Final Progress Report for Research Projects Funded by Health Research Grants

Instructions: Please complete all of the items as instructed. Do not delete instructions. Do not leave any items blank; responses must be provided for all items. If your response to an item is "None", please specify "None" as your response. "Not applicable" is not an acceptable response for any of the items. There is no limit to the length of your response to any question. Responses should be single-spaced, no smaller than 12-point type. The report **must be completed using MS Word**. Submitted reports must be Word documents; they should not be converted to pdf format. Questions? Contact Health Research Program staff at 717-783-2548.

1. **Grantee Institution:** The Pennsylvania State University

2. **Reporting Period (start and end date of grant award period):** 1/1/2009 – 12/31/2012

3. **Grant Contact Person (First Name, M.I., Last Name, Degrees):** John Anthony, MPA

4. **Grant Contact Person's Telephone Number:** 814-935-1081

5. **Grant SAP Number:** # 4100047645

6. **Project Number and Title of Research Project** 24 - Mechanisms of Microsatellite Mutagenesis in Human Cells

7. **Start and End Date of Research Project:** 07/08/2009-06/30/2010

8. **Name of Principal Investigator for the Research Project:** Kristin A. Eckert, PhD

9. **Research Project Expenses**.

   9(A) Please provide the total amount of health research grant funds spent on this project for the entire duration of the grant, including indirect costs and any interest earned that was spent:

   $ 46, 396

   9(B) Provide the last names (include first initial if multiple individuals with the same last name are listed) of **all** persons who worked on this research project and were supported with health research funds. Include position titles (Principal Investigator, Graduate Assistant, Post-doctoral Fellow, etc.), percent of effort on project and total health research funds expended for the position. For multiple year projects, if percent of effort varied from year to year, report in the % of Effort column the effort by year 1, 2, 3, etc. of the project (x% Yr 1; z% Yr 2-3).

Last Name	Position Title	% of Effort on Project	Cost
Candiello	Research Technician II	50%	$13, 255.00

9(C) Provide the names of **all** persons who worked on this research project, but who *were not* supported with health research funds. Include position titles (Research Assistant, Administrative Assistant, etc.) and percent of effort on project. For multiple year projects, if percent of effort varied from year to year, report in the % of Effort column the effort by year 1, 2, 3, etc. of the project (x% Yr 1; z% Yr 2-3).

Last Name	Position Title	% of Effort on Project
Lutzkanin	Research Technician	10 %
Baptiste	Graduate Assistant	10 %
Eckert	PI	5%

9(D) Provide a list of **all** scientific equipment purchased as part of this research grant, a short description of the value (benefit) derived by the institution from this equipment, and the cost of the equipment.

Type of Scientific Equipment	Value Derived	Cost
None		

10. Co-funding of Research Project during Health Research Grant Award Period. Did this research project receive funding from any other source during the project period when it was supported by the health research grant?

Yes_____ No___X_____

If yes, please indicate the source and amount of other funds:

11. Leveraging of Additional Funds

11(A) As a result of the health research funds provided for this research project, were you able to apply for and/or obtain funding from other sources to continue or expand the research?

Yes___X_____ No_____

If yes, please list the applications submitted (column A), the funding agency (National Institutes of Health—NIH, or other source in column B), the month and year when the application was submitted (column C), and the amount of funds requested (column D). If you have received a notice that the grant will be funded, please indicate the amount of funds to be awarded (column E). If the grant was not funded, insert "not funded" in column E.

Do not include funding from your own institution or from CURE (tobacco settlement funds). Do not include grants submitted prior to the start date of the grant as shown in Question 2. If you list grants submitted within 1-6 months of the start date of this grant, add a statement below the table indicating how the data/results from this project were used to secure that grant.

A. Title of research project on grant application	B. Funding agency (check those that apply)	C. Month and Year Submitted	D. Amount of funds requested:	E. Amount of funds to be awarded:
Computational and Biochemical Analysis of Microsatellite Life Cycle	X NIH ☐ Other federal (specify:_____) ☐ Nonfederal source (specify:_)	April 10, 2010	$2,321,062	$1,437,288 R01 GM087472
DNA Replication, DNA Repair and Microsatellite Stability	X NIH ☐ Other federal (specify:_____) ☐ Nonfederal source (specify:_)	December, 2009	$1,708,589	not funded

The purpose of this project was to provide support for a new interdisciplinary research program among experimental and computational investigators at Penn State, and to improve the competitiveness of an R01 application that we had submitted to the National Institutes of Health (NIH). Subsequent to our Health Research grant, our revised R01 application to NIH was funded, retroactive to 2009: NIH/R01 GM087472: "Computational and Biochemical Analysis of Microsatellite Life Cycle", K. A. Eckert (MPI) and K. Makova (MPI).

11(B) Are you planning to apply for additional funding in the future to continue or expand the research?

Yes__X_____ No_____

If yes, please describe your plans:

1. Planned submission to the Crohn's and Colitis Foundation of America, January 14, 2013.
2. Planned submission to the NIH of a renewal proposal for grant R01 GM087472 in Fall 2013.

12. Future of Research Project. What are the future plans for this research project?

We will test the health significance of somatic mutations arising within mature length microsatellites using an inflammatory bowel disease/colon cancer risk model. We will

continue to explore new experimental approaches to test the hypothesis in Specific aim 1, including the examination of human polymerase kappa protein levels in human tumor and pre-neoplastic tissue specimens.

13. **New Investigator Training and Development**. Did students participate in project supported internships or graduate or post-graduate training for at least one semester or one summer?

Yes____X_____ No_____

If yes, how many students? Please specify in the tables below:

	Undergraduate	Masters	Pre-doc	Post-doc
Male				
Female			1	
Unknown				
Total			**1**	

	Undergraduate	Masters	Pre-doc	Post-doc
Hispanic				
Non-Hispanic			1	
Unknown				
Total			**1**	

	Undergraduate	Masters	Pre-doc	Post-doc
White			1	
Black				
Asian				
Other				
Unknown				
Total			**1**	

14. **Recruitment of Out-of–State Researchers**. Did you bring researchers into Pennsylvania to carry out this research project?

Yes_____ No____X_____

If yes, please list the name and degree of each researcher and his/her previous affiliation:

15. **Impact on Research Capacity and Quality**. Did the health research project enhance the quality and/or capacity of research at your institution?

