

Report

Hypervariable Sites in the mtDNA Control Region Are Mutational Hotspots

Mark Stoneking

Max Planck Institute for Evolutionary Anthropology, Leipzig

Hypervariable sites in human mtDNA are readily identified in evolutionary studies and are usually assumed to represent mutational hotspots. Recently, an alternative hypothesis was proposed that holds that hypervariable sites may instead reflect ancient mtDNA mutations that have been “shuffled” among different lineages via recombination. These hypotheses can be tested by examining the evolutionary rates for sites at which new mtDNA mutations are observed; if hypervariable sites are mutational hotspots, then newly arisen mtDNA mutations should occur preferentially at hypervariable sites. Results of this study show that both germline and somatic mtDNA mutations occur preferentially at hypervariable sites, which supports the view that hypervariable sites are indeed mutational hotspots.

The existence of hypervariable sites (sites that evolve at a rate much faster than average) in the noncoding human mtDNA control region has been well documented via various analyses of human mtDNA variation (Hasegawa et al. 1993; Wakeley 1993; Excoffier and Yang 1999; Meyer et al. 1999); however, it is not known why such hypervariability exists. In the absence of any evidence, it has been generally assumed that hypervariable sites represent mutational hotspots, but recently it was suggested that hypervariable sites may actually reflect ancient mutations that have been redistributed among mtDNA lineages via recombination (Eyre-Walker et al. 1999; Hagelberg et al. 1999). Although recent claims of recombination in human mtDNA (Awadalla et al. 1999; Eyre-Walker et al. 1999; Hagelberg et al. 1999) have not been substantiated in some instances (Macaulay et al. 1999; Hagelberg et al. 2000) and remain controversial (Jorde and Bamshad 2000; Kivisild and Villems 2000; Kumar et al. 2000; Parsons and Irwin 2000), the suggestion that hypervariable sites are not mutational hotspots deserves consideration.

One way to test this hypothesis is to examine the evolutionary rate for sites at which new mutations have

been observed. If hypervariable sites are indeed mutational hotspots, then new mutations should occur preferentially at hypervariable sites. However, if hypervariable sites are not mutational hotspots but, instead, reflect recombination (or some other process), then new mutations would not be expected to occur preferentially at hypervariable sites. In this report, I analyze the evolutionary rate for sites at which new mutations are observed, to test the hypothesis that hypervariable sites are mutational hotspots.

The evolutionary rate for each site in the two hypervariable segments (HV1 and HV2; positions 16024–16383 and 57–372, respectively, numbered according to the reference sequence [Anderson et al. 1981]) of the human mtDNA control region has been estimated elsewhere (Meyer et al. 1999), from patterns of variation in worldwide populations; the overall average evolutionary rate has been found to be .60 for HV1 alone and .54 for HV1 and HV2. Two large family studies reported a total of 12 germline mtDNA mutations for HV1 and HV2 (Parsons et al. 1997; Sigurdardottir et al. 2000), and the average evolutionary rate for the sites at which these 12 mutations occurred is 2.82 (table 1), which is 5.2 times greater than the overall average evolutionary rate for HV1 and HV2. Data are also available for 12 somatic mtDNA mutations in HV1 and HV2 that were observed in tumors but not in control normal tissues from the same individuals (Fliss et al. 2000). The average evolutionary rate for the sites at which these 12

Received July 13, 2000; accepted for publication August 4, 2000; electronically published August 30, 2000.

Address for correspondence and reprints: Dr. Mark Stoneking, Max Planck Institute for Evolutionary Anthropology, Inselstrasse 22, D-04103 Leipzig, Germany. E-mail: stoneking@eva.mpg.de

© 2000 by The American Society of Human Genetics. All rights reserved. 0002-9297/2000/6704-0031\$02.00

Table 1

Sites in HV1 and HV2 at Which Germline, Tumor, and Heteroplasmic Mutations Were Observed and Their Estimated Evolutionary Rate

