Allele frequency data of 21 autosomal Short Tandem Repeat loci in Mesan and Basra provinces in south of Iraq

Article · February 2015

DOI: 10.1016/j.ejfs.2014.10.003

4 authors, including:

Imad Hadi Hameed
University of Babylon

Abeer Fauzi
University of Babylon

Ghaidaa Mohammad
University of Al-Qadisiyah/Science college

ARTICLE IN PRESS
Egyptian Journal of Forensic Sciences (2015) xxx, xxx–xxx

HOSTED BY
ELSEVIER

Contents lists available at ScienceDirect

Egyptian Journal of Forensic Sciences

journal homepage: www.ejfs.org

Allele frequency data of 21 autosomal Short Tandem Repeat loci in Mesan and Basra provinces in south of Iraq

Imad Hadi Hameed [a,*], Aamera Jaber Ommer [b], Abeer Fauzi Murad [c], Ghaidaa Jehadi Mohammed [d]

[a] *Department of Molecular Biology, Babylon University, Hilla City, Iraq*
[b] *Institute of Medico-legal in Baghdad, Ministry of Health of Iraq, Iraq*
[c] *Department of Biology, Babylon University, Faculty of Science for Women, Iraq*
[d] *Department of Molecular Biology, Qadsia University, Qadsia City, Iraq*

Received 15 March 2013; revised 23 August 2014; accepted 21 October 2014

KEYWORDS

Autosomal STR;
Forensic efficiency parameters;
Microsatellite;
Power plex21® kit;
South of Iraq

Abstract We are focusing on a sample of 100 unrelated persons from provinces of southern Iraq. This is an analysis of the allele frequency and genotyping of those STR loci in an Iraqi population and this is the first study of its kind. As such the data available could be utilized in the Iraqi database for the STR polymorphic markers. Chelex® kit was utilized to extract DNA then Power Plex21® kit (D3S1358, D13S317, PentaE, D16S539, D18S51, D2S1338, CSF1PO, Penta D, THO1, vWA, D21S11, D7S820, TPOX, D8S1179, FGA, D2S1338, D5S818, D6S1043, D12S391, D19S433) was used to amplify the isolated DNA. The mean PIC values and heterozygosity observed across 21 loci were 0.713 and 0.696 respectively. This shows high gene diversity. Those loci can be safely used to establish a DNA-based database for Iraq population because the power of discrimination values for all tested loci was from 71% to 97%.

1. Introduction

Microsatellites are a group of molecular markers chosen for a number of purposes which include forensics individual identification and relatedness testing.[1–3] Low quantities of template DNA require (10–100 ng),[4] when using microsatellites. There is a high genomic abundance of random distribution throughout the genome. There is also an abundance of polymorphism. 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.[5]

Within the DNA there are length and sequence polymorphisms.[6] Among the polymorphisms attracting most attention are those related to diseases. These may indicate that the

* Corresponding author. Tel.: +964 7716150716.
E-mail address: imad_dna@yahoo.com (I.H. Hameed).
Peer review under responsibility of Forensic Medicine Authority.

http://dx.doi.org/10.1016/j.ejfs.2014.10.003
2090-536X © 2014 Hosting by Elsevier B.V. on behalf of The International Association of Law and Forensic Sciences (IALFS).

diseases may be forthcoming. DNA can be used to study human evolution using human genome analysis regions that are not subject to selection pressure.[7,8] Besides, information from DNA typing provides vital information in medico-legal with polymorphisms allowing for more biological studies.[9] Since these STRs can be highly variable among individuals, the number of repeats in STR markers are important for human identification purposes.[10,11] The repeat pattern that STRs are often divided into three categories: The compound repeats are made up of two or more adjacent simple repeats and simple repeats contain units of similar length. Finally, sequence and complex repeats may consist of several repeat blocks of variable unit length as well as variable intervening sequences.[12]

Using 21 autosomal STR loci we plan to determine the genetic structure of Iraq population and this is the first aim of this study. The second objective of the study was to evaluate the importance of these loci for forensic genetic purposes.

2. Materials and methods

2.1. Sample preparation

Blood samples were randomly collected from one hundred unrelated Iraqi male and female persons from the Arab ethnic group of Iraqi population for DNA autosomal STRs. The Chelex® 100 (Bio-Rad) was used to extract the total DNA from blood cells.

