

ORIGINAL ARTICLE

E. Meyer · P. Wiegand · B. Brinkmann

Phenotype differences of STRs in 7 human populations

Received: 27 April 1995 / Accepted: 17 May 1995

Abstract A maximum of 6 STR systems (TH01, VWA, ACTBP2, FES, F13B, D21S11) was investigated in 7 human populations (Germans, Turks, Moroccans, Japanese, Chinese, Papuans, Ovambos). In each population no deviations from Hardy-Weinberg equilibrium were observed. Out of each population the phenotypes of 50 individuals (comprising 3 to 6 STRs) were randomly selected. Based on the phenotype frequencies interpopulation comparisons were carried out using the frequencies of each other population. Within major ethnic groups only minor differences in phenotype frequencies were found. Between major ethnic groups differences of up to several orders of magnitude could be observed. The most discriminative STRs for interpopulation comparisons were TH01, FES and F13B.

Key words Short tandem repeats (STRs) · Population comparisons · Phenotype probability

Introduction

STR polymorphisms tend to be the most promising DNA generation for forensic applications (Edwards et al. 1992; Brinkmann and Wiegand 1993; Kimpton et al. 1993). We have performed a large scale investigation including a maximum of 6 STRs and 7 ethnic groups. The investigation included phenotyping by the use of sequenced allelic ladders, sequencing of a considerable number of alleles, analysis of at least 500 meioses in each system. The advantage of this complex approach lies in the comparability of data because the same technology and definition were applied in all samples. In the following we have elaborated the question whether interracial and/or intraracial differences can be found between randomly selected

STR phenotypes. If such differences occur it would be necessary to know their influence on the evaluation of forensic casework and on evolutionary implications.

Materials and methods**Population samples**

The following Caucasian population samples were included: Germans from the Münster area (Northwestern Germany), Turks living in the Adana area (Southern Turkey) and a Moroccan subpopulation living in Brussels, Belgium. Asian (mongoloid) representatives were Chinese from the Han Shen Yong area and Japanese from the Shiga area. Ovambos (Bantus) from Namibia were chosen as a Black African population sample and the Pacific Islander population sample were Papuans from Papua/New Guinea. For DNA isolation all samples were derived from healthy and unrelated donors. A minimum of 100 individuals of each population was used to determine the allele frequencies in each STR system.

DNA extraction

DNA from EDTA blood (Germans) was extracted as described previously (Brinkmann et al. 1991). Air dried blood specimens (Turks, Moroccans, Ovambos, Papuans), dried saliva on cotton fabric (Japanese) and hair roots (Chinese) were extracted using Chelex 100 and Proteinase K (Wiegand et al. 1993a).

DNA amplification and electrophoresis

PCR was carried out in 25 µl reaction volumes, using 5 ng of template DNA, 1 U Taq-DNA-polymerase (Promega), 200 mM each deoxynucleotide, 1 mM each primer and 2 µl reaction buffer (500 mM KCl, 100 mM Tris/HCl, pH 8.8; 1% Triton X-100, 0.1% gelatine). The PCR primers and reaction conditions for the STR systems applied have been described previously.

HumTH01	(Edwards et al. 1992; Wiegand et al. 1993b)
HumVWA	(Kimpton et al. 1992; Möller et al. 1994a)
HumACTBP2	(Polymeropoulos et al. 1992; Möller and Brinkmann 1994)
HumFES	(Polymeropoulos et al. 1991; Möller et al. 1994b)
HumF13B	(Nishimura and Murray 1992; Alper et al., subm.)
HumD21S11	(Sharma and Litt 1992; Möller et al. 1994b)

E. Meyer · P. Wiegand · B. Brinkmann (✉)
Institute of Legal Medicine,
Westfälische Wilhelms-Universität, Von-Esmarch-Strasse 86,
D-48149 Münster, Germany

Electrophoretic separation of PCR products was performed by high resolution polyacrylamide gel electrophoresis according to Allen et al. (1989) as essentially described by Möller et al. (1994a, b).

Rand et al. (1992) and Wiegand et al. (1993b). An overview including more statistical evaluations and sequence data will be published elsewhere (Brinkmann et al., submitted).

