Egyptian Journal of Forensic Sciences

ORIGINAL ARTICLE Open Access

Frequency of Y-chromosome STRs using PowerPlex® Y23 System in Iraqi population

Hanan K. Mahmood[1], Nadia F. Salman[1*], Khalifa M. Salih[1], Dhurgham H. Hasan[1] and Mohammed M. Al-Zubaidi[2]

Abstract

Background: Y-chromosome STRs are valuable in the forensic identification of male DNA from sexual assault cases, and they are used to link families through genetic genealogy.

Materials and methods: For Y-STR analysis, 1032 male blood samples were used in this study, direct PCR technique was used for DNA amplification using the PowerPlex® Y23 System, and then PCR product was run with Genetic Analyzer, and the data were analyzed with the Gene Mapper ID Analysis Software. Frequency-based statistical analysis was calculated with GenAlEx 6.5-Genetic Analysis.

Results: One hundred and eighty-five alleles were detected at the 23 Y-STR loci in 1032 samples. Alleles frequency ranged from 0.002 to 0.813 and the highest allele frequency registered (0.813) for allele 11 at locus DYS392, and the mean haplotype diversity was 0.616 ± 0.027. New variants were registered for DYS458 locus.

Conclusion: The present study established the genetic information obtained by using the PowerPlex® Y23 System for the Iraqi population and also created a database of 23 Y STR markers in this population.

Keywords: Y-chromosome STRs, PowerPlex® Y23 System, Locus, Haplotype

Background

Y-chromosome short tandem repeats (STRs) are valuable in the forensic identification of male DNA from sexual assault cases; also, it has great importance in tracing paternal lineages that help in missing persons investigations and historical studies and it aid to link families through genetic genealogy (Roewer, 2009; Gusmão et al., 2006). The use of Y-STR markers has many benefits due to their genetic characteristics in paternal cases and kinship between male family members. Y-STR markers are located on the non-recombining region (NRY) of the Y chromosome; thus, they are transmitted as haplotypes in the same way as single locus alleles, and the firm paternal inheritance of Y-STRs indicates their importance in paternity and kinship tests (Gusmão et al., 2006; Jain et al., 2016).

Requests for kinship testing between different family members are sent from courts of law and police stations from all over Iraq to Paternity and Kinship division in Medico-Legal Directorate (MLD), Baghdad, to establish the identity of an unregistered people or to solve a familial dispute. In many of these cases, the use of DNA short tandem repeat (STR) is not enough to resolve complicated cases. For the past 5 years, the PowerPlex® Y23 System has been relied upon for solving paternity and kinship cases that were referred to the Paternity and Kinship division in MLD. The frequency of Y-STR haplotype was assessed in many populations due to their importance in the statistical analysis of requested cases (Parvathy et al., 2012; Roewer et al., 2008). Here in this study, we represent the frequency of alleles registered in each locus in Iraqi male samples that were collected from all over the country (except Kurdistan), which may provide more information about the genetic pool of the Iraqi population.

* Correspondence: nadia83fadhil@gmail.com
[1]Medico-legal Directorate, Baghdad, Iraq
Full list of author information is available at the end of the article

Table 1 Alleles frequency of PowerPlex® Y23 System loci for a sample of Iraqi Arab population

