Available online at www.sciencedirect.com
SCIENCE DIRECT
Forensic Science International 158 (2006) 140–148
Forensic Science International
www.elsevier.com/locate/forsciint

Iranian STR variation at the fringes of biogeographical demarcation

E.M. Shepard, R.J. Herrera *

Department of Biological Sciences, Florida International University, University Park, OE 304, Miami, FL 33199, USA

Received 11 April 2005; received in revised form 10 May 2005; accepted 10 May 2005
Available online 5 July 2005

Abstract

The integrative relationship between population genetics and forensic biology allows for a thorough genetic characterization of extant human populations. This study aimed to genetically characterize 150 unrelated healthy donors from a general population in Iran both forensically and phylogenetically. The allelic frequencies of 15 STR loci (D8S1179, D21S11, D7S820, CSF1PO, D3S1358, TH01, D13S317, D16S539, D2S1338, D19S433, vWA, TPOX, D18S51, D5S818 and FGA) were generated. This constitutes the core of polymerase chain reaction (PCR)-based DNA genetic markers in the US Combined DNA Index System (CODIS) plus two additional loci (D2S1338 and D19S433) that together are consistent with several other worldwide database requirements. There were no deviations from Hardy–Weinberg expectations. Based upon the allelic frequencies, several important forensic parameters were calculated including: gene diversity (GD) index, power of discrimination (PD), polymorphic information content (PIC) and power of exclusion (PE). *G*-tests indicate the allelic frequencies of the Iranians are statistically non-significant compared to two Turkish populations yet, statistically different from the remaining 18 groups obtained from the literature and examined in this study. This suggests that the Iranian dataset may be forensically equivalent to the dataset from the Turkish region of Eastern Anatolia and the general population from Turkey. Phylogenetic analysis of our population with the full set of 15 loci indicate the Iranians occupy an intermediate position relative to the major Caucasian and East Asian clades on a global level. A regional phylogenetic analysis using 13 of the 15 loci indicate the Iranians segregate in a more compact association with groups from southeastern Spain, Arabs from Morocco and Syria, and especially with the general population from Turkey and those from Eastern Anatolia. These groups are flanked by highly differentiated populations from northern India and a Berber group from Tunisia on opposing ends of the regional phylogram. This report also demonstrates the necessity to thoroughly characterize the genetic composition of populations located in geographic intersections in order to choose the appropriate dataset on which to base forensic calculations, not only at an intra-population level, but also at an inter-population level as well.

Keywords: Short tandem repeats; Microsatellite markers; AmpF*l*STR Identifiler; Phylogeny; Forensic; Iran

1. Introduction

Population genetics and forensic biology are inextricably linked disciplines. The relationships among groups of populations and of intra-population characteristics complement each other in phylogenetic and forensic studies. Autosomal short tandem repeats (STRs) are hyper-variable markers that can provide the fine-resolution needed to determine relationships among closely related populations in recent evolutionary history [1–4] and at the same time provide the allelic distributions used in forensic calculations [5,6].

* Corresponding author. Tel.: +1 305 348 1258; fax: +1 305 348 1259.
E-mail address: herrerar@fiu.edu (R.J. Herrera).

0379-0738/$ – see front matter © 2005 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.forsciint.2005.05.012

In this study, our goal is to investigate the genetic characteristics and phylogenetic relationships of a population at the fringes of a biogeographical zone and at the crossroads of human migrations. Southwest Asia represents a region of paramount importance to population geneticists. More specifically, Iran sits at a tri-continental crossroads of human migration between Europe, Asia and Africa. The genetic studies involving this biogeographical area to date have mainly utilized sex-linked marker systems to elucidate lineage-specific pathways of demic movements. For instance, several studies employing mtDNA [7–9] and Y-chromosome [7,8] diversity have shown haplogroup distributions consistent with an initial settlement of Southwest Asia followed by migrations into south and central Asia with limited subsequent gene flow. In a different study, an Iranian population had a significant influence on the topological relationships derived from West Asian mtDNA haplotypes when included in the phylogenetic analyses [10,11]. Thus, populations from Iran are critical to studies encompassing the southwest and central Asian corridors [9].

With this in mind, we examined 15 hyper-variable autosomal STR loci in 150 Iranian individuals. We compared the allelic distributions to a set of databases taken from the literature [12–21] to assess the Iranians in a global context using 15 loci. Due to the limited availability of datasets with the full set of 15 loci, we also ascertained the allelic distribution patterns using a smaller subset of 13 core CODIS loci [22–27] in order to assess a more regional picture of the phylogenetic relationships in the proximate area surrounding Iran. Our results are the first to report on this set of autosomal forensic markers with regards to an Iranian population. Our phylogenetic analyses highlight the Iranians as a population intermediate relative to European and East Asiatic groups with a stronger affinity for the former on a global scale. At a regional level, the Iranian population exhibits genetic homogeneity to Turkey and the Eastern Anatolia region. The data reveal that significant genetic differences exist among the other groups examined from Southwest Asia and neighboring regions. This study also illustrates the need to characterize genetic diversity of individual populations both at an intra and inter-population level. This is especially so for populations at the boundaries of biogeographical regions and/or that are subject to admixture resulting from successive demic episodes and invasions, like Iran.

2. Materials and methods

2.1. Population information

Whole blood was collected from a total of 150 unrelated healthy donors from Iran with ancestry traced back at least two generations. Donors originated from all regions within modern-day Iran and therefore are representative of the general population as a whole.

