

Sergio Tofanelli · Ilaria Boschi · Stefania Bertoneri  
Valentina Coia · Luca Taglioli  
Marcello Giovanni Franceschi · Giovanni Destro-Bisol  
Vincenzo Pascali · Giorgio Paoli

## Variation at 16 STR loci in Rwandans (Hutu) and implications on profile frequency estimation in Bantu-speakers

Received: 26 July 2002 / Accepted: 23 October 2002 / Published online: 15 February 2003

© Springer-Verlag 2003

**Abstract** A data set of 16 autosomal STRs (the 13 CODIS loci plus HumCD4, HumFES, HumF13A1) was obtained in a sample of 52 unrelated Hutus from Rwanda. Genotypes at all loci met Hardy-Weinberg expectations with the exception of HumCSF1PO. No significant evidence of association across alleles at independent loci was obtained. Statistical parameters demonstrated the forensic usefulness of the analysed systems (combined PE=0.9999996, combined PD=1:2.27×10<sup>18</sup>). Pairwise comparisons showed that the Hutu gene pool differs substantially from that of other Bantu-speaking populations suggesting the use of ethnic-specific population databases in forensic casework analysis. The introduction of a non-negligible bias was confirmed by calculating the differences between multiple-locus profile frequencies of western and eastern Bantoids using local and non-local reference databases.

**Keywords** STR-profiles · Hutu · Rwanda · CODIS · Bantu-speakers

### Introduction

Standard sets of short tandem repeat loci (STRs) have been adopted world-wide in routine forensic practice both for casework analyses and data banking. The 13 autosomal loci HumCSF1PO, D3S1358, D5S818, D7S820, D8S1179,

D13S317, D16S539, D18S51, D21S11, HumFGA, HumTH01, HumTPOX and HumVWA, are included in the most widely used commercial kits for DNA profiling (AmpF/STR Profiler Plus+ COfiler and AmpFISTR Identifier by Applied Biosystems; GenePrint Powerplex 16 System by Promega Corporation), from the U.S. CODIS database (Combined DNA Index System), have been technically validated [1], and satisfy the recommendations of the European Network of Forensic Science Institutes (ENFSI). Further STR loci have been validated for forensic applications and included in web repositories, among others HumCD4, HumFES, HumF13A1 [2, 3].

The availability of rapid and reliable genotyping procedures have facilitated the collection of reference data for many human populations as they are needed to confidently infer statistical estimates of multiple-loci DNA profiles [4, 5]. World-wide population surveys based on microsatellite loci showed that genetic variation is highest in Africans [6, 7] and that the observed linear gradient of decreasing diversity away from Africa is free of ascertainment bias [8].

Within sub-Saharan Africa, the region surrounding Lake Victoria is one of the most populated areas. However, STR-based DNA profiles of the human groups living there are virtually missing from anthropological and forensic databases, the unique exception being the Ugandans for five loci [9]. The Hutus (11.3 million according to The World Factbook [10]) are the largest Bantu-speaking group from Rwanda and Burundi (80–90% of the whole population), where they have been living for approximately 2,000 years as farmers. Even though the life expectancy rate at birth is the lowest in the world (39.3 years) and they have played the major role in terms of victims (around 1 million) and refugees (around 2 million) during the civil war of 1994, the population size has tripled in the last two generations [10, 11].

In this paper, we present the allele frequencies at 16 autosomal STR loci (those included in the CODIS database plus HumCD4, HumFES, HumF13A1) in a population sample of 52 Hutus from north-eastern Rwanda, compare them with those of other sub-Saharan populations with the

S. Tofanelli (✉) · S. Bertoneri · L. Taglioli · M. G. Franceschi  
G. Paoli

Department of Ethology, Ecology and Evolution,  
Anthropology Unit, University of Pisa,  
Via S. Maria 55, 56126 Pisa, Italy  
Tel.: +39-50-24613, Fax: +39-50-24653,  
e-mail: sergtofa@discau.unipi.it

I. Boschi · V. Pascali  
Forensic Haematology Laboratory, Institute of Legal Medicine,  
Sacro Cuore Catholic University (UCSC),  
Largo F. Vito 1, 00168 Rome, Italy

V. Coia · G. Destro-Bisol  
Department of Animal and Human Biology,  
University of Rome “La Sapienza”, P.le A. Moro, Rome, Italy

same linguistic affiliation (Niger-Kordofan/Benue-Congo/Bantoid), and estimate multi-locus profile frequency shifts using local and non-local reference databases. We aimed at enlarging the panel of STR reference data for African populations and at evaluating its use for forensic purposes.

