Federal Agency: Office of Justice Programs, National Institute of Justice, Department of Justice

Federal Grant: Award Number 2010-DN-BX-K190

Project Title: Tools for improving the quality of aged, degraded, damaged, or otherwise compromised DNA evidence

Final Report

PI: Michael M. Cox, Department of Biochemistry, University of Wisconsin-Madison

tel: 608-262-1181; email: cox@biochem.wisc.edu

Submission Date: November 4, 2013

Name of Submitting Official, Title, and Contact Information, if other than PD/PI:
Cheryl Gest
Managing Officer
preaward@rsp.wisc.edu
608-262-3822 FAX: 608-262-5111

DUNS and EIN Numbers
DUNS: 161202122
EIN Number: 396006492

Recipient Organization:
The Board of Regents of the University of Wisconsin System
Research and Sponsored Programs
21 N. Park Street, Suite 6401
Madison, WI 53715-1218

Recipient Identifying Number or Account Number:
144-PRJ42NZ

Project/Grant Period:
01/01/2011-12/31/2013

Reporting Period End Date (final report):
December 31, 2013

Report Term or Frequency:
semi-annual progress report

Final Progress Report Date: 12/31/2013

PI Signature:

Michael M. Cox Ph.D.
Evelyn M. Mercer Professor of Biochemistry

Signature of Submitting Official:

Cheryl Gest

Abstract

Almost every day, DNA samples are collected from the scenes of crimes or disasters that are too degraded for standard forensic DNA analytical procedures. This fact represents an ongoing impediment to law enforcement and victim identification efforts. Many law enforcement agencies also possess archived crime scene evidence from cold cases that are decades old, in which the DNA has become too damaged to analyze. The major problem in these samples is the presence of DNA double strand breaks. The purpose of the work carried out under grant 2010-DN-BX-K190 is to develop a new method to repair double strand breaks in forensic DNA samples, as a pre-treatment for the standard STR analysis protocols. As part of this effort, we have also developed reproducible procedures for the artificial degradation of human DNA samples, using ionizing radiation to inflict a DNA damage profile that reprises that of a typical degraded forensic sample. Using this type of DNA as a test bed, we have developed a protocol that is successful in increasing/restoring missing or substandard signals at two STR loci. The protocol utilizes the bacterial RecA protein, single-stranded DNA binding protein (SSB), and bacterial DNA polymerase I, in concert with a targeting oligonucleotide. The reactions promoted by these reagents effectively restore damaged DNA flanking a particular STR locus. With the most developed protocol, signal restoration is successful approximately 20% of the time. With a few exceptions, the restored signals are accurate. The artifacts arising in the exceptions have been traced to the targeting oligonucleotides. Efforts to further develop this technology are continuing, focused on new RecA protein variants that increase signal strength and re-designed targeting oligonucleotides.

We have also been successful in demonstrating proof of principle in efforts to recover targeted DNA segments and remove them from bulk DNA in an effort to concentrate them and eliminate conditions that could inhibit STR amplification.

Table of Contents

Executive Summary

DNA degradation remains a significant problem in forensics. Double strand breaks represent a major cause of the loss of STR signals in degraded forensic DNA samples. Accurate repair of double strand breaks has the potential to restore/enhance signals and facilitate the analysis of these samples. We have developed two protocols. One provides an avenue for the reproducible degradation of human DNA samples to use for testing. The other is a protocol for repair that utilizes enzymes derived from bacterial double strand break repair systems.

Reproducible degradation of human DNA samples can be effected using ionizing radiation. The DNA damage profile is similar to that observed in DNA subjected to oxidative damage from a variety of environmental sources. Signals derived from STR loci that generate longer PCR products are lost first, reprising the "ski slope" effect often observed in degraded forensic DNA samples.

The repair of the artificially degraded DNA samples utilizes three bacterial proteins: RecA, SSB, and DNA polymerase I. Targeting oligonucleotides that reprise the unique DNA sequence on either side of a particular STR locus are also used. The RecA protein and DNA polymerase function in succession to generate an unbroken DNA segment that permits the accurate STR signal restoration. Several hundred trials of this protocol have been carried out. Under current best practice conditions, restoration or enhancement of the correct STR signals are observed approximately 20% of the time. Artifacts arising in some trials have been traced to the presence of the targeting oligonucleotides during the subsequent PCR amplification using standard DNA analysis kits. Ongoing research utilizing RecA protein variants with enhanced activity exhibits promise for generating a more robust repair reaction. Elimination of the observed artifacts may require refinement of the protocol to remove the targeting oligonucleotides prior to PCR amplification.

A parallel effort at LSU, overseen by John R. Battista, has been directed at the recovery of repaired DNA segments, separating them from the bulk DNA. The objective is to concentrate the STR-containing DNA segments and remove them from any conditions that might inhibit STR amplification. Proof of principle has been accomplished in this effort.

I Introduction

1.1 STR typing

DNA genotyping based on the PCR amplification and electrophoretic analysis of Short Tandem Repeats (STRs) plays a prominent role in forensic science (1). A STR is a polymorphism found in mammalian DNA, a sequence of nucleotides (ranging between 2-10 base pairs) that is tandemly repeated at a locus. By examining several STR loci one can establish the unique genetic profile of an individual, linking biological evidence from a crime to the perpetrator or to other crimes by the same person. Tetranucleotide repeats are the mainstay of forensic DNA typing and criminal offender databasing (1). There are only 33 possible tetranucleotide motifs (2), and the consensus motif sequences, mostly AGAT and GATA, are ubiquitous in the human genome. The number of repeat units at these loci varies from as few as four to as many as 50.

The use of STRs in forensic analysis was pioneered in the United Kingdom in 1984 (3). In 1997 the forensic community in the United States chose thirteen STR loci to form the essential core of its Combined DNA Index System (CODIS) casework and offender databases. These loci are: D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, CSF1PO, FGA, Th01, TPOX, and vWA. There are enough different alleles at these STR loci in any major population or subpopulation to ensure that individuals will be heterozygous at most loci, enabling unambiguous identification in most cases (1).

Efforts are currently underway to expand the list of CODIS core STR loci (4, 5). This is being done for several reasons. The sheer growth in U.S. and other national databases (over 12 million profiles and growing) is revealing a need for additional loci. A number of loci, particularly on the Y chromosome, would greatly aid in kinship analyses/familiar searching. Some new loci

would also be useful for immigration testing. The additions focus generally on loci utilized in the European Standard Set (ESS) and Y STRs. The first kits with expanded sets of STR loci (24 total) were released from Promega (PowerPlex Fusion) and Applied Biosystems (GlobalFiler) in 2012.

1.2 DNA Degradation and STR Typing

Under conditions where biological samples are well-preserved, genotyping using STRs is a robust technology that can be applied with confidence. However, forensic scientists are often confronted with biological evidence in which the DNA is present in a degraded form that limits or prevents PCR amplification due to small template size, inhibition and defect (6-8). In these samples the DNA is highly fragmented and contains a large number of modified nucleotides. Commercial STR analysis protocols generate PCR fragments of between 100 – 500 base pairs. If the fragments of target DNA are smaller than this on average, effective PCR amplification will not be obtained. Even if degraded DNA is of sufficient length, damage resulting from covalent modification of the nucleobases found in the degraded DNA will inhibit amplification. Polymerases used in PCR amplifications (such as Taq) are blocked by a variety of lesions (9). A blocking lesion in the template will terminate strand extension, effectively inhibiting amplification and potentially forcing the polymerase to jump between templates, giving rise to non-specific amplification products (10, 11). A variety of DNA damages are found associated with degraded DNA: single-strand breaks, base modifications, base loss, deoxyribose damage, intra- and inter-strand cross links, and DNA-protein cross links. Restoration of degraded and damaged DNA for forensic typing is frustrated by the variety of DNA damage that can be present and the fact that no single repair process is capable of dealing with all forms of DNA damage (11). Protocols are currently available for ameliorating some kinds of damage, such as base alterations. For example, New

England Biolabs offers a kit containing enzymes that will repair abasic sites, pyrimidine dimers, and other base lesions, called the PreCR kit. However, the major problem in degraded forensic DNA samples comes from double strand DNA breaks, and no available kit or protocol addresses these lesions.

1.3 DNA Repair.

Given the fundamental role of DNA in the storage and transmission of genetic information, it is somewhat surprising that this molecule is relatively unstable *in vivo* and *in vitro* (12). DNA is susceptible to spontaneous decomposition and it can be damaged by a myriad of physical and chemical agents derived from endogenous and exogenous sources (13). The effect that DNA damage has on a cell is in large part determined by the interaction between the lesion formed and the replicative DNA polymerase (14). Lesions frequently inhibit this polymerase, blocking DNA replication, and unless the blockage is removed the cell will die. As a consequence, all characterized species express proteins that either detect and repair DNA damage or allow the offending lesion to be bypassed. The primary function of most DNA repair proteins *in vivo* is to correct DNA damage before the polymerase reaches the lesion, thus avoiding the problem. An impressive arsenal of proteins is dedicated to this function. Repair is accomplished in a few cases by direct chemical reversal of base damage. However, in most cases, repair involves excising damage and using the undamaged complementary strand to restore the original sequence (13). Direct reversion mechanisms are of necessity lesion-specific. Known direct reversion mechanisms affect only pyrimidine dimers and certain methylated bases. In contrast many lesions are targeted by excision repair systems. There are two types of excision repair: nucleotide excision repair (NER) and base excision repair (BER).

During NER, the lesion is recognized and the repair complex nicks the DNA backbone 5′ and 3′ of the lesion. When the lesion is removed by bacterial NER systems, a 12 or 13 base pair gap is generated in the strand. This gap is filled by a DNA polymerase, restoring the original DNA sequence. The proteins that catalyze NER recognize unnatural bends that are introduced by the damaged base into DNA; they do not identify specific lesions. As a consequence NER can remove many types of DNA damage and, in general, a single species encodes only one NER complex.

Base excision repair utilizes a similar strategy. However, there are many types of BER proteins, as they tend to be specific in their substrates. Each protein recognizes a limited number of lesions. The damaged nucleobase is removed during BER by cleavage of the N-glycosidic bond linking the base to the deoxyribose, leaving an apurinic/apyrimidinic (AP) site in the DNA. The sugar remaining in the AP site is then removed by the action of an AP endonuclease and deoxyribophosphodiesterase, resulting in a one base pair gap in the DNA that is filled in by a DNA polymerase.

If the lesion cannot be removed by NER or BER, or if repair is underway when a replication complex encounters the sequence in question, the replication fork either halts or collapses (Figure 1). There are back up systems that allow the cell either to repair the stalled replication fork and correct the damage found there, or to bypass it.

The accurate repair of stalled replication forks employs genetic recombination, utilizing the redundant genetic information found in sister duplexes to replace the damaged DNA and permit accurate replication restart (15-19). In many cases, this entails the repair of double strand breaks (Figure 1). It is likely that the evolution of genetic recombination systems was driven by the need to repair collapsed forks. When an arm of the replication fork is severed from the replisome by an encounter with a template discontinuity, it is the recombination system that reassembles a viable

replication fork. The double strand DNA break repair system we have set up in this project is based on the first few steps of this replication fork repair process, and utilizes some of the same bacterial enzymes.

The bypass of DNA lesions is achieved with the aid of a special class of DNA polymerases capable of translesion DNA synthesis (tls) (13, 14, 20-23). The tls DNA polymerases are much less affected by blocking lesions, and insert a nucleotide opposite the lesion to extend the growing DNA chain and effect bypass (24-26). Although these enzymes are mutagenic in the sense that they often insert incorrect nucleotides into the DNA strand being synthesized, most do not cause the insertion of fewer or extra nucleotides. Thus, some of these may be suitable for repair of forensic DNA samples, where subsequent analysis of the DNA in current protocols is based on PCR fragment lengths rather than sequences (25-28).

For the purposes of this proposal, it is instructive to recognize that the problems DNA damage causes *in vivo* are not unlike those that occur during attempts to amplify degraded DNA evidence during STR genotyping. In both situations damaged DNA bases block the movement of a DNA polymerase. By removing the offending base prior to the polymerase reaching the site of damage *in vivo*, the cell avoids a potentially lethal circumstance. If damage to DNA evidence can be repaired *in vitro* prior to PCR, it should be possible to successfully genotype that sample. It is our hypothesis that proteins utilized for DNA double strand break repair *in vivo* by bacteria can be used as reagents to facilitate the forensic typing of heavily damaged human biological samples.

1.4 Bacterial DNA repair systems.

This work is a direct outgrowth of a collaborative effort of the Cox and Battista laboratories to investigate the DNA repair systems of the bacterium *Deinococcus radiodurans*. The effort has

also included the directed evolution of *Escherichia coli* strains that are highly resistant to ionizing radiation. The bacteria in question have an extraordinary capacity to repair the DNA damage inflicted by ionizing radiation (IR) levels that are thousands of times greater than the lethal dose for a human (29-31). The damage inflicted by high (kilogray) doses of IR include many, often hundreds or thousands of chromosomal double strand breaks. In our strains, this damage is repaired routinely, and the bacterial cultures are restored to growth with repaired chromosomes in the span of about 2-4 hours. In our forensic DNA repair protocols, we are drawing on the key enzymes involved in these impressive feats of DNA repair.

Deinococcus radiodurans R1 There are a small number of bacterial species that have the capacity to survive massive amounts of DNA damage. The best studied of these is *Deinococcus radiodurans* R1, a species of particular interest in the Cox and Battista laboratories (30). The *Deinococcaceae* are distinguished by their extraordinary ability to tolerate the lethal effects of DNA damaging agents, particularly those of ionizing radiation. Whereas an IR dose of 2 Gray (Gy; 1 Gy = 100 rads) is lethal for a human cell, *Deinococcus radiodurans* can survive a dose of 5,000 Gy with no lethality. Although the physiological basis of the deinococci's extreme radiotolerance has never been adequately explained, it is clear that irradiated cells are not passively protected from the damaging effects of the incident radiation. Instead, available evidence (such as that presented in Figure 2) argues that the deinococci do suffer massive DNA damage following irradiation, and that extensive DNA repair is necessary if these cells are to survive such exposures. To generate Figure 2, a *D. radiodurans* R1 culture, grown to exponential phase, was exposed to 3000Gy of ionizing radiation, a dose that introduces approximately 110 double strand breaks (dsbs) into the chromosome of each cell in the population. Beginning at 30 minute post-irradiation, aliquots of the culture were removed at various times, and pulsed field gel electrophoresis (PFGE)

was used to provide a visual record of the cell's recovery from the dsbs. Fragmentation of the chromosome is obvious immediately after irradiation (*lane 2*). The bands of chromosomal DNA present in unirradiated cell preparations (*lane 1*) are gone, replaced by a broad smear of lower molecular weight material (*lane 2*). Within three hours, however, this smear has disappeared and the chromosome has reformed (*lane 4*). Remarkably, *D. radiodurans* survives this degree of damage without loss of viability and without evidence of induced mutation. The D37 dose (the dose that on average is required to inactivate a single cell) for *D. radiodurans* R1 is approximately 6500Gy, at least 200 fold higher than the D37 dose of *E. coli* cultures irradiated under the same conditions (30, 31). The energy deposited by 6500Gy γ radiation should introduce approximately 200 DNA double strand breaks, over 3000 single strand breaks, and greater than 1000 sites of base damage per *D. radiodurans* genome (10). Despite this overwhelming genetic insult, *D. radiodurans* can reconstitute its genome in a manner that maintains the linear continuity of the parent organism, accomplishing this feat through an error-free process. The types of DNA damage generated following the irradiation of *D. radiodurans* are similar in type and quantity to what is found in ancient DNA and presumably similar to that observed as mammalian DNA is degraded after cell death in vivo (7, 11).

Escherichia coli MG1655 Most of our detailed understanding of the mechanisms of DNA repair in bacteria was obtained from investigations of *Escherichia coli*. *E. coli* became a model for study in large part because of its genetic tractability. The genome of this organism is routinely and easily manipulated: the genes that encode DNA repair proteins are known, their regulation is understood, and the repair proteins are well characterized. In the studies we plan we will be using enzymes from *Escherichia coli* MG1655 (32) and several derivatives we have recently generated. These derivatives are ionizing radiation resistant and were selected by subjecting cultures derived

from a single colony isolate to twenty-one iterative cycles of irradiation and outgrowth. The length of each exposure was adjusted to kill greater than 99% of the population, this dose increasing from 2000Gy for the first cycle to 10000Gy on the last cycle. These cultures exhibit between 1100 to 4500 fold increase in resistance to the cellular damaged caused by ionizing radiation. Like *Deinococcus*, they all have a greatly enhanced capacity to repair their genomes after high-level exposure to IR (33). We now know that the acquired phenotype of one of the most highly evolved isolates from these populations can be almost entirely explained by three mutations, occurring in the genes *recA*, *dnaB*, and *yfjK* (34). In essence, a phenotype of extreme resistance to ionizing radiation was acquired by this strain via adaptations to the cell's DNA repair systems. The contributing mutations in the *recA* gene are under intensive investigation in the Cox laboratory.

