Homology Requirements in Mammalian
Early Homologous Recombination

by

Vatsal Desai

A Thesis
presented to
The University of Guelph

In partial fulfillment of requirements
for the degree of
Master of Science
in
Molecular and Cellular Biology

Guelph, Ontario, Canada
© Vatsal Desai, April, 2013
ABSTRACT

HOMOLOGY REQUIREMENTS IN MAMMALIAN EARLY HOMOLOGOUS RECOMBINATION

Vatsal Desai  
University of Guelph, 2013

Supervisor:  
Dr. Mark D. Baker

Homologous recombination (HR) is a precise mechanism for repairing harmful DNA double-strand breaks. The process has been extensively studied in microbial species leading to identification of the major proteins, HR models and homology requirements. Much less is known about HR in mammalian systems, especially early HR events. Our laboratory has recently developed an assay that detects the new DNA synthesis that accompanies the early homology search and strand invasion steps of HR (the 3’ extension assay). The hypothesis that homology requirements for the early steps of HR may differ from those identified in other HR assays was tested.

Plasmids bearing varying amounts of homology to the chromosomal immunoglobulin μ target locus gene were constructed and tested in the 3’ extension assay. The homology requirements for the 3’ extension assay were somewhat lower than might be expected based on other HR assays. An approximately linear relationship between homology length and 3’ extension was also established on each side of the double-strand break. The effect of excess Rad51, an essential protein involved in early HR, was also measured with respect to homology, leading to the discovery that increased Rad51 resulted in an increase in 3’ extension events independent of homology. In summary, 3’ extension generates a potentially unstable, short-lived HR intermediate that has less dependence on homology than a completed HR product. Homology plays a role in the initiation of HR, but it may be more important in the stabilization of the intermediate than the actual generation of the early HR product detected in the 3’ extension assay.
Acknowledgements

First and foremost, I would like to thank my supervisor, Dr. Mark Baker, for providing the opportunity to work with him, providing knowledge, advice and guidance, especially early on, when things seemed uncertain. I would also like to thank my committee members, Dr. Yankulov and Dr. Lu for additional guidance and advice with studies and habits.

For help with techniques, getting through day-to-day activities and the laughs I’d like to thank the members of the Baker lab: Alissa, thanks for teaching me so many things and for just being a generally cool person; Iulia, for sharing the struggle with me and for help in writing; Jenn, for being relatable and fun to talk with; Taryn for brightening the lab with your presence; Maureen, for teaching me the 3’ end extension assay and also for help with being a graduate student and creating jokes.

Finally, I’d like to thank my parents for supporting me completely throughout this endeavour and with everything else I’ve ever done.
Table of Contents

Acknowledgements ................................................................................................................ iii
List of Figures ........................................................................................................................ vii
List of Tables ........................................................................................................................ viii
List of Abbreviations .............................................................................................................. ix
Introduction ............................................................................................................................. 1

CHAPTER 1: Literature Review ............................................................................................... 2
  1.1 DNA double-strand breaks and their detection ............................................................... 2
  1.2 Pathways of double-strand break repair ........................................................................ 4
    A. Non-homologous end–joining ......................................................................................... 5
    Classical non-homologous end-joining ............................................................................. 5
    Alternative non-homologous end-joining ......................................................................... 6
    B. Homologous recombination mechanisms ...................................................................... 7
    Holliday model .................................................................................................................... 7
    Meselson-Radding model .................................................................................................. 8
    Double-strand break repair ............................................................................................... 10
    Synthesis-dependent strand annealing ............................................................................. 12
    Break-induced replication ................................................................................................. 14
    Single-strand annealing .................................................................................................... 16
  1.3 Homologous recombination proteins .......................................................................... 18
  1.4. Other pathways of DNA repair ...................................................................................... 22
    Nucleotide excision repair ................................................................................................. 22
    Base excision repair .......................................................................................................... 22
    Direct repair of DNA damage ........................................................................................... 23
    Mismatch repair ................................................................................................................ 24
  1.5 Role of homology in recombination .............................................................................. 25
  1.6 Hybridoma cell lines ..................................................................................................... 27
  1.7 3’ end extension assay ................................................................................................... 29

Hypothesis ............................................................................................................................... 32

 CHAPTER 2: Materials and Methods ..................................................................................... 33
  2.1 Hybridoma cell lines ..................................................................................................... 33
Sp6/HL .................................................................................................................. 33
igm482 .................................................................................................................. 33
51-5 ....................................................................................................................... 34

2.2 Plasmids ........................................................................................................ 34

General conditions ............................................................................................. 34

Vectors used in 3’ extension ............................................................................... 34

pTΔCμ858/2290 ..................................................................................................... 35

Left Arm Constructs .......................................................................................... 38

pTΔCμ2120/2290 .................................................................................................. 38
pTΔCμ2070/2290 .................................................................................................. 38
pTΔCμ858/2290X ................................................................................................. 39
pTΔCμ720/2290 ................................................................................................... 39
pTΔCμ517/2290 ................................................................................................... 39
pTΔCμ211/2290 .................................................................................................. 40

Right Arm Constructs ......................................................................................... 40

pTΔCμ858/1294 .................................................................................................. 40
pTΔCμ858/569 ................................................................................................... 40
pTΔBamCμ ........................................................................................................... 41

2.3 Hybridoma transfection and plasmid extraction .............................................. 43

2.4 PCR Analysis .................................................................................................. 44

CHAPTER 3: Results ............................................................................................ 47

3.1 Optimizing the amount of vector DNA for 3’ extension .............................. 47

3.2 3’ end extension events on the same side of the DSG increase linearly with homology ... 51

3.3 Effect of homology changes on one arm and their influence on 3’ extension from the opposite arm ................................................................. 57

3.4 Length of 3’ extensions .................................................................................. 57

3.5 Excess Rad51 increases 3’ extension events .................................................. 64

3.6 Effect of gap size on 3’ extensions ................................................................ 67

CHAPTER 4: Discussion ....................................................................................... 69

4.1 DNA titration studies in the 3’ extension assay ............................................... 69

4.2 The effect of homology on 3’ extension ......................................................... 70
4.3 The effect of opposite arm homology on 3’ extension............................................ 73
4.4 Effect of homology on the length of 3’ extension products...................................... 75
4.5 Effect of increased Rad51 and homology on the efficiency of 3’ extension ............... 76
4.6 3’ extension efficiency when the size of the vector-borne DSG is reduced .......... 78
4.7 A model for 3’ extensions......................................................................................... 79
CHAPTER 5: Future Studies ......................................................................................... 82
Literature Cited ........................................................................................................... 84
Appendix I .................................................................................................................... 95
  Left arm ANOVA & linear correlation....................................................................... 95
  Right arm ANOVA & linear correlation................................................................. 96
  Opposite Left arm & Opposite Right arm ANOVA ............................................... 97
  51-5 vs igm 482.................................................................................................. 98
  pTΔCμ720/2290 vs. pTΔCμ211/2290 Rad51 homology ratio comparison ............. 99
  Gap size left & right extension comparisons ..................................................... 100
  Gap ratio left vs right extensions....................................................................... 101
Appendix II ................................................................................................................ 102
  Solutions.............................................................................................................. 102
List of Figures

Figure 1.1 Holliday model. .............................................................................................................. 9
Figure 1.2 Double-strand break repair. .............................................................................................. 11
Figure 1.3 Synthesis-dependent strand annealing ............................................................................... 13
Figure 1.4 Break-induced replication. .................................................................................................. 15
Figure 1.5 Single-strand annealing. ................................................................................................... 17
Figure 1.6 Hybridoma cell lines ......................................................................................................... 28
Figure 1.7 3’ extension primer binding sites. ....................................................................................... 31
Figure 2.1 Map of pTΔCμ vector ........................................................................................................ 36
Figure 2.2 Vector homology segments used in the 3’ extension assay. ................................................. 37
Figure 2.3 Construction of a 600 bp double-stranded gap in the vector pTΔBamCμ. ............................ 42
Figure 3.1 Titrating the concentration of the pTΔCμ_{858/2290} vector ................................................ 49
Figure 3.2 Summary of pTΔCμ_{858/2290} DNA titrations .................................................................. 50
Figure 3.3 Representative gel of 3’ extension from the left invading arm ........................................... 52
Figure 3.4 Representative gel of 3’ extension from the right invading arm ........................................ 54
Figure 3.5 Role of homology on 3’ extension ....................................................................................... 56
Figure 3.6 Representative gels of left arm 3’ extensions occurring with homology changes to the right arm ................................................................................................................................. 58
Figure 3.7 Representative gels of right arm 3’ extensions occurring with homology changes to the left arm ......................................................................................................................................................... 59
Figure 3.8 3’ extensions measured opposite a homology change ......................................................... 61
Figure 3.9 Length of left arm 3’ extension products. ............................................................................ 63
Figure 3.10 Representative gels of 3’ extensions in 51-5 ................................................................. 65
Figure 3.11 Effect of excess Rad51 and homology length on 3’ extension .......................................... 66
Figure 3.12 Influence of gap size on 3’ extension events ................................................................. 68
Figure 4.1 Possible generation of newly synthesized DNA. .............................................................. 80
List of Tables

Table 2.1 Primers used in 3’ extension assay ................................................................. 45
Table 2.2 Protocol for PCR amplification of 3’ extension samples ............................... 45
Table 2.3 Summary of 3’ extension products ................................................................. 46
Table 3.1 Frequency of 3’ extension products of varying length ................................... 62
List of Abbreviations

ANOVA  analysis of variance
ATM  ataxia-telangiectasia mutated
ATR  ataxia-telangiectasia and Rad3-related
BER  base excision repair
BIR  break-induced replication
bp  basepair
BSA  bovine serum albumin
Cμ  immunoglobulin μ gene constant region
DMEM  Dulbecco’s modified Eagle’s medium
DSB  double-strand break
DSBR  double-strand break repair
dsDNA  double-stranded DNA
DSG  double-stranded gap
HJ  Holliday junction
HR  homologous recombination
IgH  immunoglobulin heavy chain
IgM  immunoglobulin M
kb  kilobase
LB  Luria-Bertani
MAT  mating-type
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMR</td>
<td>mismatch repair</td>
</tr>
<tr>
<td>NER</td>
<td>nucleotide excision repair</td>
</tr>
<tr>
<td>NHEJ</td>
<td>non-homologous end joining</td>
</tr>
<tr>
<td>PBS</td>
<td>phosphate-buffered saline</td>
</tr>
<tr>
<td>RPA</td>
<td>replication protein A</td>
</tr>
<tr>
<td>SDSA</td>
<td>synthesis-dependent strand annealing</td>
</tr>
<tr>
<td>Sμ</td>
<td>μ gene switch region</td>
</tr>
<tr>
<td>SSA</td>
<td>single strand annealing</td>
</tr>
<tr>
<td>ssDNA</td>
<td>single-stranded DNA</td>
</tr>
<tr>
<td>tk</td>
<td>thymidine kinase</td>
</tr>
<tr>
<td>TNP</td>
<td>trinitrophenyl</td>
</tr>
<tr>
<td>TLS</td>
<td>translesion DNA synthesis</td>
</tr>
<tr>
<td>V_{H}</td>
<td>heavy chain variable region</td>
</tr>
<tr>
<td>UV</td>
<td>ultraviolet</td>
</tr>
</tbody>
</table>
Introduction

Genome maintenance is integral to cell viability and future growth. DNA damage can occur through exogenous sources such as ultraviolet light and ionizing radiation, or through natural cellular processes. Many different mechanisms have evolved to repair the various types of DNA damage a cell may acquire (Pardo et al., 2009).

An unrepaired DNA double-strand break (DSB) can lead to chromosome loss, major aberrations, or apoptosis (Sonoda et al., 1998). There are two major routes for the repair of a DSB, the distinction being their dependence on homology for repair. Homologous recombination (HR) is used to repair DSBs with a homologous partner sequence serving as the template for repair, while non-homologous end-joining joins two broken DNA strands with minimal regard for homology.

Previous studies have established that mammalian gene targeting requires a minimum of approximately 500 bp of homology for efficient recombination and demonstrated a relationship between homology and HR events (Shulman et al., 1990; Hasty et al., 1991; Deng and Capecchi, 1992). Studies of ectopic HR in mammalian cells suggest that a homology of at least 1.9 kb is required (Baker et al., 1996).

The homology requirements involved in the early stages of HR are under examination in this thesis, utilizing a novel HR assay that measures new DNA synthesis during the early strand invasion step of HR in vivo (Si et al., 2010), as these homology requirements may differ from those established in previous HR assays. Additionally, the effect of gap length and the effect of HR protein Rad51 during early HR are studied.
CHAPTER 1: Literature Review

1.1 DNA double-strand breaks and cellular detection

DNA damage, in particular DNA double-strand breaks (DSB), can seriously compromise genomic integrity and by extension, cell viability and function. A DSB may result from nucleases, physical stress (e.g., during mitosis, chromosomes can be pulled to opposite poles), or the products of metabolic processes, such as reactive oxygen species (Pardo et al., 2009). Replication forks that may be stalled due to a damaging lesion can be restarted through the introduction of a DSB (Hanada et al., 2007). A DSB may also be caused directly or indirectly by exogenous sources, such as ionizing radiation, ultraviolet light or various chemicals and chemotherapeutic agents (such as mitomycin C) (Dusre et al., 1989; Pardo et al., 2009).

Programmed DSBs are also able to generate genetic diversity. For example, in the yeast Saccharomyces cerevisiae, the Spo11 meiosis-specific topoisomerase causes DSBs and is used to promote crossover during prophase I, allowing correct segregation of homologous chromosomes (Neale and Keeney, 2006). Yeast also undergo mating-type switching, which occurs via a programmed DSB induced at the mating-type (MAT) locus (Haber, 1992). Immunodiversity in B and T lymphocytes is generated in part by the repair of DSBs in immunoglobulin and T cell receptor genes, respectively (Dudley et al., 2005). One-sided DSBs, where only one free DNA end is present, such as at the end of a chromosome, are also generated in the process of telomere maintenance, upon telomere degradation and loss of telomerase (Dunham, 2000).

An unrepaired DSB is a severe form of DNA damage that can lead to chromosomal loss, major chromosomal aberrations (i.e., translocations, deletions, duplications and inversions) and problems during chromosome segregation in mitosis or apoptosis (Sonoda et al., 1998; Li and
Heyer, 2008). For cell viability, it is essential that the damage be detected and a repair pathway initiated to mend the DSB.

A cell needs to be able to detect a DSB accurately, halt the cell cycle, and establish a proper mechanism for DSB repair, or failing that, initiate a pathway for apoptosis or cell senescence. Detection of a DSB and determination of an effector pathway is a multistep process. The pathway choice for repair of a DSB may be controlled through competition; the Ku70/80 heterodimer and the Mre11-Rad50-Nbs1 (MRN) complex in mammals (the MRX complex in yeast) will both attempt to bind a DSB (Walker et al., 2001; Yang et al., 2007). The Ku70/80 heterodimer will lead to DSB repair as part of the non-homologous end joining (NHEJ) repair mechanism, while binding via the MRN complex will lead to repair by homologous recombination (HR). Both modes of recombinational repair are discussed below.

