COMPARISON STUDY OF IRAQI POPULATION IN MIDDLE-EUPHRATES WITH HAPLOTYPE REFERENCE DATABASE (YHRD) AND STUDY THEIR HAPLOGROUPS AND PHYLOGENETIC NETWORK

Mohammed zuhair Naji and Ali Hmood Al-Saadi

Biology department of science in University of Babylon, Babylon, Iraq.

Abstract

In this study, a sample of 144 persons lived in middle-Euphrates of Iraq population was analyzed using 5 Y-chromosome short tandem repeat (STR) polymorphisms. The results Comparison with (YHRD) showed that there are 44 match with 307,169 Haplotypes, the highest match was in haplotype H13 it record 4707 match with 83 countries. In study Analysis of AMOVA and MDS plots show the Kuwait (from Arabia population) and Japan (over world population) were more closed to Iraqi population with same cluster plot. The sixty two haplotypes sorted in six main haplogroups (E, G, I, J, Q and R). haplogroup J was the biggest one in compare with other haplogroup and include all Hashemites sub-population.

Key words: YHRD, Hashemites, Haplogroups, Phylogenetic network.

Introduction

The main source of information about Y-STR is the website YHRD.org. In practice, the frequency of the Y-STR profile in question can be estimated by referencing databases with large numbers of Y-STR profiles and counting the number of matching profiles within the population of interest. This provides haplotype. One of these databases is the Y-Chromosome Haplotype Reference Database (YHRD) (Willuweit and Roewer, 2015). Following the guidelines of the International Society of Forensic Genetics ISFG) for the publication of genetic population data, submissions to YHRD are required to consist of at least 17 Y-STR markers (including the eight-marker minimal Y-haplotype: DYS19, DYS389I, DYS38911, DYS390, DYS391, DYS392, DYS393, DYS319) 5 for 200 or more individuals from a population.44 As of 2015, YHRD contained more than 84,000 Y-profiles of 17 markers or greater in 572 populations (Abbai *et al.*, 2019). Polymorphisms of any type can be combined into groups of two or more tightly clustered on the chromosome-haplotypes. If two or more polymorphisms exist in close proximity in the DNA (e.g., up to a few thousands of bases apart), recombination

will be extremely infrequent and they can be studied jointly as a haplotype (Mohapatra *et al.*, 2019). A haplotype is the haploid genotype of the alleles at polymorphisms along a rather short stretch of the chromosome. When the recombination rate between any two markers is low, the history of the mutational events and random genetic drift result in a nonrandom occurrence of the alleles on chromosomes in the populations. This non-randomness is called linkage disequilibrium (LD) (Han *et al.*, 2019). Because recombination is very infrequent in molecularly very short regions, each of the possible haplotypes can be considered for analytic purposes as a distinct allele exactly analogous to alleles at any other polymorphism (SNP, Indel, etc.) (Stephens *et al.*, 2001)

Materials and Methods

Collection of samples

Samples have been collected from 144 person, they separated in two groups : Hashemites and, Commoners (72 pesons for each). These samples was collect from middle Euphrates aria (Baghadad, Hillah, Karbala, Najaf, Kut and Diwaniyah), the work done in the Biotechnology and Genetic Engineering Laboratory, Department of Biology, College of Science, University of Babylon.

Author for correspondence : E-mail: biomanearthlink@gmail.com

DNA extraction methods

Genomic DNA of blood were extract using DNA extraction kit (FAVORGEN).

Amplification Y-chromosome STR

Five Y-STR marker (Dys19, Dys390, Dys391, Dys392, Dys393) were use in this study on Iraqi population. PCR amplification was done using conventional thermocycler (Biometra - Germany)

Agarose Gel Electrophoresis

The agarose gel electrophoresis was perfomed according to the method of Robinson and Lafleche, (2000). This technique was used to detect genomic DNA extracts and PCR products.

Cycle sequencing and Sequence Analysis

The DNA Sequencing of the PCR products was done using the BigDye TM Terminator (reaction mixture consisted of: 5 uL purified products, 2 uL Big Dye Terminator, 2 ul. sequencing buffer and 1 uL reverse primer) Utilizing POP-7 polymer (Applied Biosystems) of lot number 1206453. The separation of the cycle sequencing products and detection were carried out by using the ABI 3730XL DNA Analyzer, cap array size 96, cap array length 50. The reference sequence described by (Anderson, 1981) was compared to the data observed.

