SHORT REPORT

Experimentally observed germline mutations at human micro- and minisatellite loci

Antti Sajantila¹, Matti Lukka² and Ann-Christine Syvänen²

¹*Department of Forensic Medicine, University of Helsinki*

²*Department of Human Molecular Genetics, National Public Health Institute, Helsinki, Finland*

We have analysed close to 30 000 human germline transmission events at five microsatellite loci (D3S1359, HumTH01, HumvWA, HumTPO and HumFES) and four minisatellite loci (D1S80, ApoB, Col2A1 and D17S30). At these loci the mutation rates are similar at the microsatellite and the minisatellite loci, varying from 0.2×10^{-3} to $< 3.3 \times 10^{-3}$ and from 0.5×10^{-3} to 1.5×10^{-3} , respectively. Interestingly, paternal mutations appeared to be dominant at the microsatellite loci, whilst maternal mutations are dominant at minisatellite loci. Based on our data, no unequivocal support for a strict strand-slippage mutation mechanism (gain or loss of a single repeat) was found, although the vast majority of the mutational events were small gains or losses of one to three repeats, and only few unequivocal large gains or losses were observed.

Keywords: microsatellites; minisatellites; mutation rates; human germline mutations; human evolution; paternity testing

Introduction

The micro- and minisatellite DNA loci are powerful genetic markers because they occur with a high frequency in the genomes and are highly polymorphic. Especially the microsatellite loci are commonly used in human molecular genetic analyses because of the availability of dense genetic maps, and since the allelic variation in the number of short sequence repeats of the markers can be unequivocally and rapidly analysed with the aid of PCR-based methods.^{1,2} Despite the widespread use of the micro- and minisatellites as tools in genetic mapping, individual identification and evolutionary studies,^{3,4} the mutation mechanism and mutation rates, which are the reasons for their high degree of

polymorphism, are still poorly understood. For the human microsatellite loci a step-wise replication slippage mutation model has been proposed^{5,6} and a limited number of reports on mutation rates varying from 10^{-2} to 10^{-4} have appeared.^{7–9} Some of the microsatellite loci have been reported to have even higher mutation rates, and complex gene conversion has been suggested as a mechanism in the generation of allelic variants.¹⁰

Mutation rates have been estimated either by indirect population genetic approaches^{11,12} or by experimentally observing mutations in transformed cell lines or in genomic DNA extracted from pedigree samples.^{9,13} An advantage of the direct analysis of DNA from blood samples of pedigree members is that all the observed mutations represent true *in vivo* germ line events and not somatic mutations that frequently occur in cultured cells.^{7,9} A model of the mutational mechanism at micro- or minisatellite loci would be essential for ultimate use of these markers in evolutionary and population genetic studies. Furthermore, knowledge of

Correspondence: Dr Antti Sajantila, Department of Forensic Medicine, PO Box 40, 00014 University of Helsinki, Helsinki, Finland. E-mail: antti.sajantila@helsinki.fi; Tel: +358 9 19127472; Fax: +358 9 19127518
Received 14 May 1998; revised 22 July 1998; accepted 7 August 1998

the actual mutation rates *in vivo* at particular micro- or minisatellite loci has practical value for the interpretation of the genotyping results in forensics and paternity testing.

In this report we present data on mutational events from altogether close to 30 000 parental allele transfers at five microsatellite and four minisatellite loci commonly used in population genetics studies, forensics and paternity testing.

Materials and Methods

The sample material comprised paternity testing samples analysed during 1991–1997 at the National Public Health Institute in Helsinki, Finland. DNA was extracted from 3 µl of EDTA blood by the Chelex®-resin method,¹⁴ and 1–3 µl of the 200 µl extract was used for each DNA amplification. The microsatellites D3S1359,¹⁵ HumTH01,⁵ HumvWA¹⁶ HumTPO¹⁷ and HumFES¹⁸ and the minisatellites D1S80,¹⁹ and ApoB,²⁰ Col2A1²¹ and D17S30²² were analysed. The markers D3S1359, HumTH01 and HumTPO were amplified by triplex PCR and HumFES together with HumvWA by duplex PCR.²³ The remaining markers were amplified in individual reactions at amplification conditions slightly modified from those in the original publications. The amplified alleles were separated using high resolution polyacrylamide gel electrophoresis, visualised by silver staining and genotyped against locus-specific allelic ladders constructed from known alleles as described earlier.^{24,25}

Results and Discussion

The number of parent–child allele transfers analysed at the individual microsatellite and minisatellite loci varied between 301 and 6153. In a total of 29 640 allele

transfers 18 mutations were identified as non-Mendelian inheritance of alleles differing in size (Table 1). All the observed mutations were reconfirmed by analysing new blood samples from the family members involved. In each family with a mutation at one of the loci, genotype data from several other loci were used to calculate the power of evidence.²⁶ Power of evidence for paternity ranging from 99.4% to over 99.999% verified that the analysed samples belonged to true biological families. These calculations were based solely on ‘non-exclusion’ loci.

