CROATIAN MEDICAL JOURNAL
CMJ
42(3):310-314,2001

FORENSIC SCIENCES

Pentanucleotide Short Tandem Repeat Locus DXYS156 Displays Different Patterns of Variations in Human Populations

Christian Kersting, Carsten Hohoff, Burkhard Rolf*, Bernd Brinkmann

Institut für Rechtsmedizin, Universität Münster, Münster, Germany

Aim. To establish a database for the pentameric short tandem repeat locus DXYS156 from worldwide populations for routine genotyping in forensic identity testing and evolutionary biology.

Methods. Using polymerase chain reaction with a newly designed primer pair, we analyzed 1,408 male and female samples from 28 populations representing four major geographic groups.

Results. We observed 11 different alleles, which we sequenced and used to construct an allelic ladder.

Conclusion. DXYS156 displays a contrasting pattern of X-linked and Y-linked variation among geographic regions, and between X and Y chromosomes. This complex allele distribution may be forensically useful for the ethnic differentiation of unknown stains.

Key words: *DNA; forensic medicine; genetics, population; genotype; polymerase chain reaction; microsatellite repeats*

Polymorphic short tandem repeats (STRs) are widely used in forensic and population genetic applications. Recently published data on autosomal and Y-linked minisatellites and microsatellites, mitochondrial DNA (mtDNA) (1-5), and single nucleotide polymorphisms (SNPs) have shown that there are great genetic diversities within African populations (6). Many authors have interpreted these diversities as the result of a larger population size or simply higher evolutionary age of the population.

According to certain data on DYS413, DXS8174, DXS8175 (7), and DXYS156 (8) loci, it seems that "Total Replacement" model or "Out of Africa" hypothesis – *Homo sapiens* emerging in Africa and replacing other rivaling hominids – may undergo certain changes. These loci revealed higher genetic diversity in Asian samples and are obviously at odds with the "Out of Africa" theory. There is still no explanation for this contradiction. To increase available information on this issue, we analyzed a global set of samples at both the DXYS156-X and -Y locus.

The pentanucleotide microsatellite DXYS156 was first described by Chen et al (9). It comprises two homologous regions on the short arm of the Y (Yp) and the long arm of the X chromosome (Xq) (10). This homology is the result of an X-Y transposition with following Yp inversion that dates to the time of emergence of *Homo*, about 3-4 million years ago (11). The two regions do not recombine and thus the Y-hosted portion is subject to male-specific mutation rates only (12). For some populations, Y-chromosomal STR haplotypes that include DXYS156 have been reported (13,14).

Here we present DXYS156 allele frequency data in a sample of 1,424 individuals (583 women and 841 men) from 12 geographic regions. The alleles were analyzed by polyacrylamide gel electrophoresis (PAGE) and capillary gel electrophoresis (CE) after polymerase chain reaction (PCR). We observed the occurrence of 11 different alleles, which extends the findings of Karafet et al (8): they spread over the whole range of 4-15 repeats of the $(TAAAA)_n$ motif. Sequencing of the alleles 4-14 revealed no variability in the repeat sequence of the locus or flanking regions. The distribution of alleles at the X-locus was in accordance with the Hardy-Weinberg law, as verified in a simulated exact-test.

Material and Methods

Population Samples

The samples from saliva or blood were acquired from 28 populations from 4 major geographic groups: Africans (Khoisan [Giricu, Kavango, Mbukushu, Mbunza, Sambiu, San], East Africa [Kenya, Tanzania, Uganda], West Africa [Angola, Benin, Ivory Coast, Ghana, Guinea, Cameroon, Congo, Namibia, Nigeria, Zambia, Sierra Leone, South Africa, Togo, West Indies]), Australasians (Australian Aborigines, Papua New Guinea) (15), Asians (from Han Chinese from Shenyang, Japanese from Shiga, Khalkh Mongolians), and Caucasoids (from Germany, Italy, Morocco, and Turkey). Details on sampling are provided elsewhere (15-18).