Yes___X_____ No_____

This Health Research project enhanced interdisciplinary research at the Pennsylvania State University in two ways. First, the funds allowed us to generate additional evidence demonstrating the degree to which biochemical results predict the mutational behavior of microsatellites in human cells. This evidence supported a new, funded NIH grant application between Dr. Eckert's experimental laboratory at the HY campus and Dr. Makova's computational laboratory at the UP campus from the National Institute of General Medical Sciences. We also published the results of this Health Research Grant in an interdisciplinary manuscript between the Eckert and Makova labs, further evidence of our collaborative approach to basic medical research and the productiveness of our team. Second, the funds allowed us to enhance our mechanistic understanding of how microsatellites arise, mutate, and eventually cease to exist within individual human genomes. Microsatellites can function as modifiers of gene expression, and mutations within microsatellites are directly implicated in common diseases, such as cancer. Because microsatellite lengths are highly variable among individuals, they are already in use as biomarkers in cancer diagnostics and forensics. Based on our basic science studies, supported in part by these funds, we have developed a paradigm-shifting hypothesis implicating microsatellite sequences directly as modifiers of disease risk in individuals. In the immediate future, we intend to extend our collaboration and submit new grant proposals aimed at predicting individuals who may be at risk for common diseases because of the precise DNA sequence composition of their genome. Thus, this Health Research grant has fostered a new multi-disciplinary, translational research project.

16. Collaboration, business and community involvement.

16(A) Did the health research funds lead to collaboration with research partners outside of your institution (e.g., entire university, entire hospital system)?

Yes___X_____ No_____

If yes, please describe the collaborations:

For publication of the results generated by this project, we combined our data with additional data from two collaborators at Pennsylvania State University, University Park campus:

Kateryna Makova, Ph.D.
Maria Krasilnikova, Ph.D.

16(B) Did the research project result in commercial development of any research products?

Yes_____ No____X_____

If yes, please describe commercial development activities that resulted from the research project:

16(C) Did the research lead to new involvement with the community?

Yes_____ No_____X_____

If yes, please describe involvement with community groups that resulted from the research project:

17. Progress in Achieving Research Goals, Objectives and Aims.

List the project goals, objectives and specific aims (as contained in the grant agreement). Summarize the progress made in achieving these goals, objectives and aims for the period that the project was funded (i.e., from project start date through end date). Indicate whether or not each goal/objective/aim was achieved; if something was not achieved, note the reasons why. Describe the methods used. If changes were made to the research goals/objectives/aims, methods, design or timeline since the original grant application was submitted, please describe the changes. Provide detailed results of the project. Include evidence of the data that was generated and analyzed, and provide tables, graphs, and figures of the data. List published abstracts, poster presentations and scientific meeting presentations at the end of the summary of progress; peer-reviewed publications should be listed under item 20.

This response should be a DETAILED report of the methods and findings. It is not sufficient to state that the work was completed. Insufficient information may result in an unfavorable performance review, which may jeopardize future funding. If research findings are pending publication you must still include enough detail for the expert peer reviewers to evaluate the progress during the course of the project.

Health research grants funded under the Tobacco Settlement Act will be evaluated via a performance review by an expert panel of researchers and clinicians who will assess project work using this Final Progress Report, all project Annual Reports and the project's strategic plan. After the final performance review of each project is complete, approximately 12-16 months after the end of the grant, this Final Progress Report, as well as the Final Performance Review Report containing the comments of the expert review panel, and the grantee's written response to the Final Performance Review Report, will be posted on the CURE Web site.

There is no limit to the length of your response. Responses must be single-spaced below, no smaller than 12-point type. If you cut and paste text from a publication, be sure symbols print properly, e.g., the Greek symbol for alpha (α) and beta (ß) should not print as boxes (□) and include the appropriate citation(s). DO NOT DELETE THESE INSTRUCTIONS.

Short, repetitive DNA sequences, named microsatellites, are a characteristic feature of the human genome. Mutation rates of microsatellites can be orders of magnitude higher than those of any other DNA sequences in the genome, resulting in high levels of polymorphism among individuals (Ellegran, 2004). However, mutation rates vary greatly among microsatellite loci, and the reasons behind such variability have not been completely deciphered. Length alterations in microsatellites can occur either in the germline (inherited) or in tissues (somatic), and provide a large pool of genetic variants for subsequent biologic selection. For instance, microsatellite allele length changes within exons lead to gain or loss of protein function via frameshift mutations or to accumulation of toxic proteins via in-frame mutations. Microsatellites located in promoter regions, UTRs, and introns can be important regulators of gene expression and splicing, and these effects are usually dependent on microsatellite allele length. Importantly, allele-length polymorphisms at several microsatellites are implicated as genetic risk factors in common diseases, including cancer, neurological and neurodegenerative diseases, and cardiovascular illness. The *broad objective* of this project was to elucidate the mechanisms of microsatellite emergence, mutation, and degeneration in individual human genomes, using an interdisciplinary, computational and experimental approach. The *overall goal* of this project was to provide direct experimental evidence in support of our new collaborative model in order to improve the competitiveness of an NIH-R01 application. The experiments in this Strategic Research Plan project were specifically designed to test whether our *in vitro* results accurately reflect microsatellite mutagenesis occurring in human genomes, a concern of the NIH grant reviewers. The goals of this project have been accomplished: an interdisciplinary, multiple principal investigator (MPI) NIH grant has been obtained, and a co-authored, interdisciplinary manuscript describing the results of this research has been accepted for publication.

17 (A). *Specific Aim 1:* To test the effect of cellular DNA polymerase κ (pol κ) levels on the degeneration of microsatellite alleles by the interruption pathway.

A.1. *Introduction.* The identity of the DNA polymerase synthesizing DNA significantly impacts microsatellite mutagenesis (Eckert and Hile 2009). Our published biochemical data specifically implicate pol κ in maintaining the stability of microsatellites (Hile et al 2012). We hypothesized that the frequency of this type of microsatellite interruption error occurring in human cells should be directly proportional to the levels of pol κ protein.