Allele/ *n*	DYS 576	DYS 389 I	DYS 448	DYS 389 II	DYS 19	DYS 391	DYS 481	DYS 549	DYS 533	DYS 438	DYS 437	DYS 570	DYS 635	DYS 390	DYS 439	DYS 392	DYS 643	DYS 393	DYS 458	DYS 456	YGA TAH4	DYS385 a, b haplotype frequency	
1																	**0.002**					**10, 14**	**0.002**
7						**0.002**											**0.021**					**11, 13**	**0.002**
8		**0.002**				**0.048**			**0.017**	**0.240**						**0.002**	**0.447**				**0.008**	**11, 14**	**0.061**
9		**0.004**			**0.002**	**0.557**		**0.010**	**0.084**	**0.584**					**0.070**	**0.016**	**0.238**				**0.065**	**11, 15**	**0.021**
10		**0.008**			**0.002**	**0.340**		**0.097**	**0.576**	**0.133**	**0.002**				**0.558**	**0.813**	**0.089**	**0.018**			**0.615**	**11, 16**	**0.004**
11	**0.002**	**0.115**			**0.006**	**0.048**		**0.485**	**0.271**	**0.041**					**0.297**	**0.029**	**0.134**	**0.580**	**0.006**	**0.002**	**0.241**	**11, 17**	**0.002**
12		**0.703**			**0.110**	**0.004**		**0.352**	**0.049**	**0.002**	**0.016**				**0.066**	**0.065**	**0.060**	**0.310**	**0.004**	**0.035**	**0.069**	**11, 18**	**0.002**
13	**0.008**	**0.160**			**0.634**			**0.050**	**0.002**		**0.729**	**0.008**			**0.008**	**0.049**	**0.009**	**0.078**	**0.019**	**0.346**	**0.002**	**11, 19**	**0.008**
14	**0.053**	**0.006**			**0.185**			**0.007**			**0.191**	**0.029**			**0.002**	**0.022**		**0.012**	**0.132**	**0.394**		**12, 14**	**0.017**
15	**0.114**				**0.053**						**0.063**	**0.152**	**0.002**			**0.002**		**0.002**	**0.202**	**0.154**		**12, 15**	**0.027**
16	**0.262**	**0.002**	**0.004**		**0.008**							**0.219**	**0.002**			**0.002**			**0.178**	**0.061**		**12, 16**	**0.016**
17	**0.421**		**0.029**									**0.345**							**0.268**	**0.006**		**12, 17**	**0.012**
18	**0.115**		**0.197**				**0.002**					**0.162**	**0.004**						**0.168**	**0.002**		**12, 18**	**0.016**
19	**0.017**		**0.634**				**0.010**					**0.054**	**0.081**						**0.019**			**12, 19**	**0.008**
20	**0.008**		**0.106**				**0.060**					**0.019**	**0.518**	**0.021**					**0.004**			**13, 14**	**0.019**
21			**0.029**				**0.156**					**0.010**	**0.178**	**0.095**								**13, 15**	**0.038**
22			**0.002**				**0.190**					**0.002**	**0.157**	**0.536**								**13, 16**	**0.037**
23							**0.202**						**0.032**	**0.265**								**13, 17**	**0.056**
24				**0.008**			**0.147**						**0.017**	**0.078**								**13, 18**	**0.137**

Table 1 Alleles frequency of PowerPlex® Y23 System loci for a sample of Iraqi Arab population *(Continued)*

Allele/ *n*	DYS 576	DYS 389 I	DYS 448	DYS 389 II	DYS 19	DYS 391	DYS 481	DYS 549	DYS 533	DYS 438	DYS 437	DYS 570	DYS 635	DYS 390	DYS 439	DYS 392	DYS 643	DYS 393	DYS 458	DYS 456	YGA TAH4	DYS385 a, b haplotype frequency	
25				**0.006**			**0.185**						**0.004**	**0.006**								**13, 19**	**0.133**
26				**0.010**			**0.037**						**0.002**									**13, 20**	**0.016**
27				**0.071**			**0.004**						**0.002**									**13, 21**	**0.006**
28				**0.284**			**0.008**															**14, 15**	**0.017**
29				**0.442**																		**14, 16**	**0.061**
30				**0.157**																		**14, 17**	**0.011**
31				**0.023**																		**14, 18**	**0.025**
32																	**0.002**					**14, 19**	**0.01**
																						14, 21	**0.006**
																						14, 22	**0.002**
																						15, 16	**0.027**
																						15, 17	**0.004**
																						15, 19	**0.004**
																						15, 20	**0.002**
																						16, 17	**0.035**
																						16, 18	**0.025**
																						16, 19	**0.004**
																						17, 18	**0.012**
																						17, 19	**0.002**

Table 1 Alleles frequency of PowerPlex® Y23 System loci for a sample of Iraqi Arab population *(Continued)*

Allele/ *n*	DYS 576	DYS 389 I	DYS 448	DYS 389 II	DYS 19	DYS 391	DYS 481	DYS 549	DYS 533	DYS 438	DYS 437	DYS 570	DYS 635	DYS 390	DYS 439	DYS 392	DYS 643	DYS 393	DYS 458	DYS 456	YGA TAH4	DYS385 a, b haplotype frequency	
																						18, 19	**0.027**
																						9, 16	**0.004**
																						11	**0.006**
																						12	**0.009**
																						13	**0.016**
																						14	**0.014**
																						15	**0.013**
																						16	**0.008**
																						17	**0.004**
																						18	**0.004**
																						19	**0.008**

