International Congress Series 1239 (2003) 249–258

Allele frequency distributions and other population genetic parameters for 13 STR loci in a UAE local population from Dubai

F.H. Alshamali [a], A.I. Alkhayat [a,*], B. Budowle [b], N.D. Watson [c]

[a] *Crime Laboratory, General Department of Forensic Sciences, Dubai Police GHQ, Dubai, United Arab Emirates*
[b] *Laboratory Division, FBI, Washington, D.C., USA*
[c] *Forensic Science Unit, Strathclyde University, Glasgow, UK*

Keywords: Allele frequency distributions; Population genetic parameters; 13 STR loci

1. Introduction

DNA polymorphism is a cornerstone of human identification. In the past few years, the typing of short tandem repeat (STR) loci has developed into one of the most rapid, efficient and precise methods of human identification. STR loci are the genetic markers of choice for use in routine forensic casework and for the national strategies for creating criminal intelligence databases [1–3]. The STR loci consist of tandemly repeated sequences 3 to 7 bp in size, are highly polymorphic, are abundant in the human genome, and are amenable to PCR [4–7]. Many polymorphic STR loci have been described, but the tetranucleotide STR loci exhibit properties that make them desirable as forensic markers. The small-defined size range of each amplified locus and the ability to manipulate the size of the amplicon by moving the primers has allowed development of multiplex sets of STR loci. Thus, the amount of information obtained from a single PCR is increased substantially with a concomitant reduction in template consumption and labor. By incorporating fluorescent dyes into the amplicon, automated detection is possible [8] and provides a cost-effective method to generate population studies [9–12] and to type samples to be entered in to DNA databanks or used for forensic identification.

Two multiplex kits AmpFlSTR® Profiler Plus™ kit and AmpFlSTR® Cofiler™ kit (Applied Biosystems, Foster City, CA, USA) allow amplification and typing of 13 STR

* Corresponding author. Tel.: +971-4-2666991; fax: +971-4-2014497.
E-mail address: agkhayat@emirates.net.ae (A.I. Alkhayat).

0531-5131/03 © 2003 Published by Elsevier Science B.V.
PII: S0531-5131(02)00625-8

Table 1
Observed allele frequency distributions for a population from Dubai (N=200)

Locus 1: D3S1358			
Allele	Frequency	Statistical figures	
11	0.003	Homozygosity (Obs)	20.0%
14	0.060	Homozygosity (Exp)	23.3%
15	0.253	Homozygosity test	0.273
16	0.285	Exact test	0.872
17	0.268	PD (Obs)	0.90315000
18	0.120	PD (Exp)	0.90583482
19	0.013	PE	0.542348481

Locus 2: vWA			
Allele	Frequency	Statistical figures	
11	0.003	Homozygosity (Obs)	19%
14	0.088	Homozygosity (Exp)	18.8%
15	0.140	Homozygosity test	0.956
16	0.238	Exact test	0.508
17	0.245	PD (Obs)	0.93500000
18	0.205	PD (Exp)	0.93643385
19	0.068	PE	0.62182151
20	0.015		

Locus 3: FGA			
Allele	Frequency	Statistical figures	
17	0.005	Homozygosity (Obs)	17.5%
<18	0.003	Homozygosity (Exp)	12.5%
18	0.013	Homozygosity test	0.034
19	0.088	Exact test	0.091
19.2	0.003	PD (Obs)	0.96780000
20	0.068	PD (Exp)	0.97025261
20.2	0.003	PE	0.74216393
21	0.120		
22	0.168		
22.2	0.008		
23	0.165		
24	0.168		
25	0.125		
26	0.030		
27	0.030		
28	0.005		
>30	0.003		

Locus 4: D8S1179			
D8S1179	Frequency	Statistical figures	
8	0.015	Homozygosity (Obs)	21.0%
9	0.005	Homozygosity (Exp)	16.3%
10	0.075	Homozygosity test	0.072

Table 1 (*continued*)

Locus 4: D8S1179			
D8S1179	Frequency	Statistical figures	
11	0.060	Exact test	0.698
12	0.143	PD (Obs)	0.95185000
13	0.238	PD (Exp)	0.95189295
14	0.203	PE	0.67122309
15	0.183		
16	0.068		
17	0.010		
18	0.003		

Locus 5: D21S11			
D21S11	Frequency	Statistical figures	
27	0.020	Homozygosity (Obs)	20.5%
28	0.193	Homozygosity (Exp)	16.1%
29	0.273	Homozygosity test	0.091
29.3	0.003	Exact test	0.095
30	0.148	PD (Obs)	0.95255000
30.2	0.013	PD (Exp)	0.95441389
31	0.050	PE	0.68002544
31.2	0.110		
32	0.008		
32.2	0.110		
33.2	0.050		
34.2	0.008		
35	0.015		
36	0.003		