2.2. DNA extraction

Blood was collected in EDTA Vacutainer™ tubes. Adherence to ethical guidelines was followed as stipulated by the Institutional Review Board (IRB) of Florida International University. DNA was extracted by standard phenol-chloroform protocols and followed by ethanol precipitation [28].

2.3. Polymerase chain reaction amplifications

PCR amplifications of the 15 STR loci (D8S1179, D21S11, D7S820, CSF1PO, D3S1358, TH01, D13S317, D16S539, D2S1338, D19S433, vWA, TPOX, D18S51, D5S818 and FGA) were carried out using an AmpF*l*STR Identifiler kit (Applied Biosystems, Foster City, CA). All reactions were performed as described by the manufacturer using the recommended amount of DNA (0.5–1.25 ng) in a GeneAmp PCR System 9600 thermal cycler (Applied Biosystems, Foster City, CA).

2.4. STR genotyping

DNA fragment separation and detection was completed in an ABI Prism 310 Genetic Analyzer (Applied Biosystems, Foster City, CA) CE instrument. ABI GeneScan500 LIZ was used as an internal size standard. GeneScan® 3.7 was used to determine the fragment sizes and Genotyper® 3.7 NT software was used to designate alleles by comparison with the allelic ladder provided by the manufacturer following the guidelines set forth by the DNA Commission of the International Society for Forensic Haemogenetics [29].

2.5. Statistical and phylogenetic analysis of data

For data analysis the Arlequin software package version 2.000 [30] was used to assess Hardy–Weinberg equilibrium (HWE) using Fisher's exact test [31] and to determine gene diversity indices [32]. Several forensic parameters were also examined including power of discrimination (PD), polymorphic information content (PIC) and power of exclusion (PE) using the PowerStats program version 1.2 [33,34]. Pairwise comparisons of the Iranians and all populations gathered from the literature were generated using *G*-tests in Carmody's software [35] on a global level (15 STR loci) and on a regional level (13 STR loci subset). Phylogenetic comparisons were made using Neighbor-Joining (NJ) trees with PHYLIP 3.52c software [36] as well as a multi-dimensional scaling (MDS) analyses performed using the Statistical Package for the Social Sciences (SPSS) software program [37] both based on *Fst* distances [38] for both global and regional datasets.

Table 1
Allele frequencies of 15 STR loci in Iran (n=150)

Allele	D8S1179	D21S11	D7S820	CSF1PO	D3S1358	TH01	D13S317	D16S539	D2S1338	D19S433	vWA	TPOX	D18S51	D5S818	FGA
6						0.2767									
6.2										0.0033					
7			0.0133			0.2100						0.0033			
8	0.0067		0.2167	0.0100		0.1333	0.1400	0.0400				0.5267		0.0167	
9	0.0200		0.0667	0.0233		0.2533	0.0767	0.1500				0.1067	0.0067	0.0567	
9.3						0.1067									
10	0.0867		0.2667	0.2400		0.0200	0.0567	0.0800				0.0767	0.0067	0.0967	
10.2													0.0033		
11	0.0600		0.2467	0.3333			0.3300	0.2733		0.0200		0.2367	0.0233	0.2867	
12	0.1333		0.1700	0.3433			0.3100	0.3333		0.0833		0.0500	0.1267	0.3567	
13	0.3067		0.0133	0.0433			0.0733	0.1100		0.2767	0.0033		0.1400	0.1700	
13.2										0.0300					
14	0.1633		0.0067	0.0067	0.0567		0.0100	0.0133		0.2167	0.1000		0.2033	0.0167	
14.2										0.0467					
15	0.1467				0.2367		0.0033			0.1400	0.1067		0.1367		
15.2										0.0933	0.0033				
16	0.0600				0.2900				0.0400	0.0467	0.2067		0.1433		
16.2										0.0367					
17	0.0167				0.2833				0.1900	0.0033	0.2667		0.0933		
17.2										0.0033					
18					0.1100				0.0833		0.2167		0.0500		0.0100
19					0.0233				0.0933		0.0833		0.0500		0.0533
19.2															0.0067
20									0.1167		0.0133		0.0167		0.0633
21									0.0567						0.2000
21.2															0.0067
22									0.0467						0.1367
22.2															0.0067
23									0.1833						0.2100
23.2															0.0100
24									0.0867						0.1533
24.2															0.0067
25									0.0833						0.0867
26									0.0200						0.0400
27		0.0200													0.0067
28		0.1700													0.0033
29		0.2100													
29.2		0.0033													
29.3		0.0033													
30		0.1867													
30.2		0.0233													
31		0.0367													
31.2		0.1467													
32		0.0100													
32.2		0.1333													
33.2		0.0467													
34.2		0.0067													
35		0.0033													
	Number of alleles														
	10	14	8	7	6	6	8	7	11	13	9	6	13	7	16
Ho	0.8600	0.7667	0.7200	0.7000	0.7200	0.7933	0.7800	0.7133	0.8533	0.7933	0.7867	0.7067	0.8333	0.7333	0.8333
He	0.8273	0.8556	0.7900	0.7155	0.7663	0.8054	0.7880	0.8083	0.8822	0.8370	0.8328	0.7450	0.8803	0.7511	0.8674
P-value	0.8763	0.1528	0.1626	0.9863	0.1989	0.8404	0.5660	0.3760	0.1729	0.5042	0.7396	0.6928	0.1773	0.4518	0.5892
GD	0.8273	0.8511	0.7900	0.7155	0.7663	0.7882	0.7634	0.7740	0.8821	0.8370	0.8134	0.6490	0.8721	0.7511	0.8602
PD	0.9436	0.9559	0.9215	0.8689	0.9043	0.9173	0.9002	0.9166	0.9680	0.9522	0.9360	0.8223	0.9639	0.8948	0.9593
PIC	0.8042	0.8300	0.7540	0.6568	0.7248	0.7520	0.7252	0.7378	0.8674	0.8153	0.7842	0.6024	0.8552	0.7093	0.8415
PE	0.7147	0.5387	0.4599	0.4180	0.4599	0.5867	0.5625	0.4492	0.7014	0.5867	0.5745	0.4386	0.6623	0.4817	0.6623

n: the number of individuals sampled; Ho: observed heterozygosity; He: expected heterozygosity; P-value: HWE, Fisher's exact test; GD: gene diversity index; PD: power of discrimination; PIC: polymorphic information content; PE: probability of exclusion.