1.5 Repairing degraded forensic DNA samples

Forensic samples are often subject to extensive DNA damage, resulting from exposure to fire, explosion, ionizing radiation, or simply extended exposure to atmospheric oxygen. A major degradative agent is oxidative damage, which can arise due to free radicals generated by exposure to ionizing radiation (IR) or oxygen (12). A particularly problematic type of oxidative damage is a double strand break. Radiation exposure can also lead to the formation of inter- and intra-strand crosslinks or base alkylation (35). Independent of IR exposure, DNA can be damaged via post-mortem enzymatic attack, or by exposure to heat, humidity, bacteria, or other environmental factors (36). Hydrolysis of the DNA backbone can also result in single strand or double strand breaks (DSBs) (37). All of these types of damage can result in failure of short tandem repeat (STR) genotyping of these degraded samples by blocking the DNA polymerase used in polymerase chain reaction (PCR) amplification. This generally results in amplification patterns in which smaller loci

are successfully amplified, but larger loci, which are more likely to contain a damage site, cannot be successfully typed (38).

A variety of methods have been employed by the forensic community to deal with the DNA damage associated with degraded DNA samples. All of these methods seek to increase the signal intensity in capillary electrophoresis. Many of the methods are alterations to the amplification protocol to ameliorate the effects of the damage on PCR amplification (39, 40). These include reducing the size of STR amplicons by using miniSTRs (41), using tri-alleleic single nucleotide polymorphisms (SNPs) (42), or increasing the cycle number, polymerase concentration, or purification of the STR amplicons before injection (43, 44).

Other methods attempt to actually repair the DNA damage that prevents amplification of the STR loci by DNA polymerases. One method explored was a reaction designed to repair single strand breaks in a non-specific manner through nick translation using DNA polymerase I and T4 DNA ligase (45, 46). Others include the use of cocktails of repair enzymes, including glycosylases and endonucleases (47), Y-family DNA polymerases capable of synthesizing DNA past a lesion (48), and the commercially available PreCR (New England Biolabs) and Restorase (Sigma) repair mixes (35, 42). A common theme that runs through all of these repair attempts is that none are capable of repairing a double-strand break (DSB) in the DNA backbone.

A genomic DSB is potentially the most dangerous and difficult to repair of DNA lesions (13, 37, 49). DSBs may occur as a result of exposure to ultraviolet (50-53) or ionizing radiation (54-56), or to desiccation (57-60). *In vivo*, DSBs occur most often as a byproduct of normal replication in all organisms, particularly when replication forks encounter template strand breaks created in the course of other DNA repair processes (15, 19, 37, 61-65).

The key steps for repairing DSBs *in vivo* are illustrated in Figure 3. They include the processing of a double strand end to generate a single-stranded DNA segment that terminates in a free 3' end, the invasion of that 3' end into a homologous duplex DNA, and the extension of that now-paired 3' end by DNA replication. In *Escherichia coli*, the processing of DNA ends to generate an ssDNA segment is normally carried out by the RecBCD helicase/nuclease complex. DNA end-processing is not relevant to the current study. Subsequent steps are utilized in the method described here. The bacterial recombinase, RecA, loads onto the ssDNA to form a helical filament (66-68). The RecA nucleoprotein filament begins a search for homologous sequences, followed by the invasion of the RecA-coated ssDNA into the homologous dsDNA duplex. The 3' end of the ssDNA is then extended by replication by DNA polymerases (69).

The existence of robust DSB repair capabilities *in vivo*, centered in bacteria on the bacterial RecA protein, provides inspiration for the establishment of a cognate *in vitro* system to repair DSBs. This system can be applied to the repair of DSBs near STR loci, where these DSBs would normally prevent successful amplification and genotyping.

To this end, we have sought to determine the minimum enzymatic requirements for DSB repair *in vitro*, using a simplified plasmid system. We have developed robust in vitro procedures that allow us to degrade human DNA samples with ionizing radiation, generating a reproducible degradation of STR profiles. Based on published work (7, 11), the spectrum of DNA damage in these samples should reproduce that incurred by samples exposed to the environment fairly closely. A simple and robust system for double strand break repair has been developed that requires only the bacterial RecA protein, SSB, and a DNA polymerase. We then adapted the procedure to be used in the repair of our degraded human genomic DNA samples. This repair process has resulted

in the enhancement of signal strength, and in many instances in the accurate restoration of absent signals in degraded DNA at two STR loci: D18S51 and FGA.

II Materials and Methods

2.1 Proteins, DNA, and Reagents

The native *E. coli* wild-type RecA and single-stranded DNA binding (SSB) proteins were purified as described previously (70). The RecA E38K mutant protein was purified as described previously (71) with the following modifications. After washing the protein pellet with R buffer plus 2.1 M ammonium sulfate, the pellet was resuspended in R buffer plus 1 M ammonium sulfate. The sample was loaded onto a butyl sepharose column and washed with 2 column volumes of R buffer plus 1 M ammonium sulfate and eluted with a linear gradient from R buffer plus 1 M ammonium sulfate to R buffer over 10 column volumes. Peak fractions were identified by SDS-PAGE analysis and pooled. The protein was loaded onto a hydroxyapatite column as in the previous reference, but with the linear gradient from 20 – 500 mM P buffer. The fractions were dialyzed against R buffer and loaded on to a Source 15Q column. The column was washed with 2 column volumes of R buffer, and protein was eluted with a linear gradient from no salt to 1 M KCl over 10 column volumes. Peak fractions were identified as above, pooled, and dialyzed against R buffer. The pooled fractions were loaded onto a DEAE sepharose column, washed with 2 column volumes R buffer, and eluted with a linear gradient from R buffer to R buffer plus 1 M KCl over 10 column volumes. The peak fractions were identified and pooled as above, precipitated with ammonium sulfate (45% final concentration), and resuspended in R buffer plus 1 M KCl. The protein was loaded on an S-300 sizing column, washed with 1.5 column volumes R buffer plus 1 M KCl, and fractions were identified and pooled as above. The protein was precipitated with

ammonium sulfate as above, and resuspended in R buffer plus 1 M ammonium sulfate. The protein was loaded on a butyl sepharose column and eluted in a 15 column volume linear gradient to R buffer. The fractions were identified, pooled, and precipitated as above and dialyzed against R buffer. The protein was flash frozen in liquid N_2 and stored at –80°C. The concentration of the purified RecA and SSB proteins was determined from the absorbance at 280 nm using the extinction coefficients of 2.23 x 10^4 M^{-1} cm^{-1} (72) and 2.83 x 10^4 M^{-1} cm^{-1} respectively (73). All proteins were free from detectable nuclease activity.

DNA substrates were prepared as follows. Plasmid DNA (pUC19) was purified by banding in a CsCl gradient as previously described (74) and digested with either AatII or PstI (New England Biolabs, Ispwich, MA), creating substrates that have overlapping homologies of 504 bp and 2182 bp. The digested DNA was extracted with phenol/chloroform/isoamyl alcohol (25:24:1), ethanol precipitated, and resuspended in ddH_2O. Single stranded DNA (ssDNA) was generated by heat-denaturing the AatII digested pUC19 DNA at 100 °C for 10 min and quick-chilling in an ice water bath.

2.2 Double strand break repair of plasmid substrates

Reactions contained (concentrations reported as final): RecA buffer (25 mM Tris-Acetate 80% cation, 5% glycerol, 10 mM Magnesium Acetate (all from Fisher Scientific, Fair Lawn, NJ), and 3 mM potassium glutamate (Sigma, St. Louis, MO)--(final pH = 7.4), 2 mM ATP (Sigma), 1 mM DTT (Gold Biotechnology, St. Louis, MO), 5% additional glycerol, and an ATP-regeneration system of 10 U creatine phosphokinase and 12 mM phosphocreatine (both from Sigma). SSB protein (2.4 μM), RecA protein (4 μM), AatII-digested, heat-denatured pUC19 ssDNA (9 μM nt), and PstI digested pUC19 double-stranded DNA (dsDNA) (9 μM nt) was added on ice. The reaction

was started by incubating at 37 °C for varying times. DNA Polymerase I (1 U, New England Biolabs) was added with all four dNTPs (1 mM each, Promega, Madison, WI), and allowed to incubate at 37 °C for at least 30 minutes. Reactions were deproteinized with 1 μL of 20 mg/ml predigested proteinase K (Fisher) and incubated at 37 °C for 30 minutes, followed by incubation with 1% SDS (Fisher) at 37 °C for 30 minutes. Loading buffer (20 mM EDTA, 8.3% glycerol, 0.07% bromophenol blue, in water) was added to the reaction and loaded onto a 1% agarose/1X TAE gel for electrophoresis. DNA bands were separated overnight at 15 mA, the gel stained with SYBR Green I dye (Invitrogen, Carlsbad, CA), and the DNA was visualized using a Typhoon imager model 9410 (Amersham, Piscataway, NJ).

2.3 Degradation of human cell line DNA by ionizing radiation

HeLa and Jurkat cell line DNA (New England Biolabs) was diluted to 10 ng/μL using 10:1 TE buffer (1 mM Tris-Cl pH 8.0, 0.1 mM EDTA, both from MO-BIO, Carlsbad, CA). Aliquots of 200 μL were irradiated to 5-15 kGy in a Mark I ^{137}Cs irradiator (JL Shepherd, San Fernando, CA) using the smallest turntable to place samples closest to the radioactive source. The extent of degradation was determined by using 2.5-10 μL (25-100 ng) irradiated DNA as a template for PCR amplification according to the procedure in section *2.5*.

2.4 Double strand break repair of degraded human DNA

Reactions contained (concentrations reported as final): RecA buffer (as above), dATP (3 mM, Promega), DTT (1 mM), additional glycerol (5%), and an ATP-regeneration system of creatine phosphokinase (10 U), phosphocreatine (12 mM), and upstream and downstream biotinylated targeted oligonucleotides for either the FGA or D18S51 STR locus (15 nM, Table 1,

Integrated DNA Technologies, Coralville, IA) in a total reaction volume of 20 µL. The reaction was started upon the addition of RecA protein (1-2 µM) and incubated in a 37°C water bath for the duration of the repair reaction. After 3 minutes, SSB protein (0.8 µM) was added. After 3 minutes, 50 ng of degraded HeLa or Jurkat cell line DNA (5 µL of 10 ng/µL) was added. After 30 minutes, a mixture of DNA Polymerase I (1 U) and four dNTPs (2 mM each) was added and allowed to incubate at 37 °C for an additional 30 minutes. Reactions were stopped by bringing up to 150 µL with 10:1 TE buffer and phenol:chloroform:isoamyl alcohol (25:24:1) extracted with an equal volume, followed by a back extraction with 150 µL 10:1 TE buffer. The samples were ethanol precipitated for 1 hour at -20°C and resuspended in 20 µL 10:1 TE buffer for 1 hour at 37°C.

2.5 PCR amplification and capillary electrophoresis

DNA samples were amplified using the AmpFlSTR Identifiler PCR amplification kit (Applied Biosystems, Foster City, CA) according to the manufacturer instructions on a PTC-200 thermocycler (MJ Research, Waltham, MA). All pre-amplification DNA manipulations were done in a sterile laminar flow hood, and researchers wore gloves, masks, and lab coats to prevent cross-contamination of samples with exogenous DNA. To each amplification, 25 ng of unrepaired, damaged DNA (2.5 µL of 10 ng/µL stock) or DNA from repair reactions (10 µL of each repair) was added as a template for PCR. Amplified DNA was prepared for capillary electrophoresis and run as above. For capillary electrophoresis, 1.5 µL of each amplified sample was combined with 24.5 µL Hi-Di formamide and 0.5 µL GS-500 LIZ size standard (both from Applied Biosystems) in a 96 well optical plate. With each plate, one sample of Identifiler allelic ladder was included. Samples were denatured and snap-cooled before analysis on a 3730xl DNA Analyzer (Applied

Biosystems) according to Identifiler kit protocol at the University of Wisconsin-Madison Biotechnology Center. Results were viewed and analyzed using GeneMarker HID V2.2.0 (SoftGenetics, State College, PA).

2.6 Statistical analysis of results

Signal strength (in relative fluorescence units) at each allele with a measurable signal in the repaired DNA samples was divided by signal strength at the corresponding allele in the unrepaired DNA samples to obtain the % recovery value (RV). The RVs at each allele in a sample were averaged, and the target locus was considered repaired if the RV at that allele was at least two standard deviations above the average RV for the sample. For target loci that did not have a signal in any unrepaired samples of that cell line DNA at that irradiation value, the locus was considered repaired if any correct signal was seen in repaired samples at that locus.

III Results

3.1. RecA, SSB, and DNA Pol I can repair double-strand breaks in vitro using homologous single-stranded DNA. The search for a system suitable for double strand break repair of forensic DNA samples was initiated with a simple plasmid-based assay. The small plasmid pUC19 (2686 base pairs) was cleaved either with AatII or PstI to create two full length linear duplex DNAs, with end sequences that overlapped (Fig. 4A). One of these DNA substrates was denatured. RecA protein filaments were formed on the resulting single-stranded DNAs with the aid of the single-strand DNA binding protein (SSB). These complexes were used to promote DNA strand invasion, followed by addition of DNA polymerase to extend the 3′ ends created by the strand invasion. The

overall reaction was designed to generate two different products, resulting from the merger of the two different DNAs (Fig. 4A).

The repair reaction is illustrated in Figure 4B. A successful reaction depends upon the staged addition of the RecA, SSB, and DNA polymerase proteins, along with both AatII-cut ssDNA and PstI-cut dsDNA, and deoxynucleotide triphosphates (dNTPs). Bands corresponding to the predicted product sizes of 3.2 and 4.9 kbp are prominent in lane 3. These products are the result of a double-strand break (DSB) repair reaction that joins a PstI-cleaved DNA to a AatII-cleaved DNA. If either the ssDNA or dsDNA are eliminated, the predicted repair products are not formed (lanes 4 and 5, respectively). RecA will generate joint molecule intermediates (DNAs in which strand invasion has been completed, but no polymerization has occurred) during the repair process that migrate more slowly than the starting DNA substrates and form bands distinct in size from the repaired products. When RecA is omitted, there are no bands that migrate more slowly than the starting substrates, indicating no joint molecules or repaired products are present (lane 7). The presence of SSB stimulates product formation (compare lanes 3 and 6). Homologous DNA in the presence of RecA will form DNA networks with exchanged strands (75). Absent SSB, this is clearly seen as accumulation of DNA in the well of the agarose gel (Figure 4, lane 6). Furthermore, there is no ssDNA substrate band visible when SSB is omitted, suggesting that this DNA is complexed with the homologous dsDNA. The elimination of all four dNTPs or only dGTP, both of which prevent DNA synthesis by Pol I, does not result in repair products (lanes 8 and 9, respectively). RecA will still form joint molecule intermediates which are unable to extend (*). These results indicate that the repair of DSBs *in vitro* is dependent on the presence of the *E. coli* proteins RecA, SSB, and DNA Pol I, and ssDNA homologous to the dsDNA that contains a DSB.

To further demonstrate that repair of the DSB requires DNA synthesis, which is dependent on the activity of Pol I, time points were taken during the incubation of the DNAs, RecA, and SSB, and also after the addition of Pol I and dNTPs (Figure 5). A control reaction is shown in lane 3, in which the DNAs were incubated with SSB and RecA for 30 minutes at 37°C before the addition of dNTPs and Pol I, and then incubated for an additional 30 minutes. Time points were taken during the incubation of DNAs with RecA and SSB, showing that joint molecule intermediates accumulate over time (lanes 4-6). When Pol I and dNTPs are added to reactions containing these joint intermediates, DNA synthesis by Pol I occurs and bands corresponding to the predicted repair products accumulate over time (lanes 7-9). If only dGTP is eliminated, DNA synthesis by Pol I cannot occur and joint molecules remain that are not resolved to repair products. These results indicate that DNA synthesis by Pol I is necessary for the formation of the predicted products.

Additionally, this process requires the hydrolysis of ATP by RecA (data not shown). Taken together these results indicate that at a minimum, product formation is dependent on the activities of RecA, SSB, and DNA Pol I. Various regulators of the homologous recombination function of RecA (DinI (76), RecX (77), RecFOR (78), RdgC (79)) were assessed extensively for their ability to enhance the generation of repaired products *in vitro* under a range of conditions. None were found to have a significant positive effect (data not shown). The repair process was efficient even in the presence of large amounts of non-specific genomic DNA, which led to the adaptation of the method to repair DSBs in degraded human DNA for forensic genotyping. Several variants of RecA protein were utilized in these experiments and compared directly (data not shown). The most robust reactions occurred in the presence of the RecA E38K mutant protein (also called RecA730 (80-83)), which has been utilized in all of the trials reported here unless otherwise noted. Some

additional RecA variants have been utilized in very recent trials, as described at the end of the Results section.