Locus 6: D18S51			
Allele	Frequency	Statistical figures	
8	0.003	Homozygosity (Obs)	13.0%
10	0.005	Homozygosity (Exp)	11.5%
11	0.015	Homozygosity test	0.510
12	0.118	Exact test	0.506
13	0.143	PD (Obs)	0.97165000
14	0.185	PD (Exp)	0.97477869
15	0.130	PE	0.76277196
15.2	0.003		
16	0.100		
17	0.098		
18	0.073		
19	0.078		
20	0.030		
20.2	0.003		
21	0.010		
22	0.008		
23	0.003		

(continued on next page)

Table 1 (*continued*)

Locus 7: D5S818

Allele	Frequency	Statistical figures	
8	0.018	Homozygosity (Obs)	29.5%
9	0.040	Homozygosity (Exp)	26.2%
10	0.113	Homozygosity test	0.287
11	0.320	Exact test	0.164
12	0.353	PD (Obs)	0.88945000
13	0.150	PD (Exp)	0.88743996
14	0.008	PE	0.50568854

Locus 8: D13S317

Allele	Frequency	Statistical figures	
8	0.120	Homozygosity (Obs)	21.5%
9	0.043	Homozygosity (Exp)	22.2%
10	0.090	Homozygosity test	0.811
11	0.290	Exact test	0.845
12	0.323	PD (Obs)	0.91425000
13	0.103	PD (Exp)	0.91794824
14	0.033	PE	0.57448062

Locus 9: D7S820

Allele	Frequency	Statistical figures	
7	0.025	Homozygosity (Obs)	21.0%
8	0.168	Homozygosity (Exp)	21.5%
9	0.090	Homozygosity test	0.855
10	0.315	Exact test	0.670
11	0.260	PD (Obs)	0.91615000
12	0.115	PD (Exp)	0.92101586
13	0.020	PE	0.58072485
14	0.008		

Locus 10: D16S539

Allele	Frequency	Statistical figures	
8	0.033	Homozygosity (Obs)	22.5%
9	0.138	Homozygosity (Exp)	21.8%
10	0.108	Homozygosity test	0.815
11	0.318	Exact test	0.190
12	0.273	PD (Obs)	0.91385000
13	0.115	PD (Exp)	0.91944397
14	0.018	PE	0.57683140

Locus 11: THO1

Allele	Frequency	Statistical figures	
6	0.303	Homozygosity (Obs)	25.5%
7	0.248	Homozygosity (Exp)	21.9%
8	0.173	Homozygosity test	0.224
9	0.175	Exact test	0.059
9.3	0.090	PD (Obs)	0.91235000

Table 1 (*continued*)

Locus 11: THO1			
Allele	Frequency	Statistical figures	
10	0.010	PD (Exp)	0.91602302
11	0.003	PE	0.56594945
Locus 12: TPOX			
Allele	Frequency	Statistical figures	
6	0.013	Homozygosity (Obs)	30.5%
7	0.003	Homozygosity (Exp)	29.6%
8	0.438	Homozygosity test	0.778
9	0.138	Exact test	0.695
10	0.103	PD (Obs)	0.85965000
11	0.275	PD (Exp)	0.86561853
12	0.033	PE	0.46504871
Locus 13: CSF1PO			
Allele	Frequency	Statistical figures	
7	0.005	Homozygosity (Obs)	23.5%
8	0.010	Homozygosity (Exp)	27.1%
9	0.038	Homozygosity test	0.255
9.1	0.003	Exact test	0.526
10	0.318	PD (Obs)	0.86520000
11	0.243	PD (Exp)	0.87689326
12	0.330	PE	0.48304271
13	0.050		
14	0.005		

loci: D3S1358, vWA, FGA, D8S1179, D21S11, D18S51, D5S818, D13S317, D7S820, D16S539, THO1, TPOX, and CSF1PO and the gender locus amelogenin. The reagents for typing the loci D3S1358 and D7S820 are contained in both multiplex kits and provide a further quality control mechanism when using more than one kit for typing. As multiplex kits are developed that are compatible with high throughput automated fluorescence detection, population data will be needed on the STR loci contained within the kit. This study describes a population study on Dubai Arabs for the aforementioned 13 STR loci. The data can be used for forensic and paternity analysis.