In the MRN complex, Mre11 along with Ctp1 function as 5'→3' endonucleases, creating initial single-stranded DNA overhangs, which are used in the HR repair process (Williams et al., 2007; Langerak et al., 2011). Extended DNA resection will occur due to the combined action of the Sgs1 helicase and the Dna2 endonuclease (Zhu et al., 2008). Alternatively, the Exo1 exonuclease can act independently of Sgs1 and Dna2 (often in their absence), to achieve extended resection (Bolderson et al., 2010). The role of Rad50 is to create scaffolding for the damaged DNA and along with Mre11, it is involved in binding DNA (Williams et al., 2007). Nbs1 is necessary for Rad50 foci formation, Ctp1 binding, and the recruitment and activation of the ATM (Ataxia Telangiectasia Mutated) checkpoint kinase (Su, 2006; Yang et al., 2007). ATM is necessary for the phosphorylation of histone H2AX (termed γH2AX upon phosphorylation of Serine 139), which in turn recruits more ATM, resulting in a cascade of DNA damage response proteins (Su, 2006).
While this cascade is occurring, Replication Protein A (RPA) will bind to ssDNA which has been generated at the DSB. RPA then mediates localization of the ATR (ATM and Rad3-related) kinase to the single-stranded DNA (ssDNA) at the DSB and may also be hyper-phosphorylated by the kinase (Zou and Elledge, 2003). The ATR kinase is then able to phosphorylate Rad17, which will recruit the 9-1-1 clamp repair complex to the DNA damage (reviewed in Harrison and Haber, 2006), which is involved in increasing DNA polymerase processivity during telomere maintenance (Boerckel et al., 2007). ATR and ATM signaling complexes also target cell cycle regulators, such as Chk1 and Chk2, which may halt the cell cycle during the S or G2 phase to allow for DSB repair or replication fork restart, or if the cell is unable to repair the DSB, the ATR kinase is also able to phosphorylate p53, potentially triggering the apoptosis pathway (Zou and Elledge, 2003). Various pathways have evolved to repair cellular DSBs and they are described next.

1.2 Pathways of double-strand break repair

Cells have evolved two main pathways for repairing a DNA DSB, the non-homologous end-joining (NHEJ) pathway and the homologous recombination (HR) pathway. The determination of which repair pathway is used seems to be dependent on the stage of the cell cycle, the state of the DSB and which complexes/proteins are bound to the DSB (Allen et al., 2002; Thompson, 2012).
A. Non-homologous end-joining

Two non-homologous end-joining pathways have been defined: the classical NHEJ pathway and a second alternate NHEJ pathway (ALT-NHEJ) are described below. The NHEJ recombination repair pathway is essential for cell viability, since deficiencies in NHEJ-specific proteins, such as Ku80, lead to an increase in radiation sensitivity, chromosomal aberrations and cell senescence (Difilippantonio et al., 2000). The NHEJ mechanism is also essential for generating diversity during V(D)J recombination in T and B lymphocytes (Nussenzweig et al., 1996). During the G1 phase of the cell cycle, binding of the Ku70/80 heterodimer to the DSB, prevents 3’ DNA end resection by Mre11 and Ctp1, which is a necessary step for the initiation of HR (Barlow et al., 2008).

Classical non-homologous end-joining

The NHEJ designation may be slightly misleading, as the pathway often uses a region of micro-homology (~1-4 bp) to guide repair (Roth and Wilson, 1986). This alignment is guided by the Ku70/80 heterodimer, which binds both ends of a DSB independently with high affinity forming a ring or clamp structure around the broken ends (Walker et al., 2001). It should also be noted that this interaction is sequence independent and the micro-homology alignment is not well understood (Walker et al., 2001). Next, the protein kinase DNA-PKcs is recruited by each Ku70/80 heterodimer and aids in forming the synaptic complex that brings together the two broken DNA termini. Other functions of DNA-PKcs include telomere maintenance/capping, as well as regulation of ATM in the DSB sensing pathway (Espejel et al., 2004; Shirvastav et al., 2009). If the DNA ends are not compatible, the termini will be processed by the Artemis nuclease, which cleaves overhangs and DNA hairpins (Ma et al., 2002). The Ku heterodimer then re-
cruits the XRCC4-ligase IV complex, to perform the end-joining step between the two processed termini (McElhinny et al., 2000).

It has been demonstrated that yeast cells are viable without NHEJ, but that vertebrate cells are not (Yamaguchi-Iwai et al., 1999). The reason for this is still unclear, although it may have to do with protein redundancies present in more complex eukaryotes or the suggestion that vertebrates primarily repair DSBs through NHEJ, while yeast primarily use HR (Yamaguchi-Iwai et al., 1999). NHEJ is used primarily during the G0 and G1 phases, likely because it would be problematic to enter the S phase of the cell cycle with an unrepaired DSB (Lees-Miller and Meek, 2003).

Alternative non-homologous end-joining

Micro-homology mediated end-joining has been discussed as an ALT-NHEJ pathway, guided by slightly longer tracts of homologous nucleotides than is the classical NHEJ pathway (>5 nt). ALT-NHEJ occurs by the same mechanism as classical NHEJ, but uses a different ligase (DNA ligase I or III) in cases where ligase IV is non-functional (Wang et al., 2005). An ALT-NHEJ pathway has been described as being responsible for the fusion of dysfunctional telomeres, in the absence of the NHEJ Ku70/80 heterodimer (Rai et al., 2010).

An ALT-NHEJ pathway was also characterized in the IgH class-switching mechanism. This process is Xrcc4 and ligase IV independent, since it was observed in mouse cells that were deficient in the two proteins (Wu et al., 2008; Yan et al., 2007). Generation of antigen diversity through V(D)J recombination can also be achieved by ALT-NHEJ, suggesting this mechanism may be important in maintaining immunodiversity (Corneo et al., 2007).
B. Homologous recombination mechanisms

Homologous recombination represents a second, essential DSB repair mechanism for maintaining genomic integrity and is dependent on a homologous donor sequence either on the same or a different chromosome. Several unique sub-pathways fall under the homologous recombination designation and these are used in different circumstances. In the mitotic cycle, sister chromatids are available for homologous recombination during the S and G2 phases of the cell cycle and are a factor in determining whether HR repair is initiated. Cell cycle control is also achieved through cyclin dependent kinase (CDK) phosphorylation of Sae2 (the yeast homolog to mammalian Ctp1, discussed earlier), an endonuclease required for the initial 5’→3’ resection event (Huertas et al., 2008). An overview of the various HR pathways is presented below. A separate section (Section 1.3) details the role of individual proteins common to most of the HR pathways.

Holliday model

Based on data from gene conversion events in fungi, Robyn Holliday proposed the first model for HR (Holliday, 1964). A gene conversion event is a one-way transfer of genetic information from one DNA sequence to another. In the Holliday model, two strands of DNA nicked at homologous sites invade one another, forming heteroduplex DNA (hDNA) and ultimately a Holliday junction (HJ) in which the DNA strands cross one another. Depending on how this HJ intermediate is resolved either a crossover or non-crossover product will be formed (Fig. 1.1).
Meselson-Radding model

The Holliday junction model was expanded upon by Meselson and Radding (1975) to account for observations that hDNA was only forming on one chromatid, not both, as the Holliday model would suggest. The Meselson and Radding model differs slightly in that only one of the DNA strands will be nicked, forming a 3’ end that invades a homologous template strand and primes new DNA synthesis. The non-template strand is pushed away, forming a displacement loop (D loop). A nick is later formed on the template, forming a HJ and asymmetrical hDNA preceding the HJ. This HJ can also be resolved in the same manner as in the Holliday model, leading to crossover or non-crossover products.
Figure 1.1 Holliday model. Two homologous sequences containing nicks will encounter one another (A). In the second step (B), a Holliday junction is formed after the nicked strands have invaded their homologous partner. Resolution can occur by cleaving the outer strands leading to a crossover product shown (C) or by cleaving the inner strands yielding the non-crossover products on the right (D).
Double-strand break repair

The double-strand break repair (DSBR) model was first proposed in *S. cerevisiae* to explain the insertion of a plasmid containing a fragment of DNA homologous to the yeast chromosome (Orr-Weaver *et al.*, 1981; model reviewed in Szostak *et al.*, 1983). This model builds on the foundation established by the Holliday and Meselson-Radding models, depicted above, and provides a more satisfactory explanation of crossover locations in gene conversions and the differences in the rates of gene conversions between strands (Gilbertson and Stahl, 1996).

In the DSBR model, resection occurs from the DSB in the 5′ → 3 direction, leaving long single-stranded 3′ DNA overhangs, a step that is common to all forms of HR repair. The 3′ tails engage in homology search and strand invasion of a homologous sequence (i.e. sister chromatid), collectively termed synapsis. As the invading strand initiates new DNA synthesis, a D loop is formed on the non-template strand of the chromosome, being pushed out by the invading 3′ end of the broken sequence. The D loop is captured by the non-invading 3′ tail of the broken sequence. DNA synthesis occurs simultaneously from the non-invading 3′ tail and will continue from both ends of the original DSB. Eventually, a double Holliday junction structure is formed, which will need to be resolved to complete the repair process. Assuming the HJs are resolved (inner or outer strands cleaved) in a random manner, crossover or non-crossover products will be formed equally (reviewed in Szostak *et al.*, 1983; Li and Heyer, 2008). This does not seem to be the case, as gene conversion events seem to occur predominantly from one side of a DSB (Gilbertson and Stahl, 1996; McCulloch and Baker, 2006). Crossover events are hazardous during mitosis, causing genome re-arrangements; however, during meiosis, crossover is necessary for the correct segregation of chromosomes (Nicklas, 1974).
Figure 1.2 Double-strand break repair. The DSB is processed leaving 3’ ssDNA tails (A). A homologous chromosome is located and strand invasion and new DNA synthesis (dotted line) occur by the invading 3’ end, forming a D loop on the homologous chromosome (B). The D-loop is captured by the non-invading strand, which can then undergo its own DNA synthesis, as DNA synthesis continues on the invading strand (C). Once DNA synthesis is completed, a double Holliday junction structure is formed (D). Depending on how this structure is resolved (E), either a crossover (A+B) or non-crossover product is formed (A+C).
**Synthesis-dependent strand annealing**

The synthesis-dependent strand annealing (SDSA) model of repair is another proposed HR mechanism for repair of a DSB. This model was first hypothesized by Nasmyth (1984) to explain mating type switching in yeast, which yields a non-crossover product. A similar observation was made in *Drosophila*, with HR events yielding almost exclusively non-crossover products (Gloor *et al.*, 1991). The distinguishing feature of the SDSA model is that only non-crossover products are formed and DNA synthesis is conservative, making it the favoured model for explaining non-crossover and events in mitosis and meiosis (McMahill *et al.*, 2007).

In the SDSA model, new DNA synthesis begins after resection and 3’ end invasion on one side of the DSB. The non-invading strand of the DSB is prevented from undergoing DNA synthesis by a helicase, BLM or RTEL1 (Uringa *et al.*, 2011). The helicase will disrupt the D loop structure, preventing the D loop from acting as template for the non-invading strand. This action also eventually causes dissociation of the nascent invading strand potentially allowing it to act as template for the other end of the DSB, leading to complementary strand synthesis. Finally, the newly-synthesized strand will anneal to complementary sequence on the opposite strand of the DSB and prime DNA synthesis to correct the DSB.

The original SDSA model (Fig 1.3) can be modified to include a double HJ as the DSBR model does, through extensive branch migration and capture of the D loop (Pâques and Haber, 1999). This modified version of SDSA is distinguished from the DSBR model, by the location of the HJs, which are both located on one side of the original DSB. If crossover events did occur in this modified SDSA model, crossover would be detected on only one side of the DSB.
Figure 1.3 Synthesis-dependent strand annealing. A DSB is processed by 5’ → 3’ resection on both sides of the DSB (A). Next, one of the 3’ overhang tails will invade a homologous template strand and initiate DNA synthesis (B). Eventually this strand will dissociate from the template and itself serve as a template, while also annealing to the complementary sequence on the opposite side of the DSB serving as template (C). Fill-in synthesis completes DSB repair (D).
**Break-induced replication**

Break-induced replication (BIR) is used to repair a DNA strand break that has only one free end, such as at the end of a telomere. The model was proposed in yeast to explain long tracts of one-sided, asymmetric gene conversions, based initially on the Meselson-Radding model (Esposito, 1978). This mechanism is able to lengthen telomeres, repair broken chromosomes and restart stalled replication forks, although it is also thought to cause nonreciprocal translocations, deletions, loss of heterozygosity and complex rearrangements observed in a number of diseases and cancers (Lyedard *et al.*, 2007; Haber, 2006. Hastings *et al.*, 2009).

Once again, the initial event upon DSB recognition is 5’→3’ resection creating a 3’ overhang tail (Lyedard *et al.*, 2010). The single 3’ tail invades the homologous template chromosome and initiates DNA synthesis. As DNA synthesis continues, the nascent strand is displaced and becomes available for lagging strand synthesis (Fig. 1.4). Synthesis will continue until the end of the template chromosome arm is reached and the strands will then separate (Voelkel-Meiman and Roeder, 1990).
Figure 1.4 Break-induced replication. Following end-resection (A), 3’ strand invasion and DNA synthesis (B), the newly-synthesized region will dissociate from the template strand and serve as template for lagging strand synthesis (C). BIR continues until the end of the template is reached (D).
Single-strand annealing

Single-strand annealing (SSA) was first proposed by Lin and Sternberg (1984) to explain HR between repeats on extrachromosomal plasmids. It is a homology-dependent mechanism that can be used to resolve a DSB between sequential repeats along a chromosome (Pâques and Haber, 1999). It is possible that SSA is involved in regulating the number of repeats along a chromosome. However, the simultaneous loss of one or more repeats, along with the intervening sequence, makes it a mutagenic, non-conservative repair process.

The first step in SSA is common to all the different HR mechanisms and involves 5’ → 3’ nucleolytic excision of the DSB exposing 3’ single-stranded complementary regions that can anneal (Fig. 1.5). Subsequently, the intervening sequence will be removed by a 3’-5’ flap endonuclease resulting in the loss of one of the repeats and all intervening sequence. It has been proposed that the mechanism may be controlled by the timing of the repeats’ exposure to one another, if the exposure is almost simultaneous; repair by SSA is likely (Prado and Aguilera, 1995). Another factor in determining if SSA is used is the stage of the cell cycle, as the greatest effect of SSA has been observed in the S phase (Frankenberg-Schwager et al., 2008)
Figure 1.5 Single-strand annealing. SSA occurs when a DSB is encountered between a pair of repeats (A). Resection occurs in a 5’ to 3’ direction on both sides of the DSB, exposing the repeats to one another (B). The repeats will anneal and an endonuclease will remove the intervening DNA which would be heterogeneous (C). A polymerase will fill in remaining gaps and a DNA ligase will complete the process sealing any nicks (D).
1.3 Homologous recombination proteins

In eukaryotes, the MRN complex (outlined in Section 1.1) is initially responsible for sensing and partially resecting the ends of a DSB, in a 5’ → 3’ manner. Single-stranded DNA is then bound by replication protein A (RPA), which prevents secondary structure formation and nuclease degradation, while competing for the same binding sites as Rad51 (Sugiyama et al., 1997). Rad51 is the mammalian recombinase protein, discovered to be a homolog of the critically important bacterial RecA recombinase (Shinohara et al., 1992), with homologs in all kingdoms of life (Volodin et al., 2005). Rad51 coats the 3’ ssDNA displacing RPA and forming a right-handed nucleoprotein filament. The Rad51 nucleoprotein filament is able to initiate the homology search and strand invasion steps of HR. The meiosis specific Rad51 homolog, Dmc1 interacts with Rad51 during meiotic HR events, performing a similar function (Bishop et al., 1992; Sehorn et al., 2004).

Rad51 is an ATP-dependent strand exchange protein, involved in homologous pairing (Baumann et al., 1996) and is potentially the most significant protein in early HR. Yeast and human Rad51 are able to mediate strand exchange in the absence of ATP hydrolysis, but with reduced efficiency (Sung and Stratton, 1996; Morrison et al., 1999). Rad51 knockouts are embryonic lethal in vertebrate cells (Lim and Hasty, 1996; Tsuzuki et al., 1996; Sonoda et al., 1998), suggesting that the strand exchange and the homology search steps of HR, are essential for cell viability.

