Double Strand Break Repair of Highly Damaged DNA

FINAL REPORT

October 30 2009

Department of Justice, National Institute of Justice
Award Number: 2006-DN-BX-K005
(1 June 2006 – 31 August 2009)

Principal Investigator:
Jack Ballantyne, PhD
Associate Professor
Department of Chemistry
Associate Director for Research
National Center for Forensic Science
4000 Central Florida Blvd, Bldg #5
University of Central Florida
Orlando, FL 32816-2366
Phone: (407) 823 4440
Fax: (407) 823 2252
e-mail: jballant@mail.ucf.edu

EXECUTIVE SUMMARY

- Purpose of the present work was to attempt to repair 'in the test tube' highly fragmented genomic DNA in order to retrieve a DNA profile from otherwise intractable environmentally compromised samples.

- Proposal represented an ambitious high risk-high benefit undertaking of moderate cost (total $174,000).

- Two general strategies used in our attempt at double strand break repair (dsbr)

  -'molecular biology' approach: capture repair substrates that comprised oligonucleotides that subtended a SNP locus of interest and whose function was to capture the appropriate genomic fragments from the degraded sample

  -'biochemical'approach: recapitulate the cell's biochemical machinery for double strand break repair in the test tube, specifically non homologous end joining (NHEJ).

- Successful repair of oligonucleotide substrates that mimicked some SNP loci was obtained using capture oligonucleotides. However all subsequent studies involving fragmented genomic DNA failed to show any signs of repair.

- NHEJ also yielded successful repair of oligonucleotide substrates that mimicked some SNP loci. However all subsequent studies involving fragmented genomic DNA failed to show any signs of successful NHEJ repair.

    -An alternative but related biochemical pathway, MMEJ (Microhomology-Mediated End-Joining Pathway), was reconstituted but also failed to repair genomic DNA.

- Instead of direct repair we also developed and tested a modified oligonucleotide ligation assay (OLA) that was designed to recover profiles from fragmented DNA without actually repairing the DNA. This approach also failed.

- Using current technology, sequence complexity of genomic DNA provides an insurmountable computational barrier to reconstituting two contiguous fragments of DNA back to its native state, a requirement for the successful repair of double strand breaks.

ABSTRACT

DNA extracted from biological stains is often intractable to analysis. This may be due to a number of factors including a low copy number of starting molecules, the presence of soluble inhibitors or damaged DNA templates. Remedies may be available to the forensic scientist to deal with LCN templates and soluble inhibitors but none presently exist for damaged DNA. Previous work in this laboratory has shown that double strand and single strand breaks are significant contributors to the non-typeability of damaged DNA templates extracted from forensic-type stains (i.e. dried biological stains exposed to a myriad of environmental insults). This proposal sought to repair double strand breaks, restoring sufficient genomic integrity to permit DNA typing, using single nucleotide polymorphism loci (SNPs) as a model system. Three methods were developed and tested. The first is a simple gap filling prior to strand denaturation during the DNA amplification process. The second requires the addition of in vitro synthesized repair substrates that are complementary to the sequences flanking a SNP on both DNA strands, providing a matrix for repair polymerization and facilitating the recovery of amplifiable fragments. The third method involved the biochemical reconstitution of the non-homologous end joining (NHEJ) pathway, one of the principal cellular recombinational DNA repair pathways.

While the substrate mediated gap repair system demonstrated some evidence of *in vitro* repair using artificial genomic templates, all attempts at repair of genomic DNA using any of the repair methods failed. We believe that the sequence complexity of genomic DNA provides an insurmountable computational barrier to reconstituting two

contiguous fragments of DNA back to its native state, a requirement for the successful repair of double strand breaks.

TABLE OF CONTENTS

EXECUTIVE SUMMARY 2

ABSTRACT 4

TABLE OF CONTENTS 6

CHAPTER 1: INTRODUCTION 8

CHAPTER 2: RESULTS AND DISCUSSION 14

SNP Assay Development 14

Development of Methods for DNA Degradation 15

Aim I: Double strand break repair by gap filing 15

Aim 1A. Develop a simple gap repair system 15

Aim IB. Substrate-Mediated Gap Repair 16

Description of Basic Method 16

Repair of a Synthetic Oligonucleotide Lesion : rs740598 17

Repair of a Synthetic Oligonucleotide Lesion : rs1058083 18

Repair of RE Digested Genomic DNA: rs1058083 and rs560681 20

Repair of a Synthetic Oligonucleotide Lesion Redux : rs560681 21

Repair of RE Digested Genomic DNA Redux: rs1058083 23

Repair of a Synthetic Oligonucleotide Lesion Redux Redux : rs1058083 24

Aim II. Double strand break repair by homologous recombination (HR) 29

Aim IIA. Reconstitute the HR pathway for use with human genomic DNA 29

Aim III. Double strand break repair by non-homologous end-joining 29

Aim IIIA. Develop an in vitro repair system using non-homologous end-joining 29

NHEJ of Genomic DNA: rs1058083 30

NHEJ of Synthetic Oligonucleotides: rs16896068. 32
NHEJ of RE Digested Genomic DNA Redux: rs1058083 36
Microhomology-Mediated End-Joining Pathway (MMEJ) 37
Additional Double Strand Break Repair Approaches 39
Oligonucleotide Ligation Assay (OLA) for DNA Profile Recovery 39
CHAPTER 3: CONCLUSIONS 45
APPENDIX A: FIGURES 48
APPENDIX B: TABLES 75
APPENDIX C: PRESENTATIONS AND PUBLICATIONS 81
PRESENTATIONS 81
APPENDIX D: LIST OF REFERENCES 83
REVIEWER COMMENTS 85
Reviewer #1: 85
Reviewer #2: 88
RESPONSE TO REVIEWERS 93
Reviewer #1: 93
Reviewer #2: 96

CHAPTER 1: INTRODUCTION

The ability to detect DNA polymorphisms using molecular genetic techniques has revolutionized the forensic analysis of biological evidence [1]. DNA typing now plays a critical role within the criminal justice system [1]. Numerous individuals have been convicted and falsely accused individuals exonerated based on DNA evidence. Increasing use is being made of databases of DNA profiles for criminal intelligence information. One of the limiting factors with the technology is that sometimes DNA isolated from biological stains recovered from the crime scene is found to be intractable to standard STR analysis. This may be due to a number of factors, of which the most important are likely to be the presence of PCR inhibitors, a low copy number (LCN) of starting DNA molecules, or damaged (including degraded) DNA templates. Remedies may be available to the forensic scientist to deal with soluble inhibitors or LCN templates but none presently exists for damaged DNA. Potential remedies for damaged DNA are likely to be dependent upon the precise nature of the DNA damage present in any particular sample but, unfortunately, current knowledge of the biochemical nature, and the extent, of such DNA damage in dried biological stains is rudimentary.

DNA, like all macromolecules, spontaneously decomposes and therefore has a finite, but characteristic, thermodynamic stability. The primary structure can exhibit a variety of different lesions indicative of damage, including oxidation products, single and double strand breaks, UV-induced photoproducts, DNA or protein cross-links and chemical agent-induced covalent adducts [2]. In addition, a variety of hydrolysis products caused by spontaneous depurination, depyrimidination and deamination reactions are formed in DNA over time [2]. Genomic (i.e. DNA) instability may be

endogenous in nature and caused by water and/or reactive oxygen species (ROS) including hydrogen peroxide, superoxide or the hydroxyl free radical [2]. Also, through a variety of different mechanisms, exogenous environmental agents such as UV irradiation, heat, humidity and genotoxins may cause or facilitate damage to the structural integrity of the DNA molecule [2]. *In vivo*, the organism has an extensive armamentarium of enzymes that are responsible for the continuous recognition and repair of DNA damage that occurs spontaneously as a consequence of cellular metabolism. However, once the tissue is no longer under the control of the normal cellular homeostatic processes, such as is the case for biological stains deposited at a crime scene, DNA damage cannot be repaired. Although the lack of DNA repair ability in a stain is expected to increase the formation of certain types of lesions, some of the degradative processes, such as hydrolysis, are likely to be reduced in the dry state. Thus dried biological fluid stains should experience a different rate of DNA lesion formation compared to the situation *in situ*. It is likely that environmental insults are the primary lesion-causing factors in biological stains recovered from the crime scene. The principal concern from the forensic science standpoint is that many of these environmentally induced lesions are expected to be inhibitory towards DNA polymerase-mediated primer extension and may result in amplification, and hence DNA typing, failure.

Numerous studies have assessed the effects of various environmental factors on the ability to obtain a DNA profile. For example, McNally and Kobilinsky examined the effect of UV light, heat and humidity on laboratory prepared human bloodstains exposed for periods up to five days [3]. Samples subjected to UV irradiation showed a loss of allelic signal intensity with increasing exposure, but the rate of loss was not consistent.

The same authors observed a similar loss of typing ability with increasing exposure to elevated humidity and heat. In another study, McNally and De Forest used environmentally compromised stains obtained from casework samples [4]. In these studies, DNA extracted from *bona fide* forensic specimens also exhibited varying levels of damage that affected the allelic signal intensity observed with DNA profiling. These early reports examined the effects of environmentally induced damage to VNTR (or RFLP) analysis but, significantly, not to PCR-based DNA typing systems, which have supplanted VNTR technology for forensic casework use. Empirical data from the ancient DNA field has confirmed the expectation that less damage is caused to the DNA template under conditions of lower temperatures and humidity [5].

Novel DNA typing systems undergo developmental validation studies by the forensic science community prior to use and this often includes studies of the effects of environmental insults on the ability to type DNA accurately at all genetic loci of the DNA typing system employed [6,7]. The common conclusion reached is that environmentally impacted DNA in biological samples results in a progressive loss of signal and allelic drop out with extended or intense exposure [8-12].

We previously reported the development of a model for the assessment of DNA damage in biological stains, measuring the damage done by UVC light to naked DNA in solution, naked dehydrated DNA, and the DNA extracted from dried physiological stains, from which were able to make a few generalizations [13]. First, the most extensive damage was done to naked DNA in solution, followed by naked dehydrated DNA, with stain DNA showing the least damage. The DNA is protected by the cellular milieu in a stain, but even more significantly by the state of dehydration. We extended these

experiments to assess the damage inflicted by both UVA and UVB rays singly as well as by simulated natural sunlight (unpublished observations). Surprisingly, the DNA in dried bloodstains remained typeable subsequent to prolonged UVA and/or UVB exposure, prompting us to further explore potential sources of the DNA damage that can result in amplification failure, including heat, humidity and microorganism growth.

Our (unpublished) results indicated that microorganism growth is the most important cause of DNA damage leading to the non-typeability of forensic samples. To utilize the cellular constituents for sustenance, microbiota secrete digestive enzymes that can introduce DNA double strand breaks (dsb). It is likely that dsb lesions are the main cause of degradation of DNA in forensic cases and these will often result in amplification failure. At the present time, no successful *in vitro* system has been developed to reverse double strand breaks, and the quest for such a system is the subject of this proposal.

Cells employ two primary mechanisms, homologous recombination (HR) and non-homologous end-joining (NHEJ) to deal with double strand breaks [14-22]. During HR, an intact, homologous DNA duplex is used as a template for double strand break (dsbr) repair [14-17]. The damage is faithfully repaired with no loss of genetic material, but an extensive region of homology is required. HR is depicted in **Figure A**. At the site of a double strand break, a 5' to 3' exonuclease resects the ends of the break, generating single stranded 3' ends **(Figure A(i) and (ii)).** One of the exposed 3' ends then invades an intact homologous duplex, displacing the strand. The 3' end of the invading strand serves as a primer for repair synthesis, while the displaced strand acts as a template for the repair of the remaining broken segment **(Figure A(iii)).** The newly synthesized strands are ligated, and structures known as Holliday junctions are formed. Branch

migration follows, in which the area of heteroduplex DNA is extended **(Figure A(iv)).** Finally, the structure is resolved by enzymes known as resolvases. Subsequent to a 180^{o} rotation of the junctions, the enzymes cut the crossed structure, resulting in two intact DNA duplexes **(Figure A(v)).** Depending on the cut sites, either recombinant (markers surrounding the crossover point have undergone reciprocal recombination) or patched (markers surrounding the crossover site are derived from the same initial chromosome) duplexes are formed **(Figure A(v)).**This pathway has been reconstituted *in vitro* [23], effecting the repair of plasmid DNA, but it has not been extended for use with human genomic DNA. This may be possible, however, by adding carefully controlled quantities of the multi-protein repair complexes and allowing the HR reaction to proceed.

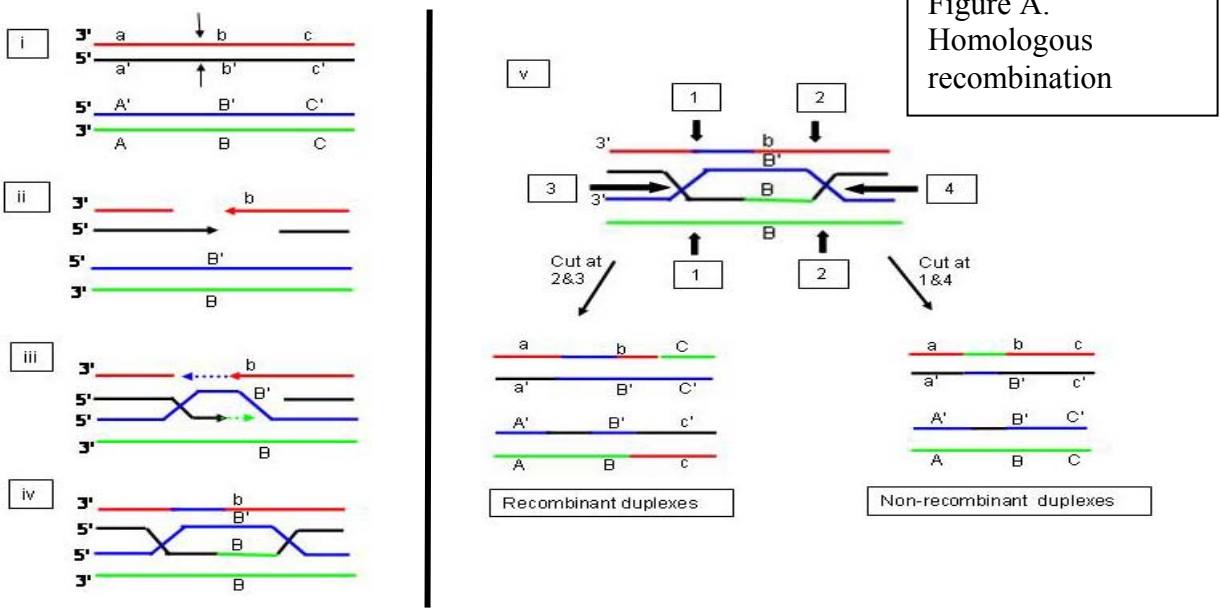


Figure A. Homologous recombination

The NHEJ pathway, also known as illegitimate recombination, is an alternative pathway that may be employed in the absence of regions of homology [24] **(Figure B).**

In human somatic cells NHEJ is believed to be the main double strand break repair mechanism [24]. The NHEJ pathway and its intermediates have been extensively studied. The RAD50/XRS/MRE11 complex has endonuclease activity and resects the ends of the break to reveal sites that may be ligated. The Ku70/Ku86 heterodimer binds DNA ends, aligning the ligatable ends, and forming the initial complex. DNA-PKcs, a serine threonine kinase that it activated by ds DNA ends, is recruited to the site to form the DNA-PK complex. The kinase also serves an important role in the processing on DNA ends. The ligation of two bases requires a 3'-OH and a 5'-phosphate. DNA-PKcs has been shown to catalyze the phosphorylation of 5'-OH groups as well the removal of 3' phosphates, and may possess other end-polishing activity. As the ends are aligned, an overhanging 'flap' may be formed, which is efficiently excised by a flap endonuclease, such as FEN-1. To complete the repair, DNA ligase IV/XRCC4 is recruited to the repair complex and ligates the broken ends held in alignment by the Ku heterodimer. Again, this pathway has been reconstituted for the *in vitro* repair of plasmid DNA as test repair substrates [25], but has not been tested on human genomic DNA. In this work we proposed to test different NHEJ formulations for their ability to restore genomic integrity, and DNA typeability, to damaged human DNA.

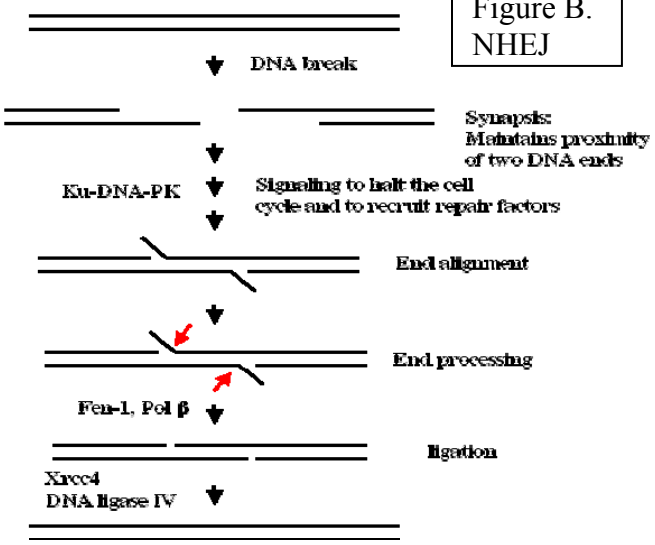


Figure B. NHEJ

CHAPTER 2: RESULTS AND DISCUSSION

SNP Assay Development

The ultimate goal of the current project was human identification from degraded DNA, indicated by the recovery of a genetic profile from otherwise non-typeable samples. DNA repair success was evaluated using selected SNPs chosen from a panel of 19 autosomal single nucleotide polymorphisms (SNPs) previously identified by K. Kidd of Yale University [26].

Five SNPs from the panel of 19 were chosen (rs279844, rs1058083, rs13182883, rs560681, and rs740598). PCR primers were designed to amplify an approximately 200bp region surrounding the SNP of interest, to permit efficient binding of the repair substrate. This was to insure that there would be enough space for the repair substrate to bind. Each primer was checked in Blast and Human BLAT to insure that it was specific to the human chromosome of the SNP of interest. Sequencing primers were also designed for subsequent Pyrosequencing.

PCR amplification of these SNPs was attempted with control K562 DNA at a concentration of 1 ng/µL. Primer concentration was 10pmol, with an annealing temperature of 56°C and 40 cycles. This produced the desired amplification products and no non-specific bands were seen in the product gel. Each SNP was amplified in 20 different individuals. The SNP genotyping data is provided in **Table 1** (see Appendix B). Thus despite the erythro-leukemic origin of K562 cell line, it appeared to be indistinguishable from a wild type phenotype for the particular SNP loci chosen for analysis.

Development of Methods for DNA Degradation

Control K562 DNA was subjected to DNase I treatment to mimic degraded DNA. DNase I treatment was carried out at 37°C for varying times followed by incubation at 75°C for 10 minutes to deactivate the enzyme. It was determined that 20ng/μL final DNA concentration and 0.1U/μL DNase I were the appropriate conditions for creating a DNA damage timescale on an agarose gel. This DNA was subjected to PCR amplification and pyrosequencing to determine at what point a genetic profile was no longer obtained.

Another control DNA (DNA T5595-1651 human female adult normal breast (76yrs old)) was used. The 1651 DNA was typed at all SNPs for which primers had been designed. Then 100ng of DNA was subjected to DNase I treatment to fragment it and then amplified at a chosen SNP of interest (rs1058083) to determine the point at which the SNP type is no longer discernible by pyrosequencing. It was determined that at a DNase I concentration of 0.5units/μL was sufficient to give a damage gradient on both agarose gels and in SNP analysis.

