Gene Therapy (2007) 14, 304–315
© 2007 Nature Publishing Group All rights reserved 0969-7128/07 $30.00
www.nature.com/gt

ORIGINAL ARTICLE

Delivery and mechanistic considerations for the production of knock-in mice by single-stranded oligonucleotide gene targeting

BR Murphy[1], HS Moayedpardazi[2], AM Gewirtz[3], SL Diamond[1] and EA Pierce[2]
[1]Department of Chemical and Biomolecular Engineering, Institute for Medicine and Engineering, University of Pennsylvania, Philadelphia, PA, USA; [2]FM Kirby Center for Molecular Ophthalmology, Scheie Eye Institute, 305 Stellar-Chance Laboratory, University of Pennsylvania, Philadelphia, PA, USA and [3]Department of Medicine, Hematology/Oncology, Abramson Cancer Center, University of Pennsylvania, Philadelphia, PA, USA

Single-stranded oligodeoxynucleotide (ssODN) gene targeting may facilitate animal model creation and gene repair therapy. Lipofection of ssODN can introduce point mutations into target genes. However, typical efficiencies in mouse embryonic stem cells (ESC) are $<10^{-4}$, leaving corrections too rare to effectively identify. We developed ESC lines with an integrated mutant neomycin resistance gene (Tyr22Ter). After targeting with ssODN, repaired cells survive selection in G418. Correction efficiencies varied with different lipofection procedures, clonal lines, and ssODN designs, ranging from 1 to 100 corrections per million cells plated. Uptake studies using cell sorting of Cy5-labelled ssODN showed 40% of the corrections concentrated in the best transfected 22% of cells. Four different basepair mismatches were tested and results show that the base-specificity of the mismatch is critical. Dual mismatch ssODN also showed mismatch preferences. These ESC lines may facilitate development of improved ssODN targeting technologies for either animal production or ex vivo *gene therapy.*
Gene Therapy (2007) **14,** 304–315. doi:10.1038/sj.gt.3302866; published online 5 October 2006

Keywords: *gene targeting; embyronic stem cells; single-stranded oligonucleotide; gene repair*

Introduction

In mouse embryonic stem cells (ESC), gene targeting has been used to create numerous knockout and knock-in mice for studies of protein function and disease etiology.[1] Gene targeting also has potential for *ex vivo* gene repair, especially of mutations with dominant negative effects which cannot be addressed by augmentation gene therapy.[2] Unfortunately, gene targeting remains a complicated and costly process, particularly for knock-in mutations, and requires significant skill, labor, and time. Ideally, gene targeting procedures would be faster, more versatile, more reliable and less expensive than current methods. The long-term goal of developing oligonucleotide-mediated gene targeting for making knock-in mice motivates this study.

Oligonucleotide gene targeting is a process in which a short (20- to 80-mer) fragment of DNA is targeted to a homologous gene and introduces a desired sequence alteration at a target site. A variety of oligonucleotide chemistries exist, including single-stranded deoxyoligonucleotides (ssODN),[3,4] triplex-forming oligonucleotides[5] and chimeric RNA–DNA molecules.[6] These technologies have all demonstrated gene targeting activity, but are subject to low efficiencies. Efficiencies of ssODN targeting vary widely with cell type. Some targets and cells give efficiencies as high as a few percent.[7,8] Unfortunately, efficiencies in mouse ESC are particularly low. Our lab has demonstrated gene targeting in mouse ESC by correcting a mutant *LacZ* marker line (efficiency $\sim 5\times10^{-5}$).[9] Dekker *et al.*[10] targeted a mutant neomycin marker gene in murine MSH2$^{-/-}$ ESC (efficiency $\sim 2\times10^{-5}$), but did not observe corrections in wild-type ESC. Episomal targeting is usually easier than chromosomal targeting, but that also has low efficiencies in mouse ESC.[11] Improvement of ssODN gene targeting to efficiencies of $\sim 10^{-3}$ would enable PCR-based screening and allow the targeting of endogenous, non-marker genes for the production of animal models.

The task of increasing oligonucleotide-mediated targeting efficiency is complicated by our limited understanding of the mechanistic steps of gene targeting. Three factors are necessary for successful targeting. The oligonucleotide must (1) enter the nucleus, (2) pair with its homologous target and (3) induce a base change at the targeted site via endogenous machinery.[12,13] It is not clear which of these barriers is most rate-limiting in ESC.

A number of studies have investigated each of these individual steps. Transfection techniques are important, and the addition of nuclear localization signals to transfection reagents can increase plasmid transfection.[14]

Correspondence: Dr EA Pierce, FM Kirby Center for Molecular Ophthalmology, Scheie Eye Institute, 305 Stellar-Chance Laboratory, 422 Curie Blvd, University of Pennsylvania, Philadelphia, PA 19104, USA.
E-mail: epierce@mail.med.upenn.edu
Received 15 January 2006; revised 30 May 2006; accepted 20 July 2006; published online 5 October 2006

Protective phosphorothioate linkages aid transfections and increase efficiencies.[15] The pairing of ssODNs to their target is more likely if the target is accessible in a single-stranded state. Both transcription[16,17] and replication[8,16,18] have been linked to higher efficiencies. Less is known about the actual conversion event. The ssODN may replace the target; it may induce a repair mechanism to alter the target strand; and/or it may become incorporated into one daughter cell during replication.[13] Gene correction may occur by more than one mechanism. There appears to be a 'strand bias'[12,16] in which targeting with antisense ssODNs often gives better results than sense ssODNs. Furthermore, some mismatches appear to be more efficiently repaired when compared to others.[7,12]

We have created a selectable mutant neomycin mouse ESC line, similar to the approach used by Dekker.[10] The goal of this study was to use the marker line to identify the most critical barriers to gene targeting. By altering conditions that are uniquely associated with delivery or the targeting mechanism and by studying the system's sensitivity to these variables, we have begun to identify the parameters with the most potential for improvement. We have demonstrated that transfection conditions, ssODN dose, clonal characteristics and the specific sequence mismatch can all affect the efficiency of gene targeting.

Results

Construction and validation of pNeoS1Pur ESC lines

Puromycin-resistant, neomycin-susceptible plasmids were created by site-directed mutagenesis (Figure 1) and stably transfected into AB2.2 mouse ESC. To confirm the neomycin resistance gene was fully inactivated, the cells were split onto two dishes and selected in puromycin (2 μg/ml) or G418 (200 μg/ml). The puromycin dishes yielded approximately 100 colonies per dish, whereas G418 selection did not yield colonies. Colonies were individually picked from the puromycin dish, and expanded into distinct clonal lines of neomycin-susceptible (NeoS) cells.

Seven pNeoS1Pur clones were functionally tested with correction ssODNs. These clones were plated at 1 million cells per dish and lipofected with Neo1Cor ssODN (Figure 1). Four of the clones were correction competent and had colonies capable of surviving in G418. Positives were marked by the formation of large colonies, visible to the naked eye after 7–9 days of selection (Figure 2a). Initial efficiencies in the pNeoS1Pur lines were $\sim$1–5 positives per million cells plated.

The positives from functional testing displayed a survival phenotype in G418 and grew into colonies with normal AB2.2 morphology (Figure 2b). Eighteen corrected colonies were picked, passed twice on feeder-free plates and then sequenced. The Neo1Cor ssODN was useful for sequencing verification because the 'corrected' genotype (TAC, Figure 2c) was distinct from the 'wild-type' (TAT) present in the feeder cells and in the original plasmid. Sixteen of the 18 colonies had the expected G$\rightarrow$C genotype. Two colonies had spontaneous G$\rightarrow$T wild-type reversions, corresponding to a G$\rightarrow$T false positive rate of $\sim 3 \times 10^{-7}$. To further quantify the background level of spontaneous reversion, 5 million pNeoS1Pur cells were transfected with a control ssODN, Neo1Cont (G$\rightarrow$G). Two colonies survived selection, giving a spontaneous reversion rate of approximately 4×10^{-7}.

As the pNeoS1Pur plasmid was not targeted to a specific locus, it integrated illegitimately in the genome. A Southern analysis of three correction competent clones (clones 6, 9 and 10) (Figure 2d), demonstrated different band sizes, consistent with a different locus of integration for each of the three clones. The dual bands for clone 6 indicated two plasmid integrations in this clone.

