GENE TARGETING AT AND DISTANT FROM DNA BREAKS IN YEAST AND HUMAN CELLS

A Dissertation
Presented To
The Academic Faculty

By

Samantha Anne Stuckey

In Partial Fulfillment
Of the Requirements for the Degree
Doctor of Philosophy in Biology

Georgia Institute of Technology

May 2013

GENE TARGETING AT AND DISTANT FROM DNA BREAKS IN YEAST AND HUMAN CELLS

Approved by:

Dr. Francesca Storici, Advisor
School of Biology
Georgia Institute of Technology

Dr. Yury Chernoff
School of Biology
Georgia Institute of Technology

Dr. Kirill Lobachev
School of Biology
Georgia Institute of Technology

Dr. Philip Santangelo
Wallace H. Coulter Department of Biomedical Engineering
Georgia Institute of Technology

Dr. Ingeborg Schmidt-Krey
School of Biology
Georgia Institute of Technology

Date approved: March 29, 2013

For my family and friends

ACKNOWLEDGMENTS

There are many people without whom I would not have made it this far.

There aren't enough words to thank the woman with whom I have been fortunate enough to conduct research with in my time at Georgia Tech, Francesca Storici. From the very first day I met her, my first day of graduate school, I was immediately drawn to her passion for her research. Her guidance has benefited me enormously, in ways I am sure I will not be able to properly convey, and to her I will forever be grateful. It is my hope that she understands how much I truly appreciate her, as both a person and mentor.

I want to show my appreciation to my committee – Dr. Yury Chernoff, Dr. Kirill Lobachev, Dr. Ingeborg (Inga) Schmidt-Krey, and Dr. Philip (Phil) Santangelo. Whether in a committee meeting, at an institute seminar, or simply in passing each other in the halls, I have always been thankful that I could turn to any of them and receive insightful, helpful, and constructive feedback.

One group of people for whom I will always be thankful is the Storici lab at Georgia Tech, including its past and present members. The teamwork and loyalty of our lab family has always meant a great deal to me, and I would be hard-pressed to find a more embracing and considerate group of people. From the beginning with (now Dr.) Ying Shen and (future Dr.) Patrick Ruff, being able to grow and figure things out as we went was an amazing experience. For the rest of my life, I'll be happy I got to share that

experience with them. I cannot remember how our lab or I survived without Kyung Duk (K.D.) Koh and Havva Keskin. And with Dr. Kuntal Mukherjee and Dr. Rekha Pai, their guidance, care, and support – moral, mental, research-related, and everything else – are things I will cherish. There are many others – including Pavan Nandi, Sathya (Sat) Balachander, Taehwan Yang, Zhiqiang Lin, and the amazing undergraduates we've seen over the years – to whom I want to express my sincerest thanks.

For the last five years, I have been fortunate to interact with several wonderful others at Georgia Tech who have helped me in various ways. These include Laura Cook Cullen, Barbara Walker, Nena Gray, Lisa Tuttle, Verene Lancaster, and Dr. David Garton. I want to thank Dr. Rupal Thazath Cutting, Dr. Jennifer Leavey, and Dr. Chrissy Spencer for allowing me to learn from them and become a better instructor, speaker, and leader. And Kevin Roman is a saint, whether he knows it or not. Additionally, I am grateful to Paramita Chatterjee, Khairat Elbaradie, Vinay Mittal, Chenyi Pan, Jianrong Wang, Andy Conley, Po-Yi Ho, Yu Zhang, Natalie Saini, Sunyoung Goo, Lively Lie, Daudi Jjingo, and Eishita Tyagi, among others. And Dr. Vidhya Narayanan, Dr. Hyun Min Kim, Dr. He Gong, Dr. Taylor Updegrove, and Dr. Gaurav Arora are fantastic individuals; I appreciate their help along the way.

In a category all of their own are Marc Pline and Angie Lessard. Our lab could not have survived without them.

I am fortunate to have a very large family, most of whom are located in Georgia. Being

able to conduct research throughout the week but still be able to see my grandmother, aunts, uncles, or cousins whenever I have needed has meant a great deal to me.

I do not think that there are enough words to describe how thankful I am to Dr. Lee Katz. Lee was the very first graduate student I met in the Biology program. Little did I know how important he would become to me. Lee is my best friend, my love, my partner, and my soul mate. I would not have made it through my graduate school career without him pushing me through it – gently but firmly. I will forever be grateful for his love and guidance. His parents, Sheri and Ian Katz, and his siblings Rebecca and Jeffrey, are fantastic and I am thankful for their support, care, and kindness.

Finally, to my parents, Wanda and James Stuckey, and my sisters, Sarah, Rachel, and Jamesan, I cannot express how much I have needed them all through everything and how much I appreciate the fact that they have all been there for me. The support of my parents with everything I have undertaken in life has been one of the highlights of my life. I cannot express my gratitude enough that I have them to call my parents. My beautiful sisters have also been there for me through everything. The words don't exist to show them how much I love and appreciate their words of encouragement and support. And to every medical person we have encountered these last few years who have helped each individual in my family, words cannot express how grateful I am that my loved ones have had such amazing care. To them I express my sincerest thanks.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

5-FOA	5-fluoroorotic acid
AAV	adeno-associated virus
ADA	adenosine deaminase
att	attachment site
BIR	break-induced replication
bp	base pair
°C	degrees Celsius
Cas	CRISPR associated
CBA	chicken β actin
cDNA	complementary DNA
cm	centimeter
CMV	cytomegalovirus
CO_2	carbon dioxide
CORE	COunterselectable REporter
CRISPR	clustered regularly interspaced short palindromic repeat
DMEM	Dulbecco's modified Eagle's medium
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
ds	double-stranded
DSB	double-strand break

DSBR	double-strand break repair
EDTA	ethylenediaminetetraacetic acid
EtOH	ethanol
FACS	fluorescence-activated cell sorting
FBS	fetal bovine serum
g	gram
g	gravity
gRNA	guide RNA
GCR	gross chromosomal repeat
GFP	green fluorescent protein
h	hour
H_3BO_3	boric acid
HEK-293	human embryonic kidney
HR	homologous recombination
in/dels	insertions/deletions
kb	kilobase
l	liter
LiOAc	lithium acetate
M	molar
mg	milligram
Mg^{2+}	magnesium ions
min	minute

ml	milliliter
mM	millimolar
MMR	mismatch repair
MRX	*MRE11*/*RAD50*/*XRS2*
mtDNA	mitochondrial DNA
NaOAc	sodium acetate
$(NH_4)_2SO_4$	ammonium sulfate
NHEJ	non-homologous end-joining
nmole	nanomole
nt	nucleotide
oligo	oligonucleotide
O/N	overnight
PARP	poly(ADP-ribose) polymerase
PCR	polymerase chain reaction
PEG	polyethylene glycol
PEI	polyethyleneimine
pmol	picomole
RFP	red fluorescent protein
RNA	ribonucleic acid
ROS	reactive oxygen species
rpm	revolutions per minute
s	second

SCID	severe combined immunodeficiency
SCID-X	x-linked SCID
SC-Leu	synthetic complete medium lacking leucine
SC-Trp	synthetic complete medium lacking tryptophan
SC-Ura	synthetic complete medium lacking uracil
SD-complete	synthetic dextrose complete
SDSA	synthesis-dependent strand annealing
ss	single-stranded
SSA	single-strand annealing
SSB	single-strand break
SSD	salmon sperm DNA
ssDNA	single-stranded DNA
TALEN	transcription activator-like effector nuclease
TBE	tris/borate/EDTA
TE	tris/EDTA
U	unit
UV	ultraviolet
V(D)J	variable, diverse, joining
wt	wild-type
YPD	yeast extract/peptone/dextrose
YPG	yeast extract/peptone/glycerol
YPGal	yeast extract/peptone/galactose

YPLac	yeast extract/peptone/lactic acid
ZFN	zinc finger nuclease
α	alpha
β	beta
Φ	phi
µl	microliter
µM	micromolar

SUMMARY

Here we developed multiple genetic systems through which genetic modifications driven by DNA breaks caused by the I-*Sce*I nuclease can be assayed in the yeast *Saccharomyces cerevisiae* and in human cells. Using the *delitto perfetto* approach for site-directed mutagenesis in yeast, we generated isogenic strains in which we could directly compare the recombination potential of different I-*Sce*I variants. By genetic engineering procedures, we generated constructs in human cells for testing the recombination activity of the same I-*Sce*I variants. Both in yeast and human cells we performed genetic correction experiments using oligonucleotides (oligos) following modification and/or optimization of existing gene targeting protocols and development of new ones. We demonstrated that an I-*Sce*I nicking enzyme can stimulate recombination on the chromosome in *S. cerevisiae* at multiple genomic loci. We also demonstrated in yeast that an I-*Sce*I-driven nick can activate recombination 10 kb distant from the initial site of the chromosomal lesion. Moreover we demonstrated that an I-*Sce*I nick can stimulate recombination at the site of the nick at episomal and chromosomal loci in human cells. We showed that an I-*Sce*I double-strand break (DSB) could trigger recombination up to 2 kb distant from the site of the break at an episomal target locus in human cells, though the same was not observed for the nick. Overall, we demonstrated the capacity for I-*Sce*I nick-induced recombination in yeast and human cells. Importantly, our findings reveal that the nick stimulates gene correction by oligos differently from a DSB lesion, as determined by genetic and molecular analyses in yeast and human cells. This research illustrates the promise of targeted gene correction following generation of a nick.

CHAPTER 1

INTRODUCTION

1.1 GENETIC MODIFICATION AND GENETIC DISORDERS

Many diseases are considered monogenic and are caused, at least primarily, by mutations in a single gene. These include sickle cell anemia, muscular dystrophy, hemophilia types A and B, forms of severe combined immunodeficiency (SCID), Rett syndrome, retinitis pigmentosa, cystic fibrosis, and phenylketonuria [1-12]. Because of the harm an afflicted individual endures if the disease is left untreated, scientists have focused on different ways in which the normal protein lost due to the genetic mutation could be provided. Thus, two strategies for gene therapy have emerged: treatment of the disease through integration of the functional gene somewhere in the genome and treatment through gene targeting to fix the mutant allele at the endogenous locus.

1.1.1 Gene therapy through integration of the functional gene in the genome

For many devastating disorders, a dysfunction occurs due to a mutation in a single gene causing a misfolded, truncated, or otherwise mutant protein [13]. While some diseases, like phenylketonuria, can be at least partially remedied by a simple dietary change, some others are much more complex [11]. In order to supplement the loss of the endogenously-encoded protein, gene therapy can be used to deliver a functional copy of the gene. Different strategies exist to deliver the nucleotide sequence into cells using viral or non-

viral vectors [14,15]. Using this type of gene therapy, successful treatments of adenosine-deaminase (ADA) SCID and x-linked SCID (SCID-X) as well as hemophilia A and hemophilia B have been reported in various studies [2,3,16-19]. A major drawback to this type of approach, however, is the risk of insertional mutagenesis. In a highly-publicized case study, three patients successfully treated for SCID-X later developed leukemia after the lentiviral vector used for therapy integrated within an oncogene [4]. As a consequence, scientists have focused their efforts on integration into so-called "safe harbor" loci – those in which targeted insertion of a gene has been shown to minimally affect the endogenous locus while allowing for adequate expression of the new genetic material [20-22]. However, the natural genetic locus retains the original, mutated sequence.

1.1.2 Gene therapy through targeted gene correction at the endogenous locus

Another type of gene therapy is gene targeting, which is a technique that can be used to directly integrate the normal or modified deoxyribonucleic acid (DNA) sequence of a gene of interest into its natural endogenous chromosomal position. It has been widely used in many organisms and cell types including the yeast *Saccharomyces cerevisiae*, the slime mold *Dictyostelium discoideum*, the moss *Physcomitrella patens*, mouse embryonic stem cells, and human cells [23-35]. The principle of this technique is that an endogenous chromosomal locus can be directly targeted for replacement with a similar sequence from an exogenous template rather than the random integration approach often used in traditional gene therapy [31]. This exchange relies on the cellular homologous recombination (HR) machinery to swap out genetic material [31]. Through gene targeting

a variety of genetic manipulations can be made including deletions, insertions, and other point mutations [31,36]. Applications include functional analysis of a protein through deletion of its genetic coding sequence, providing an added selective advantage through insertion of a selectable marker such as one encoding a fluorescent protein or one providing antibiotic resistance, and replacement of mutant alleles with corrected sequences in order to restore proper protein function [36]. In particular, the latter of these applications provides great implications for genetic disorders as it presents a treatment option to "cure" a disease at the genomic level by removing the mutant form of the culprit gene and replacing it with the functional copy.

1.2 THE ROLE OF A DNA DOUBLE-STRAND BREAK (DSB) IN HR

HR is a natural process involved in meiotic recombination and in repair of DNA mitotic DSBs [37]. It has been studied in many organisms, with much research conducted in the yeast model [38-40]. In meiotic recombination it allows for exchange between homologous chromosomes in the first meiotic division which can generate genetic variation [41]. In mitotic recombination it naturally repairs DSBs generated as a result of collapsed replication forks which can occur due to DNA damage [37]. Following generation of a DSB, the HR machinery is recruited to repair the broken DNA sequence.

Several models exist to explain repair of a DSB by HR including synthesis-dependent strand annealing (SDSA), double-strand break repair (DSBR), single-strand annealing (SSA), and break-induced replication (BIR) [39]. The SDSA and DSBR models follow the same initial steps for repair of the DSB using a homologous double-stranded template

but differ in the final product. In SDSA the result is a non-crossover event yielding a duplicated genetic sequence while in DSBR both non-crossovers as well as crossovers, in which the genetic sequence between the donor and corrected templates are swapped, can occur [42]. In SSA a DSB generated between direct repeats can be repaired by annealing between the repeat regions following resection, which results in loss of the intervening sequence [43]. BIR can be used to repair a DSB by copying the sequence from a homologous chromosome in a process which can extend all the way to the telomere, resulting in duplication of the copied chromosome end [44].

Several key proteins are involved in repair *via* these HR pathways. The Mre11/Rad50/Xrs2 (Nbs1 in humans) (MRX) complex is involved in processing the break and generating 3' overhangs following resection in a 5' to 3' manner [45]. The Rpa complex (Rfa1/Rfa2/Rfa3 in yeast; Rpa1/Rpa2/Rpa3 in humans) binds to the 3' single-stranded DNA (ssDNA) ends and prevents formation of secondary structures [41]. The Rad52 protein anneals complementary strands of ssDNA and mediates the loading of Rad51 to the resected 3' ends of the DNA, while the Rad51 protein is a recombinase involved in the exchange between ssDNA and a double-stranded molecule [37]. In yeast, Rad55 and Rad57 aid in the strand exchange initiated by Rad51, Rad54 is a chromatin remodeling factor which helps provide access to genomic regions, and Rad59 (a paralog of Rad52) is involved in gene conversion [42]. Other factors involved in one or more of these HR pathways include Sae2, Exo1, Dna2, and Sgs2, which have roles in 5' end resection; and Rad1, Rad10, Msh2, and Msh3, which have roles in trimming non-homologous ends [41,42]. In human cells, Rif1 colocalizes with other DNA damage

response factors, including ATM, in cells defective in telomerase and is highly-dependent on ATM and the human 53bp1, which is a protein involved in promoting end joining of damaged chromosome ends and which together prevent resection of the ends [46,47].

1.3 GENE TARGETING AND A DSB

As gene targeting follows the steps of HR, it requires recruitment of the HR machinery to exchange the genetic sequences. However, HR is naturally inefficient in many cell types including human somatic cells, in which spontaneous recombination occurs at levels around 10^{-6} while random integration is ~1000-fold more efficient [48,49]. Because of these low levels, researchers have sought to activate HR in order for gene targeting to be considered a practical approach. The generation of a DSB has widely been used for this purpose as a DSB is the natural target of the HR machinery.

1.3.1 A DSB can efficiently trigger HR at the site of the break

Meganucleases are proteins with large (12-40 base pairs, bp) recognition sequences and which generate a DSB at a specific DNA site [50]. Numerous studies have shown that generating a site-directed DSB with a meganuclease such as I-*Sce*I can stimulate HR up to 1,000-fold or more in yeast, human cells, zebrafish, *Xenopus*, and others [51-57]. Engineered "designer" nucleases, such as zinc finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs), have also been shown to trigger the HR machinery efficiently [58-64]. In these systems, DNA binding domains can be directed to target a specific DNA sequence by fusing ZFs or TALEs with the catalytic domain of the *Fok*I endonuclease [65]. Additionally, systems using the *Streptomyces* phage ΦC31

integrase, which stimulates recombination at recognition sequences known as attachment (att) sites, have been used for targeted gene insertion within the genome of bacterial, yeast, and human cells [66-68]. Most recently, clustered regularly interspaced short palindromic repeat (CRISPR)/CRISPR associated (Cas) systems have been shown to function as genome-engineering tools based on DSBs mediated by interactions with guide ribonucleic acid (gRNA) [69,70].

1.3.2 Recombination at a distant locus following DSB generation

Not only does generation of a DSB stimulate recombination at the site of the break, it can also stimulate recombination at sites distant from the break as has been demonstrated in yeast at distances up to 20 kb from the site of an I-*Sce*I DSB [52]. Depending on the position of the DSB relative to the targeted gene, correction after break generation was observed up to 60-fold over levels without a break [52]. Additionally, a strand bias in the correcting template for the targeted gene was observed for cases in which ss oligonucleotides (oligos) were utilized as donor sequences [52]. Resection occurs following processing of a DSB for HR repair, and this should leave the strand 3' of the break intact [45]. Because the bias observed was in favor of the template complementary to the 3' end of the break, it was proposed that generation of the DSB followed by extensive resection stimulated a large window for recombination. While observed in yeast cells with the I-*Sce*I nuclease, this approach of gene targeting distant from a DSB could also be applied to studies using other DSB-inducing systems in other cell types. For example, the optimal position for cleavage by designed nucleases such as ZFNs may be located several kb from the gene to be modified. By generating a break distant from that

locus, however, gene targeting may still be accomplished if the break is able to activate large regions for recombination.

1.3.3 A DSB can be highly threatening

While a DSB can efficiently stimulate recombination, it can do so at high cost to genomic integrity. A single unrepaired DSB can lead to cell death [71]. In yeast HR is the major DSB-repair pathway, but in human cells the error-prone non-homologous end-joining (NHEJ) pathway predominates [72]. In this process, broken DNA ends can be ligated together following minimal processing. However, if multiple DSBs are generated the ends could be rejoined in such a way that genomic segments are lost or gross chromosomal rearrangements (GCRs) are formed [73]. This could occur even when using a nuclease with a specific recognition sequence due to the propensity for "off-target" recognition of a similar genomic locus, as has been demonstrated for meganucleases and designed nucleases alike [74-76]. In multiple studies NHEJ events have been observed to occur at higher rates than the desired targeted HR events, illustrating the need for safer mechanisms of stimulating recombination for gene targeting approaches [77,78].

1.4 SINGLE-STRAND BREAKS (SSB) GENERATED BY NICKASES

An SSB, or nick, is a lesion that occurs when one strand of a double-stranded DNA molecule is cleaved. SSBs can arise from DNA damage after exposure to reactive oxygen species (ROS), during replication and DNA repair, and following cleavage by a nickase [79,80]. Various studies have demonstrated that gene targeting can be efficiently accomplished using different nickases, including natural and engineered proteins

[54,69,77,78,81-83]. However, speculation exists as to whether the SSB itself is enough to stimulate recombination or if it must first be converted to a DSB by the enzyme after prolonged incubation [84].

1.4.1 HR following generation of an SSB

In early models for recombination an SSB was thought to be the triggering lesion for the recombination event [85,86]. Though the focus has since shifted to DSBs, the finding of several groups that generation of an SSB was followed by targeted gene correction indicated that an SSB could also stimulate HR, possibly as a DSB intermediate. SSBs have been implicated as the initiating factors during V(D)J recombination in human cells [87]. Likewise, SSBs created through adeno-associated virus (AAV) Rep proteins Rep78 and Rep68 have been used to insert genetic sequences at the preferred integration site of AAV (*AAVS1*) on human chromosome 19 [88]. Studies in yeast have demonstrated that intra- and inter-chromosomal recombination increased following generation of an SSB by the Gene *II* nickase of the bacteriophage f1 [54,89]. Moreover, an engineered nicking variant of the I-*Ani*I protein was shown to stimulate recombination at episomal and chromosomal loci in human embryonic kidney (HEK-293) T cells at up to 25% the efficiency of the I-*Ani*I nuclease [81]. Most recently zinc finger nickases and CRISPR/Cas systems using nicking enzymes have proven to efficiently modify chromosomal regions in human cells [69,82,90].

1.4.2 The genetics of SSB-induced recombination

Because of the ability of an SSB to be easily re-ligated, an SSB is thought to stimulate

HR in a safer manner than a DSB [80,91]. Studies using a nicking variant of the I-*Ani*I nuclease have demonstrated not only that recombination at a genomic locus in human cells increases following generation of an SSB but that there are less accompanying NHEJ events compared to when the corresponding DSB is generated by I-*Ani*I [77,78]. While it is unclear whether the SSB itself initiated recombination or if it was first converted to a DSB, the initial presence of the SSB seemed to be less mutagenic to the cells as fewer erroneous events were detected [77,78]. In particular, the ratio of HR to NHEJ events was 50-fold greater following SSB generation than the ratio observed following DSB generation, indicating that an SSB could initiate targeted gene correction more safely than a DSB even despite lower overall levels of recombination [77]. These differences could be due to a diverse repair mechanism for a DSB arising from replication fork collapse compared to a DSB generated prior to replication [77]. These findings demonstrate that SSB generation could be a safer alternative for targeted gene correction strategies than those relying on formation of a DSB.

1.5 USE OF DNA OLIGOS FOR DNA REPAIR AND MODIFICATION

DNA oligos are short segments of deoxyribonucleotides which can be produced *in vivo* from the breakdown of larger molecules. They can also be synthesized commercially and are used for a variety of purposes including as primers during polymerase chain reaction (PCR) and DNA sequencing and as molecular probes [92-94]. Additionally, short (~20-80 nucleotide, nt) synthetic DNA oligos have been used to modify genomic DNA in various organisms by introducing a mutation at a target locus [52,53,95-100]. This is

accomplished using oligos homologous to either side of the target locus such that the HR machinery can substitute the oligo sequence information into the genomic locus [101].

While this approach can be applied to gene correction strategies, it is generally limited to loci presenting selectable phenotypes upon recombination. In yeast, the *delitto perfetto* system for site-directed mutagenesis provides a robust method for generating these types of mutations in two steps [102-104]. In the first step, a cassette containing two markers is inserted into the target locus, and upon integration recombinants can be selected for based on a phenotype provided by one of the markers. In the second step, the cassette is removed and can be replaced with a homologous template. Again a phenotype indicating recombination is used to select for gene correction. In mammalian cells, phenotypic selection is also commonly used to observe recombination at a natural or artificially-created genomic locus [60,105].

1.6 NICKING VARIANTS OF NATURAL MEGANUCLEASES

The monomeric *S. cerevisiae* I-*Sce*I meganuclease is a member of a family of nucleases characterized by the LAGLIDADG amino acid sequence motif [106]. As a homing endonuclease, I-*Sce*I catalyzes insertion of its genetic information into an allele lacking the gene but possessing the cognate recognition sequence of the enzyme [107]. It accomplishes this by generating a DSB at its 18-bp recognition sequence, which is subsequently repaired by the HR machinery, during which time the gene is inserted [106]. Because of its specificity for the target sequence, I-*Sce*I is routinely utilized for targeted gene correction approaches in yeast and has been adapted to numerous other cell

types, including mammalian cells [53,55-57,60]. However, because of the issues of off-site targeting and illegitimate repair of a DSB, other avenues have been explored. Thus, scientists have focused their attention on generating variants of I-*Sce*I and other homing endonucleases, including I-*Ani*I and I-*Cre*I, which could be used to develop highly-specific nickases for gene targeting applications [81,108-110].

1.6.1 The I-*Ani*I nickase

The *Aspergillus nidulans* meganuclease I-*Ani*I recognizes an asymmetric 19-bp target sequence [111]. A nicking variant was generated through inactivation of one of the active sites through a lysine to methionine substitution at residue 227 [81]. The I-*Ani*I K227M nickase has been used for gene targeting approaches in human cells to demonstrate the capacity for I-*Ani*I SSB-induced repair at episomal and chromosomal loci [77,78,81]. A catalytically dead mutant was also generated through an additional mutation at residue 171 to generate the non-breaking K227M/Q171K variant [81]. Another nickase, I-*Ani*I Q171K, demonstrated nicking activity but it was accompanied by subsequent DSB cleavage with extended incubation and thus was not used for additional studies [81].

1.6.2 The I-*Sce*I nickase

Nicking variants of the I-*Sce*I meganuclease were generated previously by the Gimble laboratory at Purdue [108]. Like the I-*Ani*I nickase, these variants were generated by substituting a nonpolar residue for an acidic conserved lysine adjacent to either active site (residues 122 and 223, respectively), thus destablizing coordination of ions in one of the catalytic regions of the monomeric protein [108]. However, with extended digestion each

protein demonstrated eventual cleavage on both strands, though at different rates – after 20 minutes for K122I and after 4 hours for K223I – indicating that complete inactivation of double-stranded (ds) cleavage activity was not possible through mutation of either of the residues alone [108]. Additionally, a catalytically dead mutant (I-*Sce*I D145A) was generated by substituting the aspartic acid at residue 145 with an alanine which was thought to abolish cleavage capacity due to a loss of coordination of metal ions essential for catalytic activity [108]. Despite the generation of these mutants and positive *in vitro* work, no *in vivo* findings demonstrating the capacity for I-*Sce*I SSB-induced repair has been published.

1.7 RESEARCH GOALS

In order to test recombination at and distant from the site of an SSB or DSB in yeast and human cells, we devised three major aims for carrying out these studies.

1.7.1 To develop systems for testing the recombination potential of DNA breaks induced by different I-*Sce*I variants in yeast and human cells

We sought to develop systems for testing recombination potential of different I-*Sce*I nuclease variants in yeast and in human cells. Additionally, we planned to utilize synthetic DNA oligos as recombination templates in our systems. In order to directly compare the recombination potential of different I-*Sce*I variants that could activate HR through generation of a break, we planned to develop isogenic strains in yeast and lines in human cells. In yeast we aimed to utilize the *delitto perfetto* approach for site-directed mutagenesis. This technique allows for targeted replacement of a genomic sequence in a

two-step process and can be customized for genetic manipulation at any desired locus using a variety of targeting templates. Recombination at disrupted genomic loci with the correcting sequence yields restoration of the gene which can be observed through growth on a selective medium. For construction of isogenic human cell lines, we planned to use episomal as well as chromosomal constructs to directly compare the recombination at a defined locus. In our human systems, recombination at a disrupted fluorescent protein gene locus with the correcting sequence yields restoration of the gene which can be observed through detection of fluorescence. These systems provide useful tools for testing recombination potential of breaks generated by different I-*Sce*I nucleases and can be adapted for studies in which other break-inducing proteins are assayed.

1.7.2 To demonstrate I-*Sce*I SSB-induced recombination at the site of the break *in vivo*

Following generation of our isogenic yeast strains and human cell lines, we sought to explore the capacity for and rules of I-*Sce*I SSB-induced recombination using DNA oligos as repair templates.

*1.7.2.1 Testing I-*Sce*I SSB-induced recombination in yeast*

An SSB has been demonstrated to stimulate targeted gene correction in multiple organisms when generated by various nickases. However, the mechanism by which a nick can stimulate recombination remains elusive. We planned to take advantage of a novel I-*Sce*I nickase, which was previously characterized for its *in vitro* activity, to examine its recombination potential *in vivo* and to provide insights on the mechanism by which recombination is stimulated *in vivo* by nickases. We aimed to develop systems in

S. cerevisiae in which we could directly compare I-*Sce*I SSB- and DSB-induced recombination. We planned to use two systems: a direct repeat system for repair of the lesion at the *LYS2* locus and a system using DNA oligos complementary to the region surrounding a break generated in the *TRP5* gene and capable of restoring the sequence of the gene. Moreover, we aimed to explore the rules of I-*Sce*I SSB-induced recombination including bias for a repair template, genetic requirements, and activity following arrest within the cell cycle.

*1.7.2.2 Testing I-*Sce*I SSB-induced recombination in human cells*

We sought to test the recombination potential of the I-*Sce*I nickase in episomal and chromosomal systems in HEK-293 cells to determine whether the resulting lesion was capable of stimulating HR in human cells. We planned to develop a system for detecting recombination at two reporter sequences (red fluorescent protein, *RFP*, and green fluorescent protein, *GFP*) using DNA oligo-mediated repair. We also aimed to utilize a monoclonal cell line containing a stably integrated non-functional *GFP* target locus to determine if an I-*Sce*I SSB could trigger recombination on the chromosome. If successful, the findings of this study could improve gene targeting approaches by demonstrating the capacity for an I-*Sce*I SSB to stimulate the HR machinery in a mammalian system.

1.7.3 To test I-*Sce*I SSB- and DSB-induced recombination at sites distant from the break

An I-*Sce*I DSB has previously been shown to efficiently stimulate recombination in yeast several kb distant from the site of the chromosomal break. However, little is known about

the capability of a DSB to evoke the same response in human cells and even less is known about the capacity for SSB-induced recombination at a distant locus. To understand if an I-*Sce*I SSB can trigger recombination in yeast not only at the site but also distant from the site of the break in yeast, we planned to develop several constructs for testing this activity. These constructs include isogenic strains in which the cognate recognition sequence is inserted in either orientation 10 kb upstream or downstream from the genomic locus targeted for gene correction. We also planned to develop a system for detecting HR at distant loci (up to 2 kb away) in human cells following generation of an I-*Sce*I DSB or SSB. This system includes a dual reporter (*RFP* and *GFP*) plasmid which can be used to assess distant modification at either reporter, eliminating the variability often involved with making comparisons between different plasmids.

CHAPTER 2

GENE KNOCKOUTS, *IN VIVO* SITE-DIRECTED MUTAGENESIS AND OTHER MODIFICATIONS USING THE *DELITTO PERFETTO* SYSTEM IN *SACCHAROMYCES CEREVISIAE*

The procedures detailed in Chapter 2 are modified from a publication of the same title in the online Methods Navigator series [112]:

Stuckey, S. and Storici, F. (2010) Gene Knockouts, *in vivo* Site-Directed Mutagenesis and Other Modifications Using the *Delitto Perfetto* System in *Saccharomyces cerevisiae*. Methods Navigator: Cookbook of Biomedical Labs. Online.

2.1 PURPOSE

Gene manipulation serves the purpose of providing a better understanding of the function of specific genes as well as for developing novel variants of the genes of interest. The generation of knockout genes, the alteration, depletion, or enhancement of a particular gene function through the generation of specific gene mutations, or the generation of random mutations in a gene are all essential processes for gene manipulation. The genome of the yeast *Saccharomyces cerevisiae* is relatively easy to modify, owing to its efficient HR system. Gene knockout can be a very simple, one step approach to eliminate a gene by substituting its DNA sequence with that of a genetic marker. Differently, desired mutations can be introduced into a gene by replacing the sequence of the normal

gene with that of the mutated gene. Recombinant DNA can be created *in vitro* and then introduced into cells, most often exploiting the endogenous recombination system of the cells. However, unless the desired mutation gives a particular phenotype, a bottleneck of 'recombineering' is the requirement of a selection system to identify the recombinant clones among those unmodified. Even in an organism like yeast where the level of HR is highly above the incidence of random integration, the frequency of homologous targeting is in the range of 10^{-4}-10^{-6} depending on the length of the homology used [113]. Thus, a selection system is always required to identify the targeted clones. Counterselectable markers such as *URA3*, *LYS2*, *LYS5*, *MET15*, and *TRP1* [114-117] are widely utilized in yeast and can be recycled for additional usage in the same yeast strain. If the marker is not eliminated or it is popped out via site-specific recombination between direct repeats such as in the Flp/FRT or Cre/Lox systems, a heterologous sequence is left as a scar at the modified DNA [118,119]. The presence of such scars can threaten the genomic stability of the strain and/or limit the number of successive genetic manipulations for that strain. Here we provide a refined guide, optimized protocols, and tips for using the *delitto perfetto* approach for *in vivo* mutagenesis that combines the practicality of a general selection system with the versatility of synthetic oligos for targeting [102]. The *delitto perfetto* approach allows for generation of gene knockouts and almost any sort of mutation and genome rearrangement via HR. The technique for *in vivo* mutagenesis is designed for efficient and precise manipulation of yeast strains in a two-step process spanning approximately two weeks.

2.2 THEORY

Delitto perfetto is an Italian idiom for 'perfect murder', referring to its capabilities in generating the mutation of a targeted sequence of genomic DNA through insertion and removal of two markers without leaving any trace of foreign DNA in the cells – the perfect deletion of this exogenous material [102]. Collectively these two markers are referred to as a CORE cassette, indicating the presence of both COunterselectable and REporter markers, and are housed on the CORE plasmids: pCORE, pCORE-UK, pCORE-UH, pCORE-Kp53, pCORE-Hp5, pGSKU and pGSHU (**Table 2.1**) [53,102-104,120]. Two of these, pGSKU and pGSHU, are referred to as the break-mediated system plasmids. In addition to the COunterselectable and a REporter markers, they also contain the gene for the I-*Sce*I site-specific endonuclease under regulation of the galactose inducible promoter (*GAL1* promoter) [53]. Taking into account the background of the starting strain, a CORE plasmid must first be selected from one of the seven available. Prior to introducing the chosen CORE cassette into the starting strain, polymerase chain reaction (PCR) is used to amplify the cassette from its respective plasmid. These primers should also be designed such that the tails contain 50 bases of homology upstream and downstream to the targeted genomic sequence for efficient integration.

Examples of genetic modifications that are possible using the *delitto perfetto* approach are shown here.

- *Point Mutations and Small In/Dels:* Point mutations and single or small insertions

Table 2.1 CORE Cassette Composition and Primers Used for Amplification

[1]Two systems exist for use in the *delitto perfetto in vivo* site-specific mutagenesis technique: the non-break and break-induced systems. [2]Seven CORE plasmids are available for use. [3]The sizes of the cassettes vary depending on the markers contained within that cassette. [4]Each cassette contains one COunterselectable and one REporter marker. The break-mediated system plasmids additionally contain the regulatable *GAL1* promoter and I-*Sce*I gene. [5]The sequence of primer needed to amplify each of the CORE cassettes is provided. The plasmids in the break-mediated systems require the addition of the 18 bp I-*Sce*I recognition site (**bold**) next to the *GAL1* promoter. In addition, researchers must include tails with homology to the target site for recombination of the amplified product following PCR.

SYSTEM [1]	PLASMID[2]	CASSETTE[3]	MARKERS[4]	PRIMERS TO AMPLIFY CORE[5]	
Non-break system	**pCORE**	CORE 3.2 kb	*kanMX4* *KlURA3*	*P.1* *P.2*	5' - . . .GAGCTCGTTTTCGACACTGG 5' - . . . TCCTTACCATTAAGTTGATC
	pCORE-UK	CORE-UK 3.2 kb	*KlURA3* *kanMX4*	*P.I* *P.II*	5' - . . .TTCGTACGCTGCAGGTCGAC 5' - . . .CCGCGCGTTGGCCGATTCAT
	pCORE-UH	CORE-UH 3.5 kb	*KlURA3* *hyg*	*P.I* *P.II*	5' - . . .TTCGTACGCTGCAGGTCGAC 5' - . . .CCGCGCGTTGGCCGATTCAT
	pCORE-Kp53	CORE-Kp53 3.7 kb	*kanMX4* *GAL1/10*-p53	*P.I* *P.II*	5' - . . .TTCGTACGCTGCAGGTCGAC 5' - . . .CCGCGCGTTGGCCGATTCAT
	pCORE-Hp53	CORE-Hp53 4.0 kb	*hyg* *GAL1/10*-p53	*P.I* *P.II*	5' - . . .TTCGTACGCTGCAGGTCGAC 5' - . . .CCGCGCGTTGGCCGATTCAT
Break-mediated system	**pGSKU**	GSKU 4.6 kb	*KlURA3,* *kanMX4* *GAL1*-I-*Sce*I	*P.I* *P.IIS*	5' - . . .TTCGTACGCTGCAGGTCGAC 5' . . .**TAGGGATAACAGGGTAAT** CCGCGCGTTGGCCGATTCAT
	pGSHU	GSHU 4.8 kb	*KlURA3, hyg* *GAL1*-I-*Sce*I	*P.I* *P.IIS*	5' - . . .TTCGTACGCTGCAGGTCGAC 5' . . .**TAGGGATAACAGGGTAAT** CCGCGCGTTGGCCGATTCAT

or deletions can be created through the use of a single or complementary oligo(s) as small as 40 bp in length (20 bp of homology to either side of the target region, although the total length is preferably longer) [104].

- *Random Mutations:* Random mutations can be introduced at specific genetic loci, simply by designing oligos with degenerate base sequences that are used to pop-out the CORE cassette previously integrated at the desired location.
- *Large In/Dels:* Larger insertions or deletions can be accomplished through either removal of the CORE with oligo(s) or with a PCR-amplified region of DNA containing tails of homology to the target site. These inserted or deleted sequences can be used, for example, to create gene knockouts or replacements on a particular chromosome and rely on the break-mediated system CORE cassettes for high efficiency of gene modification.
- *Gene Collage:* Gene collage is a term we use for the *in vivo* recombination of several segments of a gene or genes and their integration at the chosen locus following the *delitto perfetto* technique. For example, a selected gene and chosen promoter and terminator sequences can be amplified with primers through the use of PCR to contain regions of homology to each other as well as the target site. Co-transformation of these pieces allows for recombination to yield one continuous segment within any specified site [120].
- *Rearrangements:* Gross chromosomal rearrangements (GCR) can be generated precisely by exploiting the break-mediated *delitto perfetto* approach. These rearrangements include chromosome fusion, circularization and reciprocal translocation [104].

The principle behind the effectiveness of *delitto perfetto* is the utilization of the yeast's HR machinery [52,53,102]. Once the CORE cassette PCR product is inside the cells, the 50 bp tails of homology attached during PCR amplification drive the sequence to the specified chromosomal locus with a frequency of about one targeted cell per 10^6 cells. The resulting integration yields antibiotic resistance as well as provides a marker which can later be used for counterselection. For the plasmids used in the non-break system, these counterselectable markers are either the *URA3* homolog *KlURA3* from *Kluveromyces lactis*, which can be selected against using media containing 5-fluoroorotic acid (5-FOA), or the cDNA of human p53 mutant V122A, which can be selected against using galactose-containing media as it is toxic to yeast when overexpressed [103,104,120,121]. Antibiotic resistance is conferred by either the *hyg* (for Hygromycin B) or *kanMX4* (for Geneticin, or G418) markers [113,122]. Additionally, the plasmids in the break-mediated system contain the galactose-inducible *GAL1* promoter and the I-*Sce*I gene, which, when expressed, generates a DSB at its 18-bp recognition site that is inserted into the chosen genomic site with one of the primers used for the PCR amplification of the CORE [53] (see Table 1). A DSB at this site is utilized only in the second step of *delitto perfetto* to stimulate HR and increase the targeting efficiency more than 1,000 fold [52,53,104]. An example of the overall scheme using the break-mediated CORE plasmid pGSKU illustrates the generation of both a knockout of a chosen gene (YFG) and a site-specific mutation in the same gene (**Figure 2.1**).