Amplification and Detection

Genomic DNA was obtained from blood and saliva swabs by standard extraction methods (19). According to Chen et al (9), PCR was performed with two primer sets for a subset of the samples and with newly designed primers for all the samples:

5′-GGTGAGAATCAATTCAAGAACTCA 3′

5′-TCCTTGGCTAGGTATATTCC 3′

These primers produced smaller amplicons than the original ones (Table 1, Fig. 1) and showed a more robust and reproducible amplification pattern (data not shown). The amplicons derived from

*Present address: Institut für Rechtsmedizin, Universität München, Frauenlobstraße 7a, D-80337 München, Germany

both pairs of primers were thoroughly compared. To confirm the precision of typing by native PAGE, we performed several comparisons with denaturing capillary gel electrophoresis, using the ABI PRISM 310 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA). Finally, all 11 alleles observed (4-14) were isolated from native gels, as described elsewhere (20), and directly sequenced on both strands with the Taq Dye Deoxy Terminator Cycle Sequencing Kit (Applied Biosystems).

Chromosome Assignment

In 1994, Chen et al (21) reported that alleles 4, 7, 8, 9, and 10 (designation according to the recommendations of the International Society for Forensic Genetics [ISFG], Table 1) could be found only on the X chromosome, and alleles 11 and 12 only on the Y chromosome. We detected further alleles: 5 and 6 located on the X chromosome, and 13 and 14 located only on the Y chromosome. Unfortunately, 9 out of 1,424 (583 women and 841 men) individuals showed a discordant distribution: three women had two 'Y-specific' alleles and six men had two 'X-specific' alleles. Three other individuals had more than two alleles in total. Because of the statistically insignificant amount of exceptions (0.64%), we did not take them into account.

HWE Analysis (version 3.2, C. Puers, Münster) was used to calculate statistical parameters.

Results and Discussion

Structure and Allele Mapping of DXYS156

To verify the PAGE results, we sequenced one representative of each allele (range 4-14), using the original and the new set of primers (see methods). These data showed that sequenced amplicons differed only in the number of AAAAT repeats, with no exception. No differences were found in the flanking regions.

The alleles 4, 5, 6, 7, 8, 9, and 10 were mainly found in women (Table 2), whereas alleles 11, 12, 13, and 14 were mainly found on the Y chromosomes (Table 3). However, in 9 out of 1,408 individuals, overlapping of alleles was discovered: there were men with two "X-specific" or "Y-specific" alleles and women with one or two "Y-specific" alleles. We considered it insignificant (0.63%). The reasons for this overlap might be XYY and X0 genotypes or even other chromosomal aberrations, which were not the subject of our investigation. Nevertheless, in certain cases it might have affected the assignment of alleles to chromosomes. In fact, we found two individuals showing 2 bands and one individual with even 4 bands. Since a single band in PAGE usually indicates homozygosity in women, the total number of chromosomes examined could have been affected by X0 genotype frequency, too.

GGTGAGAATC AATTCAAGAACTCA

CAGATACCAAGGTGAGAATC AATTCAAGAACTCA ACCAC TTTTACAAAA GCTGCAAAAATAA

TAAAA TAAAA TAAAA TAAAA TAAAA TAAAA TAAAA TAAAA TAAAA TAAAA TAAAA TAAAA

TACTTA **GGAATATACCTAGCCAAGGA** GGCAAAAGACCACTAC

Figure 1. Sequence of DXYS156 (allele 12). Original primers by Chen et al (1994, ref. 9) are underlined and novel primers (this study) are shown in bold and shaded in grey. The 12 TAAAA repeats are shaded in light grey.

Table 1. Comparison of DXYS156 nomenclatures and PCR product sizes

Allele designation		Fragment size [bp]with primers	
Chen et al (1994)[a]	ISFG[b] conform	old	novel
X1	4	125	99
–	5	130	104
X2	6	135	109
X3	7	140	114
X4	8	145	119
X5	9	150	124
Y6	10	155	129
Y7	11	160	134
–	12	165	139
–	13	170	144
–	14	175	149

[a]Ref. 9.
[b]International Society for Forensic Genetics.

