academicJournals

Vol. 13(11), pp. 1210-1218, 12 March, 2014
DOI: 10.5897/AJB2013.12923
ISSN 1684-5315 ©2014 Academic Journals
http://www.academicjournals.org/AJB

African Journal of Biotechnology

Full Length Research Paper

Genetic variation of twenty autosomal STR loci and evaluate the importance of these loci for forensic genetic purposes

Imad Hadi[1]*, Mohammed Abdullah[1], Aamera Jaber[2] and Cheah Yoke[3]

[1]Department of Molecular Biology, Babylon University, Hilla City, Iraq.
[2]Institute of medico-legal in Baghdad, Ministry of Health of Iraq, Iraq.
[3]Department of Biomedical Science, University Putra Malaysia, Selangor, Malaysia.

Received 14 June, 2013; Accepted 26 February, 2014

The aim of this study was of twofold. One was to determine the genetic structure of Iraq population and the second objective of the study was to evaluate the importance of these loci for forensic genetic purposes. FTA® Technology (FTA™ paper DNA extraction) utilized to extract DNA. Twenty (20) STR loci and Amelogenin), including D3S1358, D13S317, Penta E, D16S539, D18S51, D2S1338, CSF1PO, Penta D, THO1, vWA, D21S11, D7S820, TPOX, D8S1179, FGA, D2S1338, D5S818, D6S1043, D12S391, D19S433 and Amelogenin amplified by using power plex21® kit. Polymerase chain reaction (PCR) products detected by genetic analyzer 3730xL then data analyzed by PowerStatsV1.2. Based on the allelic frequencies, several statistical parameters of genetic and forensic efficiency have been estimated. This includes the homozygosity and heterozygosity, effective number of alleles (n), the polymorphism information content (PIC), the power of discrimination (DP) and the power of exclusion (PE). The power of discrimination values for all tested loci was from 75 to 96%; therefore, those loci can be safely used to establish a DNA-based database for Iraq population. The high PIC values of the selected markers confirm their usefulness for genetic polymorphism studies and linkage mapping programs in human as well. The mean heterozygosity observed, is expected to have mean PIC values across the 20 loci which were 0.77, 0.81 and 0.78, respectively, indicating high gene diversity.

Key words: Autosomal STR, genetic variation, Iraq, statistical parameters.

INTRODUCTION

Microsatellites refer to DNA with varying numbers of short tandem repeats (Klintschar et al., 2006) between a unique sequence. DNA regions with repeat units that are 2 bp to 7 bp in length or most generally short tandem repeats (STRs) or simple sequence repeats (SSRs) are generally known as microsatellites (Ellegren, 2004).

In the core repeated bases, long repeat units may contain several hundred to thousands (Butler and Hill, 2012). Within the DNA there are length and sequence

*Corresponding author. E-mail: imad_dna@yahoo.com. Tel: 009647716150716.

Abbreviations: STRs, Short tandem repeats; **SSRs,** simple sequence repeats; **EDTA,** ethylene diamine tetraacetic acid; **TE,** tris-EDTA buffer; **PCR,** polymerase chain reaction; **Ho,** observed heterozygosity; **He,** expected heterozygosity; **PD,** power of discrimination; **PE,** probability or power of exclusion; **PIC,** polymorphism information content; **HWE,** Hardy Weinberg equilibrium; **RFLP,** restriction fragment length polymorphism; **CDP,** combined discrimination power; **PI,** paternity index; **RMP,** random match probability.

polymorphisms (Silvia et al., 2009). DNA can be used to study human evolution using human genome analysis regions that are not subjected to selection pressure (Mats et al., 2007; Imad et al., 2014). Besides, information from DNA typing provides vital information in medico-legal with polymorphisms allowing for more biological studies (Walkinshaw et al., 1996).

It has been found that microsatellites are evenly distributed in the genome on all chromosomes and all regions of the chromosome (Ensenberger et al., 2010; Imad et al., 2014). They can also be found inside gene coding regions, introns, and in the non-gene sequences. Most microsatellite loci are really small, ranging from a few to a few hundred repeats and this small size of microsatellite loci is important for PCR-facilitated genotyping. Basically, microsatellites containing a higher number of repeats are more polymorphic.

The Cooperative Human Linkage Center http://www.chlc.org evaluates the genetic markers and the loci are selected from there (Table 1) which provides particulars on the additional STR loci, chromosomal location and repeat sequence for each core STR locus (Ruitberg et al., 2001; Klintschar et al., 2004; Klintschar et al., 2005). Therefore, the repeat motif for each STR marker is listed based on this. A significant fact is that STR allele sizes are measured relative to an internal size standard during electrophoresis. This depends on the DNA strand that is labeled using a dye that may have a different apparent measured size.

