Croat Med J 2005;46(4):587-592

African Jordanian Population Genetic Database on Fifteen Short Tandem Repeat Genetic Loci

Salem R. Yasin, Mawieh M. Hamad, Ali Z. ElKarmi, Adnan S. Jaran[1]

Department of Biological Sciences, Hashmite University, Zarqa; and [1]Department of Biological Sciences, Al-alBayt University, Mefraq, Jordan

Aim To establish a genetic database of the African-Jordanian population for forensic and paternity testing purposes.

Method Allelic distribution at fifteen short tandem repeat (STR) loci was determined for 95 healthy unrelated African-Jordanians. The 15 autosomal STR loci, included within the GenePrint® PowerPlex™ 16 system, were amplified from the subset of the 95 DNA extracts isolated from the population sample. Electrophoresis for each polymerase chain reaction (PCR) product was carried out using the ABI Prism 310 Genetic Analyzer and the length of the amplified DNA fragments was determined using the Genotype 2.0 and PowerTyper 16 Macro softwares. Calculations of allelic frequencies, forensic efficiency parameters, Hardy-Weinberg departure, and quantitative analysis of the allele frequencies in various populations were determined.

Results DNA extracts were successfully amplified and the genetic database was compiled. All tested loci showed no significant statistical deviation from Hardy-Weinberg expectations. Furthermore, no significant difference was observed between the sample population under investigation and other population genetic databases.

Conclusion The loci investigated here proved to be sufficiently polymorphic for forensic purposes, since the forensic efficiency values suggest that they are very discriminating in the African-Jordanian subpopulation.

Short tandem repeat (STR) genetic loci are highly polymorphic repeat sequences of nucleotides, which are abundant in eukaryotic DNA (1,2). Their importance rises from the fact that they are the most informative genetic markers providing high statistical capability of discrimination and individualization (2-6) in various forensic and judicial settings. Their importance and polymorphic nature is widely acknowledged and documented because of extensive forensic, medical, and ethnogenetic search that was prompted by various research communities worldwide (5-11) for proper utilization of their discrimination power and polymorphic nature in human identification, medical diagnosis, and linkage studies.

We describe here our work on compiling a genetic database for 15 forensic STR loci – D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, TH01, vWA, TPOX, CSF1PO, FGA, Penta D, and Penta E in African-Jordanians. Such genetic database would provide estimates of the frequency of a DNA profile in human identity testing, particularly in Jordan, and further shed lights on the polymorphic nature of these forensic loci. Earlier, we reported the allelic frequency distribution of STR loci in Caucasian Jordanians and demonstrated significant differences between Jordanian Caucasian population and Caucasian, Hispanic-, African-Americans, and black population of Central

West Africa at TH01, FES/FPS, and D5S818 STR genetic loci (12,13). These outcomes are in compliance with previous results that indicated that genetic variation and allelic frequency distribution differences did exist between various populations and even among different communities within the same population. Indeed, in one of the earlier studies, Lins et al (5) demonstrated such variation when they reported the allelic frequency distributions for several STR loci among the three American communities, ie the Caucasian, Hispanic, and African-Americans. Such observations were also reported when comparison studies were carried out, based upon the allelic frequency distributions, between the different communities living in various parts of the world (14-17). These studies and many others reported significant discrepancies between the various human subpopulations on various bases such as race, cast, and geographical parameters. Thus, genetic variation among human subpopulations and discrepancies in the allelic frequencies for the genetic loci of forensic interest could constitute a problem during human identification or paternity testing, especially when judicial settings are considered. Therefore, this prompted us to compile an African-Jordanian-population's STR genetic database to cover such variations in the Jordanian subpopulations for accurate and proper individualization through the determination of the allele/genotype frequency distributions of the fifteen forensic STR loci in interest.

Materials and Methods

Whole blood obtained by venipuncture from 95 healthy unrelated black Jordanians residing in the Jordan valley, Jordan, was collected in EDTA vacutainer tubes (Greiner Bio-One, Kremsmünster, Austria). Genomic DNA was extracted using the Wizard® Genomic DNA Purification Kit (Promega Corporation, Madison, WI, USA) according to the manufacturer's instructions. The quantity of recovered DNA was determined using the QuantiBlot® Human DNA Quantitation Kit (Applied Biosystems, Foster City, CA, USA).