Many proteins have been shown to interact with Rad51 and influence this early stage of homologous recombination. Two Rad51 paralogs identified in S. cerevisiae, Rad55 and Rad57, form a heterodimer with the Rad51 filament that has a stabilizer effect, allowing Rad51 to main-
tain its filamentous structure on the ssDNA and resist the effects of the Srs2 helicase, which has an anti-recombination effect (Liu et al., 2011). It is also possible that Rad55 and Rad57 play a role in allowing Rad51 to overcome the inhibitory effect of RPA during strand exchange (Sung, 1997). Rad55, along with Rad52, also seems to be necessary for localization of Rad51 at a DSB in S. cerevisiae (Sugawara et al., 2003).

Rad52 has a mediator effect on Rad51 and interacts with Rad51 and RPA (Krejci et al., 2002). Rad52 aids in Rad51 filament formation by helping to overcome the inhibitory effect of RPA (Sugiyama and Kowalczykowski, 2002), as well as the Srs2 anti-recombinase by nucleating Rad51 on RPA-coated ssDNA (Burgess et al., 2009). Rad52 is also involved in Rad51-mediated DNA strand exchange in S. cerevisiae (in humans BRCA2 fulfills this role—see below), although this function has not been conclusively demonstrated in humans (Shinohara and Ogawa, 1998). In S. cerevisiae, Rad52 deletion causes severe sensitivity to DSBs caused by IR (Dudas et al., 2003), while, in higher eukaryotes, there is no increase in sensitivity to DSBs, suggesting some other protein compensates for Rad52 loss (Rijkers et al., 1998). In the absence of telomerase, Rad52 plays a role in telomere maintenance through HR (Teng and Zakian, 1999).

In mammals, the tumour suppressor BRCA2 plays an important role as a Rad51 mediator: BRCA2 is involved in displacing RPA and loading Rad51 onto ssDNA, reducing the affinity of Rad51 binding to double-stranded DNA (dsDNA), localization of Rad51 (and Dmc1) to nuclear foci and also stimulating Rad51-mediated DNA strand exchange (Venkitaraman, 2002; Jensen et al., 2010; Liu et al., 2010; Thorslund et al., 2010). Only when Rad51 has formed a filament on ssDNA is it able to carry out its function in homology search and strand invasion. Thus, BRCA2 serves an important role in the early events of HR.
The Rad54 protein is the most diverse of the HR proteins. Rad54 stabilizes the Rad51 filament, remodels chromatin and is a bidirectional motor protein with ATPase activity (Tan et al., 1999; Amitani et al., 2006). Rad54 is also involved in the transition from strand invasion to 3’ extension by stimulating Rad51 strand invasion activity, then dissociating Rad51 from double-stranded DNA (once the Rad51 nucleoprotein filament has undergone complementary pairing with the homologous sequence), allowing a DNA polymerase access to the invading 3’ end (Li and Heyer, 2009). There are other members of the Rad52 epistasis group involved in mediating Rad51 activity, such as the Rad51 paralogs, Xrcc2 and Xrcc3, in which deficiency results in a decreased HR response to DSBs, although their exact function remains unclear (Pierce et al., 1999).

DNA synthesis is carried out by a variety of DNA polymerases in mammalian cells, though precisely which are involved in homologous recombination has not been well studied. In mammalian cells, polymerases λ and μ have been shown to be involved with the NHEJ mechanism discussed earlier (McElhinny and Ramsden, 2004). These polymerases may be assumed to have low processivity, as they need to fill in 1 to 5 bp of sequence before ligation of strands. DSBs due to collapsed replication forks require DNA synthesis by replicative polymerases δ or ε in S. cerevisiae (Holmes and Haber, 1999), although they do not act alone in HR repair mechanisms.

The Y-family DNA polymerases perform translesion DNA synthesis (TLS) upon replication blockage (Friedberg et al., 2002). Translesion polymerases, such as polymerase η and ζ, function in lesion bypass at stalled replication forks (Plonsky and Woodgate, 2004), performing DNA synthesis over short tracts in these affected regions. Polymerase η is able to bypass pyrimidine dimers and perform synthesis across (6-4) photoproducts, as well as other forms of damage.
Translesion polymerases also play a role in HR as knockouts of polymerase η showed high sensitivity to DSBs (Kawamoto et al., 2005). Hypersensitivity to agents causing DSBs also became apparent in polymerase ζ depleted cells (Sharma et al., 2011). The Y family polymerases often show low fidelity and processivity, which following bypass, is compensated for by the relatively high fidelity of the main replicative polymerases, δ and ε (Garcia-Diaz and Bebenek, 2007).

Upon completion of DNA synthesis, the HJs need to be resolved to complete the DSBR process and, depending on how they are resolved, either a non-crossover or a crossover product will be formed. During mitotic events, it is usually essential that HR does not produce crossovers, since this could produce translocations. The process of dissolution, mediated by the BLM-TopoIIIα-BLAP75 complex (BTB), does not yield a crossover product. BLM is a helicase whose ATPase and unwinding activity brings together the two HJs. TopoIIIα then causes the HJs to become decatenated by causing a nick, separating the strands, such that there is no exchange of sequences, with BLAP75 acting as a cofactor in the complex (Bussen et al., 2007). In cases where crossover is necessary, the mammalian Mus81 endonuclease is able to nick the HJ strands asymmetrically (allowing for resolution via outer strands, Fig. 1.2 D) yielding a crossover event between the template and the broken strand (Osman et al. 2003). The human GEN1 (yeast, YEN1) endonuclease is able to resolve the HJ in either manner by symmetric cuts, forming crossover and non-crossover products (Rass et al., 2010).
1.4. Other pathways of DNA repair

For simple DNA damage such as altered bases or incorrect base pairing, other repair pathways are used, or the damage may be tolerated. The repair mechanisms that are available for repairing DNA damage not involving a DSB are reviewed next.

Nucleotide excision repair

Nucleotide excision repair (NER) is used to repair bulky DNA damage, such as thymine dimers, 6-4 photoproducts and benzo[a]pyrene products, among others (Petit and Sancar, 1999). The general repair pathway uses a specific set of UV-related proteins for detection of a lesion, but also proteins that are utilized in other forms of DNA repair, such as RPA (Mu et al., 1997). The general mechanism of NER includes independently-made incisions on the 3’ and 5’ sides of the lesion, followed by helicase/exonuclease removal of the ssDNA (Reardon et al., 1997). This is followed by DNA polymerase fill-in of the ssDNA gap and DNA ligase sealing the nick (reviewed in Petit and Sancar, 1999).

Base excision repair

Base excision repair (BER) is used to remove base damage that occurs due to cellular mechanisms, such as deaminations, oxidations and other base modifications. DNA glycosylases are first used to hydrolyze the N-glycosidic bond between the damaged base and the sugar creating an abasic site (Zharkov, 2008). An endonuclease then removes the abasic nucleotide, leaving a free 3’ termini which can be repaired using DNA polymerase, with DNA ligase finally sealing the nicked strand (Bennett et al., 2001).
Direct repair of DNA damage

In some cases of DNA damage, there is a simpler mode of repair, often using only one enzyme to catalyze the process rather than a complete pathway. Photoreactivation is a process that occurs in bacteria, yeast and fungi to repair UV-induced pyrimidine dimers (Sancar, 2008). The repair is carried out by the enzyme photolyase, which will first bind the pyrimidine dimer, before becoming activated by blue light in the 350-500 nm wavelength range. An energy transfer reaction then resolves the pyrimidine dimer, re-establishing proper base pairing (Sancar, 2008).

DNA ligase is used to repair single-strand nicks that often result through normal cellular processes, such as joining of Okazaki fragments during lagging strand synthesis. DNA ligase is able to seal nicks by catalyzing the formation of a phosphodiester bond between the 5’ phosphate and 3’ hydroxyl group in DNA.

Alkylation can cause cytotoxic and mutagenic lesions through the attachment of alkyl (e.g. methyl, ethyl) groups to DNA bases. This damage may be repaired via BER, alkyltransferases (a class of “suicide” enzymes) or dioxygenases. An example of a suicide enzyme is O⁶–alkyl-guanine-DNA alkyltransferase (AGT), which is widespread among prokaryotes and eukaryotes (Tubbs and Tainer, 2010). The AGT is able to transfer a damaged guanine O⁶–alkyl adduct to a cysteine residue it contains, in an irreversible manner (Tubbs and Tainer, 2010). In contrast, dioxygenases catalyze a demethylation process by hydroxylating the methylated base forming a formaldehyde group that is then removed from the base (Falnes et al., 2002).
Mismatch repair

The mismatch repair (MMR) pathway recognizes and removes incorrect base pairs and insertion/deletion loops within DNA, while also playing a role in regulating the occurrence of HR (Stone et al., 2008). In bacteria, initial recognition of the mismatched DNA (mmDNA) occurs through binding by the MutS homodimer protein complex, before recruiting the MutL homodimer in an ATP-dependent manner (Wang and Hayes, 2007). In *E. coli*, MutH proteins will then nick hemi-methylated d(GATC) before a helicase (UvrD/MutU) proceeds to remove the DNA containing the mismatch, which in turn, is removed by a ssDNA endonuclease (Viswanathan et al., 2001). The gap is then filled in from the 3’ end by a DNA polymerase before being sealed by a DNA ligase.

Eukaryotes use a different set of proteins to achieve similar functions in the MMR system. The MutS homolog, MutSα is a heterodimer protein that performs the same initial binding and recruitment of MutL-like protein heterodimer, MutLα (Harrington and Kolodner, 2007). Eukaryotes do not have an analog for the MutH protein, and thus do not recognize strands based on methylation at specific sites. Instead, cuts are made at pre-determined excision sites by MutLα in an ATP-dependent manner (Kaydrov et al., 2006). An exonuclease is then able to carry out resection (Genschel and Modrich, 2003), leaving a gap that is filled in by DNA polymerase before being sealed by DNA ligase.
1.5 Role of homology in recombination

In *E. coli*, gene targeting requires a minimum of 27 bp of homologous sequence (Shen and Huang, 1986). This length is termed the minimal efficient processing segment (MEPS) as recombinants are generated with shorter homology lengths, but at greatly reduced frequency. The recombination frequency in *E. coli* increases linearly from the MEPS (27 bp) up to approximately 300 bp (Shen and Huang, 1986). Similar studies in yeast have estimated the MEPS somewhere between 25 - 40 bp (Chino et al., 2010), again with recombination frequency increasing linearly up to 1 kb (Jinks-Robertson et al., 1993). The MEPS has not been as thoroughly studied in *Drosophila*, as different studies have reported different limits. However, it is thought to be between 172 bp (Preston and Engels, 1996) and 375 bp (Dray and Gloor, 1997).

Studies have shown that the critical length of homology for gene targeting in mouse hybridoma cells, which are used for the work presented here, is approximately 1 kb (Shulman et al., 1990). In murine embryonic stem cells, gene targeting could occur with approximately 500 bp of homology on one arm, though a minimum of 1.9 kb of total homology was necessary (Hasty et al., 1991). Embryonic mouse stem cells, used in a more extensive gene targeting experiment, showed an exponential relationship between homology and HR events using isogenic and non-isogenic DNA sources (Deng and Capecchi, 1992).

Intrachromosomal gene targeting studies between closely linked repeated sequences suggest a MEPS of 295 bp, with HR events increasing proportionally up to 1.8 kb (Liskay et al., 1987). A study measuring ectopic HR requirements in murine cells determined the MEPS between 1.9 and 4.3 kb (Baker et al., 1996).
Mismatches also have a drastic effect on HR in *E. coli*, as a sequence that had only 92% homology compared to the wild-type, reduced gene targeting by 45-fold (Shen and Huang, 1986). The effect of mismatches on intrachromosomal recombination in mouse cells showed that two single nucleotide mismatches 19 bp apart within an otherwise homologous 232 bp sequence led to a 20-fold decrease in recombination (Waldman and Liskay, 1988). Similarly, a single mismatch within a 232 bp homologous sequence caused a 2.5-fold reduction in HR (Lukacsiovich and Waldman, 1999). In these studies, single mismatches showed greater effects when less overall sequence homology was present.

The effect of gap size on HR has been studied in *Drosophila melanogaster* with gaps just under 11 kb being repaired with the same frequency as DSBs with minimal gaps (Johnson-Schlitz and Engels, 2006). However, a previous study in yeast (Pâques et al., 1998) suggested that as gap length increases up to 9.1 kb, gene targeting events decline.

During gene targeting in mouse embryonic stem cells, gaps as large as 2.5 kb are repaired efficiently, although at half the frequency of gaps that are only 200 bp (Valancius and Smithies, 1991), though the authors attribute this effect to the reduction in homology caused by the gap, rather than the gap itself. Ectopic HR studies using murine hybridoma cells have shown a trend in which vectors bearing longer gaps are repaired less efficiently than those with shorter gaps (Baker et al., 1996).
1.6 Hybridoma cell lines

The cells used in this study are derived from the Sp6/HL GK murine hybridoma cell line (Fig. 1.7). This cell line was established by fusing myeloma tumor cells with spleen cells from a trinitrophenyl (TNP)-bovine serum albumin (BSA) immunized Balb/c mouse. The chromosomal immunoglobulin μ heavy chain gene (Cμ) in Sp6/HL is able to combine with the functionally rearranged TNP-specific κ light chain gene to produce TNP-specific IgM capable of complement-dependent lysis of TNP-coupled sheep red blood cells (Baker et al., 1988). The igm482 cell line is derived from the Sp6/HL line: a 2 bp deletion in the third exon of the igm482 μ gene constant region (Cμ3), produces a non-functional IgM (Köhler et al., 1982). The 2 bp Cμ3 deletion can be corrected through HR with a wildtype μ gene, once again producing functional TNP-specific IgM.

The 51-5 hybridoma cell line was established as a derivative of the igm482 hybridoma through stable integration of a vector containing an N-terminal FLAG-tagged wildtype Rad51 cDNA (Rukšė et al., 2007). Western blot analysis reveals that the 51-5 cell line expresses the FLAG-tagged Rad51 at a 2-fold higher level than the endogenous Rad51 protein (Rukšė et al., 2007). This cell line allows for analysis of the effect of excess wild-type Rad51 on mammalian HR.

The hybridoma cell lines are advantageous in the study of HR, partly because of their rapid doubling times (18h) and their immortality. Another advantage of the hybridoma cell line is that the immunoglobulin genes which are used as recombination reporters are expressed at high levels and present in a single copy, preventing chances of unknown donor-recipient HR events (Köhler et al., 1978; Baker, 2004).
**Figure 1.6 Hybridoma cell lines.** The origin of the igm482 and 51-5 hybridoma cell lines. A myeloma cell line is fused with spleen cells from a TNP-BSA-immunized Balb/c mouse to create the initial Sp6/HLGK hybridoma cell line. For further details, refer to the text.
1.7 3’ end extension assay

Current methods for studying HR are based on the detection of the final products of the reactions. Consequently, they do not provide insight into the early stages of the HR process and so these stages in mammals are not well characterized. The basis for the understanding of early HR and its dependence on homology comes from studies done at the MAT locus in yeast (Sugawara and Haber, 1992; Sugawara et al., 2003; Wang and Haber, 2004; Coïc et al., 2011; Hicks et al., 2011). Yeast mating type switching occurs via a programmed DSB generated by the HO endonuclease at the MAT locus, followed by HR (Haber, 1992). HR studies in yeast have helped provide the basis for the models of HR detailed in section 1.2, as well as providing insight into the kinetics of the process (Sugawara et al., 2003; Hicks et al., 2011).