The PowerPlex® 21 System is compatible with automated PCR instrument and with the ABI PRISM® 3100, 3100-Avant, 3130, 3130xl, 3500 and 3500xL Applied Biosystems Genetic Analyzers. In the United States, Europe and Asia the PowerPlex® 21 System is used and it increases the discriminatory power and data-sharing possibilities by incorporating informative loci. The PowerPlex® 21 System includes the 13 CODIS core STR loci, two loci commonly used in Europe (D1S1656 and D12S391).

In China, the D6S1043 locus is commonly used. (Amelogenin, Penta D, Penta E, D2S1338 and D19S433) are several additional markers used throughout the world. In forensic casework and DNA databases, addition of new autosomal is very important to increase the discrimination power for human forensic identification.

This study was aimed at investigating the genetic variation and Forensic efficiency parameters of 20 autosomal STR loci from random unrelated individuals in the middle and south of Iraq.

MATERIALS AND METHODS

Population

Four hundred (400) healthy, randomly chosen individuals deriving from the middle and south of Iraq provinces (Baghdad, Babil, Diwania and Basrah). The number and ethnicity of individuals were chosen in order to obtain a population sample to achieve the highest possible representation of the major ethno-religious and tribal groups of the country living in these central and southern areas.

DNA extraction

DNA was extracted from all dried blood samples on FTA cards following the manufacturer's procedure as described in Whatman FTA Protocol BD01 except that the Whatman FTA purification reagent was modified to half the volume (Dobbs et al., 2002). A 1.2 mm diameter disc was punched from each FTA card with a puncher. The discs were transferred to new Eppendorf tubes and washed three times in 100 µl Whatman FTA purification reagent. Each was washed and incubated for 5 min at room temperature with moderate manual mixing and the reagent was discarded between washing steps. The discs were then washed twice in 200 µl TE buffer (10 mM Tris-HCl, 0.1 mM EDTA, pH 8.0), the buffer was discarded and the discs were left to dry at room temperature for 1 h.

PCR

Amplifications of 20 STR loci D3S1358, D13S317, Penta E, D16S539, D18S51, D2S1338, CSF1PO, Penta D, THO1, vWA, D21S11, D7S820, TPOX, D8S1179, FGA, D1S1656, D5S818, D6S1043, D12S391, D19S433 and Amelogenin were performed by using the PowerPlex®21 kit. The PCR program was follow as: 96°C for 1 min, then: 94°C for 10 s, 59°C for 1 min, 72°C for 30 s, for 25 cycles, then: 60°C for 20 min and 4°C was soak. Keep the amplified samples at -20°C in a light-protected box.

Typing

Using the ABI Prism1 3130xl Genetic Analyzer 16-capillary array system (Applied Biosystems, Foster City, CA, USA), following manufacturer's protocols, with POP-7™ Polymer and Data Collection Software, GeneMapper® V3.2 software (Applied Biosystems, Foster City, CA, USA). By comparison, the size of a sample's alleles to size the alleles in allelic ladders for the same loci are being tested in the sample, the STR genotyping was conducted.

Quality control

Experimental procedures were performed according to the guidelines of the external blind proficiency test of the GEDNAP (http://www.gednap.org) (Rand et al., 2002; Rand et al., 2004).

Statistical data analysis

The PowerStatsV1.2 (Promega, Madison, USA) was used to calculate the observed heterozygosity (Ho), power of discrimination (PD), probability of exclusion (PE), and polymorphism information content (PIC). Arlequin software program (Schneider et al., 1998) was used to conduct the exact test of population differentiation. In addition, Arlequin software program was used for the expected heterozygosity (He), Hardy Weinberg Equilibrium (HWE) and linkage equilibrium tests as well as for the F-statistics. Where test results with P-values less than 0.05 were observed, and the Bonferroni correction had to be applied to the data. The Bonferroni procedure (Weir, 1996) adjusted the rejection level for the smallest P-value at an overall level of α = 5 % to 0.05/x, where x is equal to the number of tests conducted on the data. The Ho and He values were calculated by means of the same software program.

Table 1. Information on 21 autosomal STR loci present in The PowerPlex® 21 System kits Adapted from (Cotton et al., 2000; Wiegand et al., 1993) physical positions are from (Schneider et al., 1998).