Generation of chimeric mice

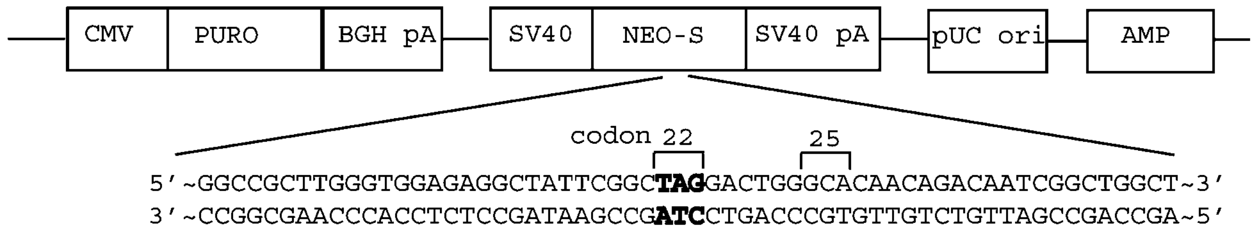
Chimeric mice were generated from both the pNeoS1Pur clone 9 cells (mutant) and corrected colonies of the pNeoS1Pur clone 9 line (corrected). The corrected ESC were made by transfection with Neo1Cor and selection in G418. Surviving clones were picked, grown and sequenced. One positive clone, εA11, was injected into blastocysts to produce chimeric mice. Four chimeric founders were generated, including three males with $\geqslant$50% chimerism.[19] To date, germline transmission has not been observed from the mutant or corrected chimeras. Additional breeding is in progress.

Correction efficiency depends on transfection procedures and DNA and lipid dosages

Procedures with ESC must be tailored to prevent differentiation. Overpopulation, death of the feeder cell layer, or serum starvation result in unhealthy cells with differentiated phenotypes (data not shown). However, lipofections are typically most efficient in reduced serum media (i.e. OptiMEM).[20] Transfection in media with 15% serum did not yield any corrections. We therefore used a 2-h transfection in OptiMEM, followed by the addition of serum-containing media to the lipoplex, and an overnight post-incubation. The time between transfection of ssODN and initiation of selection was another parameter to optimize. We have observed surviving colonies with selections as early as 12 h post-lipofection. However, waiting 36–48 h post-lipofection before G418-selection doubled the number of positives (data not shown).

It is unknown to what extent gene correction may be enhanced by improvement in ssODN delivery. The relative importance of delivery can be indirectly assessed by measuring how sensitive correction is to various formulations. Lipofections are sensitive to a number of variables; reagent concentrations are among the most critical.[21] Dose responses for plasmid expression tend to have an optimum balanced between low uptake at low reagent doses and cytotoxicity at high doses. Table 1 shows the effect of various oligonucleotide (Neo1WT) and lipid concentrations on ESC targeting. Experiments were conducted in six-well plates in triplicate, simultaneously. The highest efficiency for correction was at 2 μg lipid and 1 μg ssODN per cm^2. Higher dosages had the undesirable side effect of killing most of the feeder cells. Feeder cells are larger and tend to take up more lipoplex. They are also mitotically inactive and thus cannot regenerate. Many of the ESC colonies in higher dose wells had unusual morphology, indicative of differentiation and potential loss of pluripotency (data not shown).

a *pNeoS1Pur*



b *Correction Oligonucleotides*

Neo1Cor-SNS	5′-	gcttGGGTGGAGAGGCTATTCGGC**TAC**GACTGGGCACAACAGACAATCGgctg -3′
Neo1WT-SNS	5′-	gcttGGGTGGAGAGGCTATTCGGC**TAT**GACTGGGCACAACAGACAATCGgctg -3′
Neo1Cor	3′-	cgaaCCCACCTCTCCGATAAGCCG**ATG**CTGACCCGTGTTGTCTGTTAGCcgac -5′
Neo1WT	3′-	cgaaCCCACCTCTCCGATAAGCCG**ATA**CTGACCCGTGTTGTCTGTTAGCcgac -5′
Neo1Cont	3′-	cgaaCCCACCTCTCCGATAAGCCG**ATC**CTGACCCGTGTTGTCTGTTAGCcgac -5′
Neo1WT-Cy5	3′-	cgaaCCCACCTCTCCGATAAGCCG**ATA**CTGACCCGTGTTGTCTGTTAGCcgac-Cy5-5′
Neo1WT-9T	3′-	cgaaCCCACCTCTCCGATAAGCCG**ATA**CTGACCCG**A**GTTGTCTGTTAGCcgac -5′
Neo1WT-9C	3′-	cgaaCCCACCTCTCCGATAAGCCG**ATA**CTGACCCG**G**GTTGTCTGTTAGCcgac -5′
Neo1WT-9G	3′-	cgaaCCCACCTCTCCGATAAGCCG**ATA**CTGACCCG**C**GTTGTCTGTTAGCcgac -5′

c *Mismatches*

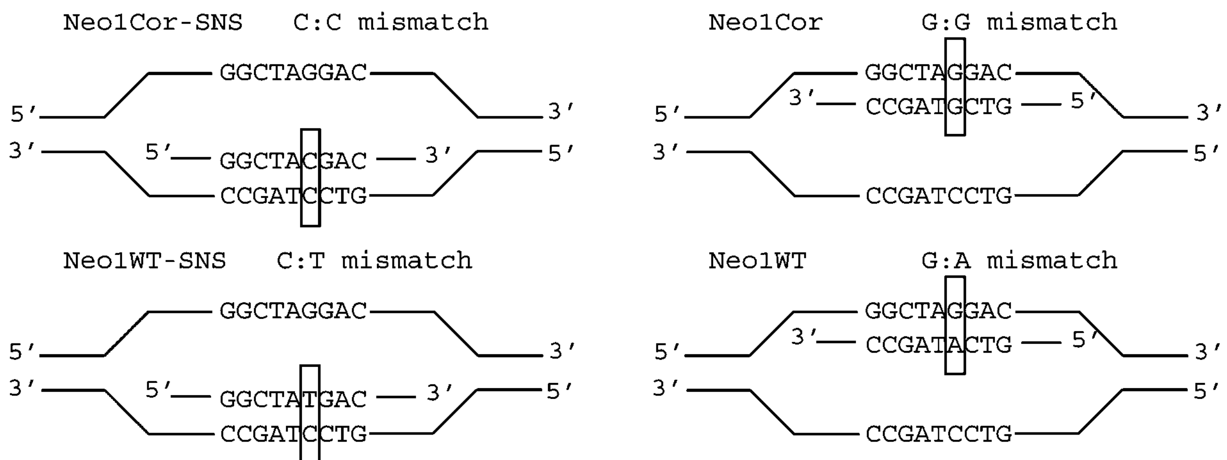


Figure 1 Plasmids and ssODNs used in ESC targeting. (**a**) The pNeoS1Pur plasmid contained a TAG stop at codon 22 of the neomycin resistance gene. (**b**) The ssODN were all 53 bases long with the four bases on each end protected with phosphorothioate linkages (represented by lowercase letters). Both antisense and sense (marked SNS) ssODN were tested. 'WT' ssODN caused a TAG→TAT correction. 'Cor' ssODN cause a TAG→TAC correction. Neo1Cont is a control, fully homologous to the TAG mutant. (**c**) Schematic representation of sense or antisense ssODNs aligned with their target strand. The mismatch nomenclature is the chromosomal base followed by the ssODN base.

Cellular uptake affects correction efficiency on a per cell basis

Within a single lipofection, the distribution of ssODN per cell is quite broad, with some cells taking up more than 100-fold more ssODN than others.[21] To determine if higher oligonucleotide uptake correlated with higher correction, we sorted ESC based on fluorescent oligonucleotide uptake. We transfected four dishes of pNeoS1Pur (clone 9) cells with a mixture of 80% Neo1WT and 20% Neo1WT-Cy5. Three of the plates (18 million cells) were trypsinized for cell sorting. Approximately 97% of the feeders were separated from the ESC by a forward-scatter/side scatter gate. The remaining 11 million cells were divided into four bins by their Cy5 fluorescence (Figure 3a–c). The bins were plated onto eight separate dishes (two for each bin) and, after 24 h, selected with G418. The fourth transfected dish, a no-sort control, began selection at the same time. There was a 1 day lag in the growth of the sorted cells, with colonies appearing on days 8 and 9 instead of the typical day 7 and 8. A replicate run had similar cytometry results as