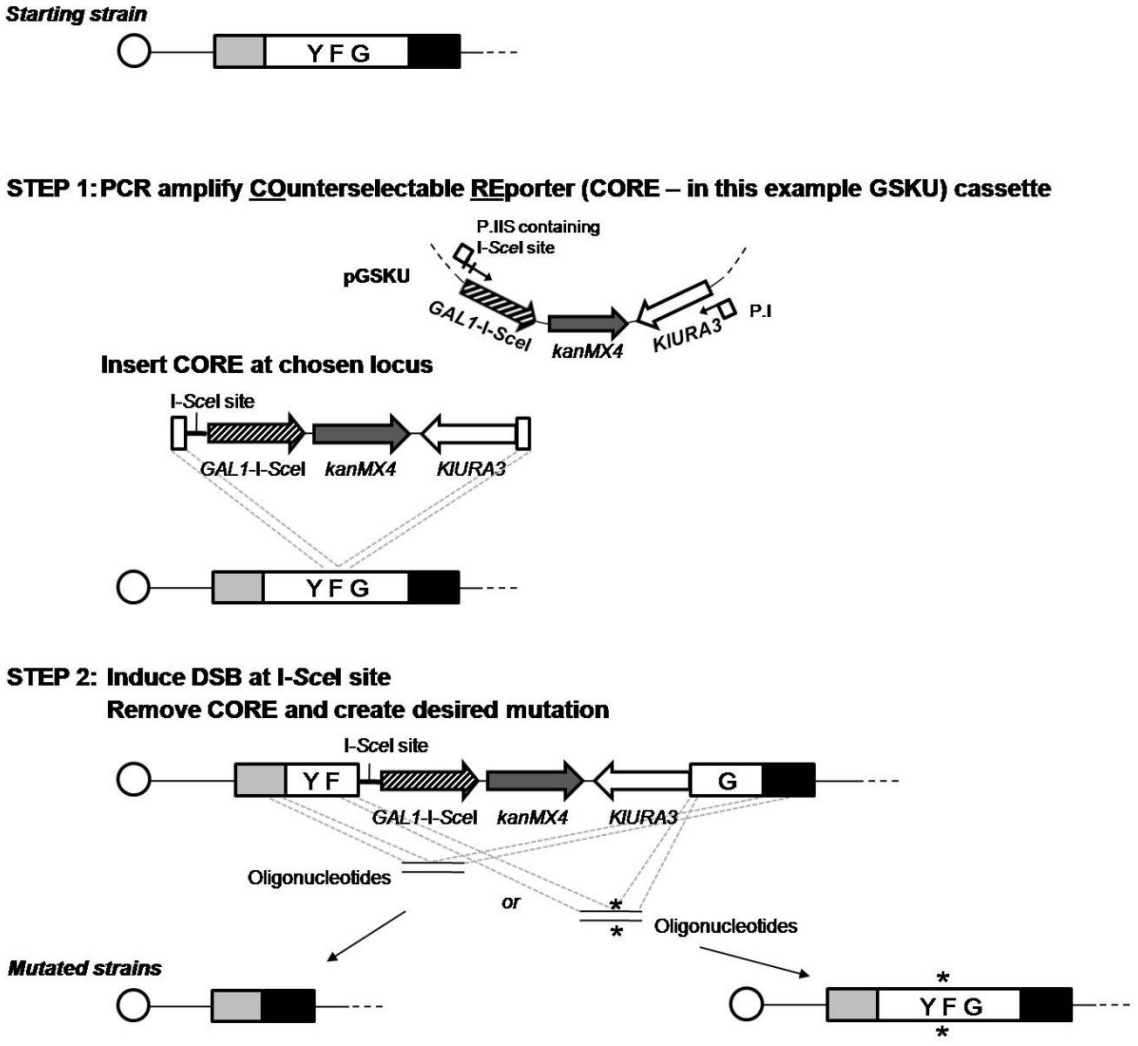

Figure 2.1. **The two-step process of *delitto perfetto* to generate a gene knockout or gene mutation.** The starting strain, containing a normal copy of your favorite gene (YFG), is illustrated at the top of the figure. The first step involves insertion of a CORE cassette (GSKU, using the break-mediated system, is shown here), which is first PCR amplified from the chosen CORE plasmid (pGSKU here). This PCR product contains at least 50 nucleotides of homology to either side of the targeted region, thereby allowing for efficient recombination and integration at this locus. The second step involves removal of the cassette to generate the desired mutation. Here, the induction of a DSB at the I-*Sce*I site following the addition of galactose to the media stimulates recombination and drives the targeting molecules (complementary oligos) to either completely remove the gene (left) or generate a desired mutation (asterisk) within the gene (right).

For successful utilization of the *delitto perfetto* technique, several points must be considered.

- One crucial factor in the *delitto perfetto* method for *in vivo* mutagenesis is working with single-colony isolates in each step of the process. Single-colony isolates are obtained by streaking cells from the initial strain and successively transformant colonies on the YPD rich growth media (recipe below). Once the single-colony isolates are obtained (after 2 days of growth on YPD), these are patched on the YPD for temporary storage, phenotypic and colony PCR testing. If the working strains do not come from single-colony isolates, the subsequent transformation results will be flawed due to a mixed population of cells.
- The *delitto perfetto* approach is designed to be used on haploid strains of yeast. In diploid strains, oligo targeting has diminished capabilities for recombination with the targeted chromosome since it is competing with the gene conversion activity from the homologous chromosome, which is ~250 fold more efficient. However, it is possible to use the *delitto perfetto* technique also in diploid strains, but through a more labor-intensive colony-PCR screening process [104].
- The frequency of spontaneous mutation in the *KlURA3* or the p53 allele is similar to the frequency of CORE pop-out by oligos. However, the *delitto perfetto* approach is designed to circumvent this issue by exploiting the use of the reporter marker. Since the likelihood of spontaneous mutation in both the counterselectable and the reporter markers is very low (expectedly ~$1/10^{12}$), cells that are deficient for both markers of the CORE following phenotypic test in practically all cases have been targeted by the oligos.

- When the *KlURA3* gene is inserted in such a way that it receives interference from another gene's promoter, there may be a delay in growth on media lacking uracil when first inserting the marker or a high level of background when selecting against its function through removal of the marker. This phenomenon, called promoter occlusion (or transcriptional interference), can be resolved through the insertion of *KlURA3* in the orientation opposite that of the conflicting promoter [104].
- Once a CORE cassette is inserted in a chosen chromosomal location, many different mutations can be generated at that locus or in its vicinity, simply by replacing the CORE using different oligos containing the desired mutaitons.

2.3 EQUIPMENT

Autoclave

Centrifuge

Microcentrifuges (with 7 cm radius for 1.5-ml tubes and 15 cm radius for 50-ml tubes)

PCR thermocycler

Agarose gel electrophoresis equipment

UV transilluminator/gel documentation system

Shaking incubator

Vortex mixer

Speed-vac

Heating block

Water bath

Incubator

Micropipettors

Micropipettor tips

1.5-ml microcentrifuge tubes

50-ml conical tubes

250-ml Erlenmeyer flasks

100 × 15 mm Petri dishes

Glass beads

Sterile velveteens (for replica-plating)

2.4 MATERIALS

CORE plasmid(s)

DNA primers for PCR

DNA oligos for transformation

Agarose

Tris base

Boric acid (H_3BO_3)

EDTA

Prestained DNA ladder

6× DNA loading dye

Ethanol (95% and 70%)

Sodium acetate (NaOAc)

Yeast extract

Soy peptone

Dextrose

Agar

Glycerol

Lactic acid

Lactose

Yeast nitrogen base (without amino acids and ammonium sulfate)

Ammonium sulfate [$(NH_4)_2SO_4$]

Amino acids*

Nucleobases

Galactose

Lithium acetate (LiOAc)

Polyethylene glycol 4000 (PEG 4000)

Salmon sperm DNA (SSD)

Geneticin (G418)

Hygromycin B (Hygro)

5-Fluoroorotic acid (5-FOA)

Sterile deionized water

Lyticase

Ex Taq DNA polymerase, 10× buffer (Mg^{2+})

Taq DNA polymerase, 10× buffer (Mg^{2+})

dNTP mix

2.5 SOLUTIONS AND BUFFERS

STEP 2.5.2

Table 2.2 Preparation of 10x TBE
Dissolve Tris and boric acid in ~ 750 ml of deionized water. Add EDTA. Adjust final volume to 1000 ml with water. There is no need to adjust the pH of this solution.
Dilute 50 ml 10× TBE in 950 ml H2O to make 0.5× TBE.

Component	Final concentration	Stock	Amount
Tris base	100 mM	N/A	108 g
EDTA, pH 8.0	1x	0.5 M	40 ml
Boric acid	N/A	N/A	220 g

STEP 2.5.3

Table 2.3 Preparation of 3 M NaOAc, pH 5.2
Filter sterilize.

Dissolve 123 g NaOAc in 450 ml deionized water. Bring volume up to 500 ml.

STEP 2.5.4

Table 2.4 Preparation of Solution 1
Add water to 10 ml.

Component	Final concentration	Stock	Amount
LiOAc	100 mM	1 M	1 ml
TE, pH 7.5	1x	10x	1 ml

Table 2.5 Preparation of Solution 2
Add 50% PEG 4000 – 10 ml.

Component	Final concentration	Stock	Amount
LiOAc	100 mM	1 M	1 ml
TE, pH 7.5	1x	10x	1 ml

Table 2.6 Preparation of YPD liquid medium
Add water to 1 l. Sterilize by autoclaving.

Component	Final concentration	Amount
Yeast extract	1%	10 g
Soy peptone	2%	20 g
Dextrose	2%	20 g

Table 2.7 Preparation of YPD solid medium
Add water to 1 l. Sterilize by autoclaving and pour plates when slightly cooled.

Component	Final concentration	Amount
Yeast extract	1%	10 g
Soy peptone	2%	20 g
Dextrose	2%	20 g
Agar	2%	20 g

Table 2.8 Preparation of YPG solid medium
Add water to 1 l. Sterilize by autoclaving and pour plates when slightly cooled.

Component	Final concentration	Amount
Yeast extract	1%	10 g
Soy peptone	2%	20 g
Glycerol	3%	30 ml
Agar	2%	20 g

Preparation of YPD solid medium containing G418

Prepare 1 l YPD solid medium and sterilize by autoclaving. After the medium has cooled to ~55-60 °C, add 4 ml of 50 mg/ml G418 solution for a final concentration of 200 µg/ml. Pour plates.

Preparation of YPD solid medium containing Hygromycin B

Prepare 1 l YPD solid medium and sterilize by autoclaving. After the medium has cooled to ~55-60 °C, add 6 ml of 50 mg/ml Hygromycin B solution for a final concentration of 300 µg/ml. Pour plates.

Preparation of YPGal solid medium

Prepare 900 ml of YPD w/o glucose medium and sterilize by autoclaving. Separately, dissolve 20 g galactose in 100 ml sterile water and filter sterilize. After the medium has cooled to ~55-60 °C, slowly add 100 ml of galactose 20% solution for a final concentration of 20 g/l galactose (2%). Pour plates.

Table 2.9 Preparation of Synthetic Complete (SC) – Ura solid medium
Add water to 1 l. Sterilize by autoclaving and pour plates when slightly cooled.
* (See [123] for composition in medium)

Component	Final concentration	Amount
Yeast nitrogen base	0.17%	1.7 g
Ammonium sulfate	0.5%	5 g
Dextrose	2%	20 g
Agar	2%	20 g

Amino and nucleic acids*: L-Tryptophan, L-Histidine hydrochloride, L-Arginine hydrochloride, L-Methionine, L-Tyrosine, L-Leucine, L-Isoleucine, L-Lysine hydrochloride, L-Phenylalanine, L-Glutamic acid, L-Aspartic acid, L-Valine, L-Threonine, L-Serine, Adenine sulfate.

Preparation of Synthetic Complete solid medium containing 5-FOA

Prepare 700 ml of Synthetic Complete medium and sterilize by autoclaving. Separately, dissolve 1 g 5-FOA in 300 ml sterile water and filter sterilize. After the medium has cooled to ~55-60 °C, slowly add 5-FOA solution for a final concentration of 1 g/l. Pour plates.

STEP 2.5.6

Table 2.10 Preparation of YPLac liquid medium
Add water to 1 l. Adjust pH to 5.5 with lactic acid. Sterilize by autoclaving.

Component	Final concentration	Amount
Yeast extract	1%	10 g
Soy peptone	2%	20 g
Lactic acid	2.7%	27 ml

Preparation of YPLac overnight (O/N) culture containing galactose

Add 5 ml of the 20% galactose solution to 50 ml YPLac O/N culture to give a final concentration of 2%.

To see the stock solutions recipes, view the Stock Solutions section in Appendix A.

2.6 PROTOCOL

Preparation

- Prepare all media beforehand (see [123]). It is recommended to check starting strain's phenotype prior to transformation.
- Based on the background of the starting strain, select the CORE plasmid to be used. Order DNA primers to amplify the chosen CORE cassette (**Table 2.1**) with tails of homology to either side of the target site.
- Depending on the desired modification, order DNA oligo(s) containing the mutated sequence or primers to amplify the sequence to be inserted with tails of

homology to either side of the target site.

- Sterilize water and velveteen squares (for replica-plating) by autoclaving.

STEP 2.6.1: PCR AMPLIFICATION OF THE CORE CASSETTE

Upon determining the desired mutation(s) to be made, the researcher must select a CORE cassette and amplify it from the plasmid using the DNA primer sequences provided in **Table 2.1**. The primers should also contain 50 bases of homology to either side of the target locus for efficient targeting and recombination. The use of the break-mediated system in the *delitto perfetto* technique allows for an increased efficiency of targeting (more than 1,000 fold), thus providing more applications for the mutagenesis system. However, the generation of a DSB is not always preferred; therefore, the plasmids for the non-break system may be more suitable.

1. Determine the CORE cassette needed based on the background of the original strain.
2. Perform PCR on the sample using Ex Taq DNA polymerase and its provided buffer and dNTP mix. Add to a thin-walled PCR tube:

   CORE plasmid (100 ng/µl) (0.5 µl), DNA primers (50 pmol/µl) (1 µl (each)), 10x Ex Taq buffer (4 µl), dNTP mix (2.5 mM) (4 µl), Ex Taq polymerase (5 U/µl) (0.2 µl), Water (to 40 µl)

3. Run PCR with the following cycling conditions:

1 cycle of	
94 °C	2 min
32 cycles of	
94 °C	30 s
57 °C	30 s
72 °C	4 min (5 min for cassettes over 4 kb in size)
1 cycle of	
72 °C	7 min
1 cycle of	
4 °C	hold

Tip

Set up six reactions so there will be enough of the PCR product for later steps.

STEP 2.6.2: AGAROSE GEL ELECTROPHORESIS TO VISULAIZE THE PRODUCT OF PCR

Following PCR, the product must be visualized by agarose gel electrophoresis to ensure that amplification occurred using the DNA primers.

1. Pour a 0.8% agarose gel in 1× TBE.
2. Mix 2 µl of PCR product with 1 µl 6× DNA loading dye and 3 µl water. Load on gel along with a prestained DNA marker.
3. Run gel at 100 V for about 30 min. Visualize band on a UV transilluminator.

4. If the PCR product is clean, combine the samples to concentrate the PCR product. If there are multiple bands, run the remaining reaction on a gel and purify the correctly sized band.

STEP 2.6.3: CONCENTRATION OF THE PCR PRODUCT

Concentrating the product is necessary prior to transformation into the cells. In this step, you will precipitate the PCR product using ethanol and sodium acetate followed by drying the pellet in a speed vac.

1. Combine the six reactions of PCR in a 1.5-ml microcentrifuge tube. Add 2.5 volumes of 95% ethanol and 0.1 volumes of 3 M NaOAc (pH 5.2).
2. Centrifuge the tube at 13 000 rpm for 10 min. A small pellet should be visible at the bottom of the tube following centrifugation.
3. Discard the supernatant and wash the pellet with 100 µl 70% ethanol without disturbing the pellet. Centrifuge the tube at 13 000 rpm for 5 min and discard the supernatant only. Remove as much ethanol as possible.
4. Dry the pellet in a speed-vac for about 15 min, or until it is completely dry. Resuspend the pellet in 50 µl of water. Five to ten microliters will be used for each transformation.

STEP 2.6.4: INSERTION OF THE CORE CASSETTE INTO THE CHOSEN GENETIC LOCUS

The first step in *delitto perfetto in vivo* mutagenesis involves insertion of the PCR-amplified CORE cassette at the desired yeast genomic locus. This is achieved through

transformation of cells with the PCR product and recombination of the CORE cassette with the cells' genomic DNA. The following transformation uses a procedure modified from the lithium acetate (LiOAc) protocol described by Wach, et al. [113]. For this procedure, LiOAc will be used to make the cells permeable while PEG 4000 is used to adhere the DNA to the cells such that the proximity will optimize the potential entry into the cells. Salmon sperm DNA (SSD) is used to act as carrier DNA and allow the targeting DNA of the PCR product to be protected from DNA degradation factors within the cell.

1. Inoculate 5 ml of YPD liquid medium with the chosen strain and shake (~ 200 rpm) at 30 °C overnight (O/N).
2. Inoculate 50 ml of fresh YPD liquid medium with 1.5 ml of the O/N culture in a 250-ml glass flask and shake vigorously at 30 °C for 3 h.
3. Prepare Solution 1 and Solution 2 immediately prior to transformation. Heat-denature the SSD at 100 °C for 5 min and immediately put on ice prior to transformation.
4. Transfer the culture to a 50-ml conical tube and spin at 3000 rpm for 2 min.
5. Discard the supernatant and wash the cells with 50 ml of sterile water. Spin at 3000 rpm for 2 min.
6. Discard the supernatant and resuspend the cells in 5 ml of Solution 1. Spin at 3000 rpm for 2 min.
7. Discard the supernatant and resuspend the cells in 250 µl of Solution 1. This amount of cells is sufficient for about seven to eight transformations.
8. Aliquot 50 µl of the cell suspension into 1.5-ml microcentrifuge tubes and add 5-10 µl of the concentrated CORE PCR product. Add 5 µl of the heat-

denatured SSD to each tube. Gently mix by tapping the tube.

9. Add 300 µl of Solution 2 for each transformation reaction and mix briefly by vortexing.
10. Incubate the transformation reactions at 30 °C for 30 min at ~ 200 rpm.
11. Heat-shock the cells at 42 °C for 15 min to drive the DNA into the cells.
12. Collect the cells by centrifugation at 5000 rpm for 4 min.
13. Remove the supernatant and resuspend the cells in 100 µl of water.
14. Plate the entire transformation mix onto one SC-Ura plate using ~ 8-12 sterile glass beads. Remove the beads when the plate is dry and incubate at 30 °C for 2-3 days.
15. After colonies are visualized, replica-plate from SC-Ura to G418- or Hygro-containing media (depending on the CORE used) using a sterile velveteen square and incubate at 30 °C O/N.
16. Typically about 5-30 colonies per plate will be observed on the antibiotic-containing media. Streak several for single colony isolates on YPD solid media. Incubate at 30 °C for 2 days.
17. Patch the original strain and the single colonies to new YPD solid media. Incubate at 30 °C O/N.
18. Replica-plate the grown patches to YPD, SC-Ura, G418 or Hygro, YPG (to select against petite cells which are defective in mitochondrial DNA), and any other various selective media depending on the background of your strain. Incubate at 30 °C O/N.
19. Following observation of the correct phenotype, the samples are now ready

for genotypic testing.

Tip

When using the CORE plasmids containing *KlURA3*, if the point of insertion for the marker places it in the same orientation as the targeted gene there may be delay in growth on the SC-Ura media. This is due to interference from the disrupted gene's promoter during transcription. For this reason, it is optimal to insert the cassette such that the insertion of *KlURA3* is oriented opposite to the direction of the gene.

Tip

If delayed growth on SC-Ura media is observed, redo the transformation and instead plate to YPD solid media. Then, replica-plate the G418- or Hygro-containing media the following day and incubate for 2-3 days before replica-plating from the antibiotic media to SC-Ura.

Tip

S. cerevisiae is an aerobic species, so do not cap lids tightly when incubating the O/N culture. Instead, loosely place the lid on and tape it down to secure it.

Tip

Each transformation reaction will require ~6 ml of Solution 1 and 300 µl Solution 2. Therefore, it is important to determine the number of transformations being performed

prior to making the working solutions so as to avoid unnecessary extra work during the procedure.

Tip

The yeast cells will be very difficult to resuspend after the final centrifugation prior to plating. First, pipette the cells to detach them from the tube, then vortex the mixture.

STEP 2.6.5: COLONY PCR OF THE TRANSFORMANTS

Following transformation of the CORE PCR product into the cells and correct phenotypic observation, colony PCR is necessary to observe the correct insertion of the CORE into the desired locus. Primer pairs should include one primer internal to the CORE and one primer external in the target sequence for both markers. Additionally, internal primers of the gene should be paired with its external primers to verify disruption of the gene [120].

1. Resuspend the cells (from a colony about 1 mm^3) in 50 µl of water containing 1 unit of lyticase.
2. Incubate the cells at room temperature for 10 min.
3. Incubate the cells in a heat block at 100 °C for 5 min.
4. Set up PCR. Add to a 0.5-ml thin-walled PCR tube:

   Cell solution (10 µl), Primers (50 pmol/µl) (1 µl (each)), 10x Taq buffer (5 µl), dNTP mix (10 mM) (1 µl), Taq polymerase (5 U/µl) (0.2 µl), Water (to 50 µl)

5. Run PCR with the following conditions:

1 cycle of	
95 °C	2 min
32 cycles of	
95 °C	30 s
55 °C	30 s
72 °C	1 min (1 min per kb)
1 cycle of	
72 °C	7 min
1 cycle of	
4 °C	hold

6. Pour a 1% agarose gel in 1× TBE. Run samples at 100 V for ~ 30 min. Visualize the PCR product using a UV transilluminator.
7. Samples are now ready to remove the CORE.

Tip

The antibiotic resistance and Ura^+ phenotype of the strain following transformation with the CORE PCR product coupled with the correct colony PCR results provide sufficient evidence for correct insertion at the desired locus. Therefore, it is not necessary to perform sequencing analysis for this part of the *delitto perfetto* technique.

Tip

Other digestive enzymes may be used to degrade the thick yeast cell wall, including zymolyase and glusulase. However, lyticase works best in our lab.

STEP 2.6.6: TRANSFORMATION USING DNA OLIGOS TO GENERATE THE DESIRED MUTATION(S)

The second step in *delitto perfetto* involves removal of the CORE cassette and generation of the desired mutation(s). This is accomplished through recombination of a variety of substrates with the cells' genomic DNA. These include a single or pair of complementary oligo(s) containing the mutation(s) or a PCR product. Again, the regions of homology to either side of the target locus drive the removal of the CORE and integration of the targeting DNA. For the following procedure, the mutation is created *via* transformation by DNA oligo(s), which act as carrier themselves at the concentration of 1 nmole/20 µl, thus the use of SSD is unnecessary.

1.a For the non-break system, inoculate 5 ml of YPD liquid medium with the chosen strain and shake (~200 rpm) at 30 °C O/N.

1.b For the break-mediated system, inoculate 50 ml of YPLac liquid medium with the chosen strain in a 250-ml glass flask and shake (~200 rpm) at 30 °C O/N (~ 18-20 h).

2.a For the non-break system, inoculate 50 ml of fresh YPD liquid medium with 1.5 ml of the O/N culture in a 250-ml glass flask and shake vigorously at 30°C for 3 h.

2.b For the break-mediated system, add 5 ml of 20% galactose to the O/N culture and shake vigorously at 30 °C for 3 h.

3. Prepare Solution 1 and Solution 2 immediately prior to transformation. Heat-denature the chosen oligo(s) at 100 °C for 2 min and immediately put on ice prior to transformation.

4. Transfer the culture to a 50-ml conical tube and spin at 3000 rpm for 2 min.
5. Discard the supernatant and wash the cells with 50 ml of sterile water. Spin at 3000 rpm for 2 min.
6. Discard the supernatant and resuspend the cells in 5 ml of Solution 1. Spin at 3000 rpm for 2 min.
7. Discard the supernatant and resuspend the cells in 250 µl of Solution 1. This provides enough cells is for about seven to eight transformations.
8. Aliquot 50 µl of the cell suspension into 1.5-ml microcentrifuge tubes and add 20 µl of oligo(s) (1 nmol). Gently mix by tapping the tube.
9. Add 300 µl of Solution 2 for each transformation reaction and mix briefly by vortexing.
10. Incubate the transformation reactions at 30 °C for 30 min, shaking at ~200 rpm.
11. Heat-shock the cells at 42 °C for 15 min to drive the DNA into the cells.
12. Collect the cells by centrifugation at 5000 rpm for 4 min.
13. Remove the supernatant and thoroughly resuspend the cells in 100 µl of water.

14.a For the non-break system, plate all cells from each transformation tube on one YPD plate using ~ 8-12 sterile glass beads. Remove the beads when the plate is dry and incubate at 30 °C O/N.

14.b For the break-mediated system, dilute cells either 100-, or 1,000-, or 10,000-fold from each transformation tube and plate on a YPD plate using ~8-12 sterile glass beads. Remove the beads when the plate is dry and incubate at 30 °C O/N.

15. Using sterile velveteen, replica-plate from YPD to 5-FOA or YPGal (if using CORE-Kp53 or CORE-Hp53) and incubate at 30 °C for 2 days. If necessary, replica-plate again on 5-FOA or YPGal medium to allow for growth of *ura3* colonies or p53-less colonies, respectively.
16. After visualizing colonies, replica-plate from 5-FOA or YPGal to YPD and G418- or Hygro-containing media (depending on the CORE used) and incubate at 30 °C O/N.
17. Mark the G418-sensitive or Hygro-sensitive colonies on the YPD media and streak several for single colonies on new YPD solid media. Incubate at 30 °C for 2 days.
18. Patch the original strain and the single colonies to new YPD solid media. Incubate at 30 °C O/N.
19. Replica-plate the grown patches to YPD, SC-Ura, G418 or Hygro, YPG (to select against petite cells which are defective in mitochondrial DNA), and any other various selective media depending on the background of your strain. Incubate at 30 °C O/N.
20. Following observation of the correct phenotype, the samples are now ready for genotypic testing using the procedures described in Step 5.
21. PCR samples containing the mutagenized region are now ready for DNA purification and sequencing analysis.

Tip

If delayed growth was seen on SC-Ura media, a high level of background cells may also

be seen on 5-FOA media when removing the CORE. A second round of replica-plating to this media, however, will sufficiently remove these excess cells.

Tip

S. cerevisiae is an aerobic species, so do not cap lids tightly when incubating the O/N culture for the non-break system. Instead, loosely place the lid on and tape it down to secure it.

Tip

When using the break-mediated system, inoculation into YPLac allows the cells to grow in a neutral carbon source medium prior to the addition of galactose. However, the cells will grow much slower so it is suggested that the O/N culture be inoculated in YPLac at least 18-20 h prior to the addition of galactose.

Tip

Each transformation reaction will require ~6 ml of Solution 1 and 300 µl Solution 2. Therefore, it is important to determine the number of transformations being performed prior to making the working solutions so as to avoid unnecessary extra work during the procedure.

Tip

The yeast cells will be very difficult to resuspend after the final centrifugation prior to plating. First, pipette the cells to detach them from the tube, then vortex the mixture.

Tip

The efficiency of the break-mediated system is much greater; therefore, dilute the cells prior to plating. The efficiency depends on the modification performed and on the oligo(s) used, whether a single strand or pairs, or short or long molecules are used.

CHAPTER 3

IN VIVO SITE-SPECIFIC MUTAGENESIS AND GENE COLLAGE USING THE *DELITTO PERFETTO* SYSTEM IN YEAST *SACCHAROMYCES CEREVISIAE*

The procedures detailed in Chapter 3 are modified from a publication of the same title in the Methods in Molecular Biology book series, DNA Recombination edition [120]:

Stuckey, S., Mukherjee, K., and Storici, F. (2011). School of Biology, Georgia Institute of Technology, Atlanta, GA. *Meth Mol Biol.* 745:173-192.

3.1 ABSTRACT

Delitto perfetto is a site-specific *in vivo* mutagenesis system that has been developed to generate changes at will in the genome of the yeast *Saccharomyces cerevisiae*. Using this technique, it is possible to rapidly and efficiently engineer yeast strains without requiring several intermediate steps as it functions in only two steps, both of which rely on HR to drive the changes to the target DNA region. The first step involves the insertion of a cassette containing two markers at or near the locus to be altered. The second step involves complete removal of this cassette with oligos and/or other genetic material and transfer of the expected genetic modification(s) to the chosen DNA locus. Here we provide a detailed protocol of the *delitto perfetto* approach and present examples of the

most common and useful applications for *in vivo* mutagenesis to generate base substitutions, deletions, insertions, as well as for precise *in vivo* assembly and integration of multiple genetic elements, or gene collage.

3.2 INTRODUCTION

The yeast *Saccharomyces cerevisiae* is the most well characterized eukaryotic organism as it has been long utilized for brewing and baking as well as being very easy to grow and manipulate in the laboratory [123]. *S. cerevisiae* was the first eukaryote to have the complete genome sequenced [124], and the genome sequencing project led to the discovery of many new yeast genes with unknown function [125,126]. Moreover, as an 'honorary mammal', *S. cerevisiae* has a large number of genes that are homologs of mammalian and human genes [127]. Thus, functional analysis studies in the yeast model organism shed light on the roles of the corresponding genes in humans and in many other higher eukaryotes. Beyond the simplest experiments of gene disruption or gene knockout, where the original sequence of a gene is replaced with that of a genetic marker [113], site-specific mutagenesis of the genes of interest is the most powerful approach of reverse genetics to reveal what phenotypes arise as a result of the presence of particular genes and to generate novel variants of the genes. Thus, the possibility to generate specific point mutations or localized random changes at will, directly *in vivo* in the DNA locus of choice without leaving behind any marker or other heterologous DNA sequence, provides the opportunity to better understand and modify the role of a given genetic element, or the structure and function of a particular protein. Without leaving any trace, as in the

‘perfect murder’, the *delitto perfetto* (Italian for perfect murder) approach to *in vivo* mutagenesis utilizes simple, precise and highly efficient tools for engineering the genome of yeast cells with the desired modifications [102,103]. Exploiting the tremendous capacity of *S. cerevisiae* to perform efficient HR even when very short regions of homology are involved (30-50 bp) [113], synthetic oligos represent the most versatile and high-throughput device for genome engineering in a homology-driven manner [103]. Moreover, taking advantage of the fact that a DSB stimulates HR 1,000-10,000 fold, using the break-mediated *delitto perfetto* system, it is possible to simultaneously generate multiple different mutants or perform more sophisticated genetic rearrangements that would otherwise be too rare to be detected [52,53,104].

The first step of *delitto perfetto* involves the insertion of a COunterselectable REporter (CORE) cassette containing two markers. Prior to initiating this step the researcher must decide which CORE cassette to use, taking into account the background of the strain (*See* **Note 3.7.1**) to be mutagenized, the markers currently present within this strain and the kind of mutation(s) desired. Seven CORE plasmids have been created (**Figure 3.1**), including those for a non-break system and break system, thereby providing the researcher various choices to utilize this technique.

Amplification of the chosen CORE cassette from its respective plasmid by polymerase chain reaction (PCR) is accomplished using primers which also contain 50-bp tails of homology to either side of the target site (**Table 3.1** and **Figure 3.2A**) to drive the

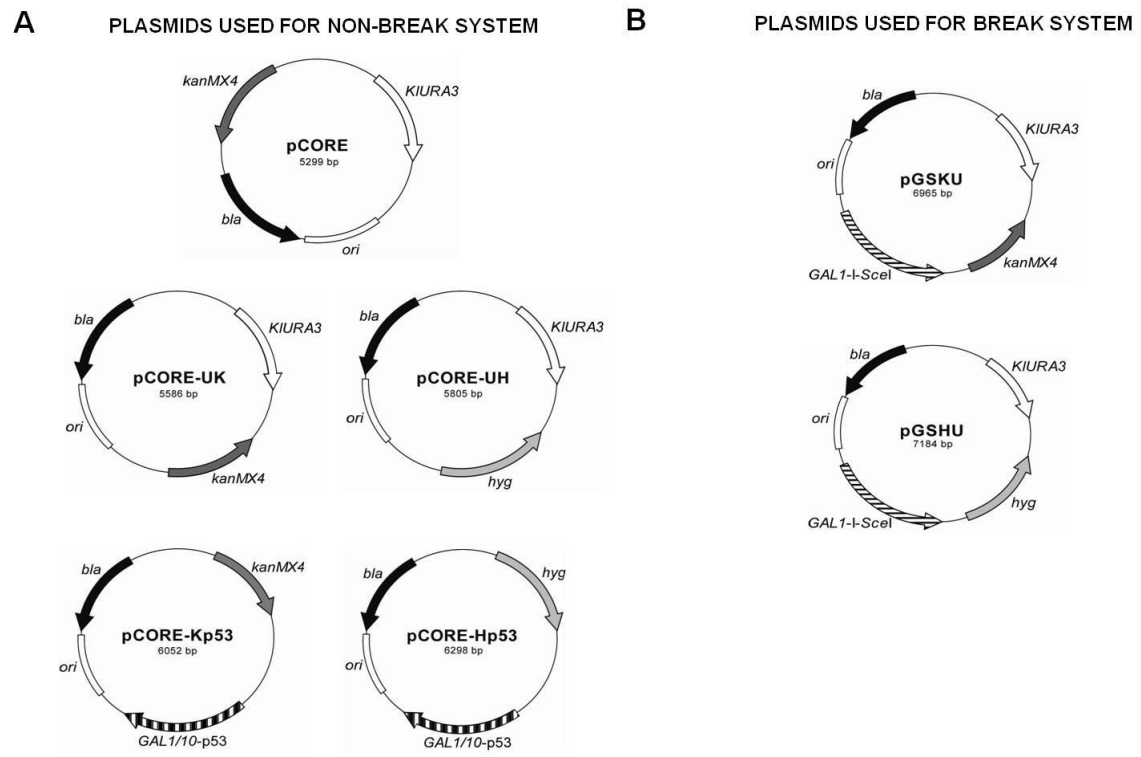

Figure 3.1 The CORE plasmids used in the *delitto perfetto* technique. Each of the five plasmids used in the non-break system (**a**) contains a counterselectable marker, either *KlURA3* from *Kluyveromyces lactis* or a mutant form (V122A) of the human p53 cDNA, and a reporter marker, either kanMX4 conveying resistance to Geneticin (G418) or hyg for resistance to the antibiotic hygromycin B. In addition to these markers, the two plasmids used in the break system (**b**) contain the inducible *GAL1* promoter and I-*Sce*I gene used to express the I-*Sce*I endonuclease and generate a DSB at the I-SceI site. The origin of replication (ori) for all CORE plasmids is indicated as well as the *bla* marker gene, which provides resistance to the β-lactam antibiotic ampicillin and is used for selection.

Table 3.1 Primers for CORE Cassette Amplification and Verification of CORE Cassette Insertion

[a]There are seven CORE plasmids available. [b]Amplification of the plasmids by PCR can be accomplished by creating primers with the above-listed sequences that are internal to the CORE and an external region homologous to the region in which the CORE will be inserted. The primers used to amplify the cassettes from pGSKU and pGSHU require the addition of the 18-bp I-*Sce*I recognition site (**bold**) next to the *GAL1* promoter. [c]The sizes and composition of the cassettes vary depending on the markers present. [d]Primers and their sequences used for verification of CORE integration and replacement are provided. Sequences of primers for testing: GalE (5' - CTAAGATAATGGGGCTCTTT); H1 (5' - CCATGGCCTCCGCGACCGGCTGC); K1 (5' - TACAATCGATAGATTGTCGCAC); K2 (5' - AGTCGTCACTCATGGTGATT); p53.2 (5' – GACTGTACCACCATCCACT); URA3.1 (5' - TTCAATAGCTCATCAGTCGA); URA3.2 (5' - AGACGACAAAGGCGATGCAT)

PLASMID[a]	PRIMERS TO AMPLIFY CORE[b]		CASSETTE[c]	MARKERS[c]	PRIMERS FOR TESTING CASSETTE INSERTION[d]
pCORE	**P.1** **P.2**	5' - . . .GAGCTCGTTTTCGACACTGG 5' - . . . TCCTTACCATTAAGTTGATC	CORE 3.2 kb	*kanMX4*, *KlURA3*	**K2** **URA3.2**
pCORE-UK	**P.I** **P.II**	5' - . . .TTCGTACGCTGCAGGTCGAC 5' - . . .CCGCGCGTTGGCCGATTCAT	CORE-UK 3.2 kb	*KlURA3*, *kanMX4*	**URA3.1** **K1**
pCORE-UH	**P.I** **P.II**	5' - . . .TTCGTACGCTGCAGGTCGAC 5' - . . .CCGCGCGTTGGCCGATTCAT	CORE-UH 3.5 kb	*KlURA3, hyg*	**URA3.1** **H1**
pCORE-Kp53	**P.I** **P.II**	5' - . . .TTCGTACGCTGCAGGTCGAC 5' - . . .CCGCGCGTTGGCCGATTCAT	CORE-Kp53 3.7 kb	*kanMX4*, *GAL1/10*-p53	**K1** **p53.2**
pCORE-Hp53	**P.I** **P.II**	5' - . . .TTCGTACGCTGCAGGTCGAC 5' - . . .CCGCGCGTTGGCCGATTCAT	CORE-Hp53 4.0 kb	*hyg*, *GAL1/10*-p53	**H1** **p53.2**

Table 3.1 (continued)

pGSKU	**P.I** **P.IIS**	5' - . . .TTCGTACGCTGCAGGTCGAC 5' . . .**TAGGGATAACAGGGTAAT** CCGCGCGTTGGCCGATTCAT	GSKU 4.6 kb	*KlURA3, kanMX4, GAL1*-I-*Sce*I	**URA3.1** **Gal.E**
pGSHU	**P.I** **P.IIS**	5' - . . .TTCGTACGCTGCAGGTCGAC 5' . . .**TAGGGATAACAGGGTAAT** CCGCGCGTTGGCCGATTCAT	GSHU 4.8 kb	*KlURA3, hyg, GAL1*-I-*Sce*I	**URA3.1** **Gal.E**

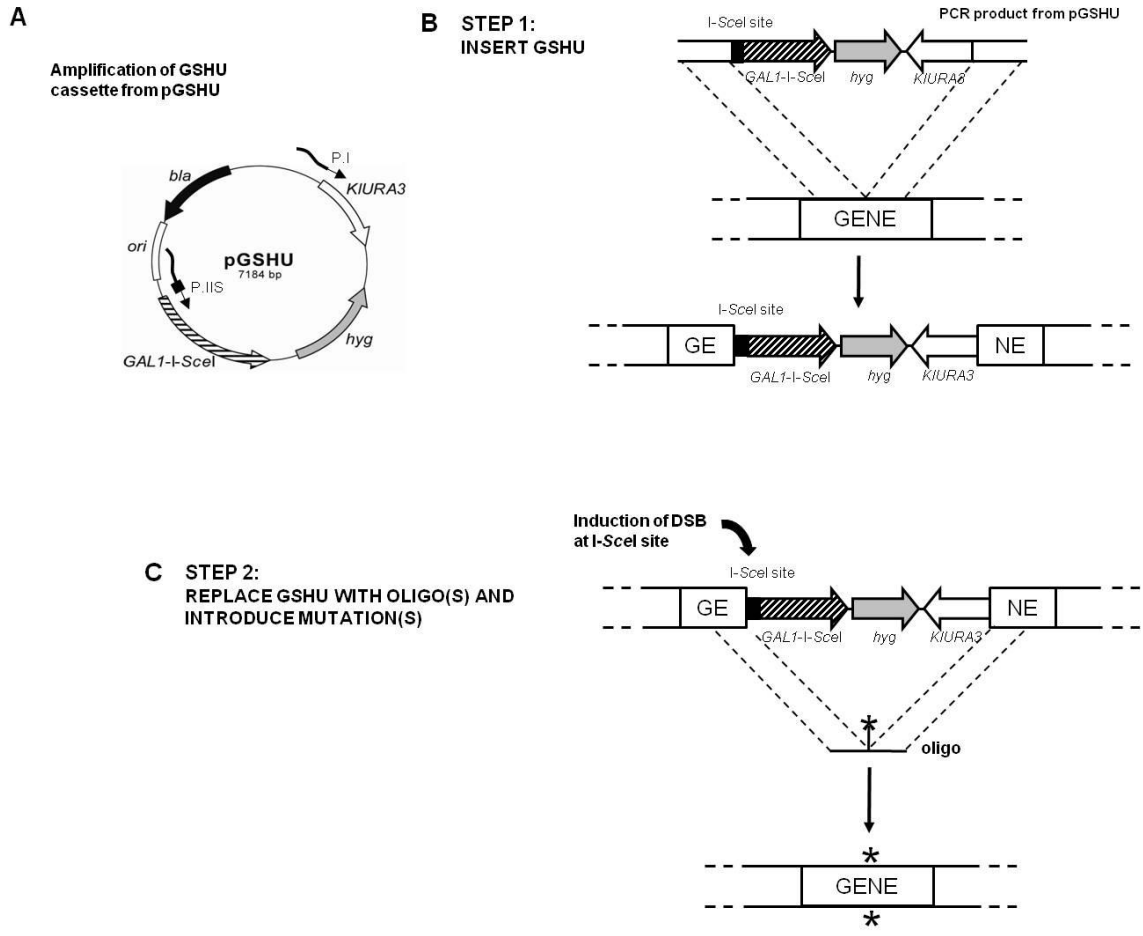

Figure 3.2 The two-step process of *delitto perfetto*. (**a**) Step one involves the amplification of a CORE cassette by PCR (portions of primers used for amplification indicated by thinner line and arrow). (**b**) The primers create tails of homology to either side of the target region (indicated by thicker line) for integration into the genome using the cell's homologous recombination machinery. In this example, the use of the break system CORE cassette GSHU is illustrated. Note that the primer amplifying from the *GAL1*-I-*Sce*I side of the cassette introduces the 18-bp I-*Sce*I recognition site (black box). This site is utilized in the second step (**c**) when the I-*Sce*I endonuclease expression is turned on with galactose to generate a DSB prior to replacement of the CORE with an oligo sequence, which introduces the desired mutation. This example uses a single-stranded oligo to enact this change; however, a pair of complementary oligos has been shown to increase the efficiency of gene targeting.

integration of the CORE to its desired location (**Figure 3.2B**) in the first step of *delitto perfetto*. The second step involves replacement of the entire cassette with oligos or larger pieces of DNA to yield the expected modification to the original segment of chromosomal DNA (**Figure 3.2C**). The generation of a DSB next to the CORE in the break system enhances the efficiency of targeting more than 1,000 fold [52,53,104], expanding the applications of the mutagenesis system. From beginning to end, *delitto perfetto* yields the final strain in less than 2 weeks and has proven to be a very useful tool in molecular biology. Examples provided in this review illustrate many changes that can be created through removal of the CORE, such as point mutations, random mutations, deletions, insertions ranging from a few nucleotides to fragments several kilobases in size, and *in vivo* gene collage.