STR Locus[a n]	Label	Physical position	Chromosomal Location1	Repeat Sequence 5´ to 3´
Amelogenin	Fluorescein	X and Y	Xp22.1-22.3 and Y	NA
D3S1358	Fluorescein	Chr 3 (45.582 Mb)	3p21.31	TCTA Complex
D1S1656	**Fluorescein**	Chr 1 (230.905 Mb)	1q42	TAGA Complex
D6S1043	**Fluorescein**	**Chr 6 (92.450 Mb)**	**6q15**	**AGAT**
D13S317	Fluorescein	**Chr 13 (82.692 Mb)**	**13q31.1**	**TATC**
Penta E	**Fluorescein**	Chr 15 (97.374 Mb)	15q26.2	AAAGA
D16S539	JOE	**Chr.16(86.386Mb)**	**16q24.1**	**GATA**
D18S51	JOE	Chr 18 (60.949 Mb)	18q21.33	AGAA
D2S1338	JOE	Chr 2 (218.879 Mb)	2q35	TGCC/TTCC
CSF1PO	JOE	Chr 5 (149.455 Mb)	5q33.1	AGAT
Penta D	**JOE**	Chr 21 (45.056 Mb)	21q22.3	AAAGA
TH01	TMR-ET	Chr 11 (2.192 Mb)	11p15.5	AATG (19)
vWA	TMR-ET	**Chr12(6.093 Mb)**	**12p13.31**	**TCTA Complex (19)**
D21S11	TMR-ET	Chr 21 (20.554 Mb)	21q21.1	TCTA Complex (19)
D7S820	TMR-ET	Chr 7 (83.789 Mb)	7q21.11	GATA
D5S818	TMR-ET	Chr 5 (123.111 Mb)	5q23.2	AGAT
TPOX	CXR-ET	Chr 2 (1.493 Mb)	2p25.3	AATG
D8S1179	CXR-ET	Chr 8 (125.907 Mb)	8q24.13 (125.976Mb)	TCTA Complex (19)
D12S391	**CXR-ET**	Chr 12 (12.450 Mb)	**12p12(12.341Mb)**	AGAT/AGAC **Complex**
D19S433	CXR-ET	**Chr19(30.416 Mb)**	(35.109Mb)	AAGG Complex
FGA	CXR-ET	Chr 4 (155.509 Mb)	4q28 (155.866Mb)	4q28 (155.866Mb)

a database of sequence-tagged sites (STSs) available on the NCBI website: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=unists.
n The 13 CODIS core loci are highlighted in bold font.

RESULTS AND DISCUSSION

Allele frequency of common autosomal genetic loci

After the samples have been collected, DNA extracted and PCR amplified were genotyped for the 20 STR loci of interest. The genotyping information was then converted into allele frequencies by counting the number of times each allele was observed. Allele frequencies for each of the 20 STR loci in the Iraq population sample are shown in (Tables 2 and 3).

Since there are some alleles which were not sampled sufficiently and an estimate of an allele frequency is uncertain if the allele is so rare that it can represented only once or a few times in a dataset, it is recommended that each allele was observed at least five times to be used in forensic calculations (Butler, 2007). The minimum allele frequency is 5/ (2n) where n is the number of individuals sampled and 2n is the number of chromosomes (as autosomes are in pairs due to inheritance of one chromosome from each parent).

In the loci D5S818 (allele 13) the highest allele frequencies are found and the lowest allele frequencies are at allele 13 as seen in D3S1358 locus. D21S11 and D18S51 loci illustrate the largest number of different alleles. The following locations of the most common alleles at the 20 loci were allele 13 for (D5S818, D8S1179 and D18S51), allele 10 for (D7S820 and Penta D), allele 16 for (D2S1338 and D2S1338), allele 11 for (TPOX, D16S539, Penta E, CSF1PO and D6S1043) loci, allele 12 for (D13S317 and D1S1656), allele 30.2 for D21S11 locus, allele 9 for THO1 locus, allele 13.2 for D21S11 locus, allele 6 for THO locus, alleles 17.3 and 16 for VWA locus, allele 30.2 for D19S433 locus, allele 17 for D3S1358 locus, allele 18 for D5S818 locus, allele 18 for D12S391 locus, allele 14.2 for D12S391 locus and allele 25 for FGA locus.

The best indicators of the genetic polymorphism within the sample are verified by the number of alleles and the expected heterozygosity is found in the Iraq population. Basically, the number of alleles is highly associated with the size of the sample. This is due to the presence of unique alleles in populations, which occur in low frequencies. The usefulness of the markers for genetic screening is verified by the number of alleles scored for each marker.

The number of alleles and the expected heterozygosities detected in Iraq population are good indicators of the genetic polymorphism within the breed. Generally, the number of alleles is highly dependent on the sample

Table 2. Allele frequencies and forensic efficiency parameters (D8S1179- D3S1358) genetic loci.

