

# Bio-Molecular Allele Frequencies Availability of 23Y-Str Markers Polymorphisms and Odontogenic Diseases in Iraqi Population Study

Mohammed Mahdi Al-Zubaidi<sup>1</sup>, Ammar Riyad Mohi<sup>2</sup>

<sup>1</sup>PhD Academic Researcher, Department of Training and Development, Forensic DNA Centre for Research and Training, Al-Nahrain University, Baghdad, Iraq, <sup>2</sup>Head and Chairman of Clinical Dentistry Department, Faculty of Dentistry, Al-Farahidi University, Baghdad, Iraq

## Abstract

Orofacial diseases of multidiscipline nature from carious lesions to tumors are significant diseases attack a large number of individuals around the world. However, In the recent years, bioinformatics methods have been reported to analyses several loci of simultaneously. Moreover, Analysis of molecular variance (AMOVA) (23) was performed to measure the genetic distance between the Iraqi population sample and reference data from populations. In this study, we analyzed the genetic polymorphisms of 23 Y-STR loci from PowerPlex® Y23 system in 313 unrelated healthy male individuals from Iraqi populations, and observed 283 different haplotypes including 269 unique haplotypes and 16 duplicate haplotypes. in which 283 different haplotypes was identified classified into 264 (93, 23%) as a unique Haplotype and 19 (6, 77%) as replicated haplotype among individuals. The highest haplotype were found in sample H45, H85, H135, H213 and H241 at frequency 0.010 ( haplotype replicated 3 times ), While 264 haplotype with 0.0033 frequency and haplotype diversity was 0.996 with a discrimination capacity 0.93 respectively. The gene diversity values ranged from 0.3942 at DYS438 to 0.9607 at DYS385a/b. Such results indicated that the 23 Y-STR loci were highly polymorphic in Bagdad population and played crucial roles in forensic application as well as population genetics. For the first time, we reported the genetic diversity of male lineages in Bagdad population at a high-resolution level of 23 Y-STR set and consequently contributed to familial searching, Odontologic science application tracking, and anthropology analysis of Bagdad population.

**Keywords:** Iraqi population ; 23Y-STR markers ; Bio-molecular allele ; patients; genetic

## Introduction

The incidence and mortality of various cancers are associated with sex-specific disparities. Men are more than twice as likely as women to develop oral cancer, according to new figures released by Cancer Research UK, particularly Odontologic malignancies such as Oral squamous cell carcinoma (OSCC).

However, amplifying of 15 Y-specific STR loci by primers, all the cancerous tissue Y chromosomes seemed to be intact. These results indicate when malignant neoplastic tissue specimens are used, that amelogenin-based gender determination should be carefully interpreted.

In same instance, fine underlying basis of the carious lesions were supported genetically but, it is still unclear. These events basically often goes undetected until advanced stages which are related to high epidemic rate, and as such it carries heavy personal and societal costs.

Iraq has a population of 30,399,572 (July 2011 estimation). Individuals for this study were sampled from Baghdad. The commercially available AmpF!STR® Identifier™ PCR Amplification Kit used for this study genotypes the thirteen Combined DNA Identification System loci (CODIS13), as recommended for use by the Federal Bureau of Investigation plus two additional loci (D19S433 and D2S1338). These thirteen CODIS loci are used world-wide by forensic laboratories and, as a result, facilitate comparisons between populations

as well as the generation of investigative leads<sup>(1)</sup>. Arabs have many ethnic-governorates, so it is important to study the inheritance of Iraqi society<sup>(2,3)</sup>.

Short tandem repeat (STR) loci are highly polymorphic and found in abundance throughout the human genome<sup>(4)</sup>. These genetic markers have proven to be particularly well suited for medical research, ethnographic studies and for discrimination/individualization in a variety of forensic and judicial settings<sup>(5)</sup>.

Human Y chromosome short tandem repeats (Y-STRs) are repeating regions with 2-7 bp long repetitive units found in the non-recombining region of Y chromosome. Y-STRs are characterized by male inheritance pattern.<sup>(6-11)</sup>

Thus, Y-STRs can be used in forensics for the investigation of sexual assault cases. This kit types 23 Y-STR loci in a tested haplotype and includes 6 new Y-STR loci when compared to the previous Y-STR commercial kits, namely DYS576, DYS481, DYS549, DYS533, DYS570, and DYS643<sup>(12,13)</sup>.