AMOVA analysis

Analysis of Molecular Variance (AMOVA) is a method of estimating population differentiation directly from mutational differences in molecular data as well as from gene frequencies and testing hypotheses about such differentiation (Excoffier *et al*., 1992). It works on such data to create a distance matrix (Fst and Rst and other analogues) between samples in order to measure the genetic structure of the population from which the samples are drawn. In statistical terms, AMOVA is a testing procedure based on permutation analysis and involves few assumptions about the statistical properties of the data. It involves a hierarchical analysis of variance in that it separates and tests levels of genetic diversity:

• Diversity among groups of populations

• Diversity among the populations within groups

• Diversity among the individuals within a population

F-Statisties

Wright' s F statistics (Wright, 1978) are measures of the difference between the mean heterozygosity among the subdivisions in a population and the potential frequency of heterozygotes if all members of the population mixed randomly (Hartl and Clark, 1997). Here F stands for fixation index, fixation being increased homozygosity resulting from inbreeding. Population subdivision results in the loss of genetic variation (measured by heterozygosity) within small subpopulations where genetic drift acts. This means that population subdivision would result in decreased heterozygosity relative to that expected under random mating as if the whole population were a single breeding unit (Weir and Cockerham, 1984). Fst measures the effect of population subdivision, which is the reduction in heterozygosity in a subpopulation due to genetic drift (fixation of alleles). It is calculated as using the subpopulation (average) heterozygosity and a total population expected heterozygosityFst ranges between 0no subdivision (random mating occurring, no genetic divergence within the population) and 1-complete isolation (extreme subdivision). However, negative values are also allowed (Balloux and Lugon-Moulin, 2002).

Results and Discussion

Y-STR Haplotype Reference Database (YHRD)

The YHRD is one of the most important Y-STR data analysis sites in the world Populations that chosen and compared to haplotype (www.yhrd.org). All 62 Y-STR haplotypes present study have been studied (Appendix II). All types of haplotype present study have been entered into this site for the purpose of data analysis and knowledge of the degree of genetic relationship with neighboring countries and the rest of the world. The results showed (Table 1) that there are 44 match with 307,169 Haplotypes, the highest match was in haplotype H13 it record 4707 match with 83 countries and the less matches were in six haplotypes (H5, H17, H21, H25, H31, H47) which record 1 match, in addition to 18 haplotypes discovered not recorded to data base yet, it will record by this study (H4, H6, H15, H18, H19, H22, H23, H26, H43, H44, H45, H46, H54, H55, H56, H57, H58, H62).

Of these results it is evident that the genetic relationship among Iraq and the surrounding countries is very large despite the database of these countries is small and this indicates that these countries descended from one origin and the same as referred to Reasons for convergence (Rodinson, 1981; Mansfield, 1992; Hitti, 2002).

Genetic relationships between populations

F statistic (Fst) values were calculated by Analysis of Molecular Variance (AMOVA) using information on comparative populations from the YHRD database (10,000 permutations) and Multidimensional Scaling (MDS) plots were constructed based on Kriskal non-metric MDS algorithm (Kruskal, 1964) using the YHRD online program available at the YHRD website(http://www.yhrd.org). Haplotype data generated in the present study was compared to the haplotype distribution from 20 countries. These countries are divided into Arab

Table 1: Matches Haplotypes with Y-HRD Data base.