Amplification by polymerase chain reaction (PCR) of the STR loci was performed in duplicates using the GenePrint® Powerplex™ 16 System (Promega) using 0.5-1.0 ng DNA in the 9600 thermal cycler (Applied Biosystems), according to the manufacturer's recommendations. PCR products were loaded on the CE310 Genetic Analyzer (Applied Biosystems) using ILS-600 (Promega) as internal lane standard. GeneScan analysis was performed on the raw data and alleles were labeled according to the international nomenclature (18) using the Genotyper 2.0 and PowerTyper™ 16 Macro Softwares by comparison of the PCR fragments with those of the allelic ladder.

Statistical Analysis

Allele frequencies and forensic efficiency parameters were calculated using the Power Stats Microsoft Excel workbook template provided by Promega Corporation (Madison, Wis., *http://www.promega.com/geneticidtools/*). Possible departure from the Hardy-Weinberg expectations at each locus was evaluated by the χ^2-test, as well as the G-test statistic at a 0.05 level of significance and the number of degrees of freedom was calculated as the number of genotypes minus the number of alleles. Percentage genotype (%GR) and percentage allele representation (%AR) were calculated as previously described (12,13). The allele frequencies between different populations were compared using the G-statistic homogeneity test STATISTICA software for Windows, 1995 version (StatSoft, Tulsa, OK, USA) was used for statistical analysis.

Results

The observed allele frequencies for the fifteen STR loci found in African-Jordanians are shown in Table 1. The data in Table 1 shows the most predominant and the least common alleles for the D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, TH01, vWA, TPOX, CSF1PO, FGA, Penta D, and Penta E STR genetic loci. Alleles 15, 12, 10, 13, 12, 11, 16, 29, 7, 16, 8, 10, 22, 13, and 8 were the most frequent alleles for the D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, TH01, vWA, TPOX, CSF1PO, FGA, Penta D, and Penta E STR genetic loci, respectively.

Possible divergence from Hardy-Weinberg expectations at each locus was evaluated and the number of degrees of freedom was calculated as the number of genotypes minus the number of alleles (Table 1). All loci showed no significant deviation ($P > 0.05$) from Hardy-Weinberg expectations.

Figure 1 shows both the %GR and %AR data for all the STR loci studied in the present

Table 1. Observed allele frequency distributions of the 15 STR loci in 95 unrelated African-Jordanians

D3S1358		*D5S818*		*D7S820*		*D8S1179*		*D13S317*	
Allele	Freq.*	Allele	Freq.	Allele	Freq.	Allele	Freq.	Allele	Freq.
14	0.068	8	0.047	7	0.016	8	0.005	8	0.063
15	0.284	9	0.021	8	0.153	9	0.005	9	0.032
16	0.274	10	0.068	9	0.111	10	0.005	10	0.053
17	0.237	11	0.216	10	0.405	11	0.068	11	0.226
18	0.137	12	0.442	11	0.253	12	0.132	12	0.489
		13	0.189	12	0.053	13	0.253	13	0.100
		14	0.005	13	0.011	14	0.237	14	0.037
		15	0.011			15	0.232		
						16	0.042		
						17	0.016		
						18	0.005		
P†	0.277		0.471		0.750		0.997		0.066

D16S539		*D18S51*		*D21S11*		*TH01*		*vWA*	
Allele	Freq.	Allele	Freq.	Allele	Freq.	Allele	Freq.	Allele	Freq.
5	0.005	9	0.005	27	0.005	6	0.205	11	0.011
8	0.068	10	0.005	28	0.163	7	0.363	12	0.005
9	0.142	11	0.011	29	0.289	8	0.168	13	0.005
10	0.132	12	0.142	30	0.247	9	0.147	14	0.089
11	0.274	13	0.068	30.2	0.011	9.3	0.089	15	0.158
12	0.216	14	0.095	31	0.021	10	0.026	16	0.242
13	0.153	15	0.116	31.2	0.068			17	0.216
14	0.011	16	0.174	32.2	0.079			18	0.179
		17	0.132	33.2	0.042			19	0.053
		18	0.089	34	0.011			20	0.037
		19	0.084	35	0.047			21	0.005
		20	0.021	36	0.005				
		21	0.026	37	0.011				
		22	0.011						
		24	0.021						
P	0.108		0.275		0.997		0.421		0.673