The possibility of error in determining an individual's sex is also a factor that cannot be completely eliminated. To exclude this source of error, we have determined the gender by analyzing amelogenin locus (22) in cases of discordant allele distribution (data not shown).

Allele Distribution

For the X locus, we found that allele 7 had the highest frequency in all studied populations. Africans showed almost Gaussian-like distribution, with a second maximum at allele 4, whereas all other populations had nearly unimodal distribution. For example,

Table 2. Allele frequencies of DXYS156-X in women

Allele	China (n = 79)	Germany (n = 58)	Japan (n = 61)	Khoisan (n = 55)	Morocco (n = 44)	Ovambo (n = 54)	Papua (n = 54)	Turkey (n = 119)	West Africa (n = 181)
4	0	0	0	0.05	0	0.32	0	0.01	0.31
5	0	0	0.01	0	0	0	0.01	0	0
6	0	0	0.01	0	0	0	0	0	0
7	1	0.80	0.91	0.05	0.76	0.21	0.96	0.80	0.21
8	0	0.10	0.04	0.27	0.05	0.30	0.03	0.03	0.29
9	0	0.070	0.03	0.05	0.19	0.14	0	0.12	0.14
10	0	0.02	0	0.41	0.01	0.02	0	0.04	0.03
11	0	0	0	0.18	0	0.01	0	0	0.01
12	0	0.01	0.01	0	0	0	0.01	0	0
13	0	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0

the Australasian group showed only two alleles (a gene diversity of 0.07 in Papua New Guinea, Table 4). On the other hand, the highest values of gene diversity (23) were found in African samples, e.g., 0.75 in West Africa. Allele 4 was highly specific for individuals of African origin, although it was also found in Moroccans (f=0 in women, f=0.02 in men) and Turks (f=0.01 in women, f=0.02 in men) but in very low frequency.

The Y locus had an even smaller overall variation. The highest gene diversity values of 0.69 were found in Chinese population samples (Fig. 2). African populations had the highest frequency of allele 11 (ranging from f=0.78 in Khoisans to f=0.98 in Ovambos); alleles 11 and 12 were the most frequent in Australopapuans (Fig. 3); in Caucasoids the peak was at allele 12 (Fig. 4); and in Asians the median was located slightly above 12. Allele 14 was observed only in Asians. The 13-repeat allele was found at highest frequency in Asians (f=0.44 in China) and at very low frequency in Caucasoids and West Africans (Fig. 5). On the other hand, alleles 8 and 9 were detected only in African and Caucasoid populations.

Differing Variation in X and Y Samples

So far, many authors reported higher genetic diversity in African than in other populations (4,24,25), which is in accordance with our findings at the X locus of DXYS156. However, the patterns of distribution of

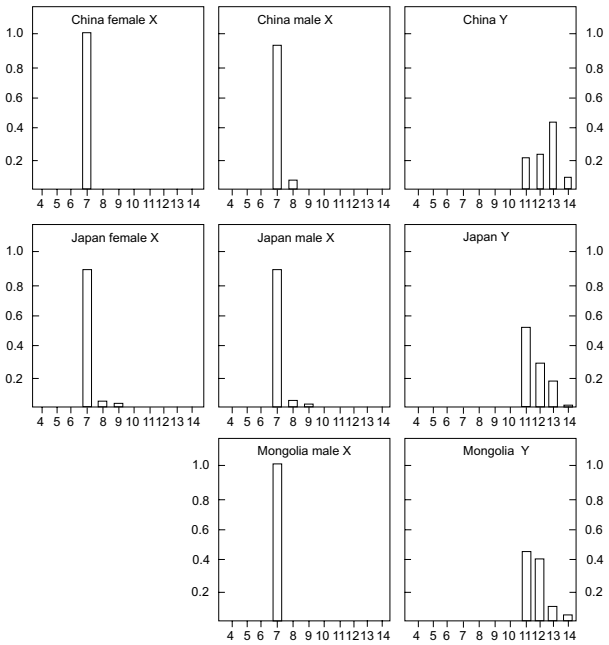

Figure 2. Allele frequency distribution for DXYS156 in Asian populations. Abscissa – alleles; ordinate – frequency of alleles.