Sample No.	Haplotype No.	No. of matches in 307,169 Haplotypes	Approximate Haplotypes match	Not typed for Y-SNP ratio
1	H1	5	1 match in 61,434 Haplotypes	100%
2	H10	1,389	1 match in 221 Haplotypes	78.1%
3	H11	127	1 match in 2,419 Haplotypes	92.1%
4	H12	7	1 match in 43,881 Haplotypes	100%
5	H13	4,707	1 match in 65 Haplotypes	91.5%
6	H14	37	1 match in 8,302 Haplotypes	94.6%
7	H15	0	-	-
8	H16	4	1 match in 76,792 Haplotypes	100%
9	H17	1	1 match in 307,169 Haplotypes	100%
10	H18	0	-	-
11	H19	0	-	-
12	H2	321	1 match in 957 Haplotypes	95.3%
13	H20	10	1 match in 30,717 Haplotypes	100%
14	H21	1	1 match in 307,169 Haplotypes	100%
15	H22	20	1 match in 15,358 Haplotypes	95.0%
16	H23	0	-	-
17	H24	0	-	-
18	H25	1	1 match in 307,169 Haplotypes	100%
19	H26	0	-	-
20	H27	6	1 match in 51,195 Haplotypes	100%
21	H28	13	1 match in 23,628 Haplotypes	100%
22	H29	170	1 match in 1,807 Haplotypes	99.4%
23	H3	821	1 match in 374 Haplotypes	92.2%
24	H30	2	1 match in 153,585 Haplotypes	100%
25	H31	1	1 match in 307,169 Haplotypes	100%
26	H32	33	1 match in 9,308 Haplotypes	97.0%
27	H33	3	1 match in 102,390 Haplotypes	100%
28	H34	88	1 match in 3,491 Haplotypes	96.6%
29	H35	1,858	1 match in 165 Haplotypes	92.1%
30	H36	429	1 match in 716 Haplotypes	93.7%
31	H37	248	1 match in 1,239 Haplotypes	94.0%
32	H38	19	1 match in 16,167 Haplotypes	94.7%
33	H39	33	1 match in 9,308 Haplotypes	97.0%
34	H4	0	-	-
35	H40	971	1 match in 316 Haplotypes	97.2%
36	H41	608	1 match in 505 Haplotypes	94.4%
37	H42	326	1 match in 942 Haplotypes	93.9%
38	H43	0	-	-
39	H44	0	-	-
40	H45	0	-	-
41	H46	0	-	-
42	H47	1	1 match in 307,169 Haplotypes	100%
43	H48	8	1 match in 38,396 Haplotypes	75.0%
44	H49	2	1 match in 153,585 Haplotypes	100%
45	H5	1	1 match in 307,169 Haplotypes	100%
46	H50	5	1 match in 61,434 Haplotypes	60.0%

Table 1 Continue...

countries and neighboring to Iraq and another division consists of the countries from Asia, Europe, Africa and America. Several methods have been used to score the variations and diversity of Y-STR in Iraq Arab male population. F statistic is one of the most used biometric methods and simply measures the lever of heterozygosity or differentiation between populations. Fst ranges from o to 1 where 0 indicating no differences in allele frequencies between two populations and 1 indicating that the two populations are fixed for alternate alleles. Fst test AMOVA table 2 shows the genetic relative between the Arabian population of present study with population from, Algeria, Iran, Egypt, Israel, Jordan, Kuwait, Lebanon, Saudi Arabia, Yemen and Turkey. The Iraq Arab population Y-STR haplotype shared relative genetic relationship with Israel Fst(0.2791), Lebanon Fst (0.2884), Egypt Fst (0.2460), Kuwait Fst (0.2314), Saudi Arabia Fst (0.2967), Yemen Fst (0.2852).

While the most distant genetic relationship countries from Iraq were Jordan Fst (0.3118), Iran Fst (0.3064), Turkey Fst (0.3531) from Asia and distant genetic the Arabs of Africa are Algeria Fst (0.3210).

The result MDS plot fig. 1 show that Iraq and Kuwait appear in same cluster plot. While Saudi Arabia, Yemen, Israel and Jordan locate in same cluster plot and finally, Iran, Turkey, Algeria, Lebanon and Egypt found together in the last Custerplot.

The genetic variation along the Y-chromosome has been shown to be affected by rapid genetic drift resulting in high levels of geographical differentiation of Y-haplotypes (Immel et al., 2004) Furthermore, the Y-chromosome contains the largest non-recombining section within the human genome, providing significant informative haplotypes for application in population genetic analyses (Underhill and Kivisild, 2007). Worldwide population data have been collected for

Table 1 Continue...