This work presents a Y-chromosomal STR haplotype analysis in the Iraqi Arab population according to 23 forensic Y-STR loci 23 Y-chromosomal STRs loci (DYS576, DYS3891, DYS448, DYS389II, DYS19, DYS481, DYS549, DYS533, DYS438, DYS437, DYS570, DYS635, DYS390, DYS439, DYS392, DYS643, DYS393, DYS458,a/b, DYS456 and YGATAH4). These markers are highly sensitive, already use in forensic practice<sup>(14,15)</sup>.

Previous studies were conducted in Iraq or in neighboring countries for Iraqi individuals from different ethnics and by using different Y-STR loci<sup>(16-19)</sup>. In this study we use 23Y-STR loci. The aim is to study the genetic structure, Haplogroup, haplotype and introduce reference database which may be used by forensic purposes.

## Materials and Methods

A total of 307 unrelated (not belonging to the same nuclear family) healthy male were sampled for the Y-STR analysis from the Iraqi populations. blood samples were collected at the ministry of health/medical legal Directorate /paternity and Kinship Section. The

collected blood used as a DNA source, placed on what man FTA cards, Air-dried and placed in paper envelopes and stored at +4°C for DNA extraction.

## Y chromosome genotyping

PCR amplification of 23 Y-chromosomal STRs loci (DYS576, DYS3891, DYS448, DYS389II, DYS19, DYS481, DYS549, DYS533, DYS438, DYS437, DYS570, DYS635, DYS390, DYS439, DYS392, DYS643, DYS393, DYS458,a/b, DYS456 and YGATAH4) was performed using PowerPlex Y 23 System (Promega Corporation) according to the manufacturer's recommendations<sup>(20)</sup>.

## Statistical Analysis

Allele and haplotype frequencies were estimated by gene counting method. The allele frequency of the multicopy locus DYS385 a/b was analyzed as a combination of two alleles<sup>(21)</sup>. Gene diversity (GD) of each locus and haplotype diversity (HD) was calculated using the formula  $D = (n/n-1)(1-\sum p_i^2)$ , where  $n$  represents the total number of alleles or haplotypes, and  $p_i$  is the frequency of the  $i$ th allele or haplotype.<sup>(22)</sup>

Haplogroups prediction was performed by The NevGen Y-DNA Haplogroup Predictor.

Analysis of molecular variance (AMOVA)<sup>(23)</sup> was performed to measure the genetic distance between the Iraqi population sample and OTER reference data from populations submitted to the Y-STR Haplotype (YHRD)<sup>(24)</sup>.

## Results

Statistical evaluations were performed to determine the suitability of the allele frequency databases in the Iraqi population was found to have the highest frequency (0.785). **Table (1)** There are many alleles that had previously been observed at very low frequency (9 and 17 allele at DYS389I with a frequency of 0.003; 17 allele at DYS448 with a frequency of 0.003; and 25 allele at DYS389II with a frequency of 0.003). The GD value for DYS385 a/b among the Baghdad of Iraq males was 0.9358.

### Table (1): Allele, Genotype frequencies and Gene

Diversity for 307 Iraqi Arabs male

Aalle	DYS576	DYS389I	DYS448	DYS389II	DYS19	DYS391	DYS481	DYS549	DYS533	DYS438	DYS437	DYS570
7										0.003		
8												
9		0.003				0.036			0.029	0.241		
10		0.007			0.007	0.606		0.003	0.104	0.616		
11		0.003				0.313		0.088	0.560	0.104	0.003	
12		0.124			0.010	0.046		0.511	0.277	0.036		
13		0.704			0.091			0.322	0.029		0.013	0.016
14	0.003	0.150			0.619			0.065			0.674	0.016
15	0.042	0.007			0.186			0.010			0.205	0.046
16	0.111				0.065						0.085	0.107
17	0.280	0.003	0.003		0.023							0.241
18	0.352		0.023									0.319
19	0.173		0.195									0.173
20	0.026		0.661				0.020					0.059
21	0.013		0.088				0.052					0.016
22			0.029				0.186					0.007
23							0.205					
24							0.163					
25				0.003			0.147					
26				0.010			0.186					
27				0.007			0.036					
28				0.075			0.007					
29				0.251								
30				0.489								
31				0.137								
32				0.029								
Std. Error	.068	.039	.041	.058	4.892	.036	.099	.331	.044	.041	.040	.086
Std.	1.188	.691	.716	1.013	85.714	.630	1.741	5.796	.767	.717	.698	1.508
GD	0.753	0.467	0.515	0.673	0.569	0.532	0.837	0.622	0.597	0.551	0.496	0.792