2. Materials and methods

2.1. Sample preparation

Whole blood from 200 unrelated UAE locals of Dubai was obtained by venipuncture and collected in EDTA vacutainer tubes. Bloodstains were prepared using sterilized cotton swatches, air dried and stored at −20 °C until use.

Table 2
Summary of Karlin correlation test

Loci	*P*-value
1 D3S1358/2 vWA	0.463
1 D3S1358/3FGA	0.166
1 D3S1358/4 D8S1179	0.831
1 D3S1358/5 D21S11	0.236
1 D3S1358/6 D18S51	0.698
1 D3S1358/7 D5S818	0.147
1 D3S1358/8 D13S317	0.193
1 D3S1358/9 D7S820	0.952
1 D3S1358/10 D16S539	0.295
1 D3S1358/11 THO1	0.889
1 D3S1358/12 TPOX	0.256
1 D3S1358/13 CSF1PO	0.502
2 vWA/3 FGA	0.576
2 vWA/4 D8S1179	0.491
2 vWA/5 D21S11	0.930
2 vWA/6 D18S51	0.535
2 vWA/7 D5S818	0.776
2 vWA/8 D13S317	0.461
2 vWA/9 D7S820	0.031
2 vWA/10 D16S539	0.403
2 vWA/11 THO1	0.602
2 vWA/12 TPOX	0.554
2 vWA/13 CSF1PO	0.216
3 FGA/4 D8S1179	0.898
3 FGA/5 D21S11	0.563
3 FGA/6 D18S51	0.463
3 FGA/7 D5S818	0.576
3 FGA/8 D13S317	0.010
3 FGA/9 D7S820	0.983
3 FGA/10 D16S539	0.036
3 FGA/11 THO1	0.256
3 FGA/12 TPOX	0.357
3 FGA/13 CSF1PO	0.569
4 D8S1179/5 D21S11	0.361
4 D8S1179/6 D18S51	0.356
4 D8S1179/7 D5S818	0.764
4 D8S1179/8 D13S317	0.545
4 D8S1179/9 D7S820	0.968
4 D8S1179/10 D16S539	0.512
4 D8S1179/11 THO1	0.193
4 D8S1179/12 TPOX	0.284
4 D8S1179/13 CSF1PO	0.969
5 D21S11/6 D18S51	0.404
5 D21S11/7 D5S818	0.214
5 D21S11/8 D13S317	0.113
5 D21S11/9 D7S820	0.964
5 D21S11/10 D16S539	0.751
5 D21S11/11 THO1	0.866
5 D21S11/12 TPOX	0.746

Table 2 (*continued*)

Loci	*P*-value
5 D21S11/13 CSF1PO	0.760
6 D18S51/7 D5S818	0.950
6 D18S51/8 D13S317	0.028
6 D18S51/9 D7S820	0.969
6 D18S51/10 D16S539	0.249
6 D18S51/11 THO1	0.487
6 D18S51/12 TPOX	0.312
6 D18S51/13 CSF1PO	0.913
7 D5S818/8 D13S317	0.281
7 D5S818/9 D7S820	0.776
7 D5S818/10 D16S539	0.731
7 D5S818/11 THO1	0.065
7 D5S818/12 TPOX	0.018
7 D5S818/13 CSF1PO	0.610
8 D13S317/9 D7S820	0.281
8 D13S317/10 D16S539	0.090
8 D13S317/11 THO1	0.174
8 D13S317/12 TPOX	0.106
8 D13S317/13 CSF1PO	0.891
9 D7S820/10 D16S539	0.887
9 D7S820/11 THO1	0.270
9 D7S820/12 TPOX	0.648
9 D7S820/13 CSF1PO	0.487
10 D16S539/11 THO1	0.580
10 D16S539/12 TPOX	1.000
10 D16S539/13 CSF1PO	0.723
11 THO1/12 TPOX	0.980
11 THO1/13 CSF1PO	0.181
12 TPOX/13 CSF1PO	0.674

DNA was extracted using an organic method (Phenol/Chloroform/Isoamyl) [13] followed by washing using a centricon-100 microconcentrator (Amicon).

The quantity of DNA was estimated by running the extracts on ethidiumbromide stained agarose gel against known standards (Gibco). Dilutions were made accordingly and subsequently, the amount of DNA was estimated using a slot-blot hybridization procedure [14] using the Quantiblot kit (Applied Biosystems).