47	H51	86	1 match in 3,572 Haplotypes	87.2%
48	H52	73	1 match in 4,208 Haplotypes	93.2%
49	H53	17	1 match in 18,069 Haplotypes	100%
50	H54	0	-	-
51	H55	0	-	-
52	H56	0	-	-
53	H57	0	-	-
54	H58	0	-	-
55	H59	61	1 match in 5,036 Haplotypes	96.7%
56	H6	0	-	-
57	H60	3	1 match in 102,390 Haplotypes	100%
58	H61	13	1 match in 23,628 Haplotypes	100%
59	H62	0	-	-
60	H7	459	1 match in 669 Haplotypes	86.7%
61	H8	30	1 match in 10,239 Haplotypes	93.3%
62	H9	9	1 match in 34,130 Haplotypes	88.9%

Y- chromosome STR (Y-STR) haplotypes. Results have shown Y-STR haplotypes are region-specific, increasing their utility in population genetic studies, forensics and paternity testing (Triki-Fendri *et al*., 2010). The YHRD database provides high–resolution and large reference sample collections of word populations for genetic analysis

Fig. 1: MDS analysis plot based on population pair wise FST value haplotype data from Iraq and Arab countries and neighbor. (Stress value = 0.04404).

(Willuweit and Roewer, 2015). These results differed with the results of Tareq United Arab Emirates, (2017). Iraq was the second closest genetic relationship country in uie United Arab Emirates after Iran, In the same study, Irag had a relative genetic relationship in Egypt (0.0003). In another study in Sudia Arabia (Share, 2014), the result showed that Iraq was the closest genetic relationship from Saudi Arabia and then to Yemen. The reason for these differences between the haplotype the present study with neighboring countries is because of the size of the database used, as the larger the size of the database there was a possibility of different results obtained Due to the geographical location of the Middle East and bidirectional dispersals, the degree of ethnic diversity in the region has complicated population genetic analyses. Discrepancies between genetic analyses of the Middle East have been observed. For example, Triki-Fendri *et al*. (2010) used Y-STR data to support the relationship between the degree of genetic diversity and geographical location. Genetic relatedness was seen between the Kuwaiti population and Arabian Peninsula populations (Yemen, Saudi Arabia and the UAE) reflecting the close geographical location of the groups (Tadmouri *et al*., 2015). However, Roewer *et al*., (2019) highlighted the importance of studying individual populations or minority groups in close-proximity in the same region and country. As the present study found genetic differences between linguistic groups residing within Iran. This latter study highlights the importance in genetic analyses of considering not only the geographical location of countries but also factors such as cultural and religious isolation of populations.
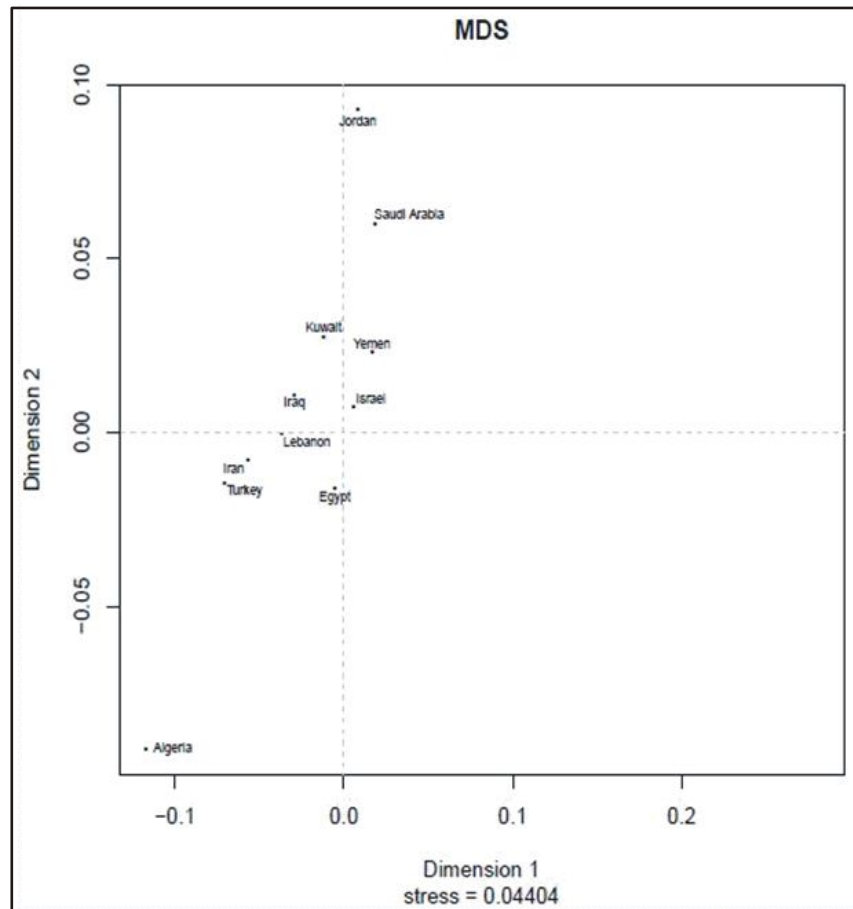
The haplotype data generated in the present study was compare to the haplotype distribution from the Countries of Asia and Europe populations. The resultant MDS plot (Fig. 2) and Fst test AMOVA table 3 shows the genetic relationship between the Iraq Arab population with Japan Fst (0.3277), Argentina Fst (0.4892),

