

Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Contents lists available at ScienceDirect

Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis

journal homepage: www.elsevier.com/locate/molmut
 Community address: www.elsevier.com/locate/mutres



Short communication

Evidence for a predominant role of oxidative damage in germline mutation in mammals

Arlin Stoltzfus^{*,1}

Center for Advanced Research in Biotechnology, 9600 Gudelsky Drive, Rockville, MD 20850, United States

ARTICLE INFO

Article history:

Received 29 February 2008
 Received in revised form 2 May 2008
 Accepted 16 May 2008
 Available online 25 May 2008

Keywords:

Mutation
 Free radicals
 Oxidative damage
 DNA repair
 Population genetics

ABSTRACT

Spontaneous copying errors in replication often are assumed to be the main source of germline mutations in humans and other mammals. However, when laboratory data on context-dependent patterns of oxidative DNA damage are compared with patterns of mutation inferred from mammalian sequence evolution, the strength of the correlation suggests that damage is the main source of mutations. Analysis of damage susceptibility holds promise for improving models of mutational specificity.

Published by Elsevier B.V.

Mutations in genomic DNA may arise by several pathways distinguished by the initiating event: a spontaneous copying error, incorporation of a damaged nucleotide precursor, or damage to DNA [1]. Oxidative damage is thought to play a major role in somatic mutation in the context of cancer, ageing, and various diseases [2]. However, in the case of germline mutation – the process underlying evolutionary changes and heritable population variation – the main source of mutations is unknown. Since Watson and Crick proposed their tautomeric nucleotide theory of mutation, it has been a widespread assumption that spontaneous copying errors are the main source of base substitution mutations (e.g., Section 9.1 of Ref. [3]). Based on the observation that variability (by species) in the apparent rate of germline mutation correlates with the number of generations (or cell divisions) per unit of time, some have argued for spontaneous copying errors during replication as the main source of germline mutations [4,5]. However, the same kind of data also correlate with species-specific metabolic rate [6], an indicator of the rate of production of damage-causing reactive oxygen species.

New systematic data on damage patterns may shed light on this issue. Recently, Greenbaum et al. [7] quantified the site-specific sensitivity of short DNA sequences to cleavage *in vitro* by hydroxyl radical, a major cause of oxidative damage in living tissues. Sensi-

tivity is largely a matter of local sequence context, such that 78% of the variability is explained by a simple model of overlapping triplets (characterized by the average cleavage sensitivity at each of the three positions); 88% of the variability is explained by a model of overlapping quadruplets [7].

To assess the potential contribution of this type of oxidative damage to germline mutation, I compared the context-dependent hydroxyl radical cleavage patterns to apparent context-dependent mutation rates, inferred from evolutionary divergence of mammalian non-coding regions on the assumption that all changes have the same (presumptively neutral) effect on fitness [8]. When data from CpG sites (which have an aberrant mutation pattern as explained below) are excluded, the hydroxyl radical cleavage intensity for the middle nucleotide of a triplet accounts for $R^2 = 50\%$ of the variability in its relative mutation rate (Fig. 1B), a remarkable correlation ($P = 2.4 \times 10^{-5}$).

In interpreting this correlation, it is important to bear in mind that the primary categories of mutation are distinguished by what *initiates* the mutation process [1]. When damage *initiates* mutation, it influences the *context* (location) of the mutation, the timing, and sometimes the direction (what it mutates to). But damage is not mutation. The involvement of repair synthesis means that differences in repair propensities also modulate the location, timing and direction of mutations that arise by a damage-induced pathway. However, DNA damage does not modulate the probabilities of spontaneous copying errors, which by definition occur in the absence of damage. The hydroxyl cleavage data represent an empirical model of how a particular type of damage influences the context

* Tel.: +1 240 314 6208; fax: +1 240 314 6255.

E-mail address: arlin.stoltzfus@nist.gov.

¹ Biochemical Science Division, NIST, 100 Bureau Drive, Gaithersburg, MD, United States.