Table 2a
G-test results for populations using 15 STR loci

Population	Abbrev.	AFA	CAB	BEL	EQG	IRN	JAP	MCH	MAL	MOZ	POL	TAI	TUT	USC
African American	AFA[a]		296.7	1261.6	250.3	801.7	2610.2	1694.7	1467.7	339.6	1040.5	2661.6	603.0	1405.6
Cabinda (Angola)	CAB[b]	0		1205.0	221.7	885.7	1814.9	1317.2	1230.1	339.9	1028.9	1776.2	439.7	1254.4
Belguim	BEL[c]	0	0		1165.1	415.7	1726.1	1186.9	1034.7	1350.3	256.3	1790.4	1410.6	215.5
Equatorial Guinea	EQG[d]	0.005*	0.021*	0		795.9	2000.3	1451.5	1264.7	316.2	994.6	2005.3	460.4	1275.2
Iran	IRN	0	0	0	0		993.7	697.9	542.2	914.6	371.2	944.3	1068.8	420.1
Japan	JAP[e]	0	0	0	0	0		500.5	789.1	2060.3	1414.9	600.7	2142.1	2075.6
Malay Chinese (Han)	MCH[f]	0	0	0	0	0	0		395.4	1593.1	1028.4	191.3	1464.4	1345.2
Malay Malay	MAL[f]	0	0	0	0	0	0	0		1453.9	909.9	575.5	1356.1	1193.5
Mozambique	MOZ[g]	0	0	0	0	0	0	0	0		1139.6	2240.8	560.5	1408.7
North Poland	POL[h]	0	0	0	0	0	0	0	0	0		1445.3	1047.4	258.6
Taiwan (Han)	TAI[i]	0	0	0	0	0	0	**0.4370**	0	0	0		2255.1	2192.1
Tutsi (Rwanda)	TUT[j]	0	0	0	0	0	0	0	0	0	0	0		1556.0
U.S. Caucasian	USC[a]	0	0	0.029*	0	0	0	0	0	0	0	0	0	

The *G* statistic and *P*-values for the populations with 15 STR loci occupy the upper and lower levels of the diagonals, respectively. The non-significant *P*-values are in bold ($\alpha = 0.05$). *P-values, which became non-significant with the Bonferroni correction for multiple comparisons ($\alpha = 0.00064$). References: [a][12]; [b][13]; [c][14]; [d][21]; [e][15]; [f][16]; [g][17]; [h][18]; [i][20]; [j][19].

3. Results

3.1. Intra-population STR diversity

There were no deviations from Hardy–Weinberg expectations detected in any of the 15 loci analyzed for this Iranian population (see Table 1). The overall combined matching probability is 1 in 2.20×10^{17} and the combined power of exclusion is 0.99999612. The combined power of discrimination for this set of markers was robust (>0.999999999999999). As expected the most polymorphic loci were also the most discriminating: D2S1338 (PD: 96.80%), D18S51 (PD: 96.39%), and FGA (PD: 95.93%). These results strongly support the use of this set of genetic markers for personal identity testing. A rare allele in one individual was encountered in locus D21S11 (29.3) that has been reported elsewhere [39]. In addition, a possible new variant in a single individual was found in locus D19S433 (6.2), which is the smallest allele in that particular locus reported thus far to the authors' knowledge. Also, a second novel variant in one sample was observed in locus vWA (15.2). All three allele designations were confirmed by re-amplification and reanalysis.

3.2. Inter-population STR diversity

In order to test for any genetic differences of statistical significance in the populations examined in this study, *G*-test comparisons were made between pairs of worldwide geographically-targeted populations based on the 15 autosomal STR loci and between regional populations from Turkey, southwestern Europe, India and Northwest Africa using the 13 autosomal STR markers (see Tables 2a and 2b). The 13 loci examined in the regional study are a subset of the 15 total markers analyzed in the worldwide set. Of a total of 78 pairwise comparisons among worldwide populations (Table 2a), non-significant genetic differences were detected only between the Han Chinese from Taiwan in relation to the

Table 2b
G-test results for populations using 13 STR loci

Population	Abbrev.	AND	EAN	IRN	JAT	KHA	MOR	SYR	TUNB	TUR
Andalusia (SW Spain)	AND[k]		281.4	249.5	389.0	477.9	214.6	227.3	279.7	328.9
East Anatolia (Turkey)	EAN[l]	0		160.2	312.9	388.8	254.2	188.3	224.6	214.8
Iran	IRN	0	**0.1810**		313.6	419.2	211.9	183.9	240.9	182.0
India - Jat	JAT[m]	0	0	0		246.1	359.6	321.4	330.5	398.3
India - Khatris	KHA[m]	0	0	0	0		448.7	386.6	325.1	557.6
Morocco - Arab	MOR[n]	0	0	0	0	0		216.3	228.2	318.7
Syria - Arab	SYR[n]	0	0.003*	0	0	0	0		247.1	232.5
Tunisia - Berber	TUNB[o]	0	0	0	0	0	0	0		304.1
Turkey (general)	TUR[p]	0	0	0.01*	0	0	0	0	0	

The *G* statistic and *P*-values for the populations with 13 STR loci occupy the upper and lower levels of the diagonals, respectively. The non-significant *P*-values are in bold ($\alpha = 0.05$). *P-values, which became non-significant with the Bonferroni correction for multiple comparisons ($\alpha = 0.00138$). References: [k][23]; [l][26]; [m][22]; [n][25]; [o][24]; [p][27].