2.2. Amplification

Samples were amplified at the loci: D3S1358, vWA, FGA, D8S1179, D21S11, D18S51, D5S818, D13S317, D7S820 and amelogenin using AmpFlSTR® Profiler Plus™ kit (Applied Biosystems) and at the loci D16S539, THO1, TPOX, CSF1PO, D3S1358, and D7S820 using the AmpFlSTR® Cofiler™ kit (Applied Biosystems). Amplification was carried out in GeneAmp PCR system 9600 using 1 ng of template DNA in a 25 μl reaction volume following the manufacturer's recommended procedure.

2.3. Detection of amplified products

Amplified products were separated by electrophoresis and detected on the ABI Prism® 377 DNA Sequencer (Applied Biosystems) according to the manufacturer's recommended protocol. A 5% long Ranger™ gel (FMC Bioproducts, Rockland, ME) with a 36-cm-well-to-read was used. Some of the PCR products were analyzed using the ABI prism™ 310 Genetic Analyzer (Applied Biosystems) according to manufacturer's recommended procedure.

2.4. Statistical analysis

Allele designations were determined by comparison of the sample fragments with those of the allelic ladders with each kit. At each locus, the frequency of each allele was calculated from the numbers of each genotype in the sample set (i.e., the gene count method). Unbiased estimates of expected heterozygosity were computed as described by Edwards et al. [15]. Possible divergence from Hardy–Weinberg expectations (HWE) was tested by calculating the unbiased estimate of the expected homozygote/heterozygote frequencies [16–19] and also by performing an exact test [20] based on 2000 shuffling experiments. An interclass correlation criterion [21] for two-locus associations was used for detecting disequilibrium between the STR loci. The probability of discrimination (PD) and probability of exclusion (PE) was calculated according to Fisher. The values for θ were determined as described by Weir and Cockerham [22].

3. Results and discussion

The distribution of observed alleles for all 13 STR loci is shown in Table 1. The observed and expected homozygosities, exact test for departures from HWE, discrimination probability (PD), and probability of exclusion (PE) are also provided. All loci are highly polymorphic with the loci D18S51 (87%), FGA (82.5%) and vWA (81%) having the highest observed heterozygosities, and the locus TPOX showing the least observed heterozygosity (69.5%).

As expected, the most polymorphic loci are the most discriminating loci: D18S51 (97.16%), FGA (96.78%) and D21S11 (95.25%). The TPOX locus was the least discriminating (85.96%). For all 13 STR loci, the combined power of discrimination was greater than 0.99999999 and the combined probability of exclusion was 0.999995.

There was no evidence for departures from Hardy–Weinberg expectations (HWE) for any loci based on exact test.

An interclass correlation test analysis was carried out to detect any departures from independence of alleles between loci (Table 2). A total of 78 pair-wise comparisons were performed and five significant departures (6.4%) were observed. This number of observations is within expectations. However, none of the empirical levels of significance for these five observations were below the adjusted Bonferroni level.

Thus, the data support that assumption of independence is valid, and a multiple-locus profile frequency can be estimated using the product rule. However, the use of Wright's F_{st}

Table 3
F_{st} values for nine STR loci for Dubai and Oman

Locus	F_{st} value
D3S1358	0.0018
vWA	0.0009
FGA	0.0008
D8S117	0.0022
D21S11	0.0029
D18S51	0.0011
D5S818	0.0027
D13S317	0.0025
D7S820	0.0019

F_{st} value over all loci is 0.0010.

estimate is suggested to correct for potential substructure in the population. The F_{st} values were calculated for 9 of the 13 loci by comparing the data with that previously published on Omanis [23]. The F_{st} value over all nine loci is 0.0010, which suggests that there is little difference at these STR loci between Dubaian and Omani Arabs (Table 3). The F_{st} value estimated from these data supports the generalized recommendations of the National Research Council (NRC) [24] for taking a value of 0.01 as a conservative value for Arab populations.

In conclusion, a database for the 13 STR loci has been established for the Dubaian population. The data can be used for both forensic casework and paternity testing in Dubai. In addition, this study supports that F_{st} value of 0.01 is conservative for the Arab population.