Statistical evaluation

We have randomly selected 50 individuals out of each population sample and carried out interpopulation comparisons using the allele frequencies of each other population. If a given allele had not been observed in the compared population, its theoretical frequency was calculated with the null-hypothesis, e.g. if 300 alleles were checked without observing this allele, its maximum frequency was calculated to be $1:300$, = 0.33%. For interpopulation comparisons the frequency pattern of person x in population y was set to one and the frequency difference to the same pattern in population z was expressed as the ratio. If the frequency was higher in population x the ratio was expressed as a figure > 1, if it was smaller as a figure < 1.

For all population samples Hardy-Weinberg equilibrium could be confirmed ($P > 0.05$) by forming allele classes according to

Results

Depending on the STRs applied, only minor allele frequency differences existed within the same ethnic group (e.g. Germans – Turks; Chinese – Japanese) but large differences could be observed between major ethnic groups (Fig. 1 a–g). For instance, in F13B allele 6 is prevalent in Ovambos with a frequency of ca. 40% and absent from both Asian populations and Papuans. Allele 10 is present in the latter populations with a frequency up to 75% in contrast to Ovambos with only 5%.

Fig. 1 Allele frequency distribution of 6 STRs in various human populations. **a** TH01; **b** VWA; **c**, **d** ACTBP2; **e** FES; **f** F13B; **g** D21S11