A.2. *Methods.* We originally proposed to use a lentiviral vector delivery technique to increase expression of pol κ in the non-tumorigenic, MMR-deficient (PMS2 homozygote) cell line, LCL-1261. However, through ongoing studies in our laboratory funded by another project, we discovered that lentiviral infection of LCL cells alone alters mutagenesis, both mutation frequency and specificity (K. Jacobs and K. Eckert, manuscript in preparation). Therefore, we performed a systematic investigation of alternative techniques, such as electroporation and lipofection, to efficiently introduce foreign DNA into LCL1261 cells, as directed by the product manufacturers. Following optimization, we compared two vectors for short-term transfection efficiency and long term retention following nucleofection: pNC1, an oriP-derived, episomal plasmid; and pMaxGFP, an origin-less, integrative vector. We also compared the efficiencies using supercoiled versus linearized forms of the two plasmids. DNA polymerase kappa cDNA was purchased from Invitrogen (pCMV-Sport6). A flag-tagged pol κ cDNA was obtained from the laboratory of Dr. Jean Sebastien Hoffman (Toulouse, France). Pol κ cDNA was cloned into

the pIRES and pcDNA3.1 expression vectors. Intracellular polκ protein levels were determined using standard immunoblot analyses, and either commercially available antibodies or a Pol κ antibody obtained from Dr. Tomoo Ogi, Nagaski University, Japan.

We also used an alternative strategy to test whether pol κ activity can be linked to the production of interruptions within microsatellite sequences. Several publications have indicated that the expression of pol κ and polη may be inducible by DNA damage, and that UV-irradiation of human cells results in the nuclear relocalization of pol κ and pol η into replication foci. Moreover, pol κ has been shown to be required for ~50% of DNA synthesis associated with the repair of UV damage. Our rationale was that UV damage produced within the HSV-tk gene sequence would recruit pol κ to the vicinity of the microsatellite sequence. We performed UV-irradiation of either MMR-proficient (LCL721) or MMR-deficient (LCL1261) cells that stably carry a microsatellite shuttle vector. The irradiated cells were cultured for 10 population doublings (2-4 weeks), after which time the shuttle vector DNA was harvested for mutational analyses, according to our published methods (Shah and Eckert 2009).

A.3. *Results*. As shown in Table 1, Amaxa nucleofection was an efficient method of LCL1261 cell transfection, as we observed green fluorescence in 70-80% of the cells. Traditional transfection methods (lipofection; electroporation) do not result in high efficiency transfection in this cell line (data not shown). Importantly, we continued to observe fluorescence in 5-10% of the cells for 2 weeks following nucleofection, suggesting that the plasmids are stably maintained in the cells.

Table 1. Nucleofection is an efficient method to transfect LCL-1261 cells

DNA/ Form	% GFP-positive cells/intensity		
	Day 1	Day 7	Day 14
pNC1 supercoiled	70-80% / high	70-80% / high	50-60% / high
linear	30-40% /medium	20-30% /medium	5-10% /medium
pMAXGFP supercoiled	70-80% / medium	70-80% / medium	50-60% / medium
linear	30-40% /medium	20-30% /medium	5-10% /medium

Initial experiments were performed to determine the UV dose range that would be mildly lethal to each cell line (Figure 1A). We did not observe an increase in the overall HSV-tk mutant frequency with UV dose for either cell line (Figure 1B), suggesting that the majority of UV-induced DNA lesions were efficiently repaired. DNA sequence analyses of independent mutants from the control and UV-irradiated cultures showed a significant change in the distribution of mutants between the HSV-tk gene and microsatellite target sequence (Figure 1C). In the absence of UV, the majority of mutations were located within the [GT/CA] microsatellite sequence in both cell lines. After UV-irradiation, we observed an increase in the proportion of mutants within the HSV-tk coding region. We conclude that the doses of UV used for this study were mutagenic, and that we did induce a low level of UV damage within the plasmid DNA.

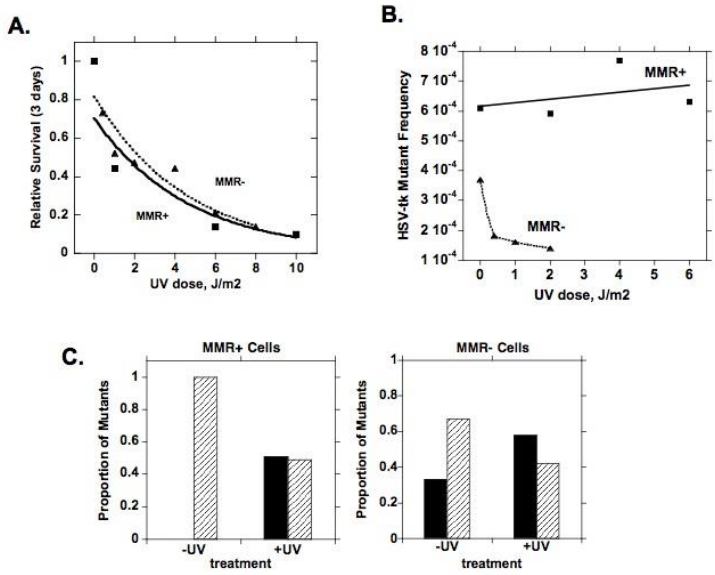

Figure 1. UV-induced mutagenesis in human LCL cells. Solid line and squares, MMR-proficient (LCL721); Dashed line and triangles, MMR-deficient (LCL-1261). **(**A). Relative survival, 3 days post irradiation. (B). overall HSV-tk mutation frequency after 10 population doublings. (C). Distribution of mutants before and after UV-irradiation. Solid bars, HSV-tk coding sequence; hatched bars, [GT/CA] microsatellite sequence.

The Sport6 plasmid purchased from Invitrogen did not encode a full length pol κ protein, as determined by DNA sequence analysis (data not shown). Due to termination of this project (time and funding), we did not test whether the Flag-tagged polκ pcDNA3 construct could be used to obtain stable pol κ overexpression in LCL1261 cells.

17 (B). *Specific Aim 2:* To test the relationship of microsatellite allele length to mutational behavior in human cells.

➢The results from this Aim have been published (Baptiste *et al*., 2012).