As can be seen in Table 1, the mutation rates varied from 0.2×10^{-3} to $< 3.3 \times 10^{-3}$ (95% confidence interval: $0.0\text{--}5.0 \times 10^{-3}$). Since mutations not appearing as Mendelian errors in the analysis based on a size difference between alleles remain undetected, the actual mutation rates may be slightly higher, eg by a factor of 1.18 as estimated by Weber and Wong.⁹ In our study the average mutation rates were similar in size (0.5×10^{-3} ; 95% confidence interval: $0.1\text{--}0.9 \times 10^{-3}$) at the minisatellite loci that have repeat units from 16 to 70 base pairs to those at the tetranucleotide repeats markers (0.6×10^{-3} ; 95% confidence interval: $0.2\text{--}1.0 \times 10^{-3}$ (Table 1).

Table 2 presents a more detailed characterisation of the observed mutations. Assuming that the smallest size change is the most probable one, more than half the mutations (10/18) involved a change of one repeat unit. This observation would support the strict step-wise replication slippage mutation model,^{5,6} according to which the majority of the mutational events at microsatellite loci involve gain or loss of one repeat unit. However, in five cases the smallest possible size

Table 1 Mutation rates observed at five microsatellite and four minisatellite loci

Locus	Location	Repeat unit	Allele size range (bp)	H ^a	Number of meioses	Number of mutations	Mutation rate $\square 10^3$ ^b
HumTPO	Chr.2	[AATG]	106–130	0.62	2457	0	<0.4 (0.1–0.6)
D3S1359	Chr. 3	[TCTA][TGTA]	193–265	0.89	4095	5	1.2 (0.1–2.2)
HumTH01	Chr. 11	[AATG]	155–179	0.79	5918	1	0.2 (0.0–0.5)
HumvWA	Chr. 12	[TCTA][TCTG]	134–166	0.82	3684	5	1.4 (0.1–2.5)
HumFES	Chr. 15	[TTTA]	211–234	0.66	301	0	<3.3 (1.6–5.0)
D1S80	Chr. 1	[16 base pairs]	387–723	0.81	6153	3	0.5 (0.0–1.0)
ApoB	Chr. 2	[30 base pairs]	570–900	0.81	2180	1	0.5 (0.2–0.6)
Col2A1	Chr. 12	[31 or 34 base pairs]	510–801	0.70	843	1	1.2 (0.5–1.7)
D17S30	Chr. 17	[70 base pairs]	170–980	0.85	4009	2	0.5 (0.0–1.1)
TOTAL							
Microsatellites				0.76	16 455	11	0.6 (0.2–1.0)
Minisatellites				0.79	13 185	7	0.5 (0.1–0.9)
All loci				0.77	29 640	18	0.6 (0.3–0.8)

^aH = heterozygosity value in the Finnish population sample studied here; ^b95% confidence intervals are given in parenthesis. For the markers where no mutations were observed, the mutation rate was approximated to $<1/\text{number of meioses}$. The 95% confidence intervals were calculated accordingly.

alteration was a change of 2–3 repeat units, and moreover, in five of the cases with a change in one repeat unit, the second smallest mutation would have a change of 2–3 repeats. Thus, an alternative interpretation of the data is that more than half of all the mutations, and about 30% of the microsatellites, are due to gain or loss of 2–3 repeats, which is not in accordance with a strict step-wise replication slippage mutation model. In three cases (17%) the observed mutation unequivocally involved the deletion or insertion of more than five repeat units. As can be seen in Table 2, we observed a similar number of gains and losses of repeat units. On this point our data agree with observations at a hypermutable locus by Talbot *et al*⁸ and do not support the suggested mutation mechanism of directional evolution towards longer repeats.^{27,28} Our data do not support the loss of heterozygosity in genotypes with a large size difference between the alleles,²⁷ although our material contained four children with homozygous genotype, but in only three of these cases were the parents heterozygous. Interestingly, all the mutations, with one exception where the phase is uncertain, at the minisatellite loci can be explained as maternal ones and all the mutations at microsatellite loci can be explained as paternal ones. The latter observation is in agreement with a finding by Weber and Wong,⁹ who report a significant preference for paternal germ line mutations at microsatellite loci. As seen in Table 2, our results do not show a mutational bias in favour of longer alleles²⁸ nor loss of hetero-