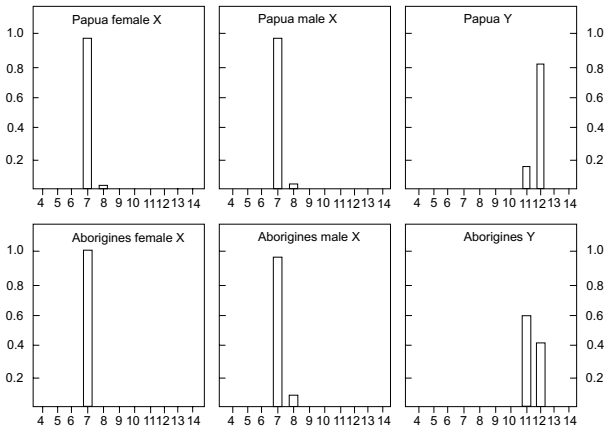

Figure 3. Allele frequency distribution for DXYS156 in Australo-Papuan populations. Abscissa – alleles; ordinate – frequency of alleles.

Table 3. Allele frequencies of DXYS156 in men

	China (n=79)		Germany (n=58)		Italy (n=100)		Japan (n=61)		Khoisan (n=55)		Morocco (n=44)	
Allele	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y
4	0	0	0	0	0	0	0	0	0.15	0	0.02	0
5	0.01	0	0	0	0	0	0	0	0.02	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0.75	0
7	0.92	0	0.83	0	0.81	0.01	0.93	0	0.25	0	0.02	0
8	0.06	0	0.05	0	0.06	0	0.05	0	0.36	0.02	0.14	0
9	0	0	0.07	0.02	0.12	0	0.02	0	0.09	0	0.02	0
10	0	0	0.05	0	0.01	0	0	0	0.09	0	0.02	0
11	0	0.22	0	0.07	0	0	0	0.51	0.04	0.78	0	0.66
12	0	0.25	0	0.88	0	0.98	0	0.3	0	0.2	0	0.32
13	0	0.44	0	0.03	0	0.01	0	0.18	0	0	0	0.02
14	0	0.09	0	0	0	0	0	0.02	0	0	0	0
	Mongolia (n=20)		East Africa (n=21)		Ovambo (n=54)		Papua (n=54)		Turkey (n=119)		West Africa (n=181)	
4	0	0	0.14	0	0.3	0	0	0	0.02	0	0.35	0
5	0	0	0	0	0	0	0	0	0	0	0.01	0
6	0	0	0.05	0	0	0	0	0	0.01	0	0	0
7	1	0	0.33	0	0.22	0	0.96	0	0.85	0	0.18	0
8	0	0	0.29	0.05	0.3	0	0.04	0	0.02	0.01	0.24	0
9	0	0	0.14	0.05	0.11	0	0	0	0.11	0	0.15	0
10	0	0	0.05	0	0.04	0	0	0	0	0	0.04	0.01
11	0	0.45	0	0.9	0.04	0.98	0	0.17	0	0.12	0.03	0.9
12	0	0.4	0	0	0	0.02	0	0.83	0	0.85	0	0.09
13	0	0.1	0	0	0	0	0	0	0	0.03	0	0.01
14	0	0.05	0	0	0	0	0	0	0	0	0	0

Table 4. Statistical parameters of DXYS156-X in population samples with n 50[a]

Parameter	Germany	Japan	Morocco	Papua	Turkey	West Africa
PD	0.56	0.31	0.59	0.13	0.58	0.89
H_{obs}	0.37	0.17	0.4	0.07	0.38	0.77
H_{exp}	0.36	0.16	0.4	0.07	0.35	0.75
± SE	0.12	0.08	0.08	0.06	0.09	0.02
Exact test (p)	0.56	1	0.97	0.01	0.93	0.66

[a]Abbreviations: PD – power of discrimination; H_{obs} – observed heterozygosity; H_{exp} – expected heterozygosity (± standard error of the mean). Calculations were performed using HWE Analysis software (version 3.2, C. Puers, Münster).