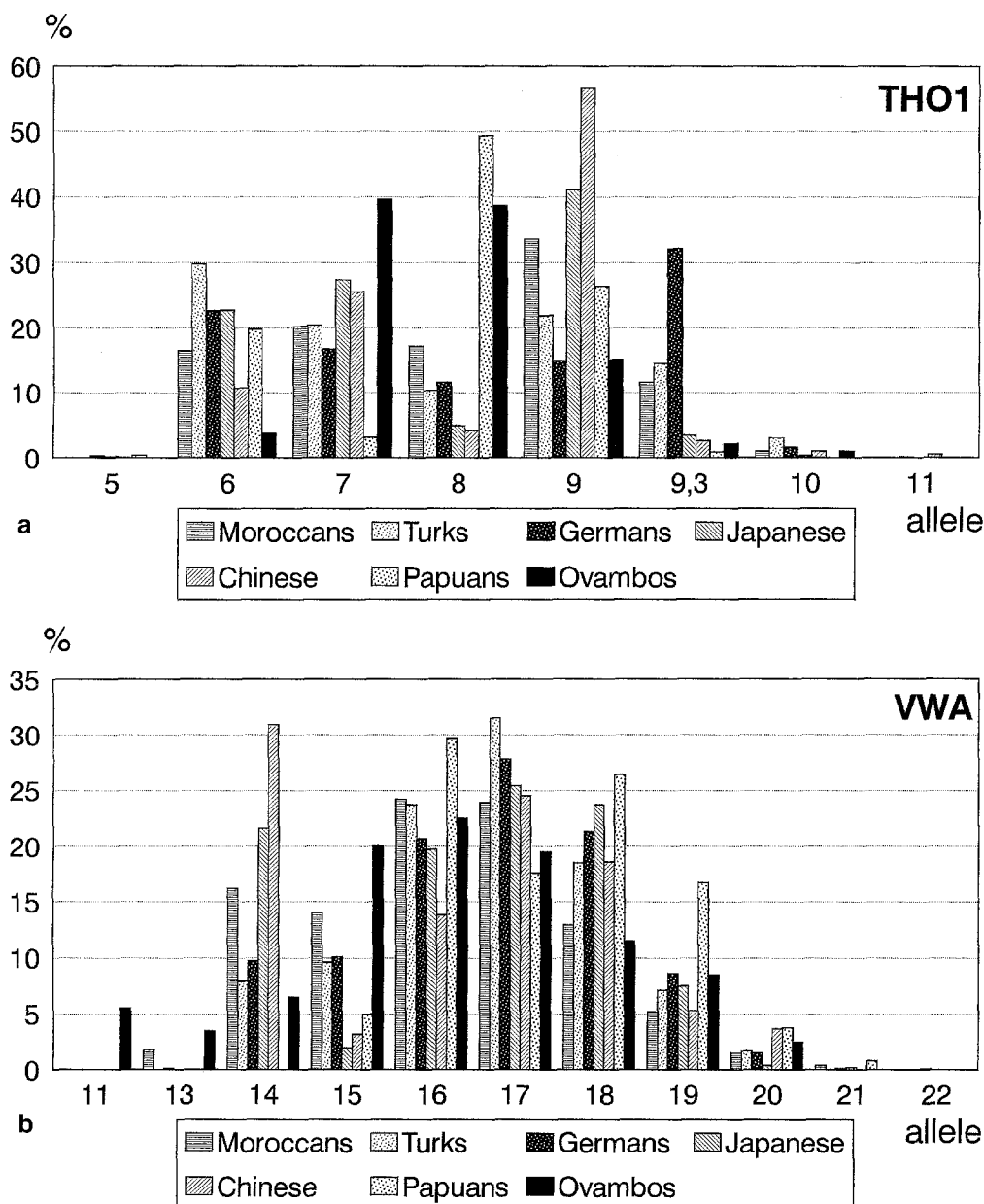


Fig. 1c-e

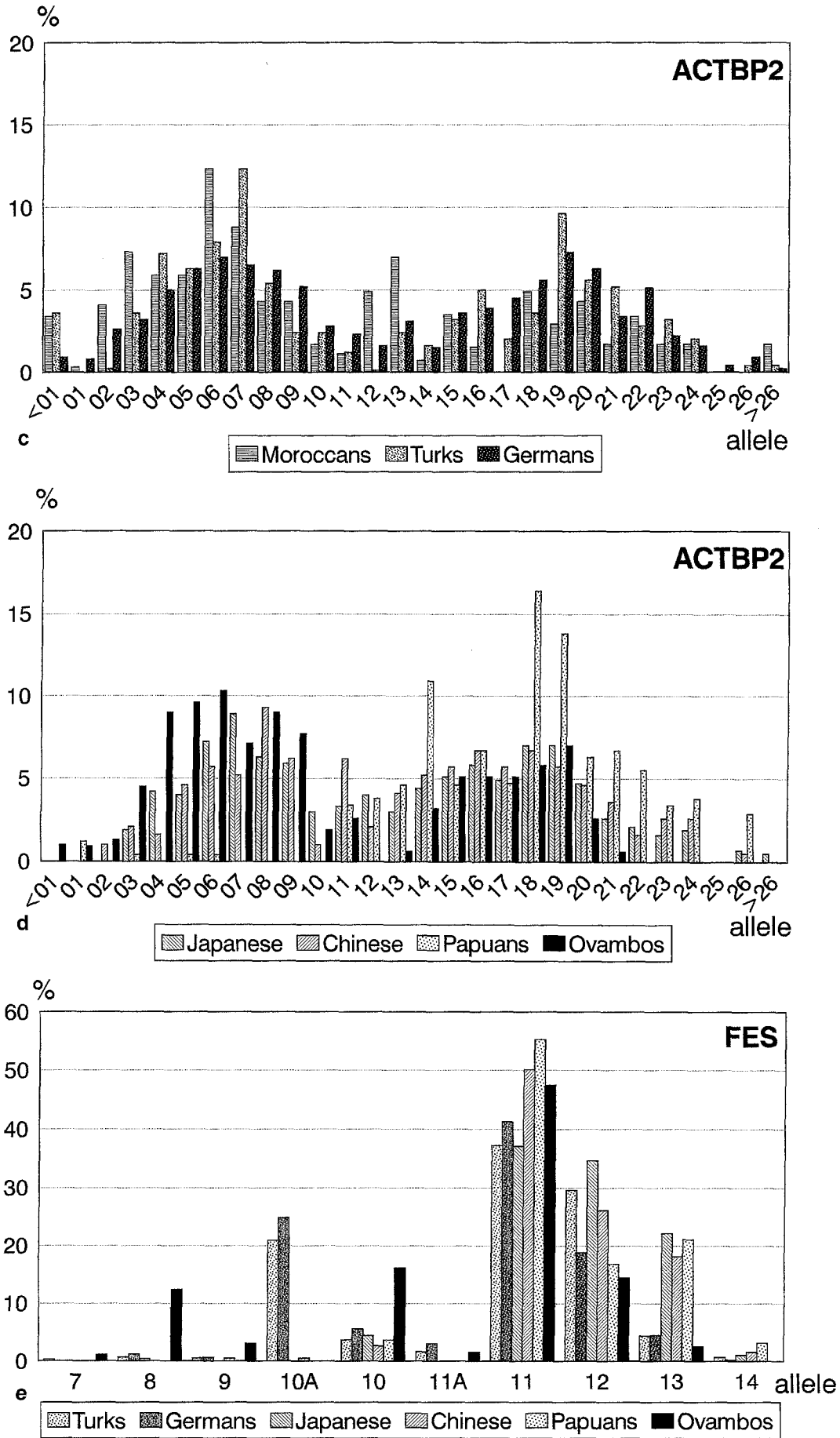
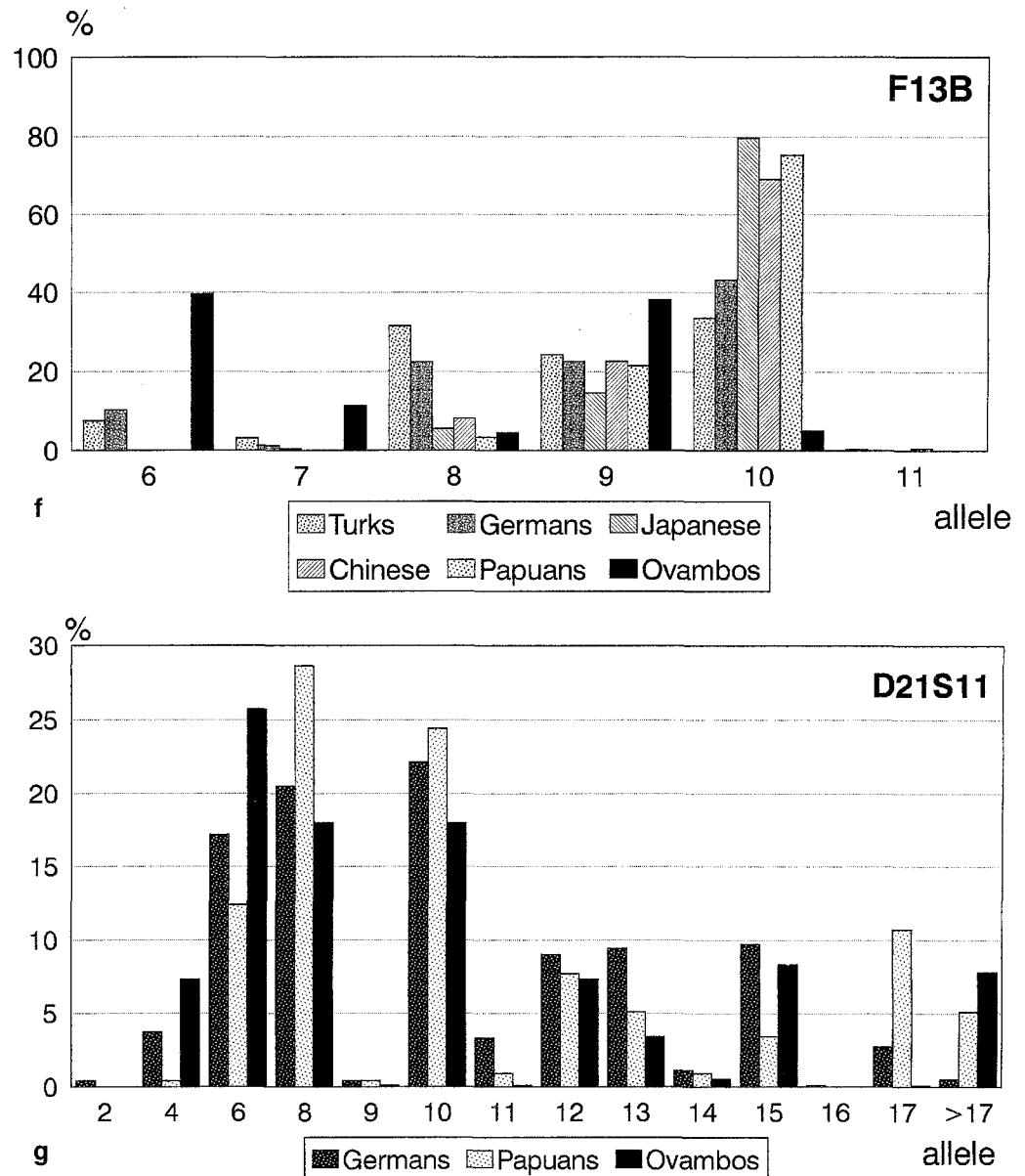


Fig. 1f–g



The calculation of the differences between 2 populations led to a wide variation of the individual figures (Table 1) and also to a considerable variation of the mean values associated to a population pair (Table 2). The STRs which are most discriminative between major ethnic groups are TH01, F13B and FES while VWA, ACTBP2 and D21S11 show smaller interpopulation differences.