Han Chinese from Malaysia when $\alpha = 0.05$. Subsequent to the application of the Bonferroni correction for multiple comparisons ($\alpha = 0.05/78 = 0.000641$), three additional pairwise combinations were found to possess statistically insignificant genetic differences: Equatorial Guinea to African Americans; Equatorial Guinea to Cabinda, Angola; and US Caucasians to Belgium. Regionally, of all 36 combinations compared, statistically non-significant differentiation exists in a single pair: Iran to Eastern Anatolia when $\alpha = 0.05$. Two *P*-values became statistically non-significant

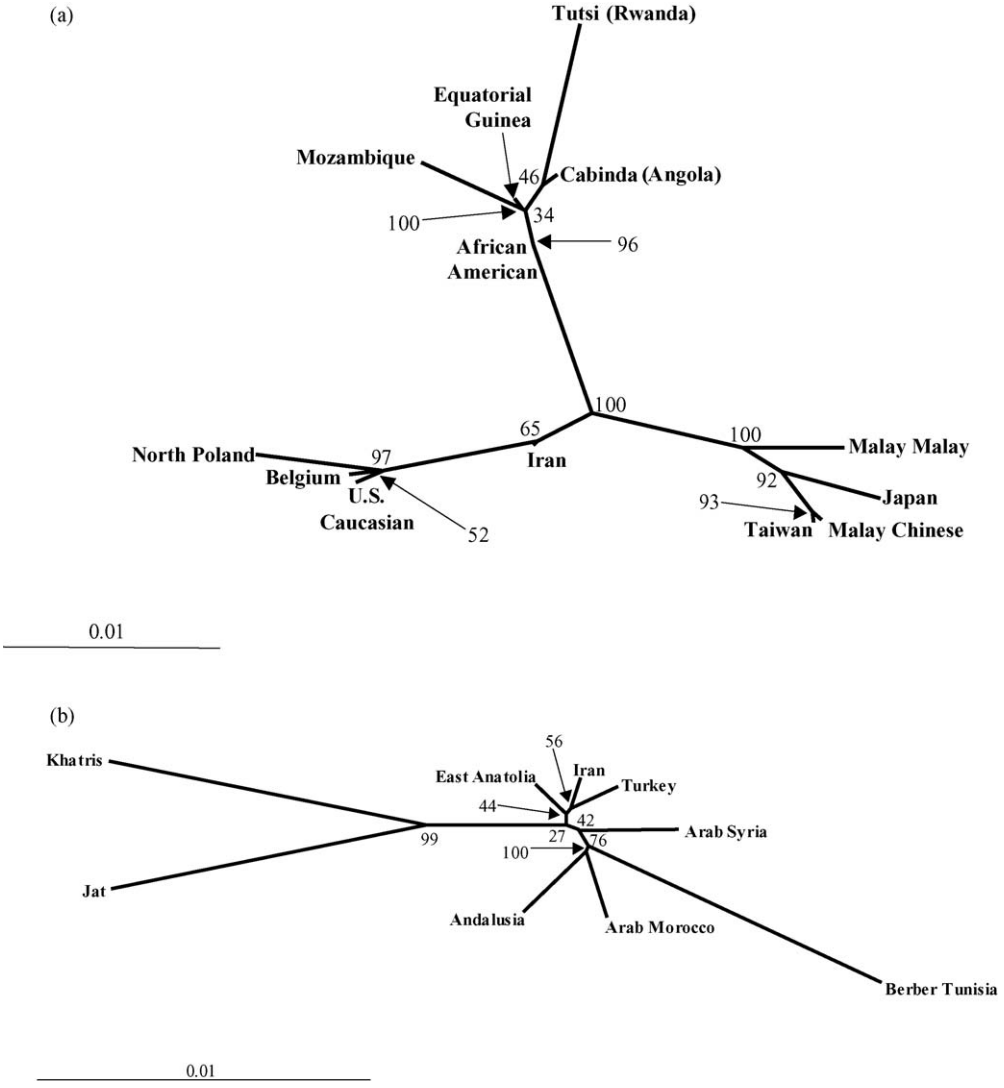

Fig. 1. (a) Neighbor-Joining (NJ) phylogenic analysis was performed using *Fst* distances generated from thirteen worldwide, geographically-targeted populations based on allele frequencies from 15 STR loci. The Gendist option of the PHYLIP software created branch distances onto which the corresponding bootstrap values (based on 1000 replications) were transferred to the corresponding nodes of the NJ tree. *Note*: In the African clade the arrow with a 96% bootstrap value corresponds to the node that separates the African Americans from the rest of the African populations. The arrow with the 100% reiteration indicates the separation of Mozambique from Equatorial Guinea while the third arrow points to the location of Equatorial Guinea. Moving clockwise, the arrow within the Asiatic clade indicates the bifurcation separating the Han Chinese from Taiwan from the Han Chinese from Malaysia with a confidence value of 93%. Finally, the fifth arrow indicates the segregation of Belgium from the US Caucasians with a confidence value of 52%. (b) Neighbor-Joining (NJ) phylogenic analysis was done using *Fst* distances utilizing nine regional populations from Turkey, Southeastern Europe, India and Northwest Africa based on allele frequencies of 13 STR loci. The Gendist option of the PHYLIP software created branch distances onto which the corresponding bootstrap values (based on 1000 replications) were transferred to the corresponding nodes of the NJ tree. The arrow indicates the split between Eastern Anatolia from both Iran and Turkey with a bootstrap value of 44% while the Iranians separate from the Turks 56% of the time. Finally Andalusia segregates from the Arabic Moroccan group in 100% of the replications.

after the application of the Bonferroni adjustment for Type I error ($\alpha = 0.05/36 = 0.00138$): Arabs from Syria compared to East Anatolia and Iran compared to a general population from Turkey.