The 3’ end extension assay is able to look at a product formed in the early stages of HR and may help elucidate the initial proteins and mechanism used during the HR repair processes (Si et al., 2010). Similar to yeast, this assay allows for characterization of early HR products and possibly the effects of HR proteins, although they have not previously been well characterized in the integrated mammalian system. The technique is similar to a PCR amplification technique used to observe early HR intermediates at the MAT locus in yeast (White and Haber, 1990). The primers used in this reaction were paired such that one primer would anneal in a region distal to the MAT locus, while the partner would anneal within the donor sequence to be copied at the MAT locus; thus, only products undergoing early DNA synthesis in HR would be produced (White and Haber, 1990).

In the 3’ extension assay, a gene targeting vector bearing homology flanking a double-stranded gap (DSG), to the endogenous target Cμ gene is electroporated into the igm482 hybrid...
The synapsis process (Fig. 1.2 B), which also includes invasion of a broken strand, will cause new DNA synthesis, that is measured by the 3’ end extension assay. To detect this DNA synthesis, plasmid DNA is extracted at various post-electroporation time points and subjected to PCR amplification using a specific pair of primers for “left” and “right” arm 3’ extension events. The primers used in the PCR amplification process are specific to the vector backbone and a region of newly synthesized DNA within the vector-borne DSG (Fig. 1.7 left and right arm extension primers). A second set of primers is used to amplify sequences within the neo resistance gene of the vector, to quantify the amount of vector present at each time point and standardize 3’ extension events per vector backbone.
Figure 1.7 3’ extension primer binding sites. The primer sites in use for quantification in the 3’ end extension assay are shown here as approximate locations. Quantification of the left arm 3’ extension events is achieved through different pairs of primers, each with one unique vector site and a site provided by the newly-synthesized DNA in the vector-borne DSG (dotted line). Right arm 3’ extension quantification will be achieved with similar primers on the other side of the gap.
Hypothesis

As reviewed above, several previous studies have examined the homology requirements for recombination, that is, the length of homology required to generate completed products of homologous recombination. Given the complexity of the entire homologous recombination process, it would be useful to understand the role of homology in the early homology search and strand invasion steps of homologous recombination.

Our laboratory has developed a plasmid-based 3’ extension assay that measures the new DNA synthesis resulting from these early steps in homologous recombination. In this thesis, I will test the hypothesis that homology requirements for 3’ extension differ from those required for the completion of homologous recombination.
CHAPTER 2: Materials and Methods

2.1 Hybridoma cell lines

Hybridoma cell lines were grown as suspension cultures in Dulbecco’s modified Eagle’s medium (DMEM) supplemented with 13% bovine calf serum and .0035% 2-mercaptoethanol, as previously described (Köhler and Shulman, 1980).

Sp6/HL

The Sp6/HL mouse hybridoma cell line (Fig. 1.6) was established by fusing a myeloma cell line with a spleen B-cell from a 2,4,6–trinitrophenyl-immunized Balb/c mouse (Köhler and Shulman, 1980). Sp6/HL has a single copy of the TNP-specific chromosomal immunoglobulin μ heavy chain gene and produces TNP-specific IgM (K chain).

igm482

The mutant igm482 (Fig. 1.6) bears a 2 bp deletion within exon 3 of the μ gene (Cμ3) (Köhler et al., 1982). The Cμ3 deletion renders the TNP-specific IgM unable to activate complement-dependent lysis of TNP-coupled sheep red blood cells (Köhler and Shulman, 1980). However, the mutant igm482 IgM genotype can be corrected by HR with the wild-type Cμ region. This makes the igm482 cell line useful in the study of mammalian homologous recombination (Baker et al., 1988).
Hybridoma cell line 51-5 is a hygromycin resistant (HYG<sup>R</sup>) igm482 derivative that expresses N-terminal FLAG-tagged mouse Rad51 in approximately a 2:1 ratio compared to endogenous levels of Rad51 and was generated as described previously (Rukšć et al., 2007). For 51-5 culture, DMEM was supplemented with hygromycin at a concentration of 700 μg/ml.

2.2 Plasmids

General Conditions

Plasmid DNA was propagated in <i>E. coli</i> DH5α cells in Luria-Bertani (LB) broth containing 50 μg/ml ampicillin (Sambrook et al., 1989). All plasmids were purified by maxi-prep using a PureLink HiPure Plasmid Purification Kit according to manufacturer specifications (Invitrogen).

Vectors used in 3’ extension

The homology segments used to construct the vectors used in the 3’ extension assay are summarized in Figure 2.1 and Figure 2.2. As indicated in the figures, the homology segments are derived from restriction enzyme cleavage fragments, originally in the chromosomal immunoglobulin μ gene target locus. Each fragment was inserted into the vector pSV<sub>2</sub>neo (Southern and Berg, 1982) in the orientation indicated “Left” and “Right” using standard DNA cloning procedures (Sambrook et al., 1989). In the vectors, BstEII digestion creates a 1.2 kb double-stranded gap (DSG) in the region of homology to the μ gene. Chromosomal μ gene sequences
excluded by the 1.2 kb vector-borne DSG serve as template for 3’ extension from the invading 3’ BstEII vector ends.

\textbf{pTΔCμ858/2290}

The pTΔCμ858/2290 vector (Fig. 2.2, Vector #1) is identical to the pTΔCμ vector used previously to investigate 3’ extension in the hybridoma system (Si \textit{et al.}, 2010; Fig. 2.1). The pTΔCμ858/2290 vector was derived from pTCμ (Baker \textit{et al.}, 1988) and consists of a 4.3 kb Xbal μ gene segment that was blunt-ended, fitted with SalI linkers and cloned into the SalI (formerly, EcoRI) site of the vector pSV2neo (Southern and Berg, 1982). Following digestion with BstEII and re-ligation, the pTΔCμ vector was generated (Si \textit{et al.}, 2010). As indicated in Figure 2.2, digestion at the unique BstEII site in pTΔCμ generates a vector bearing two homology arms to the target chromosomal immunoglobulin μ gene, an 858 bp “left” homology arm and a 2,290 bp “right” homology arm separated by a 1.2 kb double-stranded gap. To denote this arrangement and to distinguish this vector from the other vectors described below that contain different homology segments, in this thesis, pTΔCμ has been renamed pTΔCμ858/2290. As outlined below, the pTΔCμ vector was used as the basis for generating several additional vectors, each retaining the 1.2 kb BstEII DSG (unless explicitly stated), but having changes in homology of the “left” or “right” invading vector arms.
Figure 2.1 Map of pTΔCμ vector. The pTΔCμ vector is shown here with sites relevant to flanking the region homologous to the Cμ gene (SalI sites), the site containing the gap (pictured in Fig. 2.2) to the Cμ gene (BstEII) and remaining sites used for the creation of alternate homology constructs described above. The gap at BstEII encompasses a portion of Cμ exons 2 and 4, as well as the entire Cμ exon 3.
Figure 2.2 Vector homology segments used in the 3’ extension assay.

The pSV<sub>2</sub>neo plasmid backbone shown at the top is common to all vectors in this study. The vectors below indicate the homology segments used in the various plasmid constructs. To the right of these fragments, the plasmid name is given, with the left and right homology segments indicated as sub-scripts, respectively. At the bottom of the figure the endogenous immunoglobulin μ gene is shown, containing the heavy chain variable region (V<sub>H</sub>TNP), the μ gene switch region (Sμ) and the Cμ exons (1-4). Details of plasmid construction are presented in the text.
Left Arm Constructs

**pTΔCμ2120/2290**

The vector pTΔCμ2120/2290 (Fig. 2.2, vector #2) was previously established in the lab (Ng and Baker, 1999) and used to test extended homology on the left arm. This vector contains 2.12 kb homology on the left arm upstream of the Cμ exons, as well as a small portion of the μ gene switch region (Sμ), containing a series of repeats and a thymidine kinase (tk) gene. The 2.29 kb of right arm homology remains the same as in the pTΔCμ858/2290 plasmid.

**pTΔCμ2070/2290**

This vector is a derivative of pTΔCμ2120/2290 above in which the tk and Sμ regions were removed through the use of PCR-mediated deletion (Hansson et al., 2008; Fig. 2.2, vector #3). The technique was performed as described, except that Taq polymerase was used in place of Pfu polymerase. This technique uses a pair of primers with a stretch of overlapping sequence between them to amplify a desired sequence, re-creating a plasmid with the sequence between the 5’ regions of the primers deleted. A forward primer was designed just downstream of the Sμ region in pTΔCμ2120/2290 (5’ GGCGAATTAGCTGAGCAAGAGTGAGTAGAGCTGGCTG 3’) and a reverse primer upstream of the tk gene (5’ TGCTCAGCTAATTGCCAATGACAAAGACGCTGG 3’; italicized regions indicating primer overlap). The resulting vector was sequenced to confirm loss of the Sμ and tk genes, using forward primer (5’ ATCCAGCTCCCCTACACCACAGCT 3’) and reverse primer (5’ CGCCCGCGTTTCTTTCTTTCCC 3’; Laboratory Services, University of Guelph).
**pTΔCμ858/2290X**

The pTΔCμ858/2290X vector (Fig. 2.2, vector #4) contains a 4 bp insertion mutation that destroys the endogenous SacI site in the intron between Cμ1 and Cμ2, creating an EcoRV site (Ng and Baker, 1999). A fragment created by Nhel/AflII restriction enzyme digest, was used to swap the fragment containing the mutation into the pTΔCμ858/2290 vector, creating the pTΔCμ858/2290X vector.

**pTΔCμ720/2290**

The vector pTΔCμ720/2290 (Fig. 2.2, vector #5) bears a 720 bp AflII/BstEII fragment in the left homology arm and the 2290 bp BstEII/XbaI fragment from pTΔCμ858/2290 above in the right homology arm. The 720 bp left homology arm was generated by AflII/Swal digestion of pTΔCμ858/2290, which removed a 138 bp region of homology from the left-most junction (Fig. 2.2). The removal of the homology fragment was followed by blunt-ending and ligation using standard DNA cloning methods (Sambrook et al., 1989). This produces a plasmid, which upon BstEII digestion and alignment with the chromosomal μ gene, forms the 1.2 kb BstEII DSG along with the above homology arms.

**pTΔCμ517/2290**

The pTΔCμ517/2290 vector (Fig. 2.2 vector #6) bears a 517 AvaI/BstEII fragment in the left homology arm and retains the 2290 bp BstEII/XbaI pTΔCμ858/2290 right arm fragment. The 517 bp left homology arm was generated in the same manner as described for pTΔCμ720/2290 above except that an AvaI/Swal digestion was used to remove a 341 bp fragment.
\textbf{pTΔCμ}_{211/2290} 

The \( pT\Delta C\mu_{211/2290} \) vector (Fig. 2.2, vector #7) has the lowest amount of homology on the left arm, bearing a 211 bp \( SacI/BstEII \) left arm fragment and retains the right arm homology fragment of \( pT\Delta C\mu_{858/2290} \). To generate the 211 bp left arm, a 647 bp fragment was removed from \( pT\Delta C\mu_{858/2290} \) with the enzymes \( SacI/SwaI \), leaving the residual 211 bp left arm of homology.

\section*{Right Arm Constructs}

\textbf{pTΔCμ}_{858/1294} 

The \( pT\Delta C\mu_{858/1294} \) vector (Fig. 2.2, vector #8) contains a 1.29 kb \( EcoRV/BstEII \) fragment on the right arm, while the left arm contains an 858 bp fragment. A partial \( SalI \) digest was followed by an \( EcoRV \) digest, removing a 1 kb fragment from the right arm of homology. This was followed by standard blunt-ending and ligation techniques (Sambrook et al., 1989).

\textbf{pTΔCμ}_{858/569} 

The \( pT\Delta C\mu_{858/569} \) vector (Fig. 2.2, vector #9) contains a 569 bp \( FspAI/BstEII \) fragment on the right arm, while the left arm homology remains the same as in \( pT\Delta C\mu_{858/2290} \). To create the reduced right arm homology, a \( SalI \) partial digest was performed followed by an \( FspAI \) digest, that removed a 1.72 kb fragment, generating the 569 bp homology arm. Standard procedures were used for blunt ending and ligation (Sambrook et al., 1989)
**pTΔBamCμ**

The last vector generated, pTΔBamCμ, was created to bear a smaller 623 bp gap in the homology region. A BstXI/BamHI double digest of the vector pTCμ (containing the full length 4.3 kb XbaI μ gene fragment) was used to create a 623 bp gap, with left and right homology arms of 880 bp and 2.8 kb respectively (Fig. 2.3). The blunt ending and ligation process destroys the restriction sites at these locations, necessitating the creation of a new one, which could be used to linearize the plasmid and create the 623 bp gap for use in the 3’ extension assay. The blunt-ending and ligation steps created the sequence 5’ GGATCG 3’. To restore the BamHI site for vector linearization, a single primer (5’ GAAGGTGCCCAGGATC[underline]CTTTGTTCTCGATGGTC 3’) was used in conjunction with site-directed mutagenesis (QuikChange Site-Directed Mutagenesis Kit, Stratagene) to convert the guanine residue in the ligated site (the underlined G in the sequence, 5’ GGATCG 3’) to the cytosine residue indicated in the primer (the underlined cytosine). Site-directed mutagenesis was performed according to manufacturer specification.
Figure 2.3 Construction of a 600 bp double-stranded gap in the vector pTΔBamCµ. The pTΔBamCµ vector can be linearized with BamHI to form a 623 bp gap. Vector-borne SalI sites correspond to XbaI sites in the chromosome. The BstEII sites used to create the 1.2 kb gap in pTΔCµ858/2290 are shown for comparison.
2.3 Hybridoma transfection and plasmid extraction

Hybridoma transfection and plasmid extraction were performed as described (Si et al., 2010). Exponentially-growing hybridoma cells (~5x10^5 cells/ml, as determined by trypan blue staining) were used for electroporation. A 50 µg sample of BstEII or BamHI-linearized plasmid was mixed with 2x10^7 hybridoma cells, re-suspended in 750 µl permeabilization buffer in a 0.4 cm cuvette and placed in a Gene Pulser (BioRad). The sample was subjected to two consecutive 700 V, 25 µF pulses, following which the cuvettes were placed on ice for 10 minutes. After this period, 1 ml DMEM was added to the cuvette and the mixture was placed at 37°C in a 7% CO_2 incubator for 20 minutes. The contents of the cuvette were transferred to a flask containing 40 ml DMEM, which was placed at 37°C in a 7% CO_2 incubator, for the desired incubation period. In the 3’ end extension assay, several electroporations were performed simultaneously and plasmid extraction conducted at different post-electroporation time points.

For plasmid extraction, the hybridoma cells were transferred to a 50 ml conical tube and centrifuged at 800 RPM for 10 minutes at 4°C and washed twice in 1X PBS. The cell pellet was lysed and plasmid DNA extracted via column chromatography using the Qiagen Miniprep Kit (Qiagen) according to manufacturer specifications. Plasmid was eluted in 50 µl deionized ultra-fluoride (DIUF) water and stored at -20°C until use in the PCR assay.
2.4 PCR Analysis

The primers used in the 3’ end extension assay are listed in Table 2.1 (synthesized by Laboratory Services, University of Guelph) and are separated into 3 groups based on what they measure. The conditions for the PCR reactions are based on those described in Si et al. (2010) and are presented in Table 2.2.

