PennState College of Education | CSATS Center for Science and the Schools

Research Experiences for Teachers

PennState Cancer Institute

Mismatch Repair Affects Mutagenesis of Secondary Structure Forming Microsatellite Sequences

Tiffany Lewis[1], Suzanne Hile[2], Dr. Kristin Eckert[2]

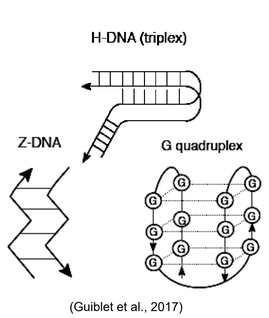
[1]Middletown Area High School, [2]Department of Pathology – Pennsylvania State University College of Medicine, Hershey, PA

Abstract

Short tandem repeats of the genome are known as microsatellites (MS), which form non-B DNA and frequently mutate due to replication difficulty. One DNA repair mechanism, known as mismatch repair, functions to correct errors in microsatellite replication caused by DNA polymerase slippage and data suggests that the efficacy of the repair is dependent on microsatellite motif and length. Comparison of the MMR-/+ ratio for dinucleotide repeats of different lengths suggests that MMR fixes shorter sequences more effectively, but comparing MMR in dinucleotides of similar length shows a dependence on sequence.

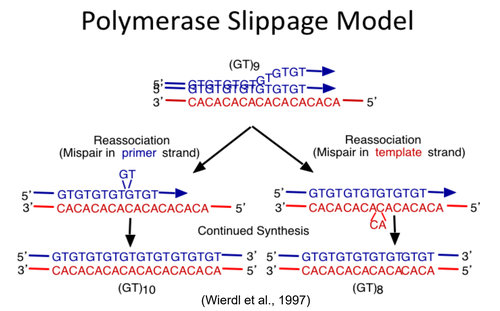
Background

DNA mutations have been linked to many different disorders, most notably cancer. Microsatellites (MS) are frequently mutated repetitive sequences in DNA that consist of 1-6 repeated nucleotide bases in tandem and have been linked to the predisposition for Lynch syndrome, a hereditary colon cancer. They are naturally present in the genomes of both prokaryotes and eukaryotes and are subject to mutation due to polymerase slippage (Fig. 2). MS sequences form non-B secondary DNA structures (Fig. 1) that make them difficult to replicate by DNA polymerase. DNA mismatch repair (MMR) proteins are able to correct errors in microsatellite regions, but their effectiveness may be based on the microsatellite sequence and length. In this study, microsatellites of varying primary sequence and length were analyzed for their ability to form single-stranded DNA (ssDNA) secondary structures, then for their mutational frequencies in both MMR +/- *E. coli*.



(Guiblet et al., 2017)

Figure 1. Examples of non-B secondary structures formed by different microsatellite motifs.



(Wierdl et al., 1997)

Figure 2 (right). During replication of repetitive DNA, strands may dissociate and reassociate with looped-out base pairs in the primer strand (expansion mutations) or the template strand (deletion mutations. MMR recognized and repairs these looped-out bases to avoid mutations.

Table 1. Characteristics of repetitive DNA sequences studied.

Motif Size	Repeat Motif	Total b. p. length	Secondary Structure Formed
Mononucleotides	$[G/C]_{10}$	10	Triplex
	$[T/A]_{19}$*	19	Triplex
Dinucleotides	$[GT/CA]_{10}$	20	Z-DNA
	$[GT/CA]_{19}$	38	Z-DNA
	$[TC/AG]_{11}$+	22	Triplex
	$[TC/AG]_{20}$	40	Triplex
Tetranucleotide	$[TTCC/AAGG]_9$	36	Triplex
Complex sequences	$[TTTC/AAAG]_3$ * $[TTCC/AAGG]_3[TC/AG]_4$	32	Triplex
	GGGGTGGGGGGAGGGGGGAGGG*	21	G-quadraplex

*Nuclease assay only
+Mutagenesis assay only

S1 Nuclease Assay
Detection of ssDNA in Secondary Structures

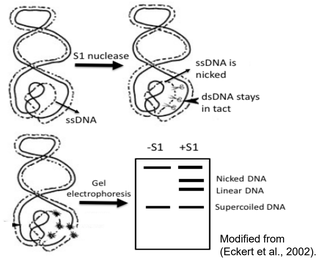


Figure 3. S1 nuclease was added to the plasmid DNA nicking only ssDNA, which detects for non-B secondary DNA structures. Gel electrophoresis analysis shows the amount of nicked, linear, and supercoiled DNA resulting from the assay.

