ISSN: 0067-2904
GIF: 0.851

Revealed of A novel Allele in Wasit – Iraqi Population

Haider K. AL-Rubai*, Mohammed M. AL-Zubaidi, Hala K. Ibrahem, Ali Mohammed, Sahar Rashed, Reem Hussam ,Sura Nabeel, Asia Abdullateef, Ali Abdulkaduhm
Forensic DNA Research and Training Center, Al-Nahrain University, Baghdad, Iraq

Abstract

The developments in forensic DNA technology have led us to perform this study in Iraqi population as reference database of autosomal Short Tandem Repeat (aSTR) DNA markers . A total of 120 unrelated individuals from Wasit province were analyzed at 15 STR DNA markers. Allele frequencies of DNA typing loci included in the AmpFlSTR1 IdentifilerTM PCR Amplification Kit panel from Applied Biosystems (D3S1358, vWA, FGA, D8S1179, D21S11, D18S51, D5S818, D13S317, D7S820, TH01, TPOX, CSF1PO, D19S433, D2S1338, D16S539) and several forensic efficiency statistical parameters were estimated from all the sample. the combined Matching Probability (CMP) using the 15 STR genetic loci in Iraqi population was estimated at 1 in 2.08286E-18 and the Combined Discrimination Power (CDP) was greater than 0.9999999 ,Combined Exclusion Probability (CEP) was 0.98350917 which should be sufficient for the identification of any individual.

Keywords: STR, Iraqi Population, Allele frequencies, human identity, statistical parameter

الكشف عن اليلات جديدة في المجتمع العراقي / محافظة واسط

حيدر كاظم الربيعي*، محمد مهدي الزبيدي، هالة خالد ابراهيم، علي محمد عبد المحسن، سهر رشيد، ريم حسام، سرى نبيل، علي عبد الكاظم، اسيا عبد الطيف
مركز الدنا العدلي للبحث والتدريب، جامعة النهرين، بغداد، العراق

الخلاصة

التطور السريع الذي حدث في التقنيات المتعلقة بالحمض النووي العدلي في العالم يدفعنا لاجراء البحوث التي تدعم هذا التوجه سيما وان الابحاث في العراق التي تخص هذا الموضوع لا تواكب التطور العالمي الحديث ، لذا بدأنا بمحافظة واسط ببناء قاعدة بيانات للكشف عن المؤشرات الجسمية الخاصة بهذه الشريحة من المجتمع العراقي autosomal Short Tandem Repeat (STR). تم جمع 120 عينة من متطوعين (لا تربطهم اي صلة قرابة) حصرا من محافظة واسط الجنوبية ، المؤشرات التي تم تحليلها هي 15 موقع منتشر على الكروموسومات الجسمية هذه المواقع هي: (D3S1358, vWA, FGA, D8S1179, D21S11, D18S51, D5S818, D13S317, D7S820, TH01, TPOX, CSF1PO, D19S433, D2S1338, D16S539). عدة قوانين اساسية استعملت في التحليل النتائج على جميع المواقع منها درجة التقارب بين العينات (CMP)، وكانت النتائج تشير الى 1 من اصل 2.08286E−18 وكذلك وقوة التطابق بين العينات 0.999999. كما كانت دقة الاحتمالية كانت 0.98350917 وهذه النتائج يجب ان تكون مطابقة عند اي عمل .

*Email: hayderkhadum@yahoo.com

Introduction

Wasit province is located in the center of Iraq. City of Kut its capital and it is 172 km away from the capital Baghdad. It has international border with Iran to the East, thus Wasit province enjoys a strategic geographic proximity to most of central and southern Provinces and the capital Baghdad figure1.The samples of this study limited only from this province. For DNA typing loci, the commercial kit was used (AmpFlSTR1 Identifile™ PCR Amplification Kit) which include the 13 DNA markers that Federal Bureau of Investigation (FBI) has recommended the forensic labs to use it, known as Combined DNA Identification System (CODIS13), as the loci of choice for forensic use. This recommendation has been accepted by forensic laboratories all over the world, thus the DNA typing can be compared to each other [1, 2].