As stated earlier (Section 1.7), primer pairs are designed such that one primer will bind in the vector backbone, near the neo or amp genes, while the other primer binding site is provided by the new DNA synthesis that has extended into the vector-borne DSG as a result of copying chromosomal μ gene sequences. The PCR primers along with the fragment sizes that they generate from the various plasmids used in the 3’ extension assays are summarized in Table 2.3. The right arm primers, AmpR20 and CμF1, amplify a 2.4 kb fragment that extends 13 bp into the gapped region. The right arm primers AmpR20 and CμF2 amplify a 3.0 kb fragment that is 580 bp into the gapped region (30 bp into the gapped region in the case of the vector pTΔBamCμ). The left arm primers NeoF20’ and CμR1 amplify a 1.2 kb fragment extending 63 bp into the gap. Primers NeoF20’ and CμR2 amplify a 1.4 kb fragment progressing 210 bp into the gap, while NeoF20’ and CμR3 amplify a 1.8 kb fragment extending 620 bp into the gap. The backbone primer pair (NeoF-1 and NeoR+6) amplify a 1.2 kb fragment within the neo gene of the vector backbone to permit quantification of vector recovery (Table 2.3).

Following PCR amplification and visualization by agarose gel analysis, the diagnostic PCR bands were quantified using a Gel Doc and Quantity One Plus Imaging software (BioRad).
### Table 2.1 Primers used in 3’ extension assay

<table>
<thead>
<tr>
<th>Sequence amplified</th>
<th>Primer designation</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Backbone</td>
<td>NeoF-1</td>
<td>5’ CTCAGAAGAACTCGTCAAGAA 3’</td>
</tr>
<tr>
<td></td>
<td>NeoR+6</td>
<td>5’ CTATTCCAGAAAGTAGTGAGGA 3’</td>
</tr>
<tr>
<td>Right arm</td>
<td>AmpR20</td>
<td>5’ AAGTGCCACCTGACCTTAA 3’</td>
</tr>
<tr>
<td></td>
<td>CμF1</td>
<td>5’ CTGGGCTTCTCAAAATGGT 3’</td>
</tr>
<tr>
<td></td>
<td>CμF2</td>
<td>5’ CTGGGCTTCTCAAAATGGT 3’</td>
</tr>
<tr>
<td>Left arm</td>
<td>NeoF20’</td>
<td>5’ ATCTGCTGACTGCAACTGTAGCA 3’</td>
</tr>
<tr>
<td></td>
<td>CμR1</td>
<td>5’ CACATTCAGGTTCAGGGCCACGAGG 3’</td>
</tr>
<tr>
<td></td>
<td>CμR2</td>
<td>5’ CACATTCAGGTTCAGGGCCACGAGG 3’</td>
</tr>
<tr>
<td></td>
<td>CμR3</td>
<td>5’ CACATTCAGGTTCAGGGCCACGAGG 3’</td>
</tr>
</tbody>
</table>

### Table 2.2 Protocol for PCR amplification of 3’ extension samples

<table>
<thead>
<tr>
<th>Stage</th>
<th>Backbone</th>
<th>Left arm</th>
<th>Right arm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>30 cycles</td>
<td>10 cycles</td>
<td>20 cycles</td>
</tr>
<tr>
<td>Initial separation</td>
<td>95° 5m</td>
<td>95° 2m</td>
<td>95° 2m</td>
</tr>
<tr>
<td><strong>DNA separation</strong></td>
<td><strong>94° 30s</strong></td>
<td><strong>95° 45s</strong></td>
<td><strong>95° 45s</strong></td>
</tr>
<tr>
<td>Annealing</td>
<td>55° 30s</td>
<td>62° 30s</td>
<td>62° 30s</td>
</tr>
<tr>
<td><strong>Extension</strong></td>
<td><strong>72° 1m</strong></td>
<td><strong>68° 3m</strong></td>
<td><strong>68° 3m+30s</strong></td>
</tr>
<tr>
<td>Final Extension</td>
<td>72° 10m</td>
<td>68° 10m</td>
<td>68° 10m</td>
</tr>
</tbody>
</table>

*Bolded sections represent repeated steps in a cycle.*
Table 2.3 Summary of 3’ extension products

<table>
<thead>
<tr>
<th>Vector</th>
<th>Size of left arm PCR product (Primer Pair)</th>
<th>Size of right arm PCR product (Primer Pair)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pTΔCμ858/2290</td>
<td>1.2 kb (neoF20’-CμR1)</td>
<td>2.4 kb (ampR20- CμF1)</td>
</tr>
<tr>
<td></td>
<td>1.4 kb (neoF20’-CμR2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.8 kb (neoF20’-CμR3)</td>
<td></td>
</tr>
<tr>
<td>pTΔCμ858/2290X</td>
<td>1.2 kb (neoF20’-CμR1)</td>
<td>-</td>
</tr>
<tr>
<td>pTΔCμ720/2290</td>
<td>1.1 kb (neoF20’-CμR1)</td>
<td>-</td>
</tr>
<tr>
<td>pTΔCμ517/2290</td>
<td>850 bp (neoF20’-CμR1)</td>
<td>-</td>
</tr>
<tr>
<td>pTΔCμ211/2290</td>
<td>550 bp (neoF20’-CμR1)</td>
<td>2.4 kb (ampR20- CμF1)</td>
</tr>
<tr>
<td>pTΔCμ2120/2290</td>
<td>3.9 kb (neoF20’-CμR1)</td>
<td>2.4 kb (ampR20- CμF1)</td>
</tr>
<tr>
<td></td>
<td>4.1 kb (neoF20’-CμR2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4.5 kb (neoF20’-CμR3)</td>
<td></td>
</tr>
<tr>
<td>pTΔCμ2070/2290</td>
<td>-</td>
<td>2.4 kb (ampR20- CμF1)</td>
</tr>
<tr>
<td>pTΔCμ858/1294</td>
<td>1.2 kb (neoF20’-CμR1)</td>
<td>1.4 kb (ampR20- CμF1)</td>
</tr>
<tr>
<td>pTΔCμ858/569</td>
<td>1.2 kb (neoF20’-CμR1)</td>
<td>680 bp (ampR20- CμF1)</td>
</tr>
<tr>
<td>pTΔBamCμ</td>
<td>1.2 kb (neoF20’-CμR1)</td>
<td>3.0 kb (ampR20- CμF2)</td>
</tr>
</tbody>
</table>

All vectors use backbone primer pair neoR+6-neoF-1

-Vectors not tested for this result in 3’ extension assay
CHAPTER 3: Results

The 3’ end extension assay permits the analysis of early homologous recombination (HR) repair events and products. Here, some aspects of its flexibility and the unique observations it allows for are explored in relation to the effect of homology on the HR repair process.

3.1 Optimizing the amount of vector DNA for 3’ extension

As a first step, the optimal quantity of vector DNA to be used in the 3’ extension assay was determined. For these experiments, the pTΔCμ858/2290 vector (Fig. 2.2, vector #1) was used and 3’ extension efficiencies were measured for the left invading vector arm (using primers neoF20’-CμR1) that generates a 1.2 kb product (Table 2.3). Previously, Si et al. (2010) showed that 3’ extension peaks between the 3 and 6 h post-electroporation time points for both the left and right invading vector arms. Therefore, for each DNA concentration, the peak 3’ extension at 3 h was chosen, with the assumption that DNA concentration did not perturb the kinetics of 3’ extensions (Fig. 3.1).

To measure the frequency of 3’ extension/vector backbone for each DNA concentration, the intensity of the various PCR bands (Fig. 3.1 A, lanes 11-15) was determined by densitometry and the number of vector copies determined relative to the intensity of the same size band in the plasmid standards (Fig. 3.1 A, lanes 2-8). Similarly, the amount of recovered vector backbone was determined by comparing the intensity of the PCR bands in the samples (Fig. 3.1 B, lanes 12-16) with the plasmid standards (Fig. 3.1 B, lanes 2-9). From this data, the frequency of 3’ extension/vector backbone was determined.
The results appear to indicate a preliminary quadratic relationship ($R^2$ value= .89) between the quantity of vector and 3’ extension events (Fig. 3.2). Above 100 μg of vector electroporated into the igm482 cells, an inhibitory effect on 3’ end extension becomes clear, with no noticeable effect on cell viability. In summary, the 50 μg quantity of vector falls in the linear phase of the curve, and along with previous 3’ extension and gene targeting experiments, is the optimal quantity for use in the 3’ extension assay.
Figure 3.1 Titrating the concentration of the pTΔCμ<sub>858/2290</sub> vector.

A titration of the pTΔCμ<sub>858/2290</sub> vector in the 3’ extension assay to determine optimal electroporation quantity. The 1.2 kb left arm 3’ extension product in (A) was quantified by densitometry and compared to the 1.2 kb vector backbone PCR product (B). Samples were extracted at the 3 h time point (Si et al., 2010). The water (lane 9 A, lane 10 B) and unelectroporated vector samples (lane 10 A, lane 11 B) are both negative controls.
Figure 3.2 Summary of $pT\Delta C_{\mu 858/2290}$ DNA titrations. The $pT\Delta C_{\mu 858/2290}$ 3’ extension data is presented here to summarize the titration of $pT\Delta C_{\mu 858/2290}$ plasmid DNA used in the 3’ extension assay. The $pT\Delta C_{\mu 858/2290}$ vector was used with left arm primers (neoF20’-CμR1) that amplify a 1.2 kb PCR product.
3.2 3’ end extension events on the same side of the DSG increase linearly with homology

Having established the optimal concentration of plasmid DNA to use in the 3’ extension assay (Section 3.1), the next set of experiments focused on determining the effect of homology length on the efficiency of 3’ extension. This analysis was performed for 3’ extensions initiated from both the left and right invading vector arms. The requirement for homology in 3’ extension utilized vectors in which the left arm homology varies, while right arm homology remained constant (i.e. vectors pTΔCμ858/2290, pTΔCμ858/2290X, pTΔCμ2120/2290, pTΔCμ720/2290, pTΔCμ517/2290, pTΔCμ211/2290) as well as vectors in which right arm homology length varied, while the left arm homology was kept constant (i.e. pTΔCμ858/2290, pTΔCμ858/1294, pTΔCμ858/569). The vectors were electroporated separately into exponentially-growing igm482 cells and the frequency of 3’ extension was determined at the 3 or 6 h time points. Representative gels of this analysis are presented in Fig. 3.3 and Fig. 3.4. The sizes of the various amplified PCR products are shown in the figures and summarized in Table 2.3, along with the primer pairs used to detect the products.

The peak 3’ extension events varied between the 3 and 6 h time points, without a noticeable trend, so for each vector multiple time points were examined and the peak 3’ extension events were determined. 3’ extension deteriorated beyond the 9 h time point, consistent with previous plasmid electroporation results (Bertling et al., 1987). Densitometry analysis of band intensity was used to determine the frequency of 3’ extension/vector backbone for each homology length and invading vector arm as summarized in Figure 3.5.
Figure 3.3 Representative gel of 3' extension from the left invading arm. (Description on following page)
**Figure 3.3** Representative gel of 3’ extension from the left invading arm. The 3’ extension analysis is presented in (A) for the various plasmid vectors differing in homology on the left invading arm. A kinetic analysis is presented for the vector, pTΔCμ858/2290 (lanes 11-17), while in the various vectors shown in lanes 21-25, only the 3’ extension products that appear at the maximum 3 or 6 h time points are shown with listed vector. Densitometric analysis of band intensity followed by comparison with plasmid copy standards (lanes 2-8) was used in determination of the number of plasmid molecules bearing 3’ extensions. These values were compared with the number of recovered vector backbone molecules as determined from a similar analysis of the data presented in (B) and from this information, the frequency of 3’ extension/vector backbone was determined. The various controls in the 3’ extension assay include in (A) water (lane 9), an unelectroporated sample (lane 10), 1 μg of igm482 genomic DNA (lane 18), 10 ng of pTΔCμ858/2290 vector (lane 19) and the mixture of igm482 genomic DNA and pTΔCμ858/2290 vector (lane 20). In (B) plasmid standards are presented in lanes 2-9, while 1/10 dilutions of pTΔCμ858/2290, vector backbone samples in lanes 12-18, with other vectors in lanes 19-23, while a water control is present in lane 10 and finally an unelectroporated control in lane 11.

**Figure 3.4** Representative gel of 3’ extension from the right invading arm. The 3’ extension analysis is presented in (A) for the various plasmid vectors differing in homology on the right invading arm. A kinetic analysis is presented for the vector, pTΔCμ858/2290 (lanes 11-16), while in the vectors shown in lanes 21-25, only the 3’ extension products that appear at the maximum 3 or 6 h time points are shown with listed vector. Densitometric analysis of band intensity followed by comparison with plasmid copy standards (lanes 2-8) were used along with the backbone vector samples in (B) as above in Fig. 3.3, in determination of the number of plasmid molecules bearing 3’ extensions/vector backbone. Controls were repeated as in Fig. 3.3, though not pictured here. In (B) plasmid standards are presented in lanes 2-9, while 1/10 dilutions of pTΔCμ858/2290, vector backbone samples in lanes 12-17 with other vectors in lanes 18-19, while a water control is present in lane 10 and a unelectroporated control in lane 11.
Figure 3.4 Representative gel of 3’ extension from the right invading arm. (Description on preceding page)
As summarized in Fig. 3.5, 3’ extension events from the left invading arm show a linear correlation (R² value=.92). Furthermore, one-way ANOVA indicates that each homology length results in a 3’ extension product which is significantly different from the others (p <.001, Tukey’s at 99% significance).

The right arm homology data (Fig. 3.5) reveal a similar trend, with unique homology lengths again producing distinct 3’ extension values (ANOVA; p <.001) as well as a high linear correlation (R² value=.91), demonstrating a relationship between homology and 3’ extensions. The apparent difference between the slopes of the linear models for the left and right arm 3’ extensions is covered further in the discussion.

Another observation in the data is that 3’ extension events were detected in the pTΔCμ211/2290 vector, which contains only 211 bp of homology on the left arm though at an approximately 10-fold lower quantity (2.86x10⁻⁵±3.76x10⁻⁶) than in the next closest homology vector pTΔCμ517/2290 (2.60x10⁻⁴±7.93x10⁻⁵). This observation along with the linear model suggest that the minimal homology segment (MEPS) for 3’ extension may be between 517 and 211 bp.

The pTΔCμ858/567 vector has 567 bp of homology on the right arm and also the lowest overall homology (1.4 kb) of any vectors tested in the 3’ extension assay. This vector has a 3’ extension frequency on the right arm of 1.71x10⁻⁴±1.45x10⁻⁴ events/vector backbone and is similar to the frequency of 3’ extensions on the left arm of the vector pTΔCμ517/2290 which has homology, 517 bp homology on the left arm but a two times higher overall homology (2.8 kb). This suggests that the homology of the individual arm involved in directing the 3’ extension product is more significant in determining new DNA synthesis than the overall vector-borne homology.
Figure 3.5 Role of homology on 3’ extension. Analysis of 3’ extension events occurring on the left and right invading vector arms. For each homology point, the mean frequency of 3’ extensions ± the standard error of the mean is presented. At a minimum, three independent electroporations were used to calculate each mean presented here. The linear equation for the right arm is = 2E-06x - 0.001, while the linear equation for the left arm is = 3E-06x - 0.0011. Fit of the linear model is discussed in the preceding section.
Previous gene targeting experiments have established that a single nucleotide change within a sequence of limited homology can cause a significant decrease in HR events (Lukacsovich and Waldman, 1999). The mutant vector (p\(\Delta\)C\(\mu\)858/2290X) with a 4 bp insertion at a site 211bp from the DSB on the left arm was tested for the effect of a microheterology on the 3’ extension assay. However, a t-test showed there was no significant difference in left arm 3’ end extensions between the p\(\Delta\)C\(\mu\)858/2290 vector (1.33x10\(^{-3}\)±5.46x10\(^{-5}\)) and the p\(\Delta\)C\(\mu\)858/2290X (1.23x10\(^{-3}\)±2.5x10\(^{-4}\)) mutant (p=.72). This indicates that a minor sequence heterology is not significant in deterring 3’ extension events.