3.3 MATERIALS

3.3.1 Amplification of CORE

1. Seven CORE plasmids are available (*See* **Figure 3.1**).
2. DNA primers (Invitrogen, Carlsbad, CA; or Alpha DNA, Montreal, Quebec, Canada), desalted and non-purified: 50 pmol/µl. Store at -20°C.
3. Ex Taq DNA polymerase, 10x buffer, 2.5 mM dNTPs (Clontech, Mountain View, CA).

3.3.2 Gel Electrophoresis

1. Agarose (Fisher, Pittsburgh, PA).

2. TBE running buffer (10x) (Fisher).
3. Prestained molecular weight marker (New England Biolabs, Ipswich, MA).
4. Loading dye (Fisher).

3.3.3 PCR Product Concentration

1. Ethanol (EtOH): 95% and 70% concentrations.
2. Sodium acetate (NaOAc, Sigma, St. Louis, MO) 3 M (pH 5.2) stock solution, filter sterilized. Store at room temperature. (*See* **Note 3.7.2**).

3.3.4 Transformation Reagents and Media

1. YPD (per 1 l): 10 g yeast extract, 20 g soy peptone, 20 g dextrose (Difco/BD, Franklin Lakes, NJ). For solid media, add 20 g agar (Difco/BD). (*See* **Note 3.7.3**).
2. YPLac liquid (per 1 l): 10 g yeast extract, 20 g soy peptone, 27 ml lactic acid (Difco/BD), pH adjusted to 5.5 with lactic acid (Fisher).
3. Stock solution of 20% high-pure galactose (Sigma) is filter sterilized and stored at room temperature.
4. Lithium Acetate (LiOAc, Sigma) stock of 1 M concentration filter sterilized and stored at room temperature.
5. TE 10x stock solution: 100 mM Tris (Fisher) (pH 7.5), 10 mM ethylenediaminetetraacetic acid (EDTA; Sigma) (pH 7.5). Filter sterilize. Store at room temperature.

6. Polyethylene glycol 4,000 (PEG 4,000; Sigma): 50% stock solution. Store at room temperature. (*See* **Note 3.7.4**).
7. Working solutions: Solution 1 (0.1 M LiOAc, TE 1x, pH 7.5) and solution 2 (0.1 M LiOAc, TE 1x, pH 7.5 in PEG 4000 50%).
8. Solution of salmon sperm DNA (SSD, Roche, Basel, Switzerland), 100 µg/ml. Store at -20°C.
9. SC-Ura (Synthetic complete media lacking uracil) solid media (Fisher).
10. Glass beads, approx. 5 mm diameter (Fisher).
11. 5-Fluoroorotic Acid (5-FOA; per 1 l): Solution of 5-FOA is prepared by dissolving 1 g 5-FOA (US Biological, Swampscott, MA) in 300 ml of water prior to filter sterilization. 700 ml SD-complete (synthetic dextrose-complete) agar media is autoclaved, then cooled to 55-60°C, and the filtered solution of 5-FOA is then mixed with media prior to pouring.
12. G418 (per 1 l): YPD agar media is autoclaved, then cooled to 55-60°C, and G418 solution (200 µg/ml; US Biological) is then mixed with media prior to pouring. Stock solution is prepared in 50 mg/ml filter-sterilized aliquots and stored at 4°C.
13. Hygromycin B (Hygro; per 1 l): YPD agar media is autoclaved, then cooled to 55-60°C, and Hygro solution (300 µg/ml; Invitrogen) is then mixed with media prior to pouring.
14. YPG (per 1 l): 10 g yeast extract, 20 g soy peptone, 30 ml glycerol (Difco/BD), 20 g agar.
15. Sterile velveteens (Fisher).

3.3.5 Genotypic Testing of Transformants

1. Lyticase (Sigma) is dissolved at 2,000 U/ml and stored in 1 ml aliquots at -20°C.
2. Taq DNA polymerase, 10x buffer, 10 mM dNTPs (Roche).

3.3.6 Design of DNA Oligos for Removal of CORE and Generation of Mutations

1. DNA oligos (Invitrogen or Alpha DNA): 50-100mers, desalted and non-purified (50 pmol/µl). Store at -20°C.

3.4 METHODS

Despite the efficiency of recombination when a DSB is induced, induction of a DSB may not be required depending on the strain being mutagenized and the type of modification. The DSB system is preferred when multiple mutations are desired simultaneously; when the modification involves gross deletions, insertions, gene fusions, or other genomic rearrangements [104]; and when the strain is deficient in homologous recombination functions [52].

Several combinations of two markers can be used for the *delitto perfetto* technique and are contained within the various CORE cassettes on plasmids (**Figure 3.1**). The two CORE markers are used for selection purposes and consist of the following: an antibiotic resistance marker (REporter)—which confers resistance to the antibiotics hygromycin B

or Geneticin (G418)—and a COunterselectable marker, either the *KlURA3* gene (a *URA3* homolog from *Kluyveromyces lactis*), which can be selected against using 5-FOA, or a marker coding for the human p53 mutant V122A, which is toxic to yeast when overexpressed and can be selected against using a galactose-containing media. In addition, the break system cassettes include the inducible *GAL1* promoter and the I-*Sce*I gene, used to induce the DSB at the 18-bp I-*Sce*I break site.

In the first step of *delitto perfetto*, the CORE is amplified through PCR to attach the tails of homology to the desired chromosomal locus (**Fig 3.2A**) and its PCR product is inserted into the cells by transformation (**Figure 3.2B**). The CORE cassette will then integrate at the desired genomic locus in approx. $1/10^6$ yeast cells via homologous recombination. Following transformation, the transformant colonies are isolated to observe for insertion of the CORE through phenotypic and genotypic testing. The second step of this technique is a transformation using oligos or other DNA to remove the entire CORE cassette and introduce the desired mutation(s) (**Figure 3.2C**). *See* **Section 3.4.6** for details on oligo design to remove the CORE.

3.4.1 Amplification of CORE from Plasmid

1. DNA primers will first be used to amplify the CORE from the chosen plasmid. These primers range from 70 to 100 bases in length with an overlap of at least 50 bp with the genomic targeting region and an overlap of 20 bp to the CORE

cassette sequence (**Table 3.1**). Additionally, in the break-induced system, the 18-bp recognition sequence for the I-*Sce*I endonuclease is included on one of the two primers (**Table 3.1**).

2. PCR conditions: Amplification of the CORE cassette from circular plasmid (about 50 ng) using 50 pmol/µl of each primer is performed with high yield in a final volume of 40 µl using Ex Taq DNA polymerase with a 2 min cycle at 94°C; 32 cycles of 30 s at 94°C, 30 s at 57°C, and 4 min at 72°C (or 5 min at 72°C for cassettes over 4 kb in size); a final extension time of 7 min at 72°C; and samples are held at 4°C. Ex Taq DNA polymerase consistently produces a higher yield of CORE cassette amplification than does Taq DNA polymerase. dNTPs (10 mM) are used for this reaction. An extension time of 1 min/kb is assumed for this reaction.
3. Following PCR, the samples are ready for gel electrophoresis and PCR product concentration.

3.4.2 Gel Electrophoresis

1. We use a dilution of 0.5x TBE running buffer, which is obtained from 10x TBE by mixing 50 ml of 10x TBE buffer with 950 ml deionized water prior to use.
2. A small aliquot (about 2 µl) of PCR product is run on a 0.8% agarose gel to observe anticipated band.

3.4.3 PCR Product Concentration

1. The product of six reactions of PCR are combined for precipitation with a 2.5x volume of 95% EtOH and 1/10 3 M NaOAc (pH 5.2) in a microcentrifuge tube. Centrifugation is carried out at maximum speed for 10 min. A small pellet should be visible on the bottom of the tube.
2. The supernatant is discarded and the pellet is washed with 100 µl of 70% EtOH, paying attention not to detach the pellet. If the pellet is detached, it is necessary to spin again for 5 min and then discard the supernatant. Then, as much as possible of the EtOH is removed without detaching the pellet.
3. The pellet is then dried in a speed vac for about 15 min and resuspended in 50 µl of water. Five to 10 µl are used for each transformation.

3.4.4 Step 1: Transformation to Insert the CORE

The following transformation protocol is used to first insert the CORE PCR product into the strain of choice, and then to drive replacement of the CORE with DNA oligos or other segments of DNA. This transformation procedure has been modified from the lithium acetate protocol described by Wach, *et al*. [113]. During the transformation, the LiOAc acts to make the cell wall permeable. The presence of PEG 4000 is used to adhere the DNA to the cells such that the proximity will allow for entry into the cells. When transforming to insert the CORE PCR product, SSD is used to act as carrier DNA and serves as a buffer between the targeting DNA from the PCR and any DNA degradation factors present within the cell. In the second transformation using oligos to remove the

CORE, the use of SSD is unnecessary as the oligos at the concentration of 1 nmole/20 µl act as carrier DNA themselves:

1. Inoculate 5 ml of YPD liquid medium with chosen strain and shake at 30°C overnight (O/N). (*See* **Note 3.7.5**).
2. Inoculate 50 ml of YPD liquid medium with 1.5 ml of the O/N culture in a 250-ml glass flask and shake vigorously at 30°C for 3 h.
3. Solution 1 and Solution 2 are prepared immediately prior to transformation.
4. Transfer culture to a 50-ml conical tube and spin at 1,562 x *g* for 2 min.
5. Remove the supernatant and wash cells with 50 ml of sterile water and spin as stated previously.
6. Remove the supernatant and resuspend cells in 5 ml of solution 1 and spin as stated previously.
7. Remove supernatant and resuspend cells in 250 µl of solution 1. This amount of cells is sufficient for approx. 7-8 transformations.
8. Aliquot 50 µl of the cell suspension in microcentrifuge tubes and add 5-10 µl of concentrated CORE PCR product and 5 µl of SSD (heat-denatured SSD for 5 min at 100°C prior to use and immediately kept on ice), then gently mix by tapping the tube.
9. Add 300 µl of solution 2 for each transformation reaction. Mix briefly by vortexing.
10. Incubate transformation reactions at 30°C for 30 min with shaking.
11. Heat shock at 42°C for 15 min to drive the DNA into the cells.
12. Collect cells by centrifugation at 4,340 x *g* for 4 min.

13. Remove supernatant and resuspend cells well in 100 µl of water.
14. Plate all cells from each transformation tube on one SC-Ura plate using approx. 8-12 sterile glass beads and incubate at 30°C for 2-3 days. (*See* **Note 3.7.6**).
15. Using sterile velveteen, replica-plate from SC-Ura to G418- or Hygro-containing media (depending on the CORE used) and incubate at 30°C O/N.
16. Once transformants are observed (typically 5-30 colonies per plate), streak for single colony isolates on YPD solid media. Incubate at 30°C for 2 days.
17. Make patches of the single colonies on new YPD solid media, along with the original strain, and incubate at 30°C O/N.
18. Replica-plate the grown patches to YPD, SC-Ura, G418, Hygro, YPG to select against petite cells, and any other various selective media depending on the background of your strain, and incubate at 30°C O/N.
19. Following observation of correct phenotype, the samples are ready for genotypic testing.

3.4.5 Colony PCR of Transformants

1. Resuspend cells (approx. 1 mm^3) in 50 µl water containing 1 U of lyticase. Incubate at room temperature for 10 min, followed by incubation in a heat block at 100°C for 5 min.
2. PCR conditions: Colony PCR of the transformant patches presenting the expected phenotypes using 10 µl of the cell resuspension solution is

accomplished with 50 pmol of each primer, with an expected amplification region between 300 bp and 1.2 kb (*See* **Figure 3.3**). dNTPs (10 mM) are used for this reaction.

3. PCR is performed in a final volume of 50 µl using Taq DNA polymerase (Roche) with a 2 min cycle at 95°C; 32 cycles of 30 s at 95°C, 30 s at 55°C, and 1 min at 72°C; a final extension time of 7 min at 72°C; and samples are held at 4°C. An extension time of 1 min/kb is assumed for this reaction. (*See* **Note 3.7.7**).
4. Following PCR, samples are run on a 1% agarose gel (*See* **Section 3.4.2**) for observation of PCR product.
5. Strains are now ready for step 2 to remove the CORE.

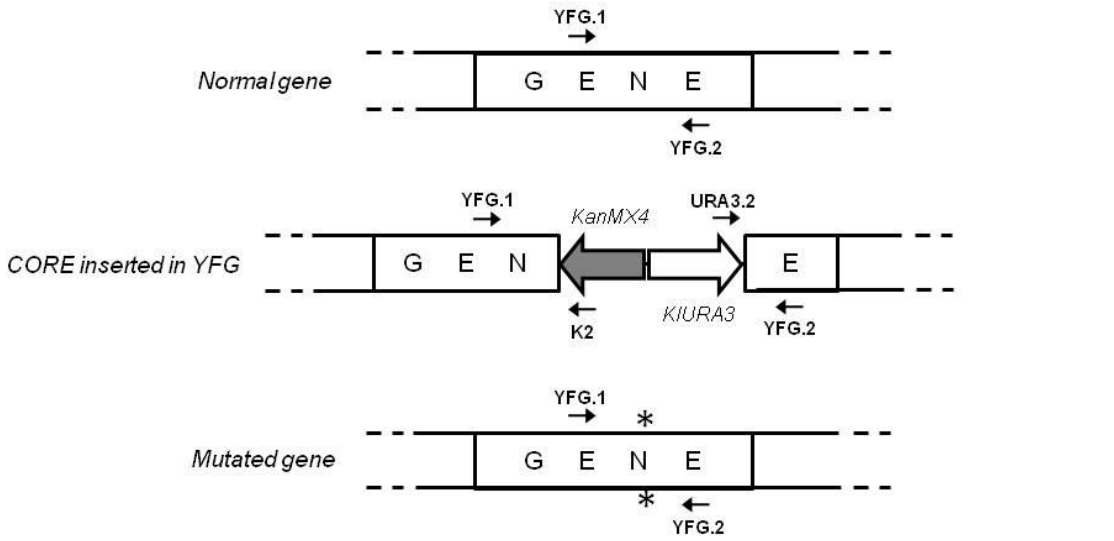

Figure 3.3 Scheme of primer pairs used for colony PCR. Primers should be designed to allow for amplification of the target region in addition to being paired with primers internal to the CORE. The sizes of colony PCR products should range between 300 bp and 1.2 kb. Using this approach, verification of the CORE's integration as well as its replacement can be made. See **Table 3.1** for a list of primers and their sequences used to verify the integration of the various CORE cassettes.

3.4.6 Design of DNA Oligos for Removal of CORE and Generation of Mutations

Numerous mutations can be accomplished through the use of the *delitto perfetto* technique. These include substitutions, insertions, the generation of random mutations through the use of degenerate oligos, and deletions. **Figure 3.4** illustrates the sequence of oligos (A-D) needed to produce all of these mutations at the genomic locus indicated in the figure. When substitutions or insertion mutations are desired, the location of the CORE insertion should be next to the region of modification. Conversely, the CORE should replace the entire targeted region when a small deletion or random mutation is

desired. If a large deletion is desired, a CORE with the break system is inserted within the region to be removed [104].

To remove the CORE and generate the mutation with DNA oligos, the following considerations should be made. The use of a single oligo is sufficient; however, a pair of complementary oligos increases the frequency of integration 5-10 fold [104]. Additionally, while shorter oligos (≈40 bp) can be used to effectively transform the strain, longer oligos approaching lengths of 80 bases are more favorable as they increase the efficiency of targeting as well as the window of mutagenesis [104]. The external 30-40 bases of the oligo or oligo pair are used for efficient homologous recombination to introduce the desired mutation and allowing for loss of the CORE. It is of note that once the CORE cassette has been integrated in a specific chromosomal locus, many gene variants can be generated by transforming the cells with oligos designed to produce different alterations.

3.4.7.1 Step 2: Transformation Using DNA Oligos in *Non-Break* System

1. Inoculate 5 ml of YPD liquid medium with chosen strain and shake at 30°C (O/N).
2. Inoculate 50 ml of YPD liquid medium with 1.5 ml of the O/N culture in a 250-ml glass flask and shake vigorously at 30°C for 3 h.
3. Solutions 1 and 2 are prepared immediately prior to transformation.

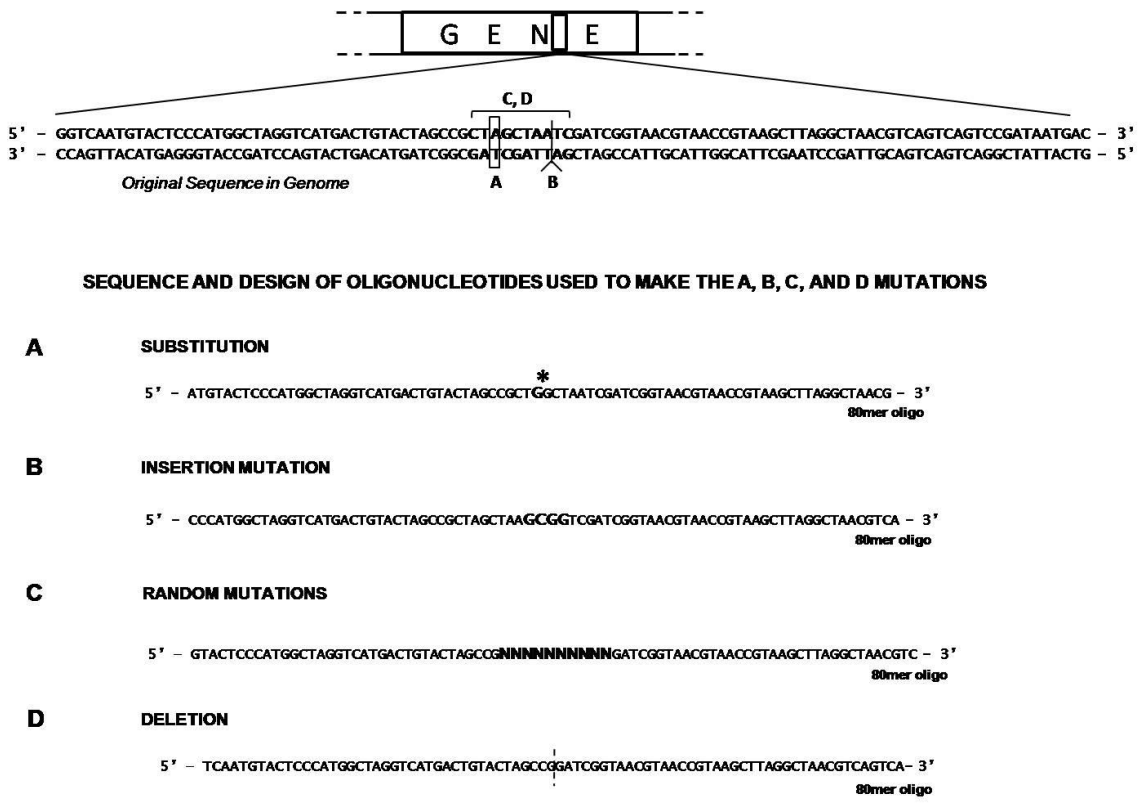

Figure 3.4 Examples of single oligo-driven mutations generated using the *delitto perfetto* technique. When a substitution or an insertion mutation is desired (A, B), the CORE should be placed next to the target region prior to replacement with a single or complementary oligo(s). In this example, the original sequence in the genome is provided as a reference at the top of the figure. (A) A substitution of a guanine, marked by an asterisk above the bolded G on the oligo, is made in place of the adenine residue on the top strand of the reference sequence (boxed). (B) An insertion mutation in the original sequence is created through the use of an oligo containing additional nucleotides (GCGG, marked in bold) which are inserted between the adenine and thymine indicated in the reference. When random mutations or small deletions (<5 kb) are desired (C, D) in a specific region, it is preferred to delete the region of interest along with the CORE insertion, as the successive targeting event with the oligos or other DNA will then eliminate the CORE and introduce the desired changes. The region of mutagenesis in the original sequence is bolded and indicated by the bracket. (C) For the generation of random mutations, the oligo sequence contains a stretch of 10 bolded Ns, which indicate that any of the four DNA bases can be used when the oligos are synthesized. The exact degree of this randomness is determined by the investigator. (D) The segment of the reference sequence indicated by the bracket can be removed through the use of oligos missing this fragment. In the example, the location of the deleted nucleotides is indicated by the dashed line.

4. Transfer culture to a 50-ml conical tube and spin at 1,562 x *g* for 2 min.
5. Remove the supernatant and wash cells with 50 ml of sterile water and spin as stated previously.
6. Remove the supernatant and resuspend cells in 5 ml of solution 1 and spin as stated previously.
7. Remove the supernatant and resuspend cells in 250 µl of solution 1. This amount of cells is sufficient for approx. 7-8 transformations.
8. Aliquot 50 µl of the cell suspension in microcentrifuge tubes and add 1 nmole of DNA oligos (heat-denatured for 2 min at 100°C, then immediately kept on ice prior to use). When using a single oligo, a 20 µl volume (at 50 pmol/µl) is used or when using a complementary DNA oligo pair, 10 µl of each is used. Gently mix the tube by tapping.
9. Add 300 µl of solution 2 for each transformation reaction. Mix briefly by vortexing.
10. Incubate transformation reactions at 30°C for 30 min with shaking.
11. Heat shock at 42°C for 15 min to drive the DNA into the cells.
12. Collect cells by centrifugation at 4,340 x *g* for 4 min.
13. Remove the supernatant and resuspend cells well in 100 µl of water.
14. Plate cells from each transformation tube on one YPD solid plate using approx. 8-12 sterile glass beads and incubate at 30°C O/N.
15. Using sterile velveteen, replica-plate from YPD to 5-FOA and incubate at 30°C for 2 days. If necessary, replica-plate again on 5-FOA media to allow for growth of Ura^- colonies clearly distinct from the background (*See* **Note 3.7.8**).

16. Using sterile velveteen, replica-plate from 5-FOA to YPD and G418- or Hygro-containing media (depending on the CORE used) and incubate at 30°C O/N.
17. Mark G418-sensitive or Hygro-sensitive colonies on the YPD media and streak for single colonies on new YPD solid media. Incubate at 30°C for 2 days.
18. Make patches of the single colonies on new YPD solid media, along with the original strain, and incubate at 30°C O/N.
19. Replica-plate patches to YPD; SC-Ura; G418; Hygro; YPG, which selects against cells with defective mtDNA; and any other various selective media depending on the background of your strain, and incubate at 30°C O/N.
20. Following observation of correct phenotype, the samples are ready for genotypic testing (*See* **Section 3.4.5**).
21. PCR samples containing the mutagenized region are now ready for DNA purification and sequencing analysis. (*See* **Note 3.7.9**).

3.4.7.2 Step 2: Transformation Using DNA Oligos in *Break* System

1. Inoculate 50 ml of YPLac liquid medium with chosen strain in a 250-ml glass flask and shake at 30°C O/N. (*See* **Note 3.7.10**).
2. Add 5 ml of galactose from a 20% solution into the O/N culture to obtain a 2% galactose solution and shake vigorously at 30°C for 3-6 h. (*See* **Note 3.7.11**).
3. Solutions 1 and 2 are prepared immediately prior to transformation.
4. Transfer culture to a 50-ml conical tube and spin at 1,562 x *g* for 2 min.

5. Remove the supernatant and wash cells with 50 ml of sterile water and spin as stated previously.
6. Remove the supernatant and resuspend cells in 5 ml of solution 1 and spin as stated previously.
7. Remove the supernatant and resuspend cells in 250 µl of solution 1. This amount of cells is sufficient for approx. 7-8 transformations.
8. Aliquot 50 µl of the cell suspension in microcentrifuge tubes and add 1 nmole of DNA oligos (heat-denatured for 2 min at 100°C, then immediately kept on ice prior to use). When using a single oligo, a 20 µl volume (at 50 pmol/µl) is used or when using a complementary DNA oligo pair, 10 µl of each is used. Gently mix the tube by tapping.
9. Add 300 µl of solution 2 for each transformation reaction. Mix briefly by vortexing.
10. Incubate transformation reactions at 30°C for 30 min with shaking.
11. Heat shock at 42°C for 15 min to drive the DNA into the cells.
12. Collect cells by centrifugation at 4,340 x *g* for 4 min.
13. Remove the supernatant and resuspend cells well in 100 µl of water.
14. Plate cells from each transformation tube on one YPD solid plate using approx. 8-12 sterile glass beads and incubate at 30°C O/N. Dilutions may be necessary prior to plating due to the efficiency of oligo recombination following DSB induction.
15. Using sterile velveteen, replica-plate from YPD to 5-FOA and incubate at 30°C for 2 days. If necessary, replica-plate again on 5-FOA media to allow for

growth of Ura$^-$ colonies clearly distinct from the background. (*See* **Note 3.7.8**).

16. Using sterile velveteen, replica-plate from 5-FOA to YPD and G418-or Hygro-containing media (depending on the CORE used) and incubate at 30°C O/N.
17. Mark G418-sensitive or Hygro-sensitive colonies on the YPD media and streak for single colonies on new YPD solid media. Incubate at 30°C for 2 days.
18. Make patches of the single colonies on new YPD solid media, along with the original strain, and incubate at 30°C O/N.
19. Replica-plate patches to YPD; SC-Ura; G418; Hygro; YPG, which selects against cells with defective mtDNA; and any other various selective media depending on the background of your strain, and incubate at 30°C O/N.
20. Following observation of correct phenotype, the samples are ready for genotypic testing. (*See* **3.4.5**).
21. PCR samples containing the mutagenized region are now ready for DNA purification and sequencing analysis. (*See* **Note 3.7.9**).

3.5 EXAMPLES

3.5.1 The *Delitto Perfetto* Approach to Insert a Large DNA Fragment

In **Figure 3.5**, the two-step process shown illustrates the insertion of a large segment of DNA, 10 kb in size. Generally, an insert of these proportions is obtained through amplification of the sequence through PCR, which, although possible, greatly increases the risk of introducing several mutations through the extension process. In our system,

the large DNA of interest is carried on a plasmid which is linearized prior to transformation. Linearization of the plasmid is required to generate free DNA ends and stimulate homologous recombination. The large segment of plasmid DNA is integrated into genomic DNA at a chosen location by *in vivo* recombination following co-transformation of the linearized plasmid carrying the fragment and two pairs of complementary oligos. Each pair contains regions of homology to either side of the target site in addition to homology with the 10-kb fragment, thereby directly driving it into its desired locus. This way, sequencing analysis is not required following integration of the large fragment.

3.5.2 The *Delitto Perfetto* Approach to Insert Multiple Sequences for Gene Collage

It is also possible to insert two or more sequences or genes next to each other simultaneously using *delitto perfetto,* as seen in **Figure 3.6**. To accomplish this, the genes or segments of interest are amplified in such a way that the primers of each PCR fragment have tails of homology to the sequence of the contiguous segment and the most external primers contain homology to the target site. Through co-transformation with these multiple PCR products, the individual pieces recombine *in vivo* as a form of gene collage, while the outlying primers drive integration into the genome at the desired locus.

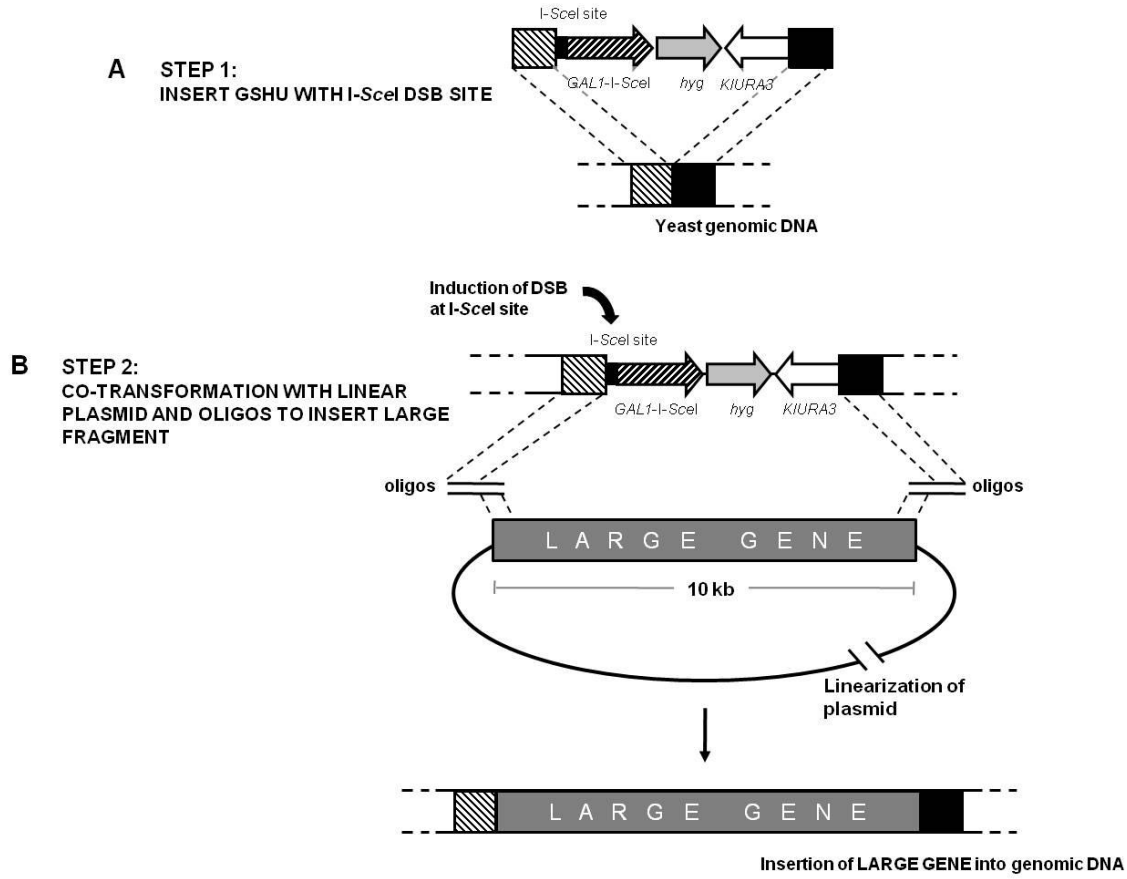

Figure 3.5 Insertion of a large segment of DNA from a plasmid. In this example, the break system is used to drive the integration of a 10-kb fragment from a plasmid into the target region. (**a**) First, the GSHU CORE cassette and the 18-bp I-*Sce*I break site are inserted into the target region. (**b**) Next, the plasmid carrying the sequence of interest to be integrated within the genome is linearized by restriction digestion outside of the fragment to be inserted. Complementary pairs of oligos have regions of homology to both the upstream and downstream portions of the sequence of interest to be integrated, as well as to either side of the target region. The linearized plasmid and oligos are co-transformed into yeast cells following DSB induction at the I-*Sce*I site. By homologous recombination, the large sequence of interest is integrated into the genomic DNA at the specific site without the need for PCR amplification, which otherwise increases the likelihood of unwanted mutations during the polymerization process.

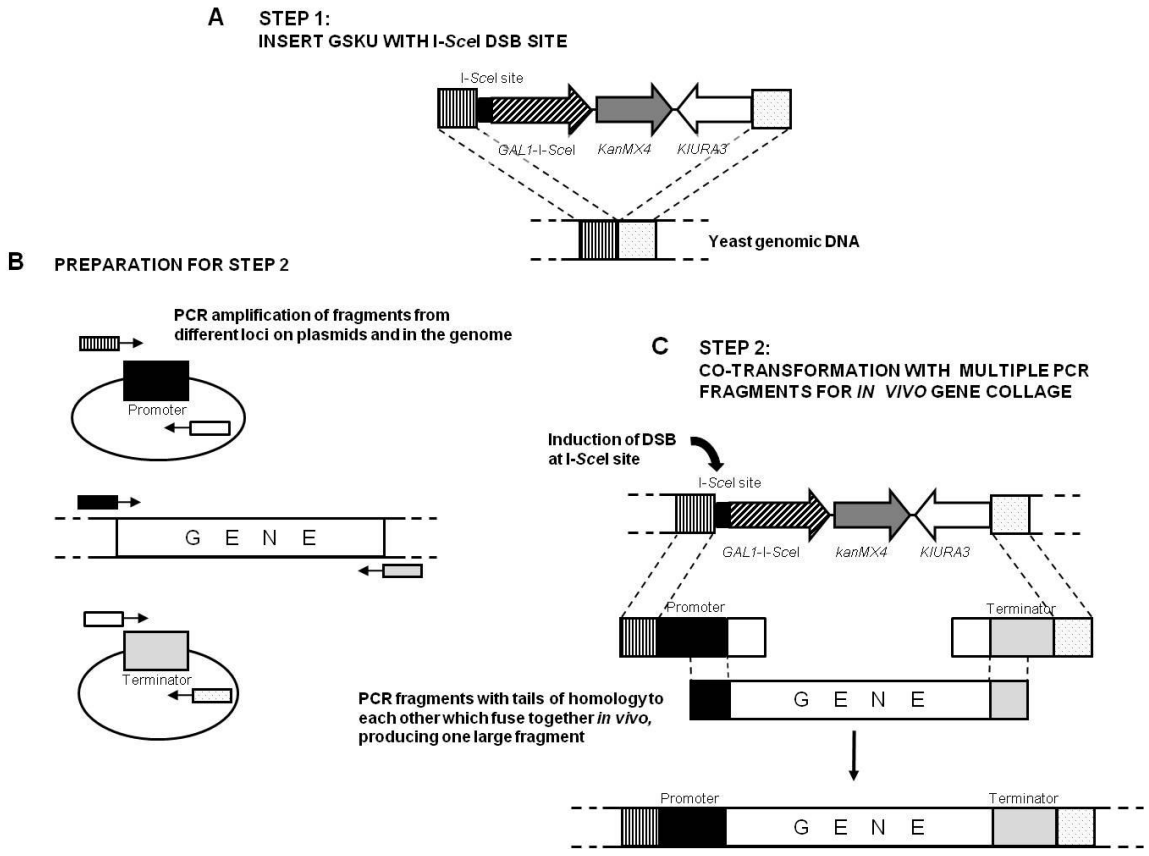

Figure 3.6 Mechanism of in vivo gene collage by the *delitto perfetto* approach. (**a**) In step one, the GSKU CORE cassette is amplified through PCR, with the 18-bp I-*Sce*I break site included within the sequence of one primer. This product is inserted into the target locus. (**b**) Prior to replacement of the cassette, a preparation step to generate PCR fragments is performed. For this example, a gene of interest will be attached to a chosen promoter and terminator sequences and all components will be inserted at a chosen locus. (**c**) In step two, the multiple PCR fragments assemble together *in vivo* by recombination to form a large fragment, which then replaces the GSKU cassette as it integrates into its specific region of the chromosome.

3.6 NOTES

Note 3.7.1

This review focuses on the generation of engineered haploid strains of yeast. For the use of *delitto perfetto* in diploid cells, refer to [104] for a detailed explanation of modifications to the protocol.

Note 3.7.2

For media preparation, deionized water is used. All other uses of the term "water" in this chapter, however, refer to deionized water that was sterilized by filtration or autoclaving.

Note 3.7.3

Unless otherwise noted, all solid media is to be stored at 4°C. Exceptions include YPD liquid and agar, which we store at room temperature.

Note 3.7.4

PEG 4000 solution will be extremely viscous, so filter sterilizing can take up to one hour depending on the volume. Autoclaving is an alternative means of sterilization.

Note 3.7.5

Using 50-ml conical tubes for O/N growth is preferred to using 15-ml tubes, as the larger size allows for greater dispersion of the nutrients in the broth to each of the cells. Additionally, *S. cerevisiae* is an aerobic species, so lids should not be capped tightly but instead loosely cover the tube and secured with tape.

Note 3.7.6

When inserting CORE, if growth on SC-Ura media is not observed after 3 days (*See* **Note 3.7.8** below), the transformation can be performed by plating onto YPD and incubating at 30°C O/N followed by replica-plating to G418- or Hygro-containing media, depending on the CORE used, and incubating at 30°C for 2-3 days until large colonies appear. This would then be followed by replica-plating to SC-Ura and incubating at 30°C O/N.

Note 3.7.7

As little as 5 µl of the cell resuspension solution can be used per reaction; however, this is not optimal when sequencing is necessary and a volume of 10 µl is suggested here.

Note 3.7.8

The following is specific to using CORE-UK, CORE-UH, GSKU, and GSHU, as this does not apply to the other cassettes. When *KlURA3* is inserted in the same orientation as the targeted gene, interference from that gene's promoter during transcription may lead to delayed growth on SC-Ura in the first step of *delitto perfetto* (*See* **Note 3.7.6** above) and may increase the number of background cells on 5-FOA in the second step. Depending on insertion orientation of the CORE, a second round of replica-plating to 5-FOA may be needed. Therefore, it is optimal to insert the cassette in such a way that *KlURA3* is oriented opposite to the gene being targeted.

Note 3.7.9

Upon successful colony PCR of transformants containing the newly-introduced CORE

sequence, sequencing analysis is not required. The resulting antibiotic resistance and Ura^+ phenotype of the strain in addition to the results of the colony PCR are sufficient to provide evidence for successful incorporation of the CORE into the targeted site. Sequencing is, however, necessary to verify the correct insertion of the desired mutation(s). Since the oligos used are non-purified, the expected additional mutations are in the range of 10-20%. Therefore, it is always better to obtain 3-5 clones for sequencing.

Note 3.7.10

YPLac is used to provide a neutral carbon source for the cells prior to addition of galactose. However, cells grow much slower in this medium. It is, therefore, optimal to inoculate cells into YPLac at least 18-20 h prior to the transformation.

Note 3.7.11

Addition of galactose activates the inducible *GAL1* promoter which regulates the I-*Sce*I gene. Experience has shown that longer induction (5-6 h) produces greater efficiency.

3.7 ACKNOWLEDGEMENTS

We thank the members of our lab for their contributions to the editing and revision of this work, notably Rekha Pai, Patrick Ruff, and Ying Shen. We also thank Lee Katz for assistance in proofreading and revision. This work was funded in part by the Georgia Cancer Coalition grant R9028 and the NIH R21EB9228.

CHAPTER 4

DSB-DRIVEN GENETIC MODIFICATION IN HUMAN CELLS USING DNA OLIGOS

The procedure and results detailed in Chapter 5 are modified from work published in Mutation Research [128]:

Shen, Y., Nandi, P., Taylor, M.B., **Stuckey, S.**, Bhadsavle, H.P., Weiss, B., and Storici, F. (2011) RNA-driven genetic changes from bacteria to human cells. *Mutat Res.* Online: 1-8.

4.1 ABSTRACT

We developed two systems for testing recombination in human cells using synthetic DNA oligos to repair an endonuclease-induced break. In one system recombination is tested at a target green fluorescent protein (*GFP*) locus on a plasmid, and in the other system the target locus has been stably-integrated into the genome of human embryonic kidney (HEK-293) cells. The break site, positioned within the coding sequence of a reporter gene, can be targeted for cleavage following expression of the nuclease. Following generation of the break, recombination using a single-stranded oligo can fix the break and restore the sequence of the gene, leading to a detectable phenotype. Here,

we demonstrate the use of these systems to efficiently repair a lesion at a plasmid and a chromosomal locus in human cells.

4.2 INTRODUCTION

Many diseases exist due to underlying genetic abnormalities. These can include multifactorial diseases, such as diabetes and certain types of cancer, as well as monogenetic disorders, such as cystic fibrosis and retinitis pigmentosa, otherwise known as tunnel vision [7-10,129,130]. While certain types of therapy could be used to mediate symptoms of some genetic diseases, a more permanent strategy to correct the defect would provide the most optimal treatment strategy. With this idea in mind, gene targeting was developed as a strategy to intentionally and precisely target a chromosomal locus for replacement with a similar sequence. Thus, through replacement with an allele carrying the correct genetic sequence a disease could be "cured" at the chromosomal level [36]. Gene targeting relies on HR, which exists naturally to allow for exchange between similar genetic sequences through such processes as meiotic recombination or to repair chromosomal breaks [36]. However, HR is highly inefficient in many organisms including humans [60]. In order to promote recombination, scientists determined that a targeted DNA DSB could efficiently be used as it is a naturally occurring substrate for the HR machinery [131,132]. Generation of a DSB has been shown to stimulate recombination in many organisms, including human cells, more than 1,000-fold [49,51,60], thus helping to establish gene targeting as an effective strategy for potentially combating genetic disorders.