Short tandem repeat (STR) genetic loci are highly polymorphic repeat sequences, Their importance rises from the fact that are the most informative genetic markers providing high statistical capability of discrimination and individualization in various forensic and judicial settings [3], Their polymorphic nature in human identification is widely acknowledged and documented because of extensive medical and ethnogenetic research that was prompted by various research communities worldwide [4]. Our presented study was performed to establish a genetic database of the wasit Iraqi population for forensic and paternity testing purposes.

Figure 1- Map of Iraq (source: One World - Nations Online all- countries of the world)

Materials and methods

Population: Buccal swabs were collected by oral stick (Sterile Omni Swab or Sterile Foam Tipped Swabs, Whatman International Ltd., Maidstone, UK) from 120 healthy, randomly chosen from wasit Iraqi Population; samples contained both genders (male & female).

DNA extraction: DNA extraction Samples were extracted using a PrepFiler Forensic DNA Extraction Kit (Applied Biosystems, Foster City, CA, for more specificity DNA concentration measured by real time PCR specific gene detection (stander curve Taq man technique) with the

commercial kit (Quntifiler ® Human DNA quantification kit) in 7500 Fast Real-Time PCR System and 7500 SDS software.

PCR amplification and DNA Typing: Fifteen autosomal STR markers (the 13 CODIS core loci and D19S433 and D2S1338) were typed along with amelogenin using the Applied Biosystems AmpFiSTR® Identifiler™ kit (3) 1±2 ng of target DNA following the protocols described in the User's Manual (applyed Biosystems). The samples were amplified by using (verity PCR System) (applyed Biosystems).

Amplification products were diluted with factor 1:15 in Hi- Di™ formamide and GS500-LIZ internal size standard (Applied Biosystems) and analyzed on the 16-capillary 3130XL Genetic Analyzer. POP™-4 (Applied Biosystems) was utilized for higher resolution separations on a 36 cm array. Data collection: was performed with Data Collection v. 2.0 software (Applied Biosystems, Foster City, CA, USA) and samples were analyzed by GeneMapper v. 3.2 software (Applied Biosystems, Foster City, CA, USA). All data Statistically Analysed depended on Hardy–Weinberg principle expectations at each locus, evaluation include: Allele frequencies, power of discrimination (PD), power of exclusion (PE), polymorphism information content (PIC), match probability (MP) and typical paternity index (TPI) were calculated using for each locus of the studied population.

Result and Discussion

The purpose of Quantifiler® Duo DNA Quantification Kit is designed to quantify the total amount of amplifiable human DNA in a sample. The results obtained using the kit can aid in determining: If the sample contains sufficient human DNA to proceed with short tandem repeat (STR) analysis, The concentration of sample to use in STR analysis applications(quantities of DNA in a sample), If PCR inhibitors are present in a sample that may require additional purification before proceeding to STR analysis[4].The range of DNA samples concentrations using Taqman technique real time PCR are (5 -40 ng/µl) .Using stander curve methods to find the unknown values of samples by compare between knowing and unknowns values [5] in figure-2.