zygosity in genotypes with a large size difference between the alleles.²⁷

Knowledge on mutational mechanism and rates *in vivo* are essential when trying to address population genetic questions, such as population history, migration or admixture, where the utility of micro- and minisatellite loci has been hampered by the difficulties in subtracting the mutation process from the population demography and population history. Furthermore, the data presented here have practical consequences for paternity testing. First, in cases with exclusion of a paternity by a single micro- or minisatellite locus, the data from experimentally observed mutation rates facilitate the statistical assessment of the finding. Second, the data can be utilised in multiple exclusion cases, where the true biological father could also be closely related to the alleged father. However, we would like to stress that it is desirable to accomplish further studies with larger data sets in order to derive better estimates of the mutation rates. Larger data sets are also essential to elucidate further the use of the knowledge of mutational behaviour of microsatellite loci in paternity testing as well as evolutionary studies.

To conclude, by analysing thousands of human germ line transmission events, we found that the mutation rates are similar at microsatellite and at minisatellite loci studied here. Paternal mutations appear to be dominant at the microsatellite loci, whilst maternal mutations are more common at minisatellite loci. Based

Table 2 Characteristics of the observed micro- and mini-satellite mutations

Locus	Genotype ^a			Mutation ^b	Origin ^b
	F	M	C		
D3S1359	6–7	2–2	2–5	–1 (–2)	Paternal
D3S1359	5–12	4–5	5–10	–2 (+5, +6)	Paternal (Maternal)
D3S1359	6–6	5–6	6–11	+5 (+6)	Paternal (Maternal)
D3S1359	3–3	2–12	2–2	–1	Paternal
D3S1359	4–7	2–6	2–8	+1 (+4)	Paternal
HumTH01	8–10	6–7	5–7	–3 (–5)	Paternal
HumvWA	18–19	17–18	17–20	+1 (+2)	Paternal
HumvWA	15–17	17–20	18–20	+1 (+3)	Paternal
HumvWA	17–19	18–18	16–18	–1 (–3)	Paternal
HumvWA	15–18	14–17	17–17	–1 (+2)	Paternal
HumvWA	17–19	15–17	15–18	+/-1	Paternal
D1S80	24–25	24–31	25–28	–3 (+4)	Maternal
D1S80	18–24	24–31	18–18	–6 (–13)	Maternal
D1S80	24–24	18–24	21–24	+3 (–3)	Maternal (Paternal)
ApoB	34–36	36–48	36–38	+2 (–10)	Maternal (Paternal)
Col2A1	19–23	19–23	23–26	+7	Phase uncertain
D17S30	3–9	4–6	5–9	+1 (–1)	Maternal
D17S30	2–3	4–11	3–3	–1 (–8)	Maternal

^aGenotype nomenclature is based on the number of repeat units. F = father, M = mother, C = child; ^bBased on the assumption that the smallest change of repeat units in the mutation is the most probable one. The other possible mutations are shown in parenthesis.

on our data, no unequivocal support for a strict strand-slippage mutation mechanism involving gain or loss of a single repeat was found, but the vast majority of the mutational events involved small gains or losses of 1–3 repeats, and only a few large unequivocal gains or losses were observed.

Acknowledgements

We thank the personnel of the Paternity Laboratory at the National Public Health Institute for excellent technical assistance. We are grateful to Dr Matti Kataja for his interest and statistical help.