3.2 The in vitro *DSB repair method can be adapted to repair DSBs in degraded human DNA samples.* Adaptation of the reaction described above to repair DSBs in degraded human DNA samples near targeted short tandem repeat (STR) loci is shown in Figure 6. Instead of generating single DNA strands via denaturation, targeting oligonucleotides (oligos) of 150-200 nt were designed that were identical to the sequences in the 3′-proximal strand immediately upstream and downstream of the D18S51 and FGA loci. These targeting oligos were designed to encompass the region between the amplification primer-binding site and one end of the STR. The targeting oligo is incubated with RecA and SSB, as the ssDNA is above. When degraded genomic dsDNA is added, strand invasion by the RecA-bound targeting oligo occurs at the location of any breaks adjacent to the targeted STR locus. As above, the addition of Pol I and dNTPs results in DNA synthesis across the DSB. The resulting product can again be PCR amplified using a standard commercial genotyping kit.

3.3 Human genomic DNA can be thoroughly degraded using ionizing radiation. Ionizing radiation (IR) from a ^{137}Cs source was used to generate stocks of degraded DNA for this study. HeLa and Jurkat cell line genomic DNA is readily available commercially, isolated from cell lines derived from a single human source. Stocks of the DNAs (10 ng/μL) were irradiated, and then 10 ng of each DNA was used as a template in the AmpFlSTR Identifiler PCR Amplification kit. Figure 7 shows the electropherogram from only the yellow color channel at each level of irradiation (indicated in kGray, where 1 Gray=100 rad). The first signals to drop out of the profile were those

generated from loci having the largest PCR product, followed by loci generating smaller PCR products as the radiation dose increases. In this particular sample, 5 kGy provided an IR dose sufficient to eliminate the signal at D18S51. After exposure to 15 kGy, no discernible signals are present at any of the loci. This pattern is repeated in the other three color channels (data not shown). The overall pattern is highly reproducible, although small variations in the extent of degradation were observed that arose from small differences in the positioning of samples in the irradiator. This figure demonstrates that IR is a useful tool for routine degradation of human genomic DNA samples for downstream repair process testing.

3.4 The in vitro *DSB repair method can be used to enhance or restore signals at the D18S51 STR locus in two different genomic DNA samples.* Targeting oligos were designed for the D18S51 locus that were homologous to the region upstream and downstream of the STR (Table 1). For all targeting oligos, the 3' end was located proximal to the STR.

The process outlined in section *3.2* was used to repair DSBs upstream and downstream of the D18S51 locus in degraded human DNA from two different cell lines (Figure 8). The top row shows electropherogram readouts from the D18S51 locus on two degraded genomic DNA samples after amplifying with the Identifiler kit.

The HeLa cell line is homozygous at the D18S51 locus, and the correct genotype is 16 repeats. Prior to repair, the correct allele at D18S51 was seen in HeLa DNA dosed with 5 kGy IR. However, at 43 RFU, it was below the analytical threshold (AT) set by the Orange County Crime Laboratory (OCCL) of 100 RFU (left column, top). After a mock repair process without targeting oligos, the allele was no longer seen due to a loss of DNA in the repair and subsequent recovery process (left column, middle). Following a full repair process using 2 μM RecA E38K, the signal

increased substantially past the OCCL AT to 218 RFU, a 411% increase from the signal before repair (left column, bottom).

Prior to repair, no signals were seen in HeLa DNA dosed with 7.5 kGy IR (Fig. 8, middle column, top). After a mock repair process without targeting oligos, the allele was again no longer seen due to a loss of DNA in the repair and subsequent recovery process (right column, middle). Following a full repair process using 1.33 μM RecA E38K, a visible signal was restored at 45 RFU (right column, bottom). In over a dozen successful trials in which signals increased substantially at D18S51, the correct allele was observed.

This process could also be used to repair DSBs flanking the D18S51 locus in DNA from the Jurkat cell line (Fig. 8, right column). The correct genotype at D18S51 for the Jurkat cell line is 13 and 21 repeats. Prior to repair, the 13 allele was seen in Jurkat DNA dosed with 7.5 kGy IR, but at 30 RFU, below the OCCL AT of 100 RFU (data not shown). The 21 allele was not present. After a mock repair process without targeting oligos, the allele was again no longer seen due to a loss of DNA in the repair and subsequent recovery process (right column, middle). Following a full repair process using 1 μM RecA E38K, the signal at the 13 allele increased to 99 RFU. The correct 21 allele peak was restored to 42 RFU. This work demonstrates that the repair process is capable of increasing low allele peak heights and restoring correct signals at the D18S51 locus in a targeting oligo-dependent manner. The correct signals at the D18S51 locus were increased or restored in these two degraded DNA samples in 39 trials carried out over the course of an extended screening effort encompassing over 200 trials in which conditions and reagents were varied. Fourteen of the successful repairs were achieved under the conditions described in the Methods section, which currently represent the best practice. Success under those best practice conditions occurred at a rate of approximately 25-30%.

Artifactual signals were not observed in any trials at the D1851 locus, utilizing these targeting oligonucleotides. However, artifacts did appear at other loci in some trials in a manner that depended only upon the presence of the targeting oligonucleotides. The DSB repair process of the D18S51 locus increased the peak height at the D18S51 locus even when an overall decrease in peak height at other loci was observed (Figure 9). The upper panel shows HeLa cell line DNA dosed with 5 kGy of IR, which does not produce an allelic peak at the D18S51 locus in this irradiated DNA sample (this is a different sample than the one used in Fig. 8), as demonstrated by the absence of black peaks between 260 and 350 bp. The average peak height across all loci was 841 RFU. The bottom panel shows the same DNA following a repair process using 1.33 μM RecA E38K. The repaired allele 16 in the D18S51 locus (denoted with an asterisk) was 144 RFU. This restoration was achieved even though the average peak height across all loci was reduced to 455 RFU, indicating a 49% recovery of the sample DNA. The overall decrease in signal strength across the entire profile highlights the relative increase in signal strength at the targeted D18S51 locus. As seen in this example, artifacts were generated in some D18S51 repair trials in the D19S433 and vWA loci (shown in this example and denoted with arrows, bottom panel). Although their specific origin is not entirely understood, they were consistently positioned in the yellow color channel across multiple DNA samples and trials when they were present. For D18S51, the artifacts were uniquely dependent on the presence of the targeting oligos, and were generated in some control samples in which neither repair proteins or degraded human DNA was added. It is important to note that the presence of these artifacts does not interfere with interpretation of the genotype at the D18S51 locus targeted for repair. However, optimal design of targeting oligonucleotides is an issue to be addressed in ongoing studies.

3.5 The in vitro *DSB repair method can be used to enhance or restore signals at the FGA STR locus in two different genomic DNA samples.* To demonstrate that successful repair was not unique to the D18S51 locus, and was capable of repairing other loci, we designed targeting oligos for the FGA locus that are homologous to the DNA sequence upstream and downstream of the STR (Table 1). The process outlined in section *3.2* was used to repair DSBs upstream and downstream of the FGA locus in degraded human DNA from two different cell lines (Figure 10). The top row shows electropherogram readouts from the FGA locus on two degraded genomic DNA samples after amplifying with the Identifiler kit.

The correct genotype for the HeLa cell line at the FGA locus is 18 and 21 repeats. Prior to repair, the correct alleles at FGA were seen in HeLa DNA dosed with 5 kGy IR, but the 18 allele is at 41 RFU, below the AT, and the 21 allele is at 108 RFU, barely above the AT (left column, top). After a mock repair process without targeting oligos, both allele peak heights decreased, due to loss of DNA during the repair and recovery process (left column, middle). Following a full repair process with 1.33 μM RecA E38K, the signal increased substantially past the AT to 310 and 406 RFU for the 18 and 21 alleles, respectively (left column, bottom). This represented a 756% increase at the 18 allele, and a 376% increase at the 21 allele. This demonstrates that the *in vitro* repair process is capable of increasing signal strength at more than one STR locus by simply changing the sequence of the targeting oligo used, and that the repair is not dependent on any particular feature of the D18S51 locus.

In some trials, the *in vitro* repair process created artifactual signals within the FGA locus. These were consistently positioned when they appeared, and their appearance spanned multiple repair attempts at this locus. The artifacts were again uniquely dependent on this particular set of targeting oligos. An example of the typical artifacts can be seen in the second example presented

in Figure 10. Prior to repair, no signals were seen at the FGA locus in Jurkat DNA dosed with 7.5 kGy IR (right column, top). The correct genotype at the FGA locus is 20 and 21 repeats when using Jurkat DNA that has not been dosed with IR (right column, middle). Following a full repair process using 1 μM RecA E38K, the correct alleles were restored with signals of 40 RFU each, although both were below the AT (middle column, bottom). However, this trial illustrates the artifacts that sometimes appeared around the location of allele 24. This artifact was present in some mock repair reactions that contained the FGA upstream and downstream targeting oligos, but did not contain any degraded or undegraded human genomic DNA or repair enzymes (data not shown). This demonstrates that the repair process is capable of restoring correct signals at the FGA locus in multiple degraded human DNA samples, but can create repeatable artifacts that are not dependent on the DNA source. Of 25 successful repairs at FGA that were achieved in our screening trials, approximately 80% included the artifacts illustrated in the right panels of Figure 10. The artifacts arise due to reactions involving the targeting oligos, and efforts are underway to test re-designed targeting oligos. Another approach to eliminating these artifacts may involve the use of targeting oligos that can be conveniently removed from the reaction when the repair process is completed. An investigation of this latter approach is currently underway.

3.6 Recent trials suggest promise for some new RecA protein variants. In the period July – Septemnber, 2013, approximately 60 additional trials were carried out. A number of these focused on the use of new RecA variants with enhance recombinase activities.

Over the past two years, a broad screen has been carried out in the Cox laboratory to identify RecA variants with these enhanced functions. The *recA* gene has been systematically mutated in the region encompassing residues 78-135. From the library of mutants created in this

region, variants with enhanced function were selected using a conjugation-based screen. Approximately 6 variants with greatly enhanced activity have been generated. These, as well as a range of RecA variants previously generated in this laboratory, were tested in the new trials. The new variants are being prepared for a patent submission, and here the two best ones will be labeled simply RecA X and RecA Y. These two variants have shown promise in a few initial trials carried out on DNA that was sufficiently degraded that no STR signals were present at loci D18S51 and FGA in a typical trial, and the result is thus preliminary. Both trials targeted the FGA locus.

Our current lot of HeLa cell line DNA that has been irradiated to 7.5 kGy does not have signals at the FGA locus in 6/7 independent trials. (The 7th trial showed the correct 18 allele only at 61 RFU). In two independent attempts to repair the FGA locus in this DNA using the RecA variant RecA X, signals appeared at allele 18 (54 RFU) and 21 (97 RFU) in one attempt, and allele 21 (62 RFU) in another (Figure 11). Further repeats will be required to ensure that this is an actual repair and not the result of normal variation in the PCR process.

Our current lot of Jurkat cell line DNA that has been irradiated to 7.5 kGy does not have signals in the FGA locus in 6/6 independent trials. In an attempt to repair the FGA locus using the RecA X variant, a signal at allele 20 (51 RFU) resulted. In an attempt to repair the FGA locus using the RecA Y variant, a signal at allele 20 (52 RFU) was also seen (Figure 12). There were no signals in the FGA locus when mock repairs were carried out without repair oligos. These are likely true repairs. However, many repeats are needed, as well as efforts to optimize conditions for these RecA variants.

We have initiated one other type of trial. In this experiment, we attempt to eliminate the STR signal at a particular locus in otherwise undamaged DNA, and then restore that signal with our repair reaction. This would in principle provide a proof of principle in DNA that did not have

a background of other types of DNA lesions present. To remove a particular signal, we utilize the RecA-dependent nuclease Ref (84). Ref, encoded by bacteriophage P1, is an HNH class endonuclease. It has the novel property that it will only cleave DNA to which RecA protein is bound. When RecA is bound to an oligonucleotide, and promotes strand invasion to form a D-loop, both strands of the target DNA are cleaved within the D-loop (84). The scheme in the current experiment is to first use the same targeting oligos used in the repair reaction to create double strand breaks on both sides of the D18S51 locus to reduce or eliminate the STR signal from just that locus. Then, using the same targeting oligonucleotides in repair mode, restore the lost signal. The first trial of this experiment is shown in Figure 13. The signal from locus D18S51 was reduced by about 50% in this trial, while the signal from FGA (shown) and all other loci (not shown) were not affected. The D18S51 signal was restored by the repair reaction. We still need to optimize this experiment so that the D18S51 signal is more thoroughly removed in the Ref/RecA stage before restoration.

3.7 Parallel effort to isolate DNA segments from bacterial and human DNA samples.

Sub-project goal. The ability to evaluate the success of the in vitro restitution protocols depends on our capacity to recover and amplify STR loci from the reaction mixtures. At the outset of these studies, we anticipated the need for an efficient system of recovering restored fragments from sample reactions – a method that would concentrate the restored fragments and simultaneously remove them from conditions that might interfere with the PCR reaction and detection. As originally envisioned, the approach uses a biotinylated oligonucleotide designed to

anneal to a conserved sequence immediately adjacent to sites in the human genome currently used to amplify the STR. Once this biotinylated oligonucleotide has annealed to its target, that oligonucleotide would be extended using a DNA polymerase. The resulting duplex and any associated DNA would be removed from solution using streptavidin-coated magnetic beads. Since the recovered genomic DNA contains the STR, it could be amplified and visualized by standard STR typing protocols. The effort would directly complement and potentially enhance the procedures described above, in that the targeting oligonucleotide would simply be biotinylated, allowing it to be extracted from the sample along with any newly repaired DNA that was contiguous with it. The effort to develop this recovery protocol was carried out at LSU, in parallel to the main repair effort described above.

Proof of Principle**.** A series of studies were implemented to determine if the proposed protocol could efficiently retrieve and concentrate a DNA fragment from isolated genomic DNA. Initially, the protocol was tested and refined with a readily available experimental system, using DNA isolated from *Escherichia coli* MG1655. The genome of strain MG1655 was chosen because it is fully sequenced, and the DNA is easily and inexpensively obtained. A biotinylated primer was designed to anneal to the sequence upstream (5′) of the MG1655 *gltS* gene (length 1206bp), which was randomly chosen as the target in these initial trials.

For the preliminary studies, the length of the biotinylated capture primer was varied from 18-200 bases long, but it was quickly determined that the length of the primer did not affect the efficiency of recovery. Sixteen picomoles of biotinylated primer was added to each reaction, a tenfold molar excess of primer relative to the highest concentration of target added to the reaction. The biotinylated primers were annealed at 58˚C. After annealing, repeated bouts of

DNA synthesis (referred to as linear amplification) were performed in a thermocycler. This step extends from the 3′ end of the biotinylated capture primer, making a complementary copy of the *gltS* gene and presumably some DNA sequence downstream of the *gltS* gene. Isolated MG1655 genomic DNA (between 1 and 500ng of material per reaction) was added to a mixture of Accu*Taq*™ DNA polymerase (2.5U), extreme thermostable single-stranded binding protein (ET SSB, 200ng), and deoxyribonucleotides (dNTPs, 200µM). Linear amplification was performed using the following conditions: (a) 95˚C for three minutes, (b) 95˚C for forty-five seconds, (c) 58^{o}C, the annealing temperature of the primer, for thirty seconds plus an additional fifteen seconds added to every cycle, and (d) 72˚C for one and a half minutes. This protocol was repeated for sixty cycles. After amplification, the reaction mixture was added to 150µg of Dynal M-280 streptavidin magnetic beads, prepared according to the manufacturer's protocol (KilobaseBINDER™ kit, Invitrogen). To elute the biotinylated product, deionized water was added and heated to 95˚C for three minutes. The resulting mixture was placed on a magnet to remove the biotinylated product. Heat breaks the streptavidin biotin bond and releases the biotinylated product into the supernatant. A 'no polymerase' control was included for each experiment and was subjected to the protocol described above except the Accu*Taq* DNA polymerase was not added to the reaction mixture. This control assisted in establishing the extent of non-specific binding during capture of the target.

The number of *gltS* sequences present in the genomic DNA were determined using quantitative polymerase chain reactions (qPCR). Upstream and downstream amplimers, designated gltS1 and glts2, were designed to be specific for the *gltS* gene 250bp downstream of the site complementary to the biotinylated primer used to capture *gltS*. Primers glts1 and glts2 allowed the accurate estimation of the number of *gltS* sequences in the sample prior to attempts

at retrieval and after retrieval. Sybr-Green master mix (Applied Biosystems) was used in conjunction with an ABI 7000 Sequence Detection System to establish a standard curve and quantify amounts of sequence recovered following capture. Since there was a possibility that genomic DNA could carry over into the qPCR reaction by non-specifically binding to the hydrophobic magnetic beads, primers were designed to amplify the 16S rDNA gene of MG1655. MG1655 encodes seven 16S rRNA genes. This locus was chosen to maximize the probability of detecting non-specific carryover. We assumed that carryover of single copy genes within the genome would be less than what was observed with the 16S gene. Amplimers were also designed to amplify a sequence 5000 bases downstream (gltsdwn1 and gltsdwn2) from the biotinylated primer binding site. These primers allowed an estimate of recovered fragment size.