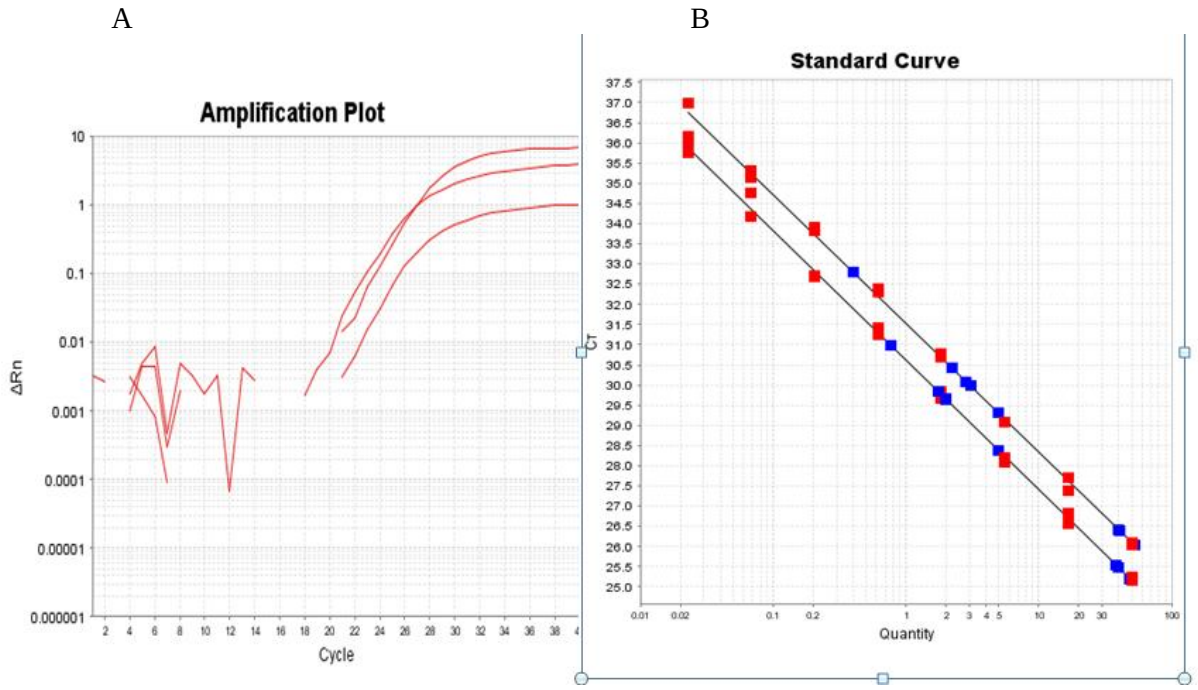

Figure 2- A- amplification plot of 3 unknown samples concentrations compare with positive DNA sample from the commercial kit (Quntifiler ® Human DNA quantification kit) on the standard curve(B).

Allele frequencies and statistical evaluations tests performed to determine the suitability of markers for forensic and paternity studies of the 15 autosomal STR loci[6].The samples analyzed from120 wasit-Iraqi population and the highest allele frequency were observed with allele 8 for TPOX locus (0.538) Table-1, this result match with the previous studies on Iraqi population [7,8]. But also this

allele considered to be the predominant allele in Arab population [6, 9, 10] it cannot represent one of the Iraqi population specificity feature. Our presented study shared most of the result with previous studies with regard to the most common allele Table-3 , minor differences in sequence of the alleles frequency may be back to the size of the data , which means more increase in number of samples ,the database will be more representative [11]. Previously unobserved alleles were detected in our study database; sixteen new alleles at 15 STR DNA markers were detected. These additional alleles ranged from one new allele at D21S11 to 7 new alleles at FGA locus Table-4. Of course, whenever there is expansion on the database there will be revealing a new, more allele [12] and may be this new alleles revealed in this study because its sampled mainly from southern local region (Wasit) compared with previous studies [7, 8]from central and southern Iraq provinces and mainly constituted by Diyala province.

According to Hardy-Weinberg equilibrium (HWE) observed heterozygosity (HO) genotype frequencies and expected heterozygosity (HE) was detected in all 15 markers Table-2, the highest heterozygosity is observed for D21S11 (86.67%) whereas the smallest heterozygosity value is obtained for TPOX (65.00%). The high level of heterozygosity for fifteen STR loci which were genotyped in this study indicated that these loci could be used in determination of individual identification [12]. STR markers show high Power of Discrimination ranged from (PD≥0.82) (TPOX) to (PD≥0.972) (D2S1338 and D18S51). Power of Exclusion (PE) value is over (0.50) except D3S1358 with PD (0.412) and TPOX with PD (0.418).

Polymorphic Information Content (PIC) for all the fifteen STR loci showed a high degree of PIC values shown in Table-2 all STR loci over 0.50 could be considered as highly polymorphic [14]. Allele frequencies were calculate and Hardy-Weinberg Equilibrium was assessed by means of Chi-square ($\chi 2$) test Table-2 and after employing modified Bonferroni correction (test was used to confirm significant differences found in the comparative analysis) the level of significance P value of all 15 loci match with HWE expectations ($P > 0.05$) except the loci : D13S317(0.017) , D18S51 (0.001) and FGA(0.001).