While a DSB can efficiently trigger the HR machinery, a great deal of variability in results has been noted between different labs due to numerous factors. For example, varying levels of recombination events have been observed based on the target locus, repair substrate, DSB-producing nuclease used, and the specificity of the enzyme to its recognition sequence [59-61,133]. For our studies, we aimed to develop a system which would allow us to directly compare different nucleases for their recombination potential at the same target plasmid or chromosomal locus in human cells. Previously, synthetic DNA oligos were shown to efficiently repair a targeted DSB in yeast cells both at and distant from the site of a DSB [53,103] and at the site of a break in human cells [98,134,135]. Additionally, a plasmid construct and subsequent stable, monoclonal cell line were developed which could be used to observe recombination at a disrupted *GFP* locus [60]. By merging these concepts, we were able to develop an assay for detecting recombination after generating a site-directed break within the *GFP* gene followed by repair using synthetic DNA oligos.

We first focused on generating a DSB using the I-*Sce*I protein, previously shown to stimulate recombination in many organisms, including human cells, up to 1,000-fold or more [51,53,54,60,133]. In the plasmid assay, a plasmid containing the 18-bp I-*Sce*I recognition site within *GFP* was transfected into human cells along with a plasmid expressing I-*Sce*I and DNA oligos complementary to either side of the disruption (**Figure 4.1** left side). Following generation of the DSB, if the sequence from the oligos was used as a repair template then the sequence of the *GFP* gene would be restored. Similarly, a

chromosomal *GFP* locus could be targeted following expression of I-*Sce*I from the plasmid and co-transfection with the oligo(s) (**Figure 4.1** right side). Our results demonstrate that an I-*Sce*I DSB can efficiently be repaired at the *GFP* locus on a plasmid or on the chromosome in human cells using either of the single oligos.

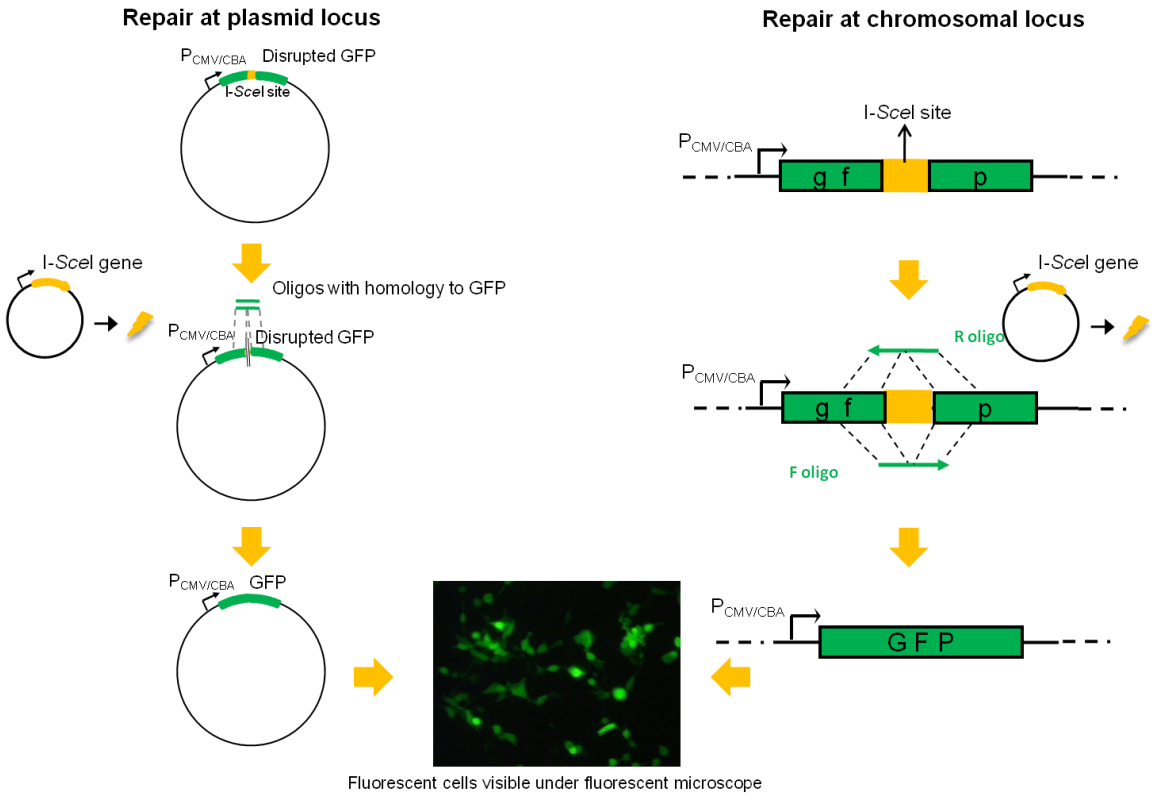

Figure 4.1 Scheme of repair at the target locus in human cells. In the system used, a disrupted *GFP* marker can be repaired after generation of a DSB within the marker followed by repair with oligos homologous to either side of the region flanking the disruption. For both the plasmid (left) and chromosomal (right) target locus, the 18-bp recognition sequence for I-*Sce*I is located within the disrupted gene. After I-*Sce*I is expressed, it generates a break which can be repaired using oligos that can knock out the disruption and restore the sequence of the gene. Cells are non-fluorescent prior to repair of *GFP*, while restoration of the gene produces the GFP^+ phenotype.

4.3 MATERIALS AND METHODS

4.2.1 Human cell lines and culture

HEK-293 cells were grown in Dulbecco's modified Eagle's medium (DMEM), supplemented with 10% heat-inactivated fetal bovine serum (FBS) and 1x penicillin/streptomycin, at 37ºC in a 5% CO_2 incubator. 658D cells were generated previously through random integration of a disrupted green fluorescent protein (*GFP*) marker into the genome [60].Within the 35-bp disruption at the *GFP* locus are two stop codons and the 18-bp recognition sequence for the I-*Sce*I endonuclease (**Figure 4.1**).

4.2.2 Plasmids

Plasmid pA658 contains a disrupted *GFP* marker under the CMV/CBA hybrid promoter and was stably integrated into HEK-293 cells to generate the 658D cell line [60]. The 35-bp disruption in the middle of *GFP* contains two stop codons and the 18-bp cognate recognition sequence for the I-*Sce*I endonuclease (**Figure 4.1**). Plasmid p67 contains the I-*Sce*I endonuclease gene under the CMV/CBA hybrid promoter [60].

4.2.3 Oligos

Synthetic DNA oligos were designed as 80mers. GFP80.F (F, 5' –GCGCACCATCTTC TTCAAGGACGACGGCAACTACAAGACGCGCGCCGAGGTGAAGTTCGAGGGC GACACCCTGGTGAACC) contains the same sequence as the coding strand of *GFP* and has 40-bases of homology to the region 5' of the break site in *GFP* and 39-bases of homology to the region 3' of the break site in the gene. In the middle of the oligo is a

nucleotide which creates a silent substitution in *GFP*, generating the *Bss*HII restriction site. GFP80.R (R, 5' – GGTTCACCAGGGTGTCGCCCTCGAACTTCACCTCGGCGC GCGTCTTGTAGTTGCCGTCGTCCTTGAAGAAGATGGTGCGC) is the reverse complement sequence of GFP80.F.

4.2.4 Transfection

Cells were seeded with approx. 50,000 cells per well in a 24-well plate 24 h prior to transfection. In transfection experiments, plasmid DNA was used in the amount of 0.5 µg of DNA per well while oligo DNA was 1.5 µg. Prior to transfection, DNA was diluted in DMEM without supplements and vortexed with polyethylenimine (PEI) transfection reagent at 5 µg of PEI per well. Mixtures were incubated at room temperature for 30 min prior to being added drop-wise into the well.

4.2.5 Detection and quantification of recombinants

Cells were observed beginning on day two following transfection using a fluorescent microscope. Frequencies of GFP^+ cells were obtained by flow cytometric analysis using the BD LSR II flow cytometer (BD Biosciences, Sparks, MD). 100,000 cells were counted for each sample.

4.2.6 Statistical analysis

All graphs were created using the GraphPad Prism 5 software (GraphPad Software, Inc.). Data are plotted as median values with the range displayed as bars. Statistical significance was determined by using two-tailed *t*-tests (Mann-Whitney U).

4.4 RESULTS

4.3.1 DNA oligos can efficiently repair a DSB in human cells

Following transfection to repair the targeted plasmid *GFP* locus, for all negative controls ≤1 green fluorescent cell per 100,000 cells read was detected. However, recombination using the complementary pair was less efficient than with either of the single oligos (p≈0.0294) (**Figure 4.2**). Thus, for subsequent experiments we used only combinations including either of the oligos as a repair template, but not both simultaneously.

Compared to when only the repair template and either oligo were provided, generation of the DSB at the target plasmid locus followed by repair with either F or R led to a significant increase in GFP^+ recombinants (p≈0.0066) (**Figure 4.2** and **Figure 4.3**). Additionally, a strand bias in favor of the F oligo was observed (p≈0.0004) (**Figure 4.3**).

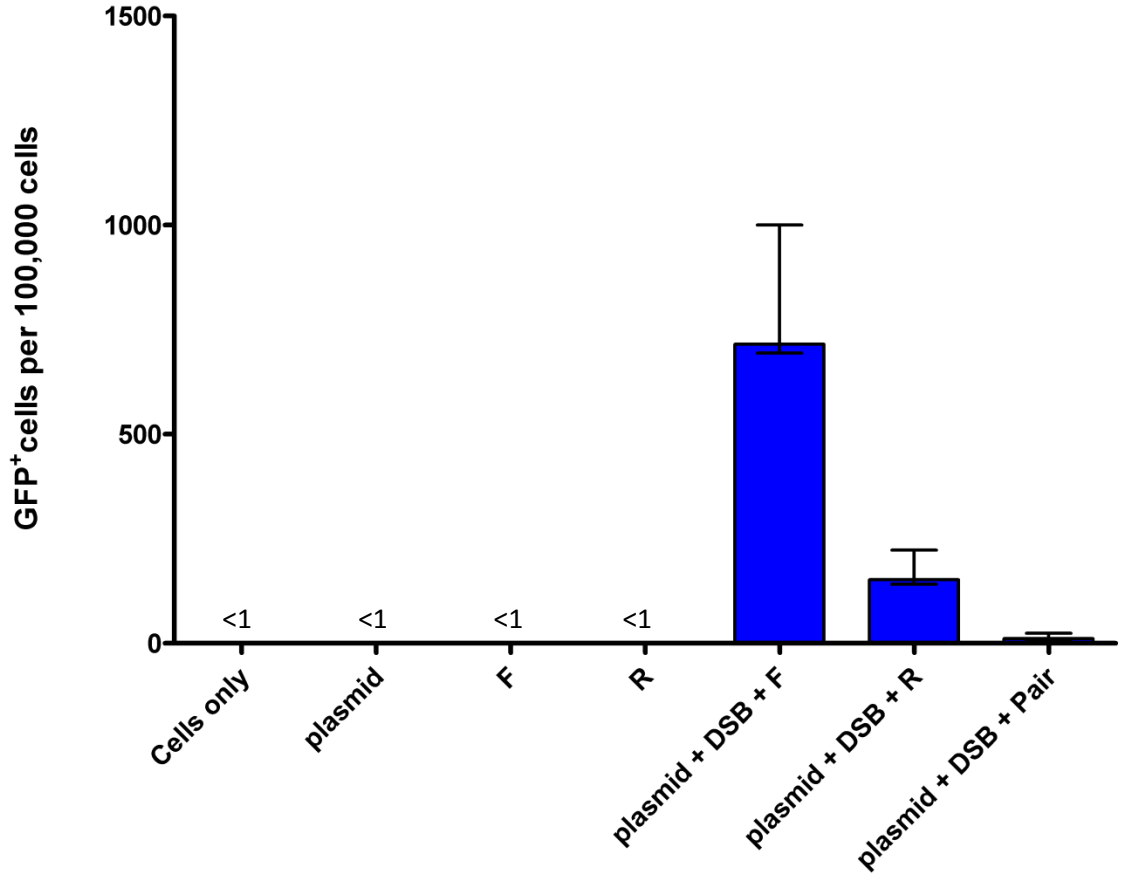

Figure 4.2 DSB-driven recombination at a target plasmid locus. Levels of GFP^+ cells following expression of I-*Sce*I oligos to repair the break in *GFP*. Cells only contains cells transfected with no additional DNA. The plasmid used is shown in the scheme depicted in **Figure 4.1**. F = oligo complementary to the anti-sense strand of the gene. R = oligo complementary to the sense strand. Pair = F + R. Plasmid = plasmid containing disrupted *GFP* marker. DSB = plasmid expressing I-*Sce*I. All data are presented as the median with range.

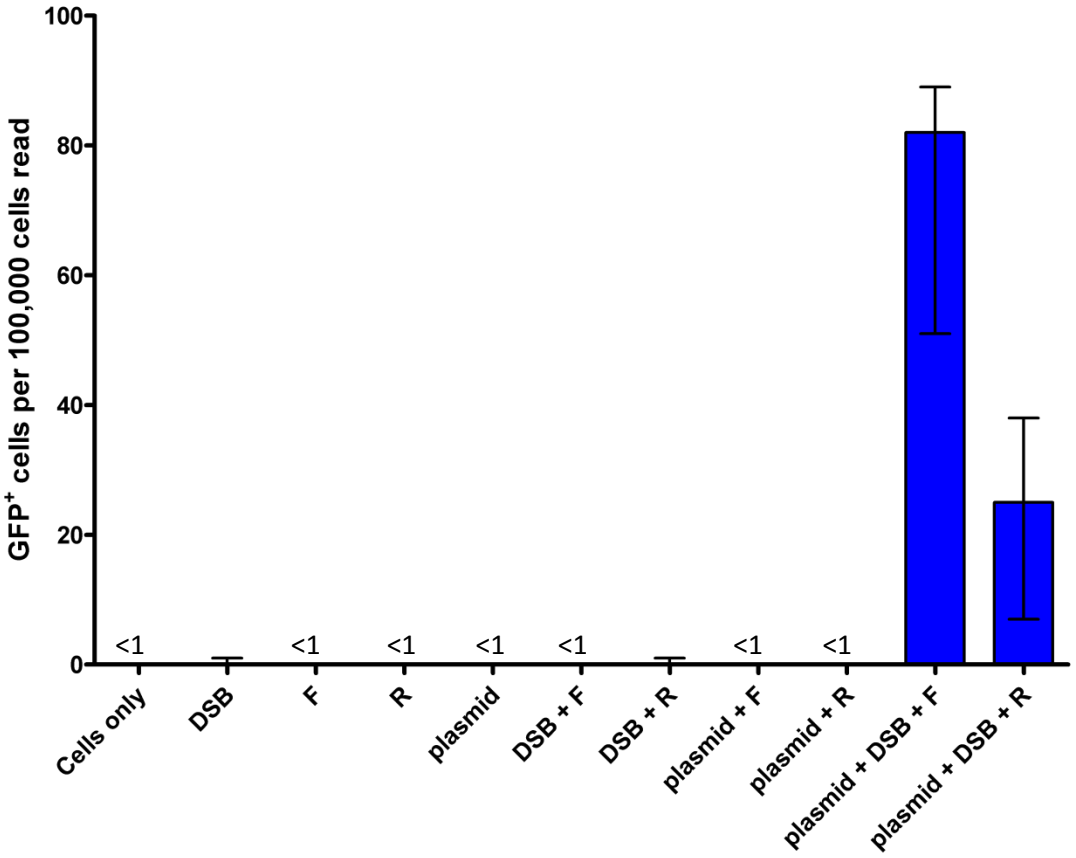

Figure 4.3 A DSB stimulates recombination on the plasmid. Levels of GFP^+ cells following expression of I-*Sce*I and transfection with oligos to repair the break in *GFP*. Cells only contains cells transfected with no additional DNA. F = oligo complementary to the anti-sense strand of the gene. R = oligo complementary to the sense strand. Plasmid = plasmid containing disrupted *GFP* marker. DSB = plasmid expressing I-*Sce*I. All data are presented as the median with range.

Similar to the results at the plasmid *GFP* locus, an increase in the frequency of GFP^+ cells was observed using both the single F and R oligo at the chromosomal *GFP* locus (**Figure 4.4**). For all negative controls ≤1 green fluorescent cell per 100,000 cells read was detected. Compared to when no DSB was induced, generation of a DSB at the target chromosomal locus followed by repair with either F or R led to a significant increase in GFP^+ recombinants ($p\approx0.0018$ and $p\approx0.0002$, respectively). Additionally, a strand bias in favor of the F oligo was observed ($p<0.0001$).

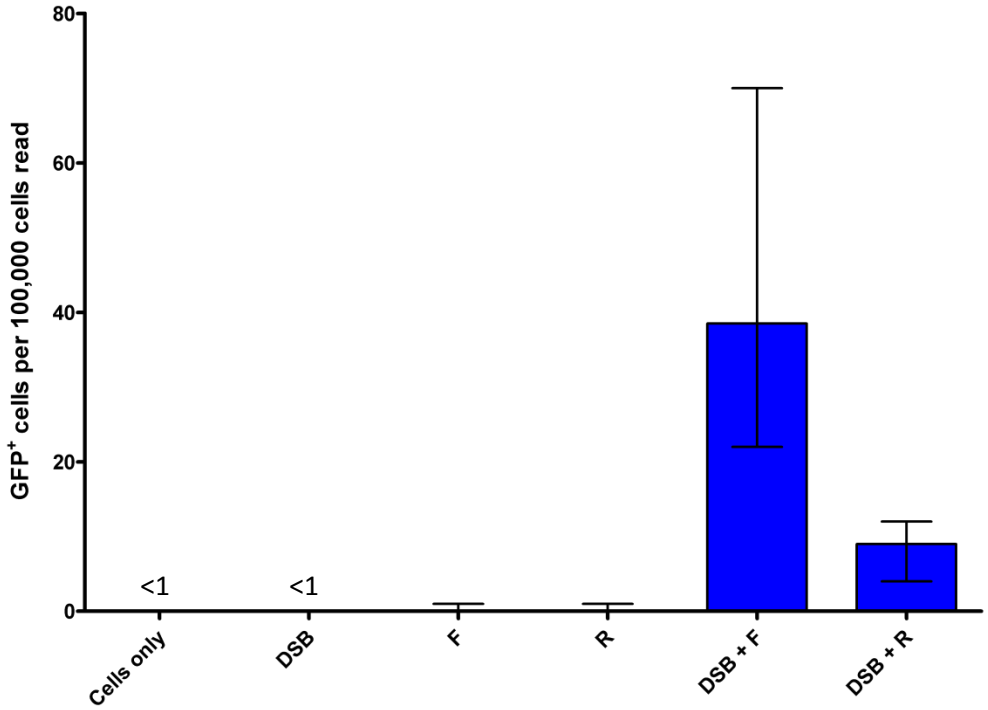

Figure 4.4 A DSB stimulates recombination on the chromosome. Levels of GFP^+ cells following expression of I-*Sce*I and transfection with oligos to repair the break in *GFP*. Cells only contains cells transfected with no additional DNA. F = oligo complementary to the anti-sense strand of the gene. R = oligo complementary to the sense strand. Plasmid = plasmid containing disrupted *GFP* marker. DSB = plasmid expressing I-*Sce*I. All data are presented as the median with range.

4.5 DISCUSSION

We demonstrated here that the system we use to stimulate recombination at the site of a DSB in human cells using synthetic DNA oligos works both on plasmid and on chromosomal DNA. For samples in which the repair template and target locus were provided and no DSB was generated, ≤1 fluorescent cell per 100,000 cells was detected (**Figures 4.3** and **4.4**). However, when the plasmid expressing I-*Sce*I was also transfected into the cells a significant increase in GFP^+ cells was observed. Each of the single oligos was able to repair the break at the *GFP* locus, both on the plasmid (**Figure 4.3**) and on the chromosome (**Figure 4.4**), as evidenced by an increase in GFP^+ cells following generation of the DSB and co-transfection with either of the repair templates. The length of homology of these oligos (40 bases upstream of the disruption and 39 bases downstream with an additional base creating a silent mismatch) was sufficient to restore the sequence of the marker. Additionally, a strand bias in favor of one oligo (F, which represents the sense strand of the gene) was observed at the locus in both systems. However, the wider implications of this cannot easily be attributed as previous findings have observed a bias in favor of either the sense or anti-sense strand oligo which varies depending on the target locus [98,134,135].

Differently from findings in yeast that the complementary pair of oligos is generally most favorable for repair of an I-*Sce*I DSB at a target locus [53], we observed that the complementary pair was least efficient compared to either of the single oligos in human cells (**Figure 4.2**). This is consistent with previous findings in mammalian cell lines that

found shorter single-stranded DNA oligos (~35-80 nt) were more favorable templates for repair of various episomal or chromosomal loci than a similar double-stranded construct [98,135]. While the exact reason for this is unclear, it was proposed that perhaps the double-stranded oligos first must unfold prior to hybridization with a target locus [98]. Additionally, it is possible that a blunt-ended double-strand sequence, such as that used for our assay, is recognized as DNA damage, triggering a checkpoint response and degradation.

Ultimately, our findings illustrate the capabilities of single-stranded synthetic DNA oligos to repair a target locus in human cells. These oligos are of sufficient length to restore the sequence of the gene, and the transfection method we use (transfection with PEI) is able to allow the cells to uptake the DNA. Finally, the system we use allows for rapid and easy detection of recombination events as fluorescent cells can be readily identified under a fluorescent microscope and quantified through FACS.

CHAPTER 5

GENETIC MODIFICATION STIMULATED BY THE INDUCTION OF A SITE-SPECIFIC BREAK DISTANT FROM THE LOCUS OF CORRECTION IN HAPLOID AND DIPLOID YEAST *SACCHAROMYCES CEREVISIAE*

The procedures detailed in Chapter 5 are modified from a publication of the same name in the Methods in Molecular Biology book series, Gene Correction edition [136]:

Stuckey, S. and Storici, F. (2013). School of Biology, Georgia Institute of Technology, Atlanta, GA. *Meth Mol Biol*

5.1 ABSTRACT

Generation of a site-specific break at a genomic locus to stimulate homologous recombination (HR) is used in many organisms to efficiently target genes for various types of genetic modification. Additionally, a site-specific chromosomal break can be used to trigger HR at genomic regions distant from the break, thereby largely expanding the region available for introducing desired mutations. In contrast to the former approach, the latter presents an alternative way in which genes can be efficiently modified also when it is not possible or desirable to introduce a break in the vicinity of the targeting locus. This type of *in vivo* site-directed mutagenesis distant from a break can be accomplished in the yeast model organism *Saccharomyces cerevisiae* because the

generation of a double-strand break (DSB) in yeast chromosomal DNA activates HR at long regions upstream and downstream from the break site. Here we provide a protocol for efficiently altering a yeast chromosomal locus following the induction of a DSB several kilobase pairs distant from the site of gene correction. The techniques described can be used in both diploid and haploid yeast strains, and we provide examples of the gene correction assays.

5.2 INTRODUCTION

Gene targeting is a genetic technique that uses homologous recombination (HR) to modify an endogenous DNA sequence. Using this technique, various genetic manipulations are possible, including gene knock-out for functional analysis, gene knock-in to provide selectable markers or for heterologous gene insertion, generation of specific point mutations to alter the function of a desired gene or regulatory sequence, introduction of particular polymorphisms, or correction of genetic defects [137]. Gene targeting has been used extensively in many cellular systems, from bacteria to human cells [27,29,32,34,49,60,63,138]. Despite the wide applications of the gene targeting technique, the generally low efficiency of HR in nearly all organisms has always been a challenge for researchers [21,139,140]. One of the most effective ways to increase gene targeting in chromosomal DNA is *via* the induction of a site-specific DNA double-strand break (DSB) near the genomic site to be changed [49,51,131]. Indeed, a DSB can stimulate HR up to several thousand-fold over natural levels [60,141,142]. What increases the efficiency of gene targeting by HR is not the formation of the break *per se*

in the genomic DNA. It is the generation of regions of single-stranded DNA, following processing of the break ends and strand resection in a 5' to 3' direction in preparation for break repair, that makes the DNA at the broken locus highly recombinogenic [52]. If a DSB occurs in DNA but formation of single-stranded DNA is impeded, gene targeting by HR is not stimulated [52]. During the process of DSB repair, the gap of single-stranded DNA that is generated can be filled in using homologous templates [52,143]. Therefore, it is possible to increase the efficiency of gene correction in chromosomal DNA also using targeting molecules that do not physically repair the break but that have complementarity with the single-stranded regions to either side of the break [52]. This approach of gene modification at sites distant from the DSB works well in the yeast *Saccharomyces cerevisiae,* in which long regions of single-stranded DNA can be generated following induction of a DSB [52,144]. Major advantages of this strategy are i) a large window for mutagenesis expanded up to 20 kb from each side of the DSB, ii) the possibility to perform gene correction in essential genes, and iii) the applicability of the approach to diploid in addition to haploid cells [51-53,59,104,145].

When applying gene targeting distant from sites of induced breaks in cellular systems different from the yeast *S. cerevisiae*, it should be taken into consideration that the efficiency of genetic modification at a locus distant from a break is limited by the length of the single-stranded DNA regions surrounding the DSB, which varies between organisms. Following generation of a DSB, 5' to 3' resection on both sides of the break leads to long segments of 3' overhangs [45]. In yeast these segments can stretch for tens

of kilobases [144], providing a large window of opportunity for modification at distant loci. Studies in mammalian cells, though, indicate the lengths of resection are much shorter, less than 2 kb [146]. Thus, while generation of a distant DSB can improve gene targeting applications by enlarging the window of opportunity for creating modifications, factors such as resection lengths can also limit this window.

Though highly efficient in yeast cells, gene targeting at regions distant from the site of DNA breakage could potentially be adapted to other cellular systems. The opportunity to exploit a wider window of mutagenesis, even a few hundred nucleotides longer, from the site of breakage may already provide a significant advantage. Generation of a DSB at a genomic locus in which the desired genetic modification will be made may not always be feasible, and in some situations it is favorable to generate a break distant from the position to be modified. For example, designed nucleases, such as zinc-finger nucleases (ZFNs) or transcription activator-like effector nucleases (TALENs), are created to recognize and cleave specific genomic loci [58,61,64,147]. In order to design the most effective designer nucleases, however, targeted genomic regions must contain highly-specific nucleotide sequences recognizable by each of the nucleases and appropriately spaced relative to each other such that dimerization of the nucleases can occur. Thus, the sequence constraints of these enzymes may require their recognition sites to be located at some distance from the genomic locus to be modified. Additionally, recent advancements with systems naturally involved in the adaptive immune systems of bacteria and archaea have demonstrated the capability of RNA-driven targeted gene correction [69,70]. Clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR associated (Cas) systems, like ZFNs and TALENs, recognize highly-specific nucleotide sequences

in the genome for targeted generation of a site-directed break [69,70]; however, as with ZFNs and TALENs, the genetic locus to be targeted for gene correction may exist at some chromosomal position distant from the most favorable sequence for recognition by these CRISPR elements. If the gene targeting sites are distant from the most optimal recognition sequences for a break to be generated but still in regions that can become single-stranded after DSB induction, the gene correction efficiency will be much higher than without DSB induction.

Here we provide a protocol for modification of a locus distant from a DSB in the yeast *S. cerevisiae* using synthetic DNA oligos as gene-correction templates. We explain in a step-wise fashion procedures in which a cassette containing an inducible I-*Sce*I endonuclease gene and its 18-bp recognition site can be integrated into the genome of haploid or diploid yeast strains and used to stimulate genetic modification of regions several kb away from the initial chromosomal break point.

5.3 MATERIALS

5.3.1 Amplification of CORE-I-*Sce*I

1. Two CORE-I-*Sce*I plasmids (pGSKU, pGSHU) are available (*See* [120] for further plasmid information).
2. DNA primers, desalted and non-purified: 50 μM stock. Stored at -20°C.

3. Ex Taq DNA polymerase, 10x buffer, 2.5 mM dNTPs (Clontech Laboratories, Inc., Mountain View, CA, USA).

5.3.2 Gel Electrophoresis

1. Agarose.
2. 10x Tris/Borate/EDTA (TBE) running buffer.
3. Prestained molecular weight marker.
4. Loading dye.

5.3.3 PCR Product Concentration

1. 95% Ethanol (EtOH).
2. 70% EtOH.
3. 3 M Sodium acetate (NaOAc; pH 5.2) solution in water: Filter-sterilized. Stored at room temperature (*See* **Note 5.6.1**).

5.3.4 Transformation Reagents and Media

1. YPD (per 1 l): 10 g yeast extract, 20 g bacto peptone, 20 g dextrose. Autoclaved for 45 min at 121ºC. 20 g agar is added for solid medium prior to autoclaving (*See* **Notes 5.6.2 & 5.6.3**).
2. YPLac liquid (per 1 l): 12 g sodium hydroxide (NaOH), 10 g yeast extract, 20 g bacto peptone, 27 ml lactic acid, pH adjusted to 5.5 with lactic acid. Autoclaved for 45 min at 121ºC.

3. 20% Highly pure galactose stock solution in water: Filter-sterilized. Stored at room temperature.
4. 1 M Lithium acetate (LiOAc) stock solution in water: Filter-sterilized. Stored at room temperature.
5. 10x TE stock solution: 100 mM Tris (pH 7.5), 10 mM ethylenediaminetetraacetic acid (EDTA) (pH 7.5) in water. Filter-sterilized. Stored at room temperature.
6. 50% Polyethylene glycol 4000 (PEG 4000) stock solution in water: Stored at room temperature (*See* **Note 4**).
7. Working solutions: Solution 1 (0.1 M LiOAc, TE 1x, pH 7.5 in water) and solution 2 (0.1 M LiOAc, TE 1x, pH 7.5 in 50% PEG 4000).
8. 100 µg/ml Salmon sperm DNA (SSD) in water: Stored at -20°C.
9. Synthetic complete medium lacking uracil (SC-Ura) solid medium.
10. Synthetic dextrose complete (SD Complete) solid medium.
11. Glass beads approximately 5 mm diameter.
12. 5-Fluoroorotic acid (5-FOA; per 1 l): Solution of 5-FOA is prepared by dissolving 1 g 5-FOA in 300 ml of water prior to filtration. 700 ml SD Complete agar medium is autoclaved for 45 min at 121ºC, then cooled to 55-60°C, and the filter-sterilized solution of 5-FOA is then mixed well with medium prior to pouring.
13. G418 (per 1 l): YPD agar medium is autoclaved for 45 min at 121ºC ,then cooled to 55-60°C, and G418 solution is then mixed well with medium

prior to pouring. Stock solution is prepared in water in 50 mg/ml filter-sterilized aliquots and stored at 4°C.

14. Hygromycin B (hygro; per 1 l): YPD agar medium is autoclaved for 45 min at 121ºC, then cooled to 55-60°C, and hygromycin B solution is then mixed well with medium prior to pouring.
15. YPG (per 1 l): 10 g yeast extract, 20 g bacto peptone, 30 ml glycerol, 20 g agar. Autoclaved for 45 min at 121ºC.
16. Sterile velveteens.

5.3.5 Genotypic Testing of Transformants

1. 2000 U/ml lyticase in water: Stored in 1 ml aliquots at -20°C.
2. Taq DNA polymerase, 10x buffer, 10 mM dNTPs.

5.3.6 Design of DNA Oligos for Modification of Distant Loci and Repair of DSB

1. 50 µM DNA oligo stock solution: 40-100mers, desalted and non-purified. Store at -20°C.

5.4 METHODS

Here we describe a method for generating an I-*Sce*I site-specific DSB in yeast chromosomal DNA which will be used to stimulate gene targeting at a region distant from the DSB site both in haploid and diploid yeast cells. The initial steps are common between the two. Briefly, a product will be amplified from plasmids available containing

a CORE-I-*Sce*I cassette. The cassettes contain COunterselectable and REporter markers for selection, hence the name. Two of these plasmids (pGSKU and pGSHU) contain the I-*Sce*I site-specific endonuclease gene, regulated by a galactose-inducible promoter, *GAL1*.The cassettes from these plasmids, hereafter referred to as the CORE-I-*Sce*I cassettes, will be amplified through polymerase chain reaction (PCR) and then transformed into the strain of interest. Additionally, during PCR the 18-bp cognate recognition sequence for I-*Sce*I is introduced at one end of the cassette by one of the primers. This site will later be used for the DSB-driven repair assay as the expressed I-*Sce*I protein will recognize it to generate the site-directed break. We further provide examples of repair assays in both haploid and diploid yeast strains. Additional material on the available CORE-I-*Sce*I cassettes has been previously published elsewhere [52,102,120].

5.4.1 Amplification of CORE-I-*Sce*I from Plasmid

1. DNA primers will first be used to amplify the CORE-I-*Sce*I cassettes (GSKU or GSHU) from the chosen plasmid. These primers range from 70 to 88 bases in length with an overlap of at least 50 bp with the genomic targeting region and an overlap of 20 bp to the CORE-I-*Sce*I cassette sequence. Additionally, the 18-bp recognition sequence for the I-*Sce*I endonuclease is included on one of the two primers (*See* [120] for further information on amplifying these cassettes).

2. PCR conditions: Amplification of the CORE-I-*Sce*I cassette from circular plasmid (about 50 ng) using 50 pmol of each primer is performed with high yield in a final volume of 40 µl using Ex Taq DNA polymerase with a 2 min cycle at 94°C; 32 cycles of 30 s at 94°C, 30 s at 57°C, and 5 min at 72°C; a final extension time of 7 min at 72°C; and samples are held at 4°C. Stock concentrations of 10 mM dNTPs are used for this reaction. An extension time of 1 min/kb is assumed for this reaction (*See* **Note 5.6.5**).
3. Following PCR, the samples are ready for gel electrophoresis and PCR product concentration.

5.4.2 Gel Electrophoresis

1. We use a dilution of 0.5x TBE running buffer, which is obtained from 10x TBE by mixing 50 ml of 10x TBE buffer with 950 ml deionized water prior to use.
2. A small aliquot (about 2 µl) of PCR product is run on a 0.8% agarose gel to observe the anticipated band.

5.4.3 PCR Product Concentration

1. The product of six reactions of PCR are combined for precipitation with a 2.5x volume of 95% EtOH and 0.1x volume of 3 M NaOAc (pH 5.2) in a microcentrifuge tube. Centrifugation is carried out at maximum speed for 10 min. A small pellet should be visible on the bottom of the tube.

2. The supernatant is discarded, and the pellet is washed with 100 µl of 70% EtOH, being careful not to detach the pellet. If the pellet is detached, it is necessary to spin again for 5 min and then discard the supernatant. Then, as much as possible of the EtOH is removed without detaching the pellet.
3. The pellet is then spun until dry in a speed vacuum device and then resuspended in 50 µl of water. Five to 10 µl are used for each transformation.

5.4.4 Transformation to Insert the CORE-I-*Sce*I Cassette

The following protocol is used to insert the CORE-I-*Sce*I PCR product into the chosen strain. This procedure applies to both haploid and diploid strains. As the CORE-I-*Sce*I cassettes, GSKU and GSHU, each contain multiple markers (the *URA3* gene from *Kluveromyces lactis* and resistance genes to the antibiotics G418 or Hygromycin B, respectively), the strain should have a non-functional *URA3* gene and should be G418- and hygromycin B-sensitive to ensure proper integration at the desired locus. This procedure has been modified from the lithium acetate protocol described by Wach *et al*. [113]. A more detailed explanation of the involved steps has been previously published by our group in the *Methods in Molecular Biology* series [120].

1. Inoculate 5 ml of YPD liquid medium with chosen strain and shake at 30°C overnight (O/N) (*See* **Note 5.6.6**).
2. When cells are in logarithmic growth (approximately 16 h later), inoculate 50 ml of YPD liquid medium with 1.5 ml of the O/N culture in a 250-ml

glass flask and shake vigorously at 30°C for 3 to 4 h.

3. Solutions 1 and 2 are prepared immediately prior to transformation.
4. Transfer culture to a 50-ml conical tube and spin at 1562 x *g* for 2 min.
5. Remove the supernatant and wash cells with 50 ml of sterile water and spin as stated previously.
6. Remove the supernatant and resuspend cells in 5 ml of solution 1 and spin as stated previously.
7. Remove supernatant and resuspend cells in 250 µl of solution 1. This amount of cells is sufficient for up to approximately 10 transformations.
8. Aliquot 50 µl of the cell suspension in microcentrifuge tubes and add 5-10 µl of concentrated CORE-I-*Sce*I PCR product and 5 µl of SSD (heat-denatured for 5 min at 100°C prior to use and immediately kept on ice).
9. Add 300 µl of solution 2 to each transformation tube. Mix briefly by vortexing.
10. Incubate transformation tubes at 30°C for 30 min with shaking.
11. Heat shock at 42°C for 15 min.
12. Collect cells by centrifugation at 2236 x *g* for 4 min.
13. Remove the supernatant and resuspend cells well in 100 µl of sterile water.
14. Plate all cells from each transformation tube on one SC-Ura plate using approximately 15 sterile glass beads and incubate at 30°C for 2-3 days until colonies appear (*See* **Note 5.6.7**).

15. Using a sterile velveteen, replica-plate from SC-Ura to G418- or hygro-containing medium (depending on the CORE used) and incubate at 30°C O/N.
16. Once transformants are observed (typically 5 to 30 colonies per plate), streak for single colony isolates on YPD solid medium. Incubate at 30°C for 2 to 3 days.
17. Make patches of the single colonies onto new YPD solid medium, along with the original strain for phenotypic comparison, and incubate at 30°C O/N.
18. Replica-plate the grown patches to YPD; SC-Ura; G418; hygro; yeast extract-peptone-glycerol (YPG) solid medium , which selects against cells with defective mitochondrial DNA (mtDNA); and any other various selective media depending on the background of your strain, and incubate at 30°C O/N.
19. Following observation of correct phenotype, the samples are ready for genotypic testing.

5.4.5 Colony PCR of Transformants

1. Resuspend cells (approximately 1 mm^3) in 50 µl water containing 1 U of lyticase. Incubate at room temperature for 10 min, then heat at 100°C for 5 min. Samples can then be left at room temperature.

2. PCR conditions: Colony PCR of the transformant patches showing the expected phenotypes using 10 µl of the cell resuspension mixture is carried out with 50 pmol of each primer and with an expected amplified product size ranging between 300 bp and 1 kb. Stock concentrations of 10 mM dNTPs are used for this reaction. PCR is performed in a final volume of 50 µl using Taq DNA polymerase with a 2 min cycle at 95°C; 32 cycles of 30 s at 95°C, 30 s at 55°C, and 1 min at 72°C; a final extension time of 7 min at 72°C; and samples are held at 4°C. An extension time of 1 min/kb is assumed for this reaction.
3. Following PCR, samples are run on a 1% agarose gel (*See* **Section 5.4.2**) for observation of PCR products of the anticipated size.
4. Strains are now ready to be sequenced to verify the sequences of the inserted I-*Sce*I gene and site.

5.4.6 Design of DNA Oligos for Modification of a Chromosomal Locus

To modify a locus distant from the DSB with oligos, a single oligo or a complementary pair can be used to generate the desired mutation. The following considerations should be made when determining which oligo(s) to use. First, while oligos as short as 40 bases in length can be used to efficiently transform the strain, the efficiency of targeting is increased as the length of oligo increases (up to 80-90 nt) [104]. Next, following generation of a DSB, there will be resection of each strand of the DNA away from the break in the 5' to 3' direction. The single oligo complementary to the intact strand following resection will be much more efficient than the other oligo [52]; therefore, it is

suggested to determine which single oligo will be more favorable to use prior to performing the experiment. Finally, a pair of complementary oligos provides the most efficient repair template; however, in many cases the use of the single oligo complementary to the intact strand will be highly efficient at modifying the DNA and thus a complementary pair may not be necessary.

5.4.7 Generation of the DSB

In **Section 5.4.1** a protocol for amplifying the CORE-I-*Sce*I cassette is detailed. One of the primers introduces the 18-bp recognition site for the I-*Sce*I endonuclease. After generation of a strain containing the integrated cassette and site, the I-*Sce*I protein will be used to create a site-directed DSB at this chromosomal position. Expression of this gene is regulated by the *GAL1* promoter, which is induced following addition of galactose to the medium. In the next two sections, we provide examples of repair assays which can be conducted using the galactose-induced system in haploid and diploid yeast strains.

5.4.8 Repair Assay: Example at the *trp5* Locus

Here we provide an example of an assay for repair of a mutated, non-functional *trp5* marker which is located 10 kb downstream from the DSB site in a haploid strain with oligos that are used to repair the break and those which correct the sequence of *trp5* (*See* **Figure 5.1**). Following correction of this marker, cells present a Trp^+ phenotype which can be selected for on the solid medium lacking tryptophan. While targeted gene correction can be used at either a selectable or a non-selectable genetic locus, the

presentation of a selectable phenotype following correction enables more rapid screening of transformants. Further information on targeting at a non-selectable genetic locus is provided in *Section 4* section below (*See* **Note 5.6.8**).