TPOX		*CSF1PO*		*FGA*		*Penta D*		*Penta E*	
Allele	Freq.	Allele	Freq.	Allele	Freq.	Allele	Freq.	Allele	Freq.
6	0.016	7	0.011	17	0.011	2.2	0.111	5	0.047
8	0.395	8	0.095	18	0.016	3.2	0.021	7	0.084
9	0.226	9	0.047	19	0.042	5	0.005	8	0.163
10	0.100	10	0.316	19.2	0.005	6	0.011	9	0.021
11	0.189	11	0.211	20	0.100	7	0.053	10	0.053
12	0.058	12	0.253	21	0.153	8	0.116	11	0.105
13	0.016	13	0.047	22	0.184	9	0.174	12	0.068
		14	0.021	22.2	0.011	10	0.142	13	0.084
				23	0.174	11	0.142	14	0.053
				23.2	0.005	12	0.105	15	0.111
				24	0.126	13	0.058	16	0.047
				25	0.063	14	0.053	17	0.079
				26	0.016	15	0.011	18	0.032
				27	0.021			19	0.021
				28	0.047			20	0.021
				29	0.021			21	0.005
				30	0.005			22	0.005
P	0.951		0.874		0.004		0.023		0.169

*Allele frequency.
†χ^2-test.

study. The percentage genotype representation for the STR genetic loci in the sample population ranged from 38.2% for D16S539 locus (10 alleles) to 5.3% for the FGA genetic locus (39 alleles), although the FGA locus was highly polymorphic locus with 17 different alleles. Disregarding the total number of alleles per each STR genetic locus, D5S818 (80% AR), D13S317 (80% AR), and Penta E (73.91% AR) alleles were the most highly expressed alleles, whereas D3S1358 alleles (38.46% AR) were the least expressed alleles in the sample population, among the 15 STR genetic loci examined (Fig. 1).

Forensic efficiency parameters, such as observed heterozygosity (H_{obs}), expected heterozygosity (H_{exp}), polymorphic information content (PIC), power of identity (P_I), power of discrimination (P_D), paternity exclusion (PE), and paternity index (PI), are shown in Table 2. The observed heterozygosity for the fifteen STR genetic loci

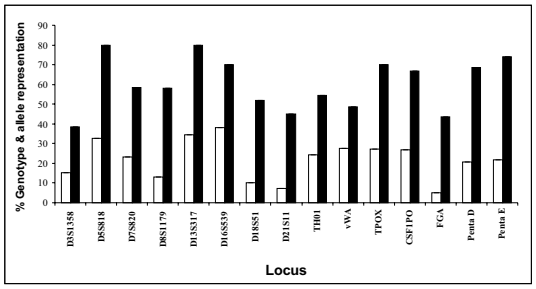

Figure 1. Percentage genotype representation (%GR) and percentage allelic representation (%AR) for the 15 STR genetic loci in African-Jordanians. Closed bars – %GR; open bars – %AR.

ranges from 65.3% for the D3S1358 to 94.7% for the FGA. The observed P_I ranges from 0.022 for the Penta E to 0.152 for the D13S317 locus. The combined P_I using the fifteen STR genetic loci in African-Jordanians was estimated as 1 in 7.818×10^{18} and hence the P_D was greater than 0.99999999999999999. The combined probability of PE value for the fifteen loci was also calculated at greater than 0.99964.

Quantitative comparisons of allele frequencies between African-Jordanians and other populations (data not shown) showed no significant statistical deviation ($P > 0.05$) when comparing African-Jordanian data with that of the USA Caucasians, African Americans, Hispanic Americans (5,6), Central West African population living in Spain (19), the black Choco population of Colombia (20), African-population of Gabon (21), and black Bubi population of Guinea (22) for the total number of markers studied. Furthermore, the African-Jordanian subpopulation did not differ significantly ($P > 0.05$) from Caucasian-Jordanians (13,23) at all STR loci tested.

Discussion

The forensic research literature provides an increasing number of STR population genetic databases, and points to the importance of compiling of own-population genetic databases, since evidence of populations and subpopulations differences at the STR genetic loci of forensic interest are mounting up. In this context, we embarked on compiling a genetic database for the African-Jordanian community in Jordan. The established genetic database has demonstrated that, according to the extremely high combined PD and PE values, the combination of the fifteen STR systems studied here is a powerful tool for forensic identification and paternity testing in African-Jordanians.