Table 2 records the numbers of target molecule detected before and after the protocol at the two downstream sites used for detection. The number of copies of the *gltS* gene, as measured by QPCR, in each reaction is reported in the first column. The numbers in parentheses refer to the number of nanograms of genomic DNA added to each reaction. The second column reports the numbers of copies of *gltS* captured using the protocol. Two different sets of numbers are recorded, corresponding to the sites used from QPCR measurements, 250 and 5000bp downstream from where the capture primer annealed. Captured fragments were detected at both sites at all DNA concentrations evaluated, but analysis did not accurately enumerate the molecules found in the samples, routinely underestimating the number of copies present in the sample. However, the values obtained from the 250bp site faithfully reproduced the linear decrease in concentration over the range evaluated. Detection at the site 5000bp from where the capture primer annealed did not exhibit a concentration dependence. Instead the number of molecules of this size detected was the same at all but the 0.1ng concentration, suggesting that

factors unrelated to the efficiency of annealing are responsible for the differences observed in the different sites used for detection. We have not pursued studies to define the reasons for this difference, but suspect that it is related to differences in the efficiency of polymerase extension. The third column records the non-specific carryover of the 16S rRNA gene at each concentration of genomic DNA tested. The protocol resulted in detection of approximately 60, 000 copies of the 16S rRNA gene when 500ng of genomic DNA was added to the reaction and 15,000 copies when 0.1ng was added, indicating that carryover occurs but that it does account for the results of the capture reaction the concentrations of genomic DNA evaluated in these studies.

One hundred nanograms of DNA was subjected to capture as described with the exception that the DNA polymerase used for linear amplification was not added to the reaction. The *gltS* copy number was then evaluated from the captured material using the site 250bp within the *gltS* gene. Only $0.047 \pm 0.01 \times 10^6$ copies were detected without linear amplification, comparable with the amount of carryover detected using the 16S rRNA gene. Clearly, linear amplification is necessary for successful capture of the target sequences.

Trials with Human Genomic DNA.

Selective capture of D18S51 locus from HeLA cell DNA**.** To date, we have used capture fragments to recover the D18S51 locus from HeLa cell DNA. The purified DNA was diluted to provide D18S51 concentrations of 150,000 copies per □L, 15,000 copies per □L, and 1500 copies per □L. After capture, quantitative PCR was used to evaluate how effective capture was at recovering the target. Values obtained after capture were 135,000 +/- 4000, 12500 +/- 650, and 1100 +/- 80, respectively (n=9, three independent studies, three replicates per study). Recovery was approximately 10% less efficient at the lower concentration relative to the highest, but the fragments were readily detected even at low copy number.

Carryover of genomic DNA, which would provide a positive signal without capture, was monitored by using quantitative PCR targeted to a region 50 bp upstream of the site that the capture primer anneals. This site should be excluded in the capture reaction. A carryover signal of 200 +/- 80 was detected for the 150,000 copies per □L sample, indicating 0.13% carryover in this sample. The signals for the more dilute samples were below the level of detection.

Use of a biotinylated primer to capture specific mitochondrial DNA sequences from a mixture of genomic and mitochondrial DNA. An 18 base biotinylated primer was designed that anneals to the conserved regions of the *cytB* gene in the human mitochondrial genome ending at position 14,724. Human DNA and mitochondrial DNA were extracted from whole blood using Qiagen Genomic 100/G kit with a final concentration of 189ng/mL. In order to determine the concentration of mitochondrial DNA within the extracted sample, qPCR was performed using primers that amplified the region between 10,360 – 10,648bp. The total DNA isolated contained 1.33 ng/mL mitochondrial DNA based on ths measurement. The biotinylated primer was used to linearly increase the amount of the hypervariable region sequence in the sample. Each 50 □l reaction mixture contained four picogram of mitochondrial DNA, Accu*Taq*™ DNA polymerase (2.5U), extreme thermostable single-stranded binding protein (ET SSB, 200ng), biotinylated primer (2μM), and deoxyribonucleotides (dNTPs, 200μM). The reaction mixture was subjected to sixty cycles of amplification as follows: 95˚C for three minutes, 58˚C for one minute to anneal the primer, and 72˚C for one and a half minutes plus an additional 15 seconds for each successive cycle. This protocol increases the amount of the hypervariable sequences by approximately 60 fold. After linear amplification, the biotinylated product is recovered using Dynal M-280 streptavidin-coated magnetic beads. Recovered sequences are washed per manufacturer's directions, and eluted in water heated to 95°C for three minutes. Magnetic beads

are removed, and the released biotinylated product is amplified by PCR using the 15996 forward primer (0.4µM) and the 151 reverse primer (0.4µM), Accu*Taq*™ DNA polymerase (2.5U), extreme thermostable single-stranded binding protein (ET SSB, 200ng), and deoxyribonucleotides (dNTPs, 200µM). Amplification was performed in a thermocycler programmed as follows: 95˚C for three minutes, 50°C for forty-five seconds, (d) 72˚C for forty-five seconds for thirty cycles, with a final extension at 72°C for five minutes. The PCR reaction was passed through a MP Bio GeneClean Spin Kit, and the eluted fragments subjected to automated Sanger sequencing using the BigDye v3.1 kit.

Carryover of mitochondrial DNA by non-specific binding to the magnetic beads was evaluated by adding beads to the reaction mixture containing all components except the biotinylated primer. Beads were recovered and attempts were made to amplify the hypervariable regions. In the three trials completed, we were unable to amplify these sequences indicating that carryover was not responsible for the sequences obtained.

The protocol was independently repeated three times. In each circumstance, the same sequence for human mitochondrial DNA hypervariable regions I and II were obtained. Although the hypervariable regions were identical in each replicate, the region of high GC content between the two hypervariable regions exhibited different amounts of GC compression.

IV Discussion

We have developed a method for the repair of DSBs flanking the STR loci that are used for forensic DNA genotyping. Our trials so far have targeted specifically the D18S51 and FGA STR loci. The repair reactions enhance fluorescent signal strength of the correct allelic peak(s) at the targeted loci (Figures 8 – 10). This system is simple in setup, and requires only the *E. coli* proteins RecA, SSB, and DNA Polymerase I. None of these proteins are capable of further damaging the DNA, as long as they are utilized as nuclease-free preparations. The protocol can be used prior to the use of standard commercial forensic DNA typing kits, requiring no significant changes in established protocols. This process was successful using a variety of different DNA sources dosed with varying levels of ionizing radiation. Although other methods exist to enhance signal strength from degraded DNA or even to repair some types of DNA damage, this is the first reported method to repair DSBs *in vitro* in a forensic context. We intend this protocol to be used in cases where degradation of the sample results in an incomplete profile due to drop-out at STR loci that generate larger PCR products. We note that the potential for successful repair declines in samples where allele dropout affects even those loci that generate relatively short PCR products. We thus anticipate that it may be useful primarily to restore complete or more complete and normal profiles where signals from substantial numbers of loci that produce larger PCR products are absent.

We began by developing this method using linear dsDNA with a piece of homologous ssDNA (Figure 4). RecA catalyzes synapsis, the pairing and exchange of homologous DNA strands (69), and this activity is absolutely required under the conditions used, as evidenced by the absence of reaction intermediates observed in its absence (Figure 1B, lane 7). Given that the stability of a RecA filament may be altered following the binding of DNA Pol I, a variety of

proteins that regulate RecA *in vivo* were tested using the plasmid-based repair system. These included the RecFOR, RecX, DinI, and RdgC proteins. None of the regulators produced any substantial positive effect on the amount of repair product formed, and all activities were consistent with previously published activities of the regulators (data not shown). Although our method is not the first *in vitro* effort to recapitulate recombination-dependent replication, ours is unique in that it uses no nucleases, gyrases, helicases, or ligases, and is not dependent on the presence of DNA supercoiling (75, 85, 86).

We decided to target regions between the amplification primer-binding site and the STR on both sides for repair. It is important to note that the method does not target the STR itself for repair through genetic recombination, as this could alter the number of repeats in the STR and produce incorrect genotyping results. Using this method, the same protein components were used as in the plasmid-based system, with the exception of the RecA variant E38K. This variant has been shown to have enhanced DNA pairing and strand exchange activities, two functions of RecA on which this method relies (87).

We chose ionizing radiation (IR) as the method for degrading human genomic DNA due to its ability to thoroughly degrade the DNA in a relatively consistent manner that approximates the type of damage seen in forensic contexts. IR is capable of creating a variety of types of DNA damage including oxidative damage, base alkylation, intra- and inter-strand crosslinks, and single-strand and double-strand breaks (88). Among the samples of DNA that we irradiated, there was some variation in the effectiveness of the DSB repair method. This is likely due to the fact that irradiation produces all of these types of DNA damage in varying ratios, and our method is designed only to repair double-strand breaks. PCR inhibition resulting from other types of DNA damage will not be affected. The PreCR repair kit (New England Biolabs, Ipswich, MA) has been

demonstrated to also increase signal strength and callable allele counts, but does not address DSBs (35). The difference in signal strength seen at the targeted loci in our samples before and after the repair process is a clear indication that DSBs present a substantial hurdle in the successful amplification of STR loci (Figures 6 and 8).

The use of the *in vitro* DSB repair process to repair DSBs flanking the D18S51 and FGA loci was successful on genomic DNA from two different sources that was dosed with two different levels of IR (Figures 8 and 10, compare top and bottom rows). Depending on the level of degradation, the method is able to either increase signal strength of the correct allele, or to restore a signal at the correct allele, by repairing DSBs between the PCR primer binding site and the STR. One caveat to this method is that the current protocol uses approximately 50 ng of genomic DNA, an amount that is greater than that present in many forensic samples. Using our current best practice protocol, repair is successful approximately 20-30% of the time. We are working to improve DNA recovery after repair and develop a consistent and reliable method that uses smaller amounts of genomic DNA. We are continuing to investigate the optimization of the concentration and order of addition of all reaction components. We are also now employing improved RecA variants, which are showing some promise in the very preliminary trials so far (Figures 11 and 12).

We have demonstrated that in certain instances, artifact signals appear in the profile (Figure 9, bottom, and Figure 10, bottom right). Artifacts that result from the D18S51 repair are all in the yellow color channel, and tend to be reproducibly positioned (Figure 9, bottom, arrows), as are the FGA artifacts. In both cases, the artifacts are dependent on the presence of the repair oligo and do not arise as a result of the repair process itself, as they appear even when PCR amplification using the Identifiler kit is attempted on the DNA recovered from a mock repair that contains targeting

oligo, but lacks genomic DNA and repair proteins. This result highlights the need for further study of the design of the targeting oligos.

An opportunity exists to enhance the double strand break repair protocol by marrying it with the DNA capture method developed at LSU. This protocol is relatively simple and inexpensive. It is also quite effective in all trials carried out to date with bacterial, human STR, and mitochondrial DNA segments. By biotinylating our targeting oligonucleotides, we may be able to enhance our overall target STR signal recovery.

V Conclusion

Degraded DNA samples are a significant hurdle for the forensic DNA typing community. Methods to increase signal strength and the number of callable alleles include procedural changes, such as miniSTRs, or actual repair of the damage, seen here and in other studies. Our method is the first that is capable of physically repairing DSBs in damaged DNA, thereby decreasing PCR inhibition and increasing the amount of reliable STR loci in the profile. Furthermore, it requires no adaptations to current DNA analysis protocols, and can simply be inserted after the first conventional amplification fails to gain additional information about missing alleles or those below the laboratory's AT.

Addendum: We plan to file an update on this report on June 30, 2014.

VI References

1. **Butler, J. M.** 2006. Forensic DNA Typing, 2 ed, vol. Academic Press, New York.
2. **Jin, L., Underhill, P. A., Buoncristiani, M., and Robertson, J. M.** 1997. Defining microsatellite alleles by genotyping global indigenous human populations and non-human primates. J Forensic Sci **42:**496-9.
3. **Jeffreys, A. J., Wilson, V., and Thein, S. L.** 1985. Hypervariable minisatellite regions in human DNA. Nature **314:**67-73.
4. **Hares, D. R.** 2012. Addendum to expanding the CODIS core loci in the United States. Forensic Science International-Genetics **6:**E135-E135.
5. **Hares, D. R.** 2012. Expanding the CODIS core loci in the United States. Forensic Science International-Genetics **6:**E52-E54.
6. **Hochmeister, M. N.** 1998. PCR analysis of DNA from fresh and decomposed bodies and skeletal remains in medicolegal death investigations. Methods Mol Biol **98:**19-26.
7. **Gilbert, M. T. P., and Willerslev, E.** 2007. Rescuing ancient DNA. Nature Biotechnology **25:**872-874.
8. **Pfeiffer, H., Huhne, J., Seitz, B., and Brinkmann, B.** 1999. Influence of soil storage and exposure period on DNA recovery from teeth. Int J Legal Med **112:**142-4.
9. **Wellinger, R. E., and Thoma, F.** 1996. Taq DNA polymerase blockage at pyrimidine dimers. Nucleic Acids Res **24:**1578-9.
10. **Paabo, S., Higuchi, R. G., and Wilson, A. C.** 1989. Ancient DNA and the polymerase chain reaction. The emerging field of molecular archaeology. J Biol Chem **264:**9709-12.
11. **Paabo, S.** 1989. Ancient DNA - extraction, characterization, molecular-cloning, and enzymatic amplification. Proceedings of the National Academy of Sciences of the United States of America **86:**1939-1943.
12. **Lindahl, T.** 1993. Instability and decay of the primary structure of DNA. Nature **362:**709-15.
13. **Friedberg, E. C., Walker, G. C., Siede, W., Wood, R. D., Schultz, R. A., and Ellenberger, T.** 2005. DNA Repair and Mutagenesis, vol. ASM Press.
14. **Battista, J. R., and Earl, A. M.** 2004. Mutagenesis and DNA repair: The consequences of error and mechanisms for remaining the same. Microbial Evolution: Gene Establishment, Survival, And Exchange. Pg.
15. **Cox, M. M., Goodman, M. F., Kreuzer, K. N., Sherratt, D. J., Sandler, S. J., and Marians, K. J.** 2000. The importance of repairing stalled replication forks. Nature **404:**37-41.
16. **Cox, M. M.** 2001. Historical overview: Searching for replication help in all of the rec places. Proceedings of the National Academy of Sciences of the United States of America **98:**8173-8180.
17. **Cox, M. M.** 2002. The nonmutagenic repair of broken replication forks via recombination. Mutation Research-Fundamental and Molecular Mechanisms of Mutagenesis **510:**107-120.