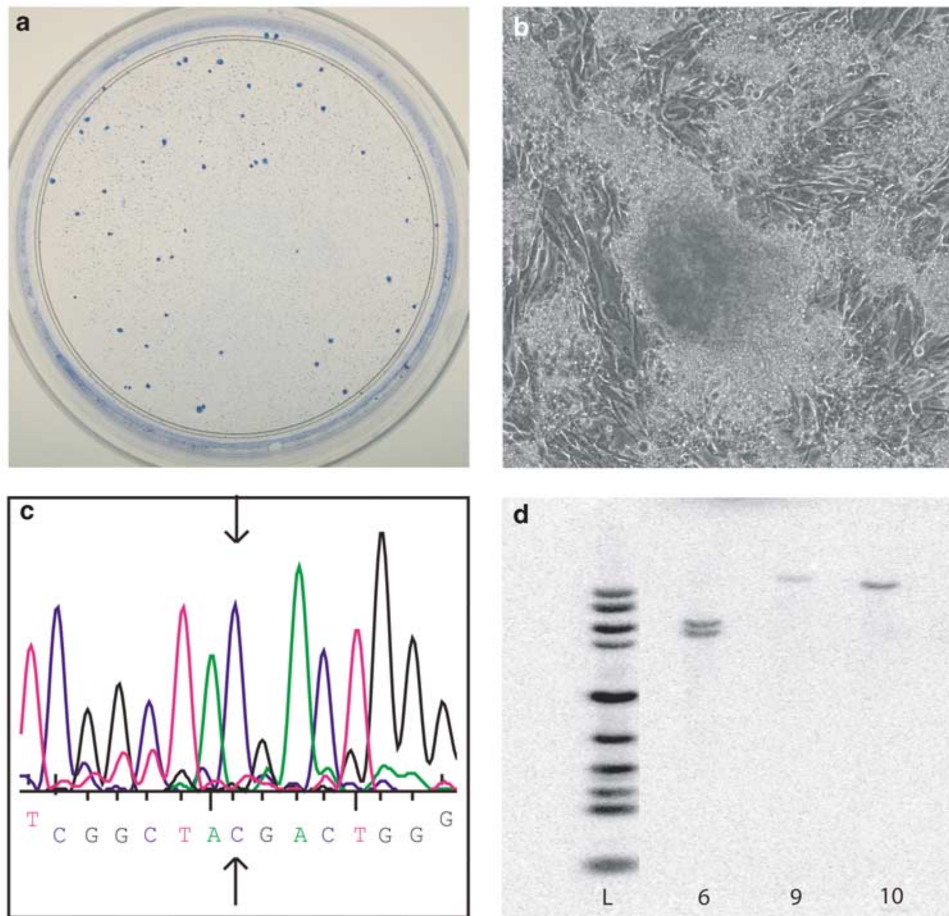


Figure 2 Validation of gene correction. (**a**) A 10-cm dish stained with Trypan blue shows 53 G418 resistant clones. (**b**) A large G418 resistant colony grew on SNLP feeder cells and was surrounded by non-surviving cells. (**c**) Sequence data showing a TAG→TAC correction from Neo1Cor ssODN. (**d**) Southern analysis revealed two integrated copies of pNeoS1Pur in clone 6 and 1 copy in clones 9 and 10.

Table 1 Transfection dose–response testing

	DNA concentration (μg/cm²)			
Lipid concentration (μg/cm²)	*1*	*2*	*4*	*8*
1	3	1	1[a]	0[a]
2	25	17[a]	3[a]	7[a]
3	13[a]	14[a]	6[a]	0[a]

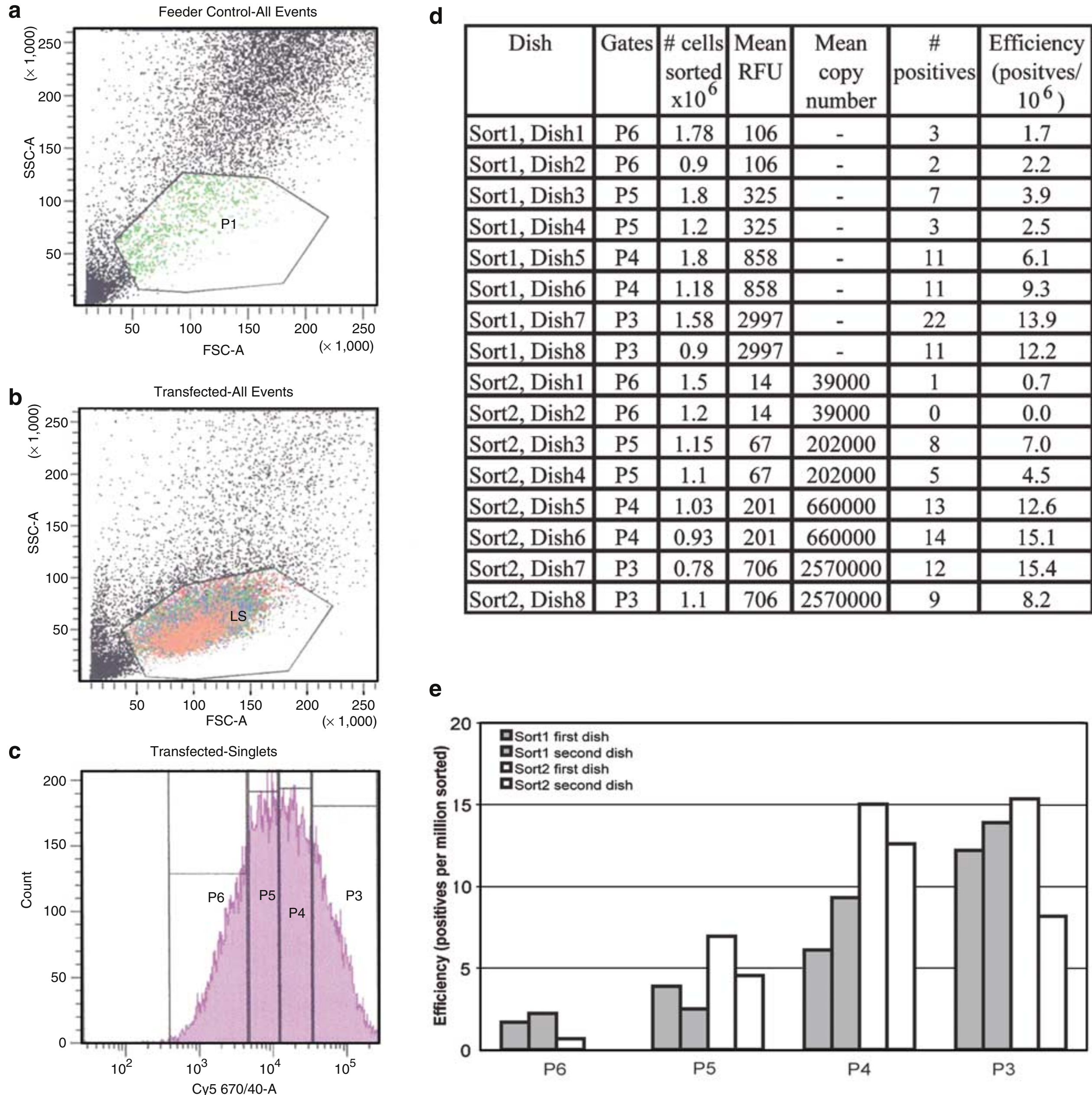
Abbreviation: ESC, embryonic stem cells.
ESC were targeted at different lipid and oligonucleotide concentrations. Dose concentrations are given as reagent mass per plate surface area. Values indicate positives per million cells plated.
[a]Transfection at this condition, followed by selection, resulted in destruction of the inactivated feeder cell layer. The surviving ESC in these wells displayed differentiated morphology.

Figure 3a–c, and was also plated on 8 dishes and selected. Results for all 16 dishes are given in Figure 3d. Note that mean RFU values can only be compared within the same sort, since the cytometer voltages were different for the two experiments. The second sort also included a set of 'molecules of equivalent soluble fluorochrome' (MESF) standards to calibrate the sample to a known number of Cy-5 molecules. The average uptake in the highest bin was 2.5 million ssODN/ cell.

About 40% of the corrections came from the 22% of cells in the bin with the highest uptake and 78% of the corrections were in the top 50% of transfected cells. ESC account for approximately 5% of the oligonucleotide uptake; the majority of ssODN (~95%) resides in feeder cells (data not shown). Overall, the number of positives was correlated with the amount of ssODN uptake (Figure 3e). While the P3 and P4 bins were not statistically different by *t*-test, they were both significantly higher ($P<0.05$) than P5, which was significantly higher ($P<0.05$) than P6. Also, the act of cell sorting ESC had some biological effects. Colonies from the unsorted control appeared on the expected day and had normal phenotypes, implying the sorting process delayed the cell growth, rather than the Cy5-labeled ssODN.