## Materials and methods

The samples were collected among the workers of the catholic parish at Nyarurema (1°21'S, 30°19'E), in north-eastern Rwanda. The study has been performed in accordance with the ethical standards, all donors gave their informed consent prior to being included in the research and personal data have been treated anonymously.

From 3 to 6 eyebrows were pulled out and stored in eppendorf tubes for 5–7 days until their arrival at the laboratory of Molecular Anthropology of the University of Pisa. DNA was isolated from hair bulbs following the method of Higuchi et al. [12]. The genotyping of the CODIS core loci (HumCSF1PO, D3S1358, D5S818, D7S820, D8 S1179, D13S317, D16S539, D18S51, D21S11, HumFGA, HumTH01, HumTPOX, HumVWA) plus the Amelogenin locus for sex typing was carried out in the Laboratory of Forensic Haematology of the Catholic University of Rome, by AmpF/STR Profiler Plus and AmpF/STR COfiler PCR amplification kits (Applied Biosystems) according to product instructions. Fragment sizes were detected by the ABI PRISM 310 Genetic Analyzer (Applied Biosystems) using Genescan 500 ROX as internal size standard and sequenced allele ladders. The genotyping of HumCD4, HumFES and HumF13A1 loci was performed in the Laboratory of Molecular Anthropology of the University “La Sapienza” of Rome. PCR conditions were as previously described [13]. The amplified products were analysed in denaturing polyacrylamide gels using a semi-automated DNA sequencer (A.L.F. express, Pharmacia Biotech). Allele and internal standards were used for typing. Allele nomenclature followed the recommendations of the European DNA profiling group (EDNAP) [14].

Departures from the Hardy-Weinberg (HWE) and linkage equilibrium (LD) were estimated by exact tests using a Markov chain method (200,000 runs) with the GENEPOP 3.1 software [15]. Pairwise population differentiations were estimated by Fisher exact tests [16] with the ARLEQUIN 2.000 software [17]. The unbiased heterozygosity (H) and relative standard errors were calculated according to equations 8.4 and 8.7 in Nei [18], the polymorphism information content (PIC) as described in Botstein et al. [19] and the power of discrimination (PD) according to Weir [20].

## Results and discussion

The distributions of the allele frequencies (analytical results are available from the authors upon request) and the statistical parameters of forensic interest computed from the 16 STR systems are given in Tables 1 and 2. Although the genotyped sample is small it has to be considered as representative of the allele diversity within Hutu population. In fact, we observed worldwide very rare variant alleles (i.e. D18S51\*15.2, HumFGA\*31.2) and an average number of alleles/locus higher than those from larger African population data sets (i.e. the Bubi from Equatorial Guinea or Blacks from Zimbabwe, see Table 3). The amelogenin typing always matched the phenotypic sex of the donors [21].

Only one probability test showed significant departures from the HWE equilibrium, at the HumCSF1PO locus. The *p*-value remained under the conventional limit of significance also after the application of the Bonferroni correction for multiple tests [22]. The use of a more sensitive test (one-tail DE test, *p*=0.021) demonstrated that such deviation was due to homozygote excess.