Table 2 Haploid Diversity by locus using PowerPlex® Y23 System

Haploid diversity by locus						
Locus	***N***	Na	Ne	I	h	uh
DYS576	1032	9	3.637	1.526	0.725	0.726
DYS389 I	1032	8	1.873	0.903	0.466	0.467
DYS448	1032	7	2.207	1.086	0.547	0.547
DYS389 II	1032	8	3.266	1.396	0.694	0.694
DYS19	1032	8	2.217	1.093	0.549	0.550
DYS391	1032	6	2.321	1.020	0.569	0.570
DYS481	1032	11	6.168	1.929	0.838	0.839
DYS549	1032	6	2.699	1.174	0.629	0.630
DYS533	1032	6	2.411	1.112	0.585	0.586
DYS438	1032	5	2.390	1.067	0.582	0.582
DYS437	1032	5	1.749	0.798	0.428	0.429
DYS570	1032	10	4.533	1.713	0.779	0.780
DYS635	1032	12	3.001	1.415	0.667	0.667
DYS390	1032	6	2.684	1.220	0.627	0.628
DYS439	1032	6	2.447	1.101	0.591	0.592
DYS392	1032	9	1.494	0.783	0.331	0.331
DYS643	1032	8	3.492	1.491	0.714	0.714
DYS393	1032	6	2.275	1.015	0.560	0.561
DYS458	1032	10	5.243	1.776	0.809	0.810
DYS456	1032	8	3.290	1.364	0.696	0.697
YGATAH4	1032	6	2.243	1.053	0.554	0.555
DYS385a,b	1032	14	7.546	2.175	0.791	0.791
Mean	**1032.000**	**7.619**	**2.935**	**1.240**	**0.616**	**0.617**
Standard error (SE)	**0.000**	**0.434**	**0.255**	**0.068**	**0.027**	**0.027**

Na no. of different alleles, *Ne* No. of effective alleles = 1/(sum pi^2), *I* Shannon's information, Index = − 1 × sum (pi × Ln (pi)), *h* diversity = 1−sum pi^2, *uh* unbiased diversity = (*N*/(*N*−1)) × *h*. Where pi is the frequency of the ith allele for the population, and sum pi^2 is the sum of the squared population allele frequencies

Table 3 Mean allelic patterns across population using PowerPlex® Y23 System

Mean allelic patterns across population	Mean values	SE
Na	**7.619**	**0.434**
Na Freq. > = 5%	**3.810**	**0.225**
Ne	**2.935**	**0.255**
I	**1.240**	**0.068**
No. private alleles	**7.619**	**0.434**
No. LComm alleles (< = 25%)	**0.000**	**0.000**
No. LComm alleles (< = 50%)	**0.000**	**0.000**
h	**0.616**	**0.027**
uh	**0.617**	**0.027**

h diversity = 1−sum pi^2, *uh* unbiased diversity = (*N*/(*N*−1)) × h. Where pi is the frequency of the ith allele for the population, and sum pi^2 is the sum of the squared population allele frequencies

Materials and methods

Sample collection

A total of 1032 blood samples were used in this study. Data were registered for each person, who were all male unrelated Iraqi nationals and identified as Iraqi Arabs. Participants were instructed about the purpose of the sampling, and interested volunteers have been enrolled, and informed consent was obtained from each subject before blood sampling. Blood samples were placed on DNA storage cards (Direct™ Classic Card-FTIZCO, USA).

DNA amplification and Y-STR typing

Samples run for DNA amplification using the direct polymerase chain reaction (PCR) method with PowerPlex® Y23 System (Promega Co., USA), according to the manufacturer's instructions using Gene Amp 9700 thermal cycler (Applied Biosystems). The amplified products were run on the ABI PRISM 3130xl Genetic Analyzer (Applied Biosystems), and the obtained data was analyzed using the Gene Mapper ID Analysis Software (Applied Biosystems, USA).