B.1. *Introduction*- We and others have described the microsatellite lifecycle in three phases: birth, when a locus acquires the necessary numbers of repeats to attain the threshold length; adulthood, a dynamic phase when a locus is above the threshold for mutagenesis; and death, when the locus repeat length dips below the threshold (AMOS and RUBINSTZEIN 1996) (KELKAR *et al.* 2010). Adult microsatellites can be further divided into "young" and "mature" based on the ability of the microsatellite to degrade below the threshold in one replication cycle. The focus of this aim was mature length dinucleotide microsatellites in the human genome, which are expected to be at increased risk for mutation. For dinucleotides, the threshold was determined to be five units (KELKAR *et al.* 2010; ANANDA *et al.* 2012). We empirically defined mature dinucleotides as repeats 10 units or greater in length, as we and others have not observed mutational events greater than four unit deletions in previous studies of dinucleotide microsatellites (KELKAR *et al.* 2010; ANANDA *et al.* 2012). Therefore, dinucleotides of 10 units are not expected to contract to lengths below the threshold. Genome-wide studies have identified directional biases in the mutational behavior of long microsatellites. For example, studies of human germline mutations at dinucleotide microsatellites indicated that expansions outnumber contractions (Ellegren 2004). Computational models have been derived to allow different rates of expansions and deletions depending on length. Understanding the mutational mechanisms operating within long microsatellite alleles is necessary to correctly model the evolution of these sequences.

Several potential mechanisms may underlie mutational biases within microsatellites, including DNA polymerase errors during synthesis and postreplication mismatch repair (MMR) (reviewed in (ECKERT and HILE 2009). Using mouse models, a complex effect of MMR on both the germline and somatic mutability of very long, disease-associated trinucleotide microsatellites has been demonstrated (MCMURRAY 2010). In this study, we quantified the types of mature microsatellite mutational biases present in human cells and examined how MMR influences mutational behavior. Our results uncovered unexpected mechanistic parallels between dinucleotide microsatellites commonly found within the human genome and rare, expanded trinucleotide microsatellite alleles. Our experimental data clearly support computational models of human genome microsatellites that allow higher mutation rates at long microsatellites and different rates of expansions and deletions depending on repeat length.

B.2. *Methods*- Mutational analyses of independent human cell clones were performed using our published *ex vivo* shuttle vector assay (SHAH and ECKERT 2009) in MMR^+ (LCL-721 cells) and MMR^- (LCL-1261 cells). Briefly, OriP-tk shuttle vectors carrying artificial microsatellite alleles of varying length were introduced into each human cell population. Plasmid-bearing cells were selected by growth in hygromycin, after which time independent clones were isolated by limited dilution technique. After 20-30 cell generations, the shuttle vector DNA was harvested and used to electroporate *E. coli* for determination of mutation frequency. Independent mutants were isolated from each clone, and the types of mutational events were determined by DNA sequence analyses.

B.3. *Results*. In MMR-proficient human cells, we observed an increase in MS mutation rate with increasing allele length (Figure 2). For the $[GT/CA]_n$ series, as the length of the allele increased ~two fold, from 10 to 19 units, the median MS mutation rate increased ~30-fold (Figure 2A), a difference that is statistically significant (p=0.0002, Kruskal-Wallis test). In comparison, the MS mutation rate of the $[TC/AG]_n$ series increased seven-fold between 11 and 20 units (Figure 2B), a difference that also is statistically significant (p=0.0052, Kruskal-Wallis test). Interestingly, the $[TC/AG]_n$ MS mutation rates did not change substantially over a span of five [TC/AG] units: 3.3 x 10^{-6}, 2.3 x 10^{-6}, and 4.1 x 10^{-6} for lengths of 11, 14, and 17 units respectively. In contrast, the median MS mutation rates of the $[GT/CA]_n$ alleles increased progressively, with a three-fold change from 10 to 13 units (2.1 x 10^{-7} and 6.9 x 10^{-7} respectively) and an additional five-fold change to 26 x 10^{-7} at 16 units. However, when similar repeat numbers are compared, the $[TC/AG]_n$ alleles are usually more mutable than the $[GT/CA]_n$ alleles. For example, the rate for a $[TC/AG]_{11}$ allele is 16-fold higher than that for the $[GT/CA]_{10}$ allele (p=0.038, Mann-Whitney test), while the rate for a $[TC/AG]_{20}$ allele is ~four-fold higher than that of a $[GT/CA]_{19}$ allele (p=0.016, Mann-Whitney test). Given the slope of the $[GT/CA]_n$ vs. unit number median mutation rate curve (not shown, derived from data in Figure 2A), the lower mutability of the [GT/CA] alleles than the [TC/AG] alleles cannot be accounted for by the one unit differences in total allele lengths.

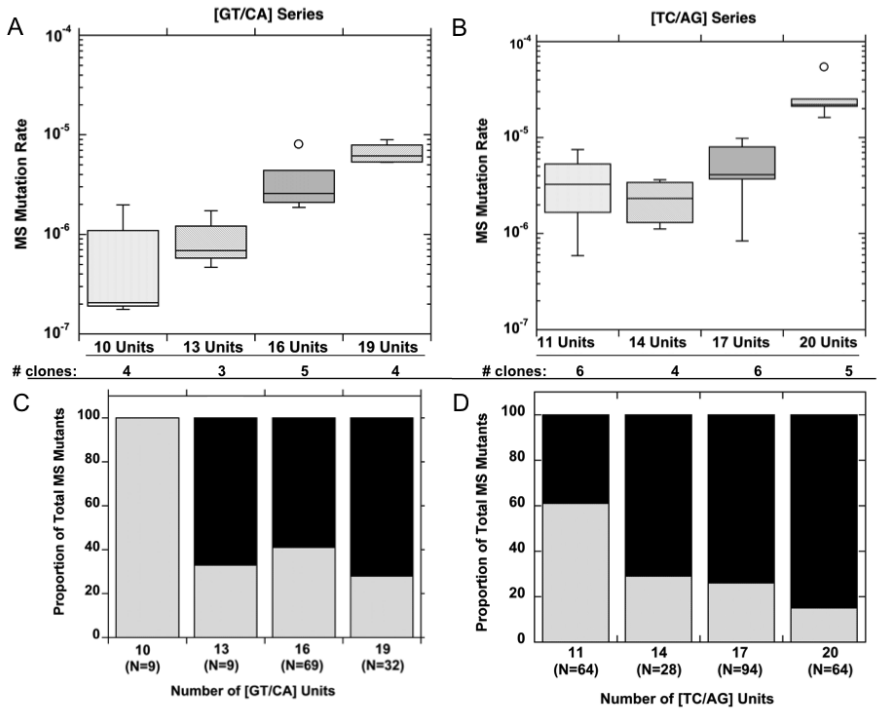

Figure 2. Microsatellite mutation rates and directional biases observed in nontumorigenic, mismatch repair-proficient human lymphoblastoid cells. The oriP-tk shuttle vector assay was performed using LCL-721 cells. (A). MS mutation rate (mutation frequency per cell generation) as a function of GT/CA units. Data are box plots for the number of clones indicated. (B). MS mutation rate as a function of TC/AG units. Data are box plots for the number of clones indicated. (C). Mutational biases within GT/CA alleles as a function of length. Total number of MS mutants observed among all clones is shown in parentheses for each allele. Solid bars, expansions of 1 or more units; gray bars, deletion of one or more units (D). Mutational biases within TC/AG alleles as a function of length. Solid bars, expansions of 1 or more units; gray bars, deletion of one or more units.

