# DNA PROFILING: A GENETIC STUDY OF TWO VNTR LOCI IN THE EAST MIDLANDS

V.J. Stinton and S.S.Mastana

Human Genetics Laboratory, Department of Human Sciences,
Loughborough University of Technology, Loughborough, Leicestershire, England.

**Introduction:** DNA fingerprinting developed by Jeffreys, Thein and Wilson(1985) is the most discriminating technique used to identify an individual. It has proved to be a powerful tool in the legal world aiding in the identification of criminals and solving paternity disputes. In recent years DNA fingerprinting has been increasingly replaced by DNA profiling. A number of problems with the application of this powerful tool have recently been highlighted in United States courts(Joyce, 1990; Bown, 1993). One particular concern in the application of genetic typing is the use of general population databases to determine DNA profile probabilities. It has already been established that there are racial differences in allele frequencies(Gill et al, 1991; Buffrey et al, 1991). This had led to the construction of databases for different racial groups. However, it has recently been proposed that minisatellite allele frequencies are different within sub-populations of a major population(Lander, 1989; Lewontin and Hartl, 1991). Population stratification would have important implications for forensic and paternity work. General databases used to calculate forensic and paternity probabilities may not accurately represent sub-population diversity. To date, studies by Budowle and Monson(1994) and Budowle et al(1994) indicate that DNA profile probabilities may not be affected by using general databases. However, further research is required to clarify and support or disprove these findings.

**Objectives:** The main aim of this study was to establish the Federal Bureau of Investigation(FBI) fixed bin allele frequencies of MS1(D1S7) and MS43a(D12S11) minisatellite loci in the Caucasian East Midlands population. The allele frequencies of MS43a were compared with other UK Caucasian sub-populations, the Forensic Science Service(FSS) Caucasian and racial databases and other Caucasian populations world-wide to determine if the populations were heterogeneous. The heterozygosity, paternity(exclusion) and forensic(inclusion) calculations were conducted on the East Midlands data to determine the efficiency of using MS1 and MS43a probes in this population.

**Population and sample size:** Caucasian East Midlands. N:102

**Methods:** DNA extracted from random blood donor samples was digested with restriction enzyme HinfI and VNTR profile was established using standard Southern blotting methodology. NICE MS1 and MS43a probes (Cellmark Diagnostics) were used in this study along with MW100 probe as a Molecular Weight Marker. The log molecular weights of MS1 and MS43a alleles were binned into FBI fixed bins(Budowle and Monson, 1994). The allele frequency in each bin was calculated and used for comparative work.

**Analysis of results:** Comparisons of population FBI bin frequencies in different populations was done using heterogeneity chi square analysis. The heterozygosities of MS1 and MS43a were calculated. The probability of match(PM), discrimination probability(DP) and probability of exclusion(PE) using MS1 and MS43a independently and combined using the Caucasian East Midlands population database.

Advances in Forensic Haemogenetics 6 (c) Springer-Verlag Berlin Heidelberg

**Results and discussion:** Significant heterogeneity chi-square values for MS43a and MS1 were found in a large number of FBI bins in UK Caucasian, world Caucasian (US, Norway and Spanish) and UK racial population comparisons. MS1 studies are relatively limited and require further studies to evaluate subpopulation heterogeneity. However, of main interest is that the study's data indicates that population stratification may exist within the United Kingdom at the MS43a minisatellite locus (Table 1). The large number of significant population heterogeneity $\chi$ 2 values in comparisons of the East Midlands, Newcastle and FSS Caucasian databases highlight this. This emphasises that a general database such as the FSS may not represent the MS43a allele frequencies found in the East Midlands and Newcastle. The study's results also challenge the findings of Budowle and Monson(1994) and Budowle et al(1994) who claimed from their data that sub-population differences were small and hence irrelevant. The implies that geographical databases may be required in the UK as the present general database may underestimate/overestimate DNA probabilities. However, further investigation would be required to support the present study's findings.