SITE ^a	RATE ^b	NO. OF MUTATIONS		
		Germline	Tumor	Heteroplasmic
16092	.69	1	0	0
16093	2.98	2	0	13
16104	.47	0	0	2
16111	2.57	1	0	1
16124	.81	0	0	1
16129	4.78	0	0	6
16172	3.76	0	1	1
16183	3	0	1	0
16185	.61	0	0	1
16187	3.67	0	1	0
16189	4.78	0	1	3
16192	3.6	0	0	3
16205	0	0	0	1
16222	.41	0	0	1
16234	.85	0	0	1
16239	.79	0	0	2
16256	2.3	1	0	2
16257	.44	0	0	1
16262	.05	0	0	1
16265	2.18	0	1	0
16292	2.32	0	1	0
16293	3.9	0	0	3
16295	.72	0	0	1
16300	2.1	0	1	0
16301	.3	0	0	1
16309	4.03	0	0	2
16311	4.78	0	0	1
16355	2.16	0	0	1
16362	4.72	0	0	2
16380	0	0	1	0
75	0	0	1	0
94	0	1	0	0
114	.17	0	1	0
150	6.12	0	1	0
153	3.99	1	0	0
185	4.86	1	0	0
189	6.22	1	0	0
195	6.22	0	1	0
207	3.29	2	0	0
234	.69	1	0	0
Total ^c		12 (2.82)	12 (2.86)	51 (2.87)

^a From Comas et al. (1995); Bendall et al. (1996, 1997); Parsons et al. (1997); Wilson et al. (1997); Hühne et al. (1998); Fliss et al. (2000); Sigurdardottir et al. (2000); Tully et al. (2000). Sites are numbered according to the published reference sequence (Anderson et al. 1981).

^b From Meyer et al. (1999).

^c Numbers in parentheses are average rates.

tumor mutations occurred is 2.86 (table 1), which is 5.3 times greater than the overall average evolutionary rate for HV1 and HV2. Finally, a number of studies have documented the existence of heteroplasmy (i.e., the existence of more than one mtDNA type) for sites in the mtDNA control region (Comas et al. 1995; Bendall et

al. 1996, 1997; Wilson et al. 1997; Hühne et al. 1998; Sigurdardottir et al. 2000; Tully et al. 2000). In principle, the heteroplasmic state of an individual could have arisen either via transmission from the mother (i.e., through the germline) or via a new somatic mutation occurring within the individual; most of the studies that have reported heteroplasmy have not attempted to determine which mechanism might be responsible. Therefore, for this analysis, all heteroplasmic mutations were grouped together, resulting in data on 51 mutations (table 1) in HV1 alone. The average evolutionary rate for the sites at which these 51 heteroplasmic mutations were observed is 2.87 (table 1), which is 4.8 times greater than the overall evolutionary rate for HV1.

The distribution of inferred evolutionary rates for all sites in HV1 and HV2 is shown in figure 1, along with the distributions for the sites at which germline, tumor, and heteroplasmic mtDNA mutations were reported. A nonparametric (Mann-Whitney) test indicates that the distribution of evolutionary rates for sites at which germline, tumor, and heteroplasmic mutations were reported are all significantly different from the distribution of evolutionary rates for all sites in HV1 and HV2 ($P < .001$ for all comparisons). Thus, germline, tumor, and heteroplasmic mtDNA mutations all occur preferentially at hypervariable sites in the human mtDNA control region.

The average evolutionary rates for sites at which germline, tumor, and heteroplasmic mutations were observed were remarkably similar (table 1). Moreover, the distribution of evolutionary rates (fig. 1) does not differ among germline, tumor, or heteroplasmic mutations (pairwise Mann-Whitney tests; $P > .05$ for all comparisons). This overall similarity among germline, tumor, and heteroplasmic mutations may indicate that the underlying mutational processes for mtDNA in germ cells and in somatic cells are similar. However, it is also possible that some (or even all) of the supposed germline mutations inferred from family studies are actually somatic mutations that occurred early during the prenatal development of the individual and drifted to homoplasmy and that heteroplasmic mutations may reflect an intermediate stage in this process. The observed close correspondence between the evolutionary rates for somatic mutations (both tumor and heteroplasmic) and for germline mutations would then arise because the supposed germline mutations are in fact somatic mutations. If not all mutations observed in family studies are true germline mutations, then this may also account for the fact that evolutionary rates that have been determined from family studies are higher than those inferred from evolutionary studies (Parsons et al. 1997; Jazin et al. 1998; Sigurdardottir et al. 2000). It would be of interest to know whether individuals who exhibit what appear to