We analyzed our data set for the directionality of mutations occurring at a microsatellite allele, namely unit-based expansions *versus* deletions. In somatic human cells, we observed that the proportion of expansion mutations is greater than deletion mutations for both dinucleotide motifs at all except the shortest alleles tested. For the $[GT/CA]_n$ series, no expansion mutations were observed at the shortest length examined (10 units), whereas greater than 50% of the MS mutations were expansions within the longer alleles (13-19 units) (Figure 2C). For the $[TC/AG]_n$ series, ~40% of the microsatellite mutations were expansions within the shortest allele (11 units) (Figure 2D). Moreover, a statistically significant increase in the proportion of expansion mutations was observed as the allele length increased (14-20 units), relative to the $[TC/AG]_{11}$ allele ($p < 0.0001$ to 0.006, Fisher's exact test). A full 80% of the microsatellite mutations arising within the $[TC/AG]_{20}$ allele were expansions (Figure 2D).

We next analyzed whether differential repair contributes to the observed mutagenesis of mature microsatellites. In the canonical pathway associated with correction of DNA synthesis errors, MMR proteins recognize premutational intermediates in which the newly synthesized DNA strand (nascent DNA) differs in sequence from the parental DNA strand (template DNA). Failure of MMR to repair the intermediates results in mutations after the next round of DNA synthesis. Microsatellite expansion mutations result from premutational intermediates containing IDLs in the nascent DNA strand, while microsatellite deletions result from premutational intermediates containing IDLs in the template DNA strand. We assessed the role of MutLα in mutational bias using the human lymphoblastoid cell line (LCL1261), which does not express the PMS2 protein, a key component of the MutLα heterodimer. This cell line does express proteins of both the MutSα and MutSβ heterodimers, as well as MLH1 and MLH3 (SHAH and ECKERT 2009), proteins of the MutLγ complex. A loss-of-function comparison with LCL721 cells allows us to determine the role that PMS2/MutLα plays in generating mutational bias (Table 2).

Table 2. Comparison of $[GT/CA]_{19}$ Mutagenesis Results

MMR heterodimer		Cell Line	
		LCL721	LCL1261
MutS	α	+	+
	β	+	+
MutL	α	+	-
	γ	+	+
Expansion rate		4.4×10^{-6}	7.2×10^{-3}
Deletion rate		1.7×10^{-6}	$<1.3 \times 10^{-4}$
Expansion:Deletion		2.6:1	>55:1

Using the assay *ori*P-*tk* shuttle vector assay, $[GT/CA]_n$–containing shuttle vectors were stably replicated in LCL1261 cells. As expected, the observed MS mutation rates were 200- to 1000-fold higher than those measured for MMR-proficient cells (Figure 3; Table 2). Across all allele lengths examined, we measured a statistically significant 180-fold increase in the median MS mutation rates for the [GT/CA] motif ($p<0.0001$, Kruskal-Wallis test)(Figure 3A). Analysis of mutational spectra revealed a striking expansion bias in LCL1261 cells, wherein 91-100% of the microsatellite mutations observed within each [GT/CA] length were expansions (Figure 3B). We observed a similar relationship between expansion/deletion microsatellite mutation rates and

length for the [TC/AG] allele in MMR-deficient cells (Figure 3C). These results demonstrate that the presence of PMS2 suppresses expansion mutations within dinucleotide alleles. Loss of MutLα (comparison of LCL721 to LCL1261) resulted in an increased frequency and proportion of expansions. Taken together, our data suggest that functional MMR generates directional biases within dinucleotide microsatellites, and identifies the MutLα complex as protecting against expansion mutations. Strikingly, of the ~200 independent mutants analyzed from LCL1261 cells deficient in MutLα, only 2% displayed deletion events within the [GT/CA] microsatellite. Therefore, our data indirectly support a role for the MutLγ MMR complex in suppressing deletion mutations within dinucleotide microsatellites. Thus, the two MutL homologues are partially redundant, possibly explaining the 40:1 ratio of MLH1 to PMS2 gene mutations seen in Lynch syndrome tumors. This redundancy also may help to explain the low prevalence and penetrance of PMS2 mutations in MMR-deficient colon cancers (LYNCH and DE LA CHAPELLE 2003).