2.2. PCR multiplex amplification

PCR is the process used to amplify a specific region of DNA. It is possible to create multiple copies from a small amount of template DNA using this process. Using the PowerPlex® Autosomal 21 kit from Promega all samples were amplified. MicroAmp® plate is put in the GeneAmp® PCR System 9700 thermal cycler with the estimated total cycle time of 1.5 h. The PCR program is as follows: 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 4 °C soak. Amplified samples are kept at −20 °C in a light-protected box.

2.3. PCR amplicon analysis (Capillary electrophoresis)

In the PowerPlex® 21 System PCR Amplification User's Manual (Promega, Madison, USA). PowerPlex® 21 System Locus-Specific Information Table 1 and the protocols for sample preparation and experimental conditions for the ABI 3130*xL* Genetic Analyzer were provided. To obtain the sample genotype the allelic ladder is the standard to which STR alleles are compared. By comparison of the size of a sample's alleles to the size of alleles in allelic ladders for the same loci being tested in the sample, the STR genotyping was conducted.[10]

2.4. 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[13] 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). Where test results with *P*-values less than 0.05 were observed, the Bonferroni correction had to be applied to the data. The Bonferroni procedure[14] adjusted the rejection level for the smallest *P*-value at an overall level of $\alpha = 5\%$ 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.

Following are the formulas used to compute various parameters for population data analyses.

2.4.1. Allele frequency

After the samples have been collected, extracted DNA and amplified PCR 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 in all giving samples. Since there are some alleles which were not sampled sufficiently and that an estimate of an allele frequency is uncertain if the allele is so rare that it is represented only once or a few times in a dataset, it is recommended that each allele be observed at least five times to be used in forensic calculations.[15] 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).

2.4.2. Observed heterozygosity and expected heterozygosity

Observed heterozygosity is calculated by dividing the number of heterozygote individuals at a locus by the total number of all individuals at that locus and describes the heterozygosity actually observed in the members of the sampling group. Expected heterozygosity is calculated as 1 minus the homozygosity (the sum of squares of all allele frequencies at a locus) and represents the number of heterozygotes that would be expected under HWE based on the observed allele frequencies in the sampling group.

Edwards et al.[16] described the following formula for calculating an unbiased estimate of the expected heterozygosity:

$$H = \frac{n}{(n-1)}\left[1 - \sum_{j=1}^{k}(n_j/n)^2\right] = \frac{n}{(n-1)}\left[1 - \sum_{j=1}^{k}(p_j)^2\right]$$

where $n1, n2, \ldots, nk$ are the allele counts of k alleles at a locus in a sample of n genes drawn from the population and p_j is the allele frequency.

2.4.3. Random match probability

A random match probability, is the probability that a person sampled randomly from the population, would have a particular STR profile. The probability of a match at locus l, PM_l, was first described from genotype data[17] which gave

$$PM_l = \sum_i \tilde{G}_{il}^2$$

where $\tilde{G}_i$ is the sample frequency of the ith genotype at locus l.

$$PM_l = \frac{\sum_{i=1}^{n}\tilde{G}_{il}^2 - 1/N_l}{1 - 1/N_l} \approx \sum_{i=1}^{n}\tilde{G}_{il}^2$$

where the first part of this equation is for a sample of size N_l at locus l.[18]

Table 1 The PowerPlex® 21 System Locus Specific Information.

STR locus	Label	Chromosomal location 1	Repeat sequence 5′ 3′
Amelogenin	Fluorescein	Xp22.1–22.3 and Y	NA
D3S1358	Fluorescein	3p21.31 (45.557 Mb)	TCTA Complex
D1S1656	Fluorescein	1q42 (228.972 Mb)	TAGA Complex
D6S1043	Fluorescein	6q15 (92.449 Mb)	AGAT
D13S317	Fluorescein	13q31.1 (81.62 Mb)	TATC
Penta E	Fluorescein	15q26.2 (95.175 Mb)	AAAGA
D16S539	JOE	16q24.1 (84.944 Mb)	GATA
D18S51	JOE	18q21.33 (59.1 Mb)	AGAA
D2S1338	JOE	2q35 (218.705 Mb)	TGCC/TTCC
CSF1PO	JOE	5q33.1 (149.436 Mb)	AGAT
Penta D	JOE	21q22.3 (43.88 Mb)	AAAGA
TH01	TMR-ET	11p15.5 (2.149 Mb)	AATG (19)
vWA	TMR-ET	12p13.31 (5.963 Mb)	TCTA Complex (19)
D21S11	TMR-ET	(19.476 Mb)	TCTA Complex (19)
D7S820	TMR-ET	7q21.11 (83.433 Mb)	GATA
D5S818	TMR-ET	5q23.2 (123.139 Mb)	AGAT
TPOX	CXR-ET	2p25.3 (1.472 Mb)	AATG
D8S1179	CXR-ET	8q24.13 (125.976 Mb)	TCTA Complex (19)
D12S391	CXR-ET	12p12 (12.341 Mb)	AGAT/AGAC Complex
D19S433	CXR-ET	(35.109 Mb)	AAGG Complex
FGA	CXR-ET	4q28 (155.866 Mb)	4q28 (155.866 Mb)