DXYS156-Y are different: the diversity in Sub-Saharan samples was low, whereas in the Japanese and Chinese it showed higher values.

The astonishing Asian Y-diversity could be explained by a higher mutation rate for Y-linked STRs in combination with a long allele to start the diversification process, because long alleles tend to mutate faster (26). The average allele size in Asians is, in fact, greater than in the African samples. Also, recent work seems to indicate that the mutation favored the formation of longer alleles. Since changes in the Y chromosome are more likely to be passed on than X changes because of the 3:1 proportion of chromosomes, the accumulation of mutations is faster in Y-chromosomal loci. But the length difference between Asian and African DXYS156Y samples is not high enough to solely explain the found patterns. Therefore, other influences must have been important, too.

Another possible explanation is a greater effective population size in Asian men, but there has been no evidence for this theory yet.

It is evident that so-called 'bottlenecks' have reduced African diversity and the lack of Gaussian distribution supports that claim. Such "bottlenecks" could have had a great influence on the reduction of diversity, as well as many sex-specific patterns of behavior, war, male sexual behavior, environmental changes, and huge migrations. Watson et al (3) and Soodyall and Jenkins (27) pointed out that linguistic data provided the evidence of such migrations: about 3,000 years ago, there was an expansion of Bantu-speaking people, who superseded other societies in ancient Africa. This may have resulted in a decrease in the variation at the DNA level. Sexual selection of certain males may also have played a role, but it is questionable whether such behavior was a characteristic of Africans only.

A small group of individuals that had a high level of diversity at the Y chromosome could have founded the Asian population. This effect is referred to as the "founder effect". The long alleles were most probably brought to Asia by those individuals, as stated above, which could explain the occurrence of the longest alleles particularly in Asian samples.

None of all these mechanisms can exclusively explain the contrasting patterns of variation, but they might have contributed together to the observed distribution. Additional studies are necessary to elucidate the evolution of this particular STR.

The low mutation rate (0 mutation in more than 400 meioses, B. Brinkmann, unpublished observation), which is in contrast with the higher mutation rate of the tetrameric STR systems, such as DYS19 and

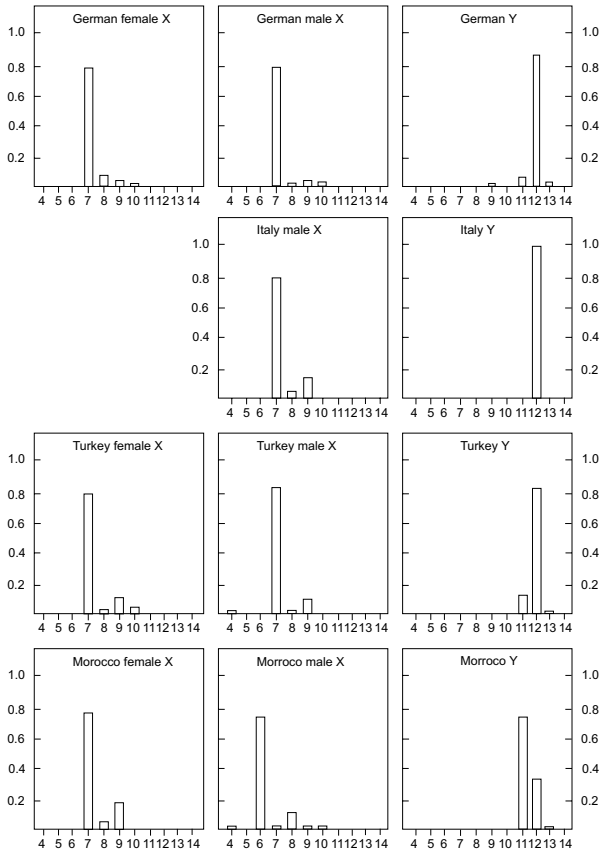

Figure 4. Allele frequency distribution for DXYS156 in Caucasoid populations. Abscissa – alleles; ordinate – frequency of alleles.