Combining the phenotype data of 3 up to 6 STRs in the case of major ethnic groups results in a great increase of interpopulation differences. Within the same ethnic group, e.g. Caucasians or Asians an accumulative effect could not be observed. On the contrary, there can occur levelling effects (Table 3).

Discussion

As has been discussed for SLPs, allele frequency differences between populations seem to become levelled after the application of several systems (Chakraborty and Kidd 1991; Budowle and Monson 1993). In contrast our results obtained with STRs indicate that differences of several orders of magnitude can be observed between major ethnic groups. There can exist extreme differences between major populations which no longer allow for any ceiling or levelling approach. To mention but a few data: After applying 5 STR systems, white Caucasians differ from Asians on average by approximately 4 orders of magnitude and after 6 systems from black Africans by 1×10^5 .

In contrast, within a major ethnic group the differences seem to become levelled. Thus the mean difference

Table 2a Mean frequency differences (population discrimination power) of single STR systems between a) Germans; b) Chinese; and c) Ovambos and other populations as indicated. (Ger = Germans; Tur = Turks; Mor = Moroccans; Ova = Ovambos; Chi = Chinese; Jap = Japanese; Pap = Papuans). The mean values (MV) were derived from 50 randomly chosen phenotypes. SD = standard deviation; Min = minimum value; Max = maximum value

		TH01	VWA	ACTBP2	FES	F13B	D21S11
<i>a</i>							
Ger-Chi	MV	23.3	1.8	1.8	267.8	15.6	
	SD	44.8	1.2	1.5	747.2	61.5	
	Min	0.17	0.36	0.3	0.03	0.39	
	Max	152.4	4.7	10.2	2460.2	424.4	
Ger-Jap	MV	11.5	1.6	1.6	707.9	40.9	
	SD	22.1	1.4	2.1	2063.1	168.9	
	Min	0.22	0.44	0.31	0.02	0.29	
	Max	75.3	4.6	12.9	6833.8	1178.8	
Ger-Mor	MV	2.2	1.3	1.9			
	SD	2.2	0.6	2.3			
	Min	0.29	0.52	0.17			
	Max	7.9	2.8	14.1			
Ger-Ova	MV	40.4	1.7	4.2	262.4	29.1	6.7
	SD	72.2	0.8	14.7	740.5	27.5	11.9
	Min	0.17	0.46	0.39	0.08	0.03	0.45
	Max	233.7	3.4	104.7	2460.2	74.7	40.5
Ger-Pap	MV	90.4	6.2	50.7	407.2	35.6	2.4
	SD	180.8	11.0	83.0	1157.8	101.8	2.5
	Min	0.13	0.2	0.17	0.01	0.33	0.18
	Max	609.7	37.9	284.3	3844.0	663.1	11.3
Ger-Tur	MV	1.6	1.0	1.5	1.2	1.1	
	SD	1.3	0.2	1.3	0.5	0.4	
	Min	0.52	0.76	0.28	0.29	0.28	
	Max	4.9	1.47	6.4	2.67	1.9	
<i>b</i>							
Chi-Ger	MV	6.8	2.0	2.2	3.9	1.9	
	SD	5.4	2.3	2.2	5.0	0.8	
	Min	0.03	0.28	0.2	0.03	0.36	
	Max	14.4	10.2	12.0	32.2	2.6	
Chi-Jap	MV	1.3	1.2	1.3	1.0	1.1	
	SD	0.5	0.7	0.7	0.2	0.4	
	Min	0.4	0.5	0.3	0.5	0.8	
	Max	1.9	4.5	3.4	1.4	2.2	
Chi-Mor	MV	1.9	1.7	8.6			
	SD	0.9	0.9	15.3			
	Min	0.05	0.3	0.07			
	Max	2.8	3.7	80.3			
Chi-Ova	MV	6.9	4.2	4.4	5.5	100.8	
	SD	5.7	5.9	9.3	7.7	90.7	
	Min	0.07	0.24	0.08	0.3	1.1	
	Max	14.1	22.6	59.0	48.5	190.4	
Chi-Pap	MV	8.4	383.5	37.1	1.2	1.1	
	SD	7.4	1425.5	88.5	0.5	0.5	
	Min	0.04	0.2	0.07	0.4	0.8	
	Max	18.0	5967.6	331.3	2.4	2.5	
Chi-Tur	MV	3.4	2.6	3.7	2.9	2.8	
	SD	2.4	3.5	3.8	2.9	1.5	
	Min	0.07	0.3	0.14	0.02	0.24	
	Max	6.7	15.3	16.8	16.9	4.2	
<i>c</i>							
Ova-Ger	MV	5.6	26.8	1.8	5.5	14.6	3.5
	SD	3.3	173.0	0.8	8.1	21.2	5.7
	Min	0.16	0.3	0.2	0.6	0.2	0.001
	Max	10.7	1225.0	3.8	29.5	90.3	23.3
Ova-Jap	MV	12.8	11.2	1.8	3.0	4555.6	
	SD	19.2	25.8	1.4	3.1	7356.2	
	Min	0.24	0.16	0.18	0.12	0.16	
	Max	62.7	136.1	7.8	11.2	17424.0	
Ova-Mor	MV	3.0	1.2	4.7			
	SD	1.7	0.7	6.7			
	Min	0.37	0.3	0.2			
	Max	4.7	3.8	27.0			