Aim I: Double strand break repair by gap filing

Aim 1A. Develop a simple gap repair system

We focused on the addition of T4 DNA ligase alone to repair simple single strand breaks in the DNA backbone of partially degraded environmental compromised DNA samples without the need for any gap filling. T4 DNA ligase was used as it is a very efficient ligase. For gap filling, T4 polymerase, DNA polymerase I, and Klenow fragment were attempted, taking into consideration the 5'-3' exonuclease activity or strand displacement activity of each enzyme (T4 pol. has no strand displacement activity

and only 3'-5' exonuclease activity, which should not be a factor since only the second assay contains a 3' overhang; DNA polymerase I shows strand displacement, but degrades the displaced strand; Klenow fragment has no exonuclease activity but has strong strand displacement activity). All attempts at this simple gap repair were unsuccessful on genomic DNA.

Given our lack of success using only T4 DNA ligase, the use of human DNA ligase IV/XRCC4 to enhance the repair success would have been worthwhile. Unfortunately, and unexpectedly, we were not able to commercially obtain the DNA LigaseIV/XRCC4 complex during the lifetime of this project and we were not able to complete any experiments that required its use.

Aim IB. Substrate-Mediated Gap Repair

Description of Basic Method

The substrate mediated gap repair method is illustrated in **Figure 1**. Two oligo 'repair substrates' (RS) complimentary to the SNP sequence between the PCR primers, but not overlapping them, were designed **(Figure 1).** One was complimentary to the sense strand of the PCR product the other was complimentary to the anti-sense strand. These are designed to be used as the repair substrates for substrate mediated gap repair of double stranded breaks deliberately induced in genomic DNA by DNase I. Oligos were designed with a mixed base at the SNP site because it cannot be determined what the SNP will be when using different DNA. A phosphate was placed on the 3' end of each oligo so that there will be no extension and thus only the gap will be repaired.

The control DNA was intentionally degraded using two different methodologies, DNase I and restriction endonuclease (RE) digestion in order to provide damaged DNA for subsequent repair. Initial attempts at repairing both DNase I and RE digested genomic DNA at the rs1058083 locus using the substrate mediated gap repair process described above were unsuccessful. Therefore a methodology utilizing synthetic oligonucleotides containing a hypothetical DNA break, and mimicking a SNP locus of interest, was developed to determine some of the critical parameters needed for successful repair.

Repair of a Synthetic Oligonucleotide Lesion : rs740598

Synthetic oligonucleotides mimicking locus rs740598 were used in order to attempt repair of a hypothetical DNA lesion. Both flanking oligonucleotides matched exactly the actual genomic DNA sequence and contained only one of the two alleles observed within the human population. The decision to repair templates containing only one of the SNPs was made because of the added simplicity to the system. Both flanking oligonucleotides were approximately equal lengths and extended eleven nucleotides beyond both the forward and reverse primer-binding sites. Numerous attempts were made at repairing these oligonucleotides using (i) various combinations of flanking oligonucleotide and substrate concentrations, (ii) different annealing buffers and (iii) altering the DNA ligase concentrations and times. None of these repair strategies were successful. Failure to successfully repair rs740598 was possibly due to adverse secondary structural motifs. In attempt to overcome this obstacle, snap cooling of all oligonucleotides was carried out prior to their mixing in attempt to keep them linear

while being "captured." This also failed to repair the lesion even after altering the ratio and concentrations of the flanking oligonucleotide and repair substrate.

Repair of a Synthetic Oligonucleotide Lesion : rs1058083

Given the persistent difficulties in repairing a strand break at locus rs740598, efforts were shifted to a different locus using the same methodology for repair. Locus rs1058083 was chosen for repair next given the relatively low amount of secondary structure present at the approximate annealing temperature of the repair substrate to the flanking oligonucleotides. Once again, the flanking oligonucleotides were approximately equal lengths with ten nucleotides extending beyond both the forward and reverse primer-binding sites. The hypothetical break in the DNA was placed in such a way that 36 nucleotides on the substrate would bind one flanking oligonucleotide (flanking oligonucleotide 1) while the other flanking oligonucleotide (flanking oligonucleotide 2) would be bound by 40 nucleotides on the repair substrate. These binding sites correspond to a melting temperature of 70^{o}C and 74^{o}C for flanking oligonucleotides 1 and 2 respectively. At 70^{o}C, secondary structure within the repair substrate is present, however, given the placement of the lesion this did not interfere with either of the flanking oligonucleotides' ability to bind the repair substrate in such a way that would allow repair to occur. The position of the secondary structure did not affect one of the flanking oligonucleotides binding to the repair substrate at all and also allowed for 13 nucleotides upstream of the cut site on the other flanking oligonucleotide to bind and facilitate repair.

As with rs740598, only one of the two rs1058083 alleles observed within the human population was initially used for repair **(Figure 2).** Multiple different repair substrate and flanking oligonucleotide concentrations were tested **(Table 2).** As a result of this optimization process some useful insights into the in vitro DNA repair process were obtained. First, the concentrations of the flanking oligonucleotides must be low enough during the ligation reaction that they are not being randomly ligated together in the absence of the repair substrate; however at the same time they must be high enough that sufficient concentrations of DNA are being repaired when the repair substrate is present so that a signal may be detected when pyrosequencing. Secondly, during PCR, low enough volumes of template must be added to the reaction to ensure that amplification is not randomly occurring when either the repair substrate or ligase is not added, but must again be high enough that sufficient concentrations of DNA are being amplified in the reactions when the repair substrate and ligase are present such that a signal can be detected during pyrosequencing. For locus rs1058083, flanking oligonucleotide and repair substrate concentrations of 50 picograms (pg) were found to effect repair of the hypothetical lesion while excluding anomalous repair in all other controls (i.e. when either of the flanking oligonucleotides, substrate, or ligase was omitted) (**Figure 3**). The repair took place using synthetic oligonucleotides. Repair was error free in that the sequence of the repaired substrates was fund to be that expected from direct ligation of the two oligonucleotides. Increasing the repair substrate concentration to two or more times that of the flanking oligonucleotides did not increase the signal intensity of the pyrogram. Similarly, removing the PCR buffer from the annealing/capture step in the repair did not affect the outcome. Next, repairing the same

locus that possesses the alternative SNP allele, as well as the heterozygote, was necessary to verify that this locus was fully repairable. Using the same oligonucleotide and repair substrate concentrations as was used for the original repair, repair of the opposite SNP as well as both SNPs mixed (heterozygote) was successful. Given this proof of concept, as well as the knowledge gained as to the intricacies of the system in terms of repair substrate to DNA concentrations, repair of in vitro fragmented genomic DNA was attempted

Repair of RE Digested Genomic DNA: rs1058083 and rs560681

Restriction enzyme (RE) DdeI was chosen to fragment genomic DNA, and locus rs1058083, since the latter provides a single cut site with a 34 and 42 nucleotide overlap between the fragmented DNA and the repair substrate on either side. This restriction enzyme was known to create sticky ends; however, it was presumed that given the large quantity of other fragments in the reaction with the same ends, the correct fragments would not be capable of rejoining themselves without the repair substrate. After multiple attempts at digesting DNA with RE DdeI and repairing it, it appeared as though enough of the fragments were in fact rejoining spontaneously and providing templates for amplification. This method was then redesigned and a suitable, blunt end generating RE (HaeIII) was used to damage another locus (rs560681) possessing the appropriate recognition sequence. This RE generates fragments that overlap the repair substrate by 62 and 37 nucleotides on the 5' and 3' sides of the cut site respectively. Numerous attempts at repairing DNA digested with this RE were made using various ratios of the fragmented DNA to repair substrate concentrations. Damaged DNA concentrations

ranged from 1ng to 200ng, per 25µL reaction, while repair substrate concentrations ranged from 0.1ng to 1ug per the same reaction volume. Ligase concentrations were also increased from the 5U/25µL reaction that was successful in repairing the synthetic oligonucleotides to 10U/25µL reaction, however, repair continued to fail. It was hypothesized that the secondary structure within either or both of the repair substrate and DNA fragments was the primary reason for the continuous failures.

The entire amplicon for locus rs560681, as well as the repair substrate by itself were analyzed using a DNA secondary structure calculator at the theoretical melting temperatures for the DNA fragments generated by HaeIII (76^{o}C and 68^{o}C for the 5' and 3' sides of the cut site respectively). As can be seen in **Figure 4**, the fragment annealing to the 3' side of the cut site is not affected by the secondary structure, however, the fragment annealing to the 5' side of the cut site is presumably prevented from binding adequately for ligation to occur. It is assumed that the fragment binding to the 5' side of the secondary structure will not be affected; however, there are only six nucleotides on the 3' side of the secondary structure, before the cut site, for the fragment to anneal and be ligated. This might be a contributing cause for repair failure at this locus since all fragments will preferentially bind to the substrate on the 5' side of the secondary structure at a higher temperature than they will on the 3' side. Even if DNA was to bind to the 3' side of the secondary structure, six nucleotides may not be sufficient to hold the fragment in place during ligation.

Repair of a Synthetic Oligonucleotide Lesion Redux : rs560681

Due to the difficulties in repairing restriction enzyme digested genomic DNA at the rs560681 locus, simpler repair substrates comprising two synthetic oligonucleotides that comprise the locus and mimic a double strand break in the rs560681 amplimer were employed. To facilitate repair a variety of different additives were incorporated into the in vitro assay such as: dimethyl sulfoxide (DMSO), single stranded DNA binding protein (SSB), urea, formamide, glycerol, betaine, 10X PCR enhancer (containing betaine) (Epicenter Technologies), 10X annealing buffer (containing Tris-HCl, NaCl, and ethylenediamine tetraacetic acid), and sodium chloride (NaCl). These additives were used alone or in combination. Additionally, three different "capturing" conditions were used in order to anneal the capturing repair substrate to the DNA fragments that were to be repaired. The first was snap-cooling of the repair reactions on ice after an initial incubation at 95^{o}C. The second consisted of incubating the repair reactions at 95^{o}C and then allowing cooling at room temperature for one hour, and the third annealing condition consisted of initially incubating the repair reactions at 95^{o}C followed by stepping down the annealing temperature from 80^{o}C to 40^{o}C at a rate of 1^{o}C per minute. For the additives, optimal results were obtained when DMSO and NaCl were used simultaneously and for the annealing conditions, best results were obtained when the step-down annealing conditions were used. DMSO reduces the secondary structures within the locus at lower temperatures while the NaCl reduces the phosphate-backbone repulsion between the DNA being repaired and the repair substrate, thus allowing the fragments to be "captured" at higher temperatures. Although repair of this locus with the synthetic oligonucleotides was successful when both DMSO and NaCl were present, there remains the requirement for high concentrations of both repair substrate and the

synthetic oligonucleotides that was not observed when repairing rs1058083. When repairing rs1058083, no additives were necessary and repair substrate and synthetic oligonucleotide concentrations were only 10% of those required for repairing rs560681. This difference is possibly due to the lower levels of secondary structure present at locus rs1058083 at the temperatures used during the repair process. One other possible factor making locus rs1058083 so easily repaired is the location of the secondary structure in relation to the lesion **(Figure 5).** Although the secondary structure is persistent beyond 80^{o}C, there is still significant base pairing available between the repair substrate and DNA fragments to allow repair.

Repair of RE Digested Genomic DNA Redux: rs1058083

Given the relative ease in repairing synthetic oligonucleotides at locus rs1058083 compared to locus rs560681, we re- focused our efforts on repairing restriction enzyme digested genomic DNA at rs1058083 as a simplified model for double stranded breaks. Although the secondary structure present within this locus is persistent beyond 80^{o}C, it might not greatly affect the repair of synthetic oligonucleotides because it is distal to the lesion. There are 18 nucleotides between the lesion and the secondary structure, which presumably allows enough base pairing to occur to facilitate repair. Initially, genomic control DNA was digested using the restriction enzyme, DdeI, to generate a double stranded lesion just upstream of the SNP site (**Figure 5**). Repair was first attempted using 1ng of damaged control DNA in a 25µL repair reaction with varying repair substrate concentrations (from 1ng to 1µg). Following the repair reaction, the entire repair product was then added to the PCR. Given the lack of success at repairing the

double strand break in this way (even when using various additives), increasing concentrations of damaged DNA were then added to the repair reaction in attempt to alter the repair substrate to DNA concentration ratio. DNA concentrations ranged from 5-50ng per 25μL repair reaction with no more than 10ng being added to the subsequent PCR. Increasing the concentration of DNA added to the PCR required the control DNA to then be double digested with the restriction enzyme, prior to repair, in order to ensure that the target DNA was sufficiently digested so not to give false positive results when the repair product (containing up to 10ng of DNA) was amplified. Again, these attempts proved unsuccessful with repair substrate concentrations ranging from 1ng to 1μg per 25μL repair reaction.

Repair of a Synthetic Oligonucleotide Lesion Redux Redux : rs1058083

Due to the repeated failures in repairing genomic DNA at this locus, we returned again to trying to 'repair' synthetic oligonucleotides instead of genomic DNA. This was done in order to better understand the importance of the ratio between the DNA being captured (synthetic oligonucleotides in this case), and the repair substrate concentrations within the repair reaction. Previous observations using this methodology suggested that the concentration of repair substrate to synthetic oligonucleotides must be in a 1:1 ratio in order for successful repair to occur. This knowledge was then adapted to account for the fact that between 1 and 10ng of damaged genomic DNA can realistically be added to the PCR without false positives occurring. The size of the repair substrate at this locus is 76bp, which corresponds to approximately $2.53X10^{-8}$ng within 1ng of total DNA. Therefore, to capture and repair this locus in 1ng of total damaged DNA, it would be

necessary to add approximately $2.53X10^{-8}$ng of our repair substrate to the repair reaction in order to achieve the desired 1:1 ratio of damaged DNA to repair substrate. In attempt to better understand the detection limits of this repair methodology, log dilutions of both the repair substrate and synthetic oligonucleotides were prepared and repair was attempted using concentrations of each ranging from 0.1ng to $1.0X10^{-9}$ng. Additionally, repair reaction volumes were reduced to 5 and 10μL in addition to the original reaction volume of 25μL. Reducing the reaction volumes aimed to increase the concentration of synthetic oligonucleotide and repair substrate in the repair reaction while keeping the concentrations unchanged in the PCR (since the entire repair product is added to the PCR). **Table 3** gives the details of the repair substrate and synthetic oligonucleotide concentrations that have been attempted to repair locus rs1058083 along with the reaction volume used and what, if any, additives were added. Even with no additives, 0.1ng of both the repair substrate and synthetic oligonucleotide could be repaired, however, DMSO and NaCl was required to repair synthetic oligonucleotides at a concentration of 0.01ng in a 25μL reaction **(Figure 6).** Reduced reaction volumes did increase the sensitivity of the repair reaction by facilitating the repair of this locus at lower concentrations without the use of additives, however, DMSO and NaCl greatly increased the signal to noise ratio during pyrosequencing and increased the levels of reproducibility for repairing 0.01ng of each of the repair substrate and synthetic oligonucleotides both.

As can also be seen in **Table 3**, we have attempted to determine whether or not a 1:1 ratio of repair substrate to synthetic oligonucleotide is absolutely essential for repair to be successful. In theory, repair substrate concentrations lower than synthetic oligonucleotide concentrations should still facilitate repair until a specific detectable limit

is reached. Repair substrate concentrations greater than synthetic oligonucleotide concentrations should not facilitate repair much beyond the 1:1 ratio. This is expected due to the fact that with a repair substrate to synthetic oligonucleotide ratio of greater than 1:1, random binding of the synthetic oligonucleotides to the repair substrates will occur without one particular repair substrate capturing both of the fragments that are to be ligated together. To determine if this 1:1 ratio is absolutely necessary for repair, we have attempted to repair synthetic oligonucleotides by capturing them with repair substrate concentrations of one order of magnitude greater and lesser. As expected, synthetic oligonucleotide concentrations of 0.001ng were captured and repaired by a repair substrate concentration of 0.0001ng in a 10μL reaction in the presence of DMSO and NaCl **(Figure 7).** Interestingly, this same repair was unsuccessful when carried out in the 5μL reaction.

To increase the level of repair at locus rs1058083 using synthetic oligonucleotides, various volume excluders were included in the repair reactions, since the smallest reaction volume achievable is ~6μL. By adding volume excluders, the hope was that the synthetic oligonucleotides and repair substrate would increase their chances of annealing through the reduction of "space" in the reaction in which repair can occur. Volume excluders that were attempted were: hexamine cobalt chloride (HCC), ficoll, polyethylene glycol 6000 (PEG-6000), and polyethylene glycol 8000 (PEG-8000).

Titrations of both synthetic oligonucleotides and repair substrate were carried out in repair reactions consisting of 5% DMSO, 100mM NaCl, 1X T4 DNA ligase buffer (Promega Corporation), and volume excluder. Titration reactions were also performed in the absence of volume excluders to determine whether the additive was contributing to

the repair reaction. When added, hexamine cobalt chloride concentrations ranged from 0.1μM-4mM and ficoll, PEG-6000 and PEG-8000 concentrations ranged from 0-20%. While HCC and ficoll showed no increase in the level of repair achievable, PEG-6000 and PEG-8000 increased the level of repair by two orders of magnitude. Without PEG-6000 and PEG-8000 present, a minimum of 400 attomoles of synthetic oligonucleotides could be repaired, however, with 12% and 5% of PEG-6000 and PEG-8000 added, respectively, a minimum of 40 attomoles of synthetic oligonucleotide was repairable.

To determine which synthetic oligonucleotide concentration could be reproducibly repaired and with what repair substrate concentration, titrations of both were carried out. Synthetic oligonucleotide and repair substrate concentrations ranged from 1 zeptomole to 10 picomoles and were carried out in the smallest reaction volume possible (~6μL). Repair of each synthetic oligonucleotide concentration was attempted with a repair substrate concentration ranging from up to 1000 times greater than and less than that of the synthetic oligonucleotide concentration. For these trials, 5% PEG-8000 was added as a volume excluder since this additive had been shown to contribute significantly to the level of repair obtainable. As shown in **Table 4**, to reproducibly repair (~100% of the time) synthetic oligonucleotides, mimicking a locus of interest, a minimum of 200 attomoles of each synthetic oligonucleotide must be present in the repair reaction with equal concentrations of repair substrate **(Figure 8).** Reducing repair substrate concentrations to 50 attomoles while maintaining 200 attomoles of each synthetic oligonucleotide resulted in only a minor decrease in reproducibility (to ~93%) **(Figure 9)** while, surprisingly, repair substrate concentrations of 100 attomoles reduced reproducibility to only ~72%. Repair substrate concentrations less than 50 attomoles

resulted in repair being significantly reduced (to ~36%) **(Figure 10),** possibly due to the required level of repair for detection with PCR based assays. Similarly, synthetic oligonucleotide concentrations of 100 attomoles or less resulted in reduced reproducibility (to <30%), regardless of the repair substrate concentration present in the reaction **(Figure 11).**

Based on the above described results, a 1:1 ratio of degraded DNA to repair substrate is required for reproducible repair of genomic DNA. Taking this into consideration, restriction enzyme digested DNA, at concentrations ranging from 1ng to 70ng, were repaired with repair substrate concentrations ranging from 40 yoctomoles to 400 zeptomoles. Using these DNA and repair substrate concentrations, a DNA to repair substrate ratio ranging from ~1000:1 to ~1:1000, respectively, is obtained. After the repair step, reactions were diluted accordingly so that no more than 4ng of total genomic DNA was being added to the PCR. This was done because DNA concentrations above 4ng are more prone to giving false positive results due to the level of degradation initially achieved. Additionally, repair reactions, as well as the untreated, digested DNA sample were amplified in replicates of eight to rule out false positives. This was done by determining whether a significantly greater number of replicates from the repair reactions were amplifiable compared to those of the unrepaired, digested DNA. These repair attempts were unsuccessful, presumably due to the fact that although a 1:1 ratio of DNA to repair substrate was achieved, there is also a minimum level of repair that is required for detection with our assays. This fact has been observed previously when repairing synthetic oligonucleotides. A minimum of 200 attomoles of synthetic oligonucleotide must be repaired with a minimum of 50 attomoles of repair substrate to be reproducibly

repaired and detected **(Table 4).** Concentrations lower than that in either the synthetic oligonucleotides or repair substrate results in repair not being reproducibly detected. Concentrations of 200 attomoles of synthetic oligonucleotides is equivalent to approximately 120.5 million copies of the locus to be repaired, which in turn would require approximately 360µg of genomic DNA. Due to this, there is a requirement for increased sensitivity in our pyrosequencing assays. To add to this sensitivity, a nested PCR was developed at locus rs1058083 to be used when repairing genomic DNA. In attempt to repair restriction enzyme digested genomic DNA, 1-4ng of genomic DNA was repaired with a titration of 40 yoctomoles to 4 attomoles of repair substrate and amplified using the previously designed nested PCR assay. Again, no repair was detected.