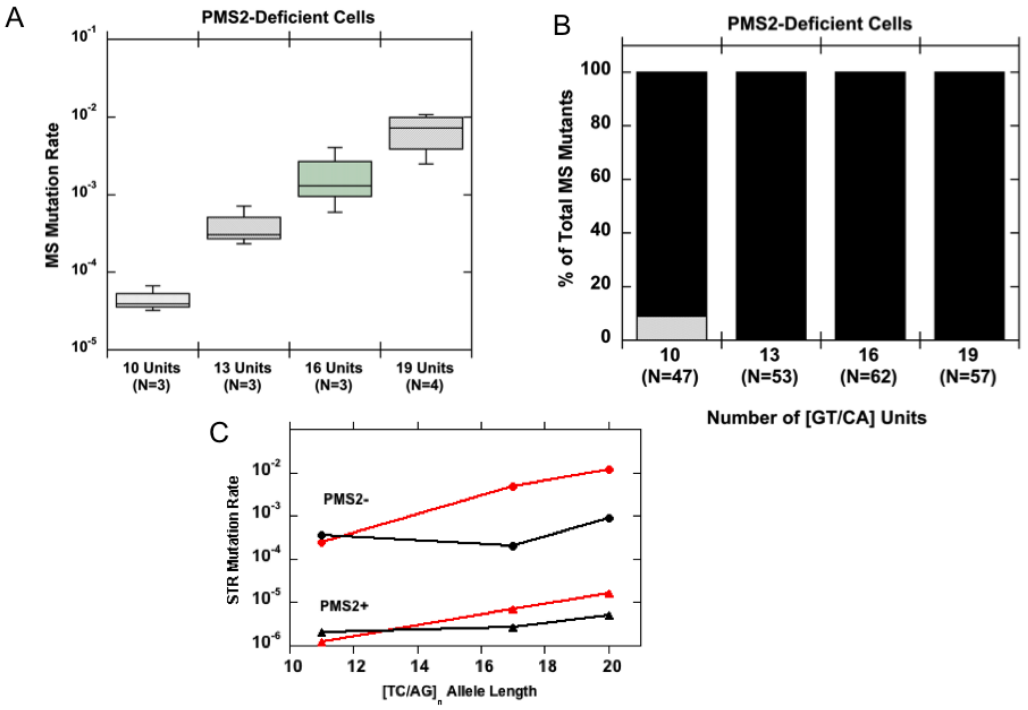

Figure 3. Dinucleotide microsatellite mutation rates and directional biases observed in mismatch repair-deficient human cells. (A). Mutation rate as a function of GT/CA allele length in PMS2-deficient human lymphoblastoid cells. Data are box plots for the number of clones shown in parentheses. (B). Mutational biases within GT/CA alleles as a function of length. Total number of MS mutants observed among all clones is shown in parentheses for each allele. Solid bars, expansions; stippled bars, deletions. (C). Mutational biases as a function of [TC/AG]allele length. Black lines, [TC/AG] deletions; Red lines, [TC/AG] expansions.

Large expansions of trinucleotide microsatellites are associated with numerous neuromuscular and neurodegenerative disorders, and models involving both DNA repair and replication have been proposed to explain these expansions (MCMURRAY 2010). We have shown here that loss of PMS2 in human cells results in a significant bias towards dinucleotide microsatellite expansion

mutations. Recent studies using Friedreich's ataxia mouse models have demonstrated that PMS2-deficient mice display an increase in expansion mutations within very long [GAA/TCC] microsatellites (BOURN *et al.* 2012). Because pre-mutational IDLs for expansions occur on the nascent strand during DNA synthesis, we propose that PMS2/MutLα-mediated MMR plays a vital role in repairing polymerase errors formed on the nascent strand (SHAH and ECKERT 2009). This model is not restricted to DNA replication, since excision repair, strand break repair, and recombination pathways all require DNA re-synthesis and involve MMR proteins.

Literature cited

Amos, W., and D. C. Rubinstzein, 1996 Microsatellites are subject to directional evolution. Nat Genet 12: 13-14.

Ananda, G., E. Walsh, T. G. P. P. S. I. A. Consortium, K. D. Jacob, M. Krasilnikova et al., 2012 Distinct mutational behaviors distinguish simple tandem repeats from microsatellites in the human genome. Gen Biol Evol, Advance Access published December 13, 2012 doi:10.1093/gbe/evs116.

Baptiste, B., G. Ananda, N. Strubczewski, Q. K. Law, K. D. Makova, M. Krasilnikova, and K. A. Eckert. Mature Microsatellites: Mechanisms Underlying Long Dinucleotide Microsatellite Mutational Biases In Human Cells, G3: Genes, Genomes, Genetics, accepted for publication December 31, 2012.

Bourn, R. L., I. De Biase, R. M. Pinto, C. Sandi, S. Al-Mahdawi et al., 2012 Pms2 Suppresses Large Expansions of the (GAA· TTC) n Sequence in Neuronal Tissues. PLoS One 7: e47085.

Eckert, K. A., and S. E. Hile, 2009 Every microsatellite is different: Intrinsic DNA features dictate mutagenesis of common microsatellites present in the human genome. Mol Carcinog 48: 379-388.

Ellegren, H., 2004 Microsatellites: simple sequences with complex evolution. Nat Rev Genet 5: 435-445.

Hile, S. E., X. Wang, M. Y. W. T. Lee and K. A. Eckert, 2012 Beyond translesion synthesis: polymerase κ fidelity as a potential determinant of microsatellite stability. Nucleic Acids Research 40: 1636-1647.

Kelkar, Y. D., N. Strubczewski, S. E. Hile, F. Chiaromonte, K. A. Eckert et al., 2010 What is a microsatellite: a computational and experimental definition based upon repeat mutational behavior at A/T and GT/AC repeats. Genome Biol Evol 2: 620-635.

Lynch, H. T., and A. de la Chapelle, 2003 Hereditary colorectal cancer. N Engl J Med 348: 919-932.

McMurray, C. T., 2010 Mechanisms of trinucleotide repeat instability during human development. Nature Reviews Genetics 11: 786-799.

Shah, S. N., and K. A. Eckert, 2009 Human postmeiotic segregation 2 exhibits biased repair at tetranucleotide microsatellite sequences. Cancer Res 69: 1143-1149.

18. Extent of Clinical Activities Initiated and Completed. Items 18(A) and 18(B) should be completed for all research projects. If the project was restricted to secondary analysis of clinical data or data analysis of clinical research, then responses to 18(A) and 18(B) should be "No."

18(A) Did you initiate a study that involved the testing of treatment, prevention or diagnostic procedures on human subjects?

_____Yes
___X__No

18(B) Did you complete a study that involved the testing of treatment, prevention or diagnostic procedures on human subjects?

_____Yes
___X__No

If "Yes" to either 18(A) or 18(B), items 18(C) – (F) must also be completed. (Do NOT complete 18(C-F) if 18(A) and 18(B) are both "No.")

18(C) How many hospital and health care professionals were involved in the research project?

_____Number of hospital and health care professionals involved in the research project

18(D) How many subjects were included in the study compared to targeted goals?

_____Number of subjects originally targeted to be included in the study
_____Number of subjects enrolled in the study

Note: Studies that fall dramatically short on recruitment are encouraged to provide the details of their recruitment efforts in Item 17, Progress in Achieving Research Goals, Objectives and Aims. For example, the number of eligible subjects approached, the number that refused to participate and the reasons for refusal. Without this information it is difficult to discern whether eligibility criteria were too restrictive or the study simply did not appeal to subjects.

18(E) How many subjects were enrolled in the study by gender, ethnicity and race?