Allele	D8S1179	D7S820	D21S11	THO1	vWA	TPOX	D13S317	D5S818	D12S391	D3S1358
2.2	-	-	-	-	-	-	-	-	-	-
5	-	-	-	-	-	-	-	-	-	-
6	-	-	-	0.0455	-	-	-	-	-	-
7	-	0.0265	-	0.1713	-	-	-	-	-	-
8	0.127	0.1896	-	0.1326	-	0.0197	0.2318	-	-	-
9	0.0064	0.1003	-	0.2362	-	0.0383	0.0373	0.0228	-	-
9.3	-	-	-	0.1344	-	-	-	-	-	-
10	0.076	0.3304	-	0.18	-	0.1013	0.0546	0.0124	-	-
10.2	-	-	-	-	-	-	-	-	-	-
11	0.1344	0.2054	-	-	-	0.2637	0.2248	0.0382	-	-
12	0.1153	0.1115	-	-	-	0.2005	0.343	0.0133	-	-
12.2	-	-	-	-	-	-	-	-	-	-
13	0.2153	0.0363	-	-	-	0.1985	0.0534	0.4375	-	0.0037
13.2	-	-	-	-	-	-	-	-	-	-
14	0.142	-	-	-	0.0917	0.178	0.0551	0.3588	-	0.04
14.2	-	-	-	-	-	-	-	-	-	-
15	0.0549	-	-	-	0.0594	-	-	0.0208	-	0.2267
15.2	-	-	-	-	-	-	-	-	-	-
15.3	-	-	-	-	-	-	-	-	-	-
16	0.0181	-	-	-	0.2824	-	-	0.0407	-	0.2394
16.2	-	-	-	-	-	-	-	-	-	-
16.3	-	-	-	-	-	-	-	-	-	-
17	-	-	-	-	0.2806	-	-	-	0.0795	0.3471
18	-	-	-	-	0.0588	-	-	-	0.3597	0.1231
19	-	-	-	-	0.0212	-	-	-	0.1827	0.02
20	-	-	-	-	0.2059	-	-	-	0.1783	-
20.2	-	-	-	-	-	-	-	-	-	-
21	-	-	-	-	-	-	-	-	0.0194	-
21.2	-	-	-	-	-	-	-	-	-	-
22	-	-	-	-	-	-	-	-	0.0615	-
22.2	-	-	-	-	-	-	-	-	-	-
23	-	-	-	-	-	-	-	-	0.042	-
23.2	-	-	-	-	-	-	-	-	-	-
24	-	-	-	-	-	-	-	-	0.0771	-
24.2	-	-	-	-	-	-	-	-	-	-
28	-	-	0.0614	-	-	-	-	-	-	-
28.2	-	-	-	-	-	-	-	-	-	-
29	-	-	0.1817	-	-	-	-	-	-	-
29.2	-	-	-	-	-	-	-	-	-	-
30	-	-	0.1393	-	-	-	-	-	-	-
30.2	-	-	0.1615	-	-	-	-	-	-	-
31	-	-	0.0596	-	-	-	-	-	-	-
31.2	-	-	0.0113	-	-	-	-	-	-	-
32	-	-	0.0485	-	-	-	-	-	-	-
32.2	-	-	0.0589	-	-	-	-	-	-	-
33	-	-	0.0589	-	-	-	-	-	-	-
33.2	-	-	0.0983	-	-	-	-	-	-	-
34	-	-	0.0029	-	-	-	-	-	-	-
34.2	-	-	0.1066	-	-	-	-	-	-	-
35	-	-	0.0112	-	-	-	-	-	-	-

Table 3. Allele frequencies and forensic efficiency parameters (Penta D - CSF1PO).

Allele	Penta D	D19S433	D2S1338	D18S51	D16S539	Penta E	FGA	D6S1043	D1S1656	CSF1PO
2.2	0.0186	-	-	-	-		-	-	-	-
5	-	-	-	-	-	0.0183	-	-	-	-
6	-	-	-	-	-	-	-	-	-	-
7	0.0383	0.0023	-	-	-	0.0412	-	-	-	-
8	0.0217	-	-	-	0.1788	0.1988	-	-	-	0.0158
9	0.1614	0.0939	-	-	0.1012	0.0417	-	-	-	0.036
9.3	-	-	-	-	-	-	-	-	-	-
10	0.3985	0.0082	-	0.0086	0.2185	0.0987	-	0.0188	-	0.2636
10.2	-	-	-	0.0208		-	-		-	-
11	0.1015	0.1834	-	0.0454	0.251	0.2415	-	0.3412	0.1388	0.2858
12	0.0587	0.157	-	0.1782	0.0987	0.1015	-	0.2385	0.2387	0.2633
12.2	-	0.1148	-	-	-	-	-	-	-	-
13	0.1016	0.0726	-	0.2459	0.1003	0.0212	-	0.1215	0.0985	0.0622
13.2	-	0.2169	-	-	-	-	-	-	-	
14	0.0413	0.0745	-	0.17	0.0515	0.0088	-	0.0783	0.1615	0.011
14.2	-	0.0424	-		-	-	-	-	-	-
15	0.0186	0.0109	0.0583	0.1044	-	0.0298	-	0.0217	0.0983	0.0633
15.2	-	0.0598	-	-	-	-	-	-	-	-
15.3	-	-	-	-	-	-	-	-	0.0417	-
16	0.0398	-	0.2015	0.1471	-	-	-	0.0587	-	-
16.2	-	-	-	-	-	-	-	-	0.2012	-
16.3	-	-	-	-	-	-	-	-	-	-
17	-	-	0.1383	0.0473	-	-	-	0.0813	0.0213	-
18	-	-	0.0785	0.0111	-	-	-	0.0898	-	-
19	-	-	0.1513	0.0064	-	-	-	0.0622	-	-
20	-	-	0.0587	0.0112	-	-	0.1014	0.112	-	-
20.2	-	-	-	-	-	-	-	-	-	-
21	-	-	0.0388	0.0112	-	-	0.0379	-	-	-
21.2	-	-	-	-	-	-	-	-	-	-
22	-	-	0.0612	-	-	-	0.1817	-	-	-
22.2	-	-	-	-	-	-	0.1389	-	-	-
23	-	-	0.0083	-	-	-	0.0787	-	-	-
23.2	-	-	-	-	-	-	-	-	-	-
24	-	-	0.0717	-	-	-	0.1373	-	-	-
24.2	-	-	-	-	-	-	-	-	-	-
25	-	-	0.0617	-	-	-	0.2429	-	-	-
26	-	-	0.0717	-	-	-	0.0812	-	-	-
27	-	-	-	-	-	-	-	-	-	-
28	-	-	-	-	-	-	-	-	-	-
28.2	-	-	-	-	-	-	-	-	-	-
29	-	-	-	-	-	-	-	-	-	-
29.2	-	-	-	-	-	-	-	-	-	-
30	-	-	-	-	-	-	-	-	-	-
30.2	-	-	-	-	-	-	-	-	-	-
31	-	-	-	-	-	-	-	-	-	-
31.2	-	-	-	-	-	-	-	-	-	-
32	-	-	-	-	-	-	-	-	-	-
32.2	-	-	-	-	-	-	-	-	-	-
33	-	-	-	-	-	-	-	-	-	-