Phylogenetic analyses were conducted using the NJ and MDS methods (see Fig. 1a and b as well as Fig. 2a and b, respectively) on the worldwide and regional sets of data (15 loci and a 13 loci subset, respectively). The topological arrangement of the populations within the worldwide NJ tree (Fig. 1a) follows partitioning along major racial and geographical lines, as expected. The phylogram contains only two bootstrap values below 50% incidence (46% and 34%). Three distinct groups of populations can be identified, namely the African, East Asian and Caucasian clusters, the latter of which includes Iran. Despite the fact that the Iranians under study separate with the Caucasian populations, they occupy an intermediate position in closer proximity to the trifurcation (bootstrap = 100%) of the three major groups. In fact, Iran splits from the remaining groups (bootstrap = 65%) in a transitional position between the Caucasian and East Asian groups while the Africans remain more distant. When a regional set of populations relative to Iran was analyzed (Fig. 1b), a clear more compact clustering of populations is evident. This group of associated popula-

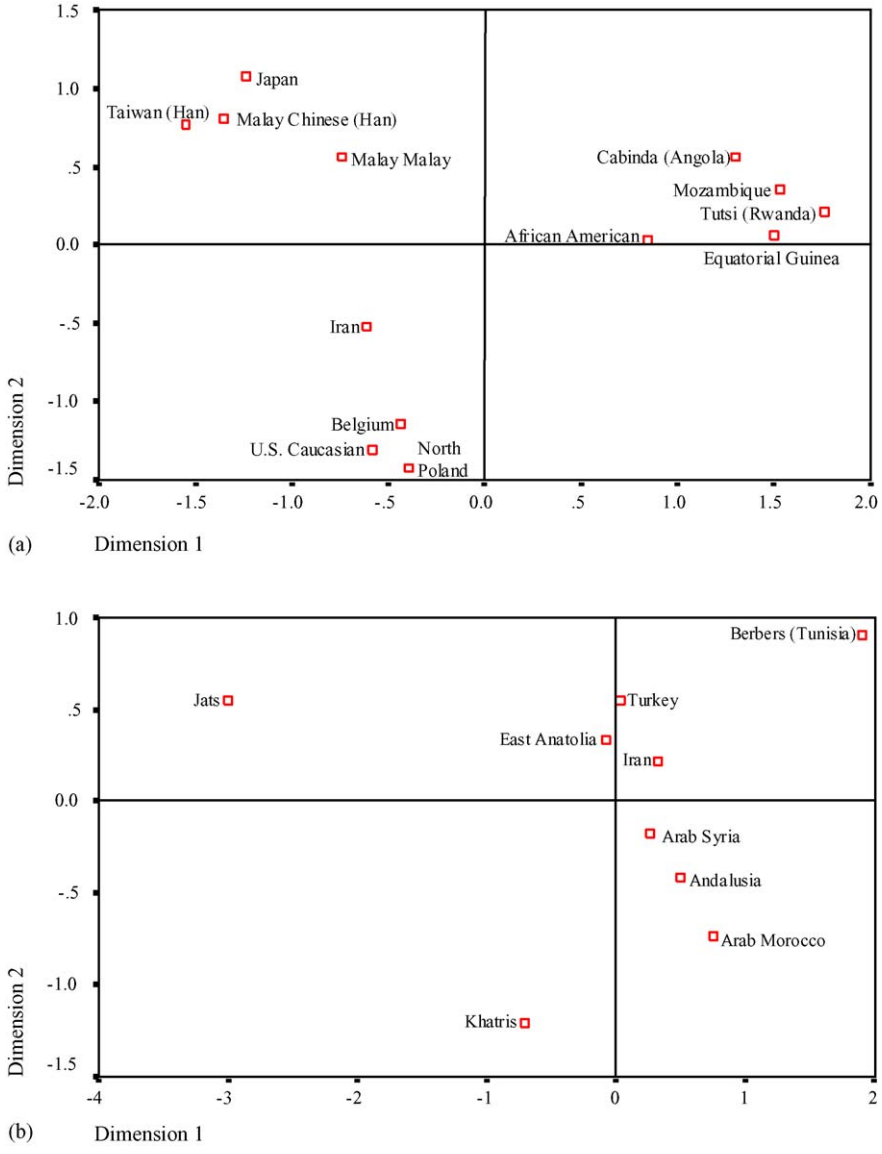

Fig. 2. (a) Multi-dimensional scaling (MDS) analyses using *Fst* distances of 13 worldwide populations based on allele frequencies of 15 STR loci. The stress value is 0.058. (b) Multi-dimensional scaling (MDS) analyses using *Fst* distances of nine regional populations based on allele frequencies of 13 STR loci. The stress value is 0.022.

tions is from southwestern Europe (Andalusia), North Africa (Moroccan Arabs) and southwestern Asia (Syrian Arabs, Turkey and the Eastern Anatolia region of Turkey as well as Iran). The two groups from the northern region of the Indian sub-continent (Khatris and Jats) and the Berber group from Tunisia represent the extreme bounds on opposing ends of the dendrogram (bootstrap values = 99% and 78%, respectively). Andalusia and Arabs from Morocco bifurcate together away from the Berber population with a bootstrap value equal to 76%. The Arabs from Syria split (bootstrap value = 42%) midway between these three groups and the cluster of East Anatolia, Iran and Turkey. Within that particular cluster East Anatolia separates initially (bootstrap = 44%) and then Iran and Turkey segregate from each other with a 56% incidence.