Table 2 (continued)

		TH01	VWA	ACTBP2	FES	F13B	D21S11
Ova-Chi	MV	16.8	7.1	2.5	17.2	1684.6	
	SD	26.5	12.0	2.5	39.7	2638.4	
	Min	0.4	0.13	0.1	0.14	0.12	
	Max	85.9	49.0	10.1	146.9	6272.6	
Ova-Pap	MV	26.2	5.2	181.0	15.6	2624.3	4.6
	SD	53.0	11.5	241.3	36.2	4127.2	9.5
	Min	0.14	0.2	0.05	0.11	0.12	0.001
	Max	167.3	76.6	795.0	134.1	9801.0	44.6
Ova-Tur	MV	5.3	4.4	2.3	11.1	11.8	
	SD	4.1	19.1	1.2	20.7	10.4	
	Min	0.24	0.38	0.12	0.23	0.2	
	Max	13.6	136.1	4.6	76.6	27.9	

Table 3 Combined phenotype frequency differences (mean values) of 3 up to 6 STR systems between different human populations. Abbreviations as given in Table 2

	3 systems (TH01, VWA, ACTBP2)	4 systems (FES)	5 systems (F13B)	6 systems (D21S11)
Mor-Ger	13	—	—	—
Mor-Tur	6.1	—	—	—
Mor-Chi	34	—	—	—
Mor-Pap	7496	—	—	—
Mor-Jap	65	—	—	—
Mor-Ova	62	—	—	—
Chi-Ger	27	116	203	—
Chi-Tur	30	106	238	—
Chi-Mor	21	—	—	—
Chi-Pap	50273	53959	53752	—
Chi-Jap	2.0	2.2	2.3	—
Chi-Ova	198	719	62199	—
Pap-Ger	149	609	729	6133
Pap-Tur	156	547	1104	—
Pap-Chi	577	745	767	—
Pap-Mor	267.3	—	—	—
Pap-Jap	273	362	349	—
Pap-Ova	187	946	149396	238595
Ger-Tur	2.5	3.0	2.9	—
Ger-Mor	4.4	—	—	—
Ger-Chi	86	9739	31281	—
Ger-Pap	6200	1790582	10490118	11749276
Ger-Jap	36	6623	31290	—
Ger-Ova	120	1446	31491	131916
Jap-Tur	14	88	293	—
Jap-Mor	9.2	—	—	—
Jap-Chi	2.3	2.7	2.8	—
Jap-Pap	82782	93351	105650	—
Jap-Ger	9.4	41	100	—
Jap-Ova	35	186	38008	—
Ova-Tur	26	268	2199	—
Ova-Mor	19	—	—	—
Ova-Chi	209	2568	278528	—
Ova-Pap	5132	77677	15703461	52645562
Ova-Jap	269	688	230938	—
Ova-Ger	140	1575	10273	14211
Tur-Ger	8.2	2.2	2.2	—
Tur-Mor	4.8	—	—	—
Tur-Chi	4.4	19209	92909	—
Tur-Pap	1054	116057	1423334	—
Tur-Jap	22	23387	235073	—
Tur-Ova	365	327147	1467681	—