In summary, extensive testing of capture repair substrates failed to repair fragmented genomic DNA.

Aim II. Double strand break repair by homologous recombination (HR)

Aim IIA. Reconstitute the HR pathway for use with human genomic DNA

The objective of this aim was to reconstitute the homologous recombination pathway, *in vitro*, using purified human repair proteins/enzymes to repair DNA double strand breaks. Due to the inability to obtain the necessary protein an enzyme reagents from commercial suppliers, we focused our attempts at biochemical reconstitution of the double strand break repair pathway using NHEJ (see below).

Aim III. Double strand break repair by non-homologous end-joining

Aim IIIA. Develop an in vitro repair system using non-homologous end-joining

NHEJ of Genomic DNA: rs1058083

The objective of this aim was to reconstitute the non-homologous end-joining (NHEJ) pathway, *in* vitro, for the repair of DNA double strand breaks in genomic DNA.. For reconstitution of the NHEJ pathway, human Ku 70/80 complex, human Ligase IV/XRCCIV complex, and various polymerases are added to damaged DNA in attempt to process, localize, and join the appropriate DNA ends. **Figure 12** shows the methodological approach taken during this project for the repair of DNA double strand breaks by reconstituting the NHEJ pathway, *in vitro*.

Initial research focused on the addition of human Ku 70/80 proteins, as well as ligase, to a simplified repair reaction which comprised Hind III digested lambda DNA fragments that were subsequently analyzed using gel electrophoresis. Repair was attempted using two separate reactions, one using Ku 70/80 with human LigaseIV/XRCCIV (the ligase used in *in vivo* NHEJ) and one with Ku 70/80 and T4 DNA ligase (a ligase not used in *in vivo* NHEJ). By observing an electrophoretic mobility shift, it was concluded that human Ku 70/80 was binding to the DNA fragments in the absence of ligase; however ligase alone was sufficient by itself to cause repair of the lambda DNA sample. Based on the fact that Hind III digested lambda DNA is readably repairable with ligase alone, even at very low DNA concentrations, genomic control DNA was used for repair instead. Control DNA was initially subjected to an array of damaging conditions, including sonication, DNase I digestion, and restriction enzyme digestion prior to NHEJ being attempted. The success of the repair was then determined using gel electrophoresis where a shift in the average molecular weight was

expected to be observed when compared to unrepaired, damaged DNA. Again, for these repair reactions, Ku 70/80 complex was added to the damaged DNA with either human LigaseIV/XRCCIV or T4 DNA ligase, and incubated under a variety of conditions to obtain optimal activity of the enzymes involved. Using this approach, no successful repair was detected. One reason for this may be because the level of repair being achieved is so low as to be undetectable when monitoring the average molecular weight by gel electrophoresis. Based on this reasoning repair was then attempted using NHEJ on DNase I digested DNA and analyzed via pyrosequencing.

For this approach, control DNA was initially degraded using DNase I and typed at locus rs1058083 to verify that no genotype was obtainable prior to repair. Repair was then attempted using only Ku 70/80 and either human LigaseIV/XRCCIV or T4 DNA ligase under a variety of conditions. After attempting repair with various different enzyme concentrations and incubation times and temperatures, no repair was observed. Hypothetically this may have been due to the fact that the Ku 70/80 complex has a high affinity for double stranded DNA with blunt (or nearly blunt) ends, but DNase I digestion may produce lengthy overhangs in DNA fragments due to its random pattern of digestion that may not be suitable binding substrates for the Ku 70/80 complex. This a DNA end-processing step was incorporated into the repair reaction. To produce blunt DNA ends, an incubation with either T4 DNA polymerase or human DNA polymerase beta was used prior to the addition of the Ku 70/80 complex and ligase. Multiple attempts at repair using a variety of different conditions (enzyme concentrations, incubation temperatures and times, and buffers) were carried out, however, successful repair was not observed.

Table 5 lists all combinations of enzymes that were attempted in the repair of DNA double strand breaks using the NHEJ pathway.

There are a number of possible reasons for the difficulties in repairing DNA double strand breaks using NHEJ. With multiple enzymes functioning in one reaction it is difficult to develop a reaction buffer that will facilitate optimal activity for all. Also, the concentrations of each enzyme must be considered along with the appropriate reaction temperature and incubation time that will be sufficient for all enzymes to function optimally. The level of detection of PCR based assays must also be taken into consideration. Repair may be occurring, however, the levels of repair may be so low as to not be detected during amplification.

NHEJ of Synthetic Oligonucleotides: rs16896068.

Given the lack of success in our prior attempts at repairing DNase I digested genomic DNA using NHEJ, we decided to simplify our repair system by utilizing synthetic oligonucleotides mimicking SNP locus rs16896068. **Figure 13** shows the methodological approach taken for the repair of synthetic oligonucleotide duplexes using the NHEJ pathway *in vitro*. For this method we acquired single stranded oligonucleotides, mimicking the rs16896068 locus and possessing a hypothetical DNA lesion. The locus was split in two with each half 34bp in length. We duplexed complementary oligonucleotides by heating to 95^{o}C for 5 minutes followed by cooling at –1^{o}C/2min. to 45^{o}C followed by a 4^{o}C hold until required. Annealing of complimentary oligonucleotides for each of the two halves of the locus was carried out in separate reactions. Once the two double stranded DNA fragments were created, they were mixed

in equal concentrations, with appropriate buffer, in attempt to ligate them together with T4 DNA ligase alone. Detection of repair products was done using a pyrosequencing assay for locus rs16896068. A titration of oligonucleotide concentrations was done in order to determine the ratios below which the blunt-end ligating activity of T4 DNA ligase was insufficient to recover the SNP genotype of the synthetic oligonucleotides. Once the lowest limit of repair was detected for T4 DNA ligase alone, other repair factors were added in attempt to increase the repair capabilities. The lowest concentration of oligonucleotide duplexes that can be repaired by T4 DNA ligase alone was found to be approximately 0.6 femtomoles of each in a 10μL reaction with a T4 DNA ligase concentration of 0.5U/μL. **Figure 14** shows the successful repair of duplex oligonucleotides at locus rs16896068 when only T4 DNA ligase is present. When T4 DNA ligase is omitted, no repair occurs.

Next, we attempted to add Ku proteins to the repair reaction to aid in localizing DNA ends for T4 DNA ligase to act upon. Using Ku concentrations of 0U, 0.01U, 0.1U, 1U, and 2U, in a 10μL reaction, no increase in repair was observed. It should be noted that the addition of Ku proteins did not reduce the level of repair. Since T4 DNA ligase does not naturally interact with Ku proteins *in vivo*, these repair reactions were repeated with the human Ligase IV/XRCC4 complex, and the DNA-dependent protein kinase catalytic subunit (DNA-PK_{cs}), which, along with Ku, are essential components of NHEJ repair. DNA-PK_{cs} functions by binding to Ku proteins previously bound to DNA ends to form the holoenzyme, DNA-dependent protein kinase (DNA-PK). This holoenzyme has been suggested to form a necessary scaffold required to tether adjacent ends of DNA together. No increase in repair was observed when Ku, Ligase IV/XRCC4 or DNA-PK_{cs}

were used while varying incubation times, temperatures or enzyme concentrations. Interestingly, even at four times the repairable concentration achievable with T4 DNA ligase alone, Ligase IV does not result in repair. Ligase IV/XRCC4 complex can, however, ligate HindIII fragments of lambda DNA with relative ease, indicating that the blunt end nature of our duplexed oligonucleotides or the length of our duplexed oligonucleotides may not be facilitating repair by this enzyme, and that it may require cohesive ends for ligation or the repair fragments to be longer.

In addition to repairing duplexed oligonucleotides in 1:1 ratios, repair was also attempted when ratios of #1 and #2 **(Figure 15)** oligonucleotides were varied from 1:10 to 10:1. From the data shown in **Figure 15**, it can be seen that concentrations of approximately 0.5 femtomoles are required to obtain repair when oligonucleotides are in 1:1 ratio, however, lower concentrations of one oligonucleotide can be repaired as long as the concentration of the other oligonucleotide increases substantially (for example 0.2 femtomoles of oligonucleotide #1 can be repaired when five times the concentration of oligonucleotide #2 is present). This is thought to be the result of the random nature by which the DNA fragments are aligned and ligated by T4 DNA ligase. When the oligonucleotides are in a 1:1 ratio (at concentrations less than 0.5 femtomoles), insufficient concentrations of the two oligonucleotides are randomly aligning in proper orientation to facilitate sufficient ligation for detection. When one of the oligonucleotides is present in excess however, the limiting oligonucleotide will be saturated by the other, resulting in higher concentrations of the properly aligned oligonucleotides for ligation. Also due to the random localization of DNA ends, the levels of repair are not always consistent and often vary from one trial to another.

Due to the inability of Ku and DNA-PK$_{cs}$ to increase the level of repair achieved with either T4 DNA ligase or Ligase IV/XRCC4, we next attempted to verify that the DNA binding activities of these proteins were working under the reaction conditions of our system. To do this we carried out electrophoretic mobility shift assays using a chemiluminescent detection of biotinylated oligonucleotides. The same oligonucleotides from locus rs16896068, used above, were used for these experiments; however, Forward Oligonucleotide 1 **(Figure 13**) was biotinylated to facilitate chemiluminescent detection. As can be seen in **Figure 16**, it appears as though our reaction conditions do facilitate DNA binding with Ku by itself (**Figure 16, Lanes 2 and 3**), as well as the DNA-PK holoenzyme (Ku and DNA-PK$_{cs}$) (**Figure 16, Lanes 4 and 5**). It also appears that ATP is necessary for the binding of DNA-PK$_{cs}$ to DNA but not Ku (compare Lanes 2, 4 and 6 to Lanes 3, 5 and 7 from **Figure 16**).

Given that all of the components of NHEJ, alone, appear to be interacting with one another under our reaction conditions, we next attempted to ligate multiple oligonucleotides together using all the components combined, and detect ligation products using Southern blotting. To do this, Oligonucleotides 1 and 2 **(Figure 13)** as well as Ligase IV/XRCC4, Ku and DNA-PK$_{cs}$ were added to a 10μL reaction, and incubated at 30^{o}C for 20 minutes to 16 hours. Multiple trials were carried out with varying enzyme/protein concentrations, incubation times and the order in which the proteins/enzymes were added (i.e. when additional Ku was incubated with the DNA-PK holoenzyme, a 15min pre-incubation of either the Ku or DNA-PK was carried out prior to the addition of the other component and ligase). To detect whether ligation was occurring using this methodology, denaturing polyacrylamide gel electrophoresis was

carried out followed by Southern blotting. These attempts, however, did not result in the successful repair/ligation of our synthetic oligonucleotide duplexes. One reason for this may be that the oligonucleotide duplexes are too short in length, and that there is some minimum fragment length required for successful ligation. The reasoning behind this is that previous reports have indicated, based on DNase I footprinting, that Ku and Ligase IV/XRCC4 each bind to approximately 25-35 nucleotides of DNA during NHEJ.

NHEJ of RE Digested Genomic DNA Redux: rs1058083

Attempts were made to repair using restriction enzyme digested PCR fragments (locus rs1058083) with NHEJ enzymes. The undigested fragment is 536bp and, when digested with RsaI, gives a 367bp and a 169bp fragment. Approximately 250ng of DNA template was added to the reaction with various NHEJ factors. After incubation at 30^{o}C for 4 hours samples were analyzed using alkaline agarose gel electrophoresis **(Figure 17).** If ligation were to occur, it would be expected that the two lower fragments disappear while larger molecular weight products are formed.

It is difficult to determine from the alkaline gel **(Figure 17)** whether or not ligation of the digested fragments is occurring or if the enzymes/proteins within the reaction are causing the digested fragments to be retarded when running through the gel. To try to gain a better understanding, these same reactions were carried out again and then the reactions were heat denatured at 75^{o}C for 30 minutes. This was done to cause the DNA binding proteins/enzymes to dissociate from the DNA prior to electrophoresis. The heat-treated reactions were then run on an agarose gel and stained with SYBR Gold. From **Figure 18**, it does not appear as though ligation is occurring since the lanes with

Ku alone, PK alone, and Lig4 alone result in the same banding pattern as Ku with Lig4 and PK with Lig4. The existence of the additional band at the bottom of the untreated RsaI digest lane is something of interest, however, the banding patterns of all lanes are unexpected (meaning that more than just the two bands for digested product and one band for undigested product are observed) and suspected as being the result of the heat denaturation prior to electrophoresis.

Microhomology-Mediated End-Joining Pathway (MMEJ)

An alternative but related approach to standard NHEJ was developed and tested, namely the microhomology-mediated end-joining pathway (MMEJ) **(Figure 19).** This pathway is similar to the NHEJ pathway, except that instead of blunting DNA ends and localizing them randomly, this pathway utilizes regions of homology at DNA ends to join the correct fragments of DNA. For this pathway, a variety of 3' to 5' exonucleases, polymerases and ligases have been used to initially chew back DNA ends at break regions, facilitating annealing, followed by gap filling and ligation.

Initial research using the MMEJ pathway to repair DNA double strand breaks was first done using sonicated control DNA with repair detection being done using gel electrophoresis. As with the NHEJ repair pathway, DNase I digested control DNA was subsequently used so that repair detection could be done using a pyrosequencing assay rather than gel electrophoresis. T4 DNA polymerase was initially used to chew-back DNA ends followed by an annealing step to join homologous DNA fragments. In addition to T4 DNA polymerase, exonuclease III was also used during the chew-back reaction as this enzyme may be more suitable for this task than exploiting the 3'-5'

exonuclease activity of T4 DNA polymerase by depriving it of dNTPs. Following DNA end chew-back and annealing, gaps in the DNA were filled with *Taq* DNA polymerase and the nicks sealed with *Taq* DNA ligase. This method is similar to that published by Gibson *et al.* [27], in which a "chew-back repair" method was developed to assemble a synthetic *Mycoplasma genitalium* genome. Difficulties using *Taq* enzymes in our system stem from the fact that because the starting fragments of DNA are extremely small, the chew-back reaction must be relatively brief; therefore, the end regions exhibiting homology will be relatively small as well. This poses a problem with fragments becoming un-annealed during the gap filling and ligation steps where the reaction is incubated at 45^{o}C.

Due to the need for lower incubation temperatures, a variety of different gap filling and ligation enzymes were attempted. T4 DNA polymerase and T4 DNA ligase were used in an attempt to fill in gaps and ligate nicks since T4 DNA polymerase retains activity at temperatures as low as 12^{o}C and T4 DNA ligase is active at temperatures as low as 4^{o}C. Using this combination of enzymes, a variety of incubation times and temperatures were attempted, however, no successful repair was observed. Once again, exonuclease III was also utilized in the chew-back reaction along with T4 DNA polymerase. Given that T4 DNA polymerase exhibits strong 3'-5' exonuclease activity, it was hypothesized that if repair were occurring, this enzyme would continue to degrade the repaired DNA, even at lower temperatures. Short periods of gap filling were also attempted after which T4 DNA polymerase was heat inactivated followed by another annealing step prior to the addition of T4 DNA ligase. Carrying out the reaction in this manner adds complexity however, due to the requirement for a second annealing step.

The next attempt at MMEJ repair used human DNA polymerase beta, a gap filling DNA polymerase that is active at relatively low temperatures, and does not exhibit exonuclease activity. T4 DNA polymerase was still maintained for the chew-back reaction and T4 DNA ligase continued to be used for ligation since it is active over a wide range of temperatures. Multiple reaction buffers were tried in order to try and ensure that all enzymes could function adequately, and a large variety of incubation times and temperatures were attempted with varying concentrations of each of the three enzymes. Repair was unsuccessful. One reason that repair was unsuccessful using this combination of enzymes could be because human DNA polymerase beta is used during NHEJ for filling only small gaps (~4nt) and is less active when filling larger gaps. This enzyme also exhibits strand displacement when filling larger gaps of DNA, which would prevent repair as one strand would overlap the other rather than create a nick between two adjacent nucleotides for ligase to act on.

Another combination of enzymes attempted involved T4 DNA polymerase (for the chew-back reaction), and T4 DNA ligase for the ligation reaction. Klenow fragment (exo^-) was used for the gap filling because it exhibits no exonuclease activity. It does however, exhibit medium strand displacement activity, so human flap endonuclease (FEN-1) was also used, which functions to cut displaced strands of DNA. This strategy also failed.

Additional Double Strand Break Repair Approaches

Oligonucleotide Ligation Assay (OLA) for DNA Profile Recovery

Due to the failures of the other approaches to double strand break repair described above, we attempted to develop a SNP assay for the recovery of a DNA profile for highly degraded DNA samples. We initially began to develop and optimize this methodology using locus rs1058083. This detection assay although similar to a conventional oligonucleotide ligation assay (OLA) possesses one significant modification. The novelty of this methodology is that the ligation of the common and allele specific oligonucleotides occurs on the DNA sample prior to amplification whereas in conventional OLAs, a pre-amplification is necessary. The difficulty with using OLA on highly degraded DNA samples is that pre-amplification is often not possible. An overview of this modified oligonucleotide ligation assay is provided in **Figure 20**. For this methodology, the common and allele specific oligonucleotides each contain unique primer-binding sequences for subsequent amplification after one or more rounds of ligation. The unique primer binding sequence within the common oligonucleotide is 5'-GGTAAACCCAGTGTCCTG-3' and the unique primer binding sequence within the allele specific oligonucleotide is 5'-CGAATAGTCGTCTAAGGC-3'.

Optimization of PCR parameters and reaction conditions was conducted. The primers are unique to this assay and were designed to not amplify human genomic DNA. Optimization was done using a synthetic oligonucleotide with the unique primer binding regions incorporated. To optimize this assay, annealing temperature gradients were run as well as $MgCl_2$ and DMSO concentration gradients. The optimal annealing temperature is 56^oC, the optimal $MgCl_2$ concentration is 1.9mM, and the addition of DMSO was not found to have any affect on the PCR. Additionally, amplifications were performed using varying concentrations of the common and allele specific oligonucleotides separately and

in combination without the addition of ligase and DNA to ensure that amplification does not occur without these two components. Oligonucleotide concentrations less than 0.5 pmoles per 25uL PCR did not result in non-specific amplification. Next, ligase was added to varying concentrations of common and allele specific oligonucleotids separately and in combination to determine if ligation occurs in the absence of template. It was observed that oligonucleotide concentrations of 50 fmoles or more in 10uL ligation reactions results in template independent ligation resulting in the generation of a SNP profile. Oligonucleotide concentrations of 10fmoles or less have not been observed as giving template independent ligation.