Correction efficiency varies with ESC clonal line and ssODN sequence

Some labs have reported different clonal efficiencies after illegitimate integration[22,23] or different efficiencies for point mutations of different mismatches.[12] We sought to systematically probe these two sources of variability. Three pNeoS1Pur clones (clones 6, 9 and 10) were each corrected with antisense ssODNs Neo1WT (TAG→TAT) and Neo1Cor (TAG→TAC) as well as sense ssODNs Neo1WT-SNS (TAG→TAT) and Neo1Cor-SNS (TAG→



Dish	Gates	# cells sorted x10^6	Mean RFU	Mean copy number	# positives	Efficiency (positves/10^6)
Sort1, Dish1	P6	1.78	106	-	3	1.7
Sort1, Dish2	P6	0.9	106	-	2	2.2
Sort1, Dish3	P5	1.8	325	-	7	3.9
Sort1, Dish4	P5	1.2	325	-	3	2.5
Sort1, Dish5	P4	1.8	858	-	11	6.1
Sort1, Dish6	P4	1.18	858	-	11	9.3
Sort1, Dish7	P3	1.58	2997	-	22	13.9
Sort1, Dish8	P3	0.9	2997	-	11	12.2
Sort2, Dish1	P6	1.5	14	39000	1	0.7
Sort2, Dish2	P6	1.2	14	39000	0	0.0
Sort2, Dish3	P5	1.15	67	202000	8	7.0
Sort2, Dish4	P5	1.1	67	202000	5	4.5
Sort2, Dish5	P4	1.03	201	660000	13	12.6
Sort2, Dish6	P4	0.93	201	660000	14	15.1
Sort2, Dish7	P3	0.78	706	2570000	12	15.4
Sort2, Dish8	P3	1.1	706	2570000	9	8.2

Figure 3 Cell Sorting. Cells were transfected with ssODN containing a Cy-5 label. (**a**) Although ESC grow on SNLP feeder cells, the two were separated by gating a region of the forward-scatter/side-scatter plot. The P1 gate excluded 97% of the cells in the feeder control. (**b**) For the gating shown: 18 million cells were sorted and 11 million were recovered. The second sort had similar gating with 16 million cells sorted and 9 million recovered. (**c**) Sample gating from Sort 1. Untransfected ESC control (not shown) ranged from 10 to 1000 relative fluorescence units, peaking at 200. Recovered cells were binned by Cy5 fluorescence (gates P3-P6) and plated onto eight dishes. After 24 h, G418 selection was initiated. (**d**) Positives were counted on day nine of selection for each dish. For Sort 2, the number of oligonucleotide copies per cell was determined by a MESF standard (standard peaks not shown). RFU values between replicates are not comparable. (**e**) For cells split into four bins, the correction efficiency (colonies per million cells sorted) increases with increased uptake.

TAC). These 12 conditions were transfected simultaneously (four oligos × three clonal lines) in 12 separate dishes. Results for three replicate experiments are shown in Figure 4a. As the total number of positives varied from 26 to 128 between replicates, the data were normalized by the total number of positives per replicate (Figure 4b). Clearly, correction efficiencies varied with both oligonucleotide and clonal line. For clonal variations, clone 9 had higher efficiencies than clones 6 or 10, and the differences were statistically significant ($P<0.05$). For oligonucleotide variations, Neo1WT was clearly the best performing sequence, producing over half of the corrections. Within clone 9, the Neo1WT ssODN was significantly better ($P<0.05$) than the other three oligonucleotides. Neo1WT resulted in sixfold more corrections than Neo1Cor, which was only one base different.

The intrinsic variability of gene correction at very low efficiencies in clone 6 and 10 does not allow the use of the *t*-test to demonstrate statistical differences from these three replicates. As such, the following two observations

should be regarded as suggestive rather than conclusive. First, Neo1WT-SNS, was so close to the spontaneous mutation rate that it might not have been correcting at all. Second, contrary to the usual strand-bias, Neo1Cor-SNS had more corrections than Neo1Cor. To rule out the possibility that arbitrary differences in oligonucleotide synthesis produced a particularly effective batch of Neo1WT or a defective batch of Neo1WT-SNS, the third replicate was transfected with separately synthesized ssODNs. Neo1WT is consistently the best corrector and Neo1WT-SNS is the worst.

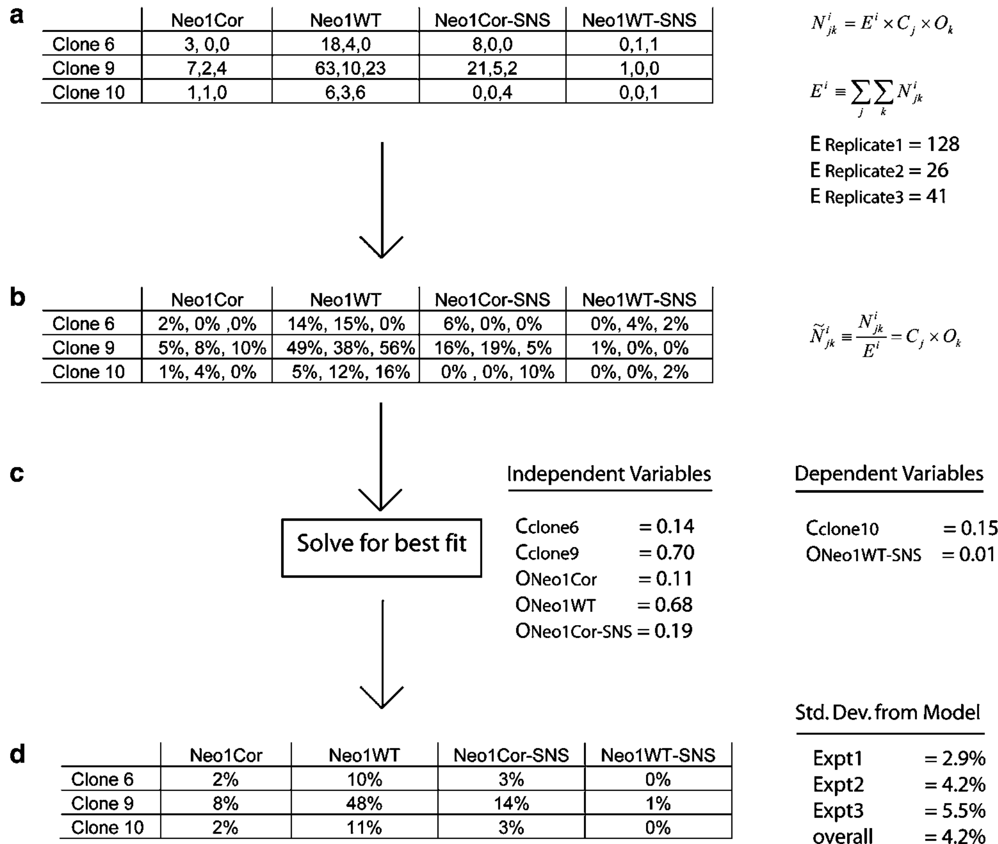
Modeling clonal and ssODN variation as independent variables

The results shown in Figure 4 motivated us to hypothesize variations due to clonal differences and variations due to oligonucleotides were separate, independent variables. The number of positives was modeled as a product of three factors: the overall transfection efficiency, a clonal probability, and an oligonucleotide probability (see Materials and methods). If this model is consistent with the data of Figure 4a, then the normalized data can be accurately described with five parameters. The 36 data points in Figure 4b were used to solve the non-linear least-squares optimization problem and the resulting C_j and O_k parameters are given in Figure 4c. Each C or O value gives the relative 'correcting strength' for a given clone or given ssODN in comparison to the other two clones or three oligonucleotides tested. The ratio of these fractions can be used to determine relative efficiencies. Neo1WT (G:A mismatch) was 6.7-fold more efficient than Neo1Cor (G:G mismatch). Despite having five adjustable parameters, this simple model of $\tilde{N}^i_{jk} = C_j \times O_k$ predicted the results for experiments spanning 12 conditions involving three different clones and four different oligonucleotides.

Using the model parameters that best fit the three experiments (Figure 4d) 10 000 Monte Carlo simulations were run to create a fully resolved event-probability distribution. The expected stochastic noise from 36 transfections was found to be 3.78%, showing a reasonable fit for the experimental deviation of 4.16%. Simulations were used to determine the 95% confidence limits of the four oligonucleotide parameters. With data from all the oligonucleotides and clones, Neo1WT was statistically better than the other three oligonucleotides.