The combined Matching Probability (CMP) using the 15 STR genetic loci in wasit-Iraqi population was estimated at 1 in 2.08286E-18 and hence the Combined Discrimination Power (CDP) was greater than 0.9999999 which should be sufficient for the identification of any individual [12]. Combined Exclusion Probability (CEP) value for the fifteen loci was also calculated at greater than 0.98350917. These results strongly support the application of this set of genetic markers for personal identity and paternity testing [13].

Table 1- Allele frequencies of 15 STR loci in the wasit- Iraqi population

Allele	D8S1179	D21S11	D7S820	CSF1PO	D3S1358	TH01	D13S317	D16S539
6						0.263		
7			0.013	0.004		0.183		
8	0.004		0.188	0.004		0.125	0.138	0.038
9	0.008		0.104	0.025		0.271	0.067	0.092
9.3						0.125		
10	0.071		0.304	0.263		0.033	0.108	0.100
10.2								
11	0.088		0.242	0.375			0.271	0.388
12	0.096		0.129	0.263			0.333	0.242
12.2								
13	0.279		0.021	0.063			0.058	0.125
13.2								
14	0.175			0.004	0.083		0.021	0.017
14.2								
15	0.204				0.238		0.004	
15.2								
16	0.067				0.288			
17	0.008				0.254			
18					0.121			
19					0.013			
20					0.004			
27		0.013						
28		0.158						
29		0.250						
30		0.183						
30.2		0.008						
31		0.038						
31.2		0.138						
32.2		0.133						
33.2		0.075						
36		0.004						

Table 1 continued:

Allele	D2S1338	D19S433	vWA	TPOX	D18S51	D5S818	FGA
6				0.008			
7				0.004			
8				0.538		0.008	
9				0.113	0.004	0.058	
9.2							
10				0.071	0.004	0.104	
10.2					0.004		
11		0.004		0.242	0.021	0.275	
12		0.083		0.025	0.108	0.358	
12.2							
13		0.213	0.004		0.196	0.175	
13.2		0.025					
14		0.258	0.096		0.158	0.017	
14.2		0.029					
15		0.175	0.146		0.150	0.004	
15.2		0.100					
16	0.050	0.058	0.221		0.092		
16.2		0.029			0.004		
17	0.146	0.025	0.254		0.092		
17.2							0.004
18	0.129		0.204		0.058		
18.2					0.004		
19	0.125		0.058		0.071		0.050
19.2							0.004
20	0.167		0.017		0.021		0.125
20.2							0.004
21	0.029				0.008		0.133
21.2							0.004
22	0.050						0.196
22.2							0.008
23	0.138						0.129
24	0.121				0.004		0.221
25	0.038						0.075
26	0.008						0.038
28							0.004
29							0.004

Table 2- Forensic statistical parameters for each of 15 autosomal STR loci (the 13 CODIS core loci and D19S433 and D2S1338) in 120wasit- Iraqi population:

Locus	MP	PD	PIC	PE	TPI	PIE	X2	P-value	HO	HE
D8S1179	0.054	0.946	0.801	0.600	2.857	2.830	27.52	0.543	82.50%	82.33%
D21S11	0.048	0.952	0.814	0.676	3.750	3.028	37.90	0.152	86.67%	83.49%
D7S820	0.078	0.922	0.754	0.495	2.069	2.334	25.60	0.222	75.83%	78.58%
CSF1PO	0.131	0.869	0.666	0.450	1.818	1.767	7.73	0.861	72.50%	71.70%
D3S1358	0.087	0.913	0.738	0.439	1.765	2.218	14.79	0.610	71.67%	77.46%
TH01	0.075	0.925	0.760	0.379	1.500	2.401	23.35	0.178	66.67%	79.18%
D13S317	0.082	0.918	0.745	0.532	2.308	2.238	35.55	0.017	78.33%	77.66%
D16S539	0.093	0.907	0.722	0.483	2.000	2.047	46.86	0.002	75.00%	75.57%
D2S1338	0.028	0.972	0.865	0.629	3.158	4.092	30.37	0.971	84.17%	87.78%
D19S433	0.047	0.953	0.814	0.585	2.727	3.015	27.01	0.860	81.67%	83.42%
vWA	0.063	0.937	0.784	0.532	2.308	2.643	12.54	0.973	78.33%	81.08%
TPOX	0.180	0.820	0.588	0.360	1.429	1.367	15.24	0.292	65.00%	63.43%
D18S51	0.028	0.972	0.864	0.545	2.400	4.035	83.42	0.001	79.17%	87.61%
D5S818	0.101	0.899	0.712	0.483	2.000	2.006	17.83	0.467	75.00%	75.07%
FGA	0.038	0.962	0.836	0.629	3.158	3.403	67.25	0.001	84.17%	85.31%
Total Alleles of each locus= **240**										