Gender:
_____Males
_____Females
_____Unknown

Ethnicity:
_____Latinos or Hispanics
_____Not Latinos or Hispanics
_____Unknown

Race:
_____American Indian or Alaska Native
_____Asian
_____Blacks or African American
_____Native Hawaiian or Other Pacific Islander
_____White
_____Other, specify:_____
_____Unknown

18(F) Where was the research study conducted? (List the county where the research study was conducted. If the treatment, prevention and diagnostic tests were offered in more than one county, list all of the counties where the research study was conducted.)

19. Human Embryonic Stem Cell Research. Item 19(A) should be completed for all research projects. If the research project involved human embryonic stem cells, items 19(B) and 19(C) must also be completed.

19(A) Did this project involve, in any capacity, human embryonic stem cells?
_____Yes
___X__ No

19(B) Were these stem cell lines NIH-approved lines that were derived outside of Pennsylvania?
_____Yes
_____ No

19(C) Please describe how this project involved human embryonic stem cells:

20. Articles Submitted to Peer-Reviewed Publications.

20(A) Identify all publications that resulted from the research performed during the funding period and that have been submitted to peer-reviewed publications. Do not list journal abstracts or presentations at professional meetings; abstract and meeting presentations should be listed at the end of item 17. **Include only those publications that acknowledge the Pennsylvania Department of Health as a funding source** (as required in the grant agreement). List the title of the journal article, the authors, the name of the peer-reviewed publication, the month and year when it was submitted, and the status of publication (submitted for publication, accepted for publication or published.). Submit an electronic

copy of each publication or paper submitted for publication, listed in the table, in a PDF version 5.0.5 (or greater) format, 1,200 dpi. Filenames for each publication should include the number of the research project, the last name of the PI, the number of the publication and an abbreviated research project title. For example, if you submit two publications for PI Smith for the "Cognition and MRI in Older Adults" research project (Project 1), and two publications for PI Zhang for the "Lung Cancer" research project (Project 3), the filenames should be:

Project 1 – Smith – Publication 1 – Cognition and MRI
Project 1 – Smith – Publication 2 – Cognition and MRI
Project 3 – Zhang – Publication 1 – Lung Cancer
Project 3 – Zhang – Publication 2 – Lung Cancer

If the publication is not available electronically, provide 5 paper copies of the publication.

Note: The grant agreement requires that recipients acknowledge the Pennsylvania Department of Health funding in all publications. Please ensure that all publications listed acknowledge the Department of Health funding. If a publication does not acknowledge the funding from the Commonwealth, do not list the publication.

Title of Journal Article:	Authors:	Name of Peer-reviewed Publication:	Month and Year Submitted:	Publication Status (check appropriate box below):
1. Mature Microsatellites: Mechanisms Underlying Dinucleotide Microsatellite Mutational Biases In Human Cells	Beverly Baptiste, Guruprasad Ananda, Noelle Strubczewski, Andrew Lutzkanin, Su Jen Khoo, Abhinaya Srikanth, Nari Kim, Kateryna D. Makova, Maria M. Krasilnikova, and Kristin A. Eckert	G3: Genes, Genomes, Genetics	December, 2012	☐Submitted ☐ Accepted X Published

20(B) Based on this project, are you planning to submit articles to peer-reviewed publications in the future?

Yes_____ No_____X_____

If yes, please describe your plans:

21. Changes in Outcome, Impact and Effectiveness Attributable to the Research Project. Describe the outcome, impact, and effectiveness of the research project by summarizing its impact on the incidence of disease, death from disease, stage of disease at time of diagnosis, or other relevant measures of outcome, impact or effectiveness of the research project. If there were no changes, insert "None"; do not use "Not applicable." Responses must be

single-spaced below, and no smaller than 12-point type. DO NOT DELETE THESE INSTRUCTIONS. There is no limit to the length of your response.

None

22. Major Discoveries, New Drugs, and New Approaches for Prevention Diagnosis and Treatment. Describe major discoveries, new drugs, and new approaches for prevention, diagnosis and treatment that are attributable to the completed research project. If there were no major discoveries, drugs or approaches, insert "None"; do not use "Not applicable." Responses must be single-spaced below, and no smaller than 12-point type. DO NOT DELETE THESE INSTRUCTIONS. There is no limit to the length of your response.

None

23. Inventions, Patents and Commercial Development Opportunities.

23(A) Were any inventions, which may be patentable or otherwise protectable under Title 35 of the United States Code, conceived or first actually reduced to practice in the performance of work under this health research grant? Yes_____ No___X_____

If "Yes" to 23(A), complete items a – g below for each invention. (Do NOT complete items a - g if 23(A) is "No.")

a. Title of Invention:

b. Name of Inventor(s):

c. Technical Description of Invention (describe nature, purpose, operation and physical, chemical, biological or electrical characteristics of the invention):

d. Was a patent filed for the invention conceived or first actually reduced to practice in the performance of work under this health research grant?
Yes_____ No____

If yes, indicate date patent was filed:

e. Was a patent issued for the invention conceived or first actually reduced to practice in the performance of work under this health research grant?
Yes_____ No____
If yes, indicate number of patent, title and date issued:
Patent number:
Title of patent:
Date issued:

f. Were any licenses granted for the patent obtained as a result of work performed under this health research grant? Yes_____ No____

If yes, how many licenses were granted?_____

g. Were any commercial development activities taken to develop the invention into a commercial product or service for manufacture or sale? Yes___ No____

If yes, describe the commercial development activities:

23(B) Based on the results of this project, are you planning to file for any licenses or patents, or undertake any commercial development opportunities in the future?