**Table 1** Allele frequencies and relevant forensic statistics computed at the 16 STR loci in the Hutu population (Rwanda)

Allele	HumCD4 (2N=100)	HumCSF1PO (2N=100)	D3S1358 (2N=104)	D5S818 (2N=102)	D7S820 (2N=100)	D8S1179 (2N=104)	D13S317 (2N=104)	D16S539 (2N=100)
3.2	-	-	-	-	-	-	-	-
5	0.2100	-	-	-	-	-	-	-
6	0.0500	-	-	-	-	-	-	-
7	0.0500	0.0800	-	-	0.0300	-	-	-
8	0.1000	0.0100	-	0.0588	0.1100	-	0.0481	0.0700
9	0.0200	0.1700	-	-	0.1200	-	-	0.1300
10	0.1900	0.2700	-	0.0294	0.4400	0.0192	0.0288	0.1300
11	0.1800	0.2200	0.0096	0.2451	0.2100	0.0577	0.2692	0.2900
12	0.0900	0.2400	0.0096	0.3824	0.0800	0.1538	0.3942	0.2300
13	0.0500	0.0100	0.0096	0.2843	0.0100	0.1538	0.1827	0.1400
14	0.0600	-	0.0673	-	-	0.2692	0.0769	-
15	-	-	0.2692	-	-	0.2788	-	0.0100
16	-	-	0.3365	-	-	0.0385	-	-
17	-	-	0.2404	-	-	0.0288	-	-
18	-	-	0.0577	-	-	-	-	-
H	0.862±0.009	0.790±0.009	0.752±0.013	0.712±0.014	0.732±0.022	0.800±0.012	0.733±0.017	0.809±0.011
PIC	0.8421	0.7515	0.7065	0.6551	0.6945	0.7675	0.6872	0.7773
PD	0.9641	0.9199	0.8948	0.8616	0.8923	0.9296	0.8844	0.9345
PE	0.7162	0.5758	0.5197	0.4534	0.516	0.6035	0.4992	0.6159
HWE	0.108±0.006	<b>0.001±0.000</b>	0.740±0.008	0.204±0.004	0.689±0.007	0.546±0.008	0.292±0.006	0.672±0.006
Exact test								

Values under the 0.05 level are in boldtype.

Table 2

Allele	D18S51 (2N=104)	HumF13A1 (2N=94)	HumFES (2N=98)	HumTH01 (2N=102)	HumTPOX (2N=98)	HumVWA (2N=100)	Allele	HumFGA (2N=104)	D21S11 (2N=102)
3,2	-	0.1809	-	-	-	-	18	0.0096	-
4	-	0.1383	-	-	-	-	19	0.0096	-
5	-	0.3298	-	-	-	-	19,2	0.0096	-
6	-	0.0426	-	0.1961	0.0510	-	20	0.0577	-
7	-	0.1489	-	0.2549	0.0102	-	21	0.0673	-
8	-	0.0957	0.0816	0.2647	0.3469	-	22	0.2692	-
9	-	0.0106	0.0612	0.1863	0.3061	-	23	0.2115	-
9,3	-	-	-	0.0686	-	-	24	0.1731	-
10	-	0.0106	0.1633	0.0098	0.1020	-	24,3	-	0.0098
10,2	0.0096	-	-	-	-	-	25	0.0481	-
11	0.0192	-	0.3878	-	0.1633	-	26	0.0192	-
12	0.0192	-	0.2347	0.0196	0.0204	-	27	0.0192	0.0294
13	0.0288	0.0319	0.0612	-	-	0.0100	27,1	0.0096	-
14	0.0288	0.0106	0.0102	-	-	0.1100	28	0.0481	0.2843
15	0.1923	-	-	-	-	0.1600	29	0.0096	0.1471
15,2	0.0096	-	-	-	-	-	30	0.0192	0.1961
16	0.1442	-	-	-	-	0.2300	30,2	-	0.0196
17	0.1923	-	-	-	-	0.1600	31	-	0.1275
18	0.1250	-	-	-	-	0.2300	31,2	-	0.0098
19	0.1154	-	-	-	-	0.1000	32	0.0192	0.0098
19,2	-	-	-	-	-	-	32,2	-	0.0784
20	0.0769	-	-	-	-	-	33	-	0.0098
21	0.0288	-	-	-	-	-	33,2	-	0.0196
22	0.0096	-	-	-	-	-	34	-	0.0196
23	-	-	-	-	-	-	34,2	-	0.0098
24	-	-	-	-	-	-	35	-	0.0196
24,3	-	-	-	-	-	-	36	-	0.0098
H	0.871±0.009	0.809±0.015	0.758±0.018	0.791±0.009	0.750±0.016	0.825±0.008	H	0.843±0.013	0.838±0.013
PIC	0.8528	0.7807	0.7193	0.7529	0.7052	0.7958	PIC	0.8205	0.8148
PD	0.9682	0.9377	0.905	0.9208	0.8948	0.9429	PD	0.956	0.9535
PE	0.7531	0.6258	0.5431	0.5783	0.5208	0.6407	PE	0.6866	0.6773
HWE Exact test	0.077±0.009	0.079±0.006	0.277±0.007	0.912±0.003	0.632±0.007	0.582±0.006	HWE Exact test	0.215±0.019	0.582±0.024