A strong parallelism exists between the NJ analysis and the topology of the MDS plots generated for both sets of populations (see Fig. 2a and b). The proportion of variance accounted for by the corresponding distances of the scaled data is 98.38% for the global and 99.84% for the regional MDS plots. Three main clusters of populations are evident in the worldwide analysis (Fig. 2a): Africans in the upper right quadrant, East Asians in the upper left quadrant and Caucasian groups with the Iranians in the lower left quadrant. Iran again falls into an intervening position between the Caucasian and East Asian cluster, in closer proximity to the former. In the regional MDS analysis (Fig. 2b) three populations, the Khatris and Jats from northern India and the Tunisian Berbers are outliers in the two dimensional plot. The Khatris and Jats lie in the outskirts of the upper and lower left quadrants and the Tunisian Berbers lie within the furthest right corner of the upper right quadrant. The remaining populations form a linear continuum running diagonally to the *X* and *Y*-axis with the Turkish and Moroccan groups at the extreme ends. The Iranian population in this study segregates in close proximity to the two Turkish groups on one side and the Syrian Arabs on the other.

4. Discussion

Autosomal STR allelic frequencies for 150 Iranian individuals are reported here for the first time. In the regional set of pairwise comparisions using a subset of the full marker set (13 of the 15 loci) Iran exhibits genetic homogeneity in relation to the Turks from Eastern Anatolia ($\alpha = 0.05$), which extends to the general population from Turkey after applying the Bonferroni correction for Type I error ($\alpha = 0.00138$). However, in the worldwide *G*-tests (full set of 15 loci) the Iranian group was found to be statistically different from all 12 databases ($\alpha = 0.05$) which persist even after the application of the Bonferroni adjustment for multiple comparisons ($\alpha = 0.000641$). These results suggest that a unique database for the Iranians will be useful in order to ascertain accurate probabilities of inclusion for the complete set of markers. In addition, Iran stands as genetically different from populations from southwestern Spain, North Africa and northern India in the regional analysis. Indeed, this last point is corroborated in the literature by several studies using both mtDNA [7–9] and Y-chromosome specific [7,8] markers. These reports have concluded that populations east of the Indus Valley have received little relative gene flow from western Asia subsequent to the initial colonization of South and Central Asia by anatomically modern humans. These findings are also congruent to an mtDNA report in which the inclusion of an Iranian population proved to cause a pivotal shift in the phylogenetic relationships among groups from the neighboring Caucasus region [11], thus underscoring the unique genetic characteristics of the Iranians.

At critical crossroads between Europe, Africa and the Far East, Iran represents a point of contact for many cultures and a corridor for demic migrations since modern man dispersed out of Africa. The worldwide set of populations in the NJ analysis indicated that the Iranians fall into an intermediary genetic position, especially relative to the East Asian and Caucasian clusters. Iran segregates within the Caucasian clade and shows no obvious differences in allelic distribution from the other three groups in that same cluster. It is likely that an accumulation of differences in allelic frequency patterns resulted in the singular separation apparent in the tree. This median position is mirrored in the topology of the MDS plot for the global set of populations.

To increase the resolution of the genetic analyses of Iran to other groups, we used a subset of the full marker system (13 of 15 STR loci) to perform a more regional phylogenetic inter-population comparison. As expected by incorporating neighboring populations in the NJ tree, Iran no longer displays an intermediary position as in the global analysis but a more compact genetic affinity towards groups from southwest Spain, Arabic populations from Morocco and Syria and especially towards populations of Turkish ancestry. The Eastern Anatolians and a general population from Turkey both segregate with the Iranians in a condensed cluster. This result in conjunction with the genetic homogeneity demonstrated in the *G*-tests involving these same three populations underscores their geographic proximity and historical socio-cultural affinities. It is probable that the commonalities that exist between Anatolia and the northern part of Iran date back to at least the birth of the agricultural revolution in this region, and therefore, they may have shared a common gene pool and/or have experienced gene flow in the past. Genetic homogeneity in the area is also reflected in a study involving mtDNA that found surprisingly little genetic variation between five Iranian groups [9]. In light of our results, this suggests some degree of genetic homogeneity in the region. Also, the results from the regional tree corroborate the *G*-tests that indicate that the northern Indian populations (Khatris and Jats) are vastly different from each other as well as from the West Asian groups. This data again substantiates the relatively small amount of recent gene flow into the Indian subcontinent from the west reported in previous studies [7–9].

The regional inter-population relationships in the MDS analysis reiterates much of the information inferred from the regional NJ tree, namely the isolation of the northern Indian groups and the Berbers from Tunisia as opposed to the genetic continuum formed by the diagonal laddering of the remaining populations, capped on the lower end by the Arabs from Morocco and by the general population of Turks on the upper end. This may be the result of common ancestry or gene flow involving the groups on the diagonal. In this regional MDS analysis, Iran remains in close proximity to the Turkish populations as in the NJ tree but is also relatively equidistant between the general population from Turkey and the Arabs from Syria. This may suggest that a region of genetic homogeneity may have included Syria as well.