Further analyses were conducted to verify that the OLA would not generate false positive results in the absence of DNA template. To do this multiple replicates were carried out where varying concentrations (10-50 fmol per 10uL reaction) of common and allele specific oligonucleotides were mixed and incubated with ligase at 4^{o}C overnight. 4^{o}C overnight was chosen for these experiments in attempt to maximize the chances of template independent ligation as a "worst case" scenario. Frequently, samples containing 50 fmoles would generate a SNP profile after amplification, however, replicates containing 30 fmoles or less of each oligonucleotide (one common oligonucleotide and two allele specific oligonucleotides) did not generate a SNP profile after amplification. Based on these results, OLA was attempted using between 1 and 50 ng of template DNA **(Figure 21**) (of T/T genotype at this locus) and 30fmol of each oligonucleotide (common and both allele specific). Annealing conditions were: 95^{o}C – 3 min; 45 or 50^{o}C – 10 min; 4^{o}C hold. SNP profiles were detected for all DNA concentrations; however, many of the samples erroneously appeared heterozygous, presumably from mismatch ligation, when

the correct genotype was homozygous T. As the input DNA concentration is increased, the correct SNP profile is more easily recovered.

OLA was next attempted by varying the annealing temperatures (from 25°C to 55°C) to make the assay generate the correct genotype and to reduce mismatch ligation. Differing concentrations of ligase (1.5U to 0.025U per 10uL reaction) were also added in attempt to reduce mismatch ligation. Annealing temperatures of 45°C and 25°C appear to work better than at higher temperatures; however, there are occasional inconsistencies where replicates annealed at higher temperatures also result in adequate SNP profile generation. Ligase concentrations less than 1U per 10uL reaction gives only minimal ligation, however, ligation could be detected down to 0.025U. OLA reactions were also carried out with two rounds of denaturation, annealing and ligation. These resulted in better signal to noise ratios on the pyrograms, however, increased mismatch ligation was also observed **(Figure 22).** To prevent mismatch ligation, NaCl gradients were run during the ligation reactions. NaCl has previously been found to prevent mismatch ligation and is a common additive to these reactions. Ligation controls were run with increasing concentrations of NaCl to determine the level at which ligation is inhibited by it. NaCl concentrations up to 200mM did not inhibit the T4 ligase during the ligation of Lambda HindIII ladder. NaCl concentrations greater than 75mM inhibited PCR, however, indicating that purification prior to amplification will be required.

Different purification methods prior to amplification were attempted to determine which would be best. Ethanol precipitation was chosen as the method of choice. Samples were purified using precipitation with 100% ethanol followed by a wash with 85% ethanol; precipitation with 100% ethanol only without a wash; and precipitation

with 90% ethanol only without a wash step. Precipitations without a wash step were done to minimize sample loss, however, slightly better results were obtained when precipitation followed by an 85% ethanol wash were used. Spermidine was also added to attempt to increase the signal to noise ratio of the pyrograms generated. Spermidine has also been reported, along with NaCl, to reduce mismatch ligation during OLA. Concentrations of 1-5mM were attempted, however, in addition to increasing the signal to noise ratio of the data, it also increased the level of detection for the wrong SNP. Based on these data, spermidine was not considered a useful additive.

When OLA was attempted using the homozygote C/C DNA template, mismatch ligation was frequently observed. To reduce this, annealing temperature gradients as well as ligase gradients were carried out. The reason for limiting the ligase concentration in the reaction is so that less enzyme will be available to ligate the mismatched product. Once ligase binds to mismatched DNA ends, it adenylates the 5' end at the junction and then dissociates. Mismatch ligation is presumed to occur when multiple binding events at the adenylated mismatch takes place, eventually leading to ligation. Also, by attempting to control the annealing temperature better, it may be possible to prevent the majority of annealing with mismatches (although this is difficult because the mismatch is at the very end of the oligo). Reducing ligase concentrations and changing annealing conditions did very little to eliminate mismatch ligation from occurring, so changes in ligation temperatures were next attempted. Ligation temperatures of 22^{o}C, 25^{o}C, and 28^{o}C were attempted, however, this had no affect on the generation of a SNP profile. Ligation temperatures were changed in attempt to alter the fidelity of the enzyme to make it less apt to ligate mismatches. Because the mismatch ligation remained persistent, OLA was

attempted using oligonucleotides specific for the opposite strand. This was done because there may be more of a tendency to ligate A/C mismatches than T/G mismatches and by directing detection to the opposite strand, the levels of mismatch ligation maybe reduced. Even when typing from the opposite strand, mismatch ligation was still frequently observed.

By increasing the amount of input DNA to 50 or 100ng and reducing the oligonucleotide concentrations to 5fmol per 10uL reaction for each of the allele specific and common oligonucleotides, superior results were obtained, however, low levels of mismatch ligation products were still observed in the C/C homozygote sample **(Figure 23**). This assay was subsequently shown to be unsuccessful for the heterozygote sample.

CHAPTER 3: CONCLUSIONS

The purpose of the present work was to attempt to repair 'in the test tube' highly fragmented genomic DNA in order to retrieve a DNA profile from otherwise intractable environmentally compromised samples. Such fragmented DNA comprises double strand breaks and, if sufficiently common throughout the sample DNA, results in a number average molecular weight of DNA below the amplimer size required for successful PCR based genetic analysis. Due to the seemingly insurmountable technical barriers to success with complex human genomic DNA samples, this proposal represented an ambitious high risk-high benefit undertaking of moderate cost (total $174,000). There have been no reports in the scientific literature showing successful *in vitro* repair of *ex vivo* human genomic DNA, thus necessitating us to take several novel approaches to the problem.

We used two general strategies in our attempt at double strand break repair. The first ('molecular biology approach') employed capture repair substrates that comprised oligonucleotides that subtended a SNP locus of interest and whose function was to capture the appropriate genomic fragments from the degraded sample to permit simple covalent ligation and subsequent DNA profiling analysis. The second ('biochemical') approach was to recapitulate the cell's biochemical machinery for double strand break repair in the test tube, specifically NHEJ. We were unable to effect repair of fragmented genomic DNA using either approach despite some preliminary encouraging results with simplified test substrates.

Successful repair of oligonucleotide substrates that mimicked some SNP loci was obtained using capture oligonucleotides. This resulted in an extensive study of the factors and conditions necessary for in vitro repair using this approach. However all subsequent studies involving fragmented genomic DNA, created by DNAseI treatment or restriction endonucleases, failed to show any signs of repair.

The biochemical approach involving NHEJ also yielded successful repair of oligonucleotide substrates that mimicked some SNP loci was obtained. However, akin to the situation with the capture repair reagents, all subsequent studies involving fragmented genomic DNA failed to show any signs of successful NHEJ repair. An alternative but related biochemical pathway, MMEJ, was reconstituted but also failed to repair genomic DNA.

Instead of direct repair we also developed and tested a modified oligonucleotide ligation assay (OLA) that was designed to recover profiles from fragmented DNA without actually repairing the DNA. This approach also failed.

It is worthwhile considering the complexities surrounding the attempt to perform in vitro repair of fragmented DNA. If the DNA is fragmented to such a degree that standard DNA typing is not possible then one assumption is that the number average molecular weight of the fragmented DNA must be less than the STR amplimer sizes (roughly 100-350 bp). Thus assuming fragment sizes of 100 bp then there might be as many as $(3 \times 10^9)/10^2 = 3 \times 10^7$ fragments of DNA that can be re-ligated to one another. Assuming that there are two ways for each double stranded fragment to reanneal with another and that two fragments have to come together to reconstitute the original STR amplimer then there is only a one in sixty million chance that the fragment in question

will find its amplimer mate by chance. Hence simple random ligation of fragmented DNA would, with a high degree of probability, not result in reconstitution of the STR amplimer. This random ligation was the basis of the biochemical methods attempted here (NHEJ and MMEJ). This understanding of the combinatorial complexities was the reason behind us trying to facilitate the alignment of amplimer mates by use of a capture repair oligonucleotide substrates ('molecular biology approach'). The approach probably failed because we weren't able to efficiently selectively enhance and analyze any captured *bona fide* amplimers.

Future studies could attempt to increase the possibility of detecting and analyzing a reconstituted amplimer by whole gene amplificatiuon methods such as WGA or mIPEP [28-30].

APPENDIX A: FIGURES

Figure 1. Substrate Mediated Gap Repair

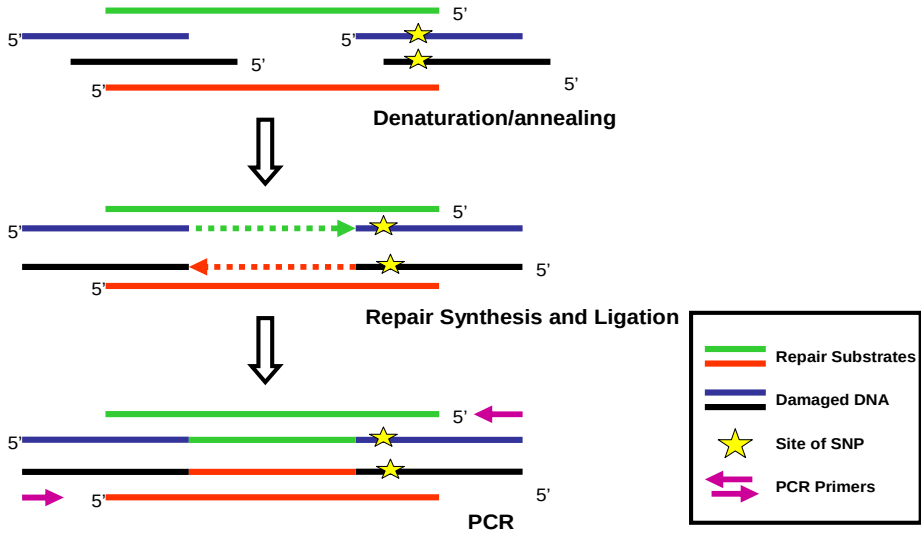


Figure 2. Repair system developed using synthetic oligonucleotides to mimic the locus of interest while excluding all other fragments found in genomic DNA extracts.

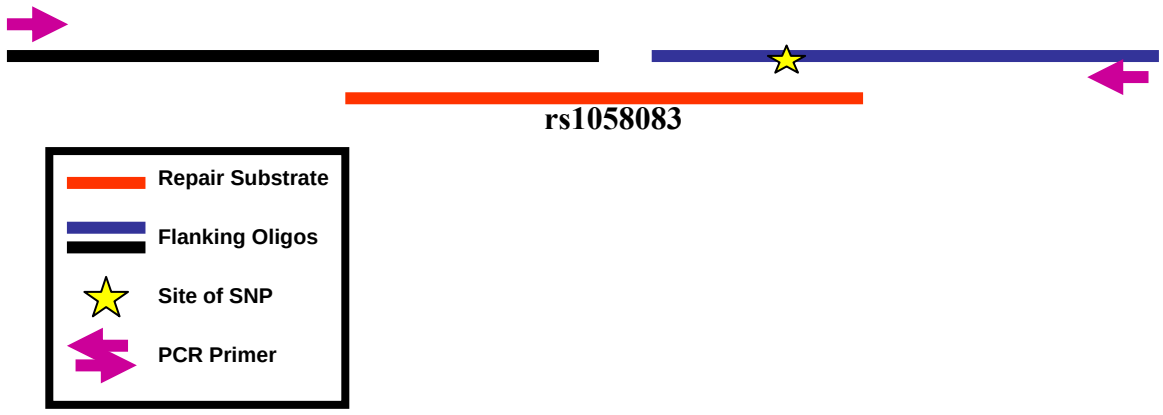


Figure 3. Successful repair of locus rs1058083. Left: Pyrogram generated after repair with 50pg of flanking oligonucleotides 1 and 2, 50pg of repair substrate, PCR Buffer II, and Ligase. Right: Pyrogram generated after repair with 50pg of flanking oligonucleotides 1 and 2, 50pg of repair substrate, and PCR Buffer II. Ligase was <u>not</u> added to this reaction. The x-axis of each program represents the nucleotides added. Incorporation of an individual nucleotide is represented by the presence of a peak (intensity of light signal). The height of each peak is proportional to the number of nucleotides incorporated.

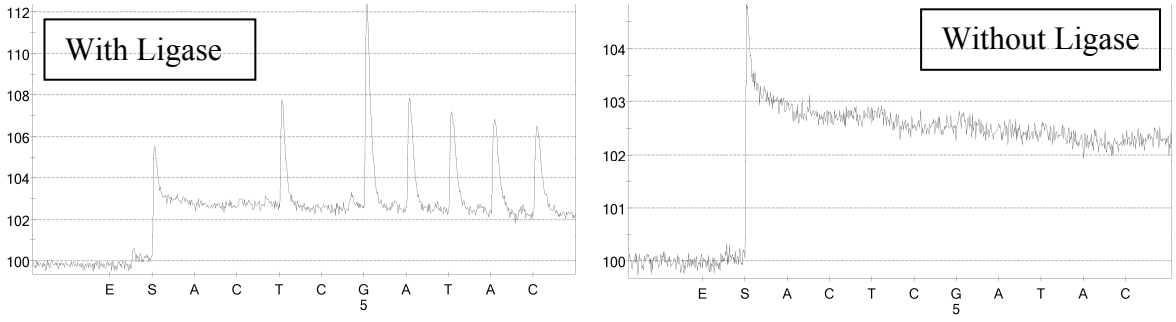


Figure 4: Secondary structure prediction at 65°C for locus rs560681 repair substrate. Complete secondary structure disappears at approximately 79°C. Scissors indicate the HaeIII cut site; A, is 62 nucleotides long and has a theoretical melting temperature of 76°C; B, is 37 nucleotides long and has a theoretical melting temperature of 68°C. 5' and 3' denote the polarity of the repair substrate.

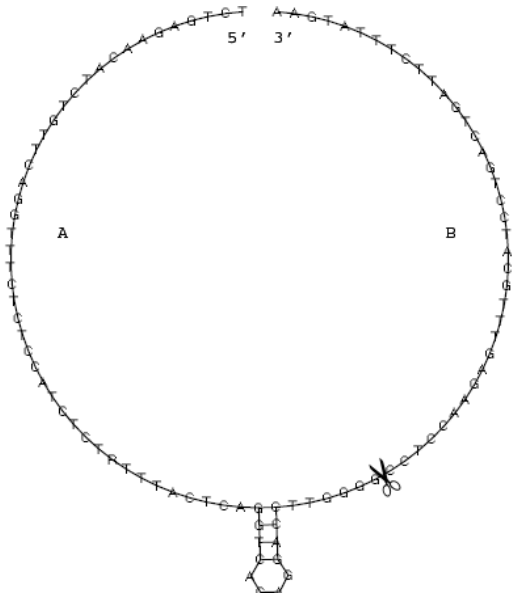


Figure 5. Predicted secondary structure at 60^{o}C for the rs1058083 repair substrate. Secondary structure persists beyond 80^{o}C. Scissors indicate the DdeI restriction enzyme cut site; A, is 31 nucleotides long and has a theoretical melting temperature of 67^{o}C; B, is 45 nucleotides long and has a theoretical melting temperature of 74^{o}C. 5' and 3' denote the polarity of the repair substrate; R, denotes the SNP.

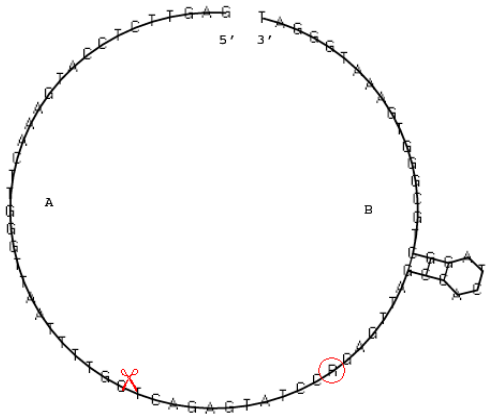


Figure 6. Successful repair of locus rs1058083. Top: Pyrogram generated after attempted repair with 0.01ng of each the repair substrate and synthetic oligonucleotides in a 25µL reaction. No additives utilized. Bottom: Pyrogram generated after repair with 0.01ng of each the repair substrate and synthetic oligonucleotides in a 25µL reaction. 5% DMSO and 50mM NaCl was added.

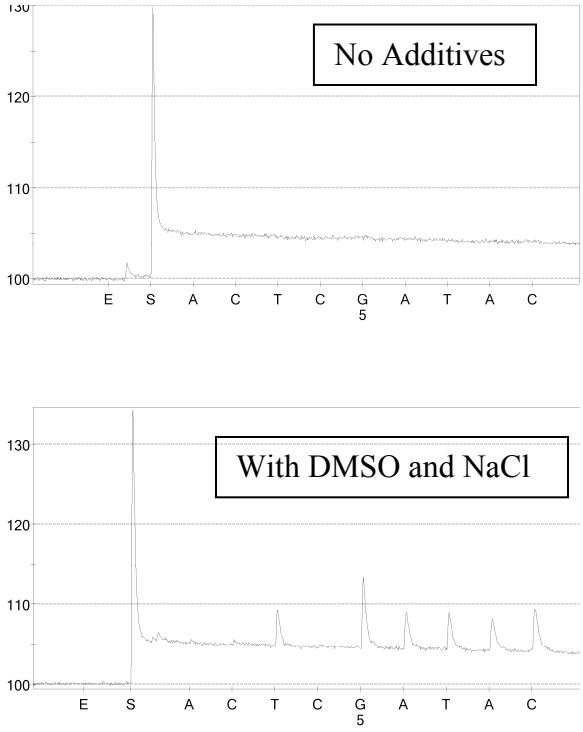


Figure 7. Successful repair of locus rs1058083. Left: Pyrogram generated after repair with 0.0001ng of the repair substrate and 0.001ng of the synthetic oligonucleotides in a 10µL reaction. 5% DMSO and 100mM NaCl were added. Right: Pyrogram generated after repair with 0.0001ng of the repair substrate and 0.001ng of the synthetic oligonucleotides in a 5µL reaction. 5% DMSO and 100mM NaCl were added.