2.4.4. Power of discrimination

Power of discrimination is defined as the probability that two individuals selected at random from the population will not have an identical genotype at the locus. Brenner and Morris[19] described the following formula for calculating the power of discrimination:

$$1 - PM$$

2.4.5. Polymorphism information content

This parameter indicates the polymorphic level of a locus. The PIC was also calculated using marker allelic frequencies using following equation:

$$\mathrm{PIC} = 1 - \sum_{i=1}^{n} p_i^2 - \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} 2p_i^2 p_j^2$$

where n is the number of alleles and p_i is the allele probability of the ith allele.[20]

2.4.6. Power of exclusion

Power of exclusion is the probability of excluding a random person as the contributor of an allele at the locus.

$$\text{Power of Exclusion(PE)} = H^2(1 - (1 - H)H^2)$$

where H = heterozygosity.

2.4.7. Paternity index

It is a calculated value generated for a single genetic marker or locus (chromosomal location or site of DNA sequence of interest) and is associated with the statistical strength or weight of that locus in favor of or against parentage.

$$\text{Paternity Index(PI)} = \frac{H+h}{2h} = \frac{(1-h)+h}{2h} = \frac{1}{2h} = \frac{1}{2\sum_{i=1}^{n} P_i^2}$$

where P_i is the frequency of ith allele in a population of n samples; h = homozygosity; H = heterozygosity.

3. Results and discussion

By noting the number of times each allele was seen the genotyping information was then converted into allele frequencies. The sum of homozygote (h) and the number of heterozygote (H) equal 100% of the samples tested. So, since $h + H = 1$, $H = 1 - h$ and $h = 1 - H$. There is less chance of a random sample matching because a higher heterozygosity means that more allele diversity would occur.

Butler,[5] states that areas containing repeated nucleotide sequences such as STR sequences are of interest to the forensic scientist. This is because they provide a lot of variation, and this makes them relevant for human identification purposes. Based on the allelic frequencies, several statistical parameters of genetic and forensic efficiency have been estimated. These include 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 total number of alleles observed across the population was found to be 158 alleles for this particular study (Table 2). It can be concluded that there is a high level of polymorphism of the selected microsatellites. In the loci D19S433 (allele 11) the highest allele frequencies are found, and the lowest allele frequencies are at allele 15 as seen in Penta E locus. D6S1043 and D18S51 loci illustrate the largest number of different alleles.

The following locations of the most common alleles at the 21 loci were allele 15 for D3S1358 locus, allele 16 for D1S1656 locus, allele 12 for D6S1043 locus and TPOX, allele 11 for D13S317, Penta E, D16S539 and D19S433 loci, allele 16 for D18S51 locus, allele 17 for D2S1338 locus, allele 10 for Penta D locus and D7S520, allele 6 for THO locus, alleles

Table 2 A summary of distribution of allele frequency results from the 21 autosomal STR loci in Iraq population.

Allele	D3S1358	D1S1656	D6S1043	D13S317	PentaE	D16S539	D18S51	D2S1338	D19S433	Penta D
2.2										0.02
5					0.02					
6										
7					0.04				0.02	0.04
8				0.26	0.20					0.02
9				0.08	0.04	0.18			0.04	0.16
9.3										
10			0.02	0.10	0.10	0.20	0.04		0.40	0.40
10.2							0.02			
11		0.14	0.34	0.30	0.24	0.32	0.04		0.51	0.10
12		0.20	0.24	0.20	0.20	0.25	0.12		0.02	0.06
13	0.02	0.10	0.12		0.10	0.02	0.07			0.10
13.2										
14	0.08	0.16	0.08	0.04	0.02	0.02	0.20			0.04
14.2										
15	0.48	0.10	0.02		0.01		0.10			0.02
15.2										
15.3		0.04								
16	0.34	0.24	0.04		0.02		0.20	0.06		0.02
16.2										
16.3		0.05								
17	0.08		0.06				0.08	0.20		
17.3										
18			0.08				0.06	0.14		
18.3										
19			0.09				0.04	0.08		
20			0.06					0.15		
21								0.06		
22								0.04		
22.2										
23								0.06		
24								0.10		
25								0.06		
26										
27										
28										
29										
30										
30.2										
31										
31.2										
32										
32.2										