CMP	2.08286E-18		Combined Matching Probability (CMP)
CDP	0.9999999		Combined Discrimination Power (CDP)
CEP	0.98350917		Combined Exclusion Probability (CEP)

Matching Probability (MP), Power of Discrimination (PD), Polymorphism Information Content (PIC),Chi square (X2) ,Power of Exclusion (PE), Typical Paternity Index (TPI), Paternity Index Expected (PIE), Observed Heterozygosity(HO),Expected Heterozygosity(HE)

Table 3- Comparison between our studies and previously published studies include the three most common allele and frequencies at each 15 STR loci:

15 STR loci	Allele &allele frequencies from our Study	Allele &allele frequencies from (Salwa J. *et al.,2014*).	Allele &allele frequencies from (Barni F. *et al*.,2007)
	# sample=280	# sample=278	# sample=206
D8S1179	13(0.279),15(0.204),14(0.175)	13(0.295),14(.0212)15(0.158)	15(0.218),13(0.213),14(0.140)
D21S11	30(0.183),28(0.158),31.2(0.138)	30(0.259),29(0.241),28(0.133)	0(0.242),29(0.222),32.2(0.148)
D7S820	10(0.304),11(0.241),8(0.188)	11(0.255),10(0.248),8(0.180)	10(0.331),11(0.207),8(0.188)
CSF1PO	11(0.343),10 &11 (0.263)	12(0.317),11(0.255),10(0.259)	12(0.328),11(0.284),10(0.264)
D3S1358	16(0.288),17(0.254),15(0.238)	16(0.266),17(0.255),15(252)	17(0.364),16(0.237),15(0.228)
TH01	9(0.271),6(0.263),7(0.183)	6(0.295),9(0.273),7(0151)	6(0.305), 9(0.237),7(0.169)
D13S317	12(0333),11(0.271),8(0.138)	12(0.342),11(0259),8(0.140)	12(0.344),8(0.233),11(0.223)
D16S539	11(0.388),12(0.242),13(0.125)	11(0.302),9(0.209),12(0.194)	11(0.354),12(0.291),9(0.135)
D2S1338	20(0.167),17(0.146),23(0.138)	19(0.187),20(162),23(0.147)	17(0.223),20(0.179),23(0.131)
D19S433	13(0.213),14(0.258),15(0.175)	13(0.255)14(0.104),15(0.137)	14(0.228),15(0.218),13(0.184)
Vwa	17(0.254),16(0.221),18(0.204)	17(0.248),16(0.223),18(0.194)	16(0.320),18(0.228),17(0.218
TPOX	8(0.538),11(0.242),9(0.113)	8(0.486),11(0.248)10(0.199)	8(0.543),11(0.242),10(0.106)
D18S51	13(0.196),14(0.158),15(0.150)	13(0.194)14(0.173),15(0.155)	14(0.201),13(0.176),12(0.147)
D5S818	12 (0.358),11(104),13(0.215)	12(0.367),11(0.277),13(0.198)	12(0.349),11(0.320),13(0.203)
FGA	24(0.221),22(0.196),21(0.133)	21(0.191)24(0.187),22,24,24.2(0.140)	24(0.208),23(0.199),21(0.165)