18. **Michel, B., Grompone, G., Flores, M. J., and Bidnenko, V.** 2004. Multiple pathways process stalled replication forks. Proceedings of the National Academy of Sciences of the United States of America **101:**12783-12788.
19. **Michel, B., Boubakri, H., Baharoglu, Z., LeMasson, M., and Lestini, R.** 2007. Recombination proteins and rescue of arrested replication forks. DNA Repair **6:**967-980.
20. **Goodman, M. F.** 2002. Error-prone repair DNA polymerases in prokaryotes and eukaryotes. Annual Review of Biochemistry **71:**17-50.
21. **Fijalkowska, I. J., Schaaper, R. M., and Jonczyk, P.** 2012. DNA replication fidelity in Escherichia coli: a multi-DNA polymerase affair. Fems Microbiology Reviews **36:**1105-1121.
22. **Hubscher, U., and Maga, G.** 2011. DNA replication and repair bypass machines. Current Opinion in Chemical Biology **15:**627-635.
23. **Klarer, A. C., and McGregor, W. G.** 2011. Replication of Damaged Genomes. Critical Reviews in Eukaryotic Gene Expression **21:**323-336.
24. **Lopes, M., Foiani, M., and Sogo, J. M.** 2006. Multiple mechanisms control chromosome integrity after replication fork uncoupling and restart at irreparable UV lesions. Mol Cell **21:**15-27.
25. **Rattray, A. J., and Strathern, J. N.** 2003. Error-prone DNA polymerases: when making a mistake is the only way to get ahead. Annu Rev Genet **37:**31-66.
26. **Woodgate, R.** 1999. A plethora of lesion-replicating DNA polymerases. Genes & Development. **13:**2191-2195.
27. **Jiang, Q., Karata, K., Woodgate, R., Cox, M. M., and Goodman, M. F.** 2009. The active form of DNA polymerase V is UmuD'2C•RecA•ATP. Nature **460:**359-363.
28. **Schlacher, K., Cox, M. M., Woodgate, R., and Goodman, M. F.** 2006. RecA acts in trans to allow replication of damaged DNA by DNA polymerase V. Nature **442:**883-887.
29. **ACADEMIES, N. R. C. O. T. N.** 2006. Health Risks from Exposure to Low Levels of Ionizing Radiation: BEIR VII Phase 2. The National Academies Press, Washington D.C.
30. **Cox, M. M., and Battista, J. R.** 2005. Deinococcus radiodurans - The consummate survivor. Nature Reviews Microbiology **3:**882-892.
31. **Blasius, M., Hubscher, U., and Sommer, S.** 2008. Deinococcus radiodurans: What belongs to the survival kit? Critical Reviews in Biochemistry and Molecular Biology **43:**221-238.
32. **Blattner, F. R., Plunkett, G. r., Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado, V. J., Glasner, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B., and Shao, Y.** 1997. The complete genome sequence of Escherichia coli K-12. Science **277:**1453-74.
33. **Harris, D. R., Pollock, S. V., Wood, E. A., Goiffon, R. J., Klingele, A. J., Cabot, E. L., Schackwitz, W., Martin, J., Eggington, J., Durfee, T. J., Middle, C. M., Norton, J. E., Popelars, M., Li, H., Klugman, S. A., Hamilton, L. L., Bane, L. B., Pennacchio, L., Albert, T. J., Perna, N. T., Cox, M. M., and Battista, J. R.** 2009. Directed evolution of radiation resistance in Escherichia coli. Journal of Bacteriology **191:**5240-5252.
34. **Byrne, R. T., Klingele, A. J., Wood, E. A., Cabot, E. L., Schackwitz, W., Martin, J., Martin, J., Wang, Z., Pennacchio, C., Pennacchio, L. A., Perna, N. T., Battista, J. R.,**

and Cox, M. M. 2013. Evolution of extreme resistance to ionizing radiation via genetic adaptation of DNA repair
. eLife **submitted**.
35. **Diegoli, T. M., Farr, M., Cromartie, C., Coble, M. D., and Bille, T. W.** 2012. An optimized protocol for forensic application of the PreCR Repair Mix to multiplex STR amplification of UV-damaged DNA. Forensic Sci Int Genet **6:**498-503.
36. **Bar, W., Kratzer, A., Machler, M., and Schmid, W.** 1988. Postmortem stability of DNA. Forensic Sci Int **39:**59-70.
37. **Cahill, D., Connor, B., and Carney, J. P.** 2006. Mechanisms of eukaryotic DNA double strand break repair. Frontiers in Bioscience **11:**1958-1976.
38. **Clayton, T. M., Whitaker, J. P., Fisher, D. L., Lee, D. A., Holland, M. M., Weedn, V. W., Maguire, C. N., DiZinno, J. A., Kimpton, C. P., and Gill, P.** 1995. Further validation of a quadruplex STR DNA typing system: a collaborative effort to identify victims of a mass disaster. Forensic Sci Int **76:**17-25.
39. **Dixon, L. A., Dobbins, A. E., Pulker, H. K., Butler, J. M., Vallone, P. M., Coble, M. D., Parson, W., Berger, B., Grubwieser, P., Mogensen, H. S., Morling, N., Nielsen, K., Sanchez, J. J., Petkovski, E., Carracedo, A., Sanchez-Diz, P., Ramos-Luis, E., Brion, M., Irwin, J. A., Just, R. S., Loreille, O., Parsons, T. J., Syndercombe-Court, D., Schmitter, H., Stradmann-Bellinghausen, B., Bender, K., and Gill, P.** 2006. Analysis of artificially degraded DNA using STRs and SNPs - results of a collaborative European (EDNAP) exercise. Forensic Science International **164:**33-44.
40. **Fondevila, M. P., C.; Naveran, M.; Cerezo, M.; Rodriguez, A.; Calvo, R.; Fernandez, L. M.; Carracedo, A.; Lareu, M. V.** 2008. Challenging DNA: assessment of a range of genotyping approaches for highly degraded forensic samples. Forensic Sci Int Genet. **Supplement Series:**26-28.
41. **Butler, J. M., Shen, Y., and McCord, B. R.** 2003. The development of reduced size STR amplicons as tools for analysis of degraded DNA. J Forensic Sci **48:**1054-64.
42. **Westen, A. A. S., T.** 2009. Degraded DNA sample analysis using DNA repair enzymes, mini-STRs and (tri-allelic) SNPs. Forensic Sci Int Genet **Supplement Series:**505-507.
43. **Gill, P., Sparkes, R., Fereday, L., and Werrett, D. J.** 2000. Report of the European Network of Forensic Science Institutes (ENSFI): formulation and testing of principles to evaluate STR multiplexes. Forensic Sci Int **108:**1-29.
44. **Lederer, T., Braunschweiger, G., Betz, P., and Seidl, S.** 2002. Purification of STR-multiplex-amplified microsamples can enhance signal intensity in capillary electrophoresis. Int J Legal Med **116:**165-9.
45. **Pusch, C. M., Giddings, I., and Scholz, M.** 1998. Repair of degraded duplex DNA from prehistoric samples using Escherichia coli DNA polymerase I and T4 DNA ligase. Nucleic Acids Res **26:**857-9.
46. **Kovatsi, L., Nikou, D., Triantaphyllou, S., Njau, S. N., Voutsaki, S., and Kouidou, S.** 2009. DNA repair enables sex identification in genetic material from human teeth. Hippokratia **13:**165-8.
47. **Nelson, J.** 2009. Repair of Damaged DNA for Forensic Analysis. NIJ, GE Global Research Center, Niskayuna, NY.
48. **Ballantyne, J.** 2006. Assessment and *in vitro* repair of damaged DNA templates.
49. **Sugawara, N., and Haber, J. E.** 2006. Repair of DNA double strand breaks: In vivo biochemistry, p. 416-429. DNA Repair, Pt A, vol. 408.

50. **Eckardt-Schupp, F., and Klaus, C.** 1999. Radiation inducible DNA repair processes in eukaryotes. Biochimie **81:**161-171.
51. **Hall, A., and Ballantyne, J.** 2004. Characterization of UVC-induced DNA damage in bloodstains: forensic implications. Anal Bioanal Chem **380:**72-83.
52. **Miguel, A. G., and Tyrrell, R. M.** 1986. Repair of near-ultraviolet (365 nm)-induced strand breaks in Escherichia coli DNA. The role of the polA and recA gene products. Biophys J **49:**485-91.
53. **Pollard, E. C., and Fugate, J. K., Jr.** 1978. Relative rates of repair of single-strand breaks and postirradiation DNA degradation in normal and induced cells of Escherichia coli. Biophysical Journal. **24:**429-37.
54. **Blasius, M., Buob, R., Shevelev, I. V., and Hubscher, U.** 2007. Enzymes involved in DNA ligation and end-healing in the radioresistant bacterium Deinococcus radiodurans. Bmc Molecular Biology **8**.
55. **Terato, H., Kobayashi, M., Yamamoto, O., and Ide, H.** 1999. DNA strand breaks induced by ionizing radiation on Rubrobacter radiotolerans, an extremely radioresistant bacterium. Microbiological Research **154:**173-178.
56. **Vilenchik, M. M., and Knudson, A. G.** 2003. Endogenous DNA double-strand breaks: Production, fidelity of repair, and induction of cancer. Proceedings of the National Academy of Sciences of the United States of America **100:**12871-12876.
57. **Jolivet, E., Lecointe, F., Coste, G., Satoh, K., Narumi, I., Bailone, A., and Sommer, S.** 2006. Limited concentration of RecA delays DNA double-strand break repair in Deinococcus radiodurans R1. Molecular Microbiology **59:**338-349.
58. **Makarova, K. S., Aravind, L., Wolf, Y. I., Tatusov, R. L., Minton, K. W., Koonin, E. V., and Daly, M. J.** 2001. Genome of the extremely radiation-resistant bacterium Deinococcus radiodurans viewed from the perspective of comparative genomics. Microbiology & Molecular Biology Reviews **65:**44-79.
59. **Mattimore, V., and Battista, J. R.** 1996. Radioresistance of *Deinococcus radiodurans*: functions necessary to survive ionizing radiation are also necessary to survive prolonged desiccation. Journal of Bacteriology **178:**633-637.
60. **Zahradka, K., Slade, D., Bailone, A., Sommer, S., Averbeck, D., Petranovic, M., Lindner, A. B., and Radman, M.** 2006. Reassembly of shattered chromosomes in Deinococcus radiodurans. Nature **443:**569-573.
61. **Friedberg, E. C.** 2005. Suffering in silence: The tolerance of DNA damage. Nature Reviews Molecular Cell Biology **6:**943-953.
62. **Haber, J. E.** 1999. DNA recombination: the replication connection. Trends in Biochemical Sciences. **24:**271-275.
63. **Kuzminov, A.** 1999. Recombinational repair of DNA damage in Escherichia coli and bacteriophage lambda. Microbiology & Molecular Biology Reviews **63:**751-813.
64. **Marians, K. J.** 2000. Replication and recombination intersect. Current Opinion in Genetics & Development **10:**151-6.
65. **McGlynn, P.** 2004. Links between DNA replication and recombination in prokaryotes. Current Opinion in Genetics & Development **14:**107-112.
66. **Cox, M. M.** 2004. The RecA Protein, p. 369-388. *In* N. P. Higgins (ed.), The Bacterial Chromosome. American Society of Microbiology, Washington, D. C.

67. **Cox, M. M.** 2007. The bacterial RecA protein: structure, function, and regulation, p. 53-94. *In* R. Rothstein and Aguilera, A. (ed.), Topics in Current Genetics: Molecular Genetics of Recombination. Springer-Verlag, Heidelberg.
68. **Lusetti, S. L., and Cox, M. M.** 2002. The bacterial RecA protein and the recombinational DNA repair of stalled replication forks. Annual Review of Biochemistry **71:**71-100.
69. **Cox, M. M.** 2007. Motoring along with the bacterial RecA protein. Nature Reviews Molecular Cell Biology **8:**127-138.
70. **Petrova, V., Chitteni-Pattu, S., Drees, J. C., Inman, R. B., and Cox, M. M.** 2009. An SOS Inhibitor that Binds to Free RecA Protein: The PsiB Protein. Molecular Cell **36:**121-130.
71. **Lusetti, S. L., Wood, E. A., Fleming, C. D., Modica, M. J., Korth, J., Abbott, L., Dwyer, D. W., Roca, A. I., Inman, R. B., and Cox, M. M.** 2003. C-terminal deletions of the Escherichia coli RecA protein - Characterization of in vivo and in vitro effects. Journal of Biological Chemistry **278:**16372-16380.
72. **Craig, N. L., and Roberts, J. W.** 1981. Function of nucleoside triphosphate and polynucleotide in Escherichia coli recA protein-directed cleavage of phage lambda repressor. Journal of Biological Chemistry **256:**8039-44.
73. **Lohman, T. M., and Overman, L. B.** 1985. Two binding modes in *Escherichia coli* single strand binding protein-single stranded DNA complexes. Modulation by NaCl concentration. J. Biol. Chem. **260:**3594-3603.
74. **Haruta, N., Yu, X. N., Yang, S. X., Egelman, E. H., and Cox, M. M.** 2003. A DNA pairing-enhanced conformation of bacterial RecA proteins. Journal of Biological Chemistry **278:**52710-52723.
75. **Xu, L. W., and Marians, K. J.** 2002. A dynamic RecA filament permits DNA polymerase-catalyzed extension of the invading strand in recombination intermediates. Journal of Biological Chemistry **277:**14321-14328.
76. **Lusetti, S. L., Voloshin, O. N., Inman, R. B., Camerini-Otero, R. D., and Cox, M. M.** 2004. The DinI protein stabilizes RecA protein filaments. Journal of Biological Chemistry **279:**30037-30046.
77. **Gruenig, M. C., Stohl, E. A., Chitteni-Pattu, S., Seifert, H. S., and Cox, M. M.** 2010. Less Is More: Neisseria gonorrhoeae RecX Protein Stimulates Recombination by Inhibiting RecA. Journal of Biological Chemistry **285:**37188-37197.
78. **Sakai, A., and Cox, M. M.** 2009. RecFOR and RecOR as Distinct RecA Loading Pathways. Journal of Biological Chemistry **284:**3264-3272.
79. **Drees, J. C., Chitteni-Pattu, S., McCaslin, D. R., Inman, R. B., and Cox, M. M.** 2006. Inhibition of RecA protein function by the RdgC protein from Escherichia coli. Journal of Biological Chemistry **281:**4708-4717.
80. **Cazaux, C., Mazard, A. M., and Defais, M.** 1993. Inducibility of the SOS response in a recA730 or recA441 strain is restored by transformation with a new recA allele. Mol Gen Genet **240:**296-301.
81. **Ennis, D. G., Levine, A. S., Koch, W. H., and Woodgate, R.** 1995. Analysis of recA mutants with altered SOS functions. Mutation Research **336:**39-48.
82. **Lavery, P. E., and Kowalczykowski, S. C.** 1992. Biochemical basis of the constitutive repressor cleavage activity of recA730 protein. A comparison to recA441 and recA803 proteins. Journal of Biological Chemistry **267:**20648-58.

83. **Wang, T. C., Chang, H. Y., and Hung, J. L.** 1993. Cosuppression of recF, recR and recO mutations by mutant recA alleles in Escherichia coli cells. Mutat Res **294:**157-66.
84. **Gruenig, M. C., Lu, D., Won, S. J., Dulberger, C. L., Manlick, A. J., Keck, J. L., and Cox, M. M.** 2011. Creating Directed Double-strand Breaks with the Ref Protein A NOVEL RecA-DEPENDENT NUCLEASE FROM BACTERIOPHAGE P1. Journal of Biological Chemistry **286:**8240-8251.
85. **Formosa, T., and Alberts, B. M.** 1986. DNA synthesis dependent on genetic recombination: characterization of a reaction catalyzed by purified bacteriophage T4 proteins. Cell **47:**793-806.
86. **Morel, P., Cherny, D., Ehrlich, S. D., and Cassuto, E.** 1997. Recombination-dependent repair of DNA double-strand breaks with purified proteins from Escherichia coli. Journal of Biological Chemistry **272:**17091-6.
87. **Lavery, P. E., and Kowalczykowski, S. C.** 1990. Properties of recA441 protein-catalyzed DNA strand exchange can be attributed to an enhanced ability to compete with SSB protein. J Biol Chem **265:**4004-10.
88. **Alaeddini, R., Walsh, S. J., and Abbas, A.** 2010. Forensic implications of genetic analyses from degraded DNA-A review. Forensic Science International-Genetics **4:**148-157.

VII Dissemination of research findings

Research discussed in this report was disseminated at several meetings:

1. The NIJ Conference 2011 (poster). The title of this poster with the authors and their affiliations are:

New Method for the Analysis of Highly Degraded Forensic DNA Samples
Michael M. Cox[1], John R. Battista[2], Jeanne C. Putinier[3], Russell Baldwin[3], John Hartmann[3], Erin A. Ronayne[1], Kelley Gwin[2], and Elizabeth A. Wood[1]
[1]Department of Biochemistry, University of Wisconsin-Madison, Madison, WI 53706; [2]Department of Biological Sciences, Louisiana State University and A&M College, Baton Rouge, Louisiana 70803; [3]DNA Section, Orange County Crime Laboratory, Santa Ana, CA 92703

2. The NIJ conference in 2012, invited oral presenations by John Battista and Mike Cox. A poster with the same title (presented chiefly by Erin Ronayne) was also presented at that same meeting.

3. The forensic DNA Symposium held in Ames, Iowa on June 14-16, 2011, sponsored by the Midwest Forensics Resource Center. Oral presentation by Mike Cox.

4. The Green Mountain DNA conference held in August 2012. (http://vfl.vermont.gov/conference) Oral presentation by Mike Cox.