	THO1	vWA	D21S11	D7S820	TPOX	D8S1179	D5S818	D12S391	CSF1PO	FGA
2.2										
5										
6	0.42									
7	0.14			0.04						
8	0.04			0.28	0.02	0.48	0.02			
9	0.32			0.12	0.04	0.12	0.02			
9.3	0.02									
10	0.02			0.32	0.10	0.04				
10.2										
11				0.12	0.26	0.30	0.04			
12				0.06	0.28	0.06	0.04		0.10	
13				0.04	0.20		0.44		0.26	
13.2					0.04				0.02	
14		0.09			0.04		0.36		0.30	
14.2									0.04	
15		0.06					0.02		0.06	
15.2									0.06	
15.3										
16		0.38					0.04		0.08	
16.2									0.04	

(*continued on next page*)

Table 2 (*continued*)

	THO1	vWA	D21S11	D7S820	TPOX	D8S1179	D5S818	D12S391	CSF1PO	FGA
16.3										
17		0.38						0.08		
17.3										
18		0.06						0.36		
18.3								0.02		
19		0.02						0.18		0.10
20								0.18		0.04
21								0.02		0.18
22								0.06		0.14
22.2										0.08
23								0.04		0.14
24								0.02		0.20
25										0.08
26										
27										
28			0.06							
29			0.18							
30			0.24							
30.2			0.26							
31			0.06							
31.2			0.10							
32			0.10							
32.2			0.08							

Table 3 Forensic efficiency parameters of the autosomal 21 genetic loci.

Locus	Homozygosity	Heterozygosity	Effective number of alleles (n)	Polymorphism information content (PIC)	Power of exclusion (PE)	Probability of identity (Matching probability)	Power of discrimination (PD)
D3S1358	0.359	0.641	2.783	0.528	0.348	0.275	0.725
D1S1656	0.166	0.833	5.992	0.807	0.613	0.057	0.943
D6S1043	0.203	0.797	4.921	0.757	0.553	0.038	0.910
D13S317	0.206	0.794	4.862	0.754	0.548	0.087	0.913
PentaE	0.153	0.847	6.553	0.824	0.638	0.047	0.952
D16S539	0.238	0.762	4.199	0.708	0.50	0.117	0.883
D18S51	0.123	0.877	8.137	0.862	0.70	0.030	0.969
D2S1338	0.115	0.885	8.734	0.872	0.712	0.026	0.974
CSF1PO	0.422	0.577	2.360	0.44	0.286	0.388	0.711
FGA	0.141	0.850	7.092	0.839	0.644	0.401	0.959
THO1	0.299	0.701	3.344	0.619	0.419	0.187	0.812
vWA	0.305	0.695	3.284	0.611	0.419	0.190	0.806
D21S11	0.191	0.808	5.236	0.773	0.570	0.074	0.926
D7S820	0.216	0.784	4.629	0.739	0.533	0.025	0.975
D5S818	0.329	0.670	3.039	0.575	0.382	0.228	0.771
TPOX	0.201	0.798	4.975	0.760	0.555	0.082	0.917
D8S1179	0.340	0.660	2.941	0.563	0.370	0.250	0.749
D12S391	0.207	0.793	4.831	0.752	0.547	0.087	0.912
D19S433	0.184	0.815	5.435	0.738	0.661	0.030	0.970
Penta D	0.214	0.786	4.673	0.742	0.536	0.093	0.906

17.3 and 16 for vWA locus, allele 30.2 for D21S11 locus, allele 8 for D8S1179 locus, allele 13 for D5S818 locus, allele 18 for D12S391 locus, allele 14.2 for CSF1PO locus and allele 25 for FGA locus.

The good indicators of the genetic polymorphism within the sample are verified by the number of alleles and the expected heterozygosity 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. Also because of bigger population size, the number of observed alleles tends to increase. The usefulness of the markers for genetic screening is verified by the number of alleles scored for each marker.