Table 4- The new alleles in Iraqi Database

STR Markers	Allele	# of Observations	Frequency
D8S1179	-	-	-
D21S11	36	1	0.004
D7S820	-	-	-
CSF1PO	7	1	0.004
D3S1358	-	-	-
TH01	-	-	-
D13S317	15	1	0.004
D16S539	-	-	-
D2S1338	-	-	-
D19S433	-	-	-
Vwa	13	1	0.004
TPOX	7	1	0.004
D18S51	9	1	0.004
	10.2	1	0.004
	18.2	1	0.004
D5S818	15	1	0.004
FGA	17.2	1	0.004
	19.2	1	0.004
	20.2	1	0.004
	21.2	1	0.004
	22.2	2	0.008
	28	1	0.004
	29	1	0.004

References

1. Christian, M. Ruilberg, Dennis, J. Reeder and Butler **.2001.** STR Base: A short tandem repeat DNA database for the human identity testing community. *Nucleic acids Research*, 29(1):320-322.
2. Untoro, E., Atmadja, D. S., Pu, C. and Wub, F. **2009**. Allele frequency of CODIS 13 in Indonesian population. *Legal Medicine*.
3. Aysim,T., Zerrin, E., Aysun, C. and Yesim, D .**2010**. Allele distribution data for 16 short tandem repeat loci in BoluTurk. *J Med Sci*, 40 (4): 659-664.
4. Yasin, R. S., Mawieh, M. H., Ali, Z. E., Adnan, S. J. **2005**.African Jordanian Population Genetic Database on Fifteen Short Tandem Repeat Genetic Loci. *Croat Med J*, 46(4):587-592.
5. Applied Biosystems Quantifiler® Duo DNA Quantification Kit. **2012.** Life Technologies Corporation User's Manual, Part Number 4391294 Rev. D 03/2012.
6. Alshamali, F., Abdul, Q. A., Bruce, B., Watson, N.D. **2005.** STR population diversity in nine ethnic populations living in Dubai. Forensic Science International, 152: 267–279.
7. Barni , F., Berti, A., Pianese, A., and Boccellino, A. **2007**. Allele frequencies of 15 autosomal STR loci in the Iraq population with comparisons to other populations from the middle-eastern region. *Forensic Science International*, 167 (1): 87-92.
8. AL- Awadi, S. J., AL- Zubaidi, M. M., Namaa, D. S., and Saleh, T. Y. **2014**. Genetic Variation of 15 Autosomal Short Tandem repeat Repeat (STR) Loci in the Diyala – Iraqi Population. *International Journal of Biological and Pharmaceutical Research*, 5(3):131-135.
9. Alenizi, M., Goodwin, W., Ismael, S. and Hadi, S. **2008.** STR data for the AmpF STR Identifiler loci in Kuwaiti population .*Legal Medicine*, 10: 321-325.
10. El-Morsi, D. A., El-Bakary, A. A. and El Baz, R. **2009**. Alleles frequency distribution of two STR loci in Egyptian population, *Mansoura J. Forensic Med. Clin. Toxicol.* XVII (2).
11. Butler, J.M. **2006.** Genetics and genomics of core short tandem repeats loci used in human identity testing. *J Forensic Sci*., 51: 253-276.
12. Sheena, M. B. Maiquilla, Salvador, J. M., Gayvelline C. Calacal, Frederick C. Delfin, Kristina A. Tabbada, Henry B. Perdigon, and Minerva S. Sagum.**2011**. Expansion of the Philippine Autosomal Short Tandem. *Philippine Journal of Science*, 140 (1): 1-6.

13. Reyhaneh, L., Vallian, S**. 2009**. Genetic variation of informative short tandem repeat (STR) loci in an Iranian population. *Iranian journal of biotechnology*, 7(3):137-141.
14. Shriver, M. D., Jin, L., Boerwinkle, E., Deka, R., Ferrell, R. E. and Chakraborty, R**. 1995**. A novel measure of genetic distance for highly polymorphic tandem repeat loci. *Mol. Biol. Evol.*, 12: 914-920.
15. AL-Zubaidi, M. M., Saleh, M. A.-D. and Alawedi, S. J**. 2014.** Genetic Variation of 15 autosomal short tandem repeat (STR) loci in samples of Palestinian population residing in Iraq. *International Journal of Development Research*, 4(1): 153-156.