Correction efficiency does not correlate with targeted gene expression

Studies with CHO cells have shown that increased transcription of a target gene can increase the correction efficiency.[17] As clone 9 had a disproportionately high



a

	Neo1Cor	Neo1WT	Neo1Cor-SNS	Neo1WT-SNS
Clone 6	3, 0,0	18,4,0	8,0,0	0,1,1
Clone 9	7,2,4	63,10,23	21,5,2	1,0,0
Clone 10	1,1,0	6,3,6	0,0,4	0,0,1

$$N^i_{jk} = E^i \times C_j \times O_k$$

$$E^i \equiv \sum_j \sum_k N^i_{jk}$$

E Replicate1 = 128
E Replicate2 = 26
E Replicate3 = 41

b

	Neo1Cor	Neo1WT	Neo1Cor-SNS	Neo1WT-SNS
Clone 6	2%, 0% ,0%	14%, 15%, 0%	6%, 0%, 0%	0%, 4%, 2%
Clone 9	5%, 8%, 10%	49%, 38%, 56%	16%, 19%, 5%	1%, 0%, 0%
Clone 10	1%, 4%, 0%	5%, 12%, 16%	0% , 0%, 10%	0%, 0%, 2%

$$\tilde{N}^i_{jk} \equiv \frac{N^i_{jk}}{E^i} = C_j \times O_k$$

c

d

	Neo1Cor	Neo1WT	Neo1Cor-SNS	Neo1WT-SNS
Clone 6	2%	10%	3%	0%
Clone 9	8%	48%	14%	1%
Clone 10	2%	11%	3%	0%

Figure 4 Correction results for three ESC clones and four ssODN. Replicates are separated by commas. (**a**) The number of positives obtained with four different ssODNs in three different pNeoS1Pur clonal lines. In each replicate experiment, the 12 conditions were transfected simultaneously. Replicates were conducted on separate days. (**b**) Data from (**a**) were normalized by the total number of positives in each independent experiment *(E^i)* to eliminate transfection variation. (**c**) Model parameters obtained from non-linear least-squares fit of normalized data. (**d**) Model predictions and s.d. from actual data sets.

correction efficiency, one might hypothesize that it had a higher level of transcription. Using RT-PCR, we compared the level of expression of the neomycin resistance gene to three different endogenous controls in each of the three clonal lines (Figure 5). Expression of Neo mRNA relative to three endogenous controls showed that clone 6 had the least neomycin phosphotransferase mRNA and clone 10 had the most, yet clone 9 displayed the best correction (Figure 4). Thus, clonal targeting efficiencies did not correlate with differences in gene expression.

TAG correction of two-mismatch targeting depends on pairing of second mismatch

The mismatch between the ssODN and the target is an important factor in correction efficiency. Our nomenclature for mismatches (N:N) lists the chromosomal target base first, followed by the base on the ssODN. One could envision making dozens of selectable marker lines to test the 24 combinations for sense and antisense mismatches. Alternatively, a different, but potentially informative methodology is to design two mismatches on the same ssODN, and examine the effects of the second mismatch on correction efficiency. Silent mutations were added to the Neo1WT ssODN nine bases downstream of the selectable target site. These ssODNs, Neo1WT-9T, Neo1WT-9C and Neo1WT-9G were designed to convert the alanine residue at codon 25 from GCA to GCT, GCC or GCG, respectively (Figures 1 and 6a). The activity of the three dual-mismatch ssODNs was compared to a one-mismatch control, Neo1WT. For these experiments, pNeoS1Pur clone 9 cells were transfected with one of the ssODNs, selected in G418, and the surviving colonies were counted (Figure 6b). As with previous experiments, the number of positives varied between replicates. This is consistent with different overall transfection efficiencies as previously observed in Figure 4. However, once normalized to the Neo1WT control, the four ssODNs show a consistent pattern. Figure 6a shows the efficiency of Neo1WT-9T (mismatch A:A), Neo1WT-9G (mismatch A:C) and Neo1WT-9C (mismatch A:G) relative to Neo1WT (basepairing A:T, not mismatched). The addition of a second mutation at base 75 consistently reduced the correction efficiency at base 66 in a basepair-dependent manner. The A:A and A:C mismatches reduced the efficiency by ~75–85%. Using the A:G basepairing at the second site has less severe consequences, incurring only a 60% penalty.

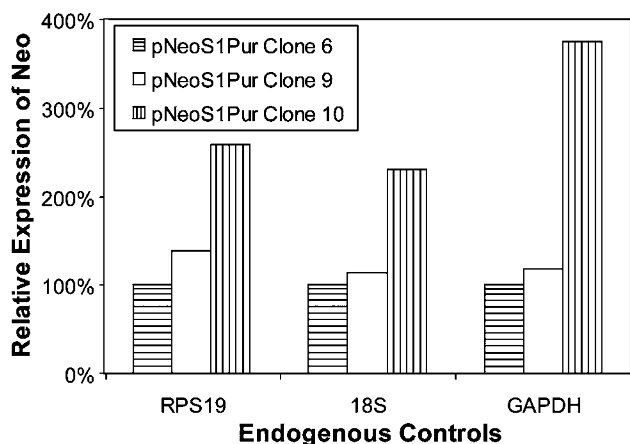


Figure 5 Expression of Neomycin resistance gene in different pNeoS1Pur lines. RNA from clones 6, 9 and 10 was measured by RT-PCR against three endogenous controls: Ribosomal structural protein 19 (RS19), 18S rRNA and glyceraldehyde-3-phosphate dehydrogenase (GAPDH).

Ala25 silent mutations also show biased mismatch effects

Colonies picked from Neo1WT-9G, Neo1WT-9T and Neo1WT-9C transfections were sequenced to determine the rate of sequence alteration at the second mismatch site, codon 25. All 178 colonies evaluated showed a successful TAG→TAT conversion at base 66, consistent with their survival in G418. The second mismatch induced a change at base 75 in 15–35% of the clones (Figure 6c). Neo1WT-9C outperformed the other two ssODNs again by having the highest frequency of second site conversion.

Discussion

Our results demonstrate oligonucleotide-mediated repair of a chromosomally integrated mutant neomycin resistance gene in ESC with efficiencies ranging from 10^{-6} to 10^{-4}. With a targeting efficiency of 10^{-3} it would be feasible to pick 1000 clones and screen them by PCR. At that point, ssODN targeting would become an attractive technology for the creation of knock-in mice and potentially for *ex vivo* gene therapies. As it is difficult to optimize such a complex process as ssODN mediated gene targeting, we have attempted to study individual parameters that impact correction. Experiments were designed to examine oligonucleotide delivery, clonal variation and basepairing effects independently.

Gene delivery affects efficiency

If gene correction were limited exclusively by the rarity of the conversion event, we might expect the process to be relatively insensitive to transfection conditions. Instead, we see a strong dependence on reagent dosing (Table 1 and Figure 3e). Cell sorting tests showed that the cells with over 500 000 ssODN copies were the ones most likely to correct. These data imply that improvements to gene delivery will result in higher efficiencies. While the upper limit on transfection doses appears to be feeder cytotoxicity (Table 1), use of more benign reagents or less sensitive feeders might allow higher doses. Also, cell sorting could be used to enhance gene targeting. As almost 4/5 positives are in the high uptake population, sorting can be used to enrich the number of positives. It is likely that more optimized gating might give even higher enrichments, although sorting may affect pluripotency.