For a given population, the quality of a polymorphism as a genetic marker was measured by the heterozygosity, the frequency of heterozygosity, or by a polymorphism information content PIC. This indicates the probability that a given offspring of a parent carrying a rare allele at a locus will allow deduction of the parental genotype at the locus. This is determined by adding the mating frequencies multiplied by the probability that an offspring will be formed.[12]

Table 4 The power of discrimination in different populations.

Locus	Iraq[a]	Turkey[b]	Iran[c]	Syria[d]	Emirates[e]	Qatar[d]	Egypt[f]	Gaza[d]
D21S11	–	–	–	–	–	–	–	–
D2S1338	0.974	–	–	–	–	0.973	–	–
D18S51	–	0.953	–	0.963	0.971	–	–	0.962
TPOX	–	0.772	–	–	–	0.855	–	–
Penta E	–	–	–	0.974	–	–	–	0.976
Penta D	–	–	–	0.951	–	–	–	0.961
FGA	–	–	–	0.957	–	–	0.973	0.967
vWA	0.806	–	–	–	–	–	0.937	–
LPL	–	–	0.924	–	–	–	–	–
F13A01	–	–	0.922	–	–	–	–	–
D5S818	–	–	–	–	0.889	–	–	–

[a] This study.
[b] Ref. 21.
[c] Ref. 22.
[d] Ref. 23.
[e] Ref. 24.
[f] Ref. 25.

The CSF1PO locus is the least polymorphic marker while D2S1338 is the most polymorphic marker based on the degree of polymorphism of every marker, expressed in heterozygosity and PIC terms (Table 3). The usefulness of the findings for genetic polymorphism studies and linkage mapping programs in humans is confirmed by the high PIC values of the selected markers.

Based on Butler,[15] the power of exclusion (PE) can be computed to demonstrate how rare it is to find a random man who could not be excluded as the biological father of the child. The PE for every locus in this study was calculated and tabled as planned. For all the microsatellites analyzed the PE ranged from 0.286 (CSF1PO) to 0.712 (D2S1338), with an average of 0.527.

In the case of the power of discrimination, values for all tested loci were 71%, 72%, 74% and 77% for the CSF1PO, D3S1357, D8S1179 and D5S818 loci respectively and 80%, 81% and 88% for the vWA, THO1 and D16S539 loci respectively, and ranged from 90% to 97%, 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 4. 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.

The possibility to find two persons with the same DNA profile if chosen at random in a population is defined as the matching possibility.

4. Conclusion

158 different alleles were observed across the population. The Penta E locus showed the lowest allele frequencies at allele 15 while the highest allele frequencies occurred in the loci D19S433 allele 11. The largest number of different alleles was seen in D6S1043 and D18S51 loci. The locus with the highest heterozygosity was D2S1338, while locus CSF1PO has the lowest heterozygosity, the heterozygosity of the 21 STR loci ranged from 57.7% to 88.5% (mean value 69.6%).

Power of discrimination values for all tested loci were 71%, 72%, 74% and 77% for the CSF1PO, D3S1357, D8S1179 and D5S818 loci, respectively and 80%, 81% and 88% for the vWA, THO1 and D16S539 loci respectively, and ranged from 90% to 97%, for the rest of the loci. Based on statistical parameters, the population of Iraq may use these 21 STR loci as a vital tool for forensic identification and paternity testing.

Funding

None.

Conflict of interest

We declare that we have no conflicts of interest in the authorship or publication of this contribution.

Acknowledgement

I sincerely wish to thank Dr. Ahmed for providing me the opportunity to work on this project. I am indebted to you for choosing the project, your enthusiasm for helping me and your patience and guidance as I progress to put this project together. I am thankful to you for helping me through the various analysis stages, and for providing helpful criticism and feedback throughout the writing process. I also would like to thank Dr. Khalifa from the Institution of medico-legal for all time put in to discuss the project and helping me to put the project together. I would also like to thank Zainab Al-Habubi from the department of biology for her guidance and help in the laboratory work.