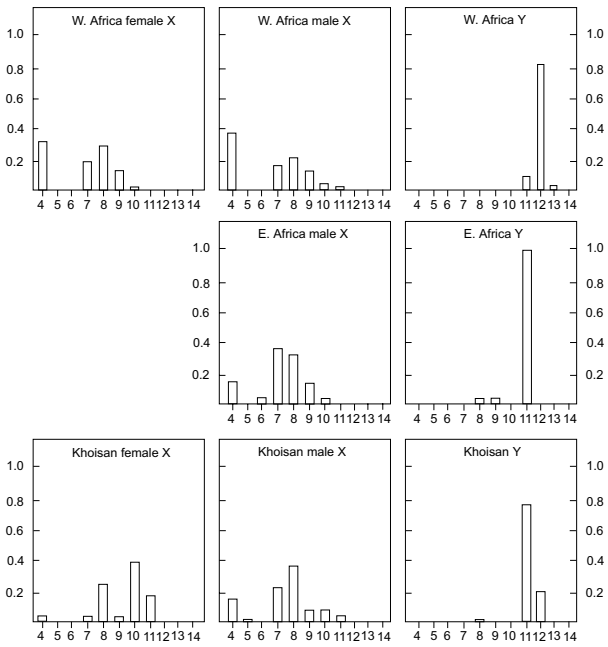

Figure 5. Allele frequency distribution for DXYS156 in African populations. Abscissa – alleles; ordinate – frequency of alleles.

DYS390 (28), makes DXYS156 particularly useful for phylogenetic analysis (29). The complex allele distribution may be forensically useful for the ethnic differentiation of the male component of unknown stains, e.g., in rape cases.

References

1 Armour JA, Anttinen T, May CA, Vega EE, Sajantila A, Kidd JR, et al. Minisatellite diversity supports a recent African origin for modern humans. Nat Genet 1996;13: 154-60.

2 Hammer MF, Spurdle AB, Karafet T, Bonner MR, Wood ET, Novelletto A, et al. The geographic distribution of human Y chromosome variation. Genetics 1997;145: 787-805.

3 Watson E, Bauer K, Aman R, Weiss G, von Haeseler A, Pääbo S. mtDNA sequence diversity in Africa. Am J Hum Genet 1996;59:437-44.

4 Vigilant L, Stoneking M, Harpending H, Hawkes K, Wilson AC. African populations and the evolution of human mitochondrial DNA. Science 1991;253:1503-7.

5 Krings M, Stone A, Schmitz RW, Krainitzki H, Stoneking M, Pääbo S. Neandertal DNA sequences and the origin of modern humans. Cell 1997;90:19-30.

6 Underhill PA, Shen P, Lin AA, Jin L, Passarino G, Yang WH, et al. Y chromosome sequence variation and the history of human populations. Nat Genet 2000;26: 358-61.

7 Scozzari R, Cruciani F, Malaspina P, Santolamazza P, Ciminelli BM, Torroni A, et al. Differential structuring of human populations for homologous X and Y microsatellite loci. Am J Hum Genet 1997;61:719-33.

8 Karafet T, de Knijff P, Wood E, Ragland J, Clark A, Hammer MF. Different patterns of variation at the X- and Y-chromosome-linked microsatellite loci DXYS156X and DXYS156Y in human populations. Hum Biol 1998; 70:979-92.

9 Chen H, Lowther W, Avramopoulos D, Antonarakis SE. Homologous loci DXYS156X and DXYS156Y contain a polymorphic pentanucleotide repeat (TAAAA)n and map to human X and Y chromosomes. Hum Mutat 1994;4:208-11.

10 Kim HS, Crow TJ. Human DXYS156 of pentanucleotide repeat (TAAAA)n: chromosomal localization by somatic hybrid mapping and sequencing analysis. Cytogenet Cell Genet 1998;83:54-5.

11 Schwartz A, Chan DC, Brown LG, Alagappan R, Pettay D, Disteche C, et al. Reconstructing hominid Y evolution: X-homologous block, created by X-Y transposition, was disrupted by Yp inversion through LINE-LINE recombination. Hum Mol Genet 1998;7:1-11.