Clonal variations affect efficiency

We also looked at clonal variations in targeting efficiency. Only four of seven clonal lines passed functional testing. Clone 9 cells consistently outperformed clones 6 and 10 by a factor of 5. The primary difference between the clonal lines is potentially the site of pNeoS1Pur plasmid integration. We hypothesize some loci are more accessible than others, or perhaps there are differences in

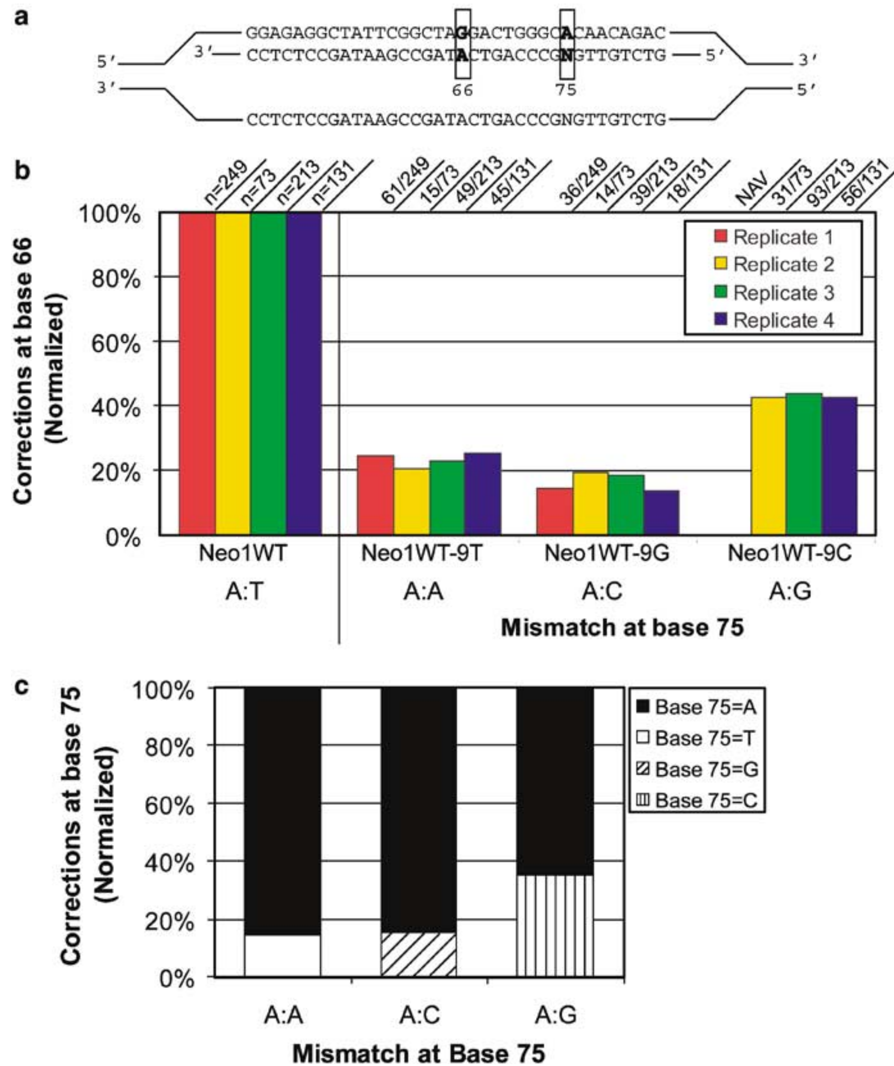


Figure 6 Efficiency of dual-mismatch gene targeting. (**a**) Targeting ssODNs were designed with two mismatches, a G:A mismatch at base 66 to correct the neomycin resistance gene and another mismatch at base 75 to induce a silent mutation. The dual-mismatch ssODNs, Neo1WT-9T (A:A), Neo1WT-9G (A:C) and Neo1WT-9C (A:G) were transfected at the same time as a Neo1WT control. (**b**) The number of corrections at base 66 was determined by the number of surviving colonies. The number of colonies for each replicate of Neo1WT is shown above the chart. The number of positives for each dual-mismatch ssODN transfection were normalized by the Neo1WT control to show the reduction of base 66 targeting efficiency caused by adding a second mismatch. The fractions shown above the bar chart are the number of colonies/number in matched Neo1WT control. (**c**) Survivors with a successful base 66 correction were sequenced to find the percent correction at base 75. Corrected cells have a T, G or C genotype, and uncorrected cells are A.

DNA methylation, histone acetylation or replication at those sites. The differences in correction efficiency were not due to different integrated copy numbers: clone 6 had two integrations, but did not display the highest correction efficiency. Nor do clonal efficiencies correlate with mutant neomycin gene expression as measured by RT-PCR: clone 10 had twice as much neomycin-resistance gene mRNA as the other cell lines and was not the best clone for correction. Locus dependence may have critical impact on future applications that target endogenous mutations.

Mismatch sequences affect efficiency

To analyze the correction event, independent from delivery, we utilized the redundancy of the genetic code to design nearly identical ssODNs that code the same amino acid, but differ by one base at the target site. These ssODNs have the same length, protection, and target homology. We expect they would have nearly identical delivery properties. Despite the similarity of Neo1WT and Neo1Cor, there was a seven-fold difference in correction efficiency. For sense ssODNs, Neo1Cor-SNS outperformed Neo1WT-SNS by 10:1. Contrary to the conventional anti-sense strand bias, Neo1Cor-SNS outperformed Neo1Cor. How does the changing of a single nucleotide so alter the efficiency? It could be caused by differences in secondary structures, changes to CpG motifs, or unknown cellular cues, but the simplest hypothesis is a mechanistic bias towards certain mismatches. Our data show a clear mismatch progression, consistent over numerous replicates and clones: *G:A > C:C* > G:G > C:T* (where * indicates sense targeting).*

This hierarchy is different from studies of TAG→TAC correction in DLD-1 cells where results show more strand bias and *G:G > C:C**.[22] It also differs from results in yeast,[7] perhaps reflecting differences between repair

machinery and correction processes in different organisms. Further tests are required to determine whether basepairing biases are conserved over different sites and sequences. Mismatch biases are important since they underscore mechanisms that are dependent on host machinery.

Dual mismatch studies underscore mismatch biases

Mismatch biases were also observed in the dual-mismatch correction experiments. The addition of a second mismatch lowered the targeting efficiency at the original site in a mismatch-specific manner. The efficiency relative to base 66 decreased as follows: A:T > A:G > A:A > A:C. Similarly, the A:G mismatch was more likely to convert base 75 than the A:A or A:C mismatches. The probability of neomycin correction was about 1×10^{-4} but the probability of a second base alteration was 15–35%. While the first base change was rare, the second was not, indicating they are not independent events. Whatever mechanism accounts for the conversion at one site is thus likely to be involved at the other. It is interesting to note that the selectable correction has a G:A mismatch and the best second site conversion had an A:G mismatch. Perhaps the favorability of the second A:G conversion is an artifact caused by our requirement of a successful G:A correction. Perhaps both G:A and A:G mismatches are intrinsically favored. Either way, there appears to be an underlying mismatch preference with mechanistic significance. The symmetry is interesting because for dual conversions to occur the G:A mismatch must resolve to T:A and the A:G must resolve to C:G. Thus, the mismatch bias was not caused by an enzyme that specifically repairs A:G mismatches to either T:A or C:G. One potential hypothesis is that the G:A mismatch and A:G mismatch are equally favored, and the bias comes not from the recruitment of a specific enzyme, but rather avoidance of undesirable repair enzymes.

Other studies have used multiple mismatches on a single ssODN.[8,24] Kenner *et al*. corrected a hamster *hprt* mutation with three-mismatch ssODNs. They simultaneously targeted a selectable: G:T mismatch, a A:G mismatch seven bases upstream, and a C:A mismatch eight bases downstream. As with our data, the A:G mismatch repaired more frequently than the C:A mismatch.

Better design rules required to address reliability issues

There are ample papers describing negative results with oligonucleotide-mediated gene targeting.[25] To those, we add that three of our seven clonal lines did not correct. Even in our best clone, one of four ssODNs does not give a signal above background. Despite repeated attempts, we have failed to achieve correction in a mutant puromycin-susceptible line (data not shown). Without predictive rules for when targeting fails and when it succeeds, oligonucleotide-mediated gene targeting will remain a highly variable process. However, determining design rules is difficult because studies often have contradictory results. Even the often-observed strand-bias can go in either direction, depending on the application.[26] Part of the diversity of results stems from the diversity of experimental protocols. Researchers have used different lengths, basepairings, and protective chemistries to target various episomal, chromosomal, and endogenous genes in bacteria, yeast and mammalian cells.[27] The development of reliable design rules specific to mouse ESC is required for functional gene targeting technologies to create new mouse models of disease.

Other complex biological technologies such as PCR[28] or transfection[29] have benefited from the use of models to break the problem into independent variables or separate specific design issues. We propose a simple model for the analysis of oligonucleotide gene targeting. The model separates correction efficiency into the product of three factors: E, which depends on experimental conditions and procedures that affect all of the samples in a given experiment, C, which depends on clone or cell-specific differences, and O, which depends on differences between ssODNs. The model is not yet validated and does not include all of the parameters that might affect gene targeting, however it is internally consistent for our data and provides a systematic framework for quantitative analysis.