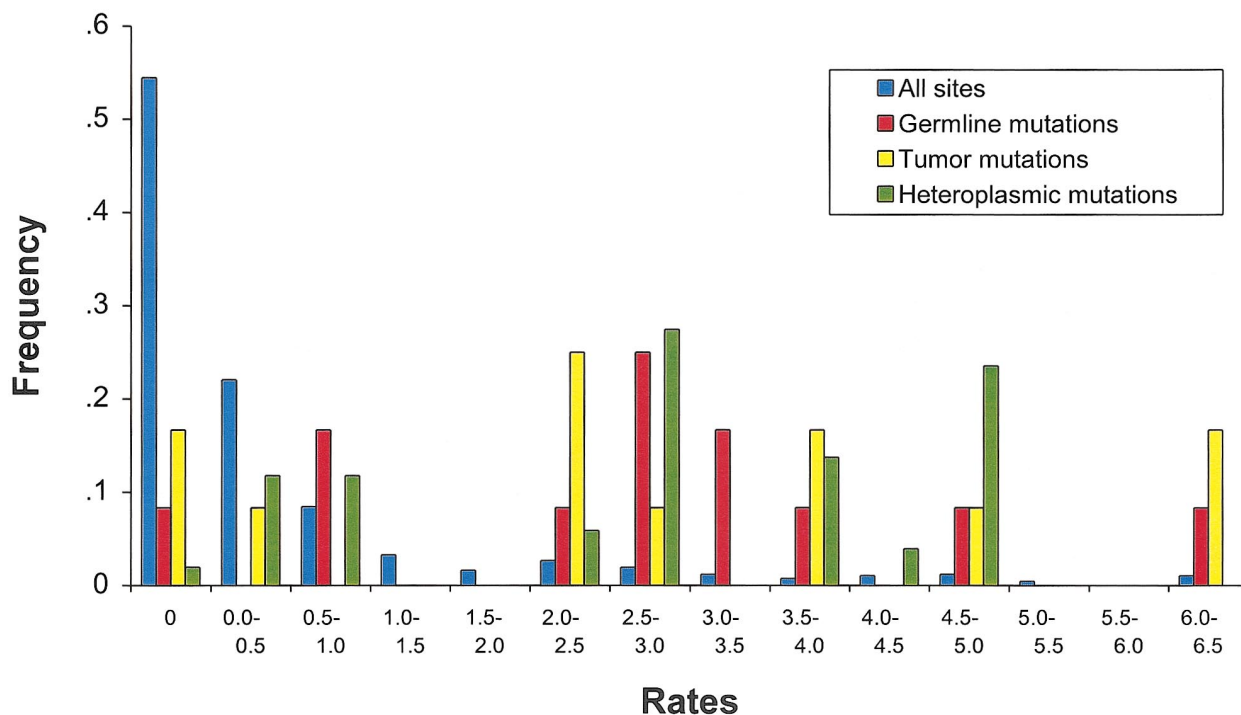


Figure 1 Frequency distribution of inferred evolutionary rates (Meyer et al. 1999) for all sites in HV1 and HV2 and for sites at which germline, tumor, and heteroplasmic mutations were observed.

be newly arisen, germline mtDNA mutations actually transmit the new mutation to their offspring.

In conclusion, both somatic and (supposed) germline mtDNA mutations occur preferentially at hypervariable sites in HV1 and HV2. This result is not consistent with the idea that such hypervariable sites may reflect ancient mutations that have spread to other mtDNA lineages via recombination (Eyre-Walker et al. 1999; Hagelberg et al. 1999). A possible counterargument would be that new mutations actually are rare recombination events between paternal and maternal mtDNA. For this to be true, however, an unusual recombination mechanism must be postulated, because observed mutant mtDNA types differ from the parental type almost exclusively by single mutations, whereas paternal and maternal mtDNA typically differ at several nucleotide positions in the control region. Thus, any supposed mtDNA recombination must instead operate more like gene conversion, involving only short stretches of mtDNA sequence, perhaps even single nucleotide positions. It is difficult to imagine that such a mechanism could be operating in both the germline and in somatic cells (especially tumor cells). Moreover, if recombination involves only single nucleotide positions, then there is no longer the expectation that linkage disequilibrium should be negatively correlated with the physical distance between pairs of sites, a position still held by some

(Awadalla et al. 2000) despite criticisms concerning the measure of linkage disequilibrium used (Jorde and Bamshad 2000; Kumar et al. 2000).

Given these difficulties with a recombinational mechanism for new mutations, it does appear that hypervariable sites are hotspots for both germline and somatic mutations. The challenge now facing mitochondrial geneticists is to determine what features of these hypervariable sites and/or the mtDNA replication machinery cause them to be mutational hotspots.

Acknowledgments

I thank S. Meyer, for providing the evolutionary rates for HV1 and HV2, and S. Pääbo, for useful discussion.