the sample size because of the presence of unique alleles in populations, which occur in low frequencies and also because the number of observed alleles tends to increase with increases in population size. The number of alleles scored for each marker is an invaluable indicator of the future usefulness of the marker for genetic screening.

Finding the same number of alleles for certain different loci in various populations (e.g., Iran, Syrian, Emirates, Qatar and Egyptian populations) may indicate common ancestries (Reyhaneh et al., 2009; Alshamali et al., 2003; Ana et al., 2006; Clotilde et al., 2007). The frequency and the number of alleles, however, may be an indication for the degree of inbreeding within each population and thus reflects the homogeneity of the population. The analysis of STR polymorphisms by PCR-based method offers certain advantages over RFLP typing.

In recent years, short tandem repeat (STR) systems have gained importance in forensic analysis of biological specimens as well as in paternity testing, as an alternative to the use of restriction fragment length polymorphism (RFLP) analysis (Edwards et al., 1991; Hammond et al., 1994; Nakamura et al., 1987). The analysis of STR polymorphisms by PCR-based method offers certain advantages over RFLP typing: (1) STR loci can be typed with a high degree of specificity and sensitivity in a short time period, (2) these loci can be successfully amplified from a limited amount of DNA even if it is degraded, and (3) typing of multiple loci can be accomplished in a single multiplex reaction (Hochmeister et al., 1991; Lins et al., 1996; Mohammad and Imad, 2013a, b).

The amelogenin locus

The amelogenin locus occurs on both the X and Y chromosome and enables sex typing (Sullivan et al., 1993) was also located within the reference human genome sequence. AMELX is located on the X chromosome at 10.676 Mb. AMELY is located on the Y chromosome at 6.441 Mb. Amplification of Amelogenin generates different length products from the X and Y-chromosomes.

Observed heterozygosity and expected heterozygosity

A higher heterozygosity means that more allele diversity exists and therefore there is less chance of a random sample matching. Observed heterozygosity and expected heterozygosity all over the 20 loci are presented in (Table 4), and the observed heterozygosity oscillated between studied populations as illustrated in (Table 5). The observed heterozygosity in a population relies on the number and the frequency of alleles of each locus. Moreover, the distribution of genotypes in a population sample may deviate from HWE expectation in a number of ways. These include the presence of an excess of homozygotes (and a corresponding lack of heterozygotes) or an excess (or deficiency) of one or more classes of heterozygotes or a combination of those states. There are populations with low heterozygosity, lower than 65% in most tested loci.

Paternity index

The potential of a randomly selected man to pass the obligate gene is determined by using a database, which lists the frequency distribution of individual alleles within a given genetic system. Combined paternity index is an odds ratio that indicates how many times more likely it is that the alleged father is the biological father than a randomly selected unrelated man of similar ethnic background. The paternity index was high for all STR analyzed it ranged from 2.651 (TPOX) to 2.864 (D21S11).

Random match probability

The match probability is the probability for a random match between two unrelated individuals drawn from the same population. It is the sum of the frequency squared of each genotype; it ranged from 0.011 to 0.168.