The pNeoS1Pur ESC lines we have developed are well suited to additional experiments to define oligonucleotide-mediated targeting parameters. For example, we tested ssODN targeting in MSH2-silenced ESC and found comparable efficiency to unsilenced cells. Targeting with non-traditional nucleotides, BrdU and 8-oxo-dG, did not yield any colonies. Transfections of ssODNs with meythlated CpG motifs were comparable to unmethylated ssODNs (data not shown). Further design issues include oligonucleotide length, the position of the mismatch, and the number and type of protective ends. Our oligonucleotides were designed as 53mers with a central mismatch because such oligonucleotides have previously succeeded in targeting stem cells, but further optimization of each variable may lead to incremental gains in efficiency.

Oligonucleotide-mediated gene targeting is appealing because ssODN synthesis is simple. Additionally, ssODN targeting reduces the technical expertise required and opens the prospect of gene targeting to the broader scientific community. While ssODN targeting still requires transfection, cell culture, and sequencing techniques it is not necessary to clone genes, construct large targeting vectors, or run numerous Southern blots. The injection of corrected pNeoS1Pur cells into blastocysts yielded chimeric mice, however their chimerism is too low to guarantee pluripotency. Additional blastocyst injections are in progress. In order to efficiently screen oligonucleotide-mediated gene targeting by PCR, a 10-fold increase in ESC correction efficiency is required. Improvements may come from advances in gene delivery, elucidation of the repair mechanism, or as with PCR, the evolution of empirically determined design rules.

Materials and methods

Materials

Oligonucleotides were synthesized by the Nucleic Acids Facility at the University of Pennsylvania or purchased from Integrated DNA Technologies (Coralville, IA, USA, www.IDTDNA.com) or Midland Certified Reagent Company, Inc (Midland, TX, USA, www.oligos.com).

Oligonucleotides were single-stranded 53-mers with the last four bases on each end protected by phosphorothioate linkages (Figure 1). The length and ends were based on the optimization in Igoucheva *et al*.[4] and successful results of Pierce *et al*.[9] AB2.2 stem cells were cultured on mitomycin-C inactivated SNLP feeder cells according to established protocols.[1,9]

Construction of pNeoS1Pur cell line

A neomycin and puromycin resistant plasmid was constructed by cloning the puromycin cassette from pPur (Clontech, Mountain View, CA, USA, www.clontech.com) into the multiple cloning site of pcDNA3.1+ (Invitrogen, Carlsbad, CA, USA, www.invitrogen.com) at the *Hind*III and *Xba*I sites. Site-directed mutagenesis (Stratagene, La Jolla, CA, USA, www.stratagene.com) was used to mutate the 22nd codon of the neomycin resistance gene from TAT to TAG (base 66, Figure 1) similar to the approach of Rice *et al*.[30] Plasmids were linearized at the *Mfe*I site, purified by phenol chloroform extraction, and electroporated into AB2.2 ESC cells at 30 μg per 8 million cells. The electroporated cells were plated in dishes for selection. After selection with 2 μg/ml puromycin, stably transformed clones were picked and placed into 24-well plates, grown and frozen.

Sequencing

Genomic DNA was PCR amplified with forward primer EAP1504 5′-CCGCCCAGTTCCGCCCATTCTC-3′ and reverse primer EAP1505 5′-GCCAGTCCCTTCCCGCTT CAGTG-3′ to produce a 453 basepair product. The PCR product was purified with a commercial kit (Qiaqen, Valencia, CA, USA, www.qiagen.com) and sequenced by dye-terminator sequencing. Samples from the two-mismatch experiments used a nested PCR. The outside primers are listed above, and a second PCR was performed with forward primer EAP1821 5′-GATGGATT GCACGCAGGTTCTC-3′ and reverse primer EAP1822 5′-GCCTCGTCCTGCAGTTCATTCA-3′.

Lipofection

A total of 1 or 2 million cells, depending on experiment, were plated on feeder cells on a 10-cm dish the day before transfection. The cells were fed non-selective media 3–5 h before transfection. Lipoplexes were formed by mixing 110 μg of ssODN and 55 μl lipofectamine2000 (Invitrogen) in 2.75 ml of Opti-MEM (Invitrogen). The lipoplexes were formed at room temperature for 30 min and diluted with OptiMEM to a total volume of 7.75 ml. The cells were rinsed with OptiMEM, incubated in lipoplex for 2 h, and then supplemented with an additional 10 ml of non-selective media. The following day the transfection solution was replaced by non-selective media. At 24, 36 or 48 h post-transfection the media was replaced with media containing 300 μg/ml G418. Selection in G418 was continued for 7–9 days post-lipofection and surviving clones were either counted or picked.

Cell sorting

Six 10-cm dishes were gelatin-coated and plated with 3.5×10^6 SNLP feeders. The next day, actively growing ESC cells (pNeoS1Pur, clone #9) were passed and plated onto five of the six feeder dishes at 2×10^6 per dish. After 24 h, four of the five ESC dishes were transfected. Transfections followed the usual procedure, except that 20% of the ssODN was labeled with Cy5 (Neo1WT-Cy5). Eleven hours after the start of transfection, the cells were washed twice with plain media then fed media with penicillin/streptomycin. An hour later, three of the four transfected dishes were rinsed with PBS and trypsinized for cell sorting. The untransfected ESC and remaining feeder dish were also trypsinized to use as sorting controls. The cells were spun down, aspirated, and resuspended at 10^7 cells/ml in PBS with 1% FBS. Cell sorting was conducted on a FACS DiVa (Becton Dickinson) at The University of Pennsylvania Flow Cytometry Core Facility (657LP excitation and 670/40 emission). Gates were constructed such that 97% of the SNLP feeder cells were separated by a FSC/SSC amplitude gate. Doublets were removed by FSC/SSC width gates. The remaining cells were binned into four separate groups based on their Cy5 signal. The sorted cells were plated in non-selective media with penicillin/streptomycin on two dishes per fluorescence bin. Selection with 300 μg/ml G418 began 24 h after sorting. The fourth transfected dish, a no-sort control, was also selected in G418. The second replicate of the sort included standardization of the Cy-5 signal to a Cy-5 Molecules of Equivalent Soluble Fluorochrome kit (MESF, p/n 822A) as described in the manufacturer's instructions (Bangs Labs, Fishers IN, USA, www.bangslabs.com).

Reverse transcriptase-polymerase chain reaction

RNA was isolated from frozen cells using the TRIZOL (Invitrogen) reagent followed by an ispropanol precipitation and 70% ethanol wash. To remove residual DNA, the RNA was hydrated in 10 mM Tris and treated with the DNA-Free kit (Ambion, Austin, TX, USA, www.ambion.com). Dilutions of pNeoS1Pur clone 9 RNA were used to make standard curves for the following Applied Biosystems (ABI, Foster City, CA, USA, www.appliedbiosystems.com) probes: Neomycin phosphotransferase (custom assay), ribosomal structural protein 19 (Rps19, assay p/n Mm00452264m1), 18S rRNA (p/n 4331182), and rodent GAPDH (p/n 4308313). Reverse transcription (ABI, p/n N808-0234) and PCR procedures (ABI, p/n 4304437) were performed using random hexamers and per the manufacturer's instructions. Samples were run in triplicate and data were analyzed by the standard curve method as described in the manufacturer's instructions.

Correction efficiency modeling

An empirical model is proposed where the predicted number of positive clones (N) depends on three independent factors: the batch transfection efficiency (E), a coefficient based on the clonal line (C), and a coefficient based on the sequence of the oligonucleotide (O):

$$N^i_{jk} = E^i \times C_j \times O_k \text{ for } i \text{ experiments, } j \text{ clones, and } k \text{ ssODNs}$$

The batch transfection efficiency depends on a number of factors. E^i is particularly sensitive to reagent and procedural conditions. To normalize the experiments, the batch efficiency of the ith experiment (E^i) was defined as the total number of positives per 12-dish transfection batch. For each experiment, a normalized variable, $\tilde{N}^i_{jk}$ was defined as the predicted positives for the jth clone

and kth ssODN divided by the total batch efficiency. It is also the fraction of positives within a given set of transfections.

$$E^i \equiv \sum_j \sum_k N^i_{jk}$$

$$\tilde{N}^i_{jk} \equiv \frac{N^i_{jk}}{E^i} = C_j \times O_k \text{ for } \sum_j C_j = 1 \text{ and } \sum_k O_k = 1$$

As the clonal dependence and oligonucleotide dependence are modeled as independent variables, we have added the additional constraints that C coefficients and O coefficients both sum to unity. For $j=3$ clones and $k=4$ ssODNs this model has five independent variables $[(j-1)+(k-1)]$. Matlab software (The MathWorks, Natick, MA, USA, www.mathworks.com) was used to find the non-linear least squares fit of the five parameters to experimental data from the three independent experiments, each involving a 12-dish transfection. The model empirically fits parameters from transfections of the 12 combinations of three clonal lines and four ssODNs. Each batch of 12 transfections was conducted at the same time with the same media and transfection reagents. The parameters C_j and O_k can be thought as the relative 'corrective strength' of a particular cell line and oligonucleotide, respectively, amongst the three cell lines and four oligonucleotides.