12 Crow JF. How much do we know about spontaneous human mutation rates? Environ Mol Mutagen 1993;21: 122-9.

13 Horst B, Eigel A, Sanguansermsri T, Brinkmann B. Human Y-chromosomal STR types in north Thailand. Int J Legal Med 1999;112:211-2.

14 Brinkmann C, Forster P, Schürenkamp M, Horst J, Rolf B, Brinkmann B. Human Y-chromosomal STR haplotypes in a Kurdish population sample. Int J Legal Med 1999;112:181-3.

15 Forster P, Kayser M, Meyer E, Roewer L, Pfeiffer H, Benkmann H, et al. Phylogenetic resolution of complex mutational features at Y-STR DYS390 in aboriginal Australians and Papuans. Mol Biol Evol 1998;15:1108-14.

16 Nata M, Brinkmann B, Rolf B. Y-Chromosomal STR haplotypes in a population from north west Germany. Int J Legal Med 1999;112:406-8.

17 Rossi E, Rolf B, Schürenkamp M, Brinkmann B. Y-chromosome STR haplotypes in an Italian population sample. Int J Legal Med 1999;112:78-81.

18 Siffert W, Forster P, Jockel KH, Mvere DA, Brinkmann B, Naber C, et al. Worldwide ethnic distribution of the G protein beta3 subunit 825T allele and its association with obesity in Caucasian, Chinese, and Black African individuals. J Am Soc Nephrol 1999;10:1921-30.

19 Banaschak S, Rolf B, Brinkmann B. Influence of different staining techniques on the DNA analysis of histological sections. Int J Legal Med 2000;113:114- 6.

20 Möller A, Brinkmann B. Locus ACTBP2 (SE33). Sequencing data reveal considerable polymorphism. Int J Legal Med 1994;106:262-7.

21 Bär W, Brinkmann B, Budowle B, Carracedo A, Gill P, Lincoln P, et al. DNA recommendations: Further report of the DNA Commission of the ISFH regarding the use of short tandem repeat systems. Int J Legal Med 1997; 110:175-6.

22 Mannucci A, Sullivan KM, Ivanov PL, Gill P. Forensic application of a rapid and quantitative DNA sex test by amplification of the X-Y homologous gene amelogenin. Int J Legal Med 1994;106:190-3.

23 Nei M, Roychoudhury AK. Sampling variances of heterozygosity and genetic distance. Genetics 1974; 76:379-90.

24 Jin L, Baskett ML, Cavalli-Sforza LL, Zhivotovsky LA, Feldman MW, Rosenberg NA. Microsatellite evolution in modern humans: a comparison of two data sets from the same populations. Ann Hum Genet 2000;64:117-34.

25 Wiegand P, Meyer E, Brinkmann B. Microsatellite structures in the context of human evolution. Electrophoresis 2000;21:889-95.

26 Brinkmann B, Klintschar M, Neuhuber F, Hühne J, Rolf B. Mutation rate in human microsatellites: influence of the structure and length of the tandem repeat. Am J Hum Genet 1998;62:1408-15.

27 Soodyall H, Jenkins T. Mitochondrial DNA polymorphisms in Negroid populations from Namibia: new light on the origins of the Dama, Herero and Ambo. Ann Hum Biol 1993;20:477-85.

28 Holtkemper U, Rolf B, Hohoff C, Forster P, Brinkmann B. Mutation rates at two human Y-chromosomal microsatellite loci using small pool PCR techniques. Hum Mol Genet 2001;10:629-33.

29 Forster P, Röhl A, Lünnemann P, Brinkmann C, Zerjal T, Tyler-Smith C, et al. A short tandem repeat-based phylogeny for the human Y chromosome. Am J Hum Genet 2000;67:182-96.

Received: May 8, 2001
Accepted: May 11, 2001

Correspondence to:
Bernd Brinkmann
Institut für Rechtsmedizin
Universität Münster
Von-Esmarch-Str. 62
D-48149 Münster, Germany
brinkma@uni-muenster.de