To our knowledge no additional Middle Eastern/western Eurasian populations have been published using this complete set of 15 forensic markers. In terms of populations genetics, the inter-population results underscore the unique characteristics of populations on the fringes of biogeographical demarcation. Statistically, our data indicates that the Iranian database may be forensically equivalent to the dataset from East Anatolia and, the general population of Turks. Yet, this report also demonstrates the need to examine the genetic structure of populations on a case-by-case basis as the Iranian database is statistically differentiated from other neighboring groups. Overall, this study suggests that populations from geographically intermediate regions require careful characterization in order to choose the appropriate dataset on which to base forensic calculations, not only at an intra-population level, but also at an inter-population level as well.

Acknowledgements

The authors gratefully acknowledge Dr. Dee Mills and Liliana Moreno of the DNA Profiling Facility (DPF) at Florida International University and especially Arif Kalantar of the Forensic Laboratory at the Dubai Police Department in the United Arab Emirates for their contribution to this study. The authors would also like to thank Tenzin Gayden and Alicia Cadenas for their technical assistance, as well as Jason Somarelli for his intellectual input.

References

[1] A. Bowcock, A. Ruiz-Linares, J. Tomfohrde, E. Minch, J.R. Kidd, L.L. Cavalli-Sforza, High resolution of human evolutionary trees with polymorphic microsatellites, Nature 368 (1994) 455–457.

[2] D.J. Rowold, R.J. Herrera, Inferring recent human phylogenies using forensic STR technology, For. Sci. Int. 133 (2003) 260–265.

[3] E. Bosch, F. Calafell, A. Perez-Lezaun, J. Clarimon, D. Comas, E. Mateu, R. Martinez-Arias, B. Morera, Z. Brakez, O. Akhayat, A. Sefiani, G. Hariti, A. Cambon-Thomsen, J. Bertranpetit, Genetic structure of north-west Africa revealed by STR analysis, Eur. J. Hum. Genet. 8 (2000) 360–366.

[4] J.K. Lum, L.B. Jorde, W. Schiefenhovel, Affinities among Melanisians, Micronesians and Polynesians: a neutral biparental genetic perspective, Hum. Biol. 74 (3) (2002) 413–430.

[5] P.J. Collins, L.K. Hennessy, C.S. Leibelt, R.K. Roby, D.J. Reeder, P.A. Foxhall, Developmental validation of a single-tube amplification of the 13 CODIS STR loci, D2S1338, D19S433 and amelogenin: the AmpFISTR Identifiler PCR Amlification kit, J. For. Sci. 49 (6) (2004) 1265–1277.

[6] C. Leibelt, B. Budowle, P. Collins, Y. Daoudi, T. Moretti, G. Nunn, D. Reeder, R. Roby, Identification of a D8S1179 primer binding site mutation and the validation of a primer designed to recover null alleles, For. Sci. Int. 133 (3) (2003) 220–227.

[7] T. Kivisild, S. Rootsi, M. Metspalu, S. Mastana, K. Kaldma, J. Parik, E. Metspalu, M. Adojaan, H.-V. Tolk, V. Stepanov, M. Golge, E. Usanga, S.S. Papiha, C. Connioglu, R. King, L. Cavalli-Sforza, P.A. Unerhill, R. Villems, The genetic heritage of the earliest settlers persist both in indian tribal and caste populations, Am. J. Hum. Genet. 72 (2003) 313–332.

[8] M. Metspalu, T. Kivisild, E. Metspalu, J. Parik, G. Hudjashov, K. Kaldma, P. Serk, M. Karmin, D.M. Behar, M.T.P. Gilbert, P. Endicott, S. Mastana, S.S. Papiha, K. Skorecki, A. Torroni, R. Villems, Most of the extant mtDNA boundaries in South and Southwest Asia were likely shaped during the initial settlement of Eurasia by anatomically modern humans, BMC Genet. 5 (2004) 26.

[9] L. Qunitana-Murci, R. Chaiz, R. Spencer Wells, D.M. Behar, H. Sayar, R. Scozzari, C. Rengo, N. Al-Zahery, O. Semino, A.S. Santachiara-Benerecetti, A. Coppa, Q. Ayub, A. Mohyuddin, C. Tyler-Smith, S.Q. Mehdi, A. Torroni, K. McElreavey, Where east meets west: the complex mtDNA landscape of the southwest and central Asian corridor, Am. J. Hum. Genet. 74 (2004) 827–845.

[10] I. Nasidze, M. Stoneking, Mitochondrial DNA variation and language replacements in the Caucasus, Proc. R. Soc. Lond. B 268 (2001) 1197–1206.

[11] I. Nasidze, E.Y.S. Ling, D. Quinque, I. Dupanloup, R. Cordaux, S. Rychkov, O. Naumova, O. Zhukova, N. Sarraf-Zadegan, G.A. Naderi, S. Asgary, S. Sardas, D.D. Farhud, T. Sarkisian, C. Asadov, A. Kerimov, M. Stoneking, Mitochondrial DNA and Y-chromosome variation in the Caucasus, Ann. Hum. Genet. 68 (2004) 205–221.

[12] J.M. Butler, R. Schoske, P.M. Vallone, J.W. Redman, M.C. Kline, Allele frequencies for 15 autosomal STR loci on US Caucasian, African American and Hispanic populations, J. For. Sci. 48 (2003) 908–911.