Power of discrimination

Values for all tested loci were 75% for TPOX. Ranged from 80 to 89% for the D3S1358, D13S317, D5S818, D12S391, vWA, Penta D, D16S539, D1S1656 and CSF1PO loci, and ranged from 91 to 96%, for the rest of the loci. This infers that a DNA-based database for Iraq population can be safely used by using these loci. The highest PD observed in some populations is presented in (Table 6). The Penta E and Penta D loci included in the PowerPlex®21 PCR amplification kits were not typed in the Turkey, Emirates, Iran or Qatari populations because they used different kits in their genotyping studies. The Combined Discrimination Power (CDP) for the Iraq population of middle and south of Iraq for the corresponding 20 STR loci used has been calculated as 0.999999972. These results mean that those loci can be safely used to establish a DNA-based database for Iraq population.

Chance of exclusion

The Power of Exclusion (PE) can be calculated to express how rare it would be to find a random man who could not be excluded as the biological father of the child (Fisher, 1951; Chakraborty and Stivers 1996; Butler,

Table 4. Forensic efficiency parameters of the autosomal 21.

Locus	Observed heterozygosity (Ho)	Expected heterozygosity (He)	paternity index (PI)	Random match probability (RMP)	Power of discrimination (PD)	Chance of exclusion (CE)	polymorphic information content (PIC)	P-value*
S1179	0.784	0.829	2.807	0.012	0.912	0.528	0.802	0.423
D7S820	0.821	0.866	2.844	0.025	0.949	0.569	0.839	0.442
D21S11	0.841	0.886	2.864	0.045	0.969	0.585	0.859	0.48
THO1	0.785	0.83	2.808	0.011	0.913	0.529	0.803	0.424
vWA	0.755	0.8	2.778	0.041	0.883	0.499	0.773	0.394
TPOX	0.628	0.673	2.651	0.168	0.756	0.372	0.646	0.267
D13S317	0.727	0.772	2.75	0.069	0.855	0.471	0.745	0.366
D5S818	0.735	0.78	2.758	0.061	0.863	0.479	0.753	0.374
D12S391	0.72	0.765	2.743	0.076	0.848	0.464	0.738	0.359
D3S1358	0.766	0.811	2.789	0.03	0.894	0.51	0.784	0.405
Penta D	0.751	0.796	2.774	0.045	0.879	0.495	0.769	0.39
D19S433	0.803	0.848	2.826	0.007	0.931	0.549	0.821	0.442
D2S1338	0.832	0.877	2.855	0.036	0.96	0.576	0.85	0.471
D18S51	0.831	0.876	2.854	0.035	0.959	0.575	0.849	0.47
D16S539	0.766	0.811	2.789	0.03	0.894	0.51	0.784	0.405
PentaE	0.807	0.852	2.83	0.011	0.935	0.551	0.825	0.446
FGA	0.823	0.868	2.846	0.027	0.951	0.567	0.841	0.462
D6S1043	0.81	0.855	2.833	0.014	0.938	0.554	0.828	0.449
D1S1656	0.746	0.791	2.769	0.05	0.874	0.49	0.764	0.385
CSF1PO	0.715	0.76	2.738	0.081	0.843	0.459	0.733	0.354

Table 5. The observed heterozygosity in different populations.

Locus	Iraq	Turkey	Iran	Syria	Emirates	Qatar	Egypt	Gaza
D21S11	0.841	-	-	0.664	-	-	0.682	0.682
TPOX	0.628	0.567	-	-	-	-	-	-
Penta E	-	-	-	0.897	-	-	-	-
FGA	-	-	-	-	-	-	0.890	
D18S51	-	0.790	-	-	0.130	-	-	0.912
D2S1338	-	-	-	-	-	0.839	-	-
vWA	-	-	-	-	-	0.542	-	-
LPL	-	-	0.780	-	-	-	-	-
F13A01	-	-	0.720	-	-	-	-	-
D5S818	-	-	-	-	0.295	-	-	-

2005). In the present study and from the genotyping data, PE for every locus was calculated and presented in (Table 4). As expected, the power of exclusion was high for all the microsatellites analyzed it ranged from 0.372 (TPOX) to 0.820 (D18S51).

Polymorphic information content

The TPOX locus is the least polymorphic marker while D21S11 is the most polymorphic marker. The usefulness of the findings of genetic polymorphism studies and linkage mapping programs in humans is confirmed by the high PIC values of the selected markers. Similar degree of polymorphism was found in Syrian population, where the most polymorphic marker was found to be Penta E, and that the least polymorphic one was TPOX (Table 7). On the contrary, the Egyptian population showed that the FGA locus is the most polymorphic marker and that the TPOX locus, as observed in our population, is the least

Table 6. The Power of discrimination in different populations.