References

1. Yamamoto T, Uchihi R, Nozawa H, Huang XL, Leong YK, Tanaka M. Allele distribution at nine STR loci-D3S1358, vWA, FGA, TH01, TPOX, CSF1PO, D5S818, D13S317 and D7S820 in

the Japanese population by multiplex PCR and capillary electrophoresis. *J Forensic Sci* 1999;**44**(1):167–70.
2. Nakamura Y. DNA variations in human and medical genetics: 25 years of my experience. *J Hum Genet* 2009;**54**(1):1–8.
3. Ellegren H. Microsatellites: simple sequences with complex evolution. *Nat Rev Genet* 2004;**5**:435–45.
4. Markoulatos P, Siafakas N, Moncany M. Multiplex polymerase chain reaction: a practical approach. *J Clin Lab Anal* 2002;**16**:47–51.
5. Butler JM, Hill CR. Biology and genetics of new autosomal STR loci useful for forensic DNA analysis. *Forensic Sci Rev* 2012;**24**:15.
6. Silvia B, Marta Campo D, Andrea Z, Dario B, Tatiana G, Luca L. Integration of genomic and gene expression data of childhood all without known aberrations identifies subgroups with specific genetic hallmarks. *Genes Chromosomes Cancer* 2009;**48**:22–38.
7. Mats O, Mo H, Erik W, Tonia S, Natasha L, Tobias U. Mating system variation and morph fluctuations in a polymorphic lizard. *Mol Ecol* 2007;**16**:5307–15.
8. Walkinshaw M, Strickland L, Hamilton H, Denning K, Gayley T. DNA profiling in two Alaskan native populations using HLA-DQA1, PM, and D1S80 loci. *J Forensic Sci* 1996;**41**:478–84.
9. Rui P, Christopher P, Cíntia A, António A, Ángel C, Leonor G. A new multiplex for human identification using insertion/deletion polymorphisms. *Electrophoresis* 2009;**30**:3682–90.
10. Butler JM, Buel E, Crivellente F, McCord B. Forensic DNA typing by capillary electrophoresis using the ABI Prism 310 and 3100 genetic analyzers for STR analysis. *Electrophoresis* 2004;**25**:1397.
11. Laurent E, Guillaume L, Stefan S. Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol Bioinform Online* 2005;**1**:47–50.
12. Botstein D, White RL, Skolnick M, Davis RW. Construction of a genetic linkage map in man using restriction fragment length polymorphism. *Am J Hum Genet* 1980;**32**:324–31.
13. Schneider S, Roessli D, Excoffier L. *Arlequin v 3.1: software for population genetics data analysis*. Available from: http://anthro.unige.ch/arlequin; 2000..
14. Weir B. The second national research council report on forensic DNA evidence. *Am J Hum Genet* 1996;**59**:497.
15. Butler JM, Buel E, Crivellente F, McCord B. Forensic DNA typing by capillary electrophoresis using the ABI Prism 310 and 3100 genetic analyzers for STR analysis. *Electrophoresis* 2009;**25**:1397–412.
16. 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.
17. Fisher RA. Standard calculations for evaluating a blood group system. *Heredity* 1951;**5**:95–102.
18. Jones DA. Blood samples: probability of discrimination. *J Forensic Sci Soc* 1972;**12**:355–8.
19. Brenner C, Morris J. Paternity index calculations in single locus hypervariable DNA probes*Proceedings of the international symposium on human identification, 1989, Promega Corporation, Long Beach, California*. p. 21–53.
20. Botstein D, White RL, Skolnick M, Davis RW. Construction of a genetic linkage map in man using restriction fragment length polymorphism. *Am J Hum Genet* 1980;**32**:324–31.
21. Dogan S, Kovacević L, Marjanović D. Genetic polymorphisms of 15 STR loci within Turkish student population living in Sarajevo, Bosnia and Herzegovina. *Coll Antropol* 2013;**37**(4):1313–9.
22. Arya H, Afrouz N, Maryam H, Amin H. Allele frequencies for 15 autosomal STR loci in Fars province population, southwest of Iran. *Leg Med* 2013;**15**(4):226–8.
23. Louai A, Ichiroh S, Bernd B, Carsten H. Analysis of 15 short tandem repeats reveals significant differences between the Arabian populations from Morocco and Syria. *Leg Med* 2003;S150–5.
24. Alshamali F, Alkhayat AQ, Budowle B, Watson ND. STR population diversity in nine ethnic populations living in Dubai. *Forensic Sci Int* 2005;**152**:267–79.
25. Ahmed A, Linacre A, Mohammed A, Vanezis P, Goodwin W. STR population data for 10 STR loci including the GenePrintTM and PowerPlexTM 1.2 kits from El-Minia (Central Egypt). *Forensic Sci Int* 2001;**117**:233–4.

Please cite this article in press as: Hameed IH et al. Allele frequency data of 21 autosomal Short Tandem Repeat loci in Mesan and Basra provinces in south of Iraq, *Egypt J Forensic Sci* (2015), http://dx.doi.org/10.1016/j.ejfs.2014.10.003