**Table 3** Genetic differentiation between Hutus and other Bantu-speaking populations plus African-Americans: probability values of Fisher exact tests and relative errors after 100,000 Markov steps

Population	African-Americans (U.S.A.)	Bubi (Eq. Guinea)	Bamileke (Cameroon)	Ewondo (Cameroon)	Sanga (C.A.R.)	Mbenzele Pygmies (C.A.R.)	Ugandans (Uganda)	Blacks (Mozambique)	Blacks (Zimbabwe)	Ovambos (Namibia)	Xhosa (South Africa)
2N	400	244–310	80–92	128–130	60–70	92–98	120–180	220	204–208	192–210	192
Reference	[28]	[29]	[13]	[13]	[13]	[13]	[9]	[30]	[31]	[9,32]	[33]
HumCD4	-	-	<b>0.0012±</b> <b>0.0005</b>	0.1872± 0.087	0.5682± 0.0078	<b>0.0002±</b> <b>0.0002</b>	<b>0.0209±</b> <b>0.0023</b>	-	-	<b>0.0468±</b> <b>0.0060</b>	-
HumCSFIPO	<b>0.0000±</b> <b>0.0000</b>	<b>0.0014±</b> <b>0.0004</b>	0.0508± 0.0047	<b>0.0000±</b> <b>0.0000</b>	<b>0.0178±</b> <b>0.0022</b>	<b>0.0000±</b> <b>0.0000</b>	-	<b>0.0007±</b> <b>0.0005</b>	<b>0.0000±</b> <b>0.0000</b>	-	<b>0.0020±</b> <b>0.0005</b>
D3S1358	0.8690± 0.0070	0.4326± 0.0146	0.6903± 0.0103	0.5345± 0.0097	0.6122± 0.0109	0.4484± 0.0172	-	0.0960± 0.0095	0.8500± 0.0054	-	-
D5S818	0.5908± 0.0119	-	-	-	-	-	-	0.1127± 0.0067	-	-	0.3286± 0.0131
D7S820	<b>0.0455±</b> <b>0.0047</b>	-	-	-	-	-	-	<b>0.0146±</b> <b>0.0023</b>	-	-	0.1945± 0.0092
D8S1179	0.3544± 0.0113	<b>0.0137±</b> <b>0.0032</b>	-	-	-	-	-	<b>0.0164±</b> <b>0.0026</b>	-	-	-
D13S317	0.5911± 0.0098	-	-	-	-	-	-	0.4170± 0.0107	-	-	0.0660± 0.0063
D16S539	0.3059± 0.0128	-	-	-	-	-	-	-	-	-	0.1341± 0.0060
D18S51	0.3617± 0.0168	<b>0.0029±</b> <b>0.0011</b>	0.9748± 0.0015	0.4761± 0.0182	<b>0.0068±</b> <b>0.0013</b>	0.0607± 0.0042	-	0.5491± 0.0194	-	-	-
HumF13A1	-	-	0.2240± 0.0079	0.1267± 0.0088	0.3798± 0.0175	<b>0.0000±</b> <b>0.0000</b>	-	-	-	-	-
HumFES	-	-	0.1264± 0.0084	0.5906± 0.0131	<b>0.0000±</b> <b>0.0000</b>	0.6522± 0.0139	0.1462± 0.0095	-	-	0.1303± 0.0076	-
HumTH01	<b>0.0013±</b> <b>0.0010</b>	<b>0.0000±</b> <b>0.0000</b>	<b>0.0013±</b> <b>0.0004</b>	<b>0.0337±</b> <b>0.0042</b>	<b>0.0025±</b> <b>0.0010</b>	<b>0.0000±</b> <b>0.0000</b>	0.5223± 0.0119	<b>0.0323±</b> <b>0.0034</b>	<b>0.0153±</b> <b>0.0026</b>	<b>0.0000±</b> <b>0.0000</b>	<b>0.0002±</b> <b>0.0001</b>
HumTPOX	0.0991± 0.0066	<b>0.0494±</b> <b>0.0045</b>	0.6795± 0.0079	<b>0.0442±</b> <b>0.0035</b>	0.4203± 0.0079	<b>0.0322±</b> <b>0.0046</b>	-	<b>0.0009±</b> <b>0.0003</b>	0.1705± 0.0094	<b>0.0021±</b> <b>0.0007</b>	<b>0.0217±</b> <b>0.0028</b>
HumVWA	0.1324± 0.0085	<b>0.0024±</b> <b>0.0007</b>	<b>0.0139±</b> <b>0.0023</b>	<b>0.0157±</b> <b>0.0031</b>	<b>0.0135±</b> <b>0.0015</b>	<b>0.0035±</b> <b>0.0012</b>	0.2508± 0.0111	0.0708± 0.0058	0.5627± 0.0111	<b>0.0144±</b> <b>0.0017</b>	0.0726± 0.0035
HumFGA	<b>0.0052±</b> <b>0.0015</b>	-	-	-	-	-	-	<b>0.0005±</b> <b>0.0003</b>	<b>0.0116±</b> <b>0.0030</b>	-	-
D21S11	0.4974± 0.0119	-	0.9099± 0.0038	0.3038± 0.0096	0.0536± 0.0032	-	-	<b>0.0016±</b> <b>0.0005</b>	-	-	-