1. Inoculate 50 ml of YPLac liquid medium with chosen strain into a 250-ml glass flask and shake vigorously at 30°C O/N (*See* **Note 5.6.9**).
2. Add 5 ml of galactose from a 20% stock solution into the O/N culture for a final 2% galactose solution and continue to shake at 30°C for 7 h (*See* **Note 5.6.10**).
3. Solutions 1 and 2 are prepared immediately prior to transformation.
4. Transfer culture to a 50-ml conical tube and spin at 1562 x *g* for 2 min.
5. Remove the supernatant and wash cells with 50 ml of sterile water and spin as stated previously.
6. Remove the supernatant and resuspend cells in 5 ml of solution 1 and spin as stated previously.
7. Remove supernatant and resuspend cells in 250-750 µl of solution 1. The pellet should be quite large. Depending on the dilution, this amount of cells is sufficient for up to approximately 30 transformations.
8. Add 300 µl of solution 2 to each transformation tube. Mix briefly by vortexing.
9. Incubate transformation tubes at 30°C for 30 min with shaking.
10. Heat shock at 42°C for 15 min.
11. Collect cells by centrifugation at 2236 x *g* for 4 min.

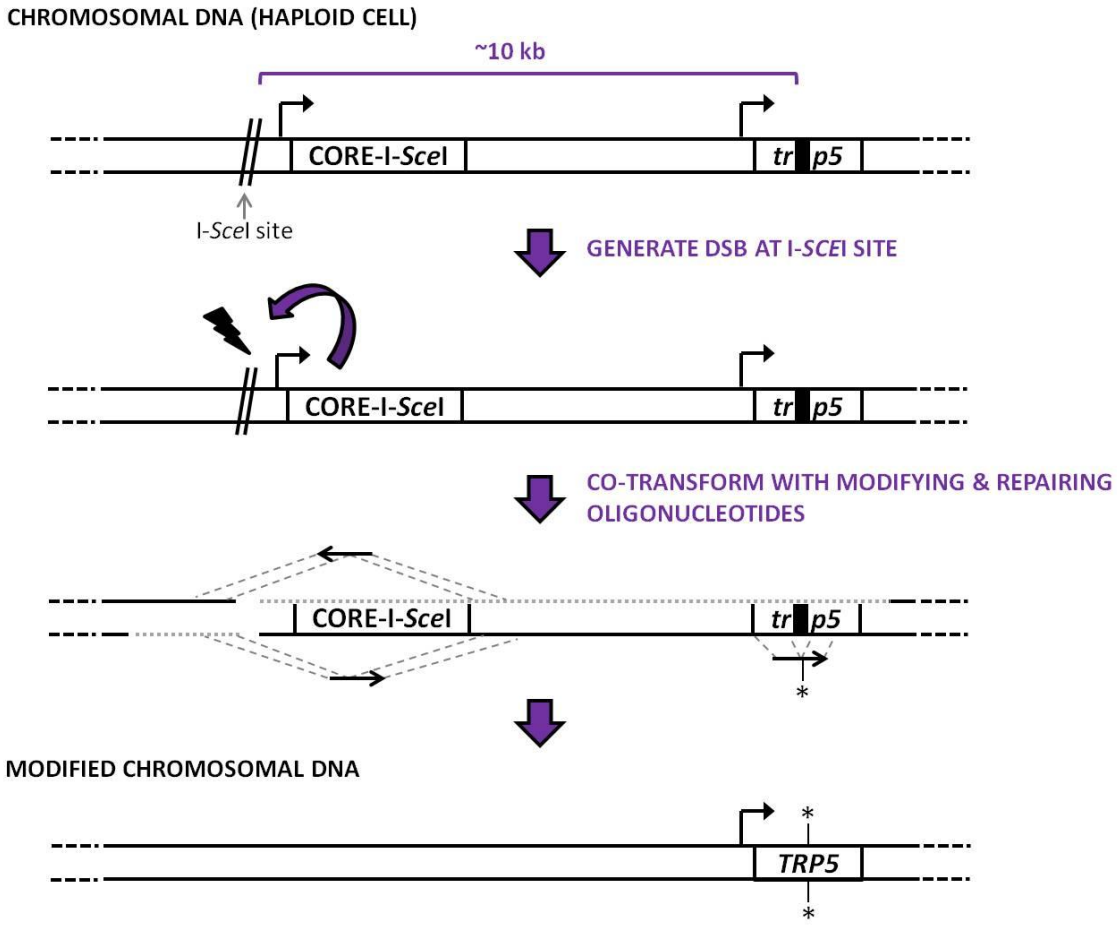

Figure 5.1 Modification of a distant locus 10 kb downstream of the DSB site in haploid yeast cells. In this example, the starting strain contains a disrupted *trp5* gene located approximately 10 kb downstream from the 18-bp recognition site for the I-*Sce*I endonuclease. Following addition of galactose, which induces expression of the nuclease (indicated by the lightning bolt), a targeted DSB is generated at the I-*Sce*I site and 5' to 3' resection occurs (indicated by the light gray dashed lines), leaving long 3' single-stranded stretches of DNA. Next, oligos are transformed into the cells. These include an oligo complementary to both sides of the disruption in *trp5* (as indicated by dashed lines between the oligos and genomic region), as well as oligos complementary to both sides of the break to repair it and simultaneously remove the CORE-I-*Sce*I cassette. Single oligos or the complementary pair can be used for each position. An asterisk on the oligo used to correct the *trp5* gene indicates the modification. Following transformation, the break is repaired while the target locus is modified (as indicated by the asterisk) to introduce the desired genetic modification in *TRP5*.

12. Remove the supernatant and resuspend cells well in 100 µl of sterile water.
13. Plate a 10^5-fold dilution per strain to SD Complete solid medium. Plate all cells from each transformation tube on one synthetic dextrose complete medium lacking tryptophan (SC-Trp) plate using approximately 15 sterile glass beads and incubate at 30°C for 3-4 days until colonies appear (*See* **Notes 5.6.12 & 5.6.13**).

5.4.9 Repair Assay: Example at the *leu2* Locus

Here we provide an example of an assay for repair of a mutated, non-functional *leu2* marker which is located 10 kb upstream from the DSB site in a diploid strain with the homologous chromosome being used to repair the break and oligos which correct the sequence of *leu2* (*See* **Figure 5.2**). Following correction of this marker, however, cells present a Leu^+ phenotype which can be selected for on solid medium lacking leucine. While targeted gene correction can be used at either a selectable or a non-selectable genetic locus, the presentation of a selectable phenotype following correction enables more rapid screening of transformants. Further information on targeting at a non-selectable genetic locus is provided in *Section 4* section below (*See* **Note 5.6.8**).

1. Inoculate 50 ml of YPLac liquid medium with chosen strain into a 250-ml glass flask and shake vigorously at 30°C O/N (*See* **Note 5.6.9**).
2. Add 5 ml of galactose from a 20% stock solution into the O/N culture for

a final 2% galactose solution and continue to shake at 30°C for 7 h (*See* **Note 5.6.10**).

3. Solutions 1 and 2 are prepared immediately prior to transformation.
4. Transfer culture to a 50-ml conical tube and spin at 1562 x *g* for 2 min.
5. Remove the supernatant and wash cells with 50 ml of sterile water and spin as stated previously.
6. Remove the supernatant and resuspend cells in 5 ml of solution 1 and spin as stated previously.

7. Remove supernatant and resuspend cells in 250-750 µl of solution 1. The pellet should be quite large. Depending on the dilution, this amount of cells is sufficient for up to approximately 30 transformations.
8. Aliquot 50 µl of the cell suspension in microcentrifuge tubes and add 1 nmole of DNA oligos (heat-denatured for 2 min at 100°C, then immediately kept on ice prior to use). When using a single oligo to modify the chromosome distant from the break, 20 µl (at 50 µM or 50 pmol/µl) is used. When using a complementary pair of oligos, 10µl of each are used.
9. Add 300 µl of solution 2 to each transformation tube. Mix briefly by vortexing.

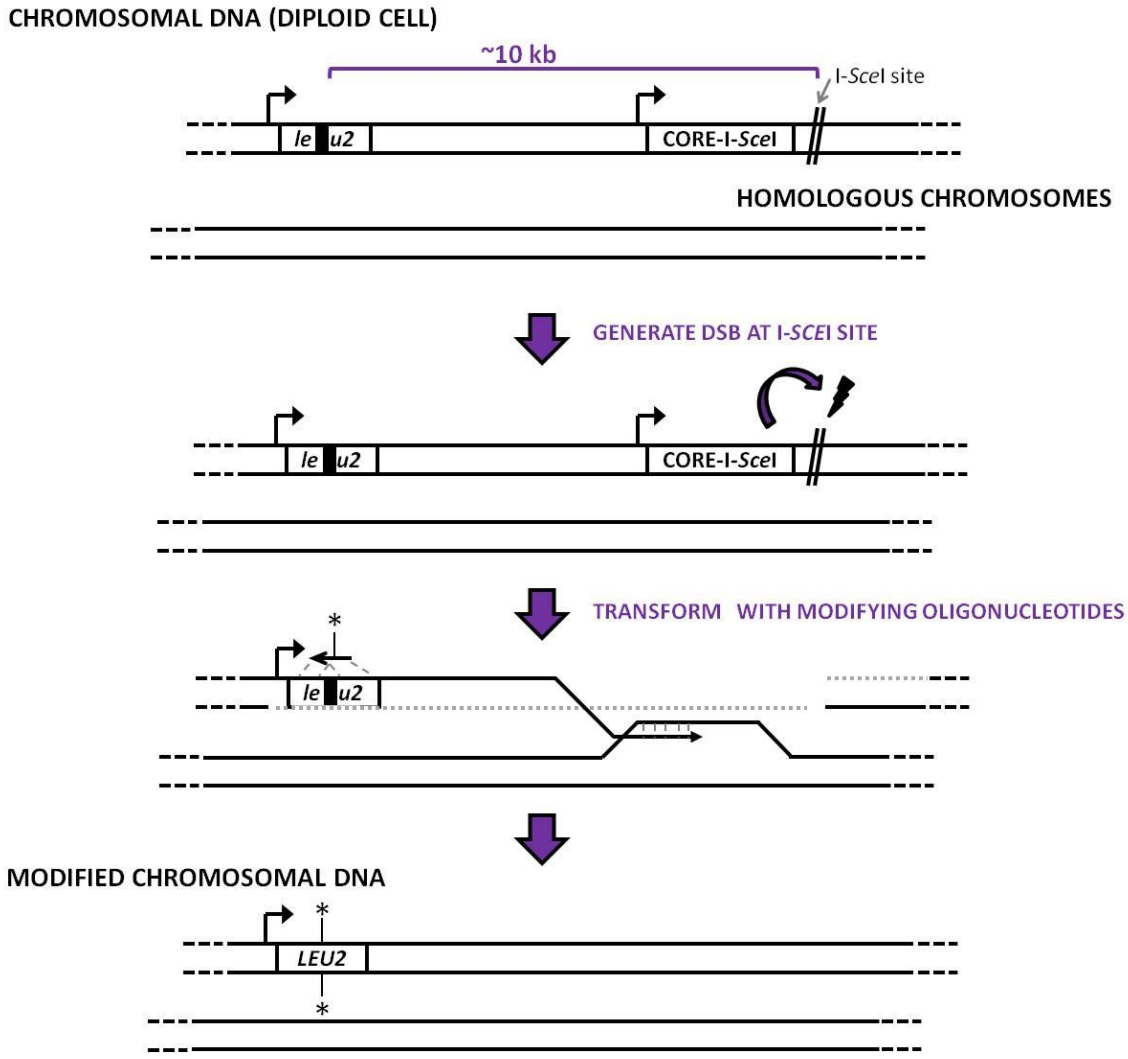

Figure 5.2 Modification of a distant locus 10 kb upstream of the DSB site in diploid yeast cells. In this example, the starting strain contains a disrupted *leu2* gene located approximately 10 kb upstream from the 18-bp recognition site for the I-*Sce*I endonuclease. Following addition of galactose, a targeted DSB is generated at the I-*Sce*I site, and 5' to 3' resection occurs (indicated by the light gray dashed lines), leaving long 3' single-stranded stretches of DNA. Next, oligos (here a single oligo is shown) are transformed into the cells with complementary to both sides of the disruption in *leu2* (as indicated by dashed lines between the oligo and genomic region). Oligos which repair the break are unnecessary here as the homologous chromosome is available for strand invasion by the 3' overhang (illustrated as the strand from the broken chromosome pairing with the complementary sequence of the homolog), which can lead to loss of the CORE-I-*Sce*I cassette (as shown here). An asterisk on the oligo used to correct the *leu2* gene indicates the modification. Following transformation, the break is repaired by the

homologous chromosome while the target locus is modified (as indicated by the asterisk) to introduce the desired genetic modification in *LEU2*.

10. Incubate transformation tubes at 30°C for 30 min with shaking.
11. Heat shock at 42°C for 15 min.
12. Collect cells by centrifugation at 2236 x *g* for 4 min.
13. Remove the supernatant and resuspend cells well in 100 µl of sterile water.
14. Plate a 10^5-fold dilution per strain to SD Complete solid medium. Plate all cells from each transformation tube on one synthetic dextrose complete medium lacking leucine (SC-Leu) plate using approximately 15 sterile glass beads and incubate at 30°C for 3-4 days until colonies appear (*See* **Notes 5.6.12 & 5.6.13**).

5.5 NOTES

Note 5.6.1

For medium preparation, deionized water is used. All other uses of the term "water" in this chapter, however, refer to water that has been sterilized by filtration or autoclaving.

Note 5.6.2

All solid media are autoclaved for at least 45 minutes at 121°C and then cooled to 55-60°C prior to pouring.

Note 5.6.3

Unless otherwise noted, all solid media are to be stored at 4°C. YPD liquid and solid preparations are exceptions that can be stored at room temperature due to frequent use.

Note 5.6.4

Filter-sterilizing the 50% PEG 4000 solution can take up to one hour depending on the volume due to the high viscosity of the mixture. Autoclaving is generally used for sterilizing this solution.

Note 5.6.5

We use Ex Taq DNA polymerase for amplification of the cassettes as it consistently produces a higher yield of CORE cassette amplification than Taq DNA polymerase.

Note 5.6.6

For O/N growth, 50-ml conical tubes are used. Additionally, lids should not be capped tightly because *S. cerevisiae* is an aerobic species. Instead, loosely cover the tube and secure with tape.

Note 5.6.7

When the *Kl-URA3* gene is inserted in the same orientation as the targeted gene, promoter occlusion may occur where interference from the target gene's promoter during transcription may lead to delayed growth of Ura^+ colonies. If colonies on SC-Ura medium are not observed after 3 days the following alternative approach can be used: First, plate onto YPD and incubate at 30°C O/N. Then, replica-plate to G418- or hygro-containing medium, depending on the CORE used, and incubate at 30°C for 2-3 days. Finally,

replica-plate to SC-Ura and incubate at 30°C O/N. If possible, it is always preferred to have the *Kl-URA3* marker gene oriented in the opposite direction relative to nearby promoters (*See* also [120]).

Note 5.6.8

This review provides examples of gene correction at a disrupted marker that is distant from the site of the DSB induced to stimulate the genetic modification. Our approach works well when the desired genetic modification results in a phenotypic change. If the desired genetic modification does not generate a change in the cell phenotype that can allow for easy selection, an alternative approach would be to screen by PCR aliquots of samples deriving from transformant cells, using primer(s) specific to the expected genetic modification. Following generation of a distant DSB, we have generally observed targeted gene correction frequencies varying from 10^{-3} to 10^{-4} ([52] and unpublished data). Using this as a guideline, we suggest testing by PCR 10 to 100 aliquots of cell samples deriving from 100 to 1,000 transformant cells. The aliquot(s) resulting positive at PCR should then be diluted and retested by PCR in order to narrow the number of cells containing the desired change. Such procedures should be repeated till the clone(s) with the desired mutation is identified.

Note 5.6.9

YPLac is used to provide a neutral carbon source for the cells prior to inoculation with galactose so as to not have inhibition of the galactose-inducible promoter. Since cells

grow much slower in this medium it is optimal to inoculate cells into YPLac at least 20-24 h prior to the transformation.

Note 5.6.10

Depending on the distance between the site of the DSB and the chromosomal locus to be targeted, the amount of time needed for resection will vary. Given the observed resection rate in yeast of 1 nt/s [144], an incubation time of 7 h following addition of galactose should provide ample opportunity for both expression of the nuclease and resection to provide lengths of single stranded DNA more than 10 kb away.

Note 5.6.11

A single or complementary pair of oligos can be designed to simultaneously repair the DSB and remove the CORE-I-*Sce*I cassette. This is advantageous as it allows for the markers present within the cassette to become available for generation of subsequent genetic modifications. If it is desired that the cassette remains in place, the DSB repairing oligo(s) can be designed accordingly.

Note 5.6.12

SD Complete medium is used here to determine the numbers of viable cells that survived the transformation. An appropriate dilution should be plated to ensure enough cells will be spread far enough for accurate counting.

Note 5.6.13

It may be necessary to plate a dilution from the transformed tube to a selective medium. This depends on the expected numbers of colonies following transformation and is up to the determination of the researcher. We generally plate both a 10^{-1} dilution and all remaining cells to a selective medium when using a pair of correcting/repairing complementary oligos.

CHAPTER 6

TO NICK OR NOT TO NICK: COMPARISON OF I-*SCE*I SINGLE- AND DOUBLE-STRAND BREAKS IN YEAST AND HUMAN CELLS

The work presented in Chapter 6 is part of a research article under the same name in preparation for submission to the journal PLOS One:

Stuckey[1], S., Gimble[2], F., and Storici[1], F. (2013) [1]School of Biology, Georgia Institute of Technology, Atlanta, GA. [2]Department of Biochemistry, Purdue University, West Lafayette, IN. (in preparation)

6.1 ABSTRACT

Repair of a disrupted genetic locus can be accomplished through targeted gene correction using the cell's HR machinery to introduce a corrected homologous sequence and replace the existing dysfunctional allele. While recombination is stimulated through generation of a site-directed DNA DSB, repair of the break through end-joining without using the homologous repairing template can lead to unwanted genomic changes such as chromosomal rearrangements. However, generation of an SSB, or nick, could lead to efficient gene correction without unwanted effects associated with a DSB as the SSB can be repaired through ligation of the nicked ends without loss of genetic sequences. Here,

we characterize a nicking variant of the I-*Sce*I endonuclease in the yeast *S. cerevisiae* and in HEK-293 cells. While an I-*Sce*I SSB stimulates gene targeting at lower levels than a DSB generated by the wild-type (wt) I-*Sce*I endonuclease, SSB-driven gene correction is proficient both at the site and distant from the site of the break at the target genomic locus in yeast and follows a different mechanism than I-*Sce*I DSB-driven correction. Moreover, an I-*Sce*I SSB efficiently stimulates gene targeting at multiple genetic positions in human cells. These findings demonstrate the capability of the I-*Sce*I nickase to stimulate recombination in multiple systems and could aid in further optimization of correction strategies for the treatment of genetic disorders.

6.2 INTRODUCTION

Gene targeting is a technique used to introduce desired genetic changes into the genome for purposes including genetic analysis through targeted deletion of a gene, addition of selectable markers, and gene replacement [36]. This process of genetic manipulation relies on HR to facilitate exchange of genetic sequences between the targeted chromosomal locus and a homologous template containing the desired mutation(s) [31]. Because many genetic disorders are caused by mutations in single genes, gene targeting has developed as a powerful technique capable of repairing the defective sequence through targeted replacement of the mutant allele with a functional copy [30,32,33]. However, because HR is naturally inefficient in many organisms, a DSB is often introduced within the vicinity of the targeted genomic locus in order to stimulate HR [49,51,131]. Numerous studies have demonstrated this activity using meganucleases,

such as I-*Sce*I and *HO* and which naturally have large cognate recognition sequences [44,49,51,53,131]. Additionally, designed endonucleases such ZFNs and TALENs, which can be designed to recognize highly-specific sequences, have widely been studied for their targeted gene correction potential [58,60,61,64,147]. Most recently, clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR associated (Cas) systems, which use RNA to target a homologous sequence for targeted cleavage and are naturally found in bacteria as part of the immune response, have shown great promise for further advancing gene targeting strategies [69,70]. However, while a DSB can efficiently stimulate HR up to 1,000-fold or more [49,51,53,54,131], the competing NHEJ pathway is favored in human cells which presents a major problem for gene targeting strategies [148]. Additionally, while an unrepaired DSB is lethal [149], repair of a DSB through NHEJ can lead to other harmful outcomes such as rearrangements and loss of essential genes [150]. Thus, other approaches to stimulate HR must be considered.

One avenue for stimulation of HR for targeted gene correction is through the generation of a site-directed SSB, or nick. As opposed to double-stranded endonuclease, a nickase cleaves a single strand of the double-stranded DNA molecule which is thought to be safer than generating a DSB due to the propensity for the SSB to be re-ligated. Recent work has also shown that an SSB can lead to less off-site targeting than a DSB [77,78]. Studies in yeast and human cells have demonstrated that an SSB can facilitate gene targeting. These include employing a bona fide nickase, such as Gene *II* from the bacteriophage f1; using nicking variants of natural meganucleases, such as I-*Ani*I; designing nickases with

zinc finger binding domains; and most recently through a CRISPR/Cas system using guide RNA (gRNA) to direct generation of an SSB at a target chromosomal locus [54,69,70,77,81-83,89,90]. Additionally, a nicking variant of I-*Sce*I was generated which demonstrated *in vitro* activity different from I-*Sce*I [108]. A lysine to isoleucine substitution at residue 223 was shown to abolish enzymatic cleavage activity at one strand of the 18-bp cognate I-*Sce*I recognition site [108]. Despite this, no *in vivo* data on I-*Sce*I K223I has been published demonstrating the capacity for an I-*Sce*I SSB to increase gene targeting.

Here, we show in multiple systems that an I-*Sce*I SSB increases targeted gene correction over non-breakage in the yeast *S. cerevisiae* and in human cells. The study reveals that I-*Sce*I K223I acts as a nickase *in vivo* and can generate an SSB both at the site and at sites distant from the SSB in yeast. Additionally, it does so following a different mechanism than wt I-*Sce*I. In yeast, genetic controls, cell cycle activity, and preference for different repairing molecules were assayed for this study. Moreover, we found that an I-*Sce*I SSB can increase gene targeting in human cells both at a target plasmid locus as well as on the chromosome. These results aid in characterization of the I-*Sce*I nickase and demonstrate its capacity for stimulating HR in different cell systems. Ultimately, these findings should further guide studies focused on developing the optimal method for targeted gene correction.

6.3 RESULTS

6.3.1 Cleavage activity *in vitro* of the I-*Sce*I nicking protein variant

Different variants of the wt I-*Sce*I protein were generated previously [108]. I-*Sce*I K223I contains a substitution of the lysine at residue 223 with an isoleucine designed to cleave only one strand at the cognate 18-bp I-*Sce*I recognition sequence, while I-*Sce*I D145A is a variant generated by substitution of the aspartate at residue 145 with an alanine and in which cleavage activity is abolished [108]. A supercoiled plasmid containing the I-*Sce*I recognition sequence was incubated with the wt, K223I, and D145A I-*Sce*I proteins to show the *in vitro* cleavage activity of K223I and D145A relative to the cleavage activity of wt I-*Sce*I. Within 20 minutes the plasmid was fully linearized by the wt I-*Sce*I protein (**Figure 6.1A**), while the D145A variant failed to cleave the plasmid sequence up to 6 hours later (**Figure 6.1B**). The K223I variant produced a nicked open circular product initially which was gradually converted into the linearized form (**Figure 6.1C**), indicating that the enzyme initially generates an SSB which later gets converted to a DSB.

6.3.2 An I-*Sce*I K223I break stimulates HR in yeast

To test the *in vivo* capacity for I-*Sce*I SSB-induced recombination in yeast, we generated plasmids containing the inducible *GAL1* promoter regulating expression of the wt I-*Sce*I, K223I, or D145A genes. For all of our repair assays, we determined HR stimulation based on the levels of transformants following expression of wt I-*Sce*I or K223I over those following expression of D145A. We generated haploid yeast strains using the *delitto perfetto* method for site-directed mutagenesis and containing the plasmids expressing the I-*Sce*I variants, as described in Chapters 2 and 3 previously.

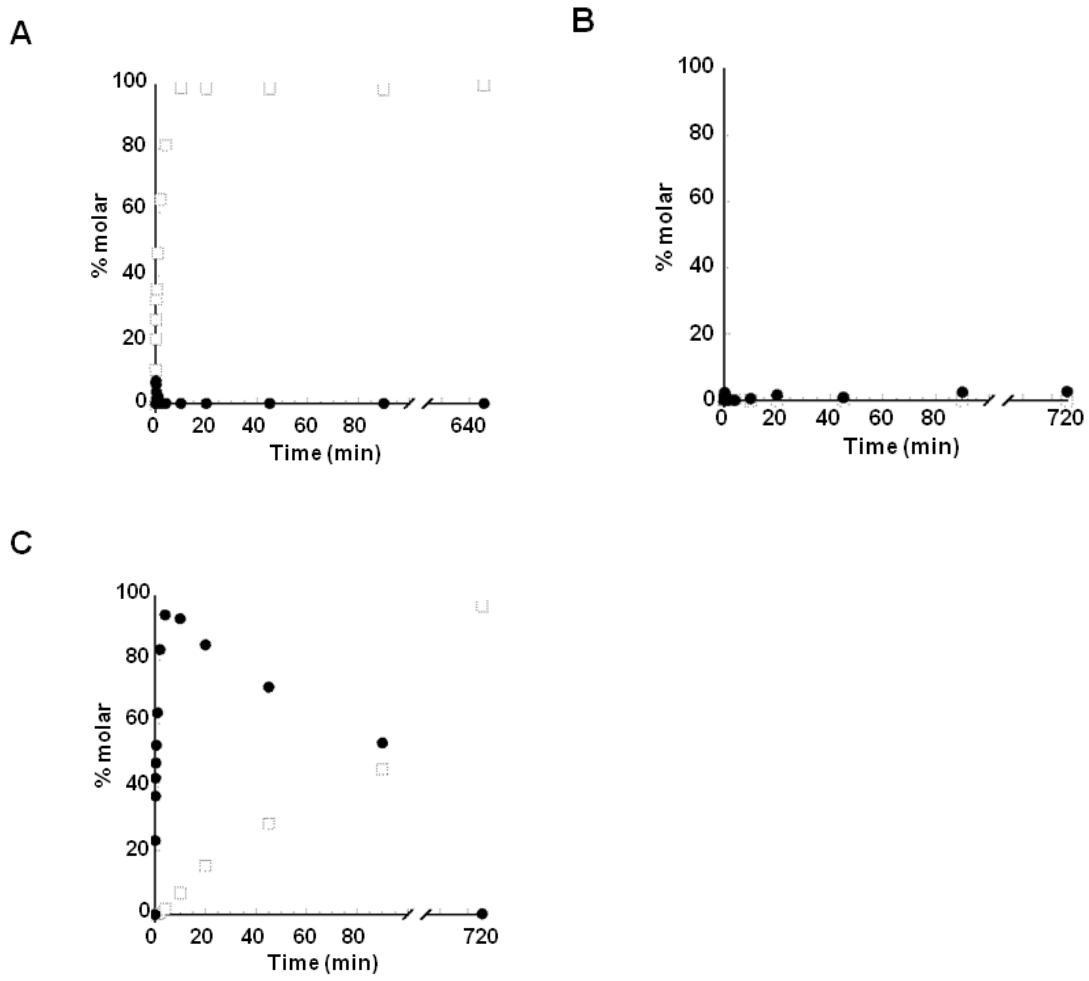

Figure 6.1 DNA cleavage activities of the wt I-*Sce*I, D145A, and K223I proteins. Supercoiled pBS I-SceI (E/H) plasmid DNA was incubated with (A) wt I-*Sce*I, (B) the D145A mutant or (C) the K223I mutant for various lengths of time and the amounts of the nicked open circle (filled circles) and linear (open squares) reaction product DNAs were plotted as a function of time. Data points represent the average values of two experiments.

We first looked at repair between direct repeats (**Figure 6.2A**). The *LYS2* genomic locus is disrupted by the I-*Sce*I recognition site which is flanked on either side by 90-bp direct repeats. Generation of a DSB between direct repeats triggers the single-strand annealing (SSA) HR pathway whereby, following resection of both ends in a 5' to 3' manner, annealing between the exposed single-stranded repeat regions and subsequent filling in of gaps leads to loss of the sequence between the repeats as well as one of the repeats itself [43]. With our system this would lead to reconstitution of the *lys2* gene, thus we expected a greater number of Lys^+ colonies with the DSB than with the non-break control. Following expression of wt I-*Sce*I and generation of the DSB, recombination was stimulated more than 2,000-fold ($p<0.0001$). While the frequency was much lower following expression of K223I compared to wt I-*Sce*I (1,000-fold less), there was a small but significant 2.4-fold increase over the non-break control ($p<0.0001$) (**Figure 6.2B**, left).

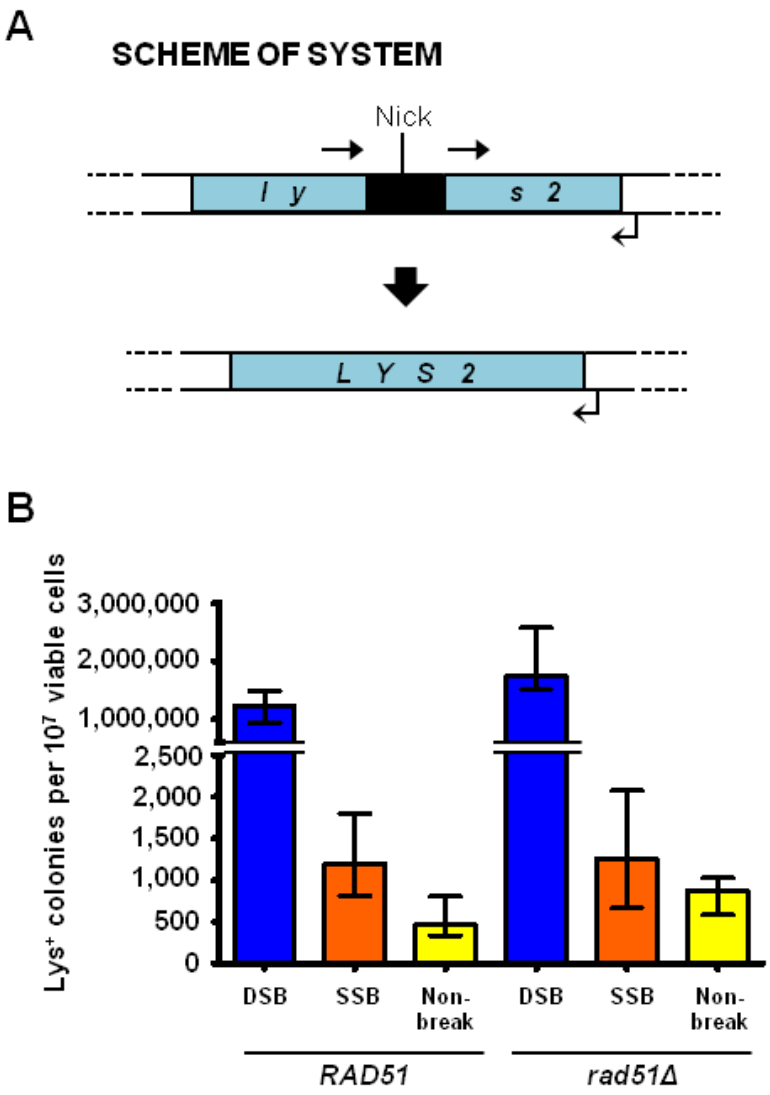

Figure 6.2 An I-*Sce*I K223I break stimulates HR between direct repeats in yeast. (A) Scheme showing disrupted yeast chromosomal locus containing the I-*Sce*I recognition sequence within 90-bp direct repeats (arrows). The position of the SSB is indicated ("Nick"). (B) Levels of Lys^+ recombinants following expression of wt I-*Sce*I (dark blue bars labeled "DSB"), K223I (orange bars labeled "SSB"), or D145A (yellow bars labeled "Non-break") in *RAD51* wild-type (left) or mutant (right) strains are presented as the median with range (n≥12).

Next, we looked at repair using oligos (**Figure 6.3A**). The genomic *TRP5* locus was disrupted by insertion of the I-*Sce*I recognition site. Synthetic DNA oligos 80-bp in length were designed with homology to either side of the disruption such that they could restore the sequence of the gene if used as a repair template, as described in Chapters 2 and 3. Previously, it was shown that oligos can efficiently repair a genetic region in yeast following generation of a DSB [53]. Using either of the single oligos (TRP5.80F (F), corresponding to the sense strand of the gene, and TRP5.80R (R), representing the antisense sequence) or the complementary pair, an increase in recombination at the *trp5* locus was observed (**Figure 6.3B**, left). Repair of the DSB was efficient for the complementary pair (190-fold increase, $p \approx 0.0022$) as well as for the single-stranded F or R oligo (80-fold increase, $p \approx 0.0022$, or 160-fold increase, $p \approx 0.0022$, respectively). With the K223I break, an increase in recombination frequency was also observed following transformation with oligos F, R, and the F+R pair (up to 9-fold increase for F, $p \approx 0.0022$; 8.2-fold for R, $p \approx 0.0022$; and 4.5-fold for pair, $p \approx 0.0087$, respectively). Using the 80-mers, recombination with K223I was 3-10% as efficient as with wt I-*Sce*I. Additionally, there was no single strand bias for repair of either type of break with the nick site positioned in the "Crick" ($p \geq 0.1797$) (**Figure 6.3B**, left) or "Watson" ($p \geq 0.3776$) (**Figure 6.3B**, right) orientation.

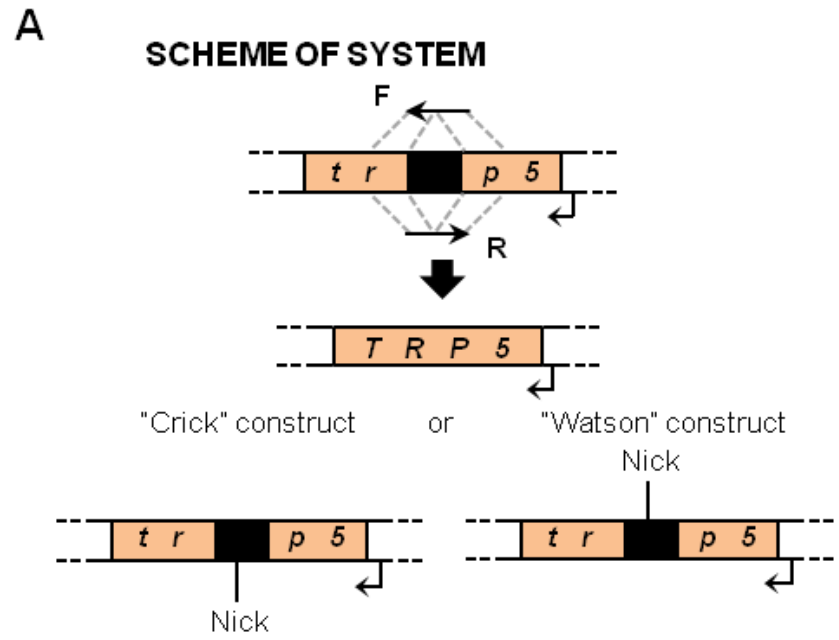

Figure 6.3 DNA oligos can efficiently be used to repair an I-*Sce*I K223I break. (A) Scheme showing disrupted yeast chromosomal locus containing the I-*Sce*I recognition sequence. The position of the SSB is indicated ("Nick"). F = oligo complementary to the anti-sense strand of the gene (indicated by dashed gray lines). R = oligo complementary to the sense strand. (B-D) Levels of Trp^+ transformants following expression of wt I-*Sce*I (dark blue bars labeled "DSB"), K223I (orange bars labeled "SSB"), or D145A (yellow bars labeled "Non-break") using either of the single or the pair of oligos to repair the break. All data are presented as the median with range ($n \geq 5$). (B) Comparison of the orientation of the site when an SSB is generated on the bottom ("Crick", left) or top ("Watson", right) chromosomal strand. (C) Levels of transformants in *RAD51* wild-type (left) or mutants (right) strains when the SSB is generated on the "Crick" strand. (D) Levels of transformants when cells were asynchronous (left) or arrested in G1 (right) prior to transformation when the SSB is generated on the "Crick" strand.

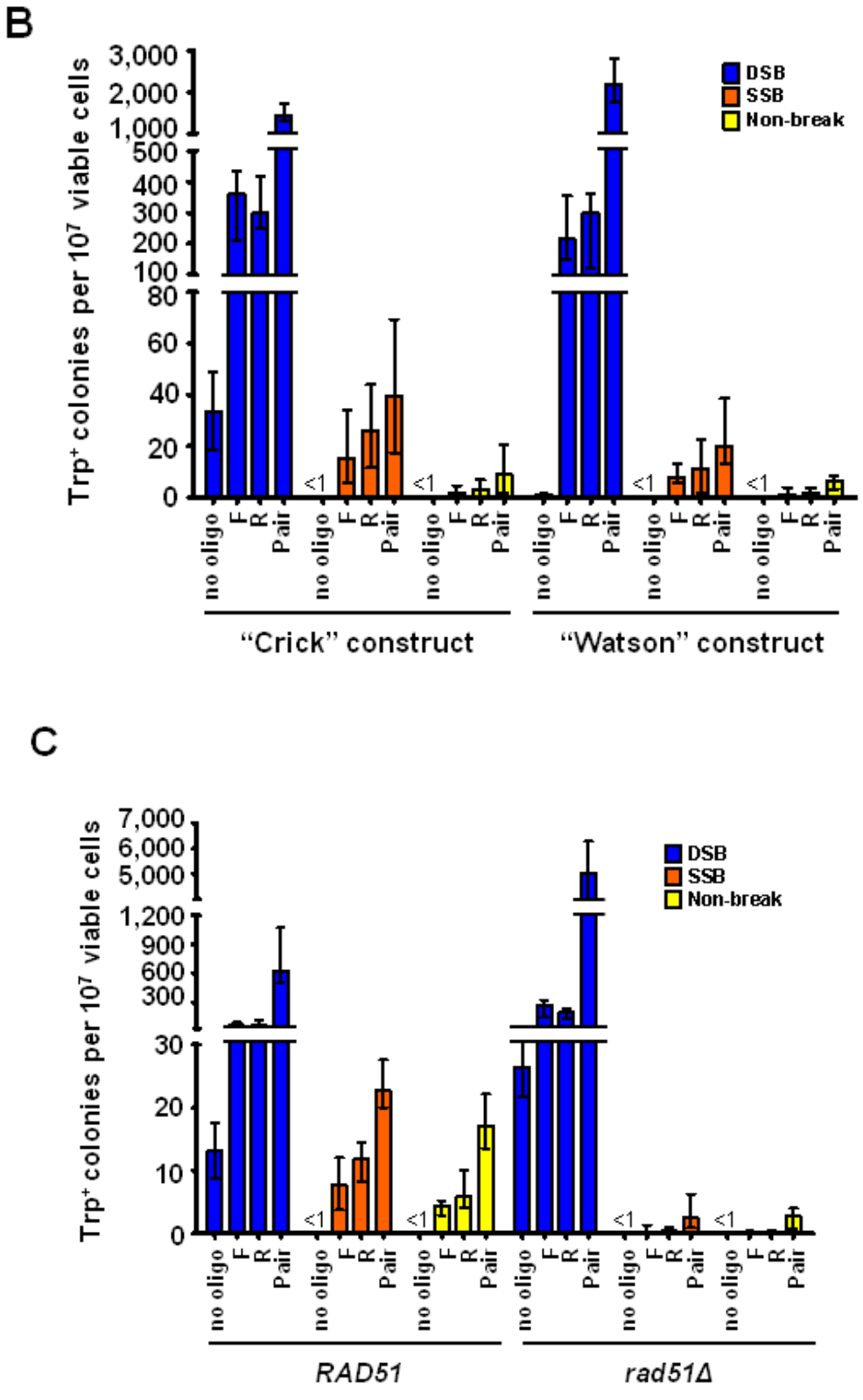
B
Trp+ colonies per 10^7 viable cells
3,000
2,000
1,000
500
400
300
200
100
80
60
40
20
0
DSB
SSB
Non-break
<1
no oligo
F
R
Pair
"Crick" construct
"Watson" construct
C
Trp+ colonies per 10^7 viable cells
7,000
6,000
5,000
1,200
900
600
300
30
20
10
0
DSB
SSB
Non-break
RAD51
rad51Δ

Figure 6.3 (continued)

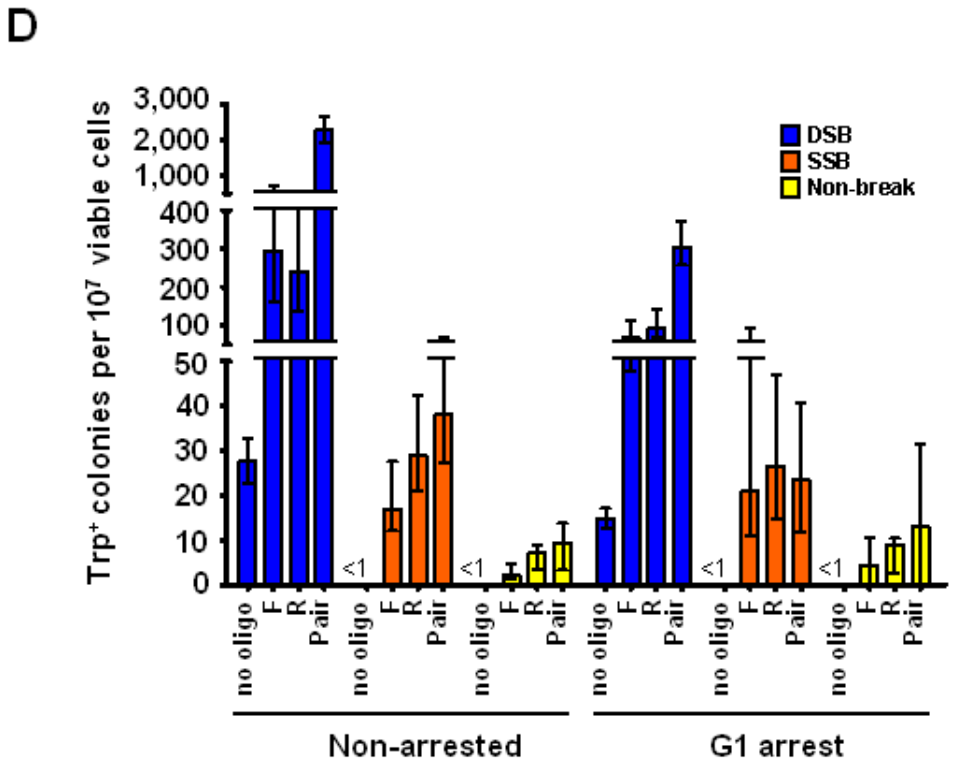

Figure 6.3 (continued)

6.3.3 Repair of a K223I break in yeast follows different genetic rules than for wt I-*Sce*I

Our *in vitro* experiment demonstrated that I-*Sce*I K223I efficiently generates a nick but also can form a DSB at the I-*Sce*I recognition site following prolonged incubation (**Figure 6.1**). In order to determine if the capacity of the K223I at stimulating gene correction was directly due to its nicking activity rather than its inefficient DSB activity, we looked into the requirements first of the strand invasion protein Rad51. DNA repair for *lys2* direct repeat recombination and gene correction by oligos at the *trp5* locus by the SSA mechanism does not require *RAD51* but rather levels of recombination are

stimulated following its deletion [151], therefore with a DSB we expected an increase in direct repeat recombination as well as with ss oligo-driven gene correction in *rad51* mutants. The frequency of Lys^+ colonies increased 1.5-fold ($p<0.0001$) in the *rad51* mutant background compared to *RAD51* wt cells following expression of wt I-*Sce*I (**Figure 6.2B**, right). A 1.8-fold increase ($p\approx0.0001$) was observed for the non-break control, as expected due to short length of the closely-spaced direct repeats [151,152]. There was no difference in levels with the K223I break ($p\approx0.4302$).