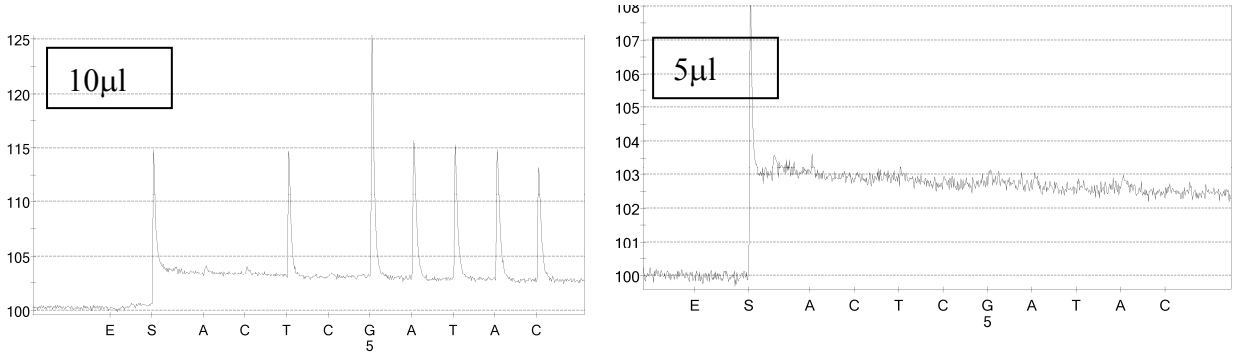
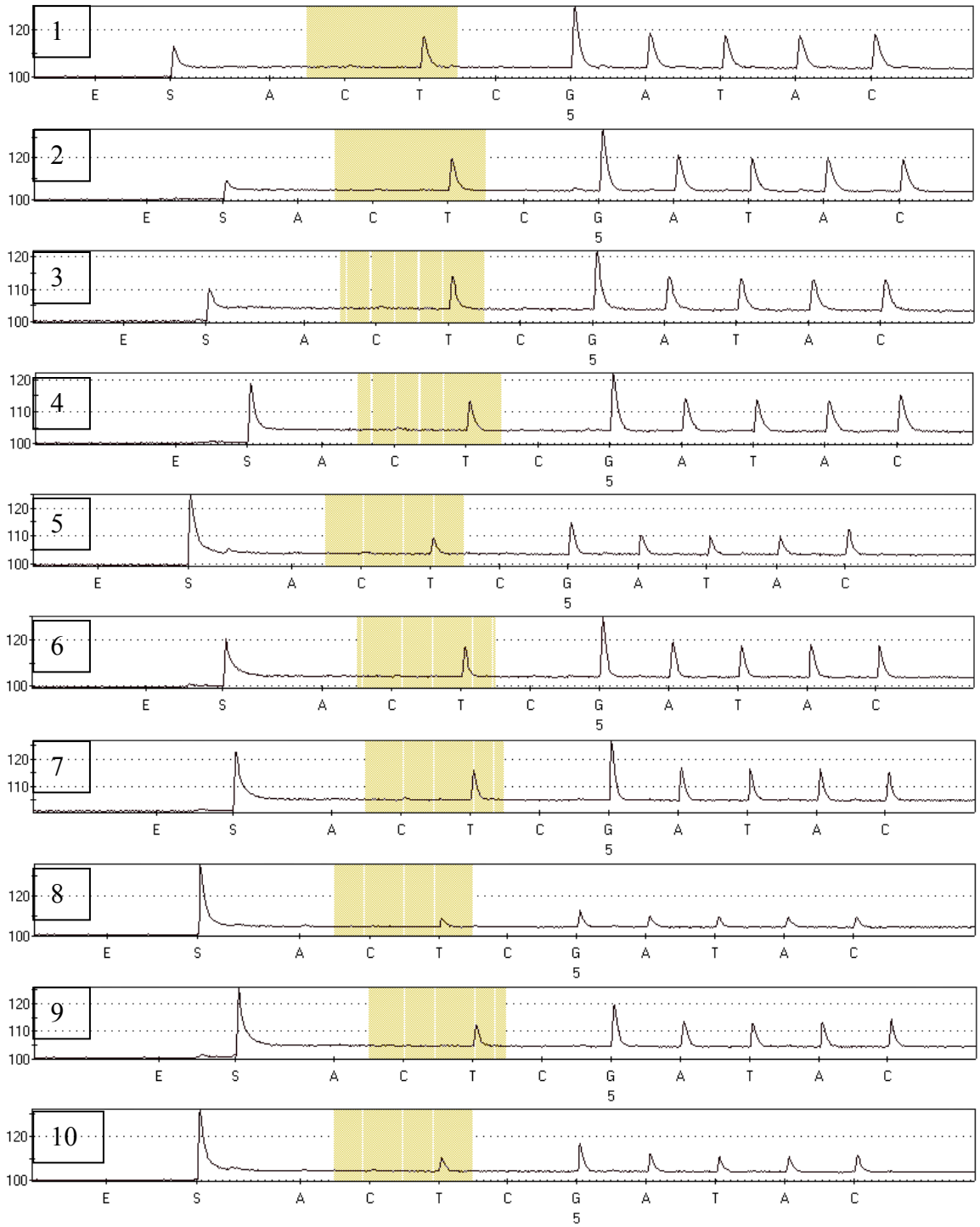


Figure 8: Pyrograms generated from the repair of 200 attomoles of synthetic oligonucleotides with 200 attomoles of repair substrate at locus rs1058083. All 14 panels show positive repair results as determined by the pyrosequencing software. Successful repair will generate a homozygous A genotype (reverse strand shown below).



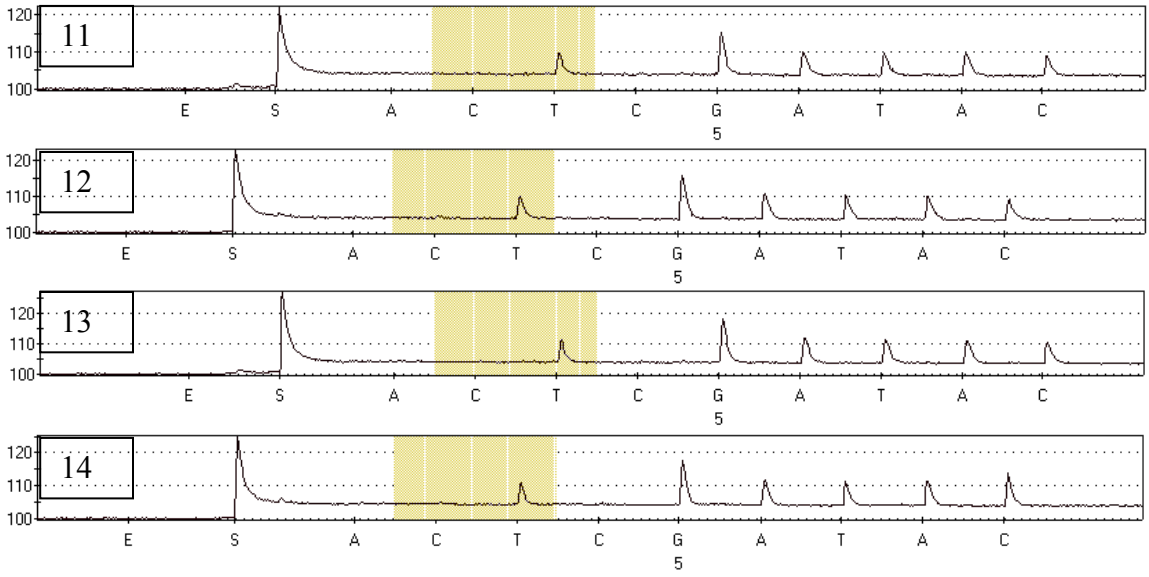
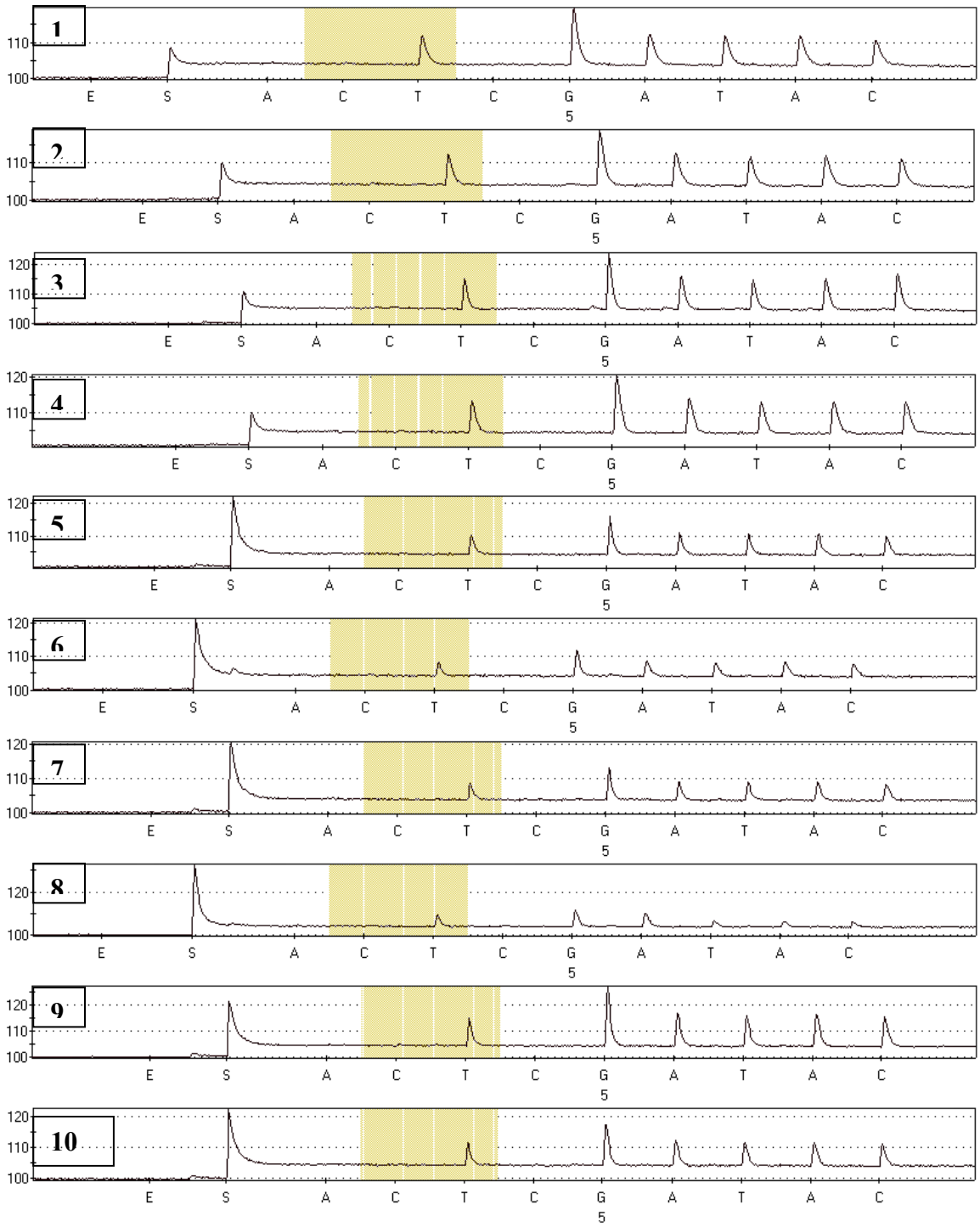

11
120
110
100
E S A C T C G A T A C
5
12
120
110
100
E S A C T C G A T A C
5
13
120
110
100
E S A C T C G A T A C
5
14
120
110
100
E S A C T C G A T A C
5

Figure 9. Pyrograms generated from the repair of 200 attomoles of synthetic oligonucleotides with 50 attomoles of repair substrate at locus rs1058083. Top 13 panels are positive repair results and the bottom panel is negative repair result. Panel 11 was deemed negative by the pyrosequencing software, however this repair was considered successful based on the recovery of a distinct profile. Successful repair will generate a homozygous A genotype (reverse strand shown below).



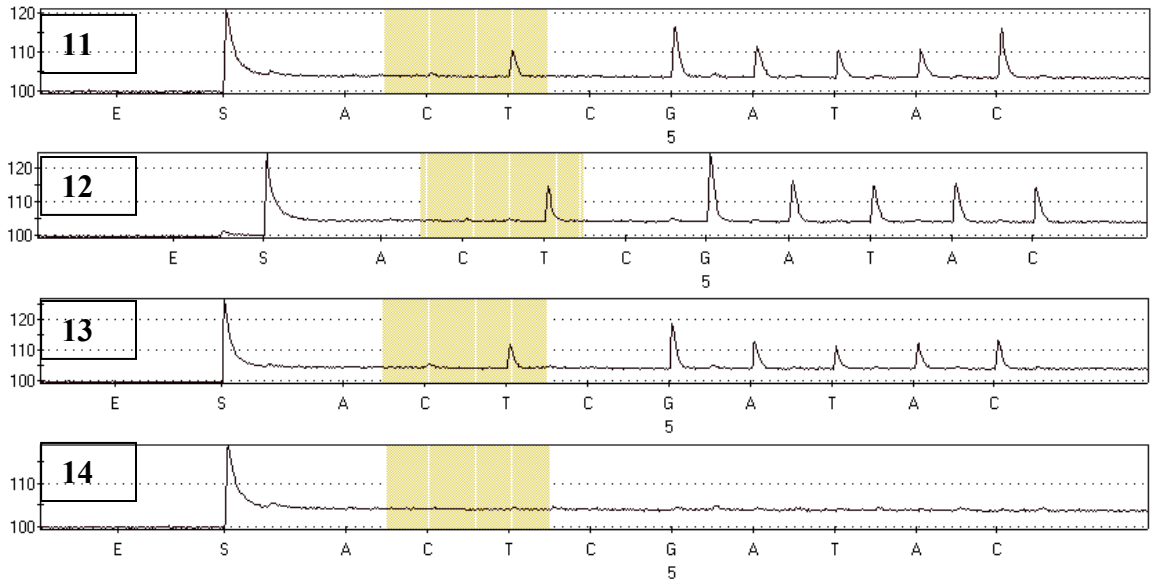
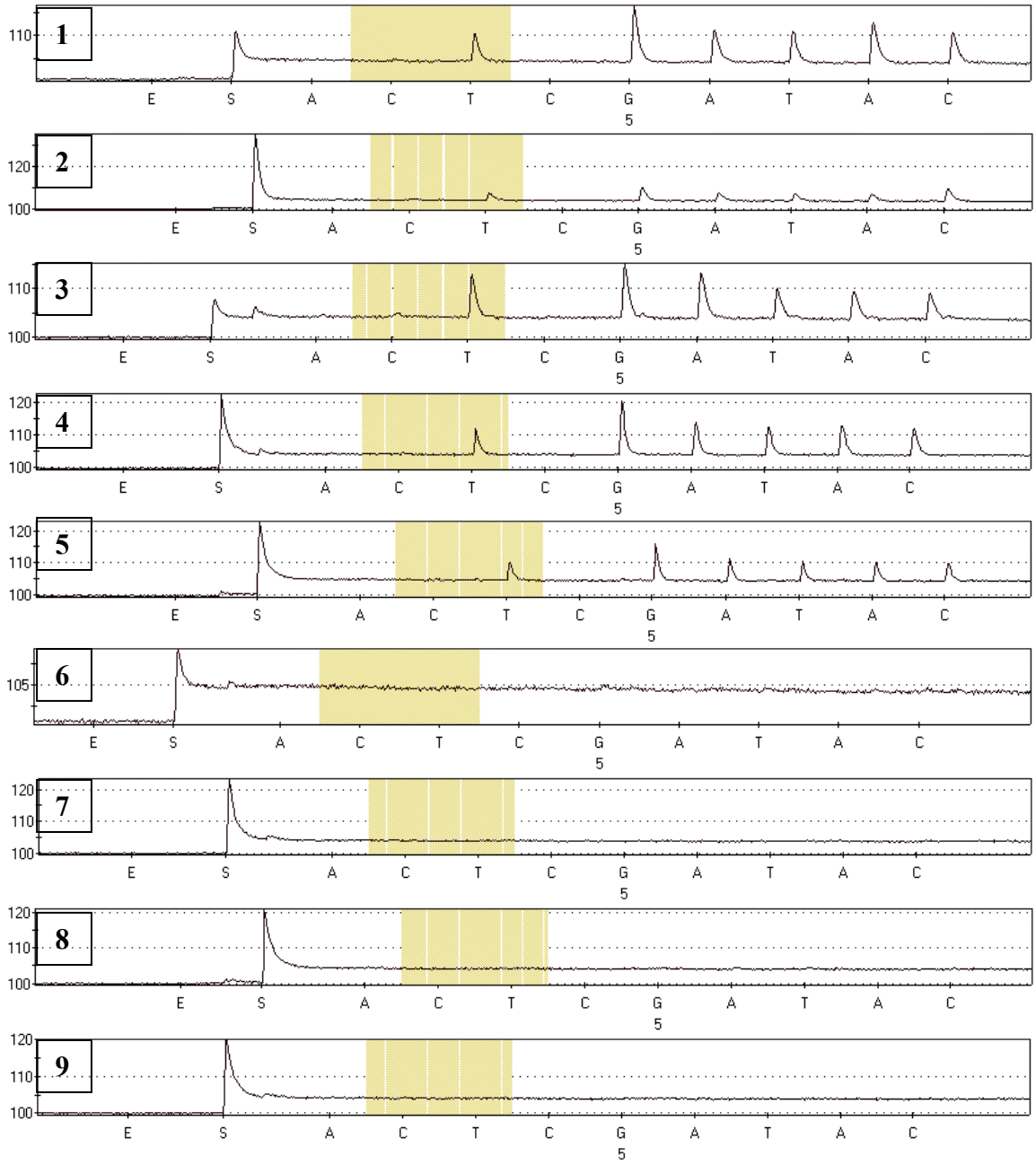

11
120
110
100
E S A C T C G A T A C
5
12
120
110
100
E S A C T C G A T A C
5
13
120
110
100
E S A C T C G A T A C
5
14
110
100
E S A C T C G A T A C
5

Figure 10. Pyrograms generated from the repair of 200 attomoles of synthetic oligonucleotides with 20 attomoles of repair substrate at locus rs1058083. Top five panels are positive repair results and the bottom nine panels are negative repair results. Panel two was deemed negative by the pyrosequencing software, however this repair was considered successful based on the recovery of a distinct profile. Successful repair will generate a homozygous A genotype (reverse strand shown below).



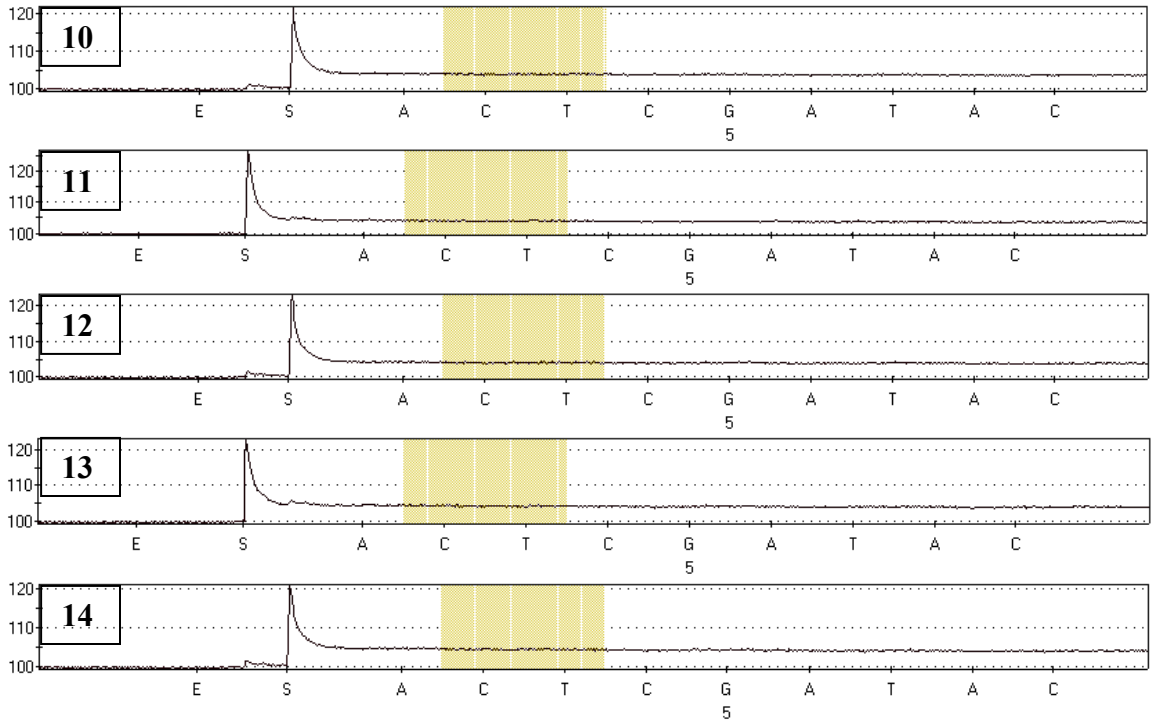
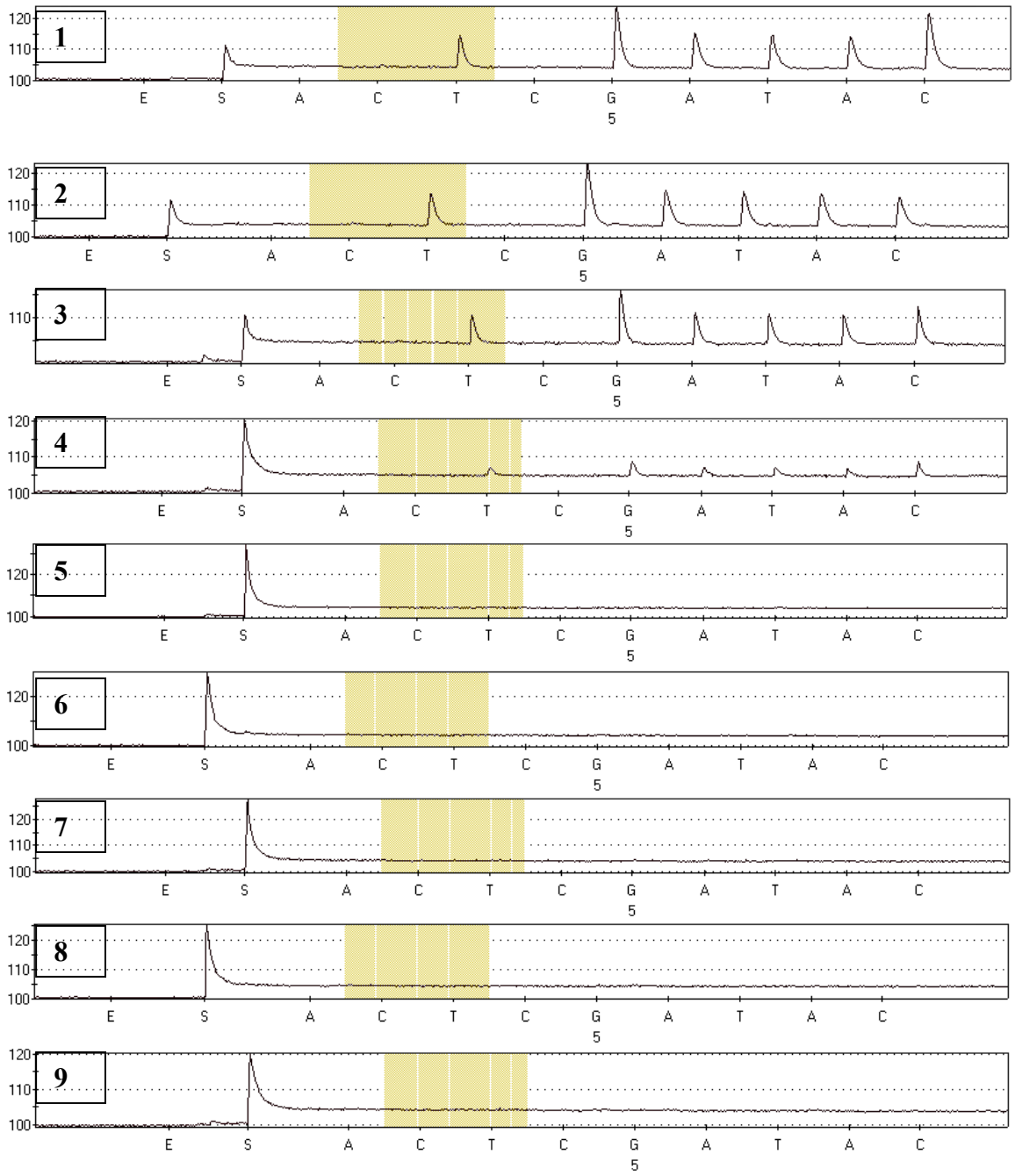