Statistical analyses

Frequency-based statistical analysis was calculated with the GenAlEx-6.5 Genetic Analysis software (Peakall & Smouse, 2006), while the haplotype diversity was calculated using the HapYdive software (http://portugene.com/hapydive.html) as mentioned previously (Parvathy et al., 2012). Analysis of molecular variance (AMOVA) was calculated by https://yhrd.org/amova/.yhrd.org.tools.

Results and discussion

Alleles frequency was calculated for each locus, in which observed alleles frequency of 23 Y chromosome STR loci in the Arab Iraqi population is summarized in Table 1. Haploid diversity by locus and the allelic patterns for haploid data as well as mean allelic patterns across population are shown in Tables 2 and 3, respectively.

Using PowerPlex® Y23 System markers, 478 different haplotypes were identified as the most frequent haplotype (detected 10 times) that is shown in Table 4. Eight profiles were unique, along with 435 profiles that found twice each, as well as 30 different profiles each repeated four times. Four different profiles were detected each within 6 individuals, which could be considered as accepted result due to the high inbreeding rate between family members or clan members across the country. DYS635 was the most polymorphic locus with 12 detected alleles, while DYS438 and DYS437 were the least polymorphic with 5 detected alleles each.

Table 4 Most frequent haplotype registered

Haplotype Code	Count	DYS576	DYS389 I	DYS448	DYS389 II	DYS19	DYS391	DYS481	DYS549	DYS533	DYS438	DYS437	DYS570	DYS635	DYS390	DYS439	DYS392	DYS643	DYS393	DYS458	DYS385	DYS385	DYS456	YGATAH4
200	10	17	14	20	31	14	10	22	13	11	10	14	20	20	24	12	11	13	13	16	18	19	15	12

One hundred and eighty-five alleles were detected at the 23 Y-STR loci in 1032 samples. Eighty-three samples (8.04%) were observed showing mono-allelic condition for the bi-allelic marker DYS385. A number of new unregistered genetic variations were recorded within DYS458 locus (using PowerPlex® Y23 System), and each sample was re-amplified and re-analyzed for confirmation; variants included (16.2:6 times, 17.2:29 times, 18.2:196 times, 19.2:143 times, 20.2:14 times, and 21.2:4 times), and these variants were previously detected by using AmpFlSTR® Yfiler™ (Applied Biosystems) and registered in STRBase (https://strbase.nist.gov/var_DYS458.htm).

Genetic diversity for each 23 Y-STR loci (h) ranged from 0.331 for DYS392 (which was the least informative locus) to 0.838 for DYS481(the most informative locus) in mono-allelic markers, and the average gene diversity was 0.616 ± 0.027; this result is similar to other neighboring populations that showed low genetic diversity due to several causes (Marchi et al., 2017; Palstra et al., 2015). The decrease in loci genetic diversity has been related to inbreeding (that is common in most of the Iraqi provinces) which at population level decreases genetic diversity eventually (Charlesworth, 2003). The firm paternal inheritance of Y-STRs as well as emigration influences the Y-haplotype diversity (Gusmão et al., 2006; Wang & Li, 2013; Singh et al., 2018). The lower effective number of Y-chromosomes in a given population indicates that Y-haplotypes/haplo groups might have a higher variation between populations than markers observed on X chromosomes or autosomes (Domingues et al., 2007). The haplotypes of Iraqi haplotypes were compared with the haplotypes of neighboring populations, according to multi-dimensional scaling plot (MDS) for genetic distance (Fig. 1), which indicates that Iraqi population are genetically closer to Lebanon and Kuwait than to the United Arab Emirates populations. Similarly, Rst values for the pairs of Iraqi and Kuwait and Lebanon populations showed much closer genetic distances than pairs of Iraqi and United Arab Emirates population (Table 5).

The loci included in the PowerPlex® Y23 System are suitable for forensic DNA fingerprinting casework, population genetic analysis, and anthropological purposes in the Iraqi population. This is the first genetic report using PowerPlex® Y23 for the Iraqi population. These markers showed a high degree of polymorphism, and these results are in concordance with previous studies regarding PowerPlex® Y23 (Purps et al., 2014). Data were registered at YHRD under accession number YA004665.