Similar to the genetic requirements for repair between direct repeats, oligo-mediated DNA repair of a DSB is both *RAD51*-independent and in fact stimulated following its deletion [52]. Accordingly, we observed a 2.9- to 8.1-fold increase in Trp^+ colonies ($p\leq0.0043$) in *rad51* mutants (**Figure 6.3C**, right). We observed the opposite for the K223I nick for which we recorded a significant decrease between 9.1- to 34-fold ($p\approx0.0022$). To demonstrate that the nicking activity of K223I, rather than its low DSB activity, could be the stimulus for the observed Rad51-dependent gene correction by oligos, we compared gene correction efficiency by K223I with that for wt I-*Sce*I using an amount of galactose that yielded frequencies of gene correction approximately equal to those observed for K223I in 2% galactose (**Figure 6.4**, left, light blue bars compared to yellow). Even at these frequencies (10-fold increase for F, $p\approx0.0043$, and 4.2-fold increase for R, $p\approx0.0022$, respectively) deletion of *RAD51* stimulated recombination (3.6-fold increase for F, $p\approx0.0152$, and 2.9-fold increase for R, $p\approx0.0411$, respectively) (**Figure 6.4**, right). These results are consistent with direct SSB-stimulated gene correction rather than gene correction promoted by low DSB levels following expression

of K223I. These findings indicate that the K223I break is different from one generated by I-*Sce*I, and we conclude that the gene correction-stimulating lesion is an SSB.

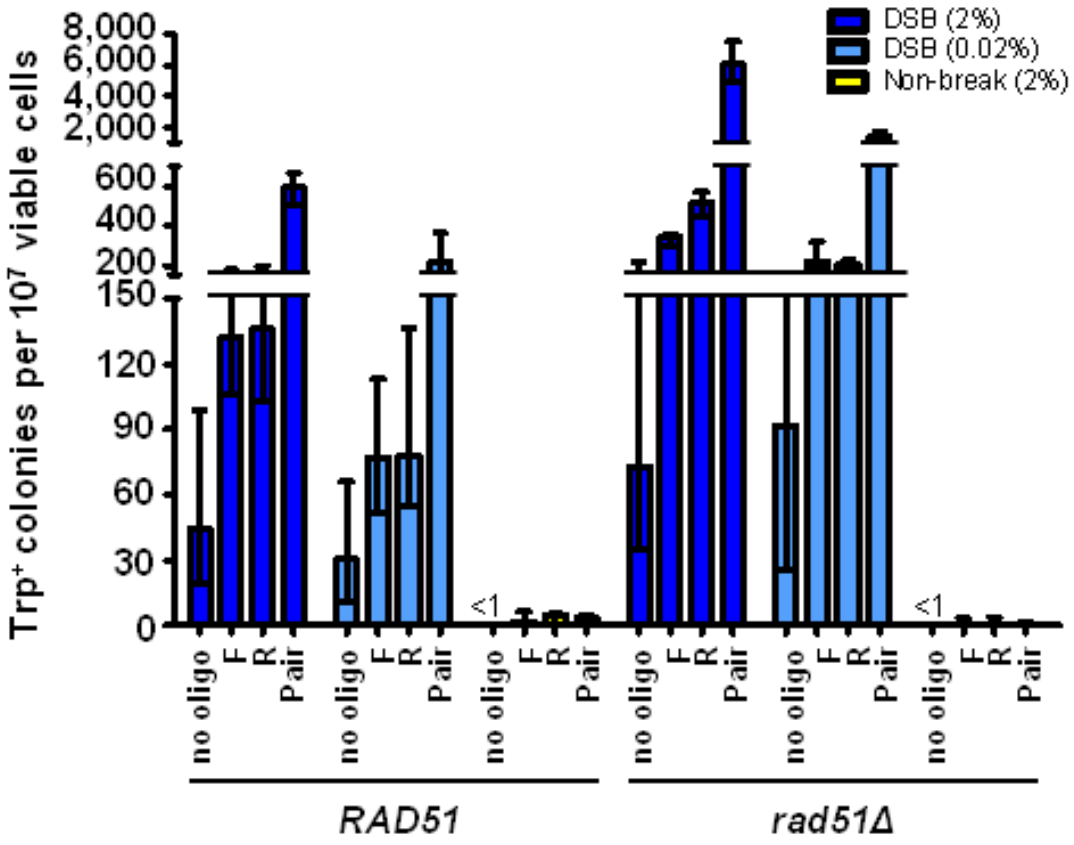

Figure 6.4 Recombination frequencies in *RAD51* and *rad51* null strains upon expressing wt I-*Sce*I using low galactose concentrations. Reducing the amount of galactose in solution lowers levels of recombination following expression of wt I-*Sce*I ("DSB (0.02%)", light blue bars, compared to "DSB (2%)", dark blue bars). Standard final concentrations of galactose in solution are 2%, while a 0.02% final concentration brought frequencies of recombination following expression of wt I-*Sce*I down to frequencies previously observed following expression of K223I (fold increase determined by comparing light blue bars on left to non-cleavage control ("Non-break") in yellow bars). Transformants were assayed in *RAD51* wild-type (left) and mutant (right) strains (n=6).

We next looked at repair in cells transformed after being arrested in the G1 phase of the cell cycle. We expected the frequencies of transformants would be lower following generation of a DSB due to the greater activity of NHEJ in G1. Following arrest in G1 with α-factor, we expressed wt I-*Sce*I, K223I, and D145A and then transformed *trp5* strains with the TRP5.80F and TRP5.80R 80-mers. Frequencies of Trp^+ colonies significantly decreased up to 12-fold following expression of wt I-*Sce*I in G1-arrested cells ($p \leq 0.0087$). Differently, frequencies were unaffected following K223I expression ($p \geq 0.0649$) (**Figure 6.3D**, right).

6.3.4 An I-*Sce*I K223I SSB stimulates HR at the site of the break in human cells

We also tested recombination at the site of the break in human cells. We generated plasmids expressing wt I-*Sce*I, K223I, or D145A under a strong CMV/CBA hybrid promoter modified from a previous study [60]. Each of these plasmids, a vector containing the I-*Sce*I site within a disrupted target locus, and 80-bp synthetic DNA oligos were transiently co-transfected into HEK-293 cells as described in Chapter 4. The target plasmid loci, *GFP* or *DsRed2* (referred to as *RFP*), were disrupted by an insert which includes the 18-bp recognition sequence for I-*Sce*I (**Figure 6.5A**). F and R oligos (GFP.80F and GFP80.R for *GFP* or DsRed2.80F and DsRed2.80R for *RFP*) were designed to restore the sequence of the disrupted gene, yielding GFP^+ or RFP^+ cells, depending on the marker, which were measured through fluorescence-activated cell sorting (FACS) approximately 8-10 days post-transfection. For both constructs, the F oligo corresponds to the sense strand of the gene and the R oligo represents the antisense sequence.

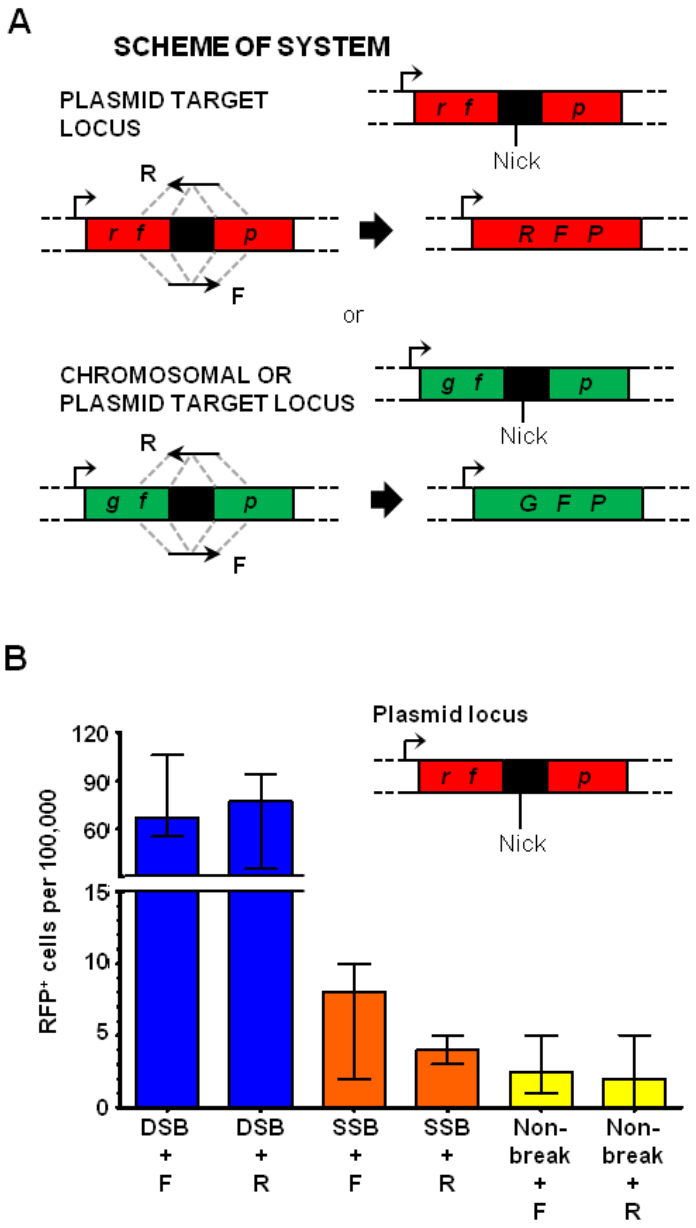

Figure 6.5 An I-*Sce*I K223I SSB can stimulate recombination in human cells. (A) Scheme showing disrupted *RFP* plasmid locus or disrupted *GFP* plasmid or chromosomal locus. The position of the SSB is indicated ("Nick"). F = oligo complementary to the anti-sense strand of the gene (indicated by dashed gray lines). R = oligo complementary to the sense strand. (B-D) Levels of fluorescent cells following expression of wt I-*Sce*I (dark blue bars labeled "DSB"), K223I (orange bars labeled "SSB"), or D145A (yellow bars labeled "Non-break") using either of the single oligos to repair the break. All data are presented as the median with range. (B) Recombination at the *RFP* target plasmid locus (n=6). (C) Recombination at the *GFP* target plasmid locus (n=9). (D) Recombination at the *GFP* target chromosomal locus (n≥8).

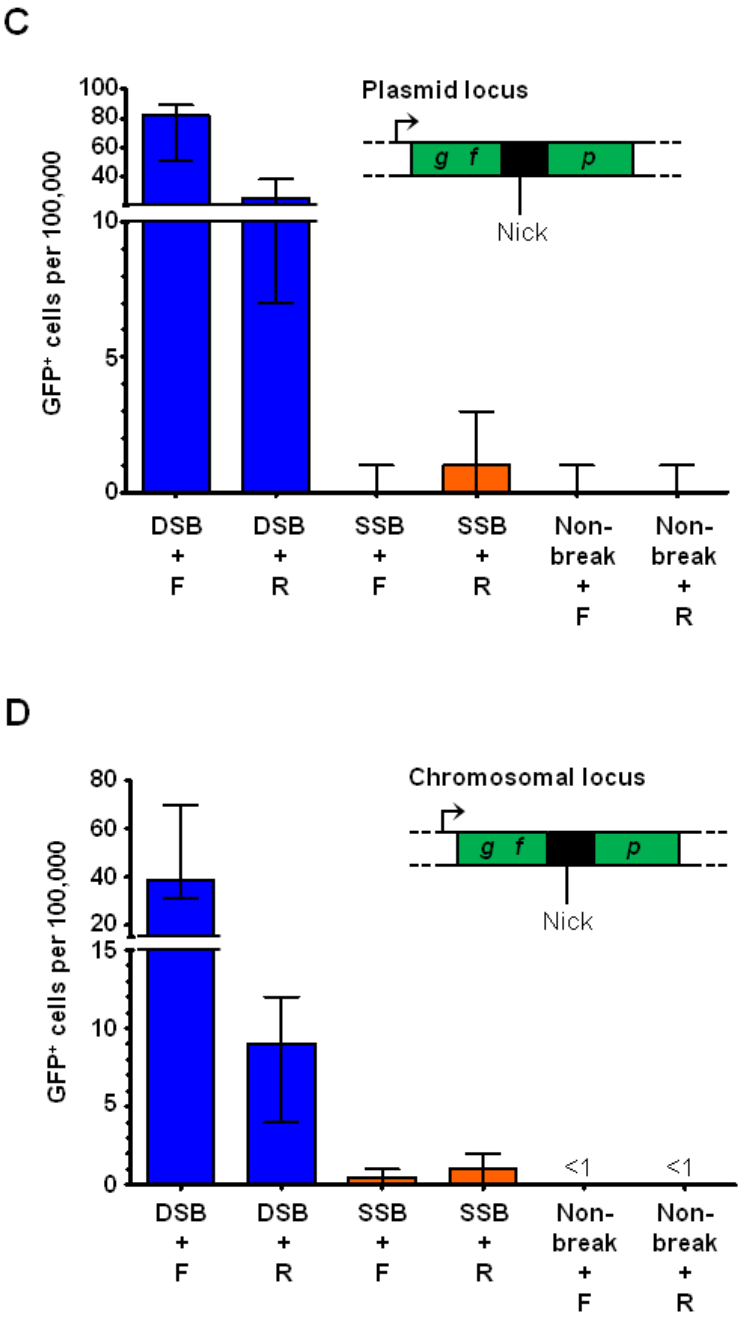

Figure 6.5 (continued)

At the *RFP* plasmid locus, higher than expected numbers of RFP^+ cells (per 100,000 cells median=5 (5-5) for F and median=3 (2-6) for R) were detected for the negative control in which the target locus and either repairing oligo was provided (no-enzyme control, data not shown). All other negative controls produced <1 fluorescent cell per 100,000 cells (<0.25-0.5, data not shown). Despite this level of background for the no-enzyme control, the DSB stimulated repair at the *RFP* locus using either F or R (13-fold increase, p≈0.0131, and 26-fold increase, p≈0.0139, respectively) (**Figure 6.5B**). Over the non-break control, levels were even higher (27-fold increase for F, p≈0.0050, and 39-fold increase for R, p≈0.0048, respectively). While there was no significant increase over the no-enzyme control using either oligo for repair of the SSB (p≥0.4500) or over the non-break control with F (p≈0.1215), the SSB did stimulate recombination at *RFP* using the R oligo 2-fold (p≈0.0423). At this plasmid locus, the SSB was ~5% as efficient as the DSB.

At the *GFP* plasmid locus, all negative controls produced <1 fluorescent cell per 100,000 cells (<0.33-1). The DSB stimulated recombination using F or R (647-fold increase, p<0.0001, and 228-fold increase, p<0.0001, respectively) (**Figure 6. 5C**). No increase was observed for repair of the SSB with F (p>0.9999) but there was a significant 12-fold increase with R (p≈0.0078). At this plasmid locus, the SSB was ~5% as efficient as the DSB.

We then looked at repair on the chromosome (**Figure 6.5A**). The cell line 658D is a monoclonal modified HEK-293 cell line in which a stably integrated copy of the same disrupted *GFP* sequence used for our plasmid assay was randomly introduced into the

genome [60]. We transfected 658D cells with the plasmid expressing the I-*Sce*I variant along with either F or R and measured fluorescence approximately 8 days later. All negative controls produced <1 fluorescent cell per 100,000 cells (<0.14-1). The DSB stimulated recombination using F or R (330-fold increase, p≈0.0002, and 77-fold increase, p<0.0001, respectively) (**Figure 6.5D**). Similar to the findings at the *GFP* plasmid locus, no increase was observed for repair of the SSB with F (p≈0.0769) but there was a significant 12-fold increase with R (p≈0.0004). On the chromosome at this target *GFP* locus, the SSB was ~16% as efficient as the DSB at stimulating recombination. At the *GFP* plasmid and chromosomal loci there was a bias in favor of the F oligo for the DSB (p<0.0001) while the R oligo was favored for the SSB (p≈0.0078 and p≈0.0334, respectively). No strand bias was observed at the *RFP* locus for either (p≈0.6991 and p≈0.3743, respectively).

6.3.5 An I-*Sce*I K223I SSB stimulates HR at distant loci in yeast

Previously it was shown that an I-*Sce*I DSB can stimulate HR at a distant locus both upstream and downstream from the site of the break in yeast [52]. In that study, a strand bias in favor of the correcting oligo complementary to the strand 3' of the DSB was observed due to the availability of that single-stranded DNA for annealing by the oligo following 5' to 3' resection [52]. We sought to determine if an I-*Sce*I SSB several kb distant from the genomic locus to be corrected could also stimulate HR at that locus and if any strand bias could be observed by using the system described in Chapter 5. We generated diploid yeast strains in which a cassette containing the wt I-*Sce*I, K223I, or D145A gene under the *GAL1* promoter, along with the 18-bp I-*Sce*I recognition site, was

integrated approximately 10 kb upstream or downstream from the *TRP5* locus, which was disrupted by a 31-bp insert described previously [52] (**Figure 6.6A**). In both constructs, the I-*Sce*I site was integrated in both orientations at the same chromosomal position such that the SSB would be generated on the top ("Watson") or bottom ("Crick") chromosomal strand. The TRP5.80F (F) and TRP5.80R (R) oligos used for repair at the site of the break were used to repair the *trp5* locus in these systems.

In the upstream system, the DSB stimulated recombination 7.1- to 35-fold with the F, R, and complementary pair of oligos ($p \approx 0.0286$) (**Figure 6.6B**). The SSB stimulated recombination 1.7- to 5.2-fold ($p \approx 0.0286$) using both of the single or the complementary pair of oligos except with the F oligo in the "Crick" construct ($p \approx 0.0571$) (**Figure 6.6B**). A strand bias in favor of the R oligo was observed following generation of the DSB in both orientations ($p \approx 0.0286$), while no strand bias was observed with the SSB for either the F or R oligo ($p \geq 0.0571$). In the downstream system, the DSB stimulated recombination 5.1- to 17-fold with the F, R, and complementary pair of oligos ($p \approx 0.0286$) (**Figure 6.6C**). The SSB stimulated recombination 2.3- to 2.8-fold with the single and complementary pair of oligos in the "Watson" construct ($p \approx 0.0286$), but no increase in transformants was observed in the "Crick" construct ($p \geq 0.0571$). A strand bias in favor of the F oligo was observed for the DSB in both orientations ($p \approx 0.0286$), while no strand bias was observed following generation of the SSB for either the F or R oligo ($p \geq 0.4857$). At distant loci in these systems, the SSB was 5-50% as efficient as the DSB at stimulating repair with oligos.

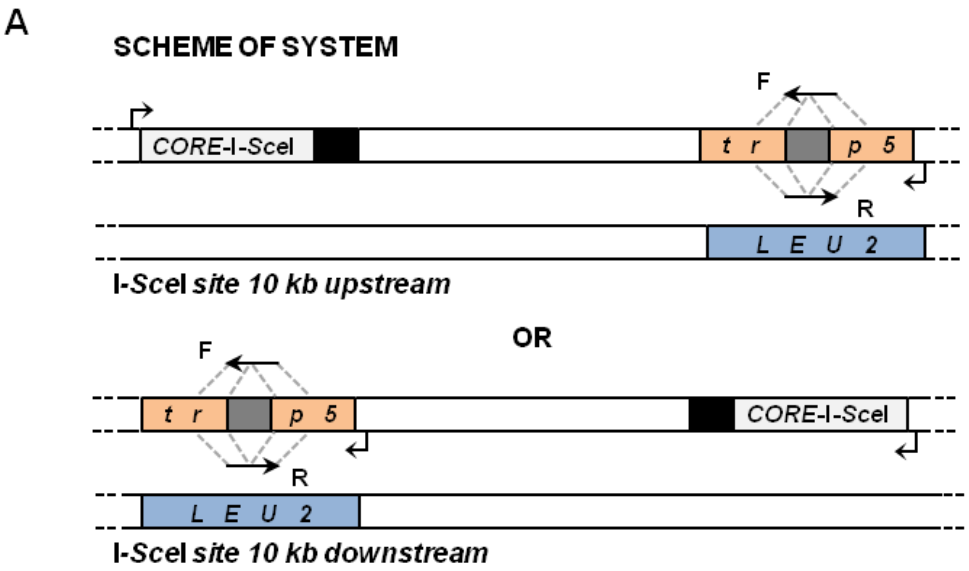

Figure 6.6 An I-*Sce*I K223I SSB stimulates HR at distant loci in yeast. (A) Scheme showing disrupted *TRP5* chromosomal locus and CORE-I-*Sce*I cassette inserted 10 kb upstream or downstream in diploid yeast cells. Black box indicates the position of the I-*Sce*I recognition sequence. F = oligo complementary to the anti-sense strand of the gene (indicated by dashed gray lines). R = oligo complementary to the sense strand. (B-C) Frequencies of Trp^+ transformants following expression of wt I-*Sce*I (dark blue bars labeled "DSB"), K223I (orange bars labeled "SSB"), or D145A (yellow bars labeled "Non-break") using either of the single or the pair of oligos to repair the break. All data are presented as the median with range ($n \geq 5$). (B) Frequencies of transformants when an SSB will be generated on the bottom ("Crick", left) or top ("Watson", right) chromosomal strand 10 kb upstream from the *trp5* locus. (C) Frequencies of transformants when an SSB is generated on the top ("Watson", left) or bottom ("Crick", right) chromosomal strand 10 kb downstream from the *trp5* locus.

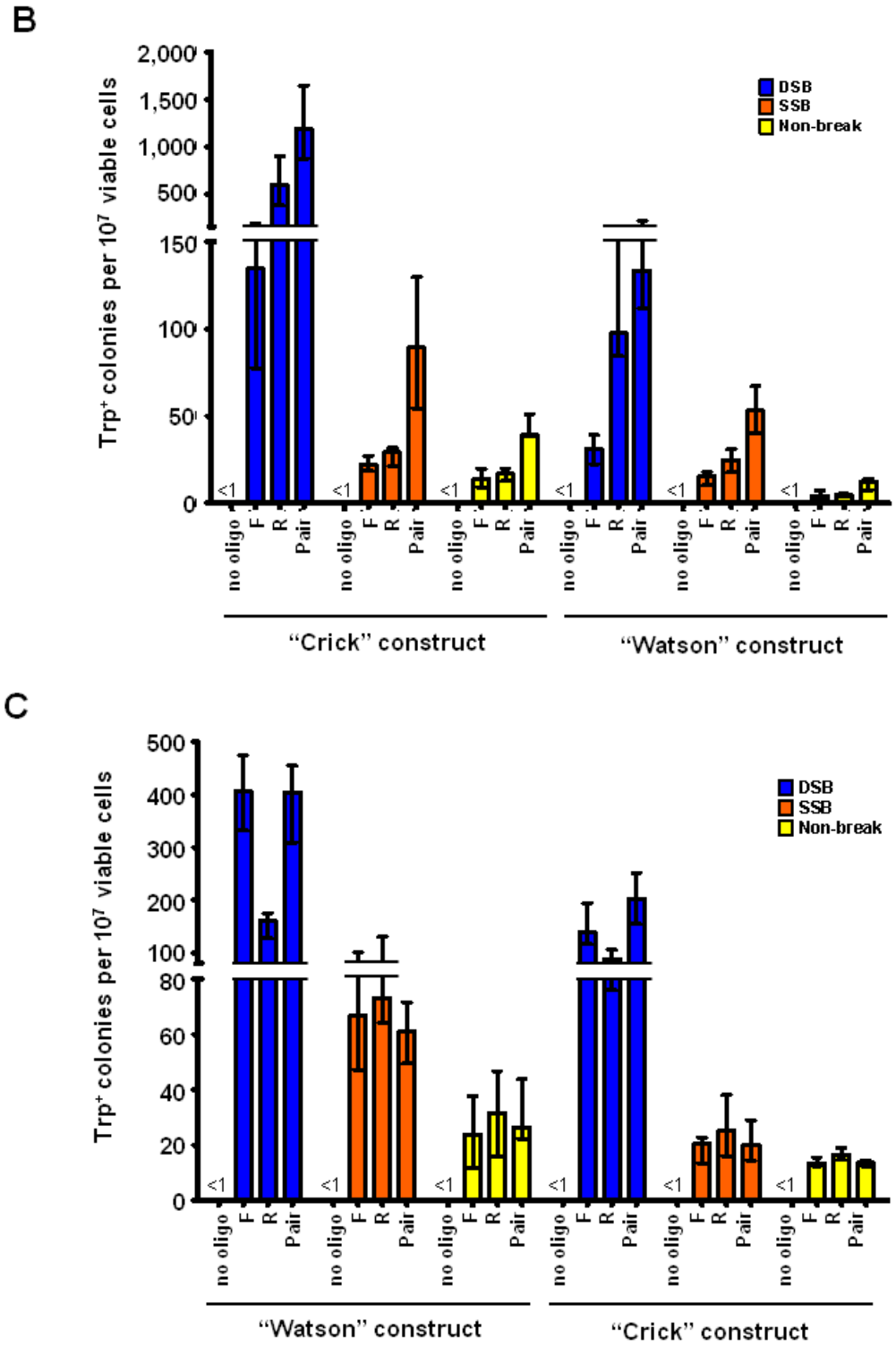

Figure 6.6 (continued)

6.3.6 An I-*Sce*I DSB, but not an I-*Sce*I SSB, stimulates HR at distant loci in human cells

To determine if the DSB or SSB could stimulate HR at distant loci in human cells, we generated a plasmid containing two disrupted fluorescent protein markers, *GFP* and *RFP*. The disruptions in each gene contain unique sites which can be used to generate a break. Following transfection with an oligo designed to restore the sequence of the marker in which the break was not generated, we could determine if the distant break was able to stimulate HR. Specifically, the disruption in the *RFP* locus contains the 18-bp I-*Sce*I recognition site which is located approximately 2.3 kb from the disruption in *GFP* (**Figure 6.7A**). We transfected this plasmid; a plasmid expressing wt I-*Sce*I, K223I, or D145A; and either of the GFP-correcting oligos into HEK-293 cells and measured green fluorescence through FACS approximately 8 days later. All negative controls produced <1 fluorescent cell per 100,000 cells (<0.25-1). While no increase in fluorescent cells was observed with the SSB using either oligo as a repair template ($p \geq 0.2731$) or with the DSB using R ($p \approx 0.0906$), a 9.6-fold increase with the DSB and the F oligo was observed at the *GFP* locus ($p \approx 0.0177$) (**Figure 6.7B**). We also generated a distant break *in vitro* ~2 kb from either *GFP* or *RFP*, but no significant increases were observed at either loci (data not shown).

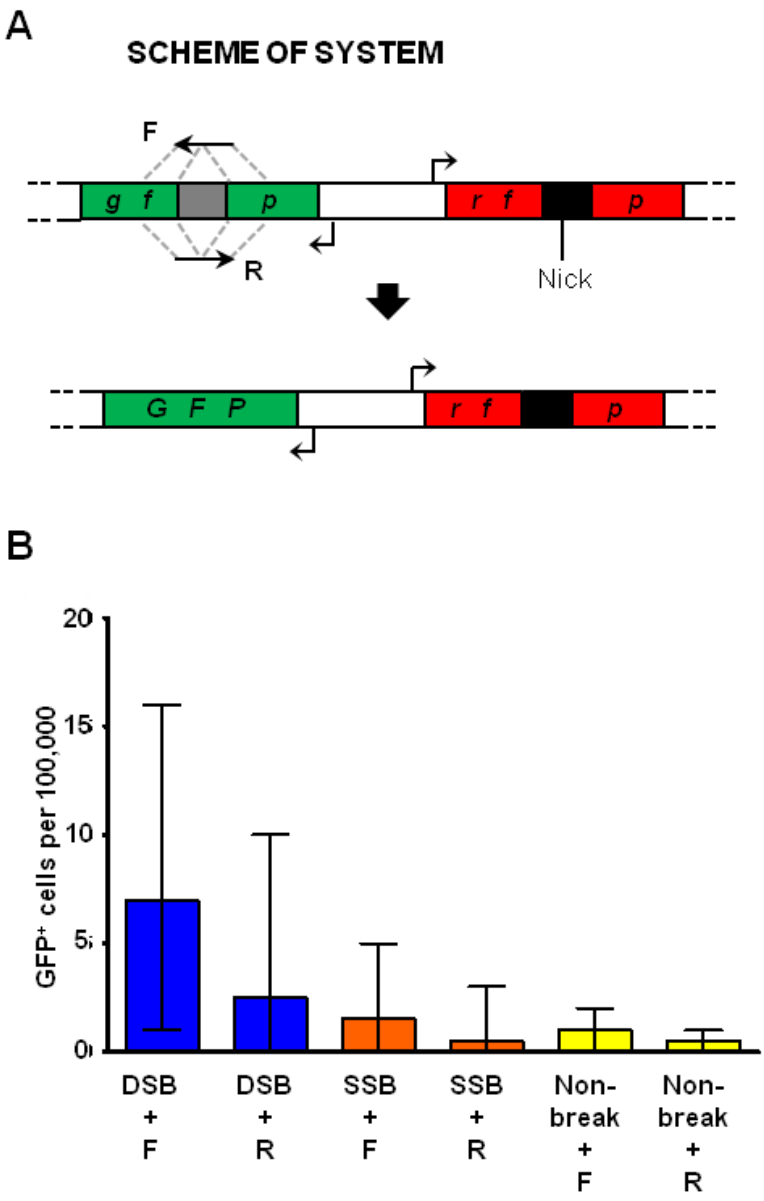

Figure 6.7 An I-*Sce*I DSB but not SSB stimulates HR at a distant plasmid locus in human cells. (A) Scheme showing disrupted *GFP* plasmid locus 2 kb distant from the I-*Sce*I recognition sequence (black box). F = oligo complementary to the anti-sense strand of the gene (indicated by dashed gray lines). R = oligo complementary to the sense strand. (B) Frequencies of GFP^+ cells following expression of wt I-*Sce*I (dark blue bars labeled "DSB"), K223I (orange bars labeled "SSB"), or D145A (yellow bars labeled "Non-break") using either of the single oligos to correct the *GFP* gene located 2 kb distant from the I-*Sce*I break. All data are presented as the median with range (n=6).

6.4 DISCUSSION

The ability to activate the HR machinery through generation of a site-directed break enables targeted correction of genes if a homologous repairing template is provided. While this is most often accomplished following generation of a DSB, the potential for off-site targeting and genomic rearrangements that accompany non-homologous repair of the break is great [150]. Though designed nucleases, such as ZFNs and TALENs, can be optimized for highly-specific recognition of a genomic locus, the sequence constraints for such precision may limit the availability of suitable target regions [153,154]. Conversely, the emergence of designed nickases, such as engineered variants of natural meganucleases as well as zinc finger nickases, as an alternative means of stimulating HR with less off-site targeting holds great promise for gene targeting and correction [77,81-83,90]. With this study, we aimed to characterize a novel variant of wt I-*Sce*I, I-*Sce*I K223I, which has been shown *in vitro* to preferentially cleave the 18-bp I-*Sce*I recognition site in only one specific position [108]. While wt I-*Sce*I has been widely-used to demonstrate the capability of DSB-driven gene targeting, no *in vivo* studies prior to this have explored the capacity for an I-*Sce*I SSB to stimulate HR. We demonstrate here that I-*Sce*I K223I generates a nick which can trigger the HR machinery in yeast and in human cells. Additionally, by directly comparing wt I-*Sce*I and K223I in various assays, we demonstrate that the lesions are repaired in different ways.

First, at the site of the break in yeast we observed an increase in recombinants following expression of K223I (**Figure 6.2B** and **Figure 6.3B**). While frequencies compared to wt I-*Sce*I were low (0.1% between direct repeats and 3-10% using oligos), there was an up

to 9-fold increase in repair with K223I in yeast cells. In human cells an up to 12-fold increase was also observed at the site of the nick at multiple plasmid loci and even on the chromosome (**Figure 6.5**, panels B, C, and D). These findings demonstrate that an I-*Sce*I K223I nick can activate HR *in vivo*, but they do not rule out the likelihood of this stimulation being due to an inefficient DSB. To explore this possibility, we looked into the requirement for the Rad51 recombinase.

Recombination between direct repeats and gene correction by oligos at the site of the break in yeast stimulated by an I-*Sce*I DSB not only does not require Rad51 but frequencies in fact increase in *rad51* mutants (**Figure 6.2B** and **Figure 6.3C**). However, there was no change with K223I in the direct repeat system (**Figure 6.2B**) and gene correction frequencies significantly decreased with the oligo assay, indicating a strong dependence on a strand exchange reaction for oligo-mediated repair with this protein (**Figure 6.3C**). The dependence on Rad51 is not due to a low level DSB being generated by I-*Sce*I K223I as reducing the recombination-inducing activity of the DSB by lowering expression of wt I-*Sce*I still stimulated recombination in the absence of Rad51 (**Figure 6.4**). Thus, we concluded that K223I generates an SSB *in vivo* which is repaired in a different manner than the I-*Sce*I DSB. Additionally, while DSB repair frequencies significantly decreased when cells were transformed following G1 arrest, likely due to inactive resection in this stage of the cell cycle, there was no difference in recombination frequency following induction of the I-*Sce*I SSB (**Figure 6.3D**), also indicating that the SSB-induced recombination is not due to an initial DSB. Finally, in human cells different strand biases for repair of the I-*Sce*I SSB or DSB were observed at the *GFP* locus

(**Figure 6.5**, panels C and D). Together, these findings indicate differences in the ways wt I-*Sce*I and I-*Sce*I K223I stimulate recombination.

We also tested whether I-*Sce*I SSB could trigger HR at sites distant from the break in yeast and human cells. For gene targeting to be successful, the recombination machinery must be activated, but this requires generation of a break which is often accomplished through expression of a nuclease with sequence recognition near the targeted genomic locus. However, the sequence specificity of these enzymes may limit the region in which a suitable break may be generated. For example, ZFNs and TALENs require a spacer region between the DNA recognition sequences for each nuclease of the pair; however, the most preferential sequence for these nucleases may be located several kb away from the targeted locus to be corrected. If it is possible to generate the distant break while still triggering the HR machinery, then gene correction approaches could explore regions within this wider window as target sites for cleavage. While this has been demonstrated in yeast following generation of a DSB [52], to our knowledge the potential for correction at a genomic locus distant from an SSB has not been explored until this study. Following generation of the SSB in yeast, we observed a significant increase in recombinants when the break was generated both 10 kb upstream or downstream from the genomic locus targeted for correction (**Figure 6.6**, panels B and C left). Additionally, while a strand bias in favor of the oligo complementary to the strand 3' of the break was observed for after generation of the DSB, no bias was observed in any system for following SSB generation for the distant assays. Despite these findings, an I-*Sce*I SSB did not stimulate recombination at a distant *GFP* locus in human cells (**Figure 6.7B**),

though an I-*Sce*I DSB generated approximately 2 kb away was able to trigger the event. It may be possible to facilitate distant modification in human cells through mutation of 53bp1 or Rif1, which function together to protect DNA ends from resection [155,156], as inactivation of this process would enhance DNA resection and expose sections of ssDNA for complementary pairing with an oligo.

Our work demonstrates that expression of the I-*Sce*I K223I variant in yeast and human cells stimulates recombination in these cells and that the molecular path by which this enzyme stimulates recombination is different from that driven by wt I-*Sce*I. We propose two general models for I-*Sce*I SSB-driven gene targeting by oligos. In the first model, an oligo can be used as a repair template to fix this SSB. As Rad51 is necessary for repair of the SSB, duplex invasion must be a required step (**Figure 6.8** panels **A** and **B**). In the second model, the oligo is used to repair the break after collapse of the fork during replication. Again, invasion into the resulting duplex is crucial for repair of the lesion (**Figure 6.8C**). Additionally, while gene targeting following generation of a distant break in human cells can be accomplished (at least within a short window) through generation of DSB, studies should explore whether other nickases could trigger HR at distant loci in human cells. Ultimately, while the SSB does not result in a great level of stimulation at the site of the break, it does stimulate significant recombination and provides further evidence that SSB-driven gene targeting is a viable mechanism through which targeted gene correction can be accomplished.