The absence of significant statistical deviation when comparing African-Jordanian data with that of some other African communities around the world for the total number of markers studied indicates a high level of homogeneity between these communities. However, further studies using other STR genetic loci might be required to assess the degree of such homogeneity. Furthermore, the finding that African-Jordanian sub-population did not differ significantly from Caucasian-Jordanians at all STR loci tested indicates that African- and Caucasian-Jordanians have similar allele distributions for these STR loci, thus suggesting a low diversity within the Jordanian population. The statistical similarities in the allelic frequency distributions between the two ethnic groups could be a reflection of the geographical

Table 2. Forensic efficiency parameters for the 15 STR loci in African-Jordanians*

	Statistical parameter						
STR loci	H_{obs}	H_{exp}	PIC	P_I	P_D	PE	PI
D3S1358	0.653	0.762	0.722	0.095	0.905	0.359	1.44
D5S818	0.737	0.712	0.672	0.145	0.855	0.488	1.90
D7S820	0.726	0.733	0.694	0.120	0.880	0.470	1.83
D8S1179	0.800	0.802	0.774	0.071	0.929	0.599	2.50
D13S317	0.695	0.904	0.654	0.152	0.848	0.420	1.64
D16S539	0.768	0.813	0.787	0.067	0.930	0.542	2.16
D18S51	0.874	0.888	0.878	0.039	0.961	0.742	3.96
D21S11	0.842	0.813	0.793	0.069	0.931	0.679	3.17
TH01	0.779	0.767	0.734	0.097	0.903	0.523	2.07
VWA	0.905	0.826	0.802	0.076	0.924	0.806	5.28
TPOX	0.768	0.743	0.706	0.103	0.897	0.542	2.16
CSF1PO	0.811	0.779	0.745	0.091	0.909	0.619	2.64
FGA	0.947	0.877	0.893	0.041	0.959	0.893	9.50
Penta D	0.874	0.883	0.872	0.039	0.961	0.742	3.96
Penta E	0.842	0.914	0.908	0.022	0.978	0.679	3.17

*Abbreviations: H_{obs} – observed heterozygosity; H_{exp} – expected heterozygosity; PIC – polymorphic information content; P_I – power of identity; P_D – power of discrimination; PE – paternity exclusion; PI – paternity index.

position of Jordan between Africa and Asia, which facilitated the migration of people in the far past, or due to the admixture of both subpopulations through inter-marriages over many centuries that allowed gene exchange to occur with no constraints, thus resulting in a certain degree of homogeneity in the genetic pool of the Jordanian population. The inter-marriage factor could have affected the genetic pool of the African subpopulation of Jordan more than that of the white population. This is due to the fact that inter-marriages occur more frequently between African men and Caucasian women, and it is less common between Caucasian men and African women. Also, allelic frequency distributions between the two ethnic groups are not identical, which suggests that each has its own separate genetic pool.

The number of genotypes increases exponentially with the number of alleles at each locus according to the formula $Genotypes = n(n+1)/2$, where n is the number of alleles (24). Thus to explore such relation between the population's pool of genotypes and the allelic windows of the tested genetic STR loci, we compared the population's observed genotypes (%GR) and the observed allelic window (%AR) of each of the tested genetic STR loci. The generated %GR and %AR data suggested that there was an inverse correlation between the %GR and %AR, total number of alleles per genetic loci and %AR, and total number of alleles per genetic loci and %GR. This generally means that, as the number of alleles per loci increases, %GR decreases. Likewise and generally, as the number of alleles per loci increases, the %AR decreases. These observations regarding the %GR and %AR of STR markers require further study using a larger number of genetic loci of different allelic windows and different populations in order to clarify this phenomenon and investigate its affect on the population's genetic pool.

In conclusion, a population database has been established for the African-Jordanians for the fifteen STR loci D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, TH01, vWA, TPOX, CSF1PO, FGA, Penta D, and Penta E. High combined power of discrimination for these loci shows their usefulness for forensic purposes.

Acknowledgements

The authors thank Mr Mosa Qandil and Miss Khawlah Salem for their technical assistance. This work was supported by a research grant from the Hashemite University, College of Graduate Studies and Scientific Research. Names of commercial manufacturers are provided for identification only and inclusion does not imply endorsement by the authors.

References

1 Weber JL, May PE. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am J Hum Genet. 1989;44:388-96.

2 Edwards A, Hammond HA, Jin L, Caskey CT, Chakraborty R. Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. Genomics. 1992;12:241-53.

3 Edwards A, Civitello A, Hammond HA, Caskey CT. DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. Am J Hum Genet. 1991;49: 746-56.

4 Hochmeister MN, Budowle B, Schumm JW, Sprecher CJ, Borer UV, Dirnhofer R. Swiss population data and forensic efficiency values on 3 tetrameric short tandem repeat loci-HUMTH01, TPOX, and CSF1PO-derived using a STR multiplex system. Int J Legal Med. 1995; 107:246-9.

5 Lins AM, Micka KA, Sprecher CJ, Taylor JA, Bacher JW, Rabbach DR, et al. Development and population study of an eight-locus short tandem repeat (STR) multiplex system. J Forensic Sci. 1998;43:1168-80.