References

- Anderson S, Bankier AT, Barrell BG, de Bruijn MHL, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJH, Staden R, Young IG (1981) Sequence and organization of the human mitochondrial genome. *Nature* 290:457–465
- Awadalla P, Eyre-Walker A, Smith JM (1999) Linkage disequilibrium and recombination in hominid mitochondrial DNA. *Science* 286:2524–2525
- (2000) Questioning evidence for recombination in human mitochondrial DNA. *Science* 288:1931

- Bendall KE, Macaulay VA, Baker JR, Sykes BC (1996) Heteroplasmic point mutations in the human mtDNA control region. *Am J Hum Genet* 59:1276–1287
- Bendall KE, Macaulay VA, Sykes BC (1997) Variable levels of a heteroplasmic point mutation in individuals' hair roots. *Am J Hum Genet* 61:1303–1308
- Comas D, Pääbo S, Bertranpetit J (1995) Heteroplasmy in the control region of human mitochondrial DNA. *Genome Res* 5:89–90
- Excoffier L, Yang Z (1999) Substitution rate variation among sites in mitochondrial hypervariable region I of humans and chimpanzees. *Mol Biol Evol* 16:1357–1368
- Eyre-Walker A, Smith NH, Smith JM (1999) How clonal are human mitochondria? *Proc R Soc Lond B Biol Sci* 266:477–483
- Fliss MS, Usadel H, Caballero OL, Wu L, Buta MR, Eleff SM, Jen J, Sidransky D (2000) Facile detection of mitochondrial DNA mutations in tumors and bodily fluids. *Science* 287:2017–2019
- Hagelberg E, Goldman N, Liò P, Whelan S, Schiefenhövel W, Clegg JB, Bowden DK (1999) Evidence for mitochondrial DNA recombination in a human population of island Melanesia. *Proc R Soc Lond B Biol Sci* 266:485–492
- (2000) Evidence for mitochondrial DNA recombination in a human population of island Melanesia [correction]. *Proc R Soc Lond B Biol Sci* 267:1595–1596
- Hasegawa M, Di Rienzo A, Kocher TD, Wilson AC (1993) Toward a more accurate time scale for the human mitochondrial DNA tree. *J Mol Evol* 37:347–354
- Hühne J, Pfeiffer H, Brinkmann B (1998) Heteroplasmic substitutions in the mitochondrial DNA control region in mother and child samples. *Int J Legal Med* 112:27–30
- Jazin E, Soodyall H, Jalonen P, Lindholm E, Stoneking M, Gyllenstein U (1998) Mitochondrial mutation rate revisited: hot spots and polymorphism. *Nat Genet* 18:109–110
- Jorde L, Bamshad M (2000) Questioning evidence for recombination in human mitochondrial DNA. *Science* 288:1931
- Kivisild T, Villems R (2000) Questioning evidence for recombination in human mitochondrial DNA. *Science* 288:1931
- Kumar S, Hedrick P, Dowling T, Stoneking M (2000) Questioning evidence for recombination in human mitochondrial DNA. *Science* 288:1931
- Macaulay V, Richards M, Sykes B (1999) Mitochondrial DNA recombination—no need to panic. *Proc R Soc Lond B Biol Sci* 266:2037–2039
- Meyer S, Weiss G, von Haeseler A (1999) Pattern of nucleotide substitution and rate heterogeneity in the hypervariable regions I and II of human mtDNA. *Genetics* 152:1103–1110
- Parsons T, Irwin J (2000) Questioning evidence for recombination in human mitochondrial DNA. *Science* 288:1931
- Parsons TJ, Muniec DS, Sullivan K, Woodyatt N, Alliston-Greiner R, Wilson MR, Berry DL, Holland KA, Weedn VW, Gill P, Holland MM (1997) A high observed substitution rate in the human mitochondrial DNA control region. *Nat Genet* 15:363–368
- Sigurdardottir S, Helgason A, Gulcher JR, Stefansson K, Donnelly P (2000) The mutation rate in the human mtDNA control region. *Am J Hum Genet* 66:1599–1609
- Tully LA, Parsons TJ, Steighner RJ, Holland MM, Marino MA, Prenger VL (2000) A sensitive denaturing gradient-gel electrophoresis assay reveals a high frequency of heteroplasmy in hypervariable region 1 of the human mtDNA control region. *Am J Hum Genet* 67:432–443
- Wakeley J (1993) Substitution rate variation among sites in hypervariable region 1 of human mitochondrial DNA. *J Mol Evol* 37:613–623
- Wilson MR, Polansky D, Replogle J, DiZinno JA, Budowle B (1997) A family exhibiting heteroplasmy in the human mitochondrial DNA control region reveals both somatic mosaicism and pronounced segregation of mitotypes. *Hum Genet* 100:167–171