**Table 1: Population heterogeneity chi-square($\chi$2) values of five population comparisons of FBI fixed bin MS43a allele frequencies.**

| | $A_1$ | $B_2$ | $C_3$ | $D_4$ | $E_5$ |
|---|---|---|---|---|---|
| FBI Fixed Bins (base Pairs) | Heterogeneity $\chi$2 | Heterogeneity $\chi$2 | Heterogeneity $\chi$2 | Heterogeneity $\chi$2 | Heterogeneity $\chi$2 |
| 3330 - 3674 | | 14.03*** | 37.39*** | 45.65*** | 40.04*** |
| 3675 - 3979 | | 3.99* | 15.60*** | | 11.88*** |
| 3980 - 4323 | | 3.99* | 16.10*** | 75.15*** | 23.28*** |
| 4324 - 4821 | 3.31 | 0.28 | 4.51 | 8.95 | 0.27 |
| 4822 - 5219 | 0.38 | 0.15 | 0.44 | 5.93 | 3.22 |
| 5220 - 5685 | 1.8 | 6.10* | 6.54* | 33.15*** | 5.03* |
| 5686 - 6368 | 6.08* | 9.97** | 40.48*** | 140.16*** | 14.24** |
| 6369 - 7241 | 6.21* | 48.39*** | 121.16*** | 228.86*** | 27.22*** |
| 7242 - 8452 | 0.01 | 2.37 | 2.71 | 40.79*** | 13.21*** |
| 8453 - 10093 | 16.07*** | 2.37 | 47.12*** | 66.21*** | 64.58*** |
| 10094 - 11368 | 9.07** | 4.76* | 57.53*** | 120.21*** | 36.56*** |
| 11369 - 12829 | 0.36 | | 24.39*** | 49.83*** | |
| 12830 - | 18.07*** | | 43.42*** | 119.08*** | |

* Significant at 5% level; ** Significant at 1% level; *** Significant at 0.1% level.

Comparative analysis:

$_1$East Midlands\Newcastle (Newcastle data obtained from Human Genetics Department , Newcastle university.)

$_2$East Midlands/Caucasian FSS(FSS data from Forensic Science Service, Birmingham)

$_3$East Midlands/Newcastle/Caucasian FSS.

$_4$East Midlands/Newcastle/Caucasian FSS/US Caucasians/Norwegians/Spanish (US Caucasian, Norwegian, and Spanish data from Budowle and Monson(1994)).

$_5$East Midlands/Afro Caribbean/Asian(Afro Caribbean and Asian data from FSS)

Advances in Forensic Haemogenetics 6 (c) Springer-Verlag Berlin Heidelberg

The heterozygosity of MS1 was 94.3% and MS43a 90.6%. These relatively high heterozygosities implies that both probes are suitable for forensic and paternity work. This is supported by the low PM(0.0016), high DP(0.9984) and high PE(0.9559) when the probes are combined(Table 2).

**Table 2 - Table showing forensic and paternity probabilities of MS1 and MS43a using the Caucasian East Midlands database.**

| Probe | Forensic Inferences | | Paternity Inferences |
|---|---|---|---|
| | Probability of Match | Discrimination Probability | Probability of Exclusion |
| MS1 | 0.0322 | 0.9678 | 0.8096 |
| MS43a | 0.0499 | 0.9501 | 0.7692 |
| MS1 + MS43a | 0.0016 | 0.9984 | 0.9559 |

**Conclusions:** This study has found that both MS1 and MS43a are efficient markers for forensic and paternity work. The study has shown that Caucasian population heterogeneity may exist within the United Kingdom at the MS43a locus. MS1 data also showed some significant differences but there are not many populations studied for this locus. This implies that present probability calculations for forensic and paternity work using general Caucasian population databases should be used with caution for MS43a probe. However, further research is required to clarify the findings.

Acknowledgements: Our sincere thanks to Mrs Alice Pacynko for technical help and Drs. S.Papiha and P.Gill for comparative data.

**References:**

Bown W (1993) DNA fingerprinting back in the dock. New Scientist 137:14-15.
Budowle B et al (1994) Evaluation of Hinf 1 generated VNTR profile frequencies determined using various ethnic databases. Journal of Forensic Science 39:988-1008.
Budowle B and Monson K L (1994) Greater differences in forensic DNA profile frequencies from racial groups than from ethnic subgroups. Clinica Chimica Acta 228:3-18.
Buffrey C et al (1991) Allele frequency distributions of four hypervariable number tandem repeat(VNTR) loci in the London area. Forensic Science International 52:53-64.
Gill P et al (1991) Population genetics of four hypervariable loci. International Journal of Legal Medicine 104:221-227.
Jeffreys A J, Wilson V and Thein S L (1985) Hypervariable "minisatellite" regions in human DNA. Nature 316:76 -79.
Joyce C (1990) High profile DNA in court again. New Scientist 127:24 -25.
Lander E S (1989) DNA fingerprinting on trial. Nature 339:501 - 505.
Lewontin R C and Hartl D L (1991) Population genetics in forensic DNA typing. Science 254:1745-1750.

Advances in Forensic Haemogenetics 6 (c) Springer-Verlag Berlin Heidelberg