3.3 Effect of homology changes on one arm and their influence on 3’ extension from the opposite arm

In the previous section, a direct relationship between homology length and the frequency of 3’ extension from individual left and right invading vector arms was observed. In this section, the influence of homology changes on one arm and how they affect 3’ extension on the opposite arm was investigated. To examine 3’ extension from the left arm, vectors p\(\Delta\)C\(\mu\)858/2290, p\(\Delta\)C\(\mu\)858/1294 and p\(\Delta\)C\(\mu\)858/567 were used. In these vectors, homology on the left arm remained constant, while right arm homology decreased. Similarly, 3’ extensions from the right arm were measured using vectors p\(\Delta\)C\(\mu\)2120/2290, p\(\Delta\)C\(\mu\)2070/2290, p\(\Delta\)C\(\mu\)858/2290 and p\(\Delta\)C\(\mu\)211/2290, in which left arm homology was decreased but right arm homology was kept constant. As above, the frequency of 3’ extension (Fig. 3.6; Fig. 3.7) was measured as a function of recovered backbone samples.
Figure 3.6 Representative gels of left arm 3’ extensions occurring with homology changes to the right arm. The 3’ extension products are presented for the left arm, when homology on the right arm is altered (A). Lanes 2-8 (A) contain pTCμ plasmid vector copy standards, while lanes 11-17 contain pTΔCμ858/2290 3’ extension vector samples. Lanes 2-9 (B) contain pTCμ plasmid copy standards, while lanes 12-18 contain samples of pTΔCμ858/2290 backbone. In (A), lanes 19-21 contain the listed vectors at the peak in 3’ extension, while (B) contain the listed vectors backbone samples. In (A), lane 18 contains 1 μg genomic igm482 DNA, lane 19, 10 ng of gapped pTΔCμ858/2290 vector and lane 20 a mix of the two. Lane 9 contains water and lane 10 an unelectroporated sample in (A), lanes 10 and 11 contain the same in (B), respectively.
Figure 3.7 Representative gels of right arm 3' extensions occurring with homology changes to the left arm. In (A), the 3’ extension products for the right arm are shown for vectors with changes to the left arm homology. Lanes 2-8 (A) contain pTCμ plasmid vector copy standards, while lanes 11-17 contain pTΔCμ_{211/2290} 3’ extension vector samples. Lanes 2-9 (B) contain pTCμ plasmid copy standards, while lanes 12-18 contain samples of pTΔCμ_{211/2290} backbone. In (A), lanes 19-21 contains the listed vectors at the peak in 3’ extension, while (B) contain the listed vectors backbone samples. In (A), lane 18 contains a mix of 1 μg genomic igm482 DNA and 10 ng of gapped pTΔCμ_{858/2290} vector negative control. Lane 9 contains water and lane 10 an unelectroporated sample in (A), lanes 10 and 11 contain the same in (B), respectively.
Each vector was digested with *BstEII*, transfected into igm482 hybridoma cells and the frequency of 3’ extension/ vector backbone at peak 3 or 6 h time points was determined. Figure 3.6/3.7 presents representative gels for each of the vectors used, while the results of this study are summarized in Figure 3.8.

The data shows that 3’ extension events are not equivalent on both sides of the DSG. One-way ANOVA reveals that changing homology on the right arm does not have a significant effect on 3’ extension events occurring from the left side of the DSG (p = .10). This suggests that 3’ extensions from the left side of the DSG occur independently of right arm homology (Fig. 3.8, circles). In contrast, one-way ANOVA reveals a significant difference (p < .001) in 3’ extensions to the right of the DSG when changes in homology are made to the left arm. Increasing left arm homology leads to a decline the frequency of 3’ extensions from the right arm (Fig. 3.8, squares). The frequency of 3’ extensions for the pTΔCμ2120/2290 vector (1.66x10^{-4}±7.04x10^{-5}) which has the greatest homology on the left arm, is approximately 14-fold lower than that of the pTΔCμ211/2290 vector (2.27x10^{-3}±3.35x10^{-4}), which has the lowest homology on the left arm.
Figure 3.8 3' extensions measured opposite a homology change. The effect of homology changes in one vector arm on 3’ extensions from the opposite vector arm. Each point represents the mean frequency of 3’ extension/vector backbone ± standard error of the mean. Each point was determined based on at least 3 independent electroporations, except for pTΔCμ2070/2290 and pTΔCμ4858/567 in which only 2 were performed. One-way ANOVA revealed there was no significant differences in 3’ extension from the left invading arm in vectors with altered right arm homology (circles).
3.4 Length of 3’ extensions

The 3’ end extension assay samples for the pTΔCμ858/2290 and pTΔCμ2120/2290 vectors were used to determine the effect that homology on the left arm had on the length of 3’ extensions into the gapped region. A PCR reaction was performed on 3’ extension samples at the peak 6 h time point (Fig. 3.9), using primer pairs that extended 211 bp (neoF20’-CμR2) and 621 bp (neoF20’-CμR3) into the gap from the BstEII site (Fig. 2.2) and compared to vector backbone recovery (Fig. 3.9).

For both vectors, 3’ extension events decline with increased distance into the gapped region (Table 3.1). However, no consistent pattern emerged as to whether one vector behaved differently than the other, therefore, all that can be concluded is that an increase in homology does not lead to an increase in 3’ extension tract length.

<table>
<thead>
<tr>
<th>Vector</th>
<th>Frequency of 3’ extension 63 bp into gap(^a)</th>
<th>Frequency of 3’ extension 211 bp into gap(^b)</th>
<th>Frequency of 3’ extension 621 bp into gap(^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pTΔCμ858/2290</td>
<td>1.33x10(^{-3})±5.46x10(^{-5})</td>
<td>1.28x10(^{-3})±2.01x10(^{-4})</td>
<td>2.51x10(^{-4})±1.63x10(^{-4})</td>
</tr>
<tr>
<td>pTΔCμ2120/2290</td>
<td>5.31x10(^{-3})±6.83x10(^{-4})</td>
<td>1.10x10(^{-3})±1.56x10(^{-4})</td>
<td>7.74x10(^{-5})±8.1x10(^{-6})</td>
</tr>
</tbody>
</table>

\(^a\)Samples representing mean 3’ extension events/vector backbone ± standard error of the mean, based on 3 independent electroporations. Based on the data presented in Fig. 3.5.

\(^b\)Samples representing mean 3’ extension events/vector backbone ± standard error of the mean, based on 2 independent electroporations
Figure 3.9 Length of left arm 3’ extension products. The 3’ extension products for the left arm, 211 bp into the DSG for pTΔCμ858/2290 (A), left arm 3’ products 621 bp into the DSG (B). Lanes 2-8 contain pTCμ plasmid standards, lanes 11-16 contain pTΔCμ858/2290 (A and B) vector samples, while lane 17 (A and B) contains peak 3’ extensions for pTΔCμ2120/2290. In (A) lanes 19-21 contains the listed vectors at the 3’ extension peaks, (B) contains the listed vectors in lanes 21-22. In (A and B), lane 9 contains water and lane 10 an unelectroporated sample. The backbone gel presented in (C), lanes 2-9 contain pTCμ standards, lanes 11-17 pTΔCμ858/2290 backbone samples, while lane 18 contains the pTΔCμ2120/2290 sample. Water and unelectroporated samples are in lanes 9 and 10.
3.5 Excess Rad51 increases 3’ extension events

As described in Section 1.3, Rad51 plays an essential role in the early homology search and strand invasion steps of the HR repair process (reviewed in Li and Heyer, 2008). Therefore, it is relevant to examine the effect of Rad51 on the 3’ extension process. For these studies, the igm482-derived cell line 51-5 was used (Rukšć et al., 2007), with comparisons made to data in igm482 (from Fig. 3.5). Cell line 51-5 expresses N-terminal FLAG-tagged mouse Rad51 at a level approximately 2-fold higher than the endogenous level of Rad51 (Rukšć et al., 2007). Two vectors differing in their left arm homologies were used to assess the efficiency of 3’ extension in 51-5, namely pTΔCμ211/2290 and pTΔCμ720/2290. As in the other experiments, the frequency of 3’ extension events/vector backbone was determined by measuring the band intensity for the various peak 3’ extensions and comparing them to the quantity of vector backbone recovered (Fig. 3.10).

As summarized in Figure 3.11, both vectors showed a significant approximately 3-fold increase in 3’ extension events in the Rad51 over-expresser 51-5 cell line compared to the control igm482 cells (taken from Fig. 3.5). A t-test comparison revealed that the frequencies of 3’ extension in the 51-5 cell line are significantly higher than in igm482 for both the pTΔCμ211/2290 (p=.02) and pTΔCμ720/2290 (p=.01) vectors.

Also given the 51-5/igm482 3’ extension ratios for the pTΔCμ211/2290 and pTΔCμ720/2290 vectors, a t-test was performed determining there was no interaction between increased Rad51 concentration and homology (p = .92). The results suggest that Rad51 and homology both play a role in 3’ extension, although they appear to affect the process independently, at least for the homology lengths under examination here.
Figure 3.10 Representative gels of 3' extensions in 51-5. In (A), the 3’ extension products for the left arm are shown for vectors in the 51-5 cell line. Lanes 2-8 (A) contain pTCμ plasmid vector copy standards, while lanes 11-17 contain pTΔCμ720/2290 3’ extension vector samples. Lanes 2-9 (B) contain pTCμ plasmid copy standards, while lanes 12-18 contain samples of pTΔCμ720/2290 backbone. In (A), lane 19 contains pTΔCμ211/2290 at the peak in 3’ extension, while (B) contains the backbone sample. In (A), lane 18 contains a mix of 1 μg genomic igm482 DNA & 10 ng of gapped pTΔCμ858/2290 vector negative control. Lane 9 contains water and lane 10 an unelectroporated sample in (A), lanes 10 and 11 contain the same in (B), respectively.
Figure 3.11 Effect of excess Rad51 and homology length on 3’ extension. The effect of the Rad51 increase on the efficiency of 3’ extensions for the pTΔCμ_{211/2290} vector (211 bp left arm homology) on the left (A), while the pTΔCμ_{720/2290} vector (720 bp left arm homology) on the right (B). The values represent the mean frequency of 3’ extension/ vector backbone ± the standard error of the mean of 3 independent electroporations.
3.6 Effect of gap size on 3’ extensions

Previous homology studies have suggested that gap length does not have a major effect on HR repair events, except in the sense that they cause a reduction in overall homology, thus indirectly affecting HR (Valancius and Smithies, 1991). To investigate the effect of gap size on 3’ extension, the vector pTΔBamCµ was generated (Fig. 2.3). This vector bears a gap length of 623 bp, along with left and right homology arms of 876 bp and 2.83 kb respectively. The frequency of 3’ extension in the pTΔBamCµ vector was determined at peak time points and then compared to previous data for pTΔCµ858/2290 (Fig. 3.5). As in the studies above, the frequency of 3’ extension for the pTΔBamCµ vector was determined relative to the vector backbone recovery.

As summarized in Figure 3.12, the 3’ extension events occur at a higher frequency on both sides of the 623 bp DSG in the pTΔBamCµ vectors. In pTΔBamCµ, 3’ extension events on the left arm are 24-fold higher due to the decreased gap size (p < .01), while on the right arm, they are 15-fold higher (p < .01). A t-test, however, shows the 24-fold and 15-fold ratios for 3’ extensions in the smaller gapped vector are not significantly different (p=.22). Therefore, smaller gap size in the pTΔBamCµ vector leads to an equal increase in the 3’ extension frequency from both sides of the DSG.
Figure 3.12 Influence of gap size on 3’ extension events

3’ extension events on the left and right side of a DSG are shown for the pTΔCμ858/2290 vector (white bars) which contains a 1165 bp gap flanked by 858 bp and 2.3 kb regions of homology on the left and right arms respectively, and the pTΔBamCμ vector (dark bars) which contains a 623 bp gap flanked by 876 bp and 2.8 kb regions of homology on the left and right arms respectively. A minimum of 3 independent electroporations were used to determine the mean 3’ extension events/vector backbone ± standard error of the mean, except for the right arm pTΔBamCμ data, which used 2 electroporations.
CHAPTER 4: Discussion

4.1 DNA titration studies in the 3’ extension assay

The 3’ extension assay provides a novel method for analysis of the early stages of mammalian HR (Si et al., 2010). A first set of experiments examined the DNA requirement for 3’ extension using the vector pTΔCμ858/2290. The process of 3’ extension demonstrates an approximately linear trend between 25 μg and 100 μg of electroporated vector, coinciding with previously reported data in electroporation studies (Bertling et al., 1987). Previous gene targeting assays have used between 5-50 μg (Baker et al., 1988; Johnson et al., 1999; Meyer et al., 2010). Similarly, the 3’ extension assay was initially developed using 50 μg of introduced vector (Si et al., 2010). Thus, these values fall within the linear range where DNA quantity is not perturbing 3’ extension, suggesting that ~50 μg of DNA is optimal.

The inhibitory effect of plasmid DNA observed in the 3’ extension assay might, as previously suggested, be related to the larger plasmid molecules undergoing less rapid diffusion into a cell (Shulman et al., 1990). The excess plasmid molecules used here may produce a clogging effect on cellular membrane pores. As cell viability did not appear to be affected by the excess plasmid DNA, the explanation for the inhibition of 3’ extension events may lie in the state of plasmid DNA electroporated into cells. Excess damaged plasmid DNA may be entering the cells causing an overwhelming effect on multiple DNA repair systems, potentially leading to the frequency of HR to be reduced.

It should be noted that the 3’ extension assay currently has no way to account for plasmid which may undergo 3’ extension, but is engaged in an intermediate that may be lost during the
extraction process. For instance, the double HJ intermediate leads to covalent bond formation between the plasmid and chromosomal template, so while this plasmid may undergo 3’ extension, it would not be recovered and detected by the assay.

4.2 The effect of homology on 3’ extension

Having established the optimal concentration of gapped vector DNA to use in the 3’ extension assay, the influence of homology on the frequency of 3’ extension was examined. The results revealed an approximately linear relationship between the frequency of 3’ extension and homology lengths ranging from approximately 500 bp-2.29 kb, although slopes between left and right arms were different (discussed further below). The approximately linear relationship between homology and 3’ extension helps explain the ~3 fold stimulation of 3’ extension for the 2.29 kb “right” arm versus the .86 kb “left” arm in the pTΔCμ_{858/2290} vector reported previously (Si et al., 2010).

Gene targeting studies in mammals have established that increased homology leads to an increase in HR in what initially appeared to be a “more than linear” relationship (Shulman et al., 1990) and was later reported as being exponential (Deng and Capecchi, 1992). Hasty et al., (1991) were able to show that shorter tracts of homology (<536 bp) were able to initiate the formation of a Holliday junction (necessary early step in HR), but if total homology was insufficient, a gene targeting event indicative of successful HR would not take place. This data suggests that homology at low levels is sufficient for the homology search during HR, but the intermediate formed is not stable enough for extended DNA synthesis to occur. For completion of HR greater homology is likely required. Larger regions of homology may be expected to have a
greater number of pairing sequences in the chromosome and thus may be expected to undergo more HR events, but the increased plasmid size may also hinder diffusion across cellular membranes (Shulman et al., 1990).

A statistical model to explain the relationship between homology and HR has been proposed previously (Fujitani et al., 1995). In this model, the HJ intermediate formed by the initial 3’ end invasion will experience random branch migration (Thompson et al., 1976), and upon migration to a region of non-homology (such as the vector backbone), the intermediate is destroyed. On the other hand, if this destructive branch migration is preceded by sufficient DNA synthesis and resolution of the HJ, then a successful HR event will have taken place. The authors state that a DSB occurring within a larger region of homology would have a lower probability of encountering the non-homologous region, thus a greater chance at a successful HR repair (Fujitani et al., 1995). This model suggests that increased homology will stimulate 3’ extension as well as HR.