To determine the level of ssDNA secondary structure formation, 1 unit of S1 nuclease was applied to 5 uL of Qiagan midi-prep purified plasmid DNA samples and the reactions were incubated at 37°C for 2, 5, and 10 minutes. The control sample without S1 nuclease was incubated at 37°C for 10 minutes. The plasmids were sequenced to confirm the presence of microsatellites.

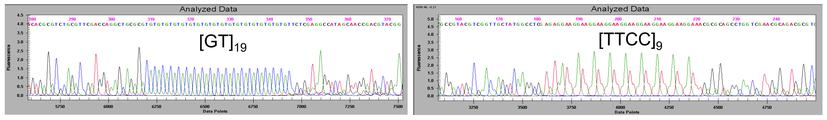


Figure 4. (A) Sequencing data confirms presence of $[GT]_{20}$ dinucleotide repeat in an unknown plasmid. (B) Sequencing data confirms presence of $[TTCC]_9$ tetranucleotide repeat in an unknown plasmid. The primer for the sequencing reaction synthesized the opposite strand in the opposite direction.

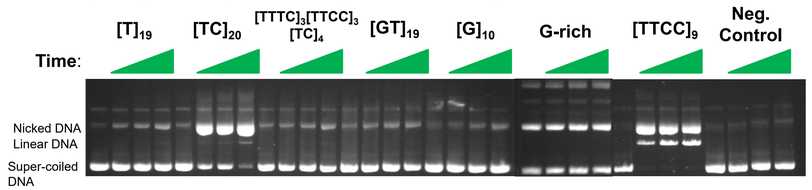


Figure 5. Agarose gel results for plasmids treated with S1 nuclease from 2 – 10 minutes. Lane 1 of each sequence contains no S1 nuclease.

In the nuclease assay, we show that:

- Triplex forming microsatellites ($[TTCC]_9$, $[TC]_{20}$, and $[T]_{19}$) are most sensitive to S1 nuclease (Fig. 6)
- For pure, triplex forming MS, longer sequences seem to be more sensitive to S1 nuclease (Fig. 6)

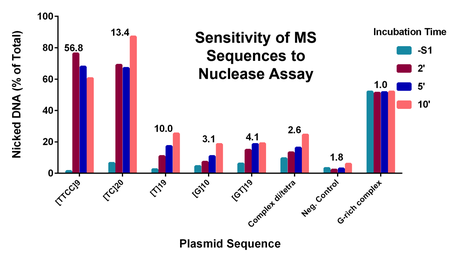


Figure 6. The percentage of nicked DNA compared to the total amount of DNA is shown. The fold difference values between minute 10 and –S1 nuclease are shown above each construct.

Mutagenesis of *rpoB* gene
Verification of MMR deficiency in PP102 *E. coli*

Mismatch repair deficiency in mutL-PP102 *E. coli* was verified by mutational analysis of the *rpoB* gene and compared to that of MMR proficient mutL+ FT334 *E. coli*. Rifampicin selection was used due to its ability to inhibit the *rpoB* gene product, RNA polymerase. Mutants of *rpoB* proliferate in the presence of rifampicin.

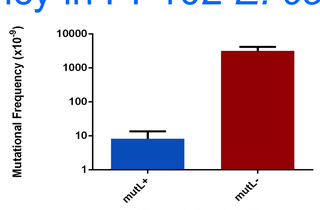


Figure 7. Effect of MMR on mutation frequencies of *rpoB* gene in E. coli. The mutagenesis assay produced a 386 fold difference between the mutL- and mutL+ *E. coli*.

By comparing mutational frequencies of mutL- ($3200x10^{-9}$) to mutlL+ ($8.3x10^{-9}$) *E. coli*, we confirm that the PP102 E. coli is mismatch repair deficient (Fig. 7).

Mutagenesis Assay
Selection for HSV-tk mutants with FUdR

To determine the mutation frequency of seven microsatellite motifs, MMR+ FT334 and MMR- PP102 *E. coli* were electroporated and plated in VBA top agar with 50 ug/ml chloramphenicol and 50 ug/ml chloramphenicol with 40 uM FUdR.