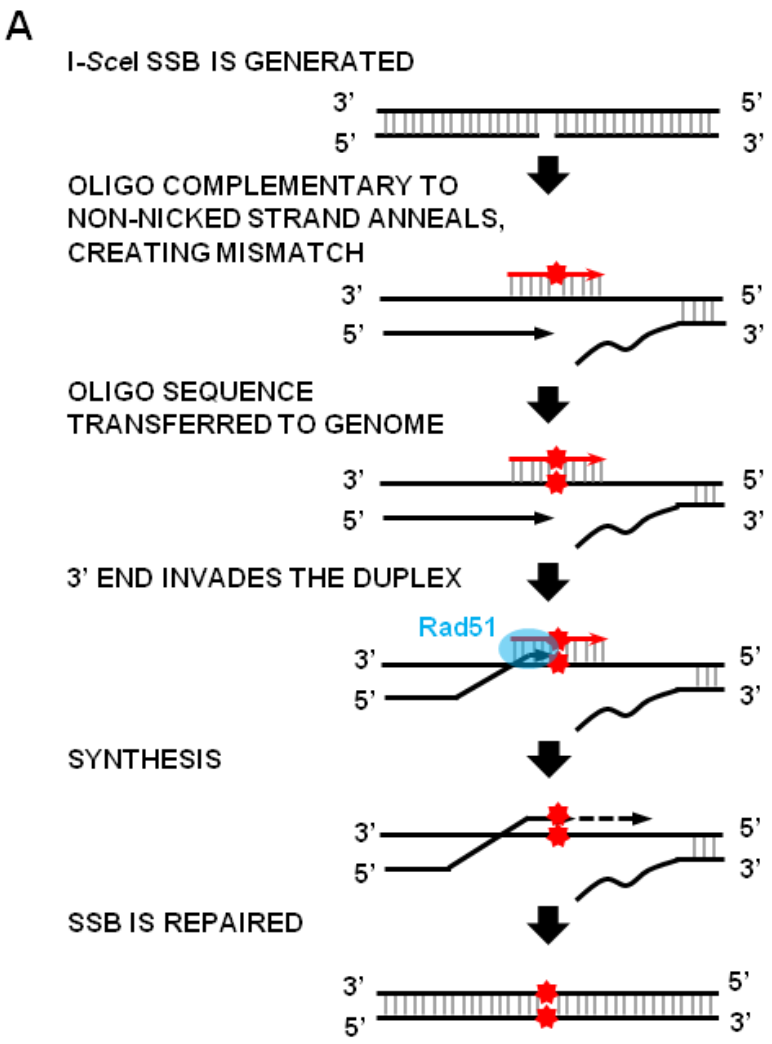

Figure 6.8 Model for I-*Sce*I K223I SSB-driven HR using single-stranded oligos. (A) Recombination using the oligo complementary to the intact strand. After the SSB is generated, the oligo (red arrow) anneals to the complementary strand, creating a mismatch which may produce a bulky lesion. This mismatch is processed, resulting in a gap which can be filled using the oligo sequence (red star), transferring it to the genome. The 3' strand of the break may then invade the duplex following unwinding. Then the 3' end is extended. (B) Recombination using the oligo complementary to the nicked strand. After the SSB is generated, the oligo sequence (red star) can be copied as it is used to extend the 3' end. This extended end may invade the duplex at the 5' end of the break. The 3' end is then extended further. (C) Recombination following collapse of the replication fork. After the SSB is generated, it persists until encountered by the replication fork. Following fork collapse, resection of the 5' ends (indicated by light gray dashed lines) produces single strands. Next, duplex formation involving annealing of the oligo allows for invasion by the 3' replicating end (dark gray). The 3' end copies the sequence of the oligo (red star) and is then extended further.

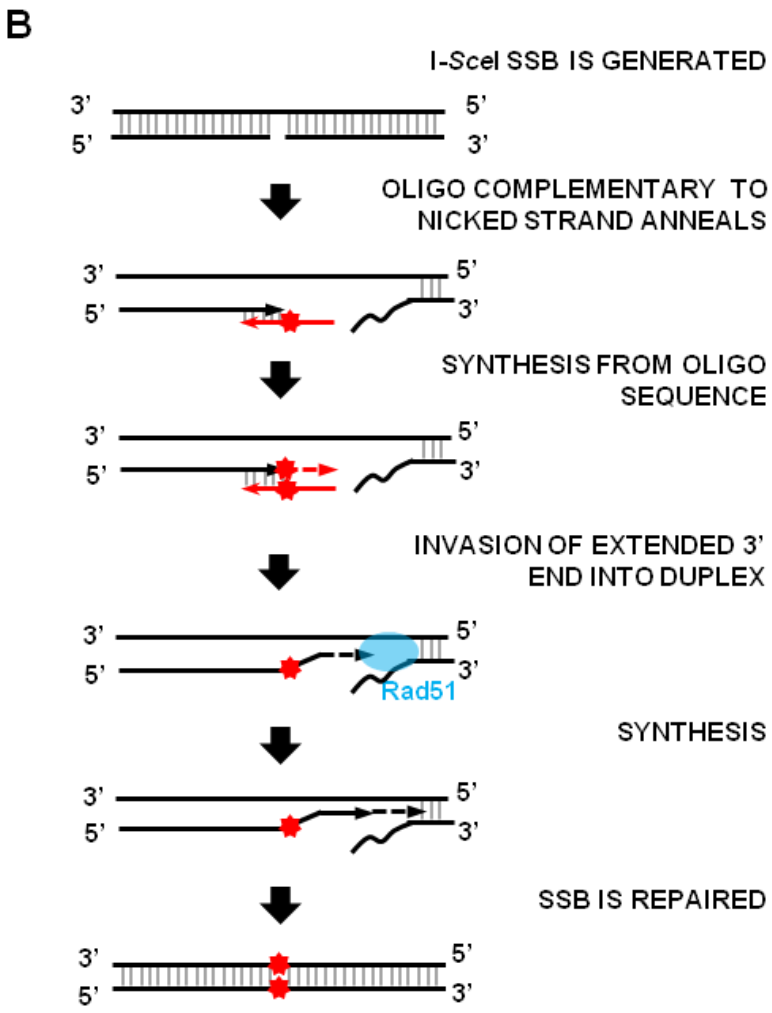

Figure 6.8 (continued)

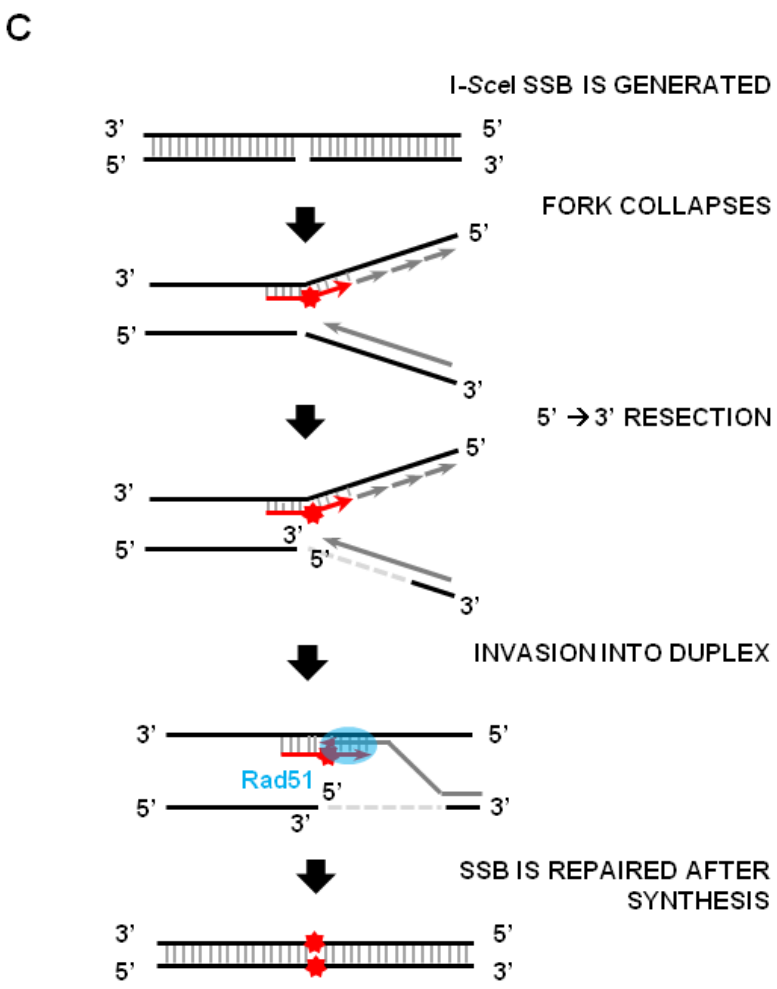

Figure 6.8 (continued)

6.5 MATERIALS AND METHODS

6.5.1 Cleavage activity

The double-strand and single-strand DNA cleavage activities of wild-type I-*Sce*I and the K223I and D145A variants were analyzed by a kinetics assay using a supercoiled plasmid (pBSI-SceI (E/H)) containing a single wild-type I-*Sce*I recognition sequence. Reaction mixtures contained I-*Sce*I cleavage buffer (10 mM Tris-HCl (pH 8.8), 1 mM DTT, and 0.1 mg/ml BSA), pBS-I-SceI (E/H) supercoiled plasmid DNA (2.5 nM), and purified wild-type, K223I or D145A I-*Sce*I (100 nM). Reactions were initiated by the addition of $MgCl_2$ (15 mM), and the mixtures were incubated for various lengths of time at 30°C before being halted by the addition of EDTA (100 mM). Supercoiled, open-circle and linear DNAs were resolved by gel electrophoresis on a 0.8% agarose gel. Gels were stained with ethidium bromide, and the fluorescence intensities of the DNA bands were determined using a Kodak EDAS 290 imager.

6.5.2 Yeast plasmids

Plasmid pGSHU, described previously [53], contains the CORE-I-*Sce*I cassette (wt I-*Sce*I gene under the *GAL1* promoter, hygromycin resistance gene, the counterselectable *KlURA3*). Plasmids pGSHU-K223I and pGSHU-D145A are derivatives of pGSHU and contain the I-*Sce*I K223I and I-*Sce*I D145A genes, respectively, which were previously described by Niu *et al*. [108] to demonstrate nicking or non-cleaving *in vitro* activity, respectively, under the *GAL1* promoter in addition to the hygromycin resistance gene and *KlURA3*. The resulting cassettes are referred to as CORE-wtI-*Sce*I, CORE-K223I, and CORE-D145A, respectively.

Plasmid pAG7 [54] is a yeast expression vector containing the *GAL1* promoter, Gene *II* gene, and selectable *LEU2* marker. Plasmids pAG7-wtI-*Sce*I, pAG7-K223I, and pAG7-D145A, are derivatives of plasmid pAG7 in which the Gene *II* gene was replaced with the wt I-*Sce*I, K223I, or D145A gene, respectively.

Mutagenesis and cloning information are provided as supporting information (Appendix B).

6.5.3 Yeast strains

Strains used for the direct repeat assay are derivatives of FRO-830 (*MATα leu2-3,112 ade5-1 his7-2 ura3*Δ *trp1-289 [lys2::Alu-DIRLEU2-lys2*Δ*50]*) [52] and contain a 27-bp insertion comprised of two stop codons and the 18-bp I-*Sce*I site in *lys2* within 90-bp direct repeats flanking either side of the insertion. Strains were generated as previously described [102-104,120]. Strains SAS-74 and -75 contain plasmid pAG7-wtI-*Sce*I, -77 and -149 contain plasmid pAG7-K223I, and -142 and -143 contain plasmid pAG7-D145A, respectively. Strains SAS-174 and -175 contain plasmid pAG7-wtI-*Sce*I, -176 and -177 contain plasmid pAG7-K223I, and -178 and -179 contain plasmid pAG7-D145A, respectively, and are *rad51* mutants in which *RAD51* is replaced with the *kanMX* cassette.

Strains used for repair at the site of the break are derivatives of FRO-1 (*MATα ade5-1 his7-2 leu2-3,112 ura3-52*) [52] and contain one stop codon and the I-*Sce*I site within a 26-bp disruption in *trp5*. Strains were generated as previously described [102-104,120].

For most assays, the site is oriented such that the I-*Sce*I SSB is on the bottom, or "Crick", chromosomal strand. The mating type of these strains was switched to mating type **a** to allow for cell cycle arrest with α-factor. Strains SAS-227 and -228 contain plasmid pAG7-wtI-*Sce*I, -229 and -230 contain plasmid pAG7-K223I, and -231 and -232 contain plasmid pAG7-D145A, respectively, and are mating type **a**. Strains SAS-235 and -236 contain plasmid pAG7-wtI-*Sce*I, -237 and -238 contain plasmid pAG7-K223I, and -239 and -240 contain plasmid pAG7-D145A, respectively, and are mating type **a** as well as *rad51* mutants in which *RAD51* is replaced with the *kanMX* cassette. For comparison of the orientation of the I-*Sce*I site at the site of the break, strains were generated such that the SSB is made on the top, or "Watson", chromosomal strand. Strains SAS-281 and -282 contain plasmid pAG7-wtI-*Sce*I, -283 and -284 contain plasmid pAG7-K223I, and -285 and -286 contain plasmid pAG7-D145A, respectively, and are mating type α. Strains which are also mating type α and in which an SSB will be made on the "Watson" strand were used in order to make comparisons of the orientation of the site between isogenic strains. For these strains, a 26-bp disruption containing the 18-bp I-*Sce*I site and two stop codons was inserted into *trp5*. Strains SAS-78 and -79 contain plasmid pAG7-wtI-*Sce*I, -80 and -148 contain plasmid pAG7-K223I, and -116 and -117 contain plasmid pAG7-D145A, respectively, and are mating type α.

Strains used for repair distant from the site of the break are derivatives of FRO-917 (*MAT*α *his3*Δ1 *leu2*Δ0 *lys2*Δ0 *ura3*Δ0 *trp5*::INS31 *lys2*::Alu IR) [52] and contain a nonfunctional *TRP5* gene disrupted by a 31-bp insert. Diploid strains SAS-150 and -151, -162 and -163, and -166 and -167 contain the CORE-wtI-*Sce*I, CORE-K223I, and CORE-

D145A cassettes, respectively, positioned approximately 10 kb upstream from the 31-bp disruption in *TRP5,* and containing the I-*Sce*I site oriented such that the SSB is made on the bottom, or "Crick", chromosomal strand. Diploid strains SAS-215 and -217, -207 and -209, and -211 and -213 contain the CORE-wtI-*Sce*I, CORE-K223I, and CORE-D145A cassettes, respectively, positioned approximately 10 kb upstream from the 31-bp disruption in *TRP5,* and containing the I-*Sce*I site oriented such that the SSB is made on the bottom, or "Watson", chromosomal strand. Diploid strains SAS-152 and -153, -154 and -156, and -158 and -160 contain the CORE-wtI-*Sce*I, CORE-K223I, and CORE-D145A cassettes, respectively, positioned approximately 10 kb downstream from the 31-bp disruption in *TRP5,* and containing the I-*Sce*I site oriented such that the SSB is made on the bottom, or "Watson", chromosomal strand. Diploid strains SAS-272 and -274, -219 and -221, and -251 and -253 contain the CORE-wtI-*Sce*I, CORE-K223I, and CORE-D145A cassettes, respectively, positioned approximately 10 kb downstream from the 31-bp disruption in *TRP5,* and containing the I-*Sce*I site oriented such that the SSB is made on the bottom, or "Crick", chromosomal strand.

The CORE-wtI-*Sce*I cassette used in this study was PCR amplified from plasmid pGSHU using primers described previously [53]. The CORE-K223I and CORE-D145A cassettes used in this study were amplified from plasmids pGSHU-K223I and pGSHU-D145A, respectively, using these same primers. The integration of the cassette into the yeast genome was carried out as previously described [53,120].

Genetic methods and standard media were described previously [112,120]. Detailed

strain construction information is provided as supporting information (Appendix B).

6.5.4 Direct repeat repair

For the assay to correct the non-functional *lys2* locus, cells were grown overnight in 5 ml yeast extract-peptone-lactic acid (YPLac) liquid medium at 30°C with shaking. Appropriate dilutions of the overnight cultures were made, and cells were plated directly to synthetic complete medium lacking lysine (SC-Lys) containing 2% galactose to express *GAL1*-I-*Sce*I and induce the break. Cells were also plated to synthetic complete (SC) medium to assess cell viability.

6.5.5 Break induction and targeting using oligos

Break induction and targeting with oligos were performed as previously described [112,120,136]. Briefly, for targeting at the *trp5* locus, 50 ml of an overnight YPLac liquid culture was inoculated with galactose (2% final concentration, unless otherwise noted) and incubated with vigorous shaking at 30°C for 4 h (7 h for experiments involving modification distant from the break) to express *GAL1*-I-*Sce*I and induce the break. After incubation in galactose, cells were prepared for transformation. Transformation was done with 1 nmol of total oligo DNA. Sequences of oligos used to repair the *trp5* gene are listed in Appendix B. Cells from each oligo transformation were diluted appropriately and spread directly to synthetic complete medium lacking tryptophan (SC-Trp) to select for oligo-mediated targeted gene correction of *trp5* or to SC to assess cell viability.

For the assay in which I-*Sce*I expression was reduced, a final concentration of 0.02%

galactose was inoculated into the overnight culture prior to shaking at 30°C for 4 h.

6.5.6 Cell cycle arrest

Mating type **a** cells were arrested in G1 as follows: α-factor (US Biological, Marblehead, MA) was transferred into a 50 ml overnight culture of YPLac for a final concentration of 0.4 μg/ml. Cultures were incubated with vigorous shaking at 30°C for 2 h. Cells were sonicated and counted. After 2 h, >80% of cells were in G1.

6.5.7 Human plasmids

Plasmid p67 (a kind gift from M. Porteus, Stanford University) contains the I-*Sce*I endonuclease gene regulated by a CMV/CBA hybrid promoter [60]. Plasmid p67-K223I, which contains the K223I gene, and plasmid p67-D145A, which contains the D145A gene, are derived from plasmid p67 and also contain the CMV/CBA hybrid promoter.

Plasmid pGRdis contains two nonfunctional reporter genes, *eGFP* and *DsRed2*, both under the CMV promoter oriented opposite from each other and separated by a 300-bp spacer. Within *eGFP* there is a 148-bp insertion containing two stop codons, a unique *Xho*I site, and the recognition sequence for the *HO* endonuclease from *S. cerevisiae*. Within *DsRed2* there is a 37-bp insertion containing two stop codons, a unique *Xba*I site, and the recognition sequence for I-*Sce*I. Additionally, the recognition sequences of *HO* and I-*Sce*I are located approximately 2.3 kb from each other. This plasmid was used for experiments assaying repair at the site of the break in *DsRed2* as well as for the human distant repair assays.

Plasmid pA658 (a kind gift from M. Porteus, Stanford University) contains a *GFP* gene disrupted by a 35-bp insert containing a single stop codon, and the recognition site for the I-*Sce*I endonuclease [60]. This plasmid was used for experiments assaying repair at the site of the break in *GFP*.

Mutagenesis and cloning information are provided as supporting information (Appendix B).

6.5.8 Human cell lines and culture

HEK-293 cells were grown in Dulbecco's modified Eagle's medium, DMEM (Mediatech, Inc., Manassas, VA), supplemented with 10% heat-inactivated fetal bovine serum (Gemini, Bio-Products, West Sacramento, CA) and 1x penicillin/streptomycin (Lonza, Walkersville, MD). Cells were grown at 37°C in a 5% CO_2 humidified incubator. Cell line 658D (kindly provided by M. Porteus, Stanford University) is a HEK-293 derivative cell line containing a randomly integrated copy of the sequence of plasmid pA658 [60]. Cells were transfected using polyethylenimine (PEI, Polysciences, Warrington, PA) transfection reagent in 24-well plates at a density of ~50,000 cells per well as previously described [128]. In all transfection experiments in HEK-293 cells, plasmid DNA was used in the amount of 0.5 μg and oligo DNA was 1.5 μg, while 1 μg each of oligo and plasmid DNA was used for transfection experiments in 658D cells. In these experiments, the oligos and the plasmid were diluted in DMEM without supplements, vortexed in the presence of PEI, and then added to the wells 20 min later. Green or red fluorescent cells were visualized by fluorescent microscopy using a Zeiss

Observer A1 microscope and an AxioCam MRm camera (Zeiss, Thornwood, NY). Frequencies of GFP^+ cells were obtained 5-8 days following transfection by flow cytometric analysis using the BD LSR II Flow Cytometer (BD Biosciences, Sparks, MD). Frequencies of RFP^+ cells were obtained 5-8 days following transfection by flow cytometric analysis using the BD FACS Aria Cell Sorter (BD Biosciences). 100,000 cells were counted for each sample. Sequences of oligos used to repair the *GFP* or *DsRed2* genes are listed in Table S1.

6.5.9 Data presentation and statistics

Graphs were created using GraphPad Prism 5 (GraphPad Software, Inc.). For all histograms, data are plotted as median values with the range shown. Statistical significance was determined by using two-tailed *t*-tests (Mann-Whitney U).

CHAPTER 7

CONCLUSIONS

Gene targeting can be effectively used to alter a genomic sequence through targeted replacement of the existing allele with a homologous sequence. This relies on the natural HR machinery of the organism. In order to trigger this machinery DNA damage can be inflicted, most commonly through generation of a DSB. However, because of the dangers of generating this type of lesion, including the potential for GCRs and gene loss, other types of DNA damage have been investigated for their recombination-inducing potential. Notably, SSBs have been proposed to efficiently stimulate HR, an activity which has been demonstrated in yeast and human cells using multiple SSB-generating enzymes, or nickases [54,81]. Moreover, generation of an SSB is thought to be a safer approach for facilitating targeted gene correction because an SSB can be easily re-ligated and produces less local mutagenesis at the site of the break than a DSB [77,78].

In this study, we aimed to characterize a novel nicking protein, I-*Sce*I K223I – which is an engineered variant of the I-*Sce*I meganuclease [108]. Because we wanted to explore the possibility of I-*Sce*I SSB-induced recombination in different organisms, we developed systems first in the yeast model organism and later in human cells. We also developed multiple types of assays in both cell types to demonstrate the recombination

potential at different episomal or chromosomal loci and at positions near or distant from the site of the break.

In yeast, strains were generated using the *delitto perfetto* system which allowed us to integrate the target recognition sequence for the I-*Sce*I enzyme into a defined genomic locus using a two-step process. Strains were isogenic in order to directly compare a break generated by the I-*Sce*I K223I nickase to a wt I-*Sce*I DSB. In Chapters 2 and 3 we detailed the *delitto perfetto* strategy, optimized steps of the *in vivo* mutagenesis approach, established new protocols, and provided examples of the types of genomic modifications which could be made using this system. These examples include gene knockouts and knockins, both of which were techniques we used to build our constructs. For example, we knocked out *RAD51* in order to demonstrate a genetic dependence on the recombinase for repair at the site of an I-*Sce*I SSB; and the knockin approach was extensively used to introduce the I-*Sce*I proteins and recognition sequence into the desired chromosomal locus.

Three types of systems were used for our yeast studies. In the first, the *LYS2* gene on chromosome II was disrupted by an insert containing the cognate 18-bp recognition sequence of I-*Sce*I between 90-bp direct repeats. Following expression of the I-*Sce*I variant and generation of the resulting break, the direct repeats could recombine to restore the sequence of the gene. The resulting phenotype was then used to score the recombination events and compare the I-*Sce*I K223I break to that of wt I-*Sce*I. In the second system, the I-*Sce*I recognition sequence was inserted into the *TRP5* gene on

chromosome VII. Following generation of the break, an exogenous repair template was introduced into the yeast cells. This template consisted of a single-stranded synthetic DNA oligo (or the complementary oligo pair) and was designed to have homology to either side of the disruption in *trp5* such that through recombination it could "pop out" the intervening nucleotides and thus restore the sequence of the gene. Again, the resulting phenotype was used to score the recombination events and compare the SSB to the DSB. In the third system, the I-*Sce*I recognition sequence was integrated into the genome ~10 kb upstream or downstream from a disrupted *trp5* gene on chromosome VII. Following generation of the distant break, oligos were introduced into cells and could recombine with the disrupted locus to restore the sequence of the gene. In Chapters 2, 3, and 5 we detailed the protocols used to test recombination at or distant from the site of the break in yeast, respectively.

In our human cell system we tested recombination at the site of the break at an episomal or chromosomal target locus as detailed in Chapter 4. Much like in the yeast system, a target gene (*RFP* or *GFP*) on a plasmid was disrupted by an insert containing the recognition sequence for I-*Sce*I. Synthetic DNA oligos were designed with homology to either side of the disruption such that they could recombine at the target locus and restore the sequence of the gene, producing a selectable phenotype in the process. We also modified this system for an assay testing recombination at a target chromosomal locus using a cell line containing a stably integrated disrupted *GFP* locus [60]. Using these systems we were able to demonstrate the recombination potential of an I-*Sce*I SSB in human cells. We also developed a system to detect modification at a target locus distant

from the site of an I-*Sce*I break in human cells. We generated a plasmid construct containing disrupted *GFP* and *RFP* genes located approximately 2 kb from each other. Each gene contains unique sites which can be cleaved by specific restriction enzymes and meganucleases. Following generation of a break at one locus, oligos transfected into the cells can recombine at the other genetic locus, restoring its sequence. The resulting phenotype can be quantified in order to determine the potential for the break to trigger recombination at the distant locus. Together these constructs and systems provide useful tools for testing the recombination potential of I-*Sce*I breaks and could be modified for other enzymes to directly compare them, as well.

Using our systems for recombination between direct repeats and for gene correction by oligos, we first demonstrated that a break generated by the I-*Sce*I K223I variant could trigger HR in yeast. In both of these systems the K223I break stimulated recombination significantly over the non-breaking control. Though the K223I variant was shown *in vitro* to generate an SSB at the cognate 18-bp recognition sequence [108], *in vivo* it was not clear if this would be the case. The *in vitro* studies demonstrated that the SSB was converted to a DSB following extended digest with the purified protein, therefore there was a possibility that the increased levels of repair we observed were not due to an I-*Sce*I SSB but instead to a DSB. To examine if the mechanism by which the I-*Sce*I SSB induced gene correction was different from that of the I-*Sce*I DSB we deleted the *RAD51* recombinase gene in both the I-*Sce*I SSB and DSB systems in yeast. Previously, deletion of *RAD51* was shown to increase levels of repair of a DSB between direct repeats and during oligo-mediated repair, likely due to suppression of DSB repair *via* strand invasion

into the corresponding sister chromatid [43,52]. We demonstrated that the break generated by I-*Sce*I K223I had a different genetic requirement than the I-*Sce*I DSB. Deletion of *RAD51* had no effect on recombination between direct repeats, and with the oligo-mediated repair levels of correction significantly decreased following expression of the K223I variant in *rad51* cells. These results indicated that recombination triggered by the I-*Sce*I SSB is stimulated by a different mechanism than recombination promoted by the I-*Sce*I DSB. Thus, we proposed that a nick can stimulate recombination without being converted into a DSB. Nevertheless, it was still possible that a low level of DSB formation, resulting from the I-*Sce*I SSB being converted into a DSB at the time as revealed by our *in vitro* data (**Figure 6.1**), stimulates recombination in a Rad51-dependent mechanism. If this hypothesis is true, we expected that reducing the expression of wt I-*Sce*I so that it stimulated recombination to the same extent of K223I would have promoted recombination in a Rad51-dependent manner. On the contrary, when we lowered expression of wt I-*Sce*I, we again observed an increase in recombinants in the absence of *RAD51*. These findings demonstrated that the break generated by I-*Sce*I K223I was different from that generated by wt I-*Sce*I, and we concluded that the I-*Sce*I SSB made by I-*Sce*I K223I directly stimulates recombination before the enzyme can convert the nick into a DSB.

Though we determined the I-*Sce*I K223I break is initially an SSB in the yeast chromosomal DNA, it still remains unknown if the lesion is subsequently converted into a DSB not by the I-*Sce*I K223I enzyme but during replication. If the SSB persists until the cells replicate, the replication fork can collapse upon encountering the unrepaired

nick and generate a DSB [12]. As expected, we observed a significant decrease in transformants following expression of wt I-*Sce*I, but this is likely due to the lack of resection while arrested in G1. Differently, we observed no change in transformants following expression of I-*Sce*I K223I. These results further supported that stimulation of recombination by I-*Sce*I K223I is not due to an inefficient DSB by the enzyme. Moreover, the G1 data could be explained by collapse of the fork as the cells progressed into S phase at the beginning of transformation, resulting in a 3-strand DSB which could then trigger recombination with the oligos. Together with the *RAD51* findings, these results indicate that an initial SSB can trigger HR in yeast and that this may be a replication-dependent event.

In the future, we could probe the DNA at different time points or during different stages of the cell cycle in order to demonstrate the condition of the lesion. We could also test other genetic requirements for repair of the I-*Sce*I K223I SSB. Candidate proteins for mutation include those involved in the checkpoint response to stalled or collapsed replication forks such as Mec1, Tel1, Ddc2, and Rad53 (the yeast homologs of the human proteins ATR, ATM, ATRIP, and CHK2, respectively) [157-159]. Comparison of I-*Sce*I SSB- to DSB-induced recombination activity in a *mec1*, *tel1*, *ddc2*, or *rad53* mutant background could yield evidence of other genetic differences between repair of the two types of lesions.

We also demonstrated that an I-*Sce*I SSB can trigger recombination in human cells at episomal and chromosomal loci. This was observed by performing numerous plasmid and

chromosomal assays in HEK-293 cells, proving that SSB-driven recombination is not limited to specific genetic sequences. Furthermore, we observed opposite strand bias targeting at one locus both at the plasmid and chromosomal position when the oligo recombination was activated by an SSB rather than by a DSB. The F oligo was favored for repair of the DSB at the *GFP* target locus, while the R oligo was favored for repair of the SSB. These findings further reveal differences in repair of the different types of lesions. In order to explore the mechanisms underlying these biases, we could change the orientation of the I-*Sce*I site such that the break generated by I-*Sce*I K223I would be on the opposite strand. Alternatively, we could test the bias within our current system using another nicking variant designed to generate a break on the opposite strand.

In addition to our findings in yeast and human cells that an I-*Sce*I SSB can stimulate recombination at the site of the break, we demonstrated the capacity of the I-*Sce*I SSB for modification at a position distant from the break. In yeast our results showed that an I-*Sce*I K223I SSB could trigger recombination at a genomic locus 10 kb away from the site of the break. Moreover, while a strand bias was observed in favor of the oligo complementary to the 3' strand when an I-*Sce*I DSB was generated, no strand bias was observed when the SSB was induced. These results again corroborated our findings that the I-*Sce*I K223I lesion that promotes recombination is, in fact, an SSB as the same bias for the oligo complementary to the strand 3' of the break would be expected if a DSB is generated by the I-*Sce*I K223I enzyme.

While we demonstrated that the I-*Sce*I SSB could trigger HR at a distant locus in yeast, we did not observe the same in human cells. It is possible that the SSB repair machinery was able to repair the lesion before any recombination with oligos could occur. We could test this hypothesis by inhibiting the poly(ADP-ribose) polymerase (PARP) family of proteins, which are involved in detecting and facilitating correction of SSBs [80], to determine if levels of recombinants increase following inhibition of SSB repair. In human cells an I-*Sce*I DSB generated *in vivo* did show a minor increase in recombinants, though an *in vitro* break did not. Perhaps introducing linearized plasmid DNA into the cells simultaneously with the oligos hinders modification at the locus distant from the broken ends. This could be due to repair of the lesion before the oligos reach the target region. In the future we could test recombination at a distant locus on the chromosome in mammalian cells similarly to our *in vivo* plasmid assay for distant modification. Following expression of the protein inside the cells, the oligos may have a better chance to reach the target locus before the break is repaired.

As gene targeting has advanced, numerous developments have been made to optimize approaches used for targeted gene correction. Overall, our findings demonstrate that an I-*Sce*I SSB can stimulate recombination in eukaryotic cells. While the nicking variant used generates breaks efficiently targeted for correction in our systems, future nickases could be designed with even greater nicking activity. However, our systems provide a foundation for the study of SSB-induced gene targeting and the findings herein should prove useful for the advancement of gene targeting approaches.

APPENDIX A

SUPPLEMENTARY MATERIALS FOR CHAPTER 2

Table A.1 Preparation of 3 M NaOAc stock solution, pH 5.2, filter sterilized
Add water to 500 ml

Component	Final concentration	Amount/500 ml
NaOAc trihydrate	3 M	204 g

Table A.2 Preparation of 10x TE stock solution, pH 7.5, filter sterilized
Add water to 500 ml

Component	Final concentration	Amount/500 ml
Tris	100 mM	6.06 g
EDTA disodium salt dihydrate	10 mM	1.86 g

Table A.3 Preparation of 1 M LiOAc stock solution, filter sterilized
Add water to 500 ml

Component	Final concentration	Amount/500 ml
LiOAc dihydrate	1 M	51.0 g

Table A.4 Preparation of 50% PEG 4000 stock solution, filter sterilized
Add water to 500 ml

Component	Final concentration	Amount/500 ml
PEG 4000	50%	250 g

Table A.5 Preparation of G418 stock solution, filter sterilized
Add water to 50 ml

Component	Final concentration	Amount/50 ml
G418	50 mg/ml	2.5 g

Table A.6 Preparation of Lyticase stock solution, stored in 1 ml aliquots
Add water to 1 ml

Component	Final concentration	Amount/1 ml
Lyticase (2000 U/ml)	2000 u/ml	1 mg

Table A.7 Preparation of 20% Galactose stock solution, filter sterilized
Add water to 1 L

Component	Final concentration	Amount/1 L
Galactose	20%	200 g

APPENDIX B

SUPPLEMENTARY MATERIALS FOR CHAPTER 6

B.1 Yeast plasmid construction

Plasmid pGSHU, described previously [102], contains the CORE-wtI-*Sce*I cassette (wt I-*Sce*I gene under the *GAL1* promoter, hygromycin resistance gene, the counterselectable *KlURA3*) and was used as a template to construct plasmids pGSHU-K223I and pGSHU-D145A. Site-directed mutagenesis was conducted using the QuikChange II Site-Directed Mutagenesis Kit (Agilent Technologies, Inc., Santa Clara, CA). Point mutations were generated by primers I-SceI K223I.F (5'-GCAGATGATGTACATTCTGCCGAACACT AT) and I-SceI K223I.R (5'-ATAGTGTTCGGCAGAATGTACATCATCTGC) to make plasmid pGSHU K223I (AAA→ATT at I-*Sce*I residue 233) or by primers I-SceI D145A.F (5'-ATACTGGTTCATGGCTGATGGTGGTAAAT) and I-SceI D145A.R (5'-ATTTACCACCATCAGCCATGAACCAGTAT) to make plasmid pI-*Sce*I D145A (GAT→GCT at I-*Sce*I residue 145) variants described by Niu *et al*. [108]. Plasmids were rescued in *Escherichia coli* and confirmed by sequence analysis.

Plasmids pAG7-wtI-*Sce*I, pAG7-K223I, and pAG7-D145A were constructed as follows. Plasmid pAG7 [54] was digested with *Bam*HI to excise the ~1.3 kb fragment containing the coding sequence of Gene *II*. A 750-bp fragment containing the coding sequence of the I-*Sce*I gene from plasmids pGSHU, pGSHU-K223I, or pGSHU-D145A was amplified through PCR using primers pAG7 clone I-SceI.F (5'-ACCCCGGATCCCGGG

GGTACATATGCATATGAAAAACATCAA) and pAG7 clone I-SceI.R (5'-CTAGAG GATCCCCGGGTACCGAGCTCGAATTCCCTTATTTCAGGAAAGTTTCGG) and cloned into the *Bam*HI site of the digested pAG7 plasmid resulting in plasmids pAG7-wtI-*Sce*I, pAG7-K223I, or pAG7-D145A, respectively. Plasmids were rescued in *E. coli* and confirmed by sequence analysis.

B.2 Yeast strain construction

The strain FRO-830 (*MATα leu2-3,112 ade5-1 his7-2 ura3Δ trp1-289 [lys2::Alu-DIRLEU2-lys2*Δ50*]*) contains the CORE-wtI-*Sce*I cassette and the 18-bp I-*Sce*I site in *lys2* within 90-bp direct repeats flanking either side of the cassette. Strains used for the direct repeat experiments are derivatives of FRO-830 and were generated as previously described [103,112,120]. Briefly, the entire CORE-wtI-*Sce*I cassette was replaced with the full I-*Sce*I site using oligos LYS2 I-SceI site.F (5'- CACTGGGTTTATCCATATGC CAAATTGAGCTAATAAATTACCCTGTTATCCCTAAAGAGAAGTGGATGGATT TGGCAAACACAGTT) and LYS2 I-SceI site.R (5'-AACTGTGTTTGCCAAATCCAT CCACTTCTCTTTAGGGATAACAGGGTAATTTATTAGCTCAATTTGGCATATGG ATAAACCCAGTG) to produce strain SAS-50. Plasmids containing the galactose-inducible *GAL1* promoter and variants of the I-*Sce*I gene were then transformed into this strain to yield the following strains: SAS-74 and -75, which contain pAG7-wtI-*Sce*I; SAS-77 and -149, which contain pAG7-K223I; and SAS-142 and -143, which contain pAG7-D145A. Additionally, in strain SAS-50 the entire coding region of *RAD51* was replaced with the *kanMX* cassette conferring antibiotic resistance to G418 for the *rad51* mutants. For the selection of G418-resistant colonies, cells were grown on yeast extract-

peptone-dextrose (YPD) medium containing 200 μg/ml of G418 (Mediatech, Inc., Manassas, VA). Plasmids were transformed into the resulting strain to yield the following *rad51* mutant strains: SAS-174 and -175, which contain pAG7-wtI-*Sce*I; SAS-176 and -177, which contain pAG7-K223I; and SAS-178 and -179, which contain pAG7-D145A. Leu^+ transformants were selected for on synthetic complete medium lacking leucine (SC-Leu) and subsequently maintained on this medium to retain the plasmid.

The strain FRO-1 (*MATα ade5-1 his7-2 leu2-3,112 ura3-52*) contains the CORE-wtI-SceI cassette and the I-*Sce*I site within *trp5*. Strains used for repair at the site of the break are derivatives of FRO-1 and were generated as previously described [103,112,120]. Briefly, the entire CORE-wtI-*Sce*I cassette was replaced with the full I-*Sce*I site oriented such that a nick created by the I-*Sce*I nickase is made on the bottom, or "Crick", chromosomal strand, using oligos TRP5 I-SceI site. F (5'-GAAGTCTTCCCAGAATGT GGGATCGATTACCCTGTTATCCCTATGATGAAAGCGACAGCTTCATCAAAAC CCTTTT) and TRP5 I-SceI site.R (5'- AAAAGGGTTTTGATGAAGCTGTCGCTTTC ATCATAGGGATAACAGGGTAATCGATCCCACATTCTGGGAAGACTT) resulting in strain SAS-59, or on the top, or "Watson", chromosomal strand, using oligos TRP5 I-SceI site opp.F (5' –AAAAGGGTTTTGATGAAGCTGTCGCATGATGAATTACCCT GTTATCCCTAACGATCCCACATTCTGGGAAGACTTC) and TRP5 I-SceI site opp.R (5' –GAAGTCTTCCCAGAATGTGGGATCGTTAGGGATAACAGGGTAATTCATC ATGCGACAGCTTCATCAAAACCCTTTT) resulting in strain SAS-278. Plasmids containing the galactose-inducible *GAL1* promoter and variants of the I-*Sce*I gene were then transformed into these strains to yield the following strains. Strains in which the I-

*Sce*I nick will be generated on the "Crick" strand: SAS-78 and -79, which contain pAG7-wtI-*Sce*I; SAS-80 and -148, which contain pAG7-K223I; and SAS-116 and -117, which contain pAG7-D145A. Strains in which the I-*Sce*I nick will be generated on the "Watson" strand: SAS-281 and -282, which contain pAG7-wtI-*Sce*I; SAS-283 and -284, which contain pAG7-K223I; and SAS-285 and -286, which contain pAG7-D145A. For strains containing the nicking orientation on the "Crick" strand, the mating type was then switched as follows: The coding region of the *BAR1* gene was replaced with an amplicon containing the *URA3* gene, and Ura^+ transformants were selected on synthetic complete medium lacking uracil (SC-Ura), yielding strain SAS-182. Plasmid pHO-LEU2, which contains the mating type switching endonuclease *HO* regulated by the inducible *GAL1* promoter and selectable *LEU2* marker, was then transformed into this strain and transformants were selected for on SC-Leu. Cells were then grown in yeast extract-peptone-galactose (YPGal) liquid medium for one hour at 30°C to allow for expression of the *HO* protein and dilutions were plated to YPD solid medium. The resulting strain, SAS-193 (*MAT***a** *ade5-1 his7-2 leu2-3,112 ura3-52, trp5*::Cr-I-SceI site), and its derivative, SAS-205, in which the entire coding region of *RAD51* was replaced with the *kanMX* cassette, were transformed with plasmids containing the galactose-inducible *GAL1* promoter and variants of the I-*Sce*I gene. Strains in which *RAD51* is wt: SAS-227 and -228, which contain pAG7-wtI-*Sce*I; SAS-229 and -230, which contain pAG7-K223I; and SAS-231 and -232, which contain pAG7-D145A. Strains in which *RAD51* is replaced: SAS-235 and -236, which contain pAG7-wtI-*Sce*I; SAS-237 and -238, which contain pAG7-K223I; and SAS-239 and -240, which contain pAG7-D145A. All strains

were grown on synthetic complete medium lacking leucine (SC-Leu) following transformation to maintain the plasmid.