Below about 500 bp, the frequency of 3’ extension events decreased significantly, suggesting that homology lengths below this value may constitute the minimal size for efficient 3’ extension (MEPS in HR studies). During gene targeting in various mammalian cell lines, MEPS values ranging from approximately 500 bp-1 kb have been suggested (Shulman et al., 1990; Hasty et al., 1991). Therefore, the MEPS values in gene targeting are in rough agreement, although perhaps slightly larger than those required for 3’ extension. As explained above, the relationship between homology length and gene targeting in mammalian cells is expected to be complicated and any determination of a MEPS value for either gene targeting or 3’ extension is likely to be highly dependent on the cell type, the degree of isogenicity of the DNA substrates and the particular locus being examined. That individual homology on each vector arm seems more im-
important in the 3’ extension assay than total overall homology is suggested from the similar frequencies of 3’ extension from the left and right invading arms in the vectors pTΔCμ₅₁₇/₂₂₉₀ and pTΔC₅₈₅₈/₅₆₉, respectively, in which overall homology to the chromosomal μ locus in the former vector is 1.4 kb, while in the latter, is 2.8 kb. Therefore, in addition to the variables highlighted above, during gene targeting, there may be a requirement for larger overall homology to stabilize a recombination intermediate.

Greater tracts of homology may play a more important role in gene targeting studies, as they need to not only initiate 3’ end DNA synthesis, but also maintain the stability of the repair complex, to allow for the gene targeting event. The 3’ extension assay on the other hand may depend less on overall stability to initiate the early extension events, leading to a more simplified linear relationship between homology and HR.

It should be noted that while 3’ extension events are detected from both the left and right invading vector arm, it is not known whether these reside on the same or different vector molecules. This has not yet been determined as there are problems associated with detecting 3’ extension from both ends of the same vector. PCR products to test extensions from both sides of the same vector would need to be at least 7 kb for even the smallest vector. Another issue is that as 3’ extension occurs from only from the 3’ end, each strand will only have amplification of one side of the extension. This would then require annealing and extension of newly-primed segments and this PCR-mediated recombination is not observed in any of the controls (unelectroporated vector plus genomic DNA), or previously (Si et al., 2010). Finally, circular products, indicating complete HR repair, were also not detected previously (Si et al., 2010).
4.3 The effect of opposite arm homology on 3’ extension

As discussed in the previous section, the various vectors used to measure 3’ extension feature one arm in which the homology varies in size, while the homology on the other arm remains constant. Using these vectors 3’ extension events can be detected on both sides of the vector-borne DSG. These vectors were used to investigate how changes in one homology arm affect 3’ extension from the arm in which homology remains constant.

The results showed that changing homology on the right arm had no discernible effect on the 3’ extensions from the left arm. This suggests that 3’ extension events are being initiated on the left arm independently of 3’ extension on the right homology arm. In contrast, 3’ extension events from the right invading vector arm change when homology on the left arm is altered: when homology on the left arm is highest, 3’ extension events on the right arm are very low. This suggests that an increase in homology on the left arm causes an inhibition of 3’ extension events from the right arm and therefore, 3’ extensions from the left invading vector arm are favored.

It has been reported in yeast that Rad51 binds preferentially to GT-rich sequences (Tracy et al., 1997), with the excess Rad51 filaments presumably leading to an increase in strand invasion and 3’ extensions (similar to the effect in Fig. 3.9). However, it is not the case here as left arm 3’ extensions are observed at a higher frequency, while the right arm actually contains a higher percentage of GT di-nucleotides than the left arm (9.9 % vs. 5.7 %).

A possible explanation for the drastically reduced right arm 3’ extensions in the pTΔCµ2120/2290 vector might be the presence of homology traps in repeated sequences that prevent useful HR repair configurations (Fulconis et al., 2006). Homology traps occur in sequences
containing repeats, whereby the substrates pair in the wrong pairing frame, leading to rather stable binding and preventing the homology search until this pairing is broken. The residual Sμ region may provide an explanation for the low 3’ extension frequency in the vector pTΔCμ2120/2290. However, the pTΔCμ2070/2290 vector, which has 50 bp lower left arm homology and does not contain the tk gene or Sμ repeats contained in the pTΔCμ2120/2290 vector, shows the same 3’ extension frequency. This suggests that Sμ homology traps are not causing the disruptive effect.

Another possibility may be that the increase in homology on the left arm, increases complementary base pairing with the target chromosomal Cμ region, such that the vector backbone, being of fixed length, may not be long enough to adequately position the right homology to permit 3’ extension. However, the fact that greater homology on the right arm does not lead to a decline in 3’ extension (Fig. 3.7) from the left arm makes this possibility less likely.

Formally, the 3’ extension events resemble gene conversion, a one-way transfer of genetic information (Pâques and Haber, 1999) Bias during gene conversion in S. cerevisiae has been observed previously (Palmer et al., 2003), with the bias being directional, in favour of polar processes, such as transcription and replication. A study of mouse meiotic recombination hotspots concluded that they were often found near transcriptional start sites (Smagulova et al., 2011). Similarly, homology bias was observed in S. cerevisiae, at the MAT locus on one side of a heterologous region that was to be replaced (Coïc et al., 2011). In this experiment the authors were able to show an increase in donor preference with increased homology on one side flanking the heterologous sequence, while the other side maintained its homology. The authors suggested that homology is selected preferentially because that side of a DSB is able to initiate DNA synthesis. Bias has also previously been observed during HJ resolution involved in integration
events of the Cμ region using the mouse hybridoma cells in the lab (McCulloch and Baker, 2006).

The above studies may help in explaining the bias in 3’ extension from the left invading vector arm, because at the immunoglobulin μ gene target, 5’ → 3’ polymerization from the left invading vector arm occurs in the same direction and from the same strand as transcription by RNA polymerase II. This bias may also account for the difference in slopes between the left and right arm linear models (Fig. 3.5). Assuming a bias in favour of the initiation of 3’ extension events from the left side of the DSB, there may be an increase in double HJ intermediates (discussed in Section 4.1) which may be indicative of 3’ extension events from the right arm but are not detected by the assay.

4.4 Effect of homology on the length of 3’ extension products

The pTΔCμ2120/2290 and pTΔCμ858/2290 vectors were tested in the 3’ extension assay to determine if homology affected the length of 3’ extension products for the left arm. Both vectors show that 3’ extensions decline rapidly with increased distance into the gapped region (Table 3.1). The results suggest that while homology has a large effect on initiating 3’ extension events, it does not necessarily affect the completion of HR gap repair.

The lack of sequence extension further into the gap may be due to the polymerases involved in HR repair. Previously, reduced polymerase processivity has been suggested during DNA repair synthesis in S. cerevisiae (Wang et al., 2004). The TLS family polymerases were mentioned earlier as playing a role in HR repair (Kawamoto et al., 2005) and they show low
processivity compared to the main replicative polymerases δ and ε. It is possible that a TLS polymerase may be responsible for short-tract DNA synthesis. It has also been suggested that Rad51 donor sequence interactions are attempted on average 4 times before a successful HR event is carried out (Coïc et al., 2011). Thus, short tracts of DNA may be synthesized by the unstable Rad51 filament-donor interaction, with longer tracts being synthesized less frequently due to the instability.

4.5 Effect of increased Rad51 and homology on the efficiency of 3’ extension

During HR repair, Rad51 will dissociate from the invading strand, following which DNA synthesis will occur (Li and Heyer, 2008). This stage of the process is expected to coincide with the product that is measured by the 3’ extension assay. Therefore, the effect of Rad51 on 3’ extension was investigated with the expectation that 3’ extension would be stimulated by excess wild-type Rad51.

In the Rad51 over-expressing 51-5 cell line, the frequency of 3’ extension was approximately 3-fold higher than control igm482 cells at the two different left arm homology lengths (211 bp and 720 bp) that were tested. Thus, the results demonstrate stimulation of 3’ extension by Rad51, with no apparent synergistic interaction between Rad51 and the two homology lengths tested here.

According to the HR models, the broken ends of the transformed vector must be resected to form 3’ ended single-stranded DNA in order to provide Rad51 binding sites (Li and Heyer, 2008). Therefore, the effect of Rad51 over-expression being homology independent, between the
211-720 bp range, might suggest that 3’ end resection is only occurring to a distance that is common to both plasmids (<211bp). Alternatively, resection may encompass all of the available homology in the two vectors with the result that they show the same dependence on Rad51 and an effect of homology may only be seen at a homology length greater than the 720 bp studied here.

The latter view would be consistent with previous studies showing that up to 1 kb ssDNA overhangs are formed in HR repair of DSBs (Cao et al., 1990; Zenvirth et al., 2003; Garcia et al., 2011). Thus, it will be important to determine whether the 3’ fold stimulation in 3’ extension is maintained when longer stretches of homology are used.

The protein Brca2 is responsible for stabilization of a filament of up to 6 Rad51 molecules and can load Rad51 in clusters onto available resected 3’ ssDNA (Jensen et al., 2010). Following the deposition of these clusters, Rad51 protein might self-associate to coat the entire filament, to carry out its function in HR repair. It is possible that the length of homology studied here is not significant enough to reveal a possible synergistic effect with Rad51 concentration. That is, a coated filament between 211 bp-720 bp might be the minimum that is needed to achieve an ~3-fold stimulation in 3’ extension.

A previous study using the murine hybridoma 51-5 cell line, used here to test Rad51 over-expression, showed gene targeting is stimulated by increased Rad51 (Rukšé et al., 2007). Studies using Chinese hamster ovary cells that moderately over-express mRad51 by 2-3 fold, showed an increase in intrachromosomal HR events and resistance to ionizing radiation, which causes DSBs in the genome (Vispé et al., 1998; Lambert and Lopez, 2000). These studies seem to agree with the results obtained in the 3’ extension assay regarding Rad51 over-expression.
However, in *S. cerevisiae* over-expression of Rad51 reduced DSB-induced HR frequency and cell growth upon DSB induction, with higher levels causing a more drastic effect (*Paffett et al.*, 2005).

The study by *Lambert and Lopez* (2000), determined that although HR events were increased, the overall number of DSBs repaired was not significantly affected by Rad51 over-expression. This suggests that Rad51 has an important role in DSB repair pathway determination, causing less DSBs to be repaired by the NHEJ mechanism and more by the HR repair mechanisms. The difference between the effects of Rad51 in yeast and mammals may be accounted for by the important role Rad51 is playing in determining whether the NHEJ or an HR mechanism will be used and the fact that NHEJ is not necessary for yeast viability (*Lees-Miller and Meek*, 2003). This means that yeast may already be operating at peak HR repair efficiency, with their wild-type levels of Rad51 and the excess may cause an inhibitory effect, possibly by becoming more difficult to remove post-synapsis, to begin DNA synthesis.

### 4.6 Increase 3’ extension efficiency when the size of the vector-borne DSG is reduced

The idea that total homology may function more efficiently if the vector contained a smaller double-stranded gap was tested in the 3’ extension assay. The results revealed that reducing the size of the DSG resulted in a drastic increase in the number of 3’ extension events. Previous studies on the effect of reducing gap length in gene targeting have shown a minimal effect, if any (*Valancius and Smithies*, 1991; *Baker et al.*, 1996), especially when compared to the approximately 20-fold increases observed with a smaller gap in the 3’ end extension assay.
The effect of gap size may be more noticeable in the 3’ extension assay, because it allows for greater stability of the early repair complex between the broken sequence and the template. One issue with discussing the potential mechanism in use is whether the left and right arm extensions are occurring on the same vector, as mentioned earlier (Section 4.3). If so, the smaller gap may allow for the effects of total homology to act on stabilizing the repair complex, increasing extension events on both sides of the DSB, possibly by more efficient alignment and annealing of the broken strand.

4.7 A model for 3’ extensions

There are potentially two mechanisms of homologous recombination, which can lead to the 3’ extension product detected by this study. Under either the SDSA (Fig. 1.3) or DSBR (Fig. 1.2) mechanism it is possible for a 3’ tail end to invade a homologous donor duplex DNA sequence and initiate DNA synthesis. Thus both models predict the priming of new DNA synthesis by DNA ends positioned to the left and right of the DSG, as detected in the 3’ extension products here (Fig. 4.1 A). A second possibility, under the SDSA mechanism, would occur when one arm has completed some degree of synthesis across the gap or DSB, then serving as template for the non-invading strand (Nasmyth, 1984; Fig. 4.1 B). The final possibility for the generation of a 3’ end extension product occurs through the DSBR mechanism. If the non-invading broken strand is able to capture the D loop, then DNA synthesis can occur, forming a 3’ extension product (Orr-Weaver et al., 1981; Fig. 4.1 C).
Figure 4.1 Possible generation of newly synthesized DNA. A DSG vector will search out a homologous template and in (A) the early stage of HR repair common to both mechanisms is shown, on one side of a sequence damaged by a DSB, the 3’ end will invade a homologous template sequence and begin synthesis. If the synthesis continues until the location of the DSB and the invading strand then detaches, the SDSA model is in use (B), the non-invading strand will then initiate synthesis from the now template, invading strand. However, if the non-invading strand captures the D-loop of the homologous sequence DSBR is in use (C), and synthesis simultaneous to the invading arm in (A) will occur. The size of the arrows indicates a relative quantity of events, DNA synthesis shown as dotted lines while the extended black arrows represent 3’ extension products detected by the assay.
A provision of both models is that new DNA synthesis may be occurring from both invading 3’ ends of a single molecule. Unfortunately, as described earlier (Section 4.2) there is not yet a way to discern whether tracts of new DNA are being synthesized from the same vector molecule. This makes determining which HR mechanism is in use during the 3’ extension assay difficult.

Some of the evidence presented in this thesis and previous study of the 3’ extension assay seems to support the notion the SDSA mechanism is favoured. The short tracts formed in the 3’ extension assay (Table 3.1), rarely forming the ~600 bp product that would be needed for second-end capture, make the DSBR model unlikely. It may also be the case that the DSBR model would require a more stable intermediate to form the double Holliday junction, again making SDSA more likely. Finally, as has also been demonstrated previously (Si et al., 2010), the products generated from the 3’ extension assay are all gene conversion products arising from templated extension of an invading 3’ end and are all non-crossover products, a trait common to the SDSA mechanism. However, it is likely that many of the 3’ extension products in this assay are produced during the phase common to both HR mechanisms discussed here (Fig. 4.1 A), with the observation of predominantly short products being formed in the 3’ extension assay (Si et al., 2010; Table 3.1). It is possible the transcriptionally active Cμ locus being used here will disrupt long tract synthesis during HR repair. It is also possible random branch migration will lead to dissociation of the invading strand before a completed HR product has been formed (Fujitani et al., 1995); supported by the observation that Rad51 donor sequence interactions occur up to 4 times before a successful HR event is carried out (Coic et al., 2011). The results here suggest that 3’ extension products undergoing extensive DNA repair will occur via the SDSA mechanism, but general 3’ products are short and occur at a step common to the HR mechanisms.
CHAPTER 5: Future Studies

The 3’ extension assay provides a novel and unique way to look at previously well established HR results. One issue with the assay mentioned earlier was the current uncertainty regarding potential double HJ intermediates, which would be indicative of 3’ extension, that are lost in the plasmid extraction process. If these samples could be identified and quantified the frequency of 3’ extension events would be higher than has been reported here and previously (Si et al., 2010). The ability to check 3’ extensions from both sides of a DSB especially has already yielded an interesting inhibitory effect on just one arm, suggesting possible bias based on the transcriptional direction of the μ heavy chain gene used here. One way to check this would be to swap the direction of the Cμ fragment with respect to the vector backbone, such that transcription would occur from the “right” side of the DSG.