Locus	Iraq	Turkey	Iran	Syria	Emirates	Qatar	Egypt	Gaza
D21S11	0.969	-	-	-	-	-	-	-
D2S1338	0.960	-	-	-	-	0.973	-	-
D18S51	0.959	0.953	-	0.963	0.971	-	-	0.962
TPOX	0.759	0.772	-	-	-	0.855	-	-
Penta E	-	-	-	0.974	-	-	-	0.976
Penta D	-	-	-	0.951	-	-	-	0.961
FGA	-	-	-	0.957	-	-	0.973	0.967
vWA	-	-	-	-	-	-	0.937	-
LPL	-	-	0.924	-	-	-	-	-
F13A01	-	-	0.922	-	-	-	-	-
D5S818	-	-	-	-	0.889	-	-	-

Table 7. Polymorphic Information Content in different populations.

Locus	Iraq	Turkey	Iran	Syria	Qatar	Egypt	Gaza
D21S11	0.859	-	-	-	-	-	-
D18S51	-	0.840	-	-	-	-	-
D2S1338	-	-	-	-	0.865	-	-
TPOX	0.646	0.540	-	0.640	0.638	0.610	0.650
Penta E	-	-	-	0.890	-	-	0.900
FGA	-	-	-	-	-	0.870	-
LPL	-	-	0.770	-	-	-	-
F13A01	-	-	0.760	-	-	-	-
D5S818	-	-	-	-	-	-	-

polymorphic marker. Even in other populations like the Iran, Qatar, and Turkey, it was found that the previously mentioned loci have a similar degree of polymorphism (Ana et al., 2006; Reyhaneh and Sadeq, 2009). The polymorphic nature of microsatellites makes them the markers of choice in characterization and genetic diversity studies. The high PIC values of the selected markers confirm their usefulness for genetic polymorphism studies and linkage mapping programs in human as well. The mean heterozygosity observed, expected and mean PIC values across the 20 loci were 0.77, 0.81 and 0.78, respectively, indicating high gene diversity.

P-value: Probability value of Chi-square test for Hardy Weinberg equilibrium

Checking for HWE, it is performed by taking the observed allele frequencies and calculating the expected genotype frequencies based on the allele frequencies. If the observed genotype frequencies are close to the expected genotype frequencies calculated from the observed allele frequencies, then the population is in Hardy-Weinberg equilibrium and allele combinations are likely to be independent of one another. The results obtained from the tests for HWE is shown in (Table 4). The null hypothesis states that all the STR loci tested are in HWE and any deviation from HWE expectations is due to sampling error. The alpha value was set at 0.05. The null hypothesis was rejected if the computed p values were below the alpha value of 0.05. Therefore having a p value above 0.05 to show that the STR alleles do not differ significantly from HWE does not imply that the samples are in HWE. Butler (2005) states that "if a p-value < 0.05 is observed with a set of alleles measured at a particular STR locus, it does not mean that a laboratory should avoid using this data because it 'failed' a test for Hardy-Weinberg equilibrium".

Conclusion

The results of the current study indicate these autosomal STRs Loci useful for DNA typing markers in Iraq and can be used for establishment of a DNA database that will be beneficial for the population in terms of resolving social and moral disputes and will contribute to improvements in

the justice system.

ACKNOWLEDGEMENTS

We sincerely wish to thank Dr. Issam for providing us the opportunity to work on this project. We are grateful for the assistance rendered through the various analysis stages, and for providing helpful criticism and feedback throughout the writing process. We also like to thank Zainab Al-Habubi from the Department of Biology for her guidance and help in the laboratory work.

REFERENCES

Alshamali FH, Alkhayat AI, Budowle B, Watson ND (2003). Allele frequency distributions and other population genetic parameters for 13 STR loci in a UAE local population from Dubai. International Congress Series. 1239:249-258.

Ana MP, Miguel AA, José AP, Rene JH (2006). Qatari DNA Variation at a Crossroad of Human Migrations. Hum. Hered. 61:67-79.

Butler JM (2005). Forensic DNA Typing. biology, technology and genetic of STR markers. Second Edition. Elsevier Academic Press.

Butler JM (2007). Short tandem repeat typing technologies used in human identity testing. Biotechniques. 43(4):2-5.

Butler JM, Hill CR (2012). Biology and genetics of new autosomal STR loci useful for forensic DNA analysis. Forensic Sci. Rev. 24(1):15-26.

Chakraborty R, Stivers DN (1996). Paternity exclusion by DNA markers: effects of paternal mutations. J. Forensic Sci. 41(4):671-677.

Clotilde C, Evelyne G, Farha E, Georges L, Jean-Michel D (2007). Allele frequencies of 15 short tandem repeats (STRs) in three Egyptian populations of different ethnic groups. Forensic Sci. Int. 169:260-265.

Cotton EA, Allsop RF, Guest JL, Frazier RR, Koumi P, Callow IP (2000). Validation of the AMPFISTR SGM plus system for use in forensic casework. Forensic Sci Int. 112:151-61.

Dobbs LJ, Madigan MN, Carter AB, Earls L (2002). Use of FTA gene guard filter paper for the storage and transportation of tumor cells for molecular testing. Arch. Pathol. Lab. Med. 126 (1):56-63.