5. Paper describing this work to be submitted to Forensic Sciences International.

VIII Participating scientists and collaborations

University of Wisconsin-Madison

Michael M. Cox PI
Evelyn M. Mercer Professor of Biochemistry

Audrey J. Klingele
Graduate Assistant

Erin A. Ronayne
Graduate Assistant

Taejin Kim
Graduate Assistant

Louisiana State University

John R. Battista
Mary Lou Applewhite Professor of Biological Sciences

Kelley Gwin
Graduate Assistant

Orange County (CA) Crime Laboratory

Elizabeth A. Thompson
DNA Laboratory Director

Russell Baldwin
Sr. Forensic Scientist – DNA Supervisor

Jeanne Putinier
Forensic Scientist

Name	Sequence	Bp from STR start/end	Length
D18S51 upstream	5'-/5Biosg/aaaaattagttgggcatggtggcacgtgcctgtagtctcagcttacttgcagggctgaggcaggaggagttcttgagcccagaaggttaaggctgcagtgagccatgttcatgccactgcacttcactctgagtgacaaattgagaccttg-3'	4	150
D18S51 downstream	5'-/5Biosg/aaagatgaaataacttactgaaattgttaatgaagtattggatagctactttaaaaataacaaacccgactaccagcaacaacacaaataaacaaaccgtcagcctaaggtggacatgttggcttctctctgttcttaacatgttaaaattaaaattaacttctctggtgtgtggagatgtcttacaataacagttgct-3'	39	200
FGA upstream	5'-/5Biosg/ttctttgggattactaattgctattaggacatcttaactggcattcatggaaggctgcagggcataacattatccaaaagtcaaatgccccataggttttgaactcacagattaaactgtaaccaaaataaaattaggcatatttacaag-3'	24	150
FGA downstream	5'-/Biosg/tcgtttcatatcaaccaactgagctctaacatttttctgcagaagctggatatgctgtacttttctatgactttgcgcttcaggacttcaattctgcttctcagatcctctgacactcggttgtaggtattatcacggtctgaaatcgaaaatatggttattgaagtagctgctgagtgatttgtctgtaattgccagc-3'	30	200

Table 1: Targeting oligonucleotide sequences and locations

Copies of *gltS* target prior to capture x 10^6	Copies of *gltS* target after capture x 10^6		Copies of 16S rRNA gene after capture x 10^6
	Detection 250bp downstream of biotinylated primer binding site	Detection 5000bp downstream of biotinylated primer binding site	
101 ± 3.1 (500ng)	63 ± 2,3	2.36 ± 0.6	0.064 ± 0.04
20.1 ± 0.8 (100ng)	11.3 ± 1.4	2.63 ± 0.8	0.061 ± 0.03
2.90 ± 0.6 (1ng)	1.5 ± 0.9	3.97 ± 1.6	0.053 ± 0.05
0.27 ± 0.02 (0.1ng)	0.92 ± 0.4	0.878 ± 0.08	0.014 ± 0.02

Table 2: Recovery of a single targeted locus from purified genomic DNA. The *gltS* gene of *E. coli* was targeted for retrieval from a solution containing genomic DNA. A biotinylated capture primer was used to specifically target this locus and the number of copies recovered quantified by QPCR conducted at sites 250 and 5000bp downstream of the capture primer binding site. Non-specific binding was evaluated by monitoring the carryover of the 16S rRNA gene after the capture reaction. Values are the mean ± standard deviation of twelve measurements obtained during four independent experiments.

Figure 1. The recombinational DNA repair of stalled and collapsed replication forks. Replication forks may encounter either a DNA template strand break (that might occur during an ongoing lesion repair reaction; path A) or an unrepaired DNA lesion (path B). In either case, recombination enzymes are needed. The RecA-mediated strand invasion reaction in path A is the key reaction relied upon in the protocol we are developing.

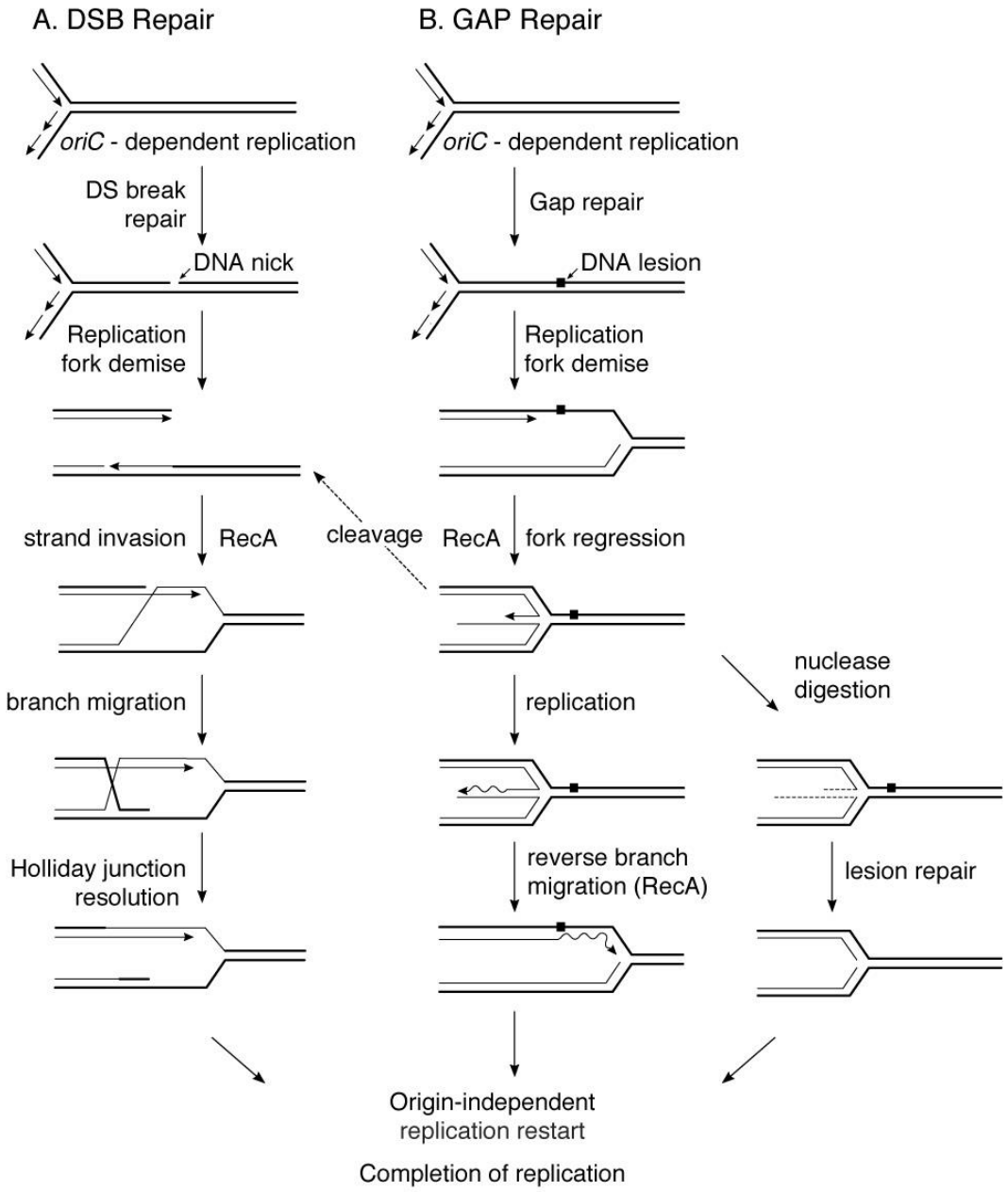

Figure 2. Genome reconstitution in *Deinococcus radiodurans*. The ability of *D. radiodurans* R1 to survive the accumulation of chromosomal DNA double strand breaks following exposure to 3000 Gy radiation. *Lane 1* is a lambda ladder size standard, *Lane 2* is chromosomal DNA prepared from an untreated culture, *lane 3* is chromosomal DNA prepared from a culture immediately after irradiation, and *lanes 4–9* are chromosomal DNA prepared from a culture 30, 60, 90, 120, 150, and 180 min post-irradiation, respectively.

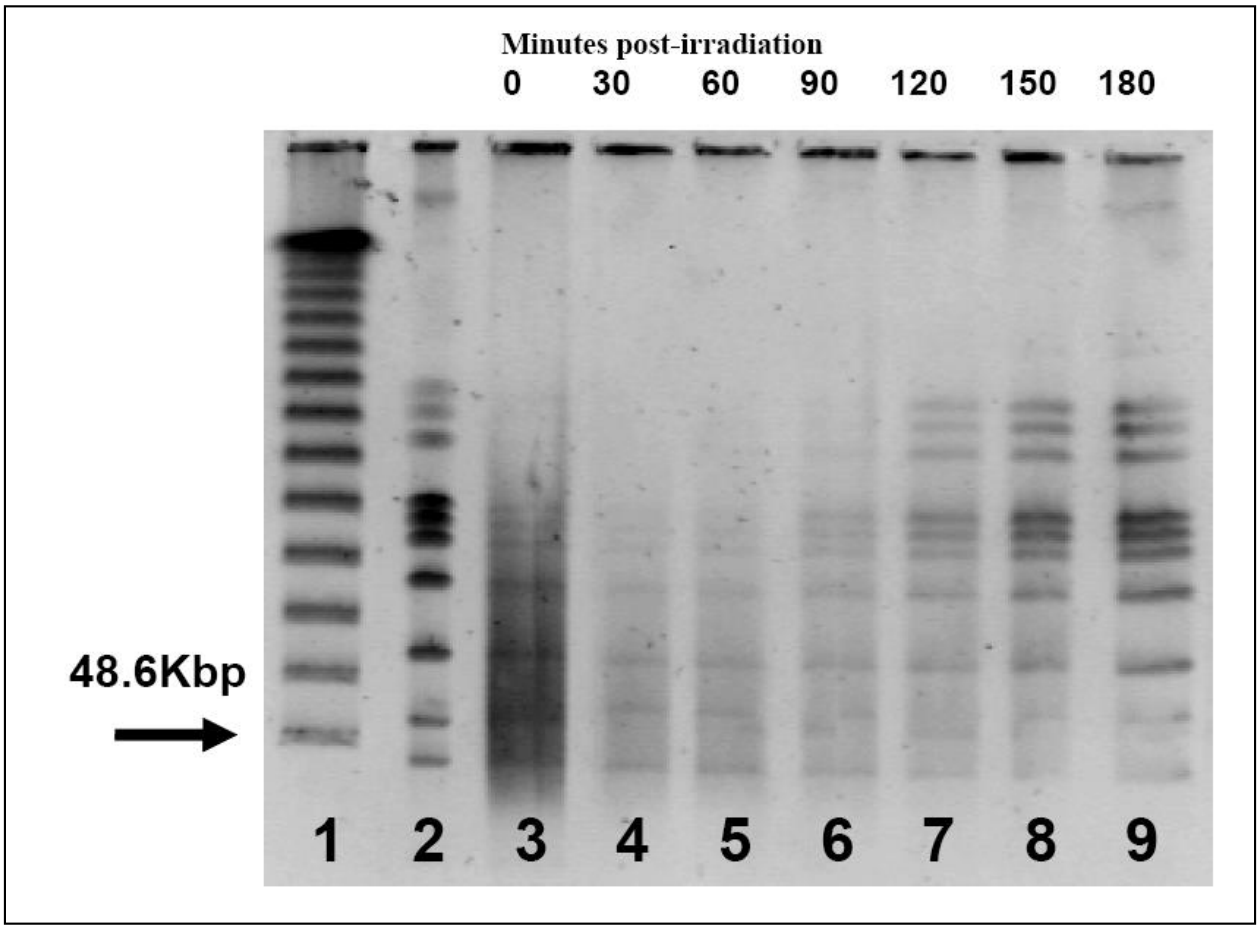

Figure 3a. Double strand break repair of highly degraded forensic DNA. A double strand break between the STR and the sequence targeted by the PCR primer will block PCR amplification. A targeting DNA segment consisting of single-stranded DNA identical in sequence to the flanking DNA of the STR will be generated, bound with RecA protein and SSB, and used to invade the damaged duplex DNA. After the invading 3′ end is extended by DNA polymerase I, the repaired DNA (with the flanking arm restored) can be amplified normally in standard STR genotyping protocols.

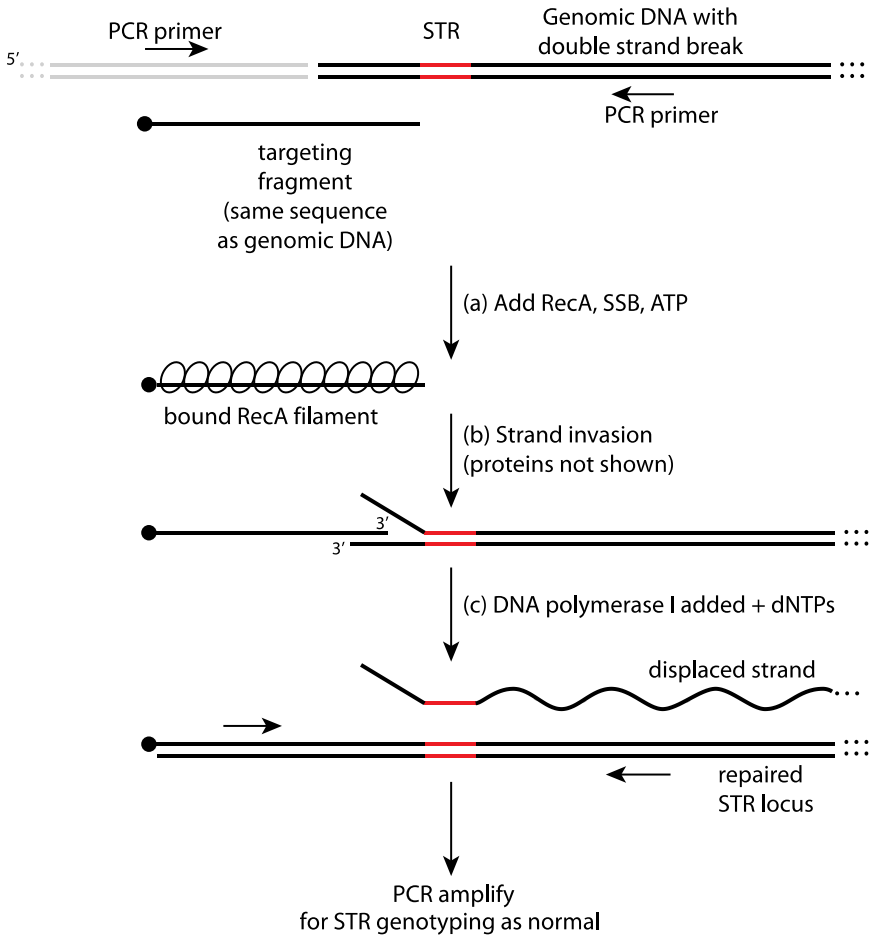

Figure 3b. Design of the targeting fragment. This will contain STR flanking DNA sequences up to approximately 2-6 bp of the STR itself. The targeting strand is the one with the 3′ end nearest the STR.

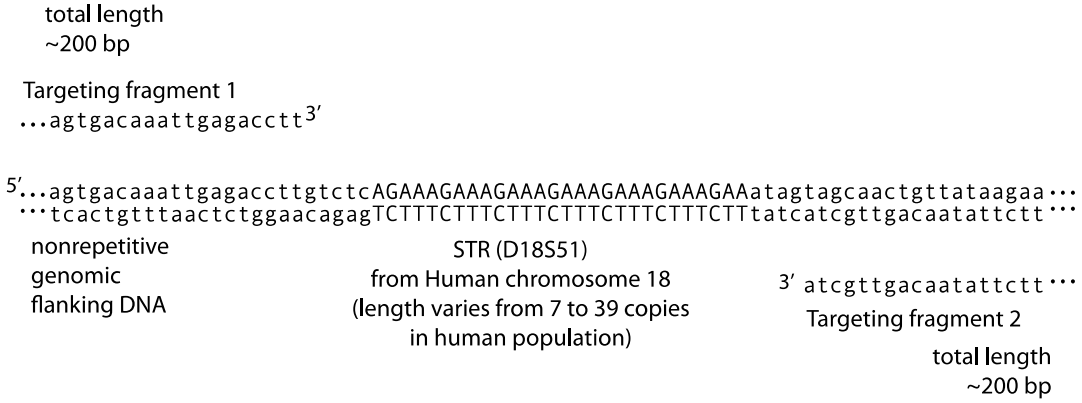

Figure 4. RecA, ATP, SSB, and DNA Pol I are both necessary and sufficient for *in vitro* double-strand break repair. *A*. Reaction scheme. Plasmid DNA is used to generate two duplex DNAs 2686 base pairs long with overlapping sequence homology. One duplex is heat denatured in water to generate two single-stranded DNAs. RecA (4 μM) is incubated with AatII-cut ssDNA (9 μMnt) and ATP (2 mM) for 3 minutes. SSB (2.4 μM) is added and incubated for 3 minutes. PstI-cut dsDNA (9 μMnt) is added and incubated for 30 minutes, and the RecA-coated ssDNA is able to invade the duplex DNA on either end to form joint molecule intermediates (*). DNA Pol I (1 U) and dNTPs (1 mM each) are added and incubated for 30 minutes and DNA synthesis fills out the 3' ends to form two repaired products of 3191 base pairs (P1) and 4869 base pairs (P2). *B*. Electrophoretic analysis of plasmid repair reactions described in *A*. Lanes 1 and 2, DNA only controls – heat-denatured AatII-cut ssDNA and PstI-cut dsDNA, respectively. Complete repair reaction, lane 3. Omitting necessary components ssDNA, dsDNA, SSB, RecA, DNA Pol I, dGTP, lanes 4-9, respectively. The New England Biolabs 2-log ladder of duplex DNA fragments is used in the marker lane to provide a point of reference between figures.