[13] S. Beleza, C. Alves, F. Reis, A. Amorim, A. Carracedo, L. Gusmao, 17 STR (AmpF/STR Identifiler and Powerplex 16 system) from Cabinda (Angola), For. Sci. Int. 141 (2004) 193–196.

[14] R. Decorte, M. Engelen, L. Larno, K. Nelissen, A. Gilissen, J.J. Cassiman, Belgian population data for 15 STR loci (AmpF/STR SGM Plus and AmpF/STR Profiler PCR amplification kit), For. Sci. Int. 139 (2004) 211–213.

[15] M. Hashiyada, Y. Itakura, T. Nagashima, M. Nata, M. Funayama, Polymorphism of 17 STRs by multiplex analysis in Japanese population, For. Sci. Int. 133 (2003) 250–253.

[16] L.H. Seah, N.H. Jeevan, M.I. Othman, P. Jaya, Y.S. Ooi, P.C. Wong, S.S. Kee, STR data for the AmpF/STR Identifiler loci in

three ethic groups (Malay, Chinese, Indian) of the Malaysian population, For. Sci. Int. 138 (2003) 134–137.
[17] C. Alves, L. Gusmao, A. Damasceno, B. Soares, A. Amorim, Contribution for an African autosomal STR database (AmpF/STR Identifiler and Powerplex 16 system) and a report on genotypic variations, For. Sci. Int. 139 (2004) 201–205.
[18] Z. Szczerkowska, E. Kapinska, J. Wysocka, L. Cybulska, Northern polish population data and forensic usefulness of 15 autosomal STR loci, For. Sci. Int. 144 (2004) 69–71.
[19] M. Reguiro, personal communication.
[20] C.W. Wang, D.P. Chen, C.Y. Chen, S.C. Lu, C.F. Sun, STR data for the AmpF/STR SGM Plus and Profiler loci from Taiwan, For. Sci. Int. 138 (2003) 119–122.
[21] C. Alves, L. Gusmao, A.M. Lopez-Parra, M. Soledad Mesa, A. Amorim, E. Arroyo-Pardo, STR allelic frequencies for an African population sample (Equatorial Guinea) using AmpF/STR Identifiler and Powerplex 16 kits, For. Sci. Int. 148 (2005) 239–242.
[22] M. Tandon, R. Trivedi, V.K. Kashyap, Genomic diverstiy at 15 fluorescent labeled short tandem repeat loci in few important populations of state of Uttar Pradesh, India, For. Sci. Int. 128 (2002) 190–195.
[23] P. Sanz, V. Prieto, I. Flores, Y. Torres, M. Lopez-Soto, N.J. Farfan, Population data of 13 STRs in southern Spain (Andalusia), For. Sci. Int. 119 (2001) 113–115.
[24] L. Cherni, B. Loueslati Yaacoubi, L. Pereira, C. Alves, H. Khodjet El Kill, A.B. Ammar El Gaaied, A. Amorim, Data for 15 autosomal STR markers (Powerplex 16 system) from two Tunisian populations: Kesra (Berber) and Zriba (Arab), For. Sci. Int. 147 (2005) 101–106.
[25] L. Abdin, I. Shimada, C. Hohoff, Analysis of 15 short tandem repeats reveals significant differences between the Arabian populations from Morocco and Syria, Legal Med. 5 (2003) S150–S155.
[26] A. Hadi Cakir, A. Celebioglu, S. Altunbas, Yardimci, Allele frequencies for 15 loci in Van-Agri districts of the East Anatolian region of Turkey, For. Sci. Int. 135 (2003) 60–63.
[27] B.S. Akbasak, B. Budowle, D.J. Reeder, J. Redman, M.C. Kline, Turkish population data with the CODIS multiplex short tandem repeat loci, For. Sci. Int. 123 (2001) 227–229.
[28] J. Sambrook, D.W. Russell, Molecular Cloning: A Laboratory Manual, third ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA, 2001.
[29] DNA recommendations, report concerning further recommendations of the DNA Commission of the ISFH regarding PCR-based polymorphisms in STR (short tandem repeat) systems, Int. J. Legal Med. 107 (1994) 159–160.
[30] S. Schneider, J.M. Kueffer, D. Roessli, L. Excoffier, Arlequin v. 2000: a software for population genetics data analysis, in: Genetics and Biometry Laboratory, University of Geneva, Switzerland, 2000.
[31] S. Guo, E. Thompson, Performing the exact test of Hardy–Weinberg proportion for multiple alleles, Biometrics 48 (1992) 361–372.
[32] M. Nei, Molecular Evolutionary Genetics, Columbia University Press, New York, USA, 1987.
[33] A. Tereba, Tools for Analysis of Population Statistics, Profiles in DNA, vol. 2, Promega Corporation, 1999, pp. 14–16
[34] C. Brenner, J. Morris, Paternity index calculations in single locus hypervariable DNA probes: validation and other studies, in: Proceedings for the International Symposium on Human Identification 1989, Promega Corporation, Madison, WI, 1990, pp. 21–53.
[35] G. Carmody, *G*-test, Carleton University, Ottawa, Canada, 1991.
[36] J. Felsenstein, Phylogeny Inference Package (PHYLIP) version 3. 6a3 distributed by author, Department of Genetics, University of Washington, Seattle, USA, 2002.
[37] SPSS for Windows, Rel. 11.0.1.2001. Chicago: SPSS, Inc.
[38] J. Reynolds, B.S. Weir, C.C. Cockerham, Estimation of the coancestry coefficient: basis for a short term genetic distance, Genetics 105 (1983) 767–779.
[39] http://www.cstl.nist.gov/biotech/strbase/var_tab.htm.