With regards to the relationship between Rad51 over-expression and homology length, it may be useful to test a vector with much greater homology than the (211 bp and 720 bp) fragments tested here, as there may indeed be a co-operative effect between homology and Rad51 that is only visible at greater lengths of homology. To complement the results observed in the 3’ extension assay when Rad51 is over-expressed, it may be interesting to determine the effects of decreased Rad51 in the assay. Testing the effect of decreased Rad51 may be more difficult to test or quantify, as Rad51 knockouts are lethal in mammals (Tsuzuki et al., 1996), though stable Rad51 siRNA knockdowns have been established in the hybridoma system and show a reduction in HR events (Malysewich, 2009). The effects of knocking down Rad51 may not have an effect with reduced homology vectors, as these vectors may already require lower amounts of Rad51 to initiate the homology search and 3’ end invasion. However, based on the results presented here
on Rad51 over-expression, it is likely that the knockdowns would show a homology independent reduction in 3’ extension events.

Reducing gap length to 623 bp in the transfected vector led to a fairly substantial increase in 3’ extension events. Therefore, it would be of interest to determine the distance 3’ extension events are proceeding into the smaller gap. This may provide some information into the mechanism of 3’ extension. If the 3’ extension products are not extending very far into the smaller gap, the increase in the 3’ extensions may be accounted for by a low processivity DNA polymerase (Wang et al., 2004) and support an SDSA mechanism of HR.
Literature Cited


Reardon, J. T., Thompson, L. H. and Sancar, A. (1997). Rodent UV-sensitive mutant cell lines in complementation groups 6-10 have normal general excision repair activity. Nucleic acids res. 5, 1015-1021.


Appendix I

Left arm ANOVA & linear correlation

<table>
<thead>
<tr>
<th>Samples</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>16</td>
</tr>
<tr>
<td>$\Sigma X$</td>
<td>0.000086</td>
<td>0.000521</td>
<td>0.002122</td>
<td>0.00533</td>
<td>0.01593</td>
<td>0.023989</td>
</tr>
<tr>
<td>Mean</td>
<td>0.000029</td>
<td>0.000281</td>
<td>0.000531</td>
<td>0.001333</td>
<td>0.00531</td>
<td>0.001499</td>
</tr>
<tr>
<td>$\Sigma X^2$</td>
<td>0</td>
<td>0</td>
<td>0.000001</td>
<td>0.000007</td>
<td>0.000087</td>
<td>0.000096</td>
</tr>
<tr>
<td>Variance</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.000004</td>
</tr>
<tr>
<td>Std.Dev.</td>
<td>0.000007</td>
<td>0.000112</td>
<td>0.000123</td>
<td>0.000108</td>
<td>0.001185</td>
<td>0.001998</td>
</tr>
<tr>
<td>Std.Err.</td>
<td>0.000004</td>
<td>0.000068</td>
<td>0.000061</td>
<td>0.000054</td>
<td>0.000684</td>
<td>0.0005</td>
</tr>
</tbody>
</table>

ANOVA Summary Table:

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment [between groups]</td>
<td>0.000057</td>
<td>4</td>
<td>0.00014</td>
<td>53.98</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Error</td>
<td>0.000003</td>
<td>11</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ss/Bl</td>
<td>0.00006</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Linear Regression Table:

<table>
<thead>
<tr>
<th>r</th>
<th>r^2</th>
<th>Slope</th>
<th>Y Intercept</th>
<th>Std. Err. of Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.959</td>
<td>0.92</td>
<td>0</td>
<td>-0.0011</td>
<td>0.0006</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>t</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>12.704</td>
<td>14</td>
<td>one-tailed &lt;.0001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>two-tailed &lt;.001</td>
</tr>
</tbody>
</table>

0.95 and 0.99 Confidence Intervals of rho

<table>
<thead>
<tr>
<th>Lower Limit</th>
<th>Upper Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.95</td>
<td>0.984</td>
</tr>
<tr>
<td>0.99</td>
<td>0.84</td>
</tr>
</tbody>
</table>

Values entered:

<table>
<thead>
<tr>
<th>Pair</th>
<th>X</th>
<th>Y</th>
<th>Residuals</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>211</td>
<td>2.55E-5</td>
<td>0.001</td>
</tr>
<tr>
<td>2</td>
<td>211</td>
<td>3.51E-5</td>
<td>0.001</td>
</tr>
<tr>
<td>3</td>
<td>211</td>
<td>2.42E-5</td>
<td>0.001</td>
</tr>
<tr>
<td>4</td>
<td>517</td>
<td>1.81E-4</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>517</td>
<td>3.4E-4</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>720</td>
<td>6.26E-4</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>720</td>
<td>6.17E-4</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>720</td>
<td>5.17E-4</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>720</td>
<td>3.62E-4</td>
<td>-0.001</td>
</tr>
<tr>
<td>10</td>
<td>858</td>
<td>1.22E-3</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>858</td>
<td>1.43E-3</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>858</td>
<td>1.42E-3</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>858</td>
<td>1.26E-3</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>2120</td>
<td>4.03E-3</td>
<td>-0.001</td>
</tr>
<tr>
<td>15</td>
<td>2120</td>
<td>5.39E-3</td>
<td>0</td>
</tr>
<tr>
<td>16</td>
<td>2120</td>
<td>6.37E-3</td>
<td>0.001</td>
</tr>
</tbody>
</table>

-Left arm homology entered as X, 3’ extensions/vector backbone as Y.
Right arm ANOVA & linear correlation

Data Summary

<table>
<thead>
<tr>
<th>Samples</th>
<th>N</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>ΣX</td>
<td></td>
<td>0.000514</td>
<td>0.003020</td>
<td>0.01645</td>
<td></td>
<td></td>
<td>0.014893</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>0.000171</td>
<td>0.00131</td>
<td>0.003483</td>
<td></td>
<td></td>
<td>0.001655</td>
</tr>
<tr>
<td>ΣX²</td>
<td></td>
<td>0</td>
<td>0.000006</td>
<td>0.006037</td>
<td></td>
<td></td>
<td>0.006043</td>
</tr>
<tr>
<td>Variance</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td>0.000002</td>
</tr>
<tr>
<td>Std.Dev.</td>
<td></td>
<td>0.000251</td>
<td>0.000543</td>
<td>0.006645</td>
<td></td>
<td></td>
<td>0.001522</td>
</tr>
<tr>
<td>Std.Err.</td>
<td></td>
<td>0.000145</td>
<td>0.000314</td>
<td>0.000372</td>
<td></td>
<td></td>
<td>0.000507</td>
</tr>
</tbody>
</table>

standard weighted-means analysis

ANOVA Summary

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>0.000017</td>
<td>2</td>
<td>0.000008</td>
<td>32.92</td>
<td>0.000583</td>
</tr>
<tr>
<td>Error</td>
<td>0.000002</td>
<td>6</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ss/Bl</td>
<td>Graph Maker</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>0.000019</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

SS/Bl = Subjects or Blocks depending on the design. Applicable only to correlated-samples ANOVA.

Data Summary

ΣX = 12462  ΣX² = 21740634
ΣY = 0.0149  ΣY² = 0
ΣXY = 29.3178

<table>
<thead>
<tr>
<th>r</th>
<th>r²</th>
<th>Slope</th>
<th>Y Intercept</th>
<th>Std. Err. of Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.054</td>
<td>0.909</td>
<td>0</td>
<td>-0.001</td>
<td>0.0005</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>t</th>
<th>df</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.385</td>
<td>7</td>
<td>one-tailed 0.0001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>two-tailed 0.0001</td>
</tr>
</tbody>
</table>

0.95 and 0.99 Confidence Intervals of rho

<table>
<thead>
<tr>
<th>Lower Limit</th>
<th>Upper Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.95</td>
<td>0.792</td>
</tr>
<tr>
<td>0.99</td>
<td>0.677</td>
</tr>
</tbody>
</table>

Values entered:

<table>
<thead>
<tr>
<th>Pairs</th>
<th>X</th>
<th>Y</th>
<th>Residuals</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>659</td>
<td>2.78E-05</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>569</td>
<td>2.56E-05</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>569</td>
<td>4.61E-04</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>1294</td>
<td>1.93E-03</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>1294</td>
<td>9.19E-04</td>
<td>-0.001</td>
</tr>
<tr>
<td>6</td>
<td>1294</td>
<td>1.08E-03</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>2291</td>
<td>2.89E-03</td>
<td>-0.001</td>
</tr>
<tr>
<td>8</td>
<td>2291</td>
<td>3.39E-03</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>2291</td>
<td>4.17E-03</td>
<td>0.001</td>
</tr>
</tbody>
</table>

-Right arm homology entered as X, 3’ extensions/vector backbone as Y.
**Opposite Left arm & Opposite Right arm ANOVA**

<table>
<thead>
<tr>
<th>Samples</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>N</strong></td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td></td>
<td>13</td>
</tr>
<tr>
<td><strong>ΣX</strong></td>
<td>0.00064</td>
<td>0.00025</td>
<td>0.00924</td>
<td>0.0968</td>
<td></td>
<td>0.01840</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>0.000166</td>
<td>0.000213</td>
<td>0.002747</td>
<td>0.00227</td>
<td></td>
<td>0.001416</td>
</tr>
<tr>
<td><strong>ΣX²</strong></td>
<td>0</td>
<td>0</td>
<td>0.000024</td>
<td>0.000022</td>
<td></td>
<td>0.000046</td>
</tr>
<tr>
<td><strong>Variance</strong></td>
<td>0</td>
<td>0</td>
<td>0.000001</td>
<td>0</td>
<td></td>
<td>0.000002</td>
</tr>
<tr>
<td><strong>Std.Dev.</strong></td>
<td>0.000141</td>
<td>0.000155</td>
<td>0.000726</td>
<td>0.000867</td>
<td></td>
<td>0.001286</td>
</tr>
<tr>
<td><strong>Std.Err.</strong></td>
<td>0.00007</td>
<td>0.00011</td>
<td>0.000419</td>
<td>0.000335</td>
<td></td>
<td>0.000357</td>
</tr>
</tbody>
</table>

**standard weighted-means analysis**

**ANOVA Summary**

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment [between groups]</td>
<td>0.000017</td>
<td>3</td>
<td>0.000006</td>
<td>20.99</td>
<td>0.00212</td>
</tr>
<tr>
<td>Error</td>
<td>0.000002</td>
<td>9</td>
<td>0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Total**

| SS/Bl | 0.000002 | 12 |

**Opposite left arm table includes** \( pT\Delta C_{\mu_{2120/2290}} \) as sample 1, \( pT\Delta C_{\mu_{2070/2290}} \) as sample 2, \( pT\Delta C_{\mu_{858/2290}} \) as sample 3 and \( pT\Delta C_{\mu_{1217/2290}} \) as sample 4.

**Opposite right arm table includes** \( pT\Delta C_{\mu_{858/2290}} \) as sample 1, \( pT\Delta C_{\mu_{858/1294}} \) as sample 2, \( pT\Delta C_{\mu_{858/569}} \) as sample 3.
51-5 vs igm 482 T-test pT\text{AC}μ_{720/2290} and pT\text{AC}μ_{211/2290}

- Sample A contains 3' extensions/vector backbone in igm482 cell line, while B contains 51-5 cell line.
pTΔCμ_{720/2290} vs. pTΔCμ_{211/2290} Rad51 homology ratio comparison

Data Entry

<table>
<thead>
<tr>
<th>Sample A</th>
<th>Sample B</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.46E+00</td>
<td>3.72E+00</td>
</tr>
<tr>
<td>3.25E+00</td>
<td>1.73E+00</td>
</tr>
<tr>
<td>2.31E+00</td>
<td>3.45E+00</td>
</tr>
</tbody>
</table>

Data Summary

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>3</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>Σx</td>
<td>9.120000</td>
<td>8.9</td>
<td>18.02000</td>
</tr>
<tr>
<td>Σx²</td>
<td>28.5302</td>
<td>28.73380</td>
<td>57.264</td>
</tr>
<tr>
<td>SS</td>
<td>0.8054</td>
<td>2.3305</td>
<td>3.1439</td>
</tr>
<tr>
<td>mean</td>
<td>3.04</td>
<td>2.9667</td>
<td>3.0033</td>
</tr>
</tbody>
</table>

Results

<table>
<thead>
<tr>
<th>Mean_a - Mean_b</th>
<th>t</th>
<th>df</th>
<th>P</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0733</td>
<td>±0.1</td>
<td>4</td>
<td>one-tailed 0.462578</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>two-tailed 0.923156</td>
<td></td>
</tr>
</tbody>
</table>

-Sample A contains pTΔCμ_{720/2290}, sample B contains pTΔCμ_{211/2290}
Gap size left & right extension comparisons

Sample A contains pTΔCμ858/2290 3’ extension/vector backbone containing 1.2 kb gap. Sample B contains pTΔBamCμ 3’ extension/vector backbone with 623 bp gap.
Gap ratio left vs. right extensions

Data Entry

<table>
<thead>
<tr>
<th>Sample A</th>
<th>Sample B</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.14E+01</td>
<td>1.26E+01</td>
</tr>
<tr>
<td>1.08E+01</td>
<td>1.76E+01</td>
</tr>
<tr>
<td>1.69E+01</td>
<td></td>
</tr>
<tr>
<td>2.00E+01</td>
<td></td>
</tr>
</tbody>
</table>

When all sample values have been entered, click the button labeled 'Calculate.'

Data Summary

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>4</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>ΣX</td>
<td>94.5</td>
<td>20.7</td>
<td>115.2</td>
</tr>
<tr>
<td>ΣX²</td>
<td>2400.49</td>
<td>476.17</td>
<td>2876.66</td>
</tr>
<tr>
<td>SS</td>
<td>167.9275</td>
<td>11.045</td>
<td>178.973</td>
</tr>
<tr>
<td>mean</td>
<td>23.625</td>
<td>5.525</td>
<td>20.033</td>
</tr>
</tbody>
</table>

Results

<table>
<thead>
<tr>
<th>Mean₁−Mean₀</th>
<th>t</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.275</td>
<td>−1.45</td>
<td>4</td>
<td>one-tailed 0.1103315</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>two-tailed 0.220663</td>
</tr>
</tbody>
</table>

-Ratio taken for small gap vs. large gap, left arm 3’ extension/vector backbone as sample A, right arm 3’ extension/vector backbone as sample B.
## Appendix II

### Solutions

#### 10X Electrode Buffer

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>dH₂O</td>
<td>1L</td>
</tr>
<tr>
<td>Glycine</td>
<td>144.0g</td>
</tr>
<tr>
<td>SDS</td>
<td>10.0g</td>
</tr>
<tr>
<td>Tris Base</td>
<td>30.3g</td>
</tr>
</tbody>
</table>

#### 1X PBS

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>dH₂O</td>
<td>4L</td>
</tr>
<tr>
<td>KCl</td>
<td>0.8g</td>
</tr>
<tr>
<td>KH₂PO₄</td>
<td>0.8g</td>
</tr>
<tr>
<td>Na₂HPO₄</td>
<td>4.6g</td>
</tr>
<tr>
<td>NaCl</td>
<td>32.0g</td>
</tr>
</tbody>
</table>

#### Permeabilization Buffer

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATP</td>
<td>2mL</td>
</tr>
<tr>
<td>0.1M CaCl₂</td>
<td>0.4mL</td>
</tr>
<tr>
<td>dH₂O</td>
<td>178.4mL</td>
</tr>
<tr>
<td>0.2M EGTA</td>
<td>1mL</td>
</tr>
<tr>
<td>1M glucose</td>
<td>2mL</td>
</tr>
<tr>
<td>HEPES (1M)</td>
<td>2mL</td>
</tr>
<tr>
<td>2M KCl</td>
<td>14mL</td>
</tr>
<tr>
<td>1M MgCl₂</td>
<td>0.2mL</td>
</tr>
</tbody>
</table>