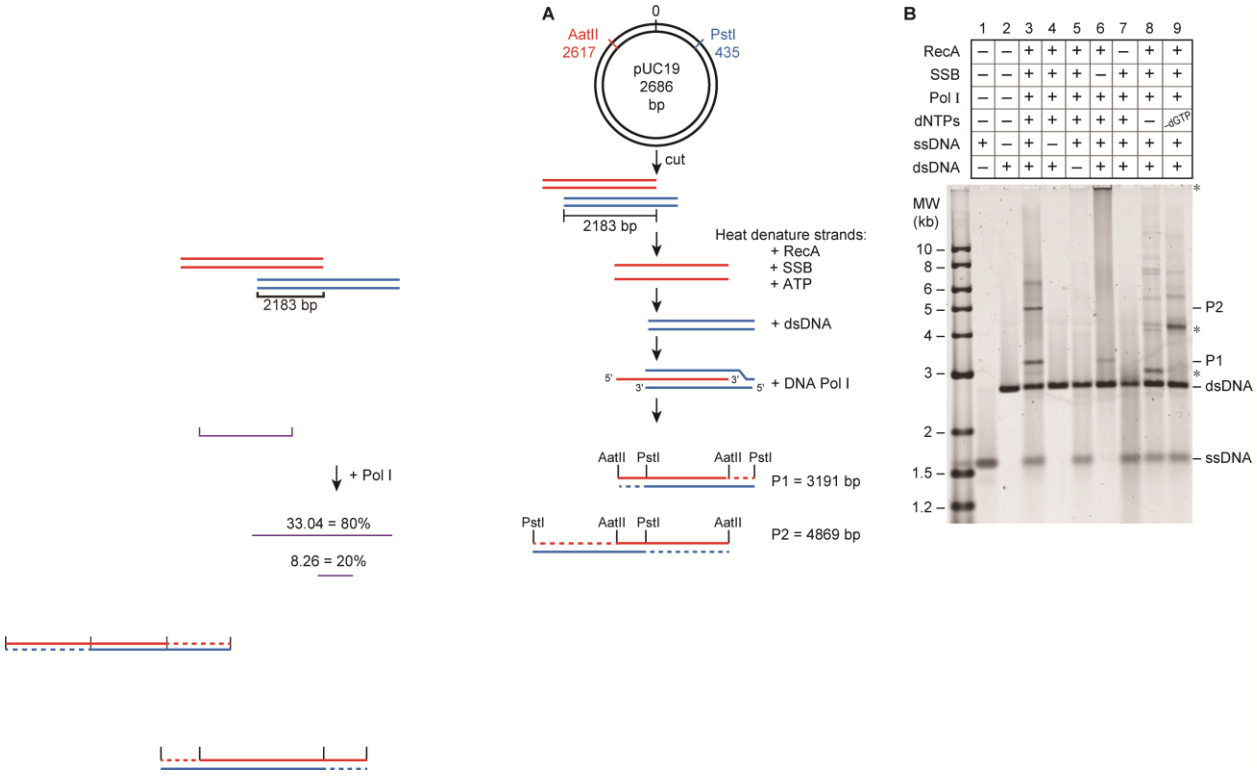

Figure 5. Repaired products accumulate over time after addition of DNA Pol I. Reactions were prepared as in Figure 1A. Lanes 4, 5, and 6 show 0, 15, and 30 minute timepoints (respectively) from the incubation of RecA and SSB with the ssDNA and dsDNA substrates at 37 °C. DNA Pol I (1U) was added with dNTPs (1mM each, lanes 7-9), or with all but dGTP (lanes 10-12) and incubated at 37°C. Timepoints were taken at 0, 15, and 30 minutes. Products of predicted length (3191 and 4869 bp) appear only when DNA Pol I and all nucleotides are present.

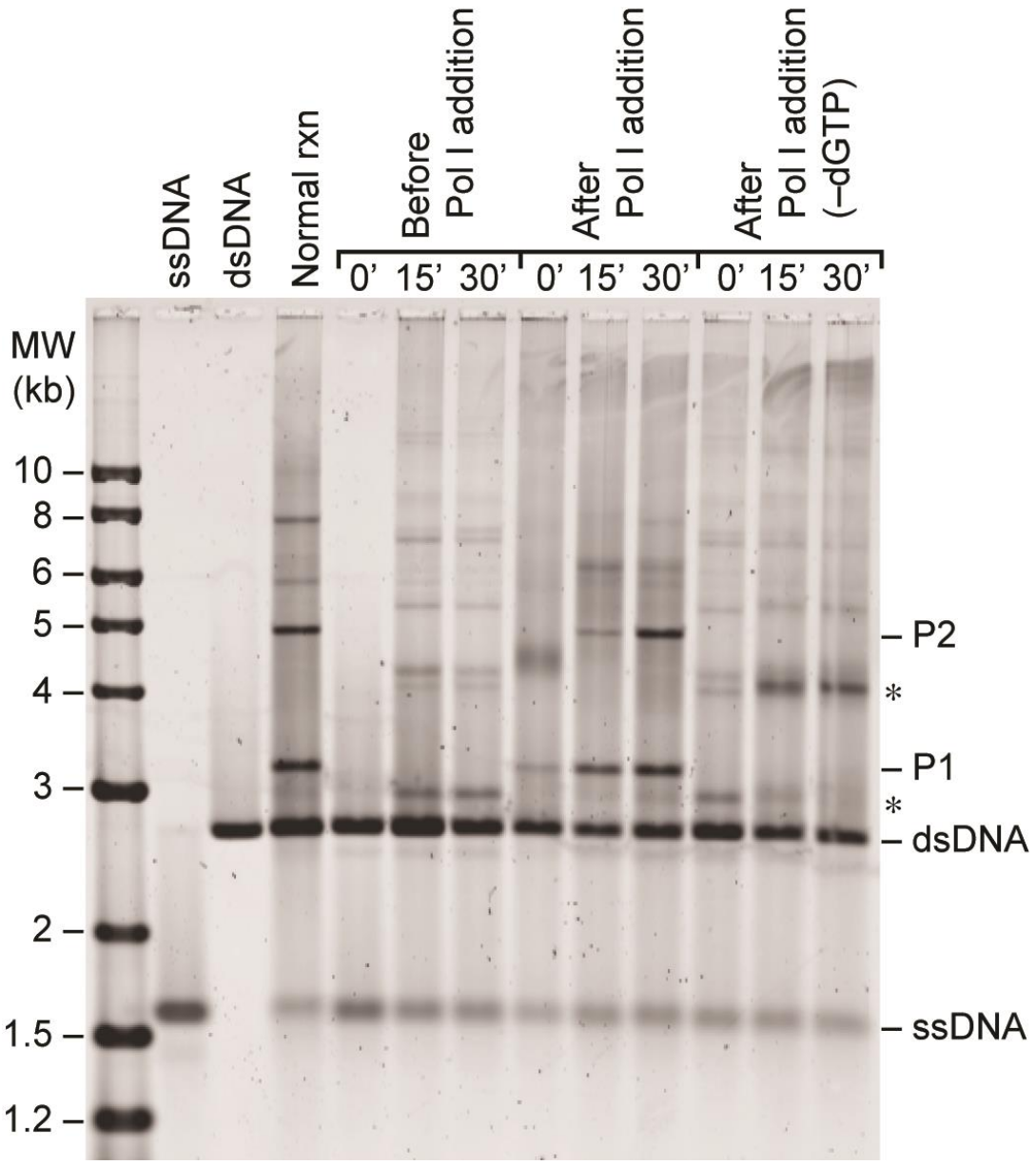

Figure 6. The double-strand break repair method can be adapted to repair DNA that flanks human short tandem repeats. Targeting oligonucleotides (150-200 nt) are designed that flank the STR of interest, and span the location of a double-strand break that will prevent PCR amplification. Upon the addition of RecA, ATP, and SSB, the RecA nucleoprotein filament invades duplex DNA at the double-strand break. DNA polymerase I then uses dNTPs to synthesize new DNA strands that span the double-strand break, allowing for PCR amplification of the STR to proceed.

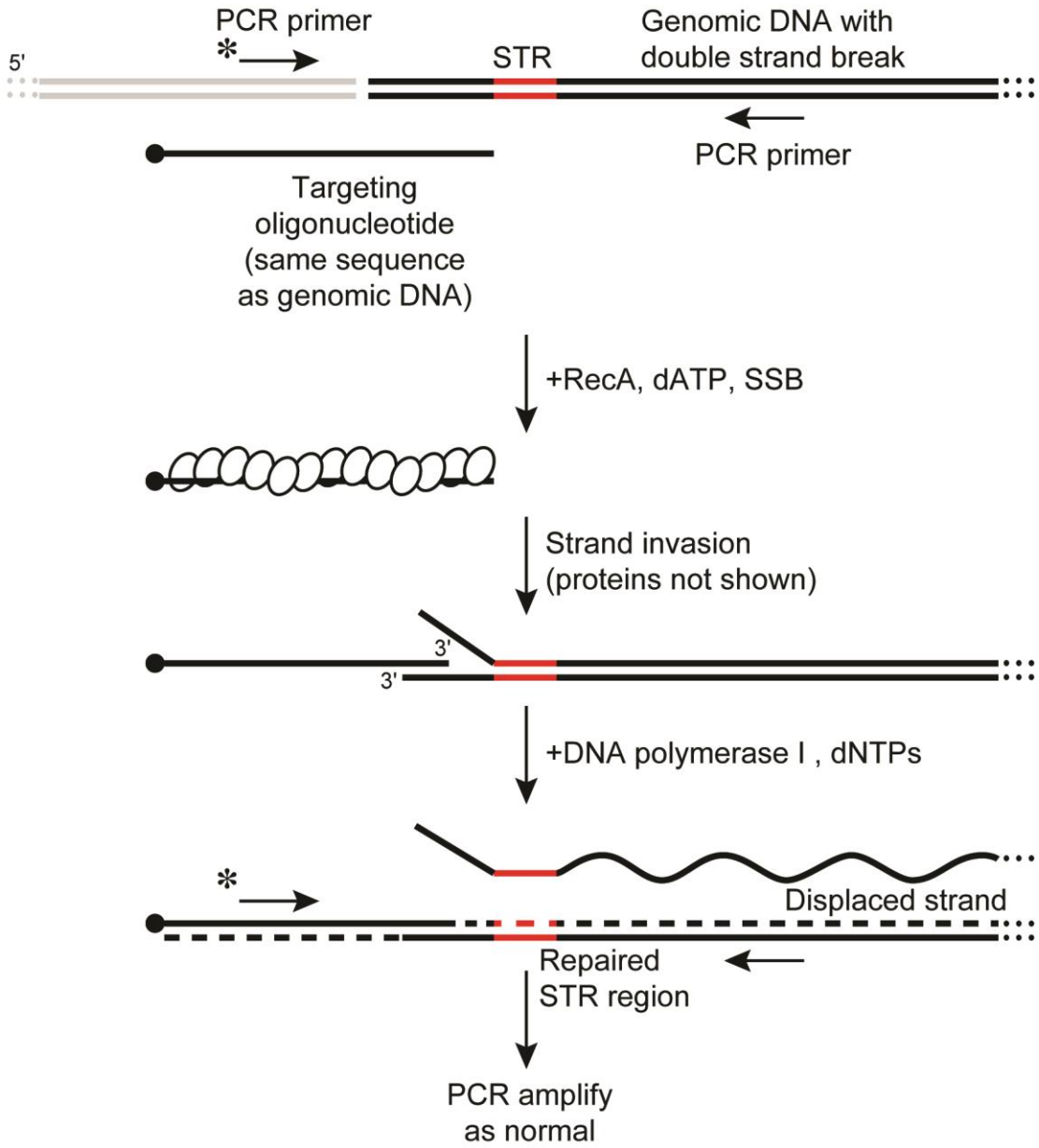

Figure 7. Human genomic DNA can be reliably and thoroughly degraded using ionizing radiation. HeLa cell line genomic DNA (10 ng/µL) is exposed to a ^{137}Cs source for approximately 12 hours/5 kGy. 10 ng of damaged DNA is used as a template for PCR amplification, as noted in Materials and Methods. Only the yellow color channel is shown for clarity, with the loci (from left to right) D19S433, vWA, TPOX, and D18S51.

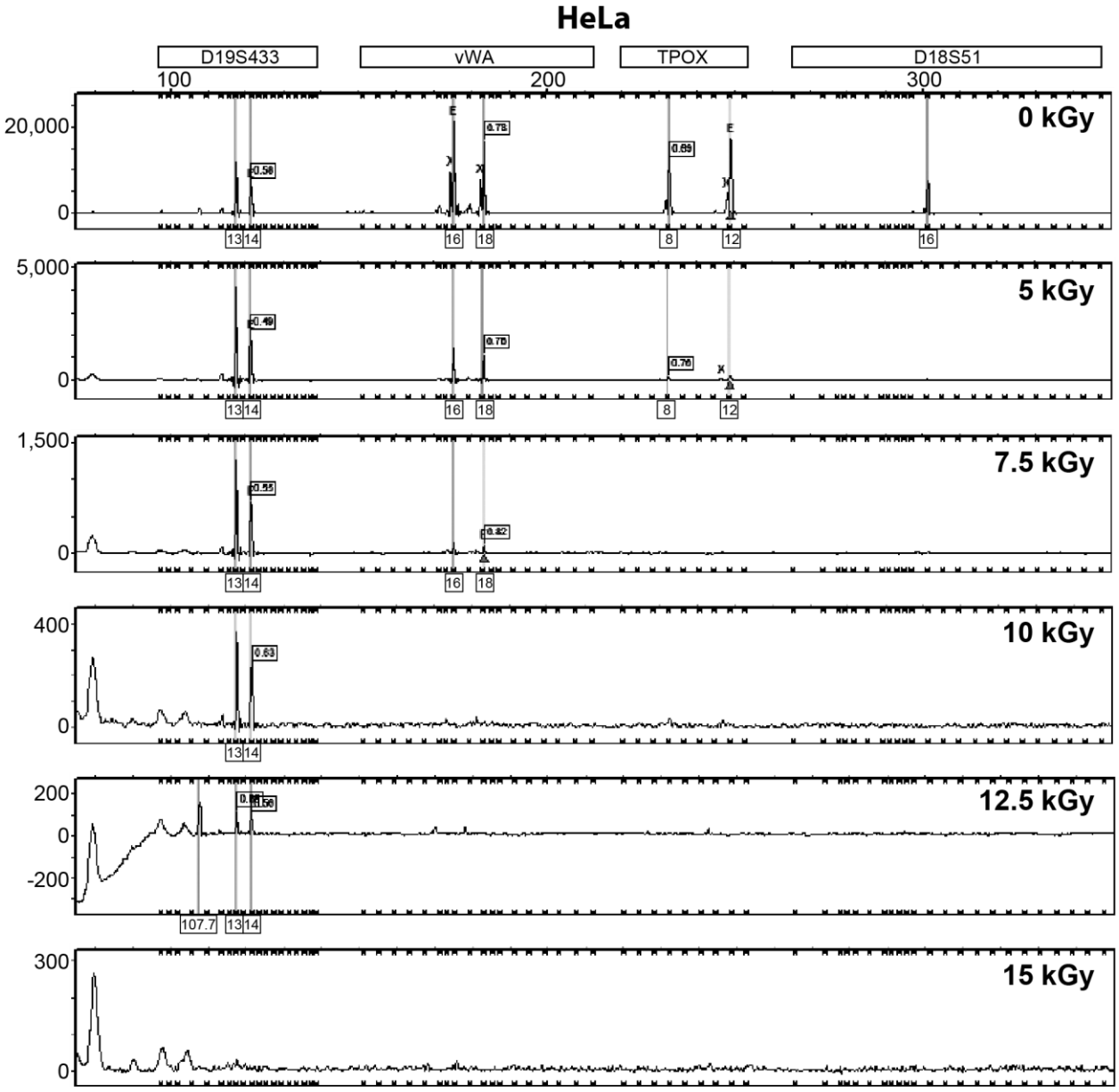

Figure 8: The double-strand break repair method can be used to enhance or restore correct allelic products at the D18S51 locus using degraded DNA. Electropherogram output of the D18S51 locus after PCR amplification and capillary electrophoresis of the following DNAs. Top row: HeLa or Jurkat cell line DNA exposed to the indicated dose of ionizing radiation in a ^{137}Cs irradiator. Middle row: Degraded DNA from top row after being subjected to the repair process as outlined in Materials and Methods, without the required targeting repair oligonucleotides. This serves as a negative control for the repair process. Bottom row: Same degraded DNA after being subjected to the repair process as outlined in Materials and Methods, with targeting repair oligonucleotides. Correct genotypes at D18S51are HeLa: 16; Jurkat: 13, 21.