10
120
110
100
E S A C T C G A T A C
5
11
120
110
100
E S A C T C G A T A C
5
12
120
110
100
E S A C T C G A T A C
5
13
120
110
100
E S A C T C G A T A C
5
14
120
110
100
E S A C T C G A T A C
5

Figure 11: Pyrograms generated from the repair of 100 attomoles of synthetic oligonucleotides with 100 attomoles of repair substrate at locus rs1058083. Top four panels are positive repair results and the bottom 10 panels are negative repair results. Panel four was deemed negative by the pyrosequencing software, however this repair was considered successful based on the recovery of a distinct profile. Successful repair will generate a homozygous A genotype (reverse strand shown below).



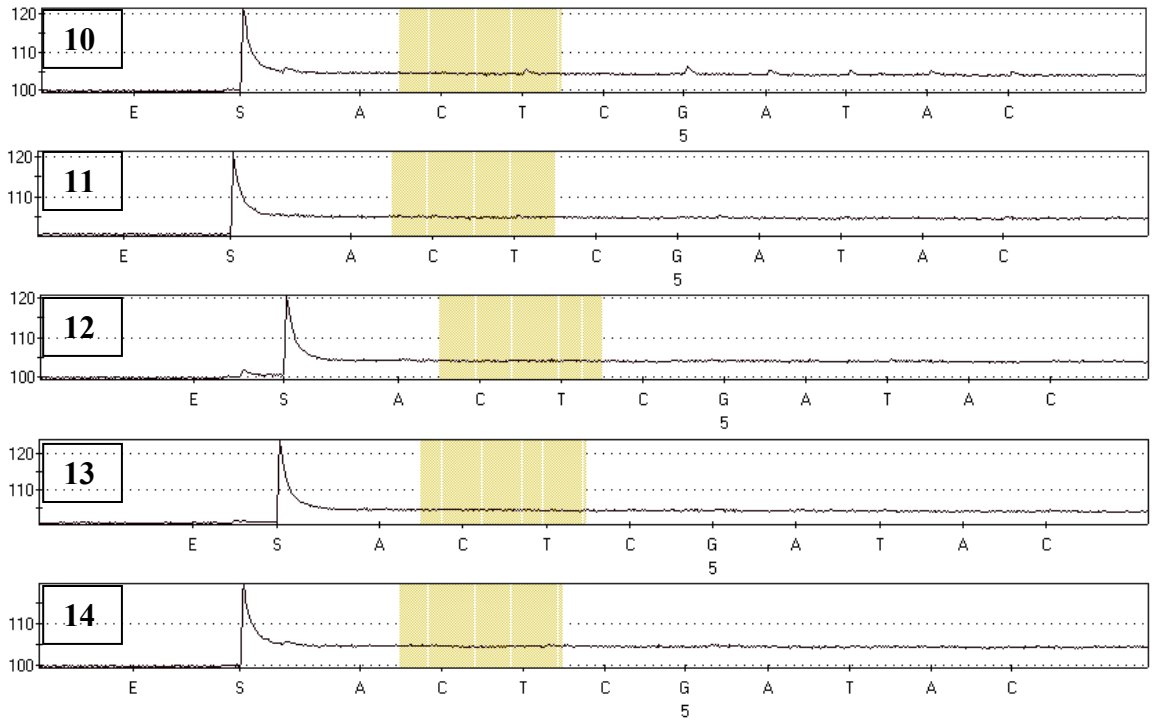

10
120
110
100
E S A C T C G A T A C
5
11
120
110
E S A C T C G A T A C
5
12
120
110
100
E S A C T C G A T A C
5
13
120
110
E S A C T C G A T A C
5
14
110
100
E S A C T C G A T A C
5

Figure 12. Diagram outlining the experimental procedure used for the repair of DNA double strand breaks via the NHEJ pathway.

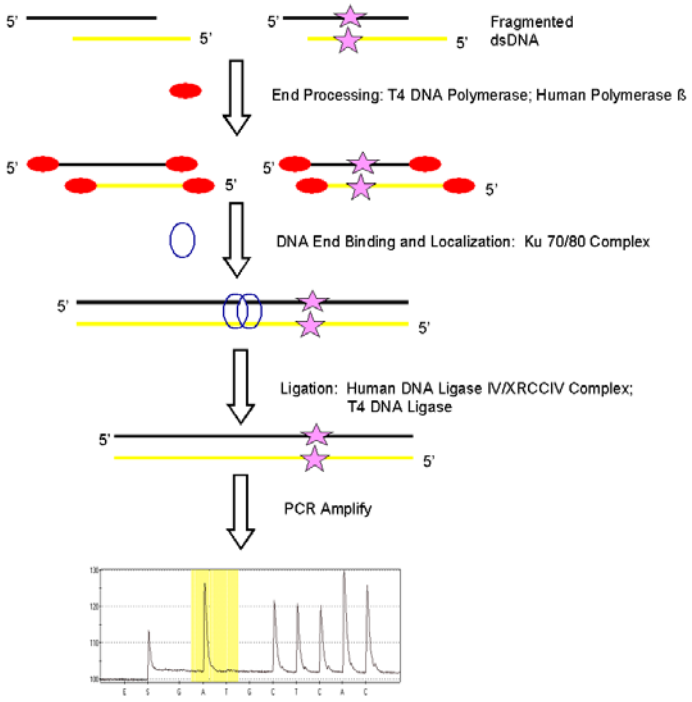


Figure 13. Diagram outlining the experimental procedure used for the repair of synthetic oligonucleotides at locus rs16896068.

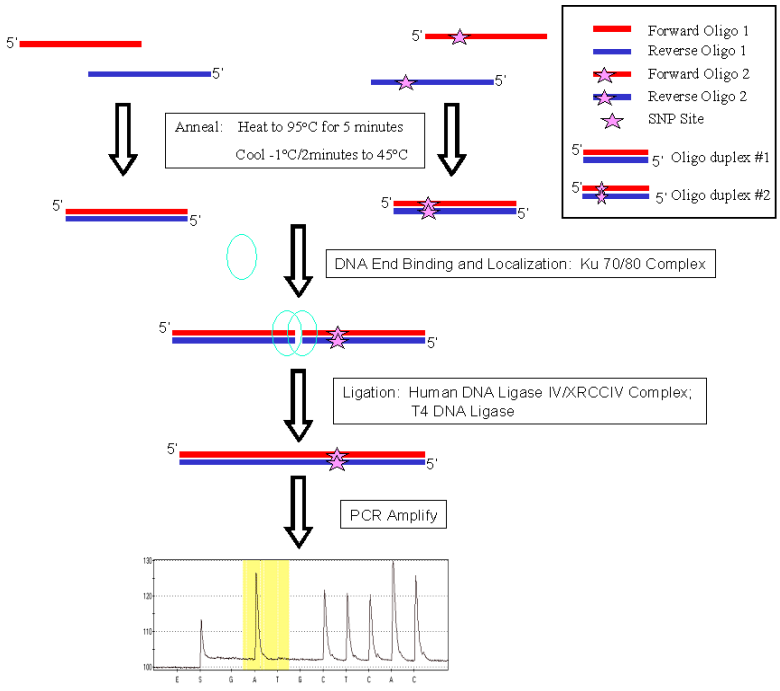


Figure 14. Repair of duplexed synthetic oligonucleotides at locus rs16896068. Top Panel: 0.6fmol. of each oligonucleotide was incubated with 5U T4 DNA Ligase in a 10μL reaction for 16hrs. Bottom Panel: 0.6fmol. of each oligonucleotide was incubated in ligation buffer without T4 DNA Ligase for 16hrs in a 10μL reaction.

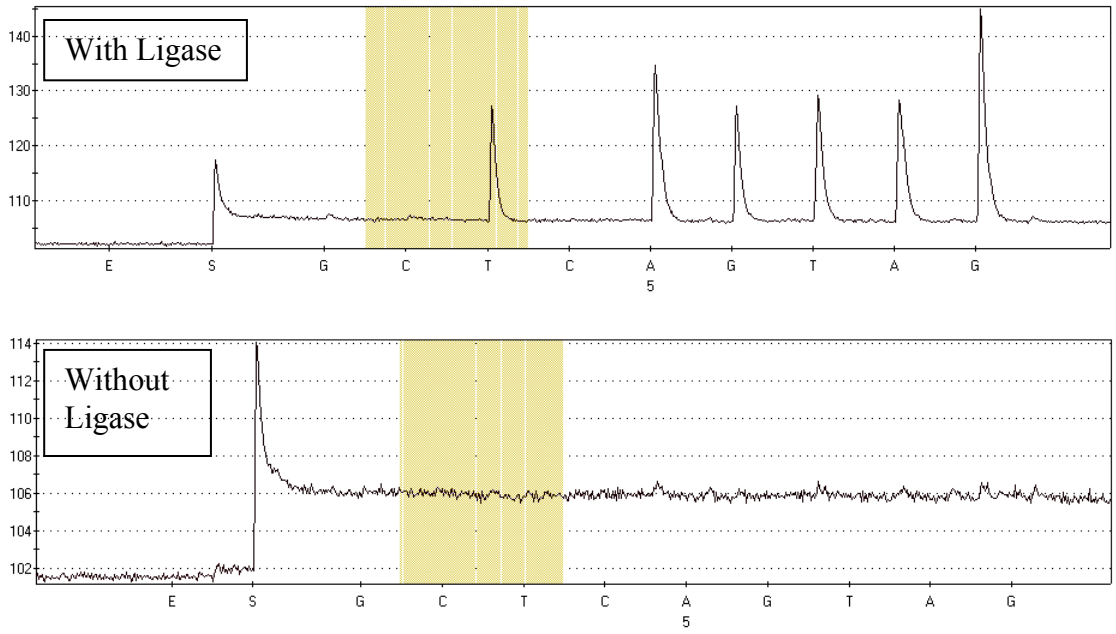


Figure 15. Graphical representation of repair success with varying oligonucleotide concentration ratios.

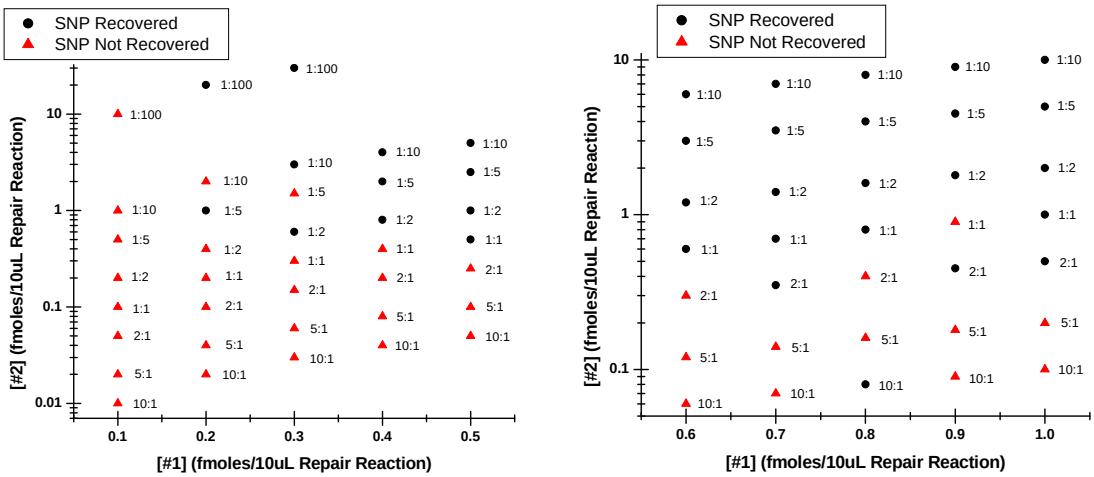


Figure 16. Electrophoretic mobility shift assay with biotinylated duplex 34mer in the presence of either Ku alone or Ku with DNA-PK$_{cs}$, with or without ATP. DNA-PK; holoenzyme containing DNA-Pk$_{cs}$ and Ku. Lanes 6 & 7 have additional Ku added to the DNA-PK holoenzyme.

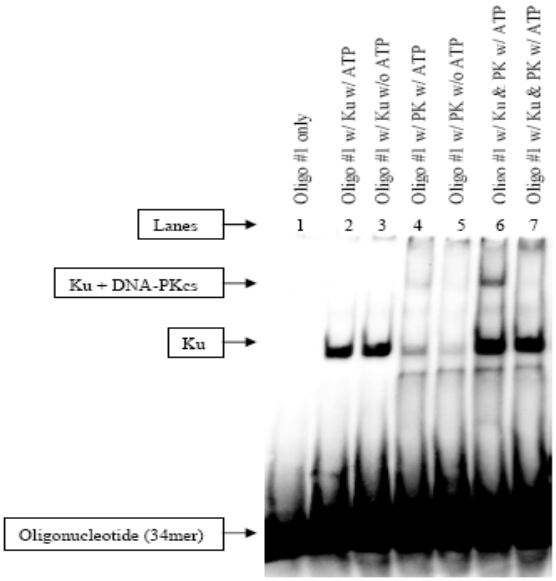


Figure 17. Alkaline agarose gel of RsaI digested PCR product treated with various NHEJ repair factors. Ku = Ku heterodimer; PK = protein kinase holoenzyme; Lig4 = LigaseIV/XRCC4.

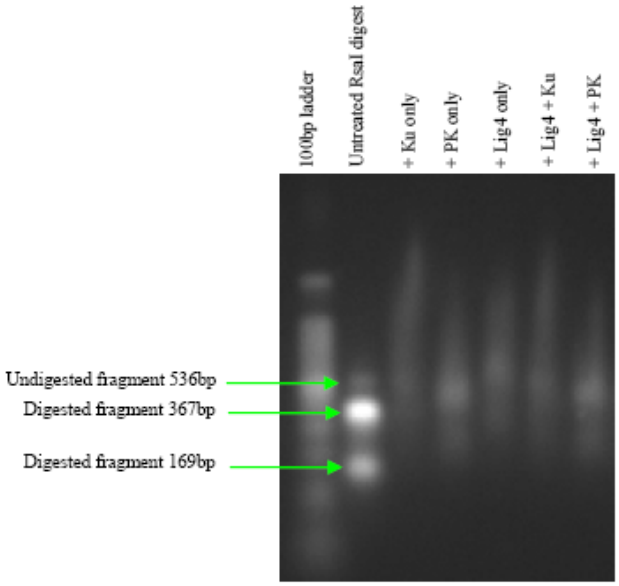


Figure 18. Agarose gel of RsaI digested PCR product treated with various NHEJ repair factors. Ku = Ku heterodimer; PK = protein kinase holoenzyme; Lig4 = LigaseIV/XRCC4.

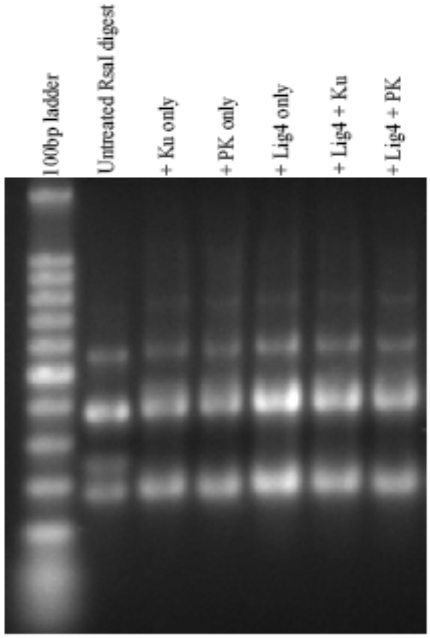


Figure 19. Diagram outlining the experimental procedure used for the repair of DNA double strand breaks via the MMEJ pathway.

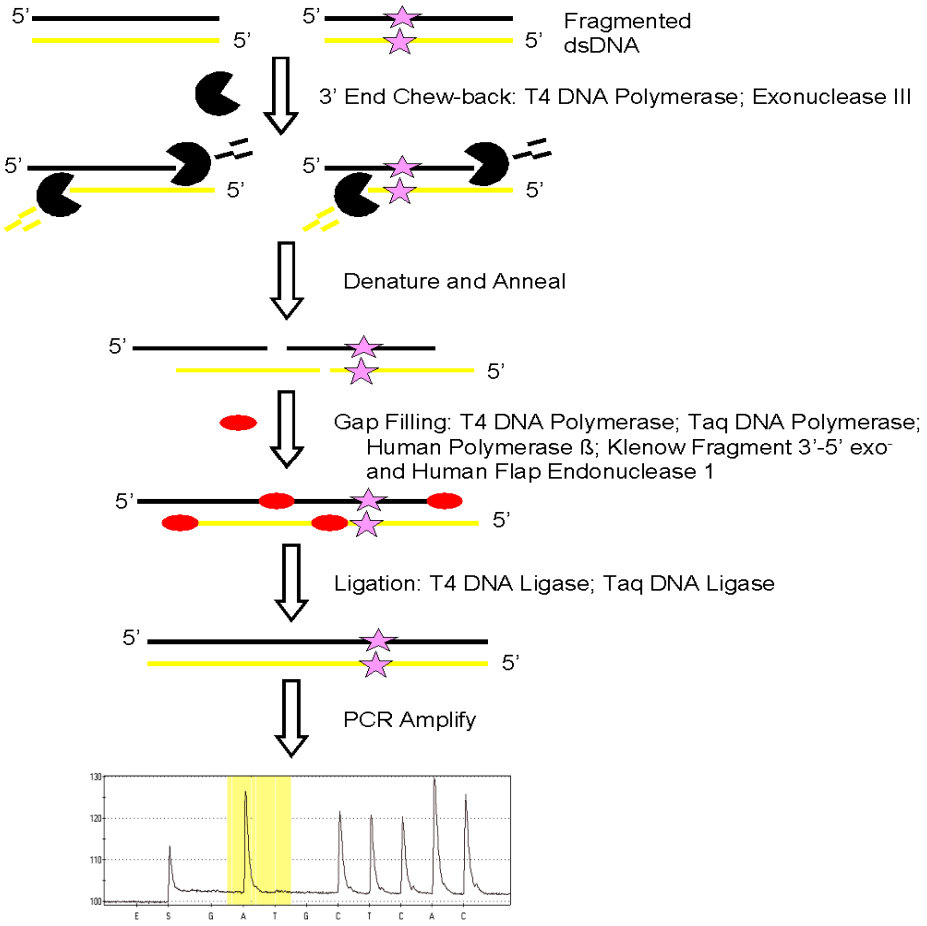


Figure 20. Diagram outlining the experimental procedure used for our modified oligonucleotide ligation SNP detection assay. Initial studies have been carried out using locus rs1058083.