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

Ellegren H (2004). Microsatellites: simple sequences with complex evolution. Nature Reviews Genet. 5:435-445.

Ensenberger MG, Thompson J, Hill B, Homick K, Kearney V, Mayntz-Press KA (2010). Developmental validation of the PowerPlex 16 HS System: an improved 16-locus fluorescent STR multiplex. Forensic. Sci. Int. Genet. 4(4):257-264.

Fisher RA (1951). Standard calculations for evaluating a blood group system. Heredity. 5:95-102.

Hammond HA, Jin L, Zhong Y, Caskey CT, Chakraborty R (1994). Evaluation of 13 short tandem repeat loci for use in personal identification applications. Am. J. Hum. Genet. 55:175-89.

Hochmeister MN, Budowle B, Jung J, Borer UV, Comey CT, Dirnhofer R (1991) PCR-based typing of DNA extracted from cigarette butts. Int. J. Leg. Med. 10:506-513.

Imad H, Abeer F, Cheah Y, Mohammed J, Aamera O (2013a). Discovery of three newly described single nucleotide polymorphisms in mitochondrial DNA Hypervariable Region I (HVI) and estimation of variants and haplotypes encompassing nucleotide positions 16024-16365. J. Forensic Res. 5(1):1-6.

Imad H, Cheah Q, Mohammad J, Aamera O (2013b). Genetic variation of 17 Y-chromosomal short tandem repeats (STRs) loci from unrelated individuals in Iraq. Int. J. Biotechnol. Mol. Biol. Res. 4(8):119-129.

Imad HH, Ameer IA, Mohammed AJ, Cheah YK, Aamera JO (2014). Haplotypes and variable position detection in the mitochondrial DNA coding region encompassing nucleotide positions 10,716-11,184. Mitochondrial DNA. 1-6.

Klintschar M, Immel U, Stiller D, Kleiber M (2004). HumTH01 and blood pressure. An obstacle for forensic application?. International Congress Series. 1261:589-591.

Klintschar M, Immel UD, Stiller D, Kleiber M (2005). TH01, a tetrameric short tandem repeat locus in the tyrosine hydroxylase gene: association with myocardial hypertrophy and death from myocardial infarction?. Dis. Markers. 21:9-13.

Klintschar M, Immela UD, Kleiber M, Wiegand P (2006). Physical location and linked genes of common forensic STR markers. International Congress Series. 1288:801-803.

Lins AM, Sprecher CJ, Puers C, Schumm JW (1996). Multiplex sets for amplification of polymorphic short tandem repeat loci- silver stain and fluorescent detection. Biotechniques. 20:882-889.

Mohammad A, Imad H (2013). Autosomal STR: From locus information to next generation sequencing technology. Res. J. Biotechnol. 8(10):92-105.

Mats O, Mo H, Erik W, Tonia S, Natasha L, Tobias U (2007). Mating system variation and morph fluctuations in a polymorphic lizard. Molecular Ecology.16:5307-5315.

Nakamura Y, Leppert M, O'Connell P, Wolff R, Holm T (1987). Variable number of repeat (VNTR) markers for human gene mapping. Science. 235:1616-1622.

Rand M, Schurenkamp B, Brinkmann (2002). The German DNA profiling group (GEDNAP) blind trial concept. Int. J. Legal Med. 116 (4):199-206.

Rand M, Schurenkamp C, Hohoff B (2004) . The GEDNAP blind trial concept part II. Trends and developments. Int. J. Legal Med. 118 (2): 83-89.

Reyhaneh L, Sadeq V (2009). Genetic variation of informative short tandem repeat(STR) loci in an Iranian population. Iran. J. Biotechnol. 7(3):137-141.

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.

Schneider HR, Rand S, Schmitter H, Weichhold G (1998). ACTBP2 - nomenclature recommendations of GEDNAP. Int. J. Legal Med. 111: 97-100.

Silvia B, Campo D, Andrea Z, Dario B, Tatiana G (2009). Integration of genomic and gene expression data of childhood ALL without known aberrations identifies subgroups with specific genetic hallmarks. Genes Chromosomes & Cancer . 48:22-38.

Sullivan KM, Mannucci A, Kimpton CP, Gill P (1993). A rapid and quantitative DNA sex test: fluorescence-based PCR analysis of X-Y homologous gene amelogenin. 15(4):636-638.

Walkinshaw M, Strickland L, Hamilton H, Denning K, Gayley T (1996). DNA Profiling in two Alaskan Native Populations Using HLA-DQA1, PM, and D1S80 Loci, J. Forensic Sci. 41:478-484.

Weir B (1996). The second national research council report on forensic DNA evidence, Am. J. Hum. Genet. 59:497-500.

Wiegand P, Budowle B, Rand S, Brinkmann B (1993). Forensic validation of the STR systems SE33 and TC11. Int. J. Legal Med.105: 315-20.