6 Budowle B, Moretti TR, Baumstark AL, Defenbaugh DA, Keys KM. 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. 1999;44:1277-86.

7 Budowle B, Shea B, Niezgoda S, Chakraborty R. CODIS STR loci data from 41 sample populations. J Forensic Sci. 2001;46:453-89.

8 Ban JD. Establishing a large DNA data bank using the PowerPlex 1.1 and 2.1 systems. Croat Med J. 2001;42: 256-9.

9 Alonso A, Martin P, Albarran C, Garcia P, Primorac D, Garcia O, et al. Specific quantification of human genomes from low copy number DNA samples in forensic and ancient DNA studies. Croat Med J. 2003;44: 273-80.

10 Smyth C, Kalsi G, Brynjolfsson J, O'Neill J, Curtis D, Rifkin L, et al. Further tests for linkage of bipolar affective disorder to the tyrosine hydroxylase gene locus on chromosome 11p15 in a new series of multiplex British affective disorder pedigrees. Am J Psychiatry. 1996; 153:271-4.

11 Reato G, Basso G, Putti MC, Cignetti A, Guarini A, Foa R. Microsatellite analysis in childhood acute lymphoblastic leukemia. Haematologica. 1998;83:403-7.

12 Hamad M, Yasin SR, Elkarmi A. Polymorphism of HUMvWA31, HUMTH01, HUMF13A1 and HUMFES/FPS STR genetic loci in Jordanians. The Korean journal of genetics. 2001;23:157-61.

13 Yasin SR. Allele frequencies at nine PCR-based STR loci in Jordanians. Korean J Genetics. 2002;24:327-334.

14 Klintschar M, Ebner A, Reichenpfader B. Population genetic studies on nine tetrameric short tandem repeat loci using fluorescence dye-labeled primers and capillary electrophoresis in the Austrian population. Electrophoresis. 1999;20:1740-2.

15 Reddy BH, Sun G, Luis JR, Crawford MH, Hemam NS, Deka R. Genomic diversity at thirteen short tandem repeat loci in a substructured caste population, Golla, of southern Andhra Pradesh, India. Hum Biol. 2001;73: 175-90.

16 Destro-Bisol G, Boschi I, Caglia A, Tofanelli S, Pascali V, Paoli G, et al. Microsatellite variation in Central Africa: an analysis of intrapopulational and interpopulational genetic diversity. Am J Phys Anthropol. 2000; 112:319-37.

17 Corte-Real F, Andrade L, Anjos MJ, Carvalho M, Vieira DN, Carracedo A, et al. Population genetics of nine STR loci in two populations from Brazil. J Forensic Sci. 2000;45:432-5.

18 Bar 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. International Society for Forensic Haemogenetics. Int J Legal Med. 1997;110: 175-6.

19 Gamero JJ, Romero JL, Gonzalez JL, Arufe MI, Cuesta MI, Corte-Real F, et al. A study on ten short tandem repeat systems: African immigrant and Spanish population data. Forensic Sci Int. 2000;110:167-77.

20 Bravo ML, Moreno MA, Builes JJ, Salas A, Lareu MV, Carracedo A. Autosomal STR genetic variation in negroid Choco and Bogota populations. Int J Legal Med. 2001;115:102-4.

21 Steinlechner M, Schmidt K, Kraft HG, Utermann G, Parson W. Gabon black population data on the ten short tandem repeat loci D3S1358, VWA, D16S539, D2S1338, D8S1179, D21S11, D18S51, D19S433, TH01 and FGA. Int J Legal Med. 2002;116:176-8.

22 Gene M, Moreno P, Borrego N, Pique E, Brandt C, Mas J, et al. The Bubi population of Equatorial Guinea characterised by HUMTH01, HUMVWA31A, HUMCSF1 PO, HUMTPOX, D3S1358, D8S1179, D18S51 and D19S253 STR polymorphisms. Int J Legal Med. 2001; 114:298-300.

23 Salem K, Yasin S, ElKarmi A, Hamad M, Jaran A. Jordanian population data on five STR forensic loci: D16S539, TPOX, CSF1PO, Penta D, and Penta E. Legal Medicine 2003;5:251-252.

24 Tracey M. Short tandem repeat-based identification of individuals and parents. Croat Med J. 2001;42:233-8.

Received: August 8, 2004
Accepted: March 14, 2005

Correspondence to:
Salem R. Yasin
Hashemite University
Biological Sciences Department
P.O. Box 330077
Al-Zarqa 13133, Jordan
omarwyazan@yahoo.com