Conclusion

The present study established the genetic information obtained by using the PowerPlex® Y23 System for the

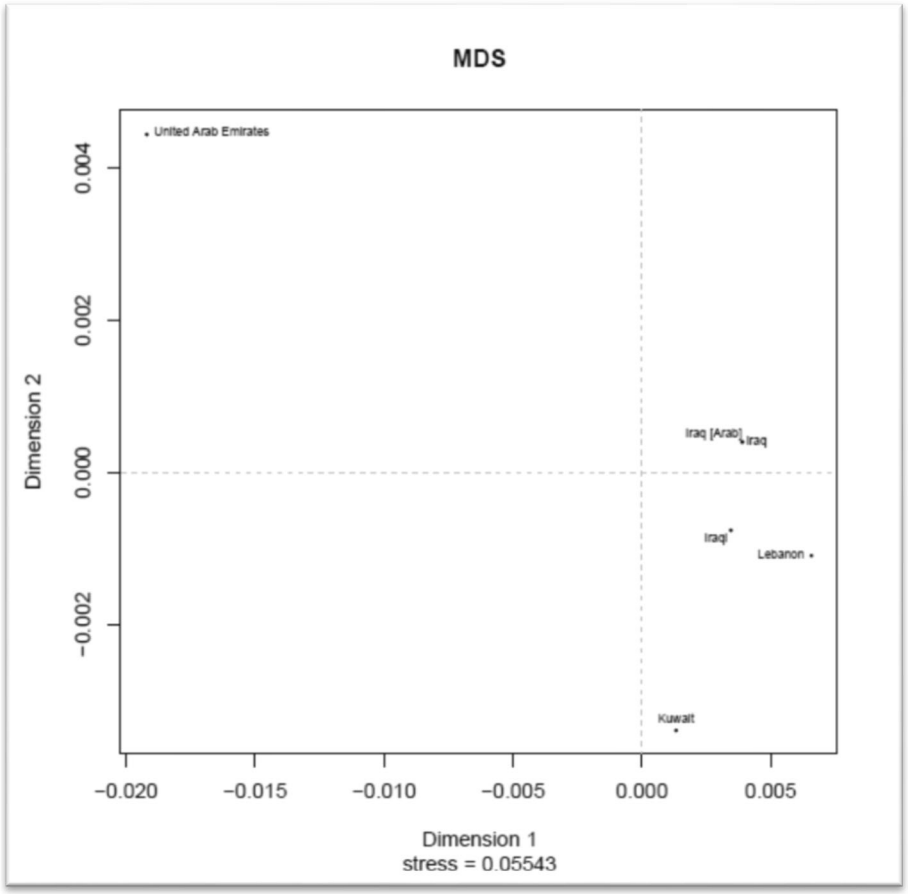

Fig. 1 Multi-dimensional scaling plot (MDS) based on pairwise Rst genetic distance values for Y-STR haplotypes for Iraqi haplotypes (current study) and other registered haplotypes of Iraqi population along with Lebanon, Kuwait, and United Arab Emirates populations

Table 5 Rst values for Iraqi and neighboring populations

Population	"Iraqi" studied group	"Iraq [Arab]"	"Iraq"	"Kuwait"	"Lebanon"	"United Arab Emirates"
"Iraqi" studied group	–	− 1.0000	− 1.0000	− 1.0000	− 1.0000	− 1.0000
"Iraq [Arab]"	0.0013	–	− 1.0000	− 1.0000	− 1.0000	− 1.0000
"Iraq"	0.0007	− 0.0013	–	− 1.0000	− 1.0000	− 1.0000
"Kuwait"	0.0017	0.0091	0.0064	–	− 1.0000	− 1.0000
"Lebanon"	0.0022	0.0061	0.0025	0.0073	–	− 1.0000
"United Arab Emirates"	0.0209	0.0212	0.0286	0.0227	0.0266	–

Iraqi population and also created a database of 23 Y STR markers in this population. It also showed that the genetic origin of the Iraqi population has a unique combination from local citizens and the ancestors that migrated to Iraq from different regions and settled in it throughout history. Current studies are ongoing to increase the size of the samples and to make the regional comparison within the country as well as compare the obtained results with other populations Y- STR data.