Table 2: Fst test AMOVA for Iraq and Arab countries and neighbor.

Population	Iraq	Algeria	Iran	Egypt	Kuwait	Jordan	Israel	Lebanon	KSA	Turkey	Yemen
Iraq	-	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Algeria	0.3210	-	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Iran	0.3064	0.0786	-	0.0000	0.0000	0.0000	0.0000	0.0034	0.0000	0.7876	0.0000
Egypt	0.2460	0.1087	0.0657	-	0.0019	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Kuwait	0.2314	0.1716	0.1053	0.0150	-	0.0000	0.0000	0.0000	0.0000	0.0000	0.0017
Jordan	0.3118	0.3509	0.2366	0.1444	0.0951	-	0.0000	0.0000	0.0000	0.0000	0.0000
Israel	0.2791	0.1766	0.0839	0.0488	0.0217	0.0439	-	0.0000	0.0000	0.0000	0.0000
Lebanon	0.2884	0.0925	0.0204	0.0241	0.0328	0.1099	0.0302	-	0.0000	0.0000	0.0000
KSA	0.2967	0.2889	0.1992	0.1054	0.0479	0.0320	0.0241	0.1009	-	0.0000	0.0000
Turkey	0.3531	0.0825	0.0027	0.0831	0.1146	0.1729	0.0861	0.0309	0.1753	-	0.0000
Yemen	0.2852	0.2126	0.1598	0.0492	0.0128	0.0746	0.0264	0.0563	0.0333	0.1414	-

Table 3: Fst test AMOVA for Iraq and Asia, Europe, Africa and America.