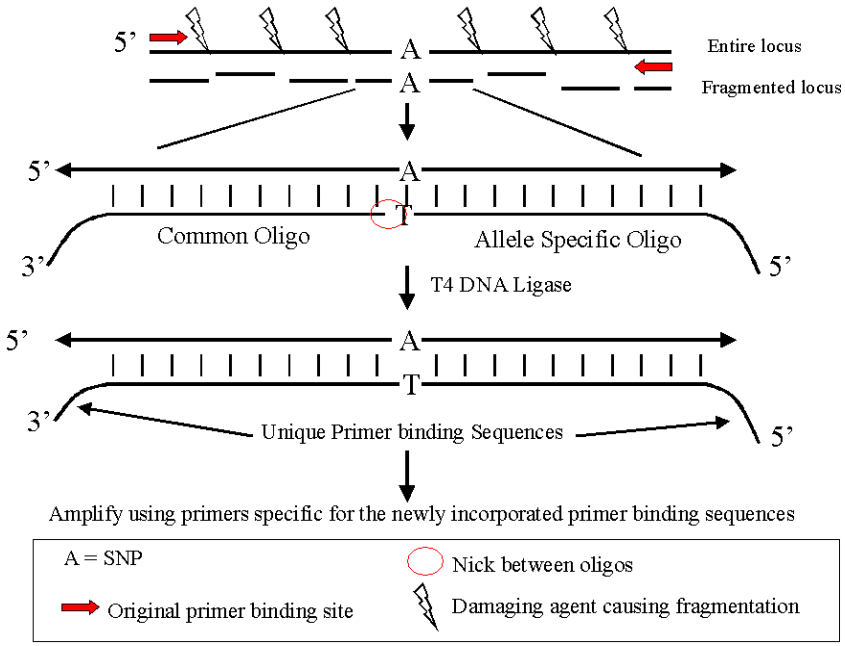


Figure 21. Pyrograms generated from OLA with A) 0ng; B) 1ng; C) 20ng; and D) 50ng of genomic DNA. The correct genotype is TT. All panels: 10uL reactions with 1U T4 DNA ligase and 30fmol of both allele specific oligonucleotides and common oligonucleotide. Annealing temperature was 45^{o}C.

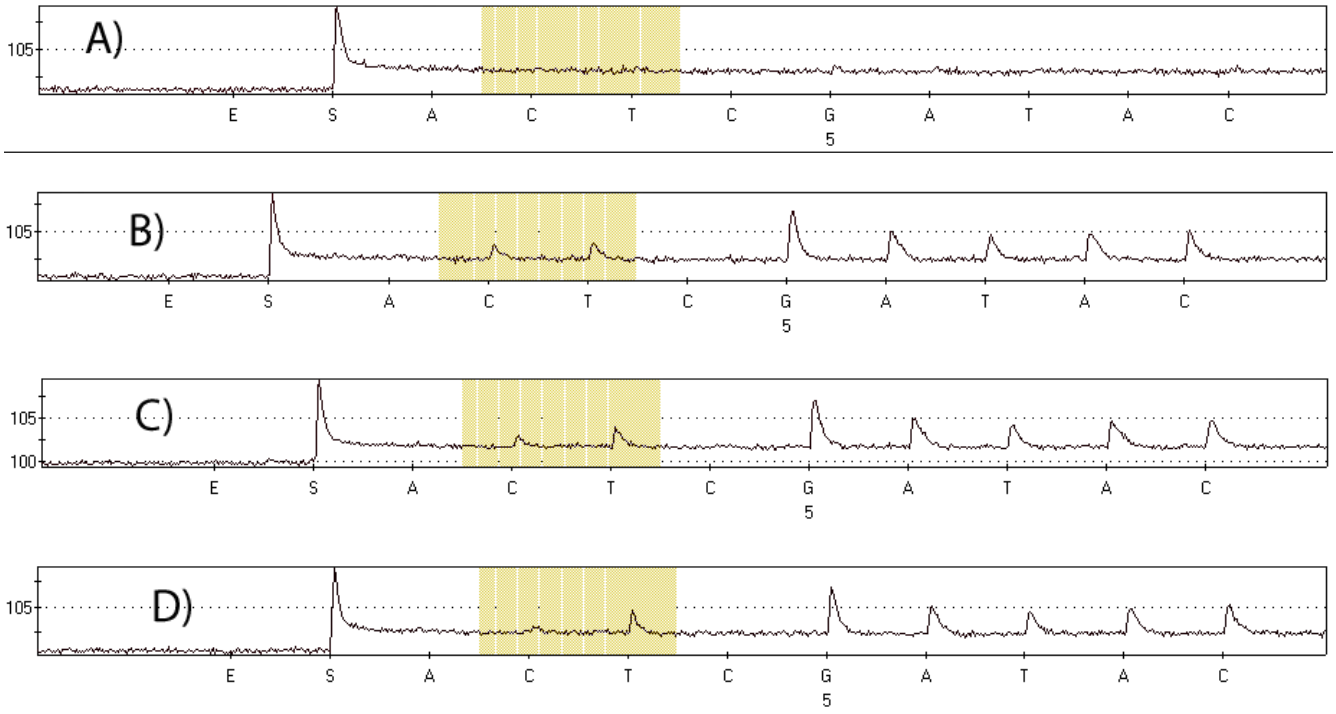


Figure 22. Pyrograms generated from OLA with either A) one round; or B) two rounds of ligation with 20ng DNA, 1U T4 DNA ligase and 30fmol of each allele specific oligonucleotide and common oligonucleotide in a 10uL reaction. Annealing temperature was at 45°C. The correct genotype is TT.

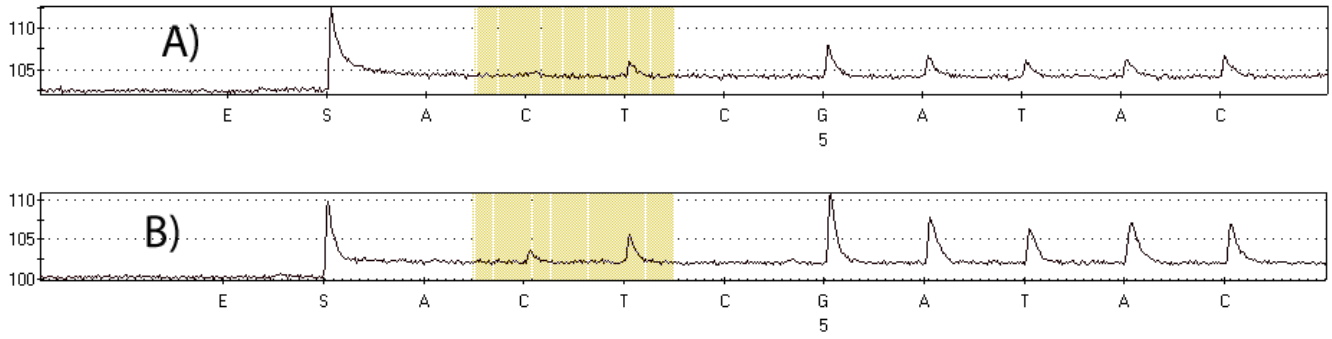
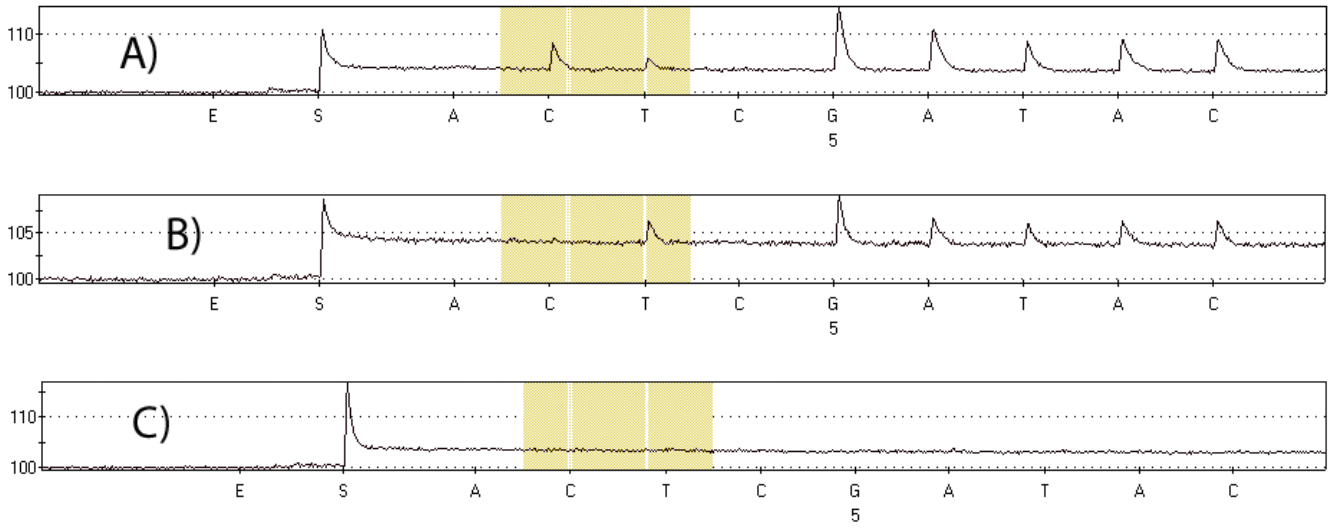


Figure 23. Pyrograms generated from 10uL OLA reactions with 1U T4 DNA ligase and 5fmol of each allele specific oligonucleotide and common oligonucleotide. Annealing temperature was at 25°C. A) Contains 100ng DNA that is homozygote CC; B) contains 100ng DNA that is homozygote TT; and C) contains no DNA, but all other components (ligase and oligonucleotides) are present.



APPENDIX B: TABLES

Table 1. SNP Genotyping Data Using 5 SNPs from K. Kidd Panel (19 SNPs)

	Individual	279844	1058083	13182883	560681	740598
1	**1651**	**A/A**	**T/T (C/T)**	**C/C**	**C/C**	**T/T**
2	**1651 (2)**	**A/T**	**T/T**	**C/C**	**T/T**	**T/T**
3	**1681**	**A/T**	**C/C**	**C/C**	**C/T**	**T/T**
4	**1657**	**T/T**	**T/T**	**T/T**	**C/T**	**C/T**
5	**1649**	**A/T**	**C/C**	**C/T**	**T/T**	**C/T**
6	**9947A**	**A/T**	**C/T**	**C/C**	**C/T**	**T/T**
7	**9948**	**A/T**	**C/C**	**C/C**	**C/T**	**T/T**
8	**100**	**A/T**	**C/C**	**C/T**	**T/T**	**C/T**
9	**101**	**A/A**	**C/T**	**C/C**	**C/T**	**T/T**
10	**102**	**A/A**	**C/C**	**C/T**	**C/C**	**T/T**
11	**103**	**T/T**	**C/C**	**C/T**	**C/T**	**C/T**
12	**104**	**A/T**	**C/C**	**C/C**	**C/T**	**T/T**
13	**CC20**	**T/T**	**T/T**	**C/T**	**T/T**	**T/T**
14	**L100**	**A/T**	**T/T**	**T/T**	**T/T**	**T/T**
15	**AA1**	**A/A**	**C/C**	**C/C**	**C/T**	**T/T**
16	**AA8**	**A/A**	**T/T**	**T/T**	**T/T**	**T/T**
17	**AA72**	**T/T**	**C/C**	**C/C**	**T/T**	**C/T**
18	**AA75**	**A/T**	**C/C**	**C/C**	**T/T**	**C/C**
19	**AA76**	**A/A**	**C/T**	**T/T**	**T/T**	**C/C**
20	**K562**	**A/T**	**T/T**	**C/C**	**C/T**	**T/T**

Table 2: List of all oligonucleotide and repair substrate concentrations used for optimizing the repair of synthetic oligonucleotides at locus rs1058083.

[1]Repair Substrate Concentration	[1]Flanking Oligonucleotide 1 & 2 Concentration	Buffer	[2]Annealing Conditions	Ligation Conditions
100ng	100ng	Water / PCR Buffer II	75°C-45°C Step-down	4°C; 20-24 hrs
10ng	10ng	Water / PCR Buffer II	75°C-45°C Step-down	4°C; 20-24 hrs
1ng	1ng	Water / PCR Buffer II	75°C-45°C Step-down	4°C; 20-24 hrs
500pg	500pg	Water / PCR Buffer II	75°C-45°C Step-down	4°C; 20-24 hrs
200pg	200pg	Water / PCR Buffer II	75°C-45°C Step-down	4°C; 20-24 hrs
200pg	100pg	Water / PCR Buffer II	75°C-45°C Step-down	4°C; 20-24 hrs
200pg	50pg	Water / PCR Buffer II	75°C-45°C Step-down	4°C; 20-24 hrs
100pg	50pg	Water / PCR Buffer II	75°C-45°C Step-down	4°C; 20-24 hrs
150pg	150pg	Water / PCR Buffer II	75°C-45°C Step-down	4°C; 20-24 hrs
50pg	50pg	Water / PCR Buffer II	75°C-45°C Step-down	4°C; 20-24 hrs
50pg	50pg	Water	75°C-45°C Step-down	4°C; 20-24 hrs
10pg	10pg	Water / PCR Buffer II	75°C-45°C Step-down	4°C; 20-24 hrs
1pg	1pg	Water / PCR Buffer II	75°C-45°C Step-down	4°C; 20-24 hrs

[1]Fifty picograms of each oligonucleotide was found to be optimal for repairing this locus.

[2]Step-down annealing conditions were programmed in such a way that each temperature beginning at 75°C was held for 2 minutes before a 3°C decrease occurred. Three degree decreases in temperature only occurred until 60°C where 5°C decreases were then programmed and held for 4 minutes until 45°C was reached.

Table 3. List of all oligonucleotide and repair substrate concentrations used for repair of synthetic oligonucleotides at locus rs1058083.

Reaction Volume (µL)	Repair Substrate Concentration (ng)	Synthetic (Flanking) Oligonucleotide Concentration	Additive and Concentration
25	0.1-1X10^{-8}	0.1-1X10^{-8}	No Additives
25	0.1-1X10^{-8}	0.1-1X10^{-8}	10% DMSO
25	0.1-1X10^{-8}	0.1-1X10^{-8}	100mM NaCl
25	0.1-1X10^{-8}	0.1-1X10^{-8}	10% DMSO & 100mM NaCl
25	0.01	0.01	None
25	0.001	0.001	None
25	0.001	0.0001	None
25	0.0001	0.0001	None
25	1X10^{-5}	0.0001	None
25	0.0001	1X10-5	None
25	1X10^{-5}	1X10-5	None
25	1X10^{-6}	1X10-5	None
25	1X10-5	1X10-6	None
25	1X10-6	1X10-6	None
25	1X10-7	1X10-6	None
25	1X10-6	1X10-7	None
25	1X10-7	1X10-7	None
25	1X10-8	1X10-7	None
25	1X10-7	1X10-8	None
25	1X10-8	1X10-8	None
25	1X10-9	1X10-8	None
25	0.01	0.01	5% DMSO & 50mM NaCl
25	0.001	0.001	5% DMSO & 50mM NaCl
25	0.001	0.0001	5% DMSO & 50mM NaCl
25	0.0001	0.0001	5% DMSO & 50mM NaCl
25	1X10^{-5}	0.0001	5% DMSO & 50mM NaCl
25	0.0001	1X10-5	5% DMSO & 50mM NaCl
25	1X10^{-5}	1X10-5	5% DMSO & 50mM NaCl
25	1X10^{-6}	1X10-5	5% DMSO & 50mM NaCl
25	1X10-5	1X10-6	5% DMSO & 50mM NaCl
25	1X10-6	1X10-6	5% DMSO & 50mM NaCl
25	1X10-7	1X10-6	5% DMSO & 50mM NaCl
25	1X10-6	1X10-7	5% DMSO & 50mM NaCl
25	1X10-7	1X10-7	5% DMSO & 50mM NaCl
25	1X10-8	1X10-7	5% DMSO & 50mM NaCl
25	1X10-7	1X10-8	5% DMSO & 50mM NaCl
25	1X10-8	1X10-8	5% DMSO & 50mM NaCl
5 & 10	0.1-1X10^{-8}	0.1-1X10^{-8}	None
5 & 10	0.1-1X10^{-8}	0.1-1X10^{-8}	5% DMSO
5 & 10	0.1-1X10^{-8}	0.1-1X10^{-8}	50mM NaCl
5 & 10	0.1-1X10^{-8}	0.1-1X10^{-8}	50% DMSO & 50mM NaCl
5 & 10	0.1-1X10^{-8}	0.1-1X10^{-8}	2.5% DMSO
5 & 10	0.1-1X10^{-8}	0.1-1X10^{-8}	25mM NaCl
5 & 10	0.1-1X10^{-8}	0.1-1X10^{-8}	12.5% DMSO & 25mM NaCl
5 & 10	0.1-1X10^{-8}	0.1-1X10^{-8}	2.5% DMSO & 25mM NaCl
5 & 10	0.01	0.001	5% DMSO & 50mM NaCl

5 & 10	0.001	0.001	5% DMSO & 50mM NaCl
5 & 10	0.0001	0.001	5% DMSO & 50mM NaCl
5 & 10	0.01	0.0001	5% DMSO & 50mM NaCl
5 & 10	0.001	0.0001	5% DMSO & 50mM NaCl
5 & 10	0.0001	0.0001	5% DMSO & 50mM NaCl
5 & 10	0.01	0.001	5% DMSO & 100mM NaCl
5 & 10	0.001	0.001	5% DMSO & 100mM NaCl
5 & 10	0.0001	0.001	5% DMSO & 100mM NaCl
5 & 10	0.01	0.0001	5% DMSO & 100mM NaCl
5 & 10	0.001	0.0001	5% DMSO & 100mM NaCl
5 & 10	0.0001	0.0001	5% DMSO & 100mM NaCl

Table 4. Substrate-mediated gap repair success rates using varying concentrations of synthetic oligonucleotides and repair substrate. Data represent replicates of 14.

Synthetic Oligonucleotide Concentration	Repair Substrate Concentration	Repair Rate
100 attomoles	20 attomoles	0%
100 attomoles	50 attomoles	0%
100 attomoles	100 attomoles	29%
100 attomoles	200 attomoles	29%
200 attomoles	20 attomoles	36%
200 attomoles	50 attomoles	93%
200 attomoles	100 attomoles	72%
200 attomoles	200 attomoles	100%

Table 5. List of all combinations of repair enzymes attempted, to date, for the repair of DNA double strand breaks using NHEJ.