Values under the 0.05 level are in boldtype.

Out of 78 comparisons only 2 exact tests for independence across loci gave probability values below the 0.05 level, less than expected on the basis of a 5% level of significance (3.9 pairs of loci). Therefore, the frequency of the multiple locus DNA profile can be estimated with a certain confidence by the product rule, namely the multiplication of individual locus profile frequencies. The estimated values for combined PE (0.9999996) and combined PD ( $1: 2.27 \times 10^{18}$ ) demonstrate the high forensic efficiency of the investigated set of STR loci in the Hutu population. The statistical support for Hutus is somewhat comparable with the combined values obtained by the Powerplex 16 System on African-Americans (PE=0.9999996, PD=1:1.42  $\times 10^{18}$ ) [23].

Linguistic and genetic evidence supports the archaeological hypothesis of a recent dispersal of Bantu farmers within the last 3000–4000 years [24, 25]. However, we observed significant differences at 11 loci between Hutu STR distributions and those of other Bantu-speakers groups (Table 3). The systems HumCSF1PO, D8S1179, HumFGA, HumTH01, HumWA and HumCD4, demonstrated to be highly discriminative as significant probability values were observed for most of the comparisons and global tests (data not shown), were always below the 0.001 significance level. This finding suggest an unfair bias (i.e. an overestimate against the suspect) in applying non-specific population databases to calculate matching probabilities when Bantu-speakers are considered as the source of DNA evidence [26].