References

[1] B. Budowle, T.R. Moretti, S.J. Niezgoda, B.L. Brown, CODIS and PCR-based short tandem repeat loci: law enforcement tools, Second European Symposium on Human Identification, Promega, Madison, WI, 1998, pp. 73–88.
[2] H.N. Dunbar, Latest developments in forensic DNA profiling—the UK national database, Second European Symposium on Human Identification, Promega, Madison, WI, 1998, pp. 55–58.
[3] W. Parson, Laboratory strategy of the Austrian DNA Intelligence database, Second European Symposium on Human Identification, Promega, Madison, WI, 1998, pp. 59–60.
[4] J.L. Weber, P.E. May, Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction, Am. J. Leg. Med. 44 (1989) 388–396.
[5] A. Edwards, A. Civileto, H.A. Hammond, C.T. Caskey, DNA typing and genetic mapping with trimeric and tetrameric tandem repeats, Am. J. Hum. Genet. 49 (1991) 746–756.
[6] A. Urquhart, C.P. Kimpton, T.J. Downes, P. Gill, Variation in short tandem repeat sequences: a survey of twelve microsatellite loci for use as forensic identification markers, Int. J. Leg. Med. 107 (1994) 13–20.
[7] H.A. Hammond, L. Jin, Y. Zhong, C.T. Caskey, R. Chakraborty, Evaluation of 13 short tandem repeat loci for use in personal identification applications, Am. J. Hum. Gen. 55 (1994) 175–189.
[8] C.P. Kimpton, P. Gill, A. Walton, A. Urquhart, E.S. Millican, M. Adams, Automated DNA profiling employing multiplex amplification of short tandem repeat loci, PCR Methods Appl. 3 (1993) 13–22.
[9] B. Brinkman, E. Meyer, Discrimination between populations using STRs, Proceedings of First European Symposium on Human Identification, Promega, Madison, WI, 1996, pp. 45–53.

[10] B. Budowle, T.R. Moretti, Al. Baumstark, D.A. Defenbaugh, K.M. Keys, Population data on the thirteen CODIS core short tandem repeat loci in African Americans, U.S. Caucasians, Hispanics, Bahamians, Jamaicans, and Trinidadians, J. Forensic Sci. 44 (1999) 1277–1286.
[11] B. Budowle, B. Shea, S. Niezgoda, R. Chakraborty, CODIS STR Loci data from 41 sample populations, J. Forensic Sci. 46 (2001) 453–489.
[12] K.M. Sullivan, S. Pope, P. Gill, J.M. Robertson, Automated DNA Profiling by fluorescent labelling of PCR product, PCR Methods Appl. 2 (1992) 34–40.
[13] C.T. Comey, B.W. Koons, K.W. Presley, J.B. Smerick, C.A. Sobieralski, D. Stanley, DNA extraction strategies for amplified fragment length polymorphism analysis, J. Forensic Sci. 39 (1994) 1254–1269.
[14] J.S. Waye, L. Presley, B. Budowle, C.G. Shutler, R.M. Fourney, A simple method for quantifying human genomic DNA in forensic specimen extracts, BioTechniques 7 (1989) 852–855.
[15] A. Edwards, H.A. Hammond, L. Jin, C.T. Caskey, R. Chakraborty, Genetic variation at five trimeric and tetrameric repeat loci in four human population groups, Genomics 12 (1992) 241–253.
[16] R. Chakraborty, P.E. Smouse, J.V. Neel, Population amalgamation and genetic variation: observation on artificially agglomerated tribal population of central and south America, Am. J. Hum. Gen. 43 (1988) 709–725.
[17] R. Chakraborty, M. Fornage, R. Guegue, E. Boerwinkle, Population genetics of hypervariable loci: analysis of PCR based VNTR polymorphism within a population, in: T. Burke, G. Dolf, A.J. Jeffreys, R. Wolff (Eds.), DNA Fingerprinting: Approaches and Applications, Birkhauser Verlag, Berlin, 1991, pp. 127–143.
[18] M. Nei, A.K. Roychouhdry, Sampling variances of heterozygosity and genetic distance, Genetics 76 (1974) 379–390.
[19] M. Nei, Estimation of average heterozygosity and genetic distance from a small number of individuals, Genetics 89 (1978) 583–590.
[20] S.W. Guo, Ea. Thompson, Performing the exact test of Hardy–Weinberg proportion for multiple alleles, Biometrics 48 (1992) 361–372.
[21] S. Karlin, E.C. Cameron, P.T. Williams, Sibling and parent–offspring correlation estimation with variable family size, Proc. Natl. Acad. Sci. U. S. A. 78 (1981) 2664–2668.
[22] B.S. Weir, C.C. Cockerham, Estimating *F*-statistics for the analysis of population structure, Evolution 38 (1984) 1358–1370.
[23] M.A. Tahir, K. Balamurugan, U.A. Tahir, M. Amjad, M.B. Owan, O.R. Chaudhary, J.E. Hamby, B. Budowle, R. Herrera, Allelic distribution of nine short tandem repeat (STR), HLA-DQA1, and Polymarker loci in an Omani sample population. Forensic Sci. Int. 109 (2000) 81–85, 38: 1358–1370.
[24] National Research Council II Report, The evaluation of forensic evidence, National Academy Press, Washington, DC, 1996.