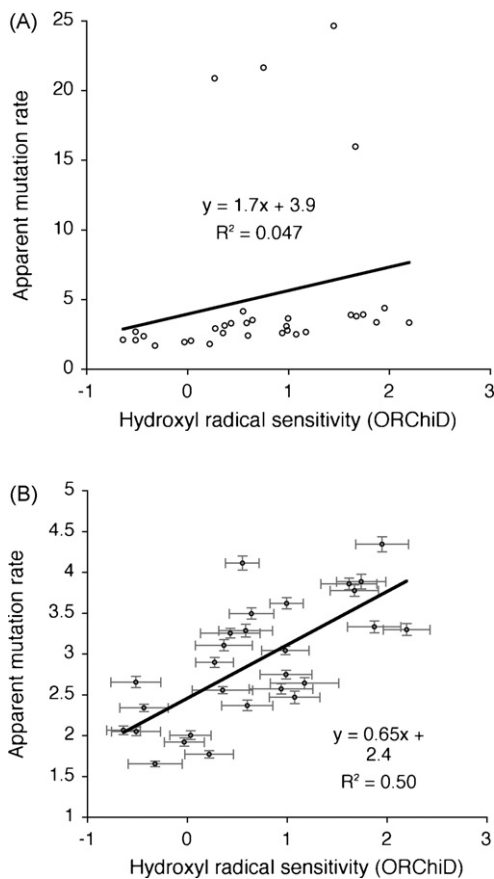


Fig. 1. Apparent mutation rate of the middle nucleotide of a triplet, as a function of its mean hydroxyl radical sensitivity. Each data point represents a double-stranded triplet, e.g., TCC on the plus strand with GGA on the minus strand (error bars, 95% confidence intervals, excluded from plot A for clarity). In (A), all 32 such triplets are shown, while in (B) the 4 triplets with the dinucleotide CG are excluded. The horizontal axis is the relative hydroxyl cleavage intensity for the central nucleotide of a given triplet, summing the strand-specific data for each strand (e.g., TCC and GGT), obtained from the ORChiD database [7]. The vertical axis is the relative apparent mutation rate for the middle nucleotide of a triplet (based on Ref. [8]), summed over all possible mutation pathways (e.g., TCC to TAC, TGC and TTC). This rate is not strand-specific, e.g., the TCC to TAC rate is also the GGA to GTA rate. The P value for the correlation in (B) is 2.4×10^{-5} .

of a mutation. Fig. 1 addresses the case of triplet contexts, relating the inferred mutation rate in a triplet context to the cleavage susceptibility of a triplet context. The most direct interpretation of the resulting correlation is that hydroxyl cleavage determines the location of most non-CpG-related substitution mutations.

However, a less direct interpretation might be possible, given that hydroxyl cleavage patterns reflect fundamental aspects of DNA biochemistry that have many implications. In particular, one may question whether the *apparent* pattern of mutation inferred from sequence divergence [8] truly reflects mutational preferences without any effect of natural selection related to these fundamental biochemical factors. Clearly the hydroxyl data show that some triplets are more damage-prone, which means they have lower fitness, which means that an adaptive evolutionary shift from damage-prone to damage-resistant triplets could be ongoing, even in non-coding regions that are otherwise assumed to be evolving neutrally. This hypothesis could account for the pattern in Fig. 1, but it also predicts an inverse correlation between hydroxyl radical cleavage intensity for a triplet and the apparent mutation rate to that triplet. In fact, as shown in Fig. 2, there is no such correlation.

Thus, the causal linkage implicated in Fig. 1 appears to be mediated by mutation and not by natural selection. It might be naïve to

implicate hydroxyl radicals directly, as the context-dependence of cleavage reflects local differences in solvent-accessible surface of the DNA backbone [9,10], presumably a key factor in many possible kinds of damage, not just hydroxyl-mediated damage. Nevertheless, a specific role for hydroxyl radicals is plausible biologically, given that they are present naturally in mammalian cells, being generated both as a side-product of certain enzyme-catalyzed reactions, and also as a product of naturally occurring radiation.

Given that mammalian DNA is thought to spend much of its time bound up in nucleosome particles, it may seem surprising that the correlation in Fig. 1B is based on hydroxyl cleavage experiments that use naked DNA fragments as substrates. DNA in nucleosomes has many subtle structural differences from naked DNA [11], and a substantial (though not major) fraction of its surface – including a substantial fraction of the backbone, where hydroxyl radicals attack [10] – is buried and thus inaccessible [12]. The fact that cleavage patterns based on naked DNA are relevant to natural damage-induced mutation might indicate simply that, in nature, most of the damage occurs while the DNA is naked. Further cleavage experiments (e.g., on DNA in reconstituted nucleosome particles) could resolve this issue.

Note that the results in Fig. 1B reflect the exclusion of data affected by CpG sites, where a 10-fold higher rate of mutation [8] also reflects oxidative damage by a known mechanism: programmed enzymatic modification results in CpG sites with 5-methyl-cytosine on one or both strands; oxidative deamination of 5-Me-C generates thymine and thus T:G mis-pairs; these T:G mis-pairs are repaired far less efficiently than the U:G mis-pairs (U, Uracil) that result from oxidative deamination of unmodified cytosine.