GD: Gene Diversity, SD: Standard Deviation

Also, the haplotype and haplotype frequency for the 160 sample each with 23 alleles were analyzed by Using Excel feature (filtering) are shown in **Table (2)**. The highest haplotype were found in sample H45, H85,

H135, H213 and H241 at frequency 0.010 ( haplotype replicated 3 times ), While 264 haplotype with 0.0033 frequency and haplotype diversity was 0.996 with a discrimination capacity 0.93, Discrimination capacity (DC). **Table (3)**

**Table 2: Y-STR haplotypes and haplotype frequencies in Iraq Arab populations**

number	DYS576	DYS3891	DYS448	DYS389II	DYS19	DYS391	DYS481	DYS549	DYS533	DYS438	DYS437	DYS570	DYS635	DYS390	DYS439	DYS392	DYS643	DYS393	DYS458	DYS385/a	DYS385/b	DYS456	YGATAH4	frequency
H8	17	13	19	30	14	10	22	13	12	12	15	17	23	24	12	13	10	12	16	11	14	15	12	0.007
H12	17	13	20	30	15	10	23	12	12	11	14	18	23	25	10	11	10	13	16	11	14	15	13	0.007
H25	18	13	20	30	17	10	23	12	12	11	14	18	23	23	11	11	10	13	15	11	16	15	13	0.007
H37	18	13	20	30	14	10	23	12	11	9	14	16	21	23	11	11	10	12	17	12	15	16	11	0.007
H45	19	12	21	28	15	10	22	12	9	10	16	19	21	23	12	11	11	15	16	12	16	15	12	0.010
H58	18	14	19	30	14	10	24	11	10	11	15	17	24	23	11	11	10	15	15	12	22	15	12	0.007
H85	16	13	20	29	14	10	24	13	11	10	15	18	24	23	10	11	11	12	16	13	16	15	11	0.010
H93	18	13	20	30	14	10	25	12	11	10	14	18	20	23	11	10	9	13	17	13	16	15	11	0.007
H135	18	13	20	30	14	11	26	13	11	10	14	18	21	23	11	11	9	12	18	13	18	14	11	0.010
H150	19	13	20	30	14	11	26	12	10	9	14	18	21	23	11	11	9	12	19	13	18	13	11	0.007
H175	18	13	20	30	14	11	25	12	11	10	14	18	21	23	11	11	9	12	18	13	19	14	11	0.007
H185	18	13	20	31	17	10	25	13	11	10	14	18	21	23	11	11	9	13	19	13	19	16	11	0.007
H195	17	13	20	30	14	11	26	13	11	10	14	18	21	23	11	11	9	12	18	13	20	14	11	0.007
H204	16	13	20	29	14	10	22	13	11	9	14	16	21	22	12	11	8	12	15	14	15	16	11	0.007
H213	17	13	21	30	15	10	20	12	12	9	16	18	21	26	11	11	9	13	16	14	16	15	11	0.010
H217	19	13	20	30	15	10	22	11	11	9	15	16	22	23	11	11	12	12	18	14	16	15	12	0.007
H228	17	13	20	31	14	10	24	14	12	10	14	18	20	23	11	11	9	12	18	14	18	15	11	0.007
H241	16	13	19	30	13	10	25	12	12	10	14	20	23	24	12	11	13	13	16	15	16	17	11	0.010
H248	17	12	20	29	13	9	26	13	11	10	14	18	20	23	12	11	11	14	17	15	17	15	12	0.007
	18	13	20	30	17	10	25	12	12	11	14	20	23	25	10	11	10	13	16	10	13	16	12	0.003

The rest of haplotype frequency has the same frequency (0.003) and appear 1 time which showed in an appendix the pattern has the same frequency