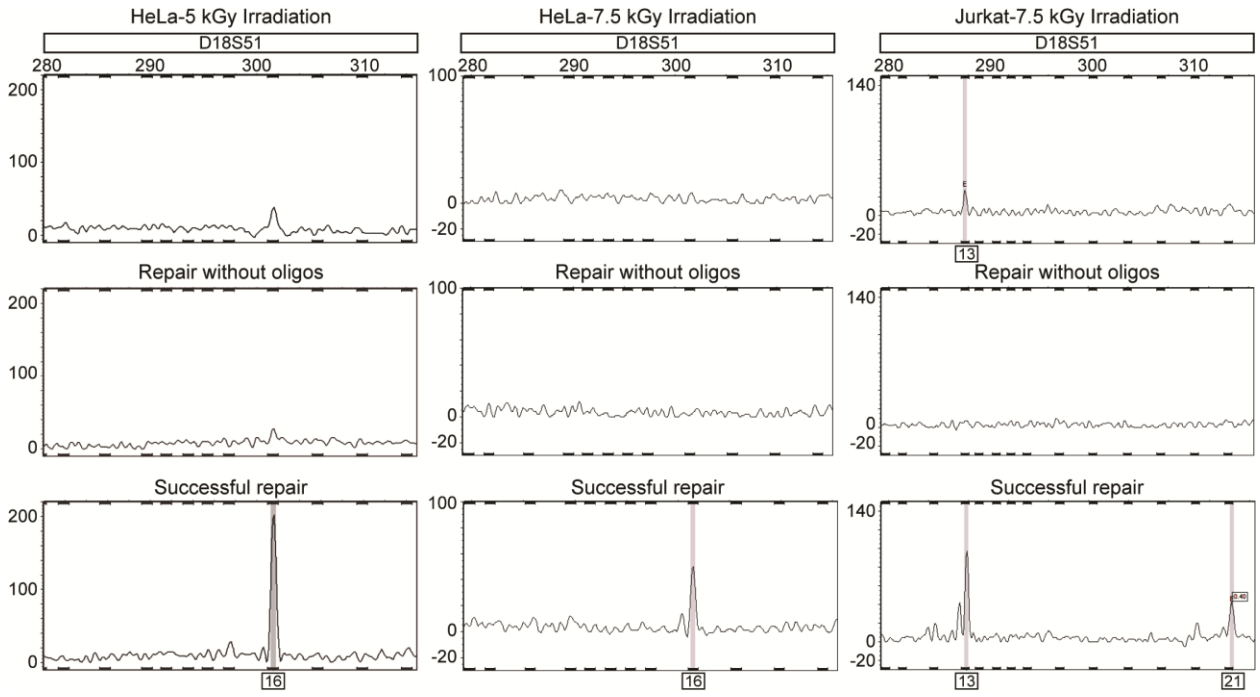

Figure 9: Peak height at the target locus increases following repair despite an overall decrease in peak height. Electropherogram output of the full Identifiler profile after PCR amplification and capillary electrophoresis of the following DNAs. Top panel: HeLa cell line DNA exposed to 5 kGy of ionizing radiation from a ^{137}Cs source. Bottom panel: Same damaged DNA after being subjected to the targeted repair of the D18S51 locus. The allelic peak restored in the D18S51 locus by the repair process is indicated by the asterisk. Artifacts are indicated with arrows.

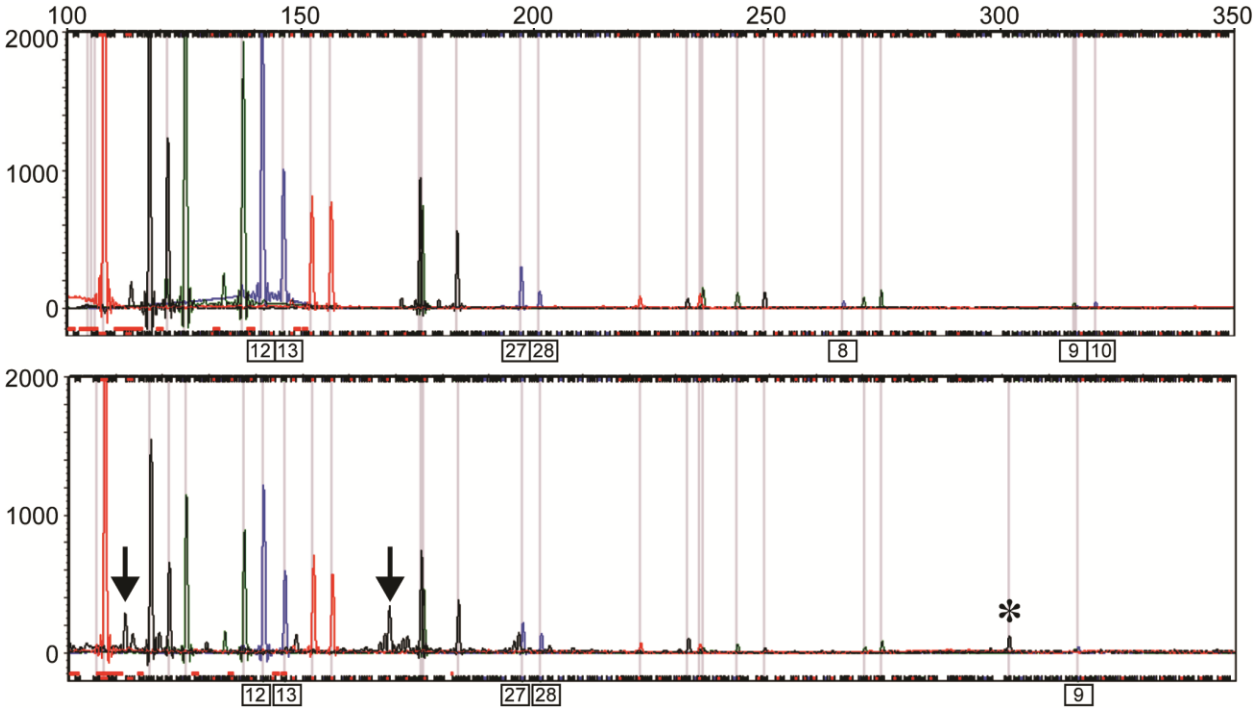

Figure 10: The method can be adapted to enhance or restore correct allelic products at the FGA locus. However, extremely low template DNA availability for repair can lead to reproducible artifacts in some samples at the FGA locus. Electropherogram output of the FGA locus after PCR amplification and capillary electrophoresis of the following DNAs. Top row: HeLa or Jurkat cell line DNA exposed to the indicated dose of ionizing radiation in a ^{137}Cs irradiator. Middle row, left panel: Degraded DNA from top row after being subjected to the repair process as outlined in Materials and Methods, without the required targeting repair oligonucleotides. Middle row, right panel: Intact, non-damaged Jurkat cell line DNA, showing the correct genotype. Bottom row: Same degraded DNA from top row after being subjected to the repair process as outlined in Materials and Methods, with targeting repair oligonucleotides. Right panel: In addition to the correct genotype, artifacts are seen near allele 24. Correct genotypes at FGA are HeLa: 18, 21; Jurkat: 20, 21.

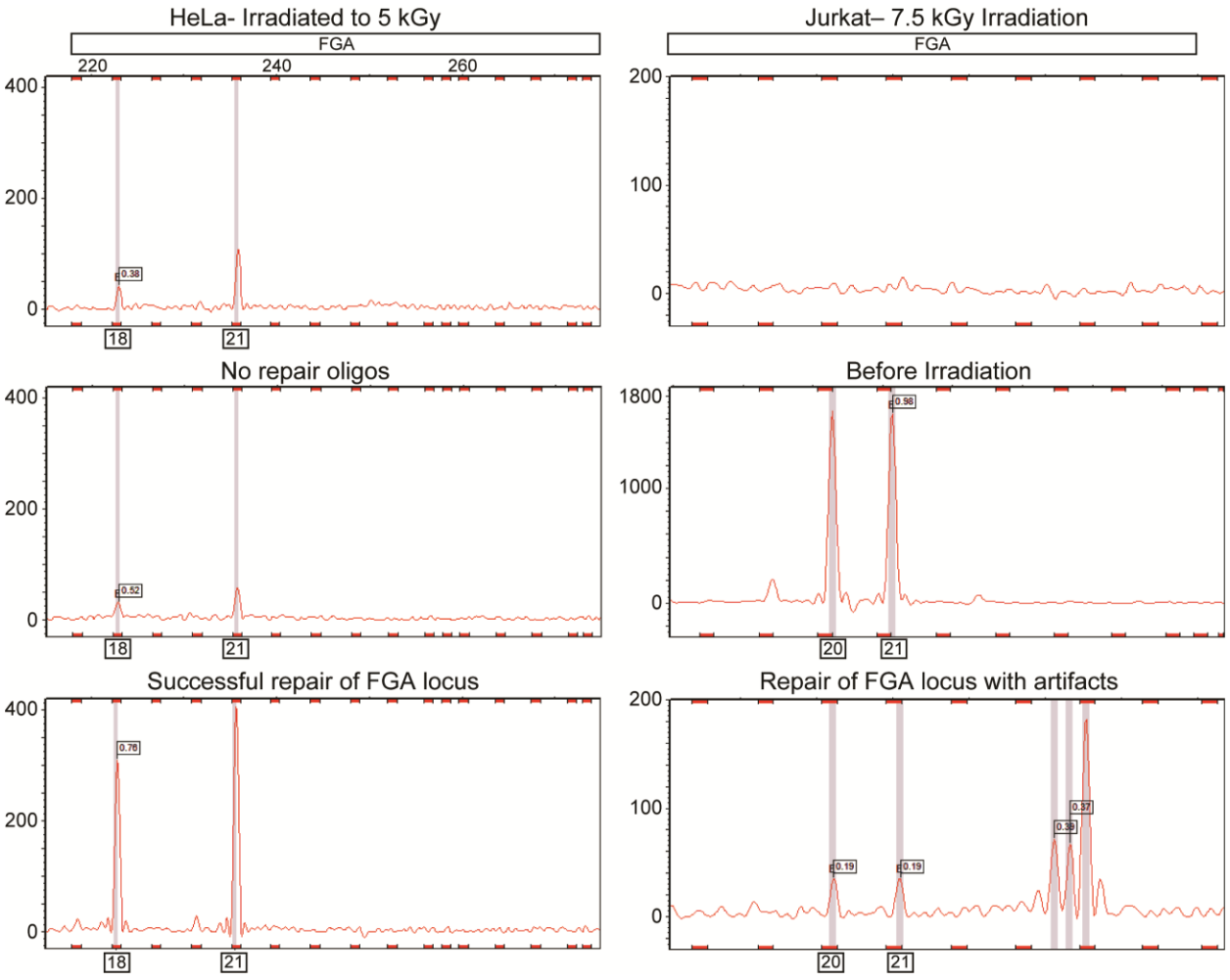

Figure 11: Repair reactions using the RecA variant V79L are successful at repairing the FGA locus in HeLa cell line DNA irradiated to 7.5 kGy. Top: HeLa cell line DNA irradiated to 7.5 kGy. No signals are seen at the FGA locus (correct signals are 18,21). Panel 2: Repair lacking repair oligonucleotides. Panel 3: Correct signals at the FGA locus (18,21) are generated when RecA V79L replaces RecA E38K in the repair reaction. Panel 4: A second independent trial confirms this result (although here only one of two correct signals is regenerated.)

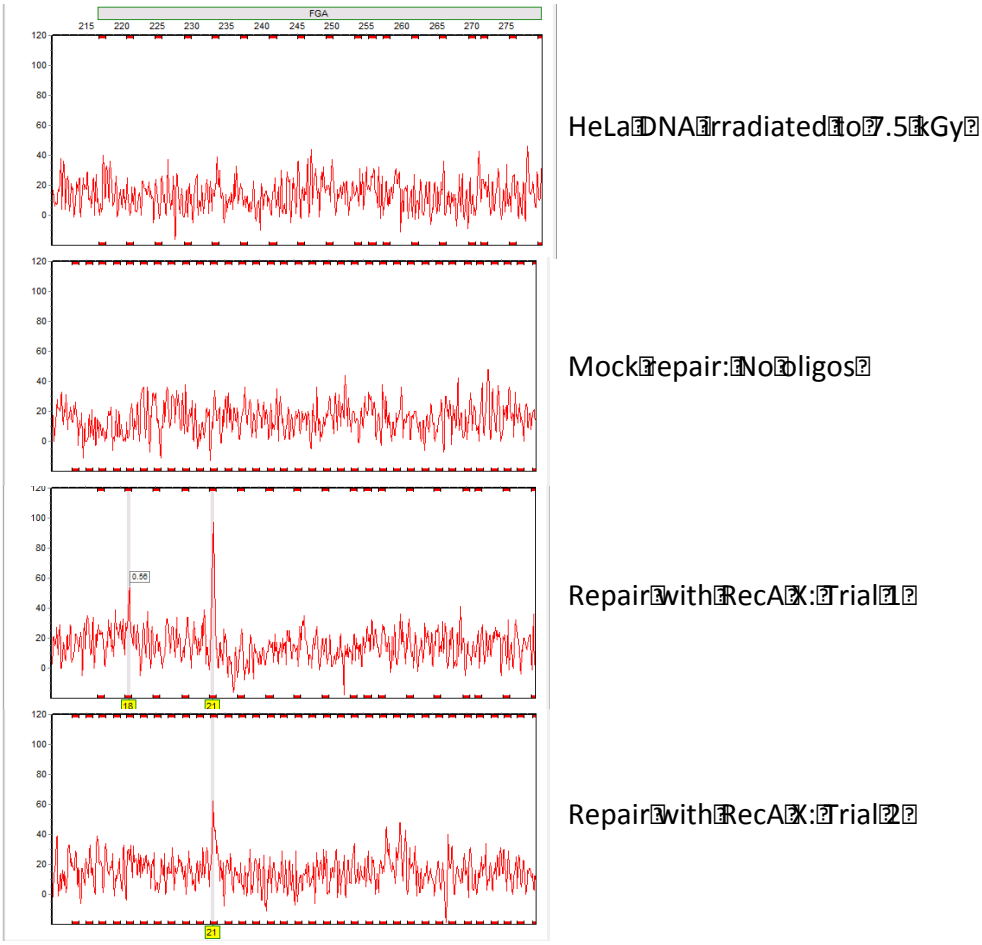

Figure 12: Repair reactions using the RecA variants V79L and I102L are successful at repairing the FGA locus in Jurkat cell line DNA irradiated to 7.5 kGy. Top: Jurkat cell line DNA irradiated to 7.5 kGy. No signals are seen at the FGA locus (correct signals are 20,21). Panel 2: Repair without repair oligonucleotides. Panel 3: One of two correct signals are regenerated at the FGA locus (20) when RecA V79L replaces RecA E38K in the repair reaction. Panel 4: One of two correct signals are regenerated at the FGA locus (20) when RecA I102L replaces RecA E38K in the repair reaction.

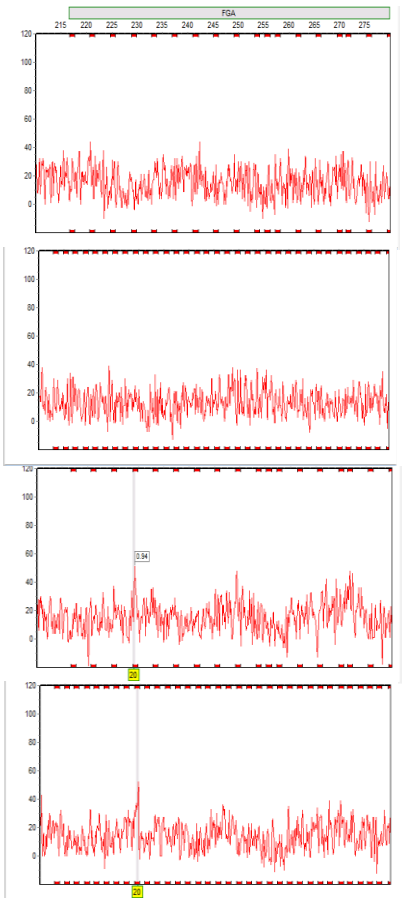

Figure 13: Pristine HeLa cell line DNA can be selectively degraded at a single locus, subsequently that signal can be restored through the repair process. Left: The D18S51 locus was targeted for degradation and subsequent repair. Right: The FGA locus as a representative locus that was not targeted for degradation or repair. Top: 4 ng of pristine HeLa cell line DNA was used as a template in the Identifiler PCR kit. Middle: The D18S51 locus was targeted for degradation using targeting oligos homologous to the regions upstream and downstream of the STR and the Ref protein. Signal strength at the D18S51 locus decreases while remaining fairly constant at the FGA locus. Bottom: The D18S51 locus was targeted for repair of the DSBs using the standard repair process without adding additional oligo. Signal strength at the D18S51 locus increases, while signal strength at the FGA locus decreases (due to loss of sample during post-repair clean-up).