Abbreviations

STRs: Short tandem repeats; MLD: Medico-Legal Directorate; PCR: Polymerase Chain Reaction; HD: Haplotype genetic diversity; h: Diversity

Acknowledgements

We gratefully acknowledge the technical support of DNA laboratory in Paternity and Kinship Division/Medico-Legal Directorate, Baghdad/Iraq.

Authors' contributions

Dr. Hanan K. Mahmood, Head of Paternity and Kinship Division, conceived of the study and participated in its design and coordination. Dr. Nadia F. Salman, corresponding author, expert in DNA fingerprinting, performed the literature search, statistical analysis, combined results of the analysis, and drafted the manuscript. Khalifa M. Salih, DNA analysis specialist, performed DNA genotyping and prepared data for statistical analysis. Dhurgham H. Hasan, DNA analysis specialist, performed DNA genotyping and prepared data for statistical analysis. Dr. Mohammed M. Al-Zubaidi reviewed the results and statistics and recording them in a YHRD database. The authors read and approved the final manuscript.

Funding

Not applicable

Availability of data and materials

Please contact the authors for data requests.

Ethics approval and consent to participate

All participants had been informed about the purpose of the sampling, and proper individuals have been enrolled.
This research was conducted based on Article 2 of the Iraqi Forensic Medicine Law of 2013.

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Medico-legal Directorate, Baghdad, Iraq. [2]Forensic DNA center Research and Training, Al-Nahrain University, Baghdad, Iraq.

Received: 23 May 2019 Accepted: 16 April 2020
Published online: 20 May 2020

References

Charlesworth, D. Efects of inbreeding on the genetic diversity of populations. Philosophical Transactions of the Royal Society of London B: Biological Sciences 358, 1051–1070 (2003).
Domingues PM, Gusmão L, da Silva DA, Amorim A, Pereira RW, de Carvalho EF (2007) Sub-Saharan Africa descendents in Rio de Janeiro (Brazil): population and mutational data for 12 Y-STR loci. Int J Legal Med 121(3):238–241
Gusmão L, Butler JM, Carracedo A, Gill P, Kayser M, Mayr WR et al (2006) DNA Commission of the International Society of Forensic Genetics (ISFG): an update of the recommendations on the use of Y-STRs in forensic analysis. Int J Legal Med 120(4):191–200
Jain T, Shrivastava P, Bansal DD, Dash HR, Trivedi VB (2016) PowerPlex Y23 system: a fast, sensitive and reliable Y-STR multiplex system for forensic and population genetic purpose. J Mol Biomark Diagn 7(281):2. https://doi.org/10.4172/2155-9929.1000287
Marchi N, Hegay T, Mennecier P, Georges M, Laurent R, Whitten M et al (2017) Sex-specific genetic diversity is shaped by cultural factors in Inner Asian human populations. Am J Phys Anthropol 162(4):627–640
Palstra FP, Heyer E, Austerlitz F (2015) Statistical inference on genetic data reveals the complex demographic history of human populations in Central Asia. Molecular biology and evolution 32(6):1411–1424
Parvathy SN, Geetha A, Jagannath C (2012) Haplotype analysis of the polymorphic 17 YSTR markers in Kerala nontribal populations. Mol Biol Rep 39(6):7049–7059
Peakall ROD, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6(1):288–295
Purps, J., Siegert, S., Willuweit, S., Nagy, M., Alves, C., Salazar, R., ... & Ayub, Q. (2014). A global analysis of Y-chromosomal haplotype diversity for 23 STR loci. Forensic Science International: Genetics, 12, 12-23.
Roewer L (2009) Y chromosome STR typing in crime casework. Forensic Sci Med Pathol 5(2):77–84
Roewer L, Willuweit S, Krüger C, Nagy M, Rychkov S, Morozowa I et al (2008) Analysis of Y chromosome STR haplotypes in the European part of Russia reveals high diversities but non-significant genetic distances between populations. Int J Med 122(3):219–223
Singh M, Sarkar A, Nandineni MR (2018) A comprehensive portrait of Y-STR diversity of Indian populations and comparison with 129 worldwide populations. Sci Rep 8(1):15421
Wang CC, Li H (2013) Inferring human history in East Asia from Y chromosomes. Invest Genet 4(1):11

Publisher's Note