References

- 1 Tautz D: Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Res* 1989; **17**: 6463–6471.
- 2 Weber JL, May PE: Abundant class of human DNA polymorphism which can be typed using the polymerase chain reaction. *Am J Hum Genet* 1989; **44**: 388–396.
- 3 Goldstein DB, Ruiz Linares A, Cavalli-Sforza LL, Feldman MW: Genetic absolute dating based on microsatellites and the origin of modern humans. *Proc Natl Acad Sci USA* 1995; **92**: 6723–6727.
- 4 Murray JC, Buetow KH, Weber JL *et al*: A comprehensive human linkage map with centimorgan density. *Science* 1994; **265**: 2049–2054.
- 5 Edwards A, Holly AH, Jin L, Caskey CT, Chakraborty R: Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. *Genomics* 1992; **12**: 241–253.
- 6 Schlötterer C, Tautz D: Slippage synthesis of simple sequence DNA. *Nucleic Acids Res* 1992; **20**: 211–215.
- 7 Banchs I, Bosch A, Guimerá J, Lázaro C, Puig A, Estivill X: New alleles at microsatellite loci in CEPH families mainly arise from somatic mutations in the lymphoblastoid cell lines. *Hum Mutat* 1994; **3**: 365–372.
- 8 Talbot CC Jr, Avramopoulos D, Gerken S *et al*: The tetranucleotide repeat polymorphism D21S1245 demonstrates hypermutability in germline and somatic cells. *Hum Mol Genet* 1995; **4**: 1193–1199.
- 9 Weber JL, Wong C: Mutation of human short tandem repeats. *Hum Mol Genet* 1993; **2**: 1123–1128.
- 10 Jeffreys AJ, Tamaki K, MacLeod A, Monckton DG, Neil DL, Armour JAL: Complex gene conversion events in germline mutation at human minisatellites. *Nature Genet* 1994; **6**: 136–145.
- 11 Chakraborty R, Kimmel M, Stivers DN, Davison LJ, Deka R: Relative mutation rates at di-, tri-, and tetranucleotide microsatellite loci. *Proc Natl Acad Sci USA* 1997; **94**: 1041–1046.
- 12 Jin L, Macaubas C, Hallmayer J, Kimura A, Mignot E: Mutation rate varies among alleles at a microsatellite locus: Phylogenetic evidence. *Proc Natl Acad Sci USA* 1996; **93**: 15285–15288.
- 13 Weissenbach J, Gyapay G, Dib C *et al*: A second-generation linkage map of the human genome. *Nature* 1992; **359**: 794–801.
- 14 Walsh PS, Metzger DA, Higuchi R: Chelex-100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *Biotechniques* 1991; **10**: 506–513.
- 15 Li H, Schmidt L, Wei MH *et al*: Three tetranucleotide polymorphisms for loci: D3S1352; D3S1358; D3S1359. *Hum Mol Genet* 1993; **2**: 1327.
- 16 Manosco DJ, Tuley EA, Westfield LA *et al*: Structure for the gene for human von Willebrand factor. *J Biol Chem* 1989; **264**: 19514–19527.
- 17 Anker R, Steinbruek T, Donis-Keller H: Tetranucleotide repeat polymorphism at the human thyroid peroxidase (hTPO) locus. *Hum Mol Genet* 1992; **1**: 137.
- 18 Polymeropoulos MH, Rath DS, Xiao H, Merrill CR: Tetranucleotide repeat polymorphism at the human c-fes/fps proto-oncogene (FES). *Nucleic Acids Res* 1991; **19**: 4018.
- 19 Kasai K, Nakamura Y, White R: Amplification of a variable number of tandem repeats (VNTR) locus (pMCT118) by the polymerase chain reaction (PCR) and its application to forensic science. *J Forensic Sci* 1990; **35**: 1196–1200.
- 20 Boerwinkle E, Xiong W, Fourest E, Chan L: Rapid typing of tandemly repeated hypervariable loci by the polymerase chain reaction: Application to the apolipoprotein B 3' hypervariable region. *Proc Natl Acad Sci USA* 1989; **86**: 212–216.
- 21 Berg ES, Olaisen B: Characterization of the COL2A1 VNTR polymorphism. *Genomics* 1993; **16**: 350–354.
- 22 Wolff RK, Nakamura Y, White R: Molecular characterization of a spontaneously generated new allele at a VNTR locus: No exchange of flanking DNA sequence. *Genomics* 1988; **3**: 347–351.
- 23 Rostedt I, Lalu K, Lukka M, Sajantila A: Genotyping of five short tandem repeat loci via triplex and duplex PCR. *Forensic Sci Int* 1996; **82**: 217–226.
- 24 Sajantila A, Puomilahti S, Johnsson V, Ehnholm C: Amplification of reproducible allele markers for amplified fragment length polymorphism (Amp-FLP) analysis. *Biotechniques* 1992; **12**: 16–22.
- 25 Sajantila A, Lukka M: Improved separation of PCR amplified VNTR alleles by a vertical polyacrylamide gel electrophoresis. *Int J Legal Med* 1993; **105**: 355–359.
- 26 Kataja M: Simulation in paternity analysis. PhD thesis. Otapaino, Espoo, Finland, 1975.
- 27 Amos W, Sawcer SJ, Feakes RW, Rubinsztein DC: Microsatellites show mutational bias and heterozygote instability. *Nat Genet* 1996; **13**: 390–391.
- 28 Primmer CR, Ellegren H, Saino N, Møller AP: Directional evolution in germline microsatellite mutations. *Nat Genet* 1996; **13**: 391–393.