Statistical methods

Difference testing was conducted using an unequal variance (heteroscedastic) two-tailed Student's t-test. Excel software (Microsoft Corporation, Redmont WA, USA, www.microsoft.com) was used for the calculation. The estimation of intrinsic stochastic model error was conducted by dividing the expected number of colonies per dish ($N^i_{jk} = E^i \times C_j \times O_k$) by the number of cells to obtain a per-cell-probability for each E^i C_j, O_k parameter. Random numbers were selected for each cell and random numbers less than or equal to a given transfection's probability were counted as 'positive' cells. The 36 transfections were then simulated 10 000 times to create a virtual data set that contains a random sampling of the stochastic error. The variance and standard deviation between each simulation and the model were calculated and compared to the variance of the experimental data.

Chimeric mouse production

The ESC from a single neomycin-selected clone were injected into blastocysts to create chimeric mice. Injections were performed by the Transgenic and Chimeric Mouse Facility at the University of Pennsylvania School of Medicine (Philadelphia, PA, USA) using standard techniques.[19] The extent of chimerism of the resulting mice was estimated by examining coat color.

Acknowledgements

The assistance of Richard Schretzenmair and Bill Murphy of the University of Pennsylvania Flow Cytometry and Cell Sorting Facility is gratefully acknowledged. We thank Gui-shuang Ying and Chengcheng Liu of the Center of Preventive Ophthalmology and Biostatistics for assistance with statistical analysis. This work was funded by Grants from NIH HL66565 (SLD), NIH PO1-CA72765, the Rosanne Silbermann Foundation, the Mackall Foundation Trust and the FM Kirby Foundation (EAP).

References

1 Ramirez-Solis R, Liu P, Bradley A. Chromosome engineering in mice. *Nature* 1995; **378**: 720–724.
2 Richardson PD, Augustin LB, Kren BT, Steer CJ. Gene repair and transposon-mediated gene therapy. *Stem Cells* 2002; **20**: 105–118.
3 Moerschell RP, Tsunasawa S, Sherman F. Transformation of yeast with synthetic oligonucleotides. *Proc Natl Acad Sci USA* 1988; **85**: 524–528.
4 Igoucheva O, Alexeev V, Yoon K. Targeted gene correction by small single-stranded oligonucleotides in mammalian cells. *Gene Ther* 2001; **8**: 391–399.
5 Havre PA, Gunther EJ, Gasparro FP, Glazer PM. Targeted mutagenesis of DNA using triple helix-forming oligonucleotides linked to psoralen. *Proc Natl Acad Sci USA* 1993; **90**: 7879–7883.
6 Yoon K, Cole-Strauss A, Kmiec EB. Targeted gene correction of episomal DNA in mammalian cells mediated by a chimeric RNA. DNA oligonucleotide. *Proc Natl Acad Sci USA* 1996; **93**: 2071–2076.
7 Brachman EE, Kmiec EB. Targeted nucleotide repair of cyc1 mutations in Saccharomyces cerevisiae directed by modified single-stranded DNA oligonucleotides. *Genetics* 2003; **163**: 527–538.
8 Olsen PA, Randol M, Krauss S. Implications of cell cycle progression on functional sequence correction by short single-stranded DNA oligonucleotides. *Gene Therapy* 2005; **12**: 546–551.
9 Pierce EA, Liu Q, Igoucheva O, Omarrudin R, Ma H, Diamond SL *et al*. Oligonucleotide-directed single-base DNA alterations in mouse embryonic stem cells. *Gene Therapy* 2003; **10**: 24–33.
10 Dekker M, Brouwers C, teRiele H. Targeted gene modification in mismatch-repair-deficient embryonic stem cells by single-stranded DNA oligonucleotides. *Nucleic Acids Res* 2003; **31**: e27.
11 Nickerson HD, Colledge WH. A comparison of gene repair strategies in cell culture using a lacZ reporter system. *Gene Therapy* 2003; **10**: 1584–1591.
12 Brachman EE, Kmiec EB. DNA replication and transcription direct a DNA strand bias in the process of targeted gene repair in mammalian cells. *J Cell Sci* 2004; **117**: 3867–3874.
13 Parekh-Olmedo H, Ferrara L, Brachman E, Kmiec EB. Gene therapy progress and prospects: targeted gene repair. *Gene Ther* 2005; **12**: 639–646.
14 Ma H, Liu Q, Diamond SL, Pierce EA. Mouse embryonic stem cells efficiently lipofected with nuclear localization peptide result in a high yield of chimeric mice and retain germline transmission potency. *Methods* 2004; **33**: 113–120.
15 Gamper HB, Parekh H, Rice MC, Bruner M, Youkey H, Kmiec EB. The DNA strand of chimeric RNA/DNA oligonucleotides can direct gene repair/conversion activity in mammalian and plant cell-free extracts. *Nucleic Acids Res* 2000; **28**: 4332–4339.
16 Li XT, Costantino N, Lu LY, Liu DP, Watt RM, Cheah KS *et al*. Identification of factors influencing strand bias in oligonucleotide-mediated recombination in Escherichia coli. *Nucleic Acids Res* 2003; **31**: 6674–6687.
17 Igoucheva O, Alexeev V, Pryce M, Yoon K. Transcription affects formation and processing of intermediates in oligonucleotide-mediated gene alteration. *Nucleic Acids Res* 2003; **31**: 2659–2670.
18 Wu XS, Xin L, Yin WX, Shang XY, Lu L, Watt RM *et al*. Increased efficiency of oligonucleotide-mediated gene repair through slowing replication fork progression. *Proc Natl Acad Sci USA* 2005; **102**: 2508–2513.
19 Nagy A. Manipulating the Mouse Embryo. In: Gertsenstein MVKBR (ed). *A Laboratory Manual*. Cold Stone Harbor Laboratory Press, 2003, pp 128–252.

20 Almofti MR, Harashima H, Shinohara Y, Almofti A, Li W, Kiwada H. Lipoplex size determines lipofection efficiency with or without serum. *Mol Membr Biol* 2003; **20**: 35–43.
21 Tseng WC, Haselton FR, Giorgio TD. Transfection by cationic liposomes using simultaneous single cell measurements of plasmid delivery and transgene expression. *J Biol Chem* 1997; **272**: 25641–25647.
22 Hu Y, Parekh-Olmedo H, Drury M, Skogen M, Kmiec EB. Reaction parameters of targeted gene repair in Mammalian cells. *Mol Biotechnol* 2005; **29**: 197–210.
23 Olsen PA, Randol M, Luna L, Brown T, Krauss S. Genomic sequence correction by single-stranded DNA oligonucleotides: role of DNA synthesis and chemical modifications of the oligonucleotide ends. *J Gene Med* 2005; **7**: 1534–1544.
24 Kenner O, Lutomska A, Speit G, Vogel W, Kaufmann D. Concurrent targeted exchange of three bases in mammalian hprt by oligonucleotides. *Biochem Biophys Res Commun* 2004; **321**: 1017–1023.
25 van der Steege G, Schuilenga-Hut PH, Buys CH, Scheffer H, Pas HH, Jonkman MF. Persistent failures in gene repair. *Nat Biotechnol* 2001; **19**: 305–306.
26 Yamamoto T, Moerschell RP, Wakem LP, Komar-Panicucci S, Sherman F. Strand-specificity in the transformation of yeast with synthetic oligonucleotides. *Genetics* 1992; **131**: 811–819.
27 Igoucheva O, Alexeev V, Yoon K. Oligonucleotide-directed mutagenesis and targeted gene correction: a mechanistic point of view. *Curr Mol Med* 2004; **4**: 445–463.
28 Rubin E, Levy AA. A mathematical model and a computerized simulation of PCR using complex templates. *Nucleic Acids Res* 1996; **24**: 3538–3545.
29 Banks GA, Roselli RJ, Chen R, Giorgio TD. A model for the analysis of nonviral gene therapy. *Gene Ther* 2003; **10**: 1766–1775.
30 Rice MC, Bruner M, Czymmek K, Kmiec EB. *In vitro* and *in vivo* nucleotide exchange directed by chimeric RNA/DNA oligonucleotides in Saccharomyces cerevisae. *Mol Microbiol* 2001; **40**: 857–868.