In order to quantify the bias introduced when the Hutu database is used with non-Hutu Bantoids, we calculated the difference between 9-locus (HumCSF1PO, D3S1358, D18S51, D21S11, HumTH01, HumTPOX, HumVWA, HumFES, HumF13A1) profile frequencies of genotyped individuals belonging to the eastern (Hutus) and the western clades of Bantu dialects inferred by using local and non-local databases. To obtain the most conservative estimate of profile frequency shifts we chose the population with the least genetic differentiation from Hutus (Ewondo and Bamileke from Cameroon, see Table 3) and selected only those genotypes having no allele missing in the reference database. The number of individuals considered thus ranged from 50, comparing Ewondo genotypes against the Bamileke database, and 21, comparing Hutus genotypes against the Ewondo database (76.9% and 40.4%, respectively of the whole available data set). The presence of private alleles within each data set due to both low sample sizes and different population histories, accounts for this reduction.

When profile frequencies for Bamileke or Ewondo individuals were estimated against the other western Bantu database, average frequencies were 16 and 17 times lower, respectively than the frequencies estimated against the ethnic-specific database and the largest differences were about 2 orders of magnitude (from  $9.24 \times 10^{-12}$  to  $7.99 \times 10^{-14}$  and from  $1.02 \times 10^{-13}$  to  $1.19 \times 10^{-15}$ ). Conversely, the average matching probabilities for western Bantu genotypes were significantly lower when using the Hutu database and, vice versa, when using a Bamileke or Ewondo

database for Hutu genotypes. In this cases they were from 36 to 358 times lower than the frequencies obtained by using population-specific databases, the largest shift exceeded 3 orders of magnitude (from  $6.34 \times 10^{-14}$  to  $2.01 \times 10^{-17}$ ), and often the reciprocal of the upper frequency estimate approached the size of the world population ( $\sim 7 \times 10^9$ ).

Therefore, our conclusions are different from those recently drawn for Arabic speaking populations [27]: the use of non-local reference population data in forensic casework involving Bantu speakers or, more generally, sub-Saharan Africans, may yield matching probability estimates which differ several orders of magnitude, with non-negligible impact on the jury.

**Acknowledgements** This research was supported by the MIUR COFIN 2001 grant. We are indebted to the Hutu donors, whose participation in the project made this research possible.

## References

- Moretti TR, Baumstark AL, Defenbaugh DA, Keys KM, Smerick JB, Budowle B (2001) Validation of short tandem repeats (STRs) for forensic usage: performance testing of fluorescent multiplex STR systems and analysis of authentic and simulated forensic samples. *J Forensic Sci* 46:647–660
- Lygo JE, Johnson PE, Holdaway DJ, Woodroffe S, Whitaker JP, Clayton TM, Kimpton CP, Gill P (1994) The validation of short tandem repeat (STR) loci for use in forensic casework. *Int J Legal Med* 107:77–89
- Klitsch M, Crevenna R (1997) HumCD4: validation of a STR system for forensic purposes in an Austrian Caucasian population sample. *J Forensic Sci* 42:907–910
- Osier MV, Cheung KH, Kidd JR, Pakstis AJ, Miller PL, Kidd KK (2001) ALFRED: an allele frequency database for diverse populations and DNA polymorphisms. *Nucleic Acids Res* 29: 317–319
- Ruitberg CM, Reeder DJ, Butler JM (2001) STRBase: a short tandem repeat DNA database for the human identity testing community. *Nucleic Acids Res* 29:320–322
- Jorde LB, Rogers AR, Bamshad M, Watkins WS, Krakowiak P, Sung S, Kere J, Harpending HC (1997) Microsatellite diversity and the demographic history of modern humans. *Proc Natl Acad Sci USA* 94:3100–3103
- Calafell F, Shuster A, Speed WC, Kidd JR, Kidd KK (1998) Short tandem repeat polymorphism evolution in humans. *Eur J Hum Genet* 6:38–49
- Harpending H, Rogers A (2000) Genetic perspectives on human origins and differentiation. *Annu Rev Genomics Hum Genet* 1:361–385
- Brinkmann B, Junge A, Meyer A, Wiegand P (1998) Population genetic diversity in relation to microsatellite heterogeneity. *Hum Mutat* 11:135–144
- The World Factbook (2001) <http://www.cia.gov/cia/publications/factbook/>
- U.S. Census Bureau, <http://www.census.gov/>
- Higuchi R, Beroldingen CH von, Sensabaugh GF, Erlich HA (1988) DNA typing from single hairs. *Nature* 332:543–546
- Destro Bisol G, Boschi I, Caglià A, Tofanelli S, Pascali V, Paoli G, Spedini G (2000) Microsatellite variation in Central Africa: an analysis of intrapopulation and interpopulation genetic diversity. *Am J Phys Anthropol* 112:319–337
- Gill P, Brinkmann B, d'Aloja E, et al. (1997) Considerations from the European DNA profiling group (EDNAP) concerning STR nomenclature. *Forensic Sci Int* 87:185–192