between Turks and Germans is extremely small (Tur → Ger; 5 systems, mean value (MV) = 2.2) and quite small differences exist between Germans and Moroccans (Ger → Mor; 3 systems, MV = 4.4). Furthermore, the difference comparing Japanese and Chinese is also extremely small irrespective of which population the phenotypes were derived from (Jap → Chi; 5 systems, MV = 2.8); but between Japanese and Germans for example there exist different mean values depending on the direction of comparison (Ger → Jap; 5 systems, MV = 31 290; Jap → Ger; 5 systems, MV = 100). This can be explained by the fact that all alleles present in the Japanese population also exist in Germans but not all alleles present in Germans can be observed in Japanese.

For forensic application our results suggest that STR typing of medicolegal evidence requires knowledge of which ethnic group a suspect belongs to and the allele frequencies in this population.

Using the wrong data base could lead to a miscalculation of probability of up to 5 orders of magnitude (Ger → Ova; 6 STRs, MV = 130 000; Ger → Chi; 5 STRs, MV = 31 000). Within a major ethnic group such as Caucasians, the differences in allele frequencies are low and differences of combined phenotypes among Germans, Turks and Moroccans nearly become levelled. A comparison of German and other Caucasians populations also reveals minor differences (TH01: Germans → Americans MV = 1.0; Germans → Croatians MV = 1.28; Germans → Danes MV = 1.02; Puers et al. 1993; Kubat et al. 1995; Nellemann et al. 1994; VWA: Germans → Italians MV = 1.04; Germans → Finns MV = 1.25; allele frequencies derived from Buscemi et al. 1995; Sajantila et al. 1994). These data suggest that also further Caucasian populations will show this levelling effect. Our data further suggest that in the case of biological evidence typed with a set of STR markers the ethnic origin of a phenotype pattern can be predicted sometimes with considerable probability. But this would always necessitate precise information of existing alternatives. For instance this could especially be important in countries like the USA where Americans of African origin and Caucasians are abundant and equally represented. Of course the allele frequencies between Black Americans and Ovambos are not identical (TH01: Ovambos → US Blacks, MV = 1.8; Puers et al. 1993; VWA: Ovambos → US Blacks, MV = 2.1; Sajantila et al. 1994) but these relatively small differences will only have a minor influence on the particular phenotype probabilities.

As already mentioned, STRs such as TH01, FES and F13B have a higher power of discrimination between major ethnic groups compared to VWA, ACTBP2 and D21S11.

The question arises whether these differences are due to different mutation rates, genetic drift or admixture. So far we have investigated more than 500 meioses in each of the STR systems and have detected 2 mutations in VWA (0.3%) and 3 mutations in ACTBP2 (0.4%) only. Therefore, the most probable explanation for the observed differences is genetic drift rather than new mutations. Thus we assume that the phenotype frequency differences are due to differences within the founder groups of the major

ethnic groups. Each small founder group must have had an original STR pattern or pool which made it different from the other ones.

Additional changes in population size and genetic admixture from other populations could have modified the allele composition compared to the original distributions (Brinkmann et al., submitted).