Blunting Reaction Enzymes		End Binding Enzymes	Ligation Reaction Enzymes	
T4 DNA Polymerase	Human DNA polymerase beta	Human Ku 70/80 complex	Human LigaseIV/XRCCIV	T4 DNA ligase
-	-	+	+	-
-	-	+	-	+
+	-	+	+	-
+	-	+	-	+
-	+	+	+	-
-	+	+	-	+
+	+	+	+	-
+	+	+	-	+

APPENDIX C: PRESENTATIONS AND PUBLICATIONS

PRESENTATIONS

2007 Biochemistry of Dry State DNA: Comparison of Depurination Rates of Bases. Pope, A. and Ballantyne, J. Forensic Sciences Symposium, Nova Southeastern University, Ft. Lauderdale, FL.

2007 Getting Blood form a Rock: Getting More and More from Less and Less. International Society for Optical Engineering (SPIE) Defense and Security Symposium, Orlando, FL

2007 Double Strand Break Repair of Damaged DNA Templates. Lamers, R and Ballantyne, J. The NIJ Conference. Arlington, VA

2007 Getting Blood from a Stone: Getting More and More Forensic Evidence form Less and Less. Qiagen Corporation, Gaithersburg, MD.

2008 Double Strand Break Repair of Damaged DNA Templates. Ballantyne J. and Lamers, R. The NIJ Conference, Crystal City, VA.

2008 The Fundamental Biochemistry of Dry State DNA: Hydrolytic Reactions. Marrone, A. and Ballantyne, J. 19th International Symposium on Human Identification, Hollywood, CA.

2008 Forensic Biology Research at the National Center for Forensic Science. The 20th EDNAP (European DNA Profiling Group) Meeting, Zurich, Switzerland (NR/I/I)

APPENDIX D: LIST OF REFERENCES

[1] M.A.Jobling and P.Gill, Encoded evidence: DNA in forensic analysis, Nat.Rev.Genet. 5 (2004) 739-751.

[2] T.Lindahl, Instability and decay of the primary structure of DNA, Nature 362 (1993) 709-715.

[3] L.McNally, R.C.Shaler, M.Baird, I.Balazs, F.P.De, and L.Kobilinsky, Evaluation of deoxyribonucleic acid (DNA) isolated from human bloodstains exposed to ultraviolet light, heat, humidity, and soil contamination, J Forensic Sci 34 (1989) 1059-1069.

[4] L.McNally, R.C.Shaler, M.Baird, I.Balazs, L.Kobilinsky, and F.P.De, The effects of environment and substrata on deoxyribonucleic acid (DNA): the use of casework samples from New York City, J Forensic Sci 34 (1989) 1070-1077.

[5] M.Hoss, P.Jaruga, T.H.Zastawny, M.Dizdaroglu, and S.Paabo, DNA damage and DNA sequence retrieval from ancient tissues, Nucleic Acids Res. 24 (1996) 1304-1307.

[6] SWGDAM. Revised Validation Guidelines. Forensic Science Communications 6(3). 2003.
Ref Type: Journal (Full)

[7] TWGDAM and TWGoDAM. Guidelines for a Quality Assurance Program for DNA Analysis. Crime Laboratory Digest 22(2), 21-43. 1995.
Ref Type: Journal (Full)

[8] D.E.Adams, L.A.Presley, A.L.Baumstark, K.W.Hensley, A.L.Hill, K.S.Anoe, P.A.Campbell, C.M.McLaughlin, B.Budowle, A.M.Giusti, and ., Deoxyribonucleic acid (DNA) analysis by restriction fragment length polymorphisms of blood and other body fluid stains subjected to contamination and environmental insults, J Forensic Sci 36 (1991) 1284-1298.

[9] B.Budowle, T.R.Moretti, K.M.Keys, B.W.Koons, and J.B.Smerick, Validation studies of the CTT STR multiplex system, J Forensic Sci 42 (1997) 701-707.

[10] S.Cosso and R.Reynolds, Validation of the AmpliFLP D1S80 PCR Amplification Kit for forensic casework analysis according to TWGDAM guidelines, J Forensic Sci 40 (1995) 424-434.

[11] A.M.Lins, K.A.Micka, C.J.Sprecher, J.A.Taylor, J.W.Bacher, D.R.Rabbach, R.A.Bever, S.D.Creacy, and J.W.Schumm, Development and population study of an eight-locus short tandem repeat (STR) multiplex system, J Forensic Sci 43 (1998) 1168-1180.

[12] J.M.Wallin, M.R.Buoncristiani, K.D.Lazaruk, N.Fildes, C.L.Holt, and P.S.Walsh, TWGDAM validation of the AmpFISTR blue PCR amplification kit for forensic casework analysis, J Forensic Sci 43 (1998) 854-870.

[13] A.Hall and J.Ballantyne, Characterization of UVC-induced DNA damage in bloodstains: forensic implications, Anal.Bioanal.Chem. 380 (2004) 72-83.

[14] Y.Aylon and M.Kupiec, New insights into the mechanism of homologous recombination in yeast, Mutat.Res. 566 (2004) 231-248.

[15] A.J.Bishop and R.H.Schiestl, Homologous recombination as a mechanism for genome rearrangements: environmental and genetic effects, Hum.Mol Genet. 9 (2000) 2427-334.
[16] D.O.Ferguson and F.W.Alt, DNA double strand break repair and chromosomal translocation: lessons from animal models, Oncogene 20 (2001) 5572-5579.
[17] M.L.Hefferin and A.E.Tomkinson, Mechanism of DNA double-strand break repair by non-homologous end joining, DNA Repair (Amst) 4 (2005) 639-648.
[18] W.D.Heyer and J.Kohli, Homologous recombination, Experientia 50 (1994) 189-191.
[19] R.D.Johnson and M.Jasin, Double-strand-break-induced homologous recombination in mammalian cells, Biochem.Soc Trans. 29 (2001) 196-201.
[20] R.Kanaar, J.H.Hoeijmakers, and G.van, Molecular mechanisms of DNA double strand break repair, Trends Cell Biol 8 (1998) 483-489.
[21] P.Karran, DNA double strand break repair in mammalian cells, Curr.Opin.Genet.Dev. 10 (2000) 144-150.
[22] M.R.Lieber, Y.Ma, U.Pannicke, and K.Schwarz, Mechanism and regulation of human non-homologous DNA end-joining, Nat.Rev.Mol Cell Biol 4 (2003) 712-720.
[23] M.J.McIlwraith, D.E.Van, J.Y.Masson, A.Z.Stasiak, A.Stasiak, and S.C.West, Reconstitution of the strand invasion step of double-strand break repair using human Rad51 Rad52 and RPA proteins, J Mol Biol 304 (2000) 151-164.
[24] E.Pastwa and J.Blasiak, Non-homologous DNA end joining, Acta Biochim.Pol. 50 (2003) 891-908.
[25] Y.Ma, H.Lu, B.Tippin, M.F.Goodman, N.Shimazaki, O.Koiwai, C.L.Hsieh, K.Schwarz, and M.R.Lieber, A biochemically defined system for mammalian nonhomologous DNA end joining, Mol Cell 16 (2004) 701-713.
[26] K.K.Kidd, A.J.Pakstis, W.C.Speed, E.L.Grigorenko, S.L.Kajuna, N.J.Karoma, S.Kungulilo, J.J.Kim, R.B.Lu, A.Odunsi, F.Okonofua, J.Parnas, L.O.Schulz, O.V.Zhukova, and J.R.Kidd, Developing a SNP panel for forensic identification of individuals, Forensic Sci Int 164 (2006) 20-32.
[27] D.G.Gibson, G.A.Benders, C.ndrews-Pfannkoch, E.A.Denisova, H.Baden-Tillson, J.Zaveri, T.B.Stockwell, A.Brownley, D.W.Thomas, M.A.Algire, C.Merryman, L.Young, V.N.Noskov, J.I.Glass, J.C.Venter, C.A.Hutchison, III, and H.O.Smith, Complete chemical synthesis, assembly, and cloning of a Mycoplasma genitalium genome, Science 319 (2008) 1215-1220.
[28] E.K.Hanson and J.Ballantyne, Whole genome amplification strategy for forensic genetic analysis using single or few cell equivalents of genomic DNA, Anal.Biochem. 346 (2005) 246-257.
[29] F.B.Dean, S.Hosono, L.Fang, X.Wu, A.F.Faruqi, P.Bray-Ward, Z.Sun, Q.Zong, Y.Du, J.Du, M.Driscoll, W.Song, S.F.Kingsmore, M.Egholm, and R.S.Lasken, Comprehensive human genome amplification using multiple displacement amplification, Proc Natl Acad Sci U S A 99 (2002) 5261-6.
[30] J.F.Holbrook, D.Stabley, and K.Sol-Church, Exploring whole genome amplification as a DNA recovery tool for molecular genetic studies, J.Biomol.Tech. 16 (2005) 125-133.

REVIEWER COMMENTS

Reviewer #1:

Award number: 2006-DN-BX-K005

Project title: Double Strand Break Repair of Highly Degraded DNA

Practitioner need addressed by project *(to be provided by the NIJ Program Manager)***:**
The need for tools that increase the success rate of obtaining DNA profiles from compromised (damaged) DNA evidence.

Review of Draft Final Report for Project 2006-DN-BX-K005

Substantive Quality:

This project investigated three methods for repair of double stranded breaks in DNA to determine if a method could be developed that would allow subsequent DNA typing of repaired samples. None of the three methods evaluated, gap repair, nonhomologous end joining, and a modified oligo ligation assay proved successful in repairing a double stranded break in human genomic DNA. There was partial success using synthetic oligos during the experiments using the gap repair mechanism, but these were not scalable to genomic DNA samples.

This final report is fairly well written, with only a moderate number of grammatical and spelling errors. The executive summary accurately describes the contents of the full

report and is written so that a general audience would be capable of understanding the results of the research project. The data contained within the report supports the results and conclusions reached during the project. All of the methods tested were based upon current knowledge of DNA repair mechanisms. The experimental design was appropriate, but could have been further improved through additional optimization of conditions enzyme utilization and cofactors with synthetic DNA strands prior to attempts at genomic repair. Too much emphasis was placed on obtained in producing degraded DNA and screening of various SNP locations prior to ensuring any of the methods had a reasonable chance of success. The theoretical calculations included in the conclusion of the report would have been more appropriately done prior to beginning actual experimentation, as this could have potentially prevented many ineffective studies.

Implications of the Research:

The studies conducted as part of this project provide minimal additional information to the basic knowledge regarding DNA repair mechanisms. Although the results obtained during this project may prove of value in future experiments once a better understanding of *in vivo* DNA repair mechanisms has been developed, the only current significant finding of interest to the forensic community is that repair of extremely degraded DNA is not a viable option in analysis of samples. Further research into this area should be focused on optimization of the repair conditions using theoretical calculations and synthetic DNA prior to actual experimentation on human genomic DNA.

Relevance:

This report would be graded as FAIR regarding its relevance for policy or practice based upon the following reasons: 1) this research did not lead to any new methodology that can assist local, state or federal laboratories in performing casework DNA analysis; 2) this research did not lead to improved methods that could be used in analysis of samples recovered from mass disasters; and 3) the studies conducted as part of this project added a minimal new information to the basic scientific understanding of DNA repair mechanisms.

Publication of the results obtained during this project would provide a baseline for additional studies in the future, but a more appropriate mode of dissemination would be presentation at various scientific meetings.

Reviewer #2:

Review of project technical report/deliverables

Award number: 2006-DN-BX-K005

Project title: Double Strand Break Repair of Highly Degraded DNA

Practitioner need addressed by project *(to be provided by the NIJ Program Manager)***:**

The need for tools that increase the success rate of obtaining DNA profiles from compromised (damaged) DNA evidence.

Reviewer provides information below:

1. **Substantive Quality**
   - What are the significant findings of the research?
   - Are the findings supported by the research? Was the methodology appropriate and sound?
   - Does the Executive Summary adequately describe the full report?
   - Is the report well-written in terms of style, organization, and format?
   - Classify the overall quality as one of the following: Poor, Fair, Good, and Excellent.

The research described in the report entitled "Double Strand Break Repair of Highly Damaged DNA" represented a fair-to-good approach toward devising an effective strategy for the repair of highly damaged DNA. This work may be considered a logical extension of the subject of some of the principal investigator's previous research efforts

which included study of the nature and extent of the effect of DNA damage on the suitability of biological evidence for forensic genetic identity analysis.

The stated purpose of the report was "to attempt to repair in the test tube highly fragmented DNA in order to retrieve a DNA profile from otherwise intractable environmentally compromised samples". The specific focus of the research was on devising protocols to affect the successful in vitro repair of DNA double-strand breaks. The approaches adopted by the investigator toward that end were well-informed. The multiple attempts to modify and optimize the various enzyme assays designed to achieve double-strand break repair were thorough and well-reasoned.

The investigator approached the problem using multiple strategies. The first strategy was to employ synthetic oligonucleotides as "molecular scaffolding" to capture and hold the damaged DNA in place in order to allow subsequent repair by appropriate enzymes or enzyme complexes. The second attempt was to try to functionally reconstitute in vitro the cellular mechanism for non-homologous end-joining along with the related approach of microhomolgy end-joining. A third attempt involved a somewhat novel modification of a DNA ligation assay to directly detect SNP alleles. Each of these approaches was explored in depth with diverse attempts to measure the effect of various reaction parameters on the underlying mechanisms in order to optimize the repair assay. The principal investigator's draft report on this research was generally well-written and concise; the primary focus remained throughout on the methods employed and the results obtained during the attempt to affect the repair of double-strand breaks of genomic DNA.

Unfortunately, in all cases, the strategies to directly repair damaged genomic DNA were unsuccessful. As follow-up, to attempt to devise model systems for further study, the investigator resorted to simplified direct repair assays employing synthetic oligonucleotides to capture targets for amplifications or as substrates for ligation. These model systems, in both cases, were shown to result in successful rejoining of synthetic double-stranded DNA but the "lessons learned" could not be extended to the more challenging task of repairing genomic DNA. One additional approach involved a modified oligonucleotide ligation assay modified to provide SNP typing without a need for prior repair of the damaged temples. This strategy was also shown to be non-viable.

2. Implications of the Research

- Do the findings make a significant contribution to existing knowledge in the area?
- What are the implications, if any, for further research, program development, and evaluation efforts?

Overall, due to the investigator's inability to affect the desired repair of damaged genomic DNA in the test tube, this report does not make a significant contribution toward the stated goal of the grant which was to increase the success rate of obtaining DNA profiles from compromised DNA evidence. None of the findings can be considered to offer the promise of helping the practicing forensic scientist to overcome the inability to type DNA that has suffered double-stranded breaks. As such, this reviewer does not believe that the report would be of general interest to the forensic community and would not recommend it for publication. The research did indicate that, despite present limitations, at some point rare events such as a reconstituted amplimer might be detected

by whole genome amplification. The utility of whole genome amplification for forensic purposes would seem to be a potentially useful area for further research.

Relevance

- Summarize your overall rating of the report's/deliverable's relevance for policy or practice as one of the following: Poor, Fair, Good, Excellent
- To whom would this report/deliverable be of greatest interest?
- Would you recommend this information for publication?
- What specific dissemination vehicles would be particularly appropriate for publicizing the research (e.g., conferences, scientific literature, etc.)?
- How well does this report/deliverable address the practitioner need as stated above?

The overall rating of this report's relevance for policy or practice would have to be considered to be "poor" because of the failure, despite considerable effort, to achieve the research objective of reconstituting highly degraded DNA samples. As noted by the author, this was a high risk undertaking with little chance of success. The extent of the risk for failure was evidently quite clear at the outset and an inability to affect repair was, in fact, the eventual outcome. To selectively repair a small portion of a badly fragmented genome is widely recognized as a daunting task and the various novel approaches with attendant modifications described in this report were not suitable to yield either repair or a good model system for future research. Because the results were ultimately uniformly negative, I could not recommend that the study be submitted for publication. This report might be of interest to those individuals considering similar research efforts and, in that regard, might be suitable for delivery at a professional conference involving forensic DNA analysts.

Recommended revisions

> Describe fully any revisions or changes that should be made to improve the quality of the report or increase its usefulness. Your anonymous comments and suggestions will be forwarded to the author of the report, who will make appropriate revisions in order to improve the report.

This reviewer does take exception with the principal investigator's conclusion in the Executive Summary. Although the author's findings resulting from this "high risk, low cost" attempt to repair may well indicate the impossibility of the task, they do not entirely preclude eventual technological advances that might make such a daunting task possible. Thus, the ultimate conclusion might have been qualified as follows, "*Using current technology*, sequence complexity of genomic DNA provides an insurmountable computational barrier to reconstituting two contiguous fragments of DNA back to its native state, a requirement for successful repair of breaks".

RESPONSE TO REVIEWERS:

Reviewer #1:

1. **Executive Summary Page 2: In the third bullet the abbreviation 'dsbr' is used without first providing the meaning.**

   Response: The meaning of the abbreviation 'dsbr' has been included in the third bullet of the executive summary.

2. **Executive Summary Page 3: Abbreviation MMEJ is used without first providing the meaning.**

   Response: The meaning of the abbreviation "MMEJ" has been included in the executive summary.

3. **Suggestion: Replace all instances of the abbreviation 'dsbr' with 'double strand break(s)'. This will improve the readability of the report.**

   Response: All 'dsbr' abbreviations have been replaced with 'double strand break repair' to improve the readability of the report.

4. **Figure 3: The labels on the axes of the charts need to be explained. Also a control pyrogram for the rs1058083 must be shown for comparative purposes or a better explanation of how it can be determined that the pyrogram shown actually represents accurate repair of the sequence. The**

manuscript does not clearly demonstrate with data at this point that the repair mechanism tested generated a DNA sequence identical to the starting DNA.

Response: An explanation of the program axis has been provided in the figure legend for Figure 3.

The following has been added to the text that references Figure 3 (p 19): "The repair took place using synthetic oligonucleotides. Repair was error free in that the sequence of the repaired substrates was fund to be that expected from direct ligation of the two oligonucleotides"**.**

5. **Figure 16: Although it may seem obvious, a control lane of Ku only should be included to demonstrate that the mobility shift only occurs in the presence of the labeled oligo.**

   Response: Ku is unlabeled and therefore would not show up on the mobility shift assay. The oligo on its own (lane 1) and the oligo with Ku (lane 3) clearly shows that Ku is binding to the oligo. This additional control was not considered necessary at the time.

6. **On all figures depicting results of replicate analysis using a specific set of conditions, the individual panels should be labeled. Also there was a label, presumably indicating the SNP mutation above many of the panels, but no**

explanation as to the significance relative to the individual experiment results depicted.

Response: Additional labels have been added to figures with multiple panels in order to improve the readability of the report.

Confusing labels that describe the SNP mutation were superfluous and have been removed since they do not aid in the understanding of the figure.

7. **It was noted that the reference to Figure 22 was out of order in the draft report. The report could be slightly improved through the grouping of all experiments related to the synthetic oligonucleotide sequences using the NHEJ method despite the fact that this is not necessarily the time sequence in which they occurred.**

   Response: The reference to Figure 22 that appeared on p. 36 was incorrect. The reference was actually to Figure 17. This has been corrected in the manuscript. The authors feel that the order of experiments provided in the report is sufficient and does not require a re-grouping of experiments.

Reviewer #2:

1. **This reviewer does take exception with the principal investigator's conclusion in the Executive Summary. Although the author's findings resulting from this "high risk, low cost" attempt to repair may well indicate the impossibility of the task, they do not entirely preclude eventual technological advances that might make such a daunting task possible. Thus, the ultimate conclusion might have been qualified as follows, "*Using current technology*, sequence complexity of genomic DNA provides an insurmountable computational barrier to reconstituting two contiguous fragments of DNA back to its native state, a requirement for successful repair of breaks".**

   Response: The conclusion provided in the executive summary has been revised and now states: "Using current technology, sequence complexity of genomic DNA provides an insurmountable computational barrier to reconstituting two

contiguous fragments of DNA back to its native state, a requirement for successful repair of breaks".

.