Thus, given that mutations at CpG sites reflect oxidative damage and account for about 1/4 of nucleotide substitution mutations in humans [13], and given that the remaining 3/4 of mutations show a context-dependence that correlates strongly with the context-dependence of hydroxyl radical cleavage, it would seem that oxidative damage is the predominant source of germline substitution mutations in mammals.

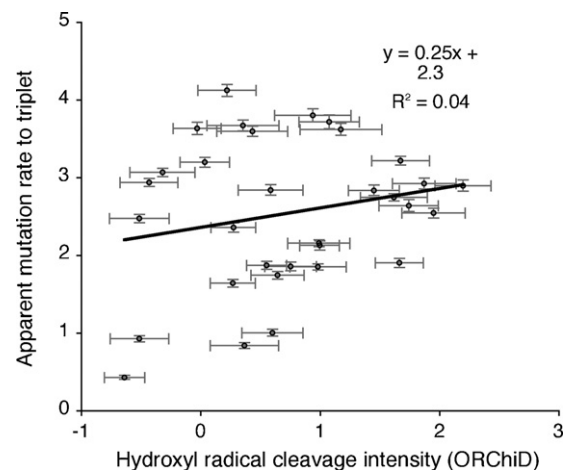


Fig. 2. Apparent mutation rate to a triplet as a function of hydroxyl radical sensitivity. Each data point represents a double-stranded triplet, as in Fig. 1, based on the same sources of data. The horizontal axis is as in Fig. 1. The vertical axis is the relative mutation rate to a given triplet, summing over all three possible mutation pathways (e.g., to TCC from TAC, TGC or TTC), but excluding any pathways from triplets with CpG. Under the hypothesis that the correlation in Fig. 1 is due to selection (see text), the apparent mutation rate is actually a rate of evolution that reflects selection on triplet stability, so that the apparent rate of mutation to triplets should correlate negatively with damage sensitivity. In fact, the correlation is positive and not significant ($P=0.29$).

Acknowledgements

I thank Dick Hwang for providing context-dependent mutation rate results from Ref. [8]. The identification of specific commercial software products in this paper is for the purpose of specifying a protocol, and does not imply a recommendation or endorsement by the National Institute of Standards and Technology.

References

- [1] H. Maki, Origins of spontaneous mutations: specificity and directionality of base-substitution, frameshift, and sequence-substitution mutageneses, *Annu. Rev. Genet.* 36 (2002) 279–303.
- [2] M.D. Evans, M. Dizdaroglu, M.S. Cooke, Oxidative DNA damage and disease: induction, repair and significance, *Mutat. Res.* 567 (2004) 1–61.
- [3] T. Strachan, A.P. Read, *Human Molecular Genetics*, vol. 2, Wiley-Liss, New York, 1999.
- [4] M.A. Baker, R.J. Aitken, Reactive oxygen species in spermatozoa: methods for monitoring and significance for the origins of genetic disease and infertility, *Reprod. Biol. Endocrinol.* 3 (2005) 67.
- [5] W.-H. Li, *Molecular Evolution*, Sinauer, Sunderland, MA, 1997.
- [6] A.P. Martin, S.R. Palumbi, Body size, metabolic rate, generation time, and the molecular clock, *Proc. Natl. Acad. Sci. U.S.A.* 90 (1993) 4087–4091.
- [7] J.A. Greenbaum, B. Pang, T.D. Tullius, Construction of a genome-scale structural map at single-nucleotide resolution, *Genome. Res.* 17 (2007) 947–953.
- [8] D.G. Hwang, P. Green, Bayesian Markov chain Monte Carlo sequence analysis reveals varying neutral substitution patterns in mammalian evolution, *Proc. Natl. Acad. Sci. U.S.A.* 101 (2004) 13994–14001.
- [9] T.D. Tullius, J.A. Greenbaum, Mapping nucleic acid structure by hydroxyl radical cleavage, *Curr. Opin. Chem. Biol.* 9 (2005) 127–134.
- [10] B. Balasubramanian, W.K. Pogozelski, T.D. Tullius, DNA strand breaking by the hydroxyl radical is governed by the accessible surface areas of the hydrogen atoms of the DNA backbone, *Proc. Natl. Acad. Sci. U.S.A.* 95 (1998) 9738–9743.
- [11] T.J. Richmond, C.A. Davey, The structure of DNA in the nucleosome core, *Nature* 423 (2003) 145–150.
- [12] K. Luger, A.W. Mader, R.K. Richmond, D.F. Sargent, T.J. Richmond, Crystal structure of the nucleosome core particle at 2.8 Å resolution, *Nature* 389 (1997) 251–260.
- [13] M.W. Nachman, S.L. Crowell, Estimate of the mutation rate per nucleotide in humans, *Genetics* 156 (2000) 297–304.