**Table (3): Diversity statistics of Y-STR haplotypes.**

	23Y-STR haplotype
Sample size	307
Number of haplotypes	283
Unique haplotype	264
Repeating haplotype	19
Haplotype Diversity	0.996
Discrimination capacity	0.933

Haplogroup prediction suggests that Haplogroup J1 is the most common in the Iraqi population encompassing (39%) of men, followed by J2 (20%), R1,E1 (10.8%), and G2(6%) with other predicted Haplogroups (T,L1,I2,R2,G1,Q,H1,N1,O2,I1,C2 AND O2) occurring at progressively lower frequencies table (4).

**Table (4): Frequencies of predicted Haplogroups for Iraqi Population .**

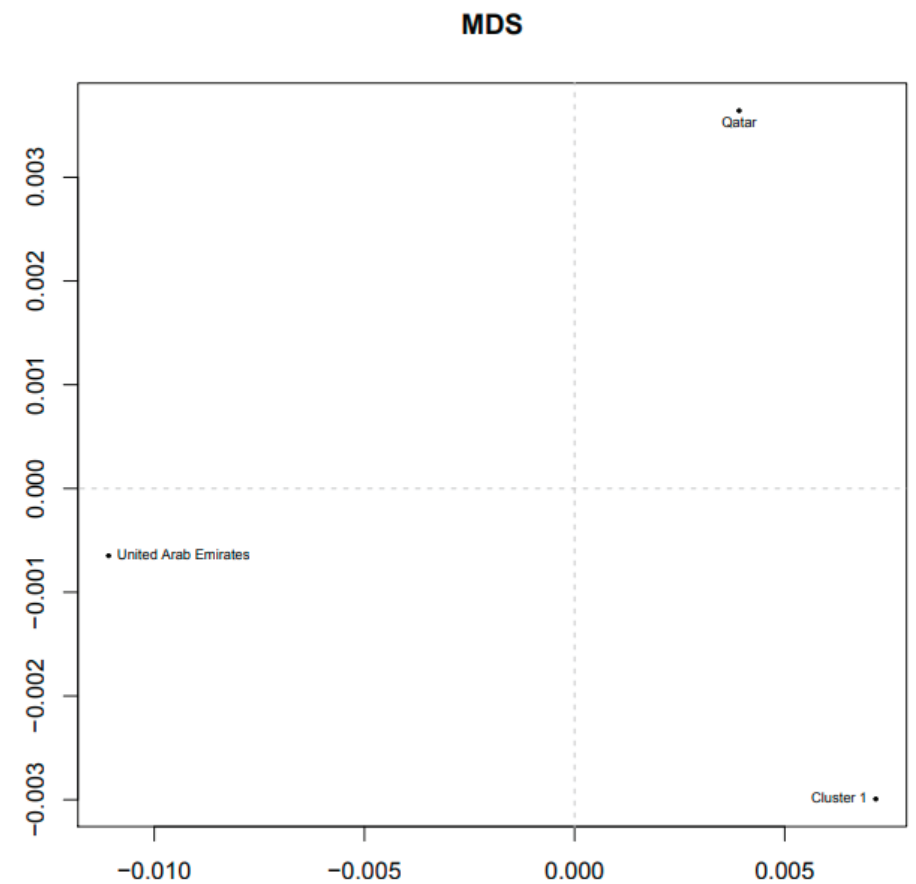
Haplogroup	number	Freq.
J1a	112	0.396
J2a	57	0.201
R1	30	0.106
E1	30	0.106
G2	17	0.060
T	10	0.035
L1	5	0.018
I2c	4	0.014
R2	3	0.011
G1	3	0.011
Unsupported-subclade	3	0.011
Q	2	0.007
H1	2	0.007
N1a1	1	0.004
O2	1	0.004
I1	1	0.004
C2b	1	0.004
O1	1	0.004

AMOVA was used to explore population differentiation with pairwise genetic distances ( $R_{st}$ ) visualized by multidimensional scaling (MDS), Iraq population samples with Some Arab countries obtained from YHRD Release (Table 5 and Figure 1).