References

- Allen RC, Graves G, Budowle B (1989) Polymerase chain reaction amplification products separated on rehydratable polyacrylamide gels and stained with silver. *Biotechniques* 7:736–744
- Alper B, Schürenkamp M, Meyer E (1995) HumFES/FPS and HumF13B: population genetic study on a Turkish and German population survey. *Int J Legal Med* (submitted)
- Brinkmann B, Rand S, Wiegand P (1991) Population and family data of RFLP's using selected single- and multi-locus systems. *Int J Legal Med* 104:81–86
- Brinkmann B, Wiegand P (1993) Medicolegal implications of PCR-based VNTRs. In: *Proceedings from the Fifth International Symposium on Human Identification*. Promega Corporation, Madison, USA, pp 149–160
- Brinkmann B, Sajantila A, Goedde HW, Matsumoto H, Nishi K, Wiegand P (1995) Population genetic comparisons among eight different populations using allele frequency and sequence data from three microsatellite loci. *Am J Hum Genet* (submitted)
- Budowle B, Monson K (1993) The forensic significance of various reference population databases for estimating the rarity of variable number of tandem repeat (VNTR) loci profiles. In: *Pena SDJ, Chakraborty R, Epplen JT, Jeffreys AJ (eds) DNA Fingerprinting: State of science*. Birkhäuser, Basel
- Buscemi L, Cucurachi N, Mencarelli R, Tagliabracci A, Wiegand P, Ferrara SD (1995) PCR analysis of the short tandem repeat (STR) system HUMVWA31. Allele and genotype frequencies in an Italian population sample. *Int J Legal Med* 107:171–173
- Chakraborty R, Kidd KK (1991) The utility of DNA typing in forensic work. *Science* 254:1735–1739
- Edwards A, Hammond HA, Jin L, Caskey CT, Chakraborty R (1992) Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. *Genomics* 12:241–253
- Kimpton CP, Walton A, Gill P (1992) A further tetranucleotide repeat polymorphism in the vWF gene. *Hum Mol Genet* 1:28
- Kimpton CP, Gill P, Walton A, Urquhart A, Millican ES, Adams M (1993) Automated DNA profiling employing multiplex amplification of short tandem repeat loci. *PCR Methods Appl* 8(3):13–22
- Kubat M, Wiegand P, Brinkmann B (1995) Population genetic study from the Zagreb area using 3 STR systems. *Int J Legal Med* 107:219–221
- Möller A, Wiegand P, Grischow C, Seuchter SA, Baur MP, Brinkmann B (1994a) Population data and forensic efficiency values for the STR systems HumVWA, HumMBP and HumFABP. *Int J Legal Med* 106:183–189
- Möller A, Meyer E, Brinkmann B (1994b) Different types of structural variation in STRs: HumFES/FPS, HumVWA and HumD21S11. *Int J Legal Med* 106:319–323
- Möller A, Brinkmann B (1994) Locus ACTBP2 (SE33) – Sequencing data reveal considerable polymorphism. *Int J Legal Med* 106:262–267
- Nellemann LJ, Möller A, Morling N (1994) PCR typing of DNA fragments of the short tandem repeat (STR) system HUMTH01 in Danes and Greenland Eskimos. *Forensic Sci Int* 68:45–51
- Nishimura DY, Murray JC (1992) A tetranucleotide repeat for the F13B locus. *Nucleic Acids Res* 20:1167

- Polymeropoulos MH, Rath DS, Xiao H, Merrill CR (1991) Tetranucleotide repeat polymorphism at the human c-fes/fps proto-oncogene (FES). *Nucleic Acids Res* 19:4791
- Polymeropoulos MH, Rath DS, Xiao H, Merrill CR (1992) Tetranucleotide repeat polymorphism at the human beta-actin related pseudogene H-beta-Ac-psi-2 (ACTBP2). *Nucleic Acids Res* 20:1432
- Puers C, Hammond HA, Jin L, Caskey T, Schumm J (1993) Identification of repeat sequence heterogeneity at the polymorphic short tandem repeat locus HUMTH01 (AATG)_n and reassignment of alleles in population analysis by using a locus-specific allelic ladder. *Am J Hum Genet* 53:953–958
- Rand S, Puers C, Skowasch K, Wiegand P, Budowle B, Brinkmann B (1992) Population genetics and forensic efficiency data of 4 AMPFLPs. *Int J Legal Med* 104:329–333
- Sajantila A, Pacek P, Lukka M, Syvänen AC, Nokelainen P, Sistonen P, Peltonen L, Budowle B (1994) A microsatellite polymorphism in the von Willebrand Factor gene: comparison of the allele frequencies in different population samples and evaluation for forensic medicine. *Forensic Sci Int* 68:91–102
- Sharma V, Litt M (1992) Tetranucleotide repeat polymorphism at the D21S11 locus. *Hum Mol Genet* 1:67
- Wiegand P, Bajanowski T, Brinkmann B (1993a) DNA typing of debris from fingernails. *Int J Legal Med* 106:81–84
- Wiegand P, Budowle B, Rand S, Brinkmann B (1993b) Forensic validation of the STR systems SE33 and TC11. *Int J Legal Med* 105:315–320