603-7.
33. **Rosenkranz HS, Levy JA.** (1965) Hydroxyurea: A Specific Inhibitor of Deoxyribonucleic Acid Synthesis. *Biochim Biophys Acta.* 95, 181-3.
34. **Radecke F, Peter I, Radecke S, Gellhaus K, Schwarz K, Cathomen T.** (2006) Targeted chromosomal gene modification in human cells by single-stranded oligodeoxynucleotides in the presence of a DNA double-strand break. *Mol Ther.* 14, 798-808.
35. **Brachman EE, Kmiec EB.** (2003) Targeted nucleotide repair of *cyc1* mutations in *Saccharomyces cerevisiae* directed by modified single-stranded DNA oligonucleotides. *Genetics.* 163, 527-38.
36. **Aladjem MI.** (2007) Replication in context: dynamic regulation of DNA replication patterns in metazoans. *Nat Rev Genet.* 8, 588-600.
37. **Courbet S, Gay S, Arnoult N, Wronka G, Anglana M, Brison O, Debatisse M.** (2008) Replication fork movement sets chromatin loop size and origin choice in mammalian cells. *Nature.* 455, 557-60.
38. **Walter J, Newport J.** (2000) Initiation of eukaryotic DNA replication: origin unwinding and sequential chromatin association of Cdc45, RPA, and DNA polymerase alpha. *Mol Cell.* 5, 617-27.
39. **Byun TS, Pacek M, Yee MC, Walter JC, Cimprich KA.** (2005) Functional uncoupling of MCM helicase and DNA polymerase activities activates the ATR-dependent checkpoint. *Genes Dev.* 19, 1040-52.
40. **Hooper M, Hardy K, Handyside A, Hunter S, Monk M.** (1987) HPRT-deficient (Lesch-Nyhan) mouse embryos derived from germline colonization by cultured cells. *Nature.* 326, 292-5.
41. **Brummelkamp TR, Bernards R, Agami R.** (2002) A system for stable expression of short interfering RNAs in mammalian cells. *Science.* 296, 550-3.

## SUPPLEMENTARY INFORMATION



### Supplementary Figure 1 Opposite orientation of a frameshift neo reporter gene.

(A) Sequence of the mutant neomycin (*neo*) reporter gene that carries a +2 nt frameshift mutation (indicated in grey) after the ATG start codon. Fwd and Rev *neo* cell lines contain a single copy of the mutant *neo* gene integrated in opposite directions at the same chromosomal position in the *Rosa26* locus of *Msh2*-deficient ES cells. The direction of the *Rosa26* promoter is indicated by an arrow. Activity of the *neo* gene can be restored by single-stranded oligodeoxyribonucleotides (ssODNs) that insert four extra nucleotides (underlined). Sequences of sense ssODNs complementary to the transcribed strand are shown. (B) Correction of the *neo* gene by sense ssODNs (black bars) or anti-sense ssODNs (white bars) in the Fwd and Rev *neo* cell lines. Targeting efficiency is the number of G418-resistant colonies per 10<sup>5</sup> cells that were plated after ssODN exposure. Error bars represent the s.d. of at least three independent experiments.