**Table (5):  $R_{ST}$  values for pairwise comparisons of haplotypes between Iraq population with Some Arab countries.**

Population	Qatar	United Arab Emirates	“Cluster 1”	Cluster 1
Qatar	-	0.0013		0.0044
United Arab Emirate	0.0156	-		0.0000
Cluster 1	0.0074	0.0184		-

Cluster 1 Iraq; Iraqi; Kuwait; Lebanon



Cluster 1: Iraq; Iraqi; Kuwait; Lebanon

**Figure (1): Multidimensional scaling (MDS) plot based on pairwise  $F_{st}$  genetic distance among Iraq population samples with Some Arab countries (Qatar, United Arab Emirates, Kuwait and Lebanon) population using the 23Y-STR power plex.**

### Discussion

Human identification is an important forensic application and other various strategies have been adopted to utilise the minimum quantity of sample to detect maximum variability in a single PCR reaction. Our study using one such kit (AmpF/STR Profiler Plus, Perkin Elmer) showed a deletion of the amelogenin gene on the Y chromosome in five Indian males.

Statistical validation of the data is essential for implementation to central reference data of Iraq. According to our information, All the population individuals have a high number of allelic ranges nearly in all loci. High allelic ranges result in high gene diversity value which has been observed across all the Iraqi population. are graphical representations of allele frequency distributions among the populations studied

for All samples .

All Y-STR loci show a unimodal distribution modal distribution with one frequent allele and with less frequent adjacent alleles, differing by one repeat unit from the most frequent allele. (25) Most males of all governorates were within the predicted Haplogroups J1,, then J2, E1 and others with lower incidence. These also highlighted the evidence of the similarity them in genetically based manner. Analysis of genetic s using Rst AMOVA and visualized by multidimensional plot showed high similarities among population.

Even though the frequency of considering the fact that the future of odontologic diseases and forensics, is based on the reliability of this test, one should make a conclusions about the gender susceptibility by using alone. We, therefore, suggest the inclusion of additional

Y chromosome markers such as SRY, STR, STS and/or other Y chromosome markers in the existing multiplex STR kits for tumor early detection also in manner of public cariology susceptibility identification studies.

### Conclusion

Gene diversity values for each Y-STR loci have been presented in this study, is likely emphasized that There are some polymorphic signals among Iraqi populations with significantly different allelic frequencies among healthy control peoples, indicating that belongings to these lineages would potentially increase the level of susceptibility or resistance to Odontologic problems.

**Ethical Clearance:** The Research Ethical Committee at scientific research by ethical approval of both MOH and MOHSER in Iraq