Yes_____ No___X_____

If yes, please describe your plans:

24. Key Investigator Qualifications. Briefly describe the education, research interests and experience and professional commitments of the Principal Investigator and all other key investigators. In place of narrative you may insert the NIH biosketch form here; however, please limit each biosketch to 1-2 pages. *For Nonformula grants only – include information for only those key investigators whose biosketches were not included in the original grant application.*

BIOGRAPHICAL SKETCH

Provide the following information for the Senior/key personnel and other significant contributors.
Follow this format for each person. **DO NOT EXCEED FOUR PAGES.**

NAME Eckert, Kristin A.	POSITION TITLE Professor of Pathology
eRA COMMONS USER NAME (credential, e.g., agency login) KAEckert	

EDUCATION/TRAINING *(Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable.)*

INSTITUTION AND LOCATION	DEGREE *(if applicable)*	MM/YY	FIELD OF STUDY
Pennsylvania State University, University Park, PA	B.S.	05/82	Microbiology
University of Wisconsin-Madison, Madison, WI	Ph.D.	12/88	Oncology
National Institute of Environmental Health Sci, RTP, NC	Postdoctoral	06/93	Molecular Genetics

A. Personal Statement

The experimental focus of my research career has been mechanisms of mutagenesis in somatic human cells, particularly in relation to neoplastic cell evolution. I have 24 years of basic research experience studying damage-induced mutagenesis, DNA repair, DNA polymerase fidelity, and repetitive DNA sequence biochemistry. My Ph.D. training focused on the relationship of chemical carcinogenesis, mutagenesis and tumor initiation. As a postdoctoral fellow, I studied DNA polymerase structure/function in relation to DNA polymerase error discrimination mechanisms. At the Penn State College of Medicine, my lab developed complementary *in vitro* and *ex vivo* approaches to study mutagenesis in human cells caused by endogenous DNA lesions (alkylation and oxidative lesions), using an HSV-tk gene-based assays. Beginning in 1995, the primary focus of my independent laboratory became mechanisms of microsatellite DNA mutagenesis. Microsatellite DNA instability is a well-known genetic signature found in colorectal cancer. We have used both nontumorigenic and tumorigenic human cell lines to elucidate the contributions of DNA polymerase fidelity and mismatch repair to microsatellite instability, within the colorectal cancer model. My research program includes an interdisciplinary, collaborative research project with Dr. Makova's computational group (Penn State University Park) to study microsatellite DNA evolution in the human genome. Recently, my laboratory embarked on the in-depth examination of replicative DNA polymerases and microsatellites present within common fragile sites (CFS), regions of the genome prone to chromosomal instability.

B. Positions

Positions and Employment

1993-1999 Assistant Professor, Departments of Pathology and Biochemistry and Molecular Biology, Pennsylvania State University College of Medicine, Hershey, PA
1999-2007 Associate Professor (with tenure), Penn State College of Medicine
2007-present Professor, Penn State College of Medicine
2006-present Member, Penn State Hershey Cancer Institute

Other Experience and Professional Memberships

2000-2008 Member, Genetics SSSY Study Section, NIH/CSR
2000-2006 Member, American Cancer Society CNE Study Section (Chair, 2006)
2006-2010 Member, Radiation Therapeutics and Biology Study Section, NIH/CSR (Acting Co-chair)

C. Selected Peer-reviewed Publications (Selected from over 50 peer-reviewed publications)

1. Sweasy, J.B., J.M. Lauper and **Eckert KA**. DNA Polymerases and human disease, *Rad Res* 2006; 166: 693-714.
2. Hile SE and **Eckert KA.** DNA Polymerase κ produces interrupted mutations and displays polar pausing within mononucleotide microsatellite sequences. *Nucleic Acids Res*. 2008; 36: 688-696
3. Shah SN, **Eckert KA**. Human postmeiotic segregation 2 exhibits biased repair at tetranucleotide microsatellite sequences. *Cancer Res*. 2009; 69(3):1143-9.
4. **Eckert KA**, Hile SE. Every microsatellite is different: Intrinsic DNA features dictate mutagenesis of common microsatellites present in the human genome. *Mol Carcinog*. 2009; 48(4):379-88.
5. Shah SN, Hile SE, **Eckert KA**. Defective mismatch repair, microsatellite mutation bias, and variability in clinical cancer phenotypes. *Cancer Res*. 2010; 70(2):431-5.
6. Shah SN, Opresko PL, Meng X, Lee MY, **Eckert KA**. DNA structure and the Werner protein modulate human DNA polymerase delta-dependent replication dynamics within the common fragile site FRA16D. *Nucleic Acids Res*. 2010; 38:1149-1162.
7. Damerla RR, Knickelbein K, Kepchia D, Jackson A, Armitage B, **Eckert KA**, and Opresko PL. Telomeric repeat mutagenicity in human somatic cells is modulated by repeat orientation and G-quadruplex stability, *DNA Repair* 2010; 9: 1119-1129.
8. Kelkar, YD, Strubczewski, N. Hile, SE, Chiaromonte, F**, Eckert KA***, and **Makova KD***. What is a microsatellite: A computational and experimental definition based upon repeat mutational behavior at A/T and GT/CA repeats, *Genome Biology and Evolution*. 2010; 2: 620-635. ***Co-corresponding authors**
9. Abdulovic, AL, Hile SE, Kunkel TA, and **Eckert KA**. The in vitro fidelity of yeast DNA polymerase δ and polymerase ε holoenzymes during dinucleotide microsatellite DNA synthesis, *DNA Repair* 2011; 10: 497-505.
10. Hile, SE, Wang XX, Lee MYWT, and **Eckert KA.** Beyond Translesion Synthesis: Polymerase κ Fidelity as a Potential Determinant of Microsatellite Stability, *Nucleic Acids Res.* 2012; 40: 1636-47.
11. Kelkar, YD., **Eckert KA**, Chiaromonte F, and Makova KD. A matter of life and death: How microsatellites emerge in and vanish from the human genome. *Genome Res.* 2011; 21: 2038-48.
12. Fungtammasan, A., Walsh E, Chiaromonte F*, **Eckert KA***, and **Makova** KD* A Genome-wide Analysis of Common Fragile Sites: What Features Determine Chromosomal Instability in the Human Genome? *Genome Research* 2012; **22**: 993-1005. ***Co-corresponding**
13. Baptiste, B. and **Eckert KA**. DNA Polymerase kappa microsatellite synthesis: two distinct mechanisms of slippage-mediated errors. *Environ. Mol. Mutagenesis* 2012; **53:** 787-796
14. Walsh, E., X. Wang, M. Lee and **K.A. Eckert**. Mechanism of replicative DNA polymerase δ pausing and a potential role for DNA polymerase κ in common fragile site replication. *J Mol Biol.* 2012 Nov 19. doi: 10.1016/j.jmb.2012.11.016. [Epub ahead of print]
15. Hile, S.E., S. Shabashev, and **K.A. Eckert**. Tumor-specific microsatellite instability: Do distinct mechanisms underlie the MSI-L and EMAST phenotypes? *Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis*, 2012; Dec 1. doi:pii: S0027-5107(12)00219-9. 10.1016 [Epub ahead of print] (invited review).