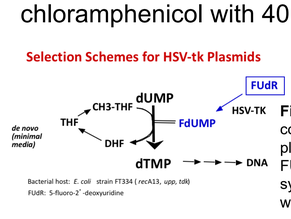


Figure 8. E. coli containing a mutant *tk* plasmid are selected using FUdR because dTTP synthesis can occur without thymidine kinase.

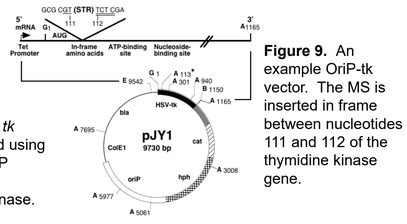


Figure 9. An example OriP-tk vector. The MS is inserted in frame between nucleotides 111 and 112 of the thymidine kinase gene.

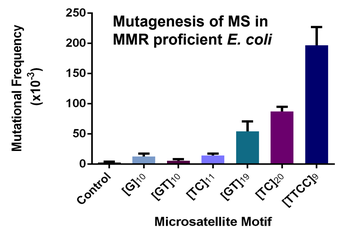


Figure 10. The frequency of mutants produced in FT334 MMR proficient *E. coli* shows the effect of MS length and sequence on MS mutability. Mutational frequency is $FUdR^RChlor^R/Chlor^R$.

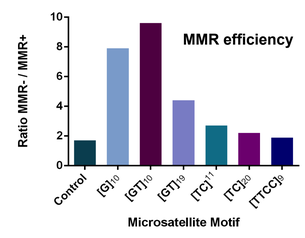


Figure 11. The effect of MMR on MS mutability. Length and sequence contribute the efficacy of repair.

In the mutagenesis assay, we show that:

- $[TTCC]_9$ was the most mutagenic, with a mutational frequency of $200x10^{-5}$, followed by $[TC]_{20}$, $87x10^{-5}$, and $[GT]_{19}$, $54x10^{-5}$ (Fig. 10), but MMR plays little role in repair of the tetranucleotide sequence (Fig. 11)
- In general, longer total repeat lengths were more mutagenic (Fig. 10)
- A comparison of the MMR-/+ ratio for $[GT]_{10}$, 9.6, and $[GT]_{19}$, 2.7, suggests that MMR fixes shorter sequences more effectively (Fig. 11)
- Comparing dinucleotides of similar length, $[GT]_{10}$ and $[TC]_{11}$, $[GT]_{10}$ is fixed more effectively, suggesting a dependence on sequence (Fig. 11)

Overall Conclusions and Future Research

There may be a correlation between nuclease sensitivity and microsatellite mutability, as $[TTCC]_9$ produced the most nicked DNA and was also the most mutagenic. This trend carried through to the $[TC]_{20}$, $[GT]_{19}$, and $[G]_{10}$ motifs. The most mutable sequences are also least likely to be repaired by MMR, which suggests that DNA structure can influence MMR recognition and repair. In the future, Subsequent assays would be performed to determine validity of the MMR efficiency. Additionally, independent mutants would be plated and sequenced for both MMR+ and MMR- plasmids to determine the effect of mismatch repair specifically on microsatellite sequences. Current mutational frequencies are based on the entire *HSVtk* gene (~1000 base pairs) and its inserted artificial MS sequence.

Acknowledgements

The presenter would like to thank Dr. Kristin Eckert, Suzanne Hile, and the Penn State Cancer Institute. The presenter would also like to thank Dr. Kathleen Hill, Dr. Matthew Johnson, Amanda Smith, Amber Cesare, and Taylor Wood from the Center for Science and the Schools, Pennsylvania State University.

References

Eckert, K. A., Yan, G., Hile, S. E. (2002). Mutation Rate and Specificity Analysis of Tetranucleotide Microsatellite DNA alleles in Somatic Human Cells. Molecular Carcinogenesis, 34: 140-150.

Guiblet, W. M., Cremona, M. A., Cechona, M., Harris, R. S., Kejnovska, I., Kejnovsky, E., Eckert, K., Chiaromonte, F., and Makova, K. D. (2017). Long-read sequencing technology indicates genome-wide effects of non-B DNA on polymerization speed and error rate. *BioRXiv,* 1-37. http://dx.doi.org/10.1101/237461

Wierdl, M., Dominska, M., Petes, T.D. (1997). Microsatellite instability in yeast: dependence on the length of the microsatellite. *Genetics*, 46(3): 769-779. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/9215886

The RET for Science Teachers is funded by Penn State Cancer Institute.