The strain FRO-917 is a derivative of BY4742 (*MATα his3*Δ1 *leu2*Δ0 *lys2*Δ0 *ura3*Δ0 *trp5*::INS31 *lys2*::Alu IR) and contains a nonfunctional *trp5* gene disrupted by a 31-bp insert [52]. Strains FRO-872 [52], SAS-138, and -140 contain a CORE-wtI-*Sce*I, CORE-K223I, or CORE-D145A cassette (with the hygromycin resistance gene and I-*Sce*I site on the "Crick" strand), respectively, which was amplified as a PCR product with primers 10KbUP.IS (5-AGATAATTTACCCTTGCTTTAAGCTGCGTATATCAAGTGCATTT GCTGTCTAGGGATAACAGGGTAATTTCGTACGCTGCAGGTCGAC) and 10KbUP.II (5'-TCGTTCGTTATCCGAAGCTGGCCAATTGATACAATTAATTGAC ATCAGCATTGGATGGACGCAAAGAAGT) and cloned approximately 10 kb upstream from the 31-bp insert in *trp5*. Strains SAS-199 and -200, -195 and -196, and -197 and -198 contain a CORE-wtI-*Sce*I, CORE-K223I, or CORE-D145A cassette (with the hygromycin resistance gene and I-*Sce*I site on the "Watson" strand), respectively, which was amplified as a PCR product with primers 10KbUP.opp.IS (5' –AGATAATT TACCCTTGCTTTAAGCTGCGTATATCAAGTGCATTTGCTGTCATTACCCTGTTA TCCCTATTCGTACGCTGCAGGTCGAC) and 10KbUP.II and cloned approximately 10 kb upstream from the 31-bp insert in *trp5*. The diploid strains SAS-150 and -151, -162 and -163, -166 and -167, -215 and -217, -207 and -209, and -211 and -213 were generated by mating FRO-872 with FRO-879, SAS-138 and FRO-879, SAS-140 with FRO-879, SAS-199 or -200 with FRO-879, SAS-195 or -196 with FRO-879, and SAS-197 or -198 with FRO-879 respectively. FRO-879 is BY4741 (*MAT***a** *his3*Δ1 *leu2*Δ0 *met15*Δ0

*ura3*Δ0), where *TRP5* has been replaced by *LEU2* [52]. Strains FRO-876 [52], SAS-134 and -135, and -136 and -137 contain a CORE-wtI-*Sce*I, CORE-K223I, or CORE-D145A cassette (with the hygromycin resistance gene and I-*Sce*I site on the "Watson" strand), respectively, which was amplified as a PCR product with primers 10KbDW.IS (5'-CCGAGGTATTGTTGTATATACCACTACTCTGTGATTTTTTTTCACTCTTGTAGG GATAACAGGGTAATTTCGTACGCTGCAGGTCGAC) and 10KbDW.II (5'-TTTGT CATTAGACACTTACCAGTTGATGTTTTCACTTTTTCTTTCCTTCCTTGGATGGA CGCAAAGAAGT) and cloned approximately 10 kb downstream from the 31-bp insert in *trp5*. Strains SAS-269 and -270, SAS-201 and -202, and -245 and -246 contain a CORE-wtI-*Sce*I, CORE-K223I, or CORE-D145A cassette (with the hygromycin resistance gene and I-*Sce*I site on the "Crick" strand), respectively, which was amplified as a PCR product with primers 10KbDW.opp.IS (5' –CCGAGGTATTGTTGTATATAC CACTACTCTGTGATTTTTTTTCACTCTTGATTACCCTGTTATCCCTATTCGTAC GCTGCAGGTCGAC) and 10KbDW.II and cloned approximately 10 kb downstream from the 31-bp insert in *trp5*. The diploid strains SAS-152 and -153, -154 and -156, -158 and -160, -272 and -274, -219 and -221, and -251 and -253 were generated by mating FRO-876 with FRO-879, SAS-134 or -135 and FRO-879, SAS-136 or -137 with FRO-879, SAS-269 or -270 with FRO-879, SAS-201 or -202 with FRO-879, and SAS-245 or -246 with FRO-879, respectively.

B.3 Human plasmid construction

Plasmids pET15b ISceI and pET15b ISceI-K223I (kindly provided by F. Gimble, Purdue University) contain the wtI-*Sce*I and K223I genes described previously [108]. These

plasmids were digested with *Nsi*I and *Bam*HI to produce a ~740-bp fragment containing the gene which was cloned into pFLAG-CMV-6c (Sigma Aldrich, St. Louis, MO) between the *Bam*HI and *Pst*I sites. This yielded plasmids pSD1 (wtI-*Sce*I) and pSS1 (K223I). Plasmid pSD1 was then used as a template for site-directed mutagenesis using primers I-SceI D145A.F and I-SceI D145A.R to generate plasmid pSBO1, which contains the D145A mutation yielding the D145A gene. Plasmid p67 (a kind gift from M. Porteus, Stanford University) contains the wtI-*Sce*I endonuclease gene regulated by a CMV/CBA hybrid promoter [60] and was used as a template to construct plasmids p67-K223I and p67-D145A. First, fragments of the K223I and D145A genes from the internal *Hin*dIII site to the stop codon were amplified with primers p67clone HindIII.F (5'–TGTAGTACAGTCATACGCGCATACAAGCTTTCAACAAACTGGCTAACCTG) and p67 Clone SceI SalI.R (5'–ATCGTAGTCGTATGTCGACGAATTCTTATTATTT CAGGAAAGTTTCGGAG) and blunt-end cloned into the pCR-Blunt II-TOPO vector (Invitrogen, Carlsbad, CA). Next, these plasmids were digested with *Hin*dIII and *Sal*I to produce 400-bp fragments. Plasmid p67 was also digested with these two enzymes, and the 400-bp fragments were cloned into the vector. All plasmids were rescued in *E. coli* and confirmed by sequence analysis.

Plasmid pGRdis contains two nonfunctional reporter genes, *eGFP* and *DsRed2*, both under the CMV promoter oriented opposite from each other and separated by a 300-bp spacer. It was constructed as follows: A fragment, containing the entire coding sequence of *eGFP* and amplified as a ~720-bp amplicon through PCR using primers GFP clone pSILENCE.F (5'–AAAAAGGATCCATGGTGAGCAAGGGCGAGGA) and GFP clone

pSILENCE.R (5 –AAAAAAAGCTTTTACTTGTACAGCTCGTCCA), was digested with *Bam*HI and *Hin*dIII and cloned into plasmid pSilencer (Invitrogen) between the *Bam*HI and *Hin*dIII sites to produce plasmid pEGFP17. Primers EGFP disXho.F (5'–ACGACGGCAACTACAAGACCTGATAAGGCTCGAGCGCGCCGAGGTGAAGTT CGA) and EGFP disXho.R (5'–TCGAACTTCACCTCGGCGCGCTCGAGCCTTATCA GGTCTTGTAGTTGCCGTCGT) were used to disrupt *eGFP* between residues 109 and 110 to insert two stop codons, an *Xho*I site, and 2 frameshift bases through *in vitro* mutagenesis, producing plasmid pdisEGFP. Next, using primers 300 bp BamHI buff.F (5'–TTTCTTCTCAGGATCCGGGTCCATACATTTGCCTTT) and 300 Bam buff.R (5'–AAGCTCGATCGGATCCATCAATAATCCCCTTGGTTC) a 300-bp amplicon with *Bam*HI tails generated through PCR of part of the *URA3* gene from *Kluveromyces lactis* was inserted into plasmid pdisEGFP3 at the *Bgl*II site through compatible cloning, generating plasmid pdisEGFP300. Separately, plasmid pDsRed2-N1 (Clontech, Mountain View, CA) was modified to generate the disrupted *DsRed2* marker. Part of the MCS was removed through compatible cloning following double digestion with *Bam*HI and *Bgl*II. Next, the plasmid was linearized at *Sbf*I, and using primers ISceIstop SbfI.F (5'-ACAGTGCTAAGTCCTGCAGGACTAATAGTCTAGAATTACCCTGTTATCCCTAA GCTTCCTGCAGGTAGGAACTCAAT) and ISceIstop_SbfI.R (5'-ATTGAGTTCCTA CCTGCAGGAAGCTTAGGGATAACAGGGTAATTCTAGACTATTAGTCCTGCAG GACTTAGCACTGT) a fragment containing multiple stop codons and the 18-bp I-*Sce*I site was cloned in. Next, site-directed mutagenesis was used to remove one of the flanking *Sbf*I sites with primers DeltaSbfI.F (5' –AGAATTACCCTGTTATCCCTAAG CTTACGGCTGCTTCATCTACAAGGTGAAG) and DeltaSbfI.R (5'–CTTCACCTTGT

AGATGAAGCAGCCGTAAGCTTAGGGATAACAGGGTAATTCT). The construct was amplified through PCR with primers DsRed2-EGFPclone.F (5'–ACAGTGCTAAG TGGATCCACCGTATTACCGCCATGCAT) and DsRed2-EGFPclone.R (5'–ATTGAG TTCCTAGAATTCATCTCGGTCTATTCTTTTGA) and blunt-end cloned into the pCR-Blunt II-TOPO vector. A ~1.8 kb fragment containing the CMV promoter and disrupted *DsRed2* gene was excised by digestion with *Eco*RI and cloned into plasmid pdisEGFP300, linearized by *Eco*RI, to produce plasmid pdisEGFP300-disDsRed2. Finally, a ~150-bp fragment, containing the *HO* endonuclease recognition site and amplified through PCR using primers HO site XhoI.F (5'-ATCAAGCTAGCTCGAGG GTACGGGGATCTAAATAAATTCGTTTTCA) and HO site SalI.R (5'–CGATATCG AAGTCGACTACAACCACTCTACAAAACCAAAACCAGGG), was digested with *Xho*I and *Sal*I. Plasmid pdisEGFP300-disDsRed2 was linearized by *Xho*I, and the fragment was cloned into this site, leaving a unique *Xho*I site on the plasmid. All plasmids were rescued in *E. coli* and confirmed by sequence analysis.

Table B.1 Oligos used for repair assays

Gene	Oligo	Size	Sequence
TRP5	TRP5.80F	80-mer	5’ –GTCTAAGAGAGTTGGAAAAGGGTTTTGATGAAGCTGTC GCGGATCCCACATTCTGGGAAGACTTCAAATCCTTGTATTCT
TRP5	TRP5.80R	80-mer	5’ – AGAATACAAGGATTTGAAGTCTTCCCAGAATGTGGGAT CCGCGACAGCTTCATCAAAACCCTTTTCCAACTCTCTTAGAC
GFP	GFP80.F	80-mer	5’ –GCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGAC GCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACC
GFP	GFP80.R	80-mer	5’ – GGTTCACCAGGGTGTCGCCCTCGAACTTCACCTCGGCGC GCGTCTTGTAGTTGCCGTCGTCCTTGAAGAAGATGGTGCGC
DsRed2	DsRed2.80F	80-mer	5’ – GGCGGCGTGGCGACCGTGACCCAGGACTCCTCCCTGCAG GACGGCTGCTTCATCTACAAGGTGAAGTTCATCGGCGTGAA
DsRed2	DsRed2.80R	80-mer	5’ – TTCACGCCGATGAACTTCACCTTGTAGATGAAGCAGCCG TCCTGCAGGGAGGAGTCCTGGGTCACGGTCGCCACGCCGCC

REFERENCES

1. Saiki RK, Scharf S, Faloona F, Mullis KB, Horn GT, et al. (1985) Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science 230: 1350-1354.

2. Roth DA, Tawa NE, Jr., O'Brien JM, Treco DA, Selden RF, et al. (2001) Nonviral transfer of the gene encoding coagulation factor VIII in patients with severe hemophilia A. N Engl J Med 344: 1735-1742.

3. Kay MA, Rothenberg S, Landen CN, Bellinger DA, Leland F, et al. (1993) In vivo gene therapy of hemophilia B: sustained partial correction in factor IX-deficient dogs. Science 262: 117-119.

4. Hacein-Bey-Abina S, von Kalle C, Schmidt M, Le Deist F, Wulffraat N, et al. (2003) A serious adverse event after successful gene therapy for X-linked severe combined immunodeficiency. N Engl J Med 348: 255-256.

5. Blaese RM, Culver KW, Miller AD, Carter CS, Fleisher T, et al. (1995) T lymphocyte-directed gene therapy for ADA- SCID: initial trial results after 4 years. Science 270: 475-480.

6. Renieri A, Meloni I, Longo I, Ariani F, Mari F, et al. (2003) Rett syndrome: the complex nature of a monogenic disease. J Mol Med (Berl) 81: 346-354.

7. Sieving P, Caruso R (2008) Retinitis pigmentosa and related disorders. In: Yanoff M, Duker J, editors. Ophthalmology. 3rd ed. Maryland Heights, Mo: Mosby Elsevier.

8. Rommens JM, Iannuzzi MC, Kerem B, Drumm ML, Melmer G, et al. (1989) Identification of the cystic fibrosis gene: chromosome walking and jumping. Science 245: 1059-1065.

9. Riordan JR, Rommens JM, Kerem B, Alon N, Rozmahel R, et al. (1989) Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. Science 245: 1066-1073.

10. Kerem B, Rommens JM, Buchanan JA, Markiewicz D, Cox TK, et al. (1989) Identification of the cystic fibrosis gene: genetic analysis. Science 245: 1073-1080.

11. Scriver CR, Waters PJ (1999) Monogenic traits are not simple: lessons from phenylketonuria. Trends Genet 15: 267-272.

12. Hoffman EP, Brown RH, Jr., Kunkel LM (1987) Dystrophin: the protein product of the Duchenne muscular dystrophy locus. Cell 51: 919-928.

13. Yue P, Li Z, Moult J (2005) Loss of protein structure stability as a major causative factor in monogenic disease. J Mol Biol 353: 459-473.

14. Zufferey R, Nagy D, Mandel RJ, Naldini L, Trono D (1997) Multiply attenuated lentiviral vector achieves efficient gene delivery in vivo. Nat Biotechnol 15: 871-875.

15. Pfeifer A, Verma IM (2001) Gene therapy: promises and problems. Annu Rev Genomics Hum Genet 2: 177-211.

16. Candotti F, Shaw KL, Muul L, Carbonaro D, Sokolic R, et al. (2012) Gene therapy for adenosine deaminase-deficient severe combined immune deficiency: clinical comparison of retroviral vectors and treatment plans. Blood 120: 3635-3646.

17. Nathwani AC, Tuddenham EG, Rangarajan S, Rosales C, McIntosh J, et al. (2011) Adenovirus-associated virus vector-mediated gene transfer in hemophilia B. N Engl J Med 365: 2357-2365.

18. Carbonaro DA, Jin X, Wang X, Yu XJ, Rozengurt N, et al. (2012) Gene therapy/bone marrow transplantation in ADA-deficient mice: roles of enzyme-replacement therapy and cytoreduction. Blood 120: 3677-3687.

19. Hacein-Bey-Abina S, Le Deist F, Carlier F, Bouneaud C, Hue C, et al. (2002) Sustained correction of X-linked severe combined immunodeficiency by ex vivo gene therapy. N Engl J Med 346: 1185-1193.

20. Papapetrou EP, Lee G, Malani N, Setty M, Riviere I, et al. (2011) Genomic safe harbors permit high beta-globin transgene expression in thalassemia induced pluripotent stem cells. Nat Biotechnol 29: 73-78.

21. Silva G, Poirot L, Galetto R, Smith J, Montoya G, et al. (2011) Meganucleases and other tools for targeted genome engineering: perspectives and challenges for gene therapy. Curr Gene Ther 11: 11-27.

22. DeKelver RC, Choi VM, Moehle EA, Paschon DE, Hockemeyer D, et al. (2010) Functional genomics, proteomics, and regulatory DNA analysis in isogenic settings using zinc finger nuclease-driven transgenesis into a safe harbor locus in the human genome. Genome Res 20: 1133-1142.

23. Struhl K (1983) Direct selection for gene replacement events in yeast. Gene 26: 231-241.

24. Orr-Weaver TL, Szostak JW, Rothstein RJ (1981) Yeast transformation: a model system for the study of recombination. Proc Natl Acad Sci U S A 78: 6354-6358.

25. De Lozanne A, Spudich JA (1987) Disruption of the Dictyostelium myosin heavy chain gene by homologous recombination. Science 236: 1086-1091.

26. Manstein DJ, Titus MA, De Lozanne A, Spudich JA (1989) Gene replacement in Dictyostelium: generation of myosin null mutants. EMBO J 8: 923-932.

27. Schaefer DG, Zryd JP (1997) Efficient gene targeting in the moss Physcomitrella patens. Plant J 11: 1195-1206.

28. Strepp R, Scholz S, Kruse S, Speth V, Reski R (1998) Plant nuclear gene knockout reveals a role in plastid division for the homolog of the bacterial cell division protein FtsZ, an ancestral tubulin. Proc Natl Acad Sci U S A 95: 4368-4373.

29. Doetschman T, Maeda N, Smithies O (1988) Targeted mutation of the Hprt gene in mouse embryonic stem cells. Proc Natl Acad Sci U S A 85: 8583-8587.

30. Doetschman T, Gregg RG, Maeda N, Hooper ML, Melton DW, et al. (1987) Targetted correction of a mutant HPRT gene in mouse embryonic stem cells. Nature 330: 576-578.

31. Capecchi MR (1989) Altering the genome by homologous recombination. Science 244: 1288-1292.

32. Thomas KR, Capecchi MR (1987) Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. Cell 51: 503-512.

33. Robertson E, Bradley A, Kuehn M, Evans M (1986) Germ-line transmission of genes introduced into cultured pluripotential cells by retroviral vector. Nature 323: 445-448.

34. Kuehn MR, Bradley A, Robertson EJ, Evans MJ (1987) A potential animal model for Lesch-Nyhan syndrome through introduction of HPRT mutations into mice. Nature 326: 295-298.

35. Yanez RJ, Porter AC (1998) Therapeutic gene targeting. Gene Ther 5: 149-159.

36. Capecchi MR (2005) Gene targeting in mice: functional analysis of the mammalian genome for the twenty-first century. Nat Rev Genet 6: 507-512.

37. Symington LS (2002) Role of RAD52 epistasis group genes in homologous recombination and double-strand break repair. Microbiol Mol Biol Rev 66: 630-670, table of contents.

38. McKim KS, Hayashi-Hagihara A (1998) mei-W68 in Drosophila melanogaster encodes a Spo11 homolog: evidence that the mechanism for initiating meiotic recombination is conserved. Genes Dev 12: 2932-2942.

39. Shinohara A, Ogawa T (1995) Homologous recombination and the roles of double-strand breaks. Trends Biochem Sci 20: 387-391.

40. Dernburg AF, McDonald K, Moulder G, Barstead R, Dresser M, et al. (1998) Meiotic recombination in C. elegans initiates by a conserved mechanism and is dispensable for homologous chromosome synapsis. Cell 94: 387-398.

41. Krogh BO, Symington LS (2004) Recombination proteins in yeast. Annu Rev Genet 38: 233-271.

42. Paques F, Haber JE (1999) Multiple pathways of recombination induced by double-strand breaks in Saccharomyces cerevisiae. Microbiol Mol Biol Rev 63: 349-404.

43. Ivanov EL, Sugawara N, Fishman-Lobell J, Haber JE (1996) Genetic requirements for the single-strand annealing pathway of double-strand break repair in Saccharomyces cerevisiae. Genetics 142: 693-704.

44. Malkova A, Ivanov EL, Haber JE (1996) Double-strand break repair in the absence of RAD51 in yeast: a possible role for break-induced DNA replication. Proc Natl Acad Sci U S A 93: 7131-7136.

45. Ira G, Pellicioli A, Balijja A, Wang X, Fiorani S, et al. (2004) DNA end resection, homologous recombination and DNA damage checkpoint activation require CDK1. Nature 431: 1011-1017.

46. Silverman J, Takai H, Buonomo SB, Eisenhaber F, de Lange T (2004) Human Rif1, ortholog of a yeast telomeric protein, is regulated by ATM and 53BP1 and functions in the S-phase checkpoint. Genes Dev 18: 2108-2119.

47. DiTullio RA, Jr., Mochan TA, Venere M, Bartkova J, Sehested M, et al. (2002) 53BP1 functions in an ATM-dependent checkpoint pathway that is constitutively activated in human cancer. Nat Cell Biol 4: 998-1002.

48. Bollag RJ, Waldman AS, Liskay RM (1989) Homologous recombination in mammalian cells. Annu Rev Genet 23: 199-225.

49. Liang F, Han M, Romanienko PJ, Jasin M (1998) Homology-directed repair is a major double-strand break repair pathway in mammalian cells. Proc Natl Acad Sci U S A 95: 5172-5177.

50. Paques F, Duchateau P (2007) Meganucleases and DNA double-strand break-induced recombination: perspectives for gene therapy. Curr Gene Ther 7: 49-66.

51. Donoho G, Jasin M, Berg P (1998) Analysis of gene targeting and intrachromosomal homologous recombination stimulated by genomic double-strand breaks in mouse embryonic stem cells. Mol Cell Biol 18: 4070-4078.

52. Storici F, Snipe JR, Chan GK, Gordenin DA, Resnick MA (2006) Conservative repair of a chromosomal double-strand break by single-strand DNA through two steps of annealing. Mol Cell Biol 26: 7645-7657.

53. Storici F, Durham CL, Gordenin DA, Resnick MA (2003) Chromosomal site-specific double-strand breaks are efficiently targeted for repair by oligonucleotides in yeast. Proc Natl Acad Sci U S A 100: 14994-14999.

54. Galli A, Schiestl RH (1998) Effects of DNA double-strand and single-strand breaks on intrachromosomal recombination events in cell-cycle-arrested yeast cells. Genetics 149: 1235-1250.

55. Choulika A, Perrin A, Dujon B, Nicolas JF (1995) Induction of homologous recombination in mammalian chromosomes by using the I-SceI system of Saccharomyces cerevisiae. Mol Cell Biol 15: 1968-1973.

56. Thermes V, Grabher C, Ristoratore F, Bourrat F, Choulika A, et al. (2002) I-SceI meganuclease mediates highly efficient transgenesis in fish. Mech Dev 118: 91-98.

57. Pan FC, Chen Y, Loeber J, Henningfeld K, Pieler T (2006) I-SceI meganuclease-mediated transgenesis in Xenopus. Dev Dyn 235: 247-252.

58. Maeder ML, Thibodeau-Beganny S, Osiak A, Wright DA, Anthony RM, et al. (2008) Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification. Mol Cell 31: 294-301.

59. Urnov FD, Miller JC, Lee YL, Beausejour CM, Rock JM, et al. (2005) Highly efficient endogenous human gene correction using designed zinc-finger nucleases. Nature 435: 646-651.

60. Porteus MH, Baltimore D (2003) Chimeric nucleases stimulate gene targeting in human cells. Science 300: 763.

61. Hockemeyer D, Wang H, Kiani S, Lai CS, Gao Q, et al. (2011) Genetic engineering of human pluripotent cells using TALE nucleases. Nat Biotechnol 29: 731-734.

62. Huang P, Xiao A, Zhou M, Zhu Z, Lin S, et al. (2011) Heritable gene targeting in zebrafish using customized TALENs. Nat Biotechnol 29: 699-700.

63. Bibikova M, Beumer K, Trautman JK, Carroll D (2003) Enhancing gene targeting with designed zinc finger nucleases. Science 300: 764.

64. Christian M, Cermak T, Doyle EL, Schmidt C, Zhang F, et al. (2010) Targeting DNA double-strand breaks with TAL effector nucleases. Genetics 186: 757-761.

65. Pan H, Zhang W, Zhang W, Liu GH (2011) Find and replace: editing human genome in pluripotent stem cells. Protein Cell 2: 950-956.

66. Ortiz-Urda S, Thyagarajan B, Keene DR, Lin Q, Fang M, et al. (2002) Stable nonviral genetic correction of inherited human skin disease. Nat Med 8: 1166-1170.

67. Held PK, Olivares EC, Aguilar CP, Finegold M, Calos MP, et al. (2005) In vivo correction of murine hereditary tyrosinemia type I by phiC31 integrase-mediated gene delivery. Mol Ther 11: 399-408.

68. Groth AC, Fish M, Nusse R, Calos MP (2004) Construction of transgenic Drosophila by using the site-specific integrase from phage phiC31. Genetics 166: 1775-1782.

69. Cong L, Ran FA, Cox D, Lin S, Barretto R, et al. (2013) Multiplex Genome Engineering Using CRISPR/Cas Systems. Science.

70. Mali P, Yang L, Esvelt KM, Aach J, Guell M, et al. (2013) RNA-Guided Human Genome Engineering via Cas9. Science.

71. Obe G, Johannes C, Schulte-Frohlinde D (1992) DNA double-strand breaks induced by sparsely ionizing radiation and endonucleases as critical lesions for cell death, chromosomal aberrations, mutations and oncogenic transformation. Mutagenesis 7: 3-12.

72. Sonoda E, Hochegger H, Saberi A, Taniguchi Y, Takeda S (2006) Differential usage of non-homologous end-joining and homologous recombination in double strand break repair. DNA Repair (Amst) 5: 1021-1029.

73. Pastink A, Eeken JC, Lohman PH (2001) Genomic integrity and the repair of double-strand DNA breaks. Mutat Res 480-481: 37-50.

74. Petek LM, Russell DW, Miller DG (2010) Frequent endonuclease cleavage at off-target locations in vivo. Mol Ther 18: 983-986.

75. Bibikova M, Golic M, Golic KG, Carroll D (2002) Targeted chromosomal cleavage and mutagenesis in Drosophila using zinc-finger nucleases. Genetics 161: 1169-1175.

76. Gupta A, Meng X, Zhu LJ, Lawson ND, Wolfe SA (2011) Zinc finger protein-dependent and -independent contributions to the in vivo off-target activity of zinc finger nucleases. Nucleic Acids Res 39: 381-392.

77. Davis L, Maizels N (2011) DNA nicks promote efficient and safe targeted gene correction. PLoS One 6: e23981.

78. Metzger MJ, McConnell-Smith A, Stoddard BL, Miller AD (2011) Single-strand nicks induce homologous recombination with less toxicity than double-strand breaks using an AAV vector template. Nucleic Acids Res 39: 926-935.

79. Okano S, Lan L, Caldecott KW, Mori T, Yasui A (2003) Spatial and temporal cellular responses to single-strand breaks in human cells. Mol Cell Biol 23: 3974-3981.

80. Caldecott KW (2001) Mammalian DNA single-strand break repair: an X-ra(y)ted affair. Bioessays 23: 447-455.

81. McConnell Smith A, Takeuchi R, Pellenz S, Davis L, Maizels N, et al. (2009) Generation of a nicking enzyme that stimulates site-specific gene conversion from the I-AniI LAGLIDADG homing endonuclease. Proc Natl Acad Sci U S A 106: 5099-5104.

82. Wang J, Friedman G, Doyon Y, Wang NS, Li CJ, et al. (2012) Targeted gene addition to a predetermined site in the human genome using a ZFN-based nicking enzyme. Genome Res 22: 1316-1326.

83. Kim E, Kim S, Kim DH, Choi BS, Choi IY, et al. (2012) Precision genome engineering with programmable DNA-nicking enzymes. Genome Res 22: 1327-1333.

84. O'Driscoll M, Jeggo PA (2008) The role of the DNA damage response pathways in brain development and microcephaly: insight from human disorders. DNA Repair (Amst) 7: 1039-1050.

85. Radding CM (1982) Homologous pairing and strand exchange in genetic recombination. Annu Rev Genet 16: 405-437.

86. Meselson MS, Radding CM (1975) A general model for genetic recombination. Proc Natl Acad Sci U S A 72: 358-361.

87. McBlane JF, van Gent DC, Ramsden DA, Romeo C, Cuomo CA, et al. (1995) Cleavage at a V(D)J recombination signal requires only RAG1 and RAG2 proteins and occurs in two steps. Cell 83: 387-395.

88. van Nierop GP, de Vries AA, Holkers M, Vrijsen KR, Goncalves MA (2009) Stimulation of homology-directed gene targeting at an endogenous human locus by a nicking endonuclease. Nucleic Acids Res 37: 5725-5736.

89. Strathern JN, Weinstock KG, Higgins DR, McGill CB (1991) A novel recombinator in yeast based on gene II protein from bacteriophage f1. Genetics 127: 61-73.

90. Ramirez CL, Certo MT, Mussolino C, Goodwin MJ, Cradick TJ, et al. (2012) Engineered zinc finger nickases induce homology-directed repair with reduced mutagenic effects. Nucleic Acids Res 40: 5560-5568.

91. Dugle DL, Gillespie CJ, Chapman JD (1976) DNA strand breaks, repair, and survival in x-irradiated mammalian cells. Proc Natl Acad Sci U S A 73: 809-812.

92. Pease AC, Solas D, Sullivan EJ, Cronin MT, Holmes CP, et al. (1994) Light-generated oligonucleotide arrays for rapid DNA sequence analysis. Proc Natl Acad Sci U S A 91: 5022-5026.

93. Hardenbol P, Baner J, Jain M, Nilsson M, Namsaraev EA, et al. (2003) Multiplexed genotyping with sequence-tagged molecular inversion probes. Nat Biotechnol 21: 673-678.

94. Alm EW, Oerther DB, Larsen N, Stahl DA, Raskin L (1996) The oligonucleotide probe database. Appl Environ Microbiol 62: 3557-3559.

95. Yu D, Ellis HM, Lee EC, Jenkins NA, Copeland NG, et al. (2000) An efficient recombination system for chromosome engineering in Escherichia coli. Proc Natl Acad Sci U S A 97: 5978-5983.

96. Zhang Y, Muyrers JP, Rientjes J, Stewart AF (2003) Phage annealing proteins promote oligonucleotide-directed mutagenesis in Escherichia coli and mouse ES cells. BMC Mol Biol 4: 1.

97. Knauert MP, Glazer PM (2001) Triplex forming oligonucleotides: sequence-specific tools for gene targeting. Hum Mol Genet 10: 2243-2251.

98. Radecke F, Radecke S, Schwarz K (2004) Unmodified oligodeoxynucleotides require single-strandedness to induce targeted repair of a chromosomal EGFP gene. J Gene Med 6: 1257-1271.

99. Radecke F, Peter I, Radecke S, Gellhaus K, Schwarz K, et al. (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.

100. Igoucheva O, Alexeev V, Yoon K (2001) Targeted gene correction by small single-stranded oligonucleotides in mammalian cells. Gene Ther 8: 391-399.

101. 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.

102. Storici F, Lewis LK, Resnick MA (2001) In vivo site-directed mutagenesis using oligonucleotides. Nat Biotechnol 19: 773-776.

103. Storici F, Resnick MA (2003) Delitto perfetto targeted mutagenesis in yeast with oligonucleotides. Genet Eng (N Y) 25: 189-207.

104. Storici F, Resnick MA (2006) The delitto perfetto approach to in vivo site-directed mutagenesis and chromosome rearrangements with synthetic oligonucleotides in yeast. Methods Enzymol 409: 329-345.

105. Kanai F, Shiratori Y, Yoshida Y, Wakimoto H, Hamada H, et al. (1996) Gene therapy for alpha-fetoprotein-producing human hepatoma cells by adenovirus-mediated transfer of the herpes simplex virus thymidine kinase gene. Hepatology 23: 1359-1368.

106. Belfort M, Roberts RJ (1997) Homing endonucleases: keeping the house in order. Nucleic Acids Res 25: 3379-3388.

107. Dalgaard JZ, Klar AJ, Moser MJ, Holley WR, Chatterjee A, et al. (1997) Statistical modeling and analysis of the LAGLIDADG family of site-specific endonucleases and identification of an intein that encodes a site-specific endonuclease of the HNH family. Nucleic Acids Res 25: 4626-4638.

108. Niu Y, Tenney K, Li H, Gimble FS (2008) Engineering variants of the I-SceI homing endonuclease with strand-specific and site-specific DNA-nicking activity. J Mol Biol 382: 188-202.

109. Arnould S, Delenda C, Grizot S, Desseaux C, Paques F, et al. (2011) The I-CreI meganuclease and its engineered derivatives: applications from cell modification to gene therapy. Protein Eng Des Sel 24: 27-31.

110. Chan SH, Stoddard BL, Xu SY (2011) Natural and engineered nicking endonucleases--from cleavage mechanism to engineering of strand-specificity. Nucleic Acids Res 39: 1-18.

111. Haugen P, Simon DM, Bhattacharya D (2005) The natural history of group I introns. Trends Genet 21: 111-119.

112. Stuckey S, Storici F (2010) Gene knockouts, *in vivo* site-directed mutagenesis and other modifications using the *delitto perfetto* system in *Saccharomyces cerevisiae*. In: Lorsch J, editor. Methods Navigator. Oxford, UK: Elsevier Press.

113. Wach A, Brachat A, Pohlmann R, Philippsen P (1994) New heterologous modules for classical or PCR-based gene disruptions in Saccharomyces cerevisiae. Yeast 10: 1793-1808.

114. Bach ML, Lacroute F (1972) Direct selective techniques for the isolation of pyrimidine auxotrophs in yeast. Mol Gen Genet 115: 126-130.

115. Chattoo BB, Sherman F, Azubalis DA, Fjellstedt TA, Mehnert D, et al. (1979) Selection of lys2 Mutants of the Yeast SACCHAROMYCES CEREVISIAE by the Utilization of alpha-AMINOADIPATE. Genetics 93: 51-65.

116. Singh A, Sherman F (1974) Characteristics and relationships of mercury-resistant mutants and methionine auxotrophs of yeast. J Bacteriol 118: 911-918.

117. Toyn JH, Gunyuzlu PL, White WH, Thompson LA, Hollis GF (2000) A counterselection for the tryptophan pathway in yeast: 5-fluoroanthranilic acid resistance. Yeast 16: 553-560.

118. Storici F, Coglievina M, Bruschi CV (1999) A 2-microm DNA-based marker recycling system for multiple gene disruption in the yeast Saccharomyces cerevisiae. Yeast 15: 271-283.

119. Sauer B (1987) Functional expression of the cre-lox site-specific recombination system in the yeast Saccharomyces cerevisiae. Mol Cell Biol 7: 2087-2096.

120. Stuckey S, Mukherjee K, Storici F (2011) In vivo site-specific mutagenesis and gene collage using the delitto perfetto system in yeast Saccharomyces cerevisiae. Methods Mol Biol 745: 173-191.

121. Inga A, Resnick MA (2001) Novel human p53 mutations that are toxic to yeast can enhance transactivation of specific promoters and reactivate tumor p53 mutants. Oncogene 20: 3409-3419.

122. Goldstein AL, McCusker JH (1999) Three new dominant drug resistance cassettes for gene disruption in Saccharomyces cerevisiae. Yeast 15: 1541-1553.

123. Sherman F (2002) Getting started with yeast. Methods Enzymol 350: 3-41.

124. Dujon B (1996) The yeast genome project: what did we learn? Trends Genet 12: 263-270.

125. Oliver SG (1996) From DNA sequence to biological function. Nature 379: 597-600.

126. Winzeler EA, Davis RW (1997) Functional analysis of the yeast genome. Curr Opin Genet Dev 7: 771-776.

127. Resnick MA, Cox BS (2000) Yeast as an honorary mammal. Mutat Res 451: 1-11.

128. Shen Y, Nandi P, Taylor MB, Stuckey S, Bhadsavle HP, et al. (2011) RNA-driven genetic changes in bacteria and in human cells. Mutat Res 717: 91-98.

129. Lakhani SR, Jacquemier J, Sloane JP, Gusterson BA, Anderson TJ, et al. (1998) Multifactorial analysis of differences between sporadic breast cancers and cancers involving BRCA1 and BRCA2 mutations. J Natl Cancer Inst 90: 1138-1145.

130. Bodmer W, Bonilla C (2008) Common and rare variants in multifactorial susceptibility to common diseases. Nat Genet 40: 695-701.

131. Richardson C, Moynahan ME, Jasin M (1998) Double-strand break repair by interchromosomal recombination: suppression of chromosomal translocations. Genes Dev 12: 3831-3842.

132. Helleday T, Lo J, van Gent DC, Engelward BP (2007) DNA double-strand break repair: from mechanistic understanding to cancer treatment. DNA Repair (Amst) 6: 923-935.

133. Porteus MH, Cathomen T, Weitzman MD, Baltimore D (2003) Efficient gene targeting mediated by adeno-associated virus and DNA double-strand breaks. Mol Cell Biol 23: 3558-3565.

134. Lu IL, Lin CY, Lin SB, Chen ST, Yeh LY, et al. (2003) Correction/mutation of acid alpha-D-glucosidase gene by modified single-stranded oligonucleotides: in vitro and in vivo studies. Gene Ther 10: 1910-1916.

135. Nickerson HD, Colledge WH (2003) A comparison of gene repair strategies in cell culture using a lacZ reporter system. Gene Ther 10: 1584-1591.

136. Stuckey S, Storici F (2013) Genetic Modification Stimulated by the Induction of a Site-Specific Break Distant from the Locus of Correction in Haploid and Diploid Yeast *Saccharomyces cerevisiae*. Methods Mol Biol: In press.

137. Rothstein R (1991) Targeting, disruption, replacement, and allele rescue: integrative DNA transformation in yeast. Methods Enzymol 194: 281-301.

138. Rong YS, Golic KG (2000) Gene targeting by homologous recombination in Drosophila. Science 288: 2013-2018.

139. Vasquez KM, Marburger K, Intody Z, Wilson JH (2001) Manipulating the mammalian genome by homologous recombination. Proc Natl Acad Sci U S A 98: 8403-8410.

140. Kumar S, Allen GC, Thompson WF (2006) Gene targeting in plants: fingers on the move. Trends Plant Sci 11: 159-161.

141. Takata M, Sasaki MS, Sonoda E, Morrison C, Hashimoto M, et al. (1998) Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells. EMBO J 17: 5497-5508.

142. Schaefer DG (2001) Gene targeting in Physcomitrella patens. Curr Opin Plant Biol 4: 143-150.

143. Yang Y, Sterling J, Storici F, Resnick MA, Gordenin DA (2008) Hypermutability of damaged single-strand DNA formed at double-strand breaks and uncapped telomeres in yeast Saccharomyces cerevisiae. PLoS Genet 4: e1000264.

144. Lee SE, Pellicioli A, Demeter J, Vaze MP, Gasch AP, et al. (2000) Arrest, adaptation, and recovery following a chromosome double-strand break in Saccharomyces cerevisiae. Cold Spring Harb Symp Quant Biol 65: 303-314.

145. Rahn JJ, Rowley B, Lowery MP, Coletta LD, Limanni T, et al. (2011) Effects of varying gene targeting parameters on processing of recombination intermediates by ERCC1-XPF. DNA Repair (Amst) 10: 188-198.

146. Perez C, Guyot V, Cabaniols JP, Gouble A, Micheaux B, et al. (2005) Factors affecting double-strand break-induced homologous recombination in mammalian cells. Biotechniques 39: 109-115.

147. Miller JC, Holmes MC, Wang J, Guschin DY, Lee YL, et al. (2007) An improved zinc-finger nuclease architecture for highly specific genome editing. Nat Biotechnol 25: 778-785.

148. Karran P (2000) DNA double strand break repair in mammalian cells. Curr Opin Genet Dev 10: 144-150.

149. Bennett CB, Lewis AL, Baldwin KK, Resnick MA (1993) Lethality induced by a single site-specific double-strand break in a dispensable yeast plasmid. Proc Natl Acad Sci U S A 90: 5613-5617.

150. Jackson SP, Bartek J (2009) The DNA-damage response in human biology and disease. Nature 461: 1071-1078.

151. McDonald JP, Rothstein R (1994) Unrepaired heteroduplex DNA in Saccharomyces cerevisiae is decreased in RAD1 RAD52-independent recombination. Genetics 137: 393-405.

152. Kim PM, Paffett KS, Solinger JA, Heyer WD, Nickoloff JA (2002) Spontaneous and double-strand break-induced recombination, and gene conversion tract lengths, are differentially affected by overexpression of wild-type or ATPase-defective yeast Rad54. Nucleic Acids Res 30: 2727-2735.

153. Lam KN, van Bakel H, Cote AG, van der Ven A, Hughes TR (2011) Sequence specificity is obtained from the majority of modular C2H2 zinc-finger arrays. Nucleic Acids Res 39: 4680-4690.

154. Kleinstiver BP, Wolfs JM, Kolaczyk T, Roberts AK, Hu SX, et al. (2012) Monomeric site-specific nucleases for genome editing. Proc Natl Acad Sci U S A 109: 8061-8066.

155. Di Virgilio M, Callen E, Yamane A, Zhang W, Jankovic M, et al. (2013) Rif1 prevents resection of DNA breaks and promotes immunoglobulin class switching. Science 339: 711-715.

156. Zimmermann M, Lottersberger F, Buonomo SB, Sfeir A, de Lange T (2013) 53BP1 regulates DSB repair using Rif1 to control 5' end resection. Science 339: 700-704.

157. Cha RS, Kleckner N (2002) ATR homolog Mec1 promotes fork progression, thus averting breaks in replication slow zones. Science 297: 602-606.

158. Cortez D, Guntuku S, Qin J, Elledge SJ (2001) ATR and ATRIP: partners in checkpoint signaling. Science 294: 1713-1716.

159. Lopes M, Cotta-Ramusino C, Pellicioli A, Liberi G, Plevani P, et al. (2001) The DNA replication checkpoint response stabilizes stalled replication forks. Nature 412: 557-561.