15. Raymond M, Rousset F (1995) GENEPOP (version 3.1) is an updated version of GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *J Hered* 86:248–249
16. Fisher RA (1951) Standard calculations for evaluating a blood group system. *Heredity* 5:95–102
17. Excoffier L, Roessli D, Schneider S (2000) Arlequin (ver. 2.000): a software for population genetic data analysis. <http://anthropologie.unige.ch/arlequin>
18. Nei M (1987) *Molecular evolutionary genetics*. Columbia University Press, New York, pp 178–179
19. Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet* 32:314–31
20. Weir BS (1996) *Genetic data analysis II*. Sinauer Associates, Sunderland MA, p 210
21. Brinkmann B (2002) Is the amelogenin sex test valid? *Int J Legal Med* 116:63
22. Hockberg Y (1988) A sharper Bonferroni procedure for multiple tests of significance. *Biometrika* 75:800–802
23. Promega GenePrint Powerplex 16 System, Technical Manual <http://www.shpromega.com.cn/tmd012.pdf>
24. Greenberg J (1979) *Introduzione alla linguistica*. Boringhieri, Torino
25. Cavalli-Sforza LL, Menozzi P, Piazza A (1994) *The history and geography of human genes*. Princeton University Press, New Jersey, pp 162–171
26. Gallo JC, Thomas E, Novick GE, Herrera RJ (1997) Effects of subpopulation structure on probability calculations of DNA profiles from forensic PCR analysis. *Genetica* 101:1–12
27. Klitschar M, al-Hammadi N, Reichenpfader B (2001) Significant differences between Yemenite and Egyptian STR profiles and the influence on frequency estimations in Arabs. *Int J Legal Med* 114:211–214
28. Budowle B, Moretti TR, Baumstark AL, Defenbaugh DA, Keys KM (1999) Population data on the thirteen CODIS core short tandem repeat loci in African Americans, US Caucasians, Hispanics, Bahamians, Jamaicans and Trinidadians. *J Forensic Sci* 44:1277–1286
29. Gené M, Moreno P, Borrego N, Piqué E, Brandt C, Mas J, Luna M, Corbella J, Hugué E (2001) The Bubi population of Equatorial Guinea characterised by HUMTH01, HUMCSF1PO, HUMTPOX, D3S1358, D8S1179, D18S51 and D19S253 STR polymorphisms. *Int J Legal Med* 114:298–300
30. Alves C, Gusmão L, Amorim A (2001) STR data (AmpFISTR Profiler Plus and GenePrint CTTv) from Mozambique. *Forensic Sci Int* 119:131–133
31. Budowle B, Nhari LT, Moretti TR, Kanoyangwa SB, Masuka E, Defenbaugh DA, Smerick JB (1997) Zimbabwe black population data on the six short tandem repeat loci—CSF1PO, TPOX, TH01, D3S1358, VWA and FGA. *Forensic Sci Int* 90:215–221
32. Takeshita H, Meyer E, Brinkmann B (1997) The STR loci HumTPO and HumLPL: population genetic data in eight populations. *Int J Legal Med* 110:331–333
33. Wang B, Pang H, Koda Y, Soejima M, Kimura H (2000) Polymorphisms of eight STR loci in Chinese and African (Xhosa) populations. *Forensic Sci Int* 125:279–280