**Conflict of Interest:** None

**Funding:** Self-funding

### References

- [1] Untoro E, Atmadja DS, Pu C-E, Wu F-C. Allele frequency of CODIS 13 in Indonesian population. *Int J Legal Med*. Elsevier; 2009;11:S203–S205.
- [2] -Central\_Intelligence\_Agency. The World Factbook: Central Intelligence Agency; (2016): 2016.
- [3] Kirmanj S. Identity and Nation in Iraq: Lynne Rienner Publishers; 2013.
- [4] Yasin SR, Hamad MM, ElKarmi AZ, Jaran AS. African Jordanian population genetic database on fifteen short tandem repeat genetic loci. *Croatian Medical Journal. MEDICINSKA NAKLADA VLASKA* 69, HR-10000 ZAGREB, CROATIA; 2005.
- [5] Tuğ A, Erkol Z, Çetinyürek A, Alakoç YD, Elma C, Büken B, et al. Allele distribution data for 16 short tandem repeat loci in Bolu. *Turkish J Med Sci. The Scientific and Technological Research Council of Turkey*; 2010;40(4):659–64.
- [6] Serin a, canan h, alper b, Sertdemir Y. haplotype frequencies of 17 Y-chromosomal short tandem repeat loci from the cukurova region of turkey. *croat Med J*. 2011;52:703-8.
- [7] cenanović M, Pojskić n, Kovacević I, Dzehverović M, cakar J, Musemić Dz, et al. Diversity of Y-short tandem repeats in the representative sample of the population of canton Sarajevo residents, bosnia and herzegovina. *coll antropol*. 2010;34:545-50.
- [8] grsković b, Mršić g, Vrdoljak a, Merkas S, andelinović S. Population genetic analysis of haplotypes based on 17 short tandem repeat loci on Y-chromosome in population sample from eastern croatia. *croat Med J*. 2010;51:202-8.
- [9] butler JM. *Fundamentals of forensic Dna typing*. 3rd ed. burlington, Ma: elsevier; 2010.
- [10] butler JM. *advanced topics in forensic Dna typing: Methodology*. 1st ed. burlington, Ma: elsevier; 2011.
- [11] Kovačević I, Fatur-cerić V, hadžić n, Čakar J, Primorac D, Marjanović D. haplotype Data for 23 Y-chromosome markers in a reference sample from bosnia and herzegovina. *croat Med J*. 2013;54:286-90.
- [12] coble MD, hill cr, butler JM. haplotype data for 23 Y-chromosome markers in four uS population groups. *Forensic Sci int genet*. 2013;7:e66-8.
- [13] Dogan, S., Primorac, D., & Marjanović, D. Genetic analysis of haplotype data for 23 Y-chromosome short tandem repeat loci in the Turkish population recently settled in Sarajevo, Bosnia and Herzegovina. *Croatian medical journal*, (2014). 55(5), 530.
- [14] Kayser ,M.; Krawczak, M.; Excoffier, L.; Dieltjes, P.; Corach, D.; Pascali, V. et al. An extensive analysis of Y-chromosomal microsatellite haplotypes in globally dispersed human populations. *Am J Hum Genet*. (2001). 68(4): 990–1018.
- [15] Roewer, L.; Kayser, M.; de Knijff P, Anslinger, K.; Betz, A.; Caglià ,A. et al. A new method for the evaluation of matches in non-recombining genomes: application to Y-chromosomal short tandem repeat (STR) haplotypes in European males. *Forensic Sci Int*. (2000). 114: 31–43.
- [16] Imad ,H.;Cheah, Q.; Mohammad, J. and Aamera, J. Genetic variation of 17 Y-chromosomal short tandem repeats (STRs) loci from unrelated individuals in Iraq. *Int. J. Biotechnol. Mol. Biol. Res*. (2013). 12: 120-130.
- [17] Al-Zahery, N.; Pala, M.; Battaglia, V. *et al*. In search of the genetic footprints of Sumerians: a survey of Y-chromosome and mtDNA variation in the Marsh Arabs of Iraq. *BMC Evol Biol* (2011). 11, 288 .

- [18] Dogan, S.; Gurkan ,C.; Dogan, M.; Balkaya, H.E.; Tunc, R.; Demirdov, D.K. et al. A glimpse at the intricate mosaic of ethnigovernorates from Mesopotamia: Paternal lineages of the Northern Iraqi Arabs, Kurds, Syriacs, Turkmens and Yazidis. *PLoS ONE* (2017). 12(11).
- [19] Nei M. *Molecular evolutionary genetics*. New York: Columbia University Press; 1987.
- [20] PromegaPowerPlex® Y23 system – technical manual (tMD035). Madison, Wi: Promega corporation; 2012.
- [21] Butler, J. M. Recent developments in Y-short tandem repeat and Y-single nucleotide polymorphism analysis. *Forensic Sci Rev.* (2003). 15(2): 91-111.
- [22] Kayser, M .; Caglià, A.; Corach, D.; Fretwell, N.; Gehrig, C.; Graziosi, G. et al. Evaluation of Y-chromosomal STRs: a multicenter study. *Int J Leg. Med* (1997). 110: 125–133.
- [23] Roewer ,L.; Kayser, M.; Dieltjes, P.; Nagy, M.; Bakker, E. and Krawczak, M. et al. Analysis of molecular variance (AMOVA) of Y-chromosome-specific microsatellites in two closely related human populations. *Hum Mol Genet.* (1996). 5(7): 1029–33.
- [24] Willuweit, S. and Roewer, L. Y chromosome haplotype reference database YHRD: Update. *Forensic Science International: Genetics*, (2007). 1: 83-87.
- [25] De Knijff, P.; Kayser, M.; Caglià, A.; Corach, D.; Fretwell, N.; Gehrig, C.; Graziosi, G. and Heidorn, F. Chromosome Y microsatellites: population genetic and evolutionary aspects. *Int. J. Legal Med.*, (1997). 110: 134-140.