| Population	Iraq	Japan	Argentina	Australia	Brazil	China	Italy	South Africa	Spain	US	Zimbabwe
Iraq	-	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Japan	0.3277	-	0.0000	0.0000	0.0000	0.0000	0.0011	0.0000	0.0000	0.0000	0.0000
Argentina	0.4892	0.0828	-	0.0000	0.3850	0.0000	0.0179	0.0000	0.0000	0.0049	0.0000
Australia	0.5675	0.1658	0.0502	-	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Brasilia	0.4905	0.0741	0.0001	0.0484	-	0.0000	0.0215	0.0000	0.0000	0.0078	0.0000
China	0.4502	0.0790	-0.1244	0.1491	0.1184	-	0.0000	0.0000	0.0000	0.0000	0.0000
Italy	0.4166	0.0299	0.0102	0.0899	0.0103	0.0982	-	0.0000	0.0000	0.0021	0.0000
South Africa	0.4448	0.1670	0.3132	0.3220	0.3101	0.1384	0.2622	-	0.0000	0.0000	0.0000
Spain	0.6362	0.2277	0.0371	0.0244	0.0411	0.2181	0.0988	0.4406	-	0.0000	0.0000
US	0.4920	0.0830	0.0129	0.0296	0.0138	0.1316	0.0267	0.3042	0.0387	-	0.0000
Zimbabwe	0.4196	0.2932	0.5027	0.5680	0.4954	0.4456	0.4462	0.4787	0.6527	0.5303	-
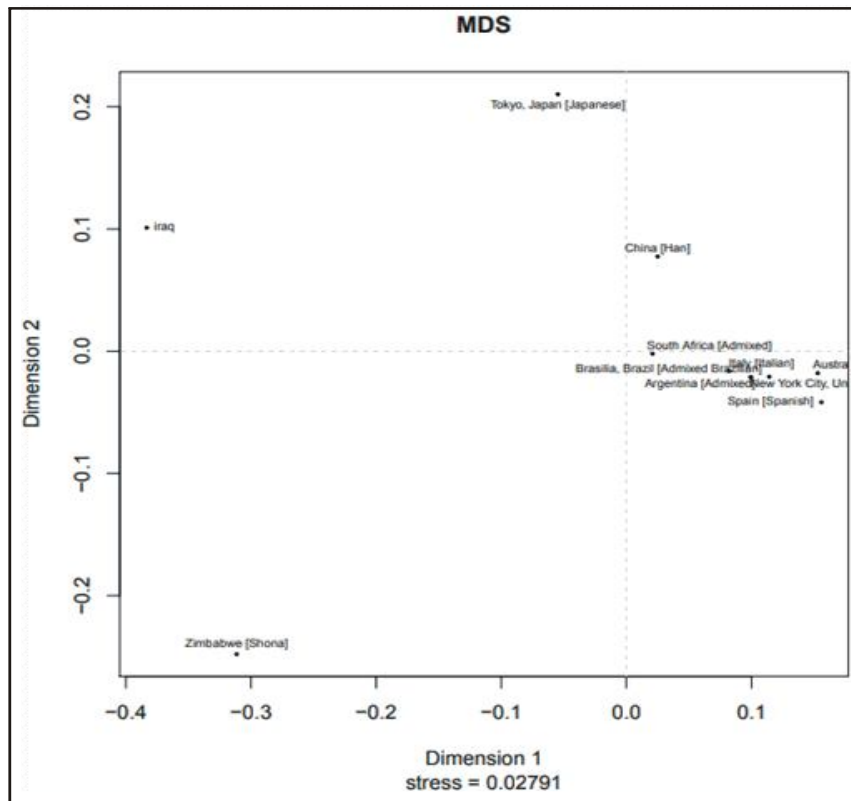
Fig. 2: MDS analysis plot based on population pair wise FST value haplotype data from Iraq and Asia, Europe, Africa and America. (Stress value = 0.02791).

Australia Fst (0.5675), Brasilia Fst (0.4905), China Fst (0.4502), Italy Fst (0.4166), South Africa Fst (0.4448), Spain Fst (0.6362), US Fst (0.4920), Zimbabwe Fst (0.4196).

The result MDS plot fig. 2 show that Iraq and Japan appear in same cluster plot. While all other counties located together in the other Custer plot.

The haplogroups of Iraqi population

In the study of molecular evolution, a haplogroup is a group of similar haplotypes that share a common ancestor with a single nucleotide polymorphism (SNP) mutation. Haplogroups pertain to deep ancestral origins dating back thousands of years. The most commonly studied human haplogroups depended on Y-DNA haplotypes or Y-SNP, both of which can be used to define genetic populations. Y-DNA is passed solely along the patrilineal line, from father to son. The Y-DNA may change by chance

Table 4: The haplogruops of Iraqi population.

Haplo group	Sub-Haplogroup	Haplo-type	Prob-ability
E	E1b1a V38> M329	H29	47.15
	E1b1b	H46	33.26
	E1b1b M123* (xM34)	H51	32.57
	E1b1b M123>M34> Z841>> Y4971> Y4972	H34	69.32
	E1b1b V13>>Z5017>> Z16988	H31	50.54
	E1b1b V13>>Z5018> Y145455	H28	62.03
	E1b1b V257>M81>> PF2546> CTS12227	H32	34.72
	E1b1b V257>M81>> PF2546> MZ2	H30	58.16
	E2 M75	H36	62.87
	E2 M75	H49	60.07
G	G1 M342	H37	53.34
I	I2a2a M223>L701> L699> S12195	H40	49.07
J	J1a>> L620>> FGC6064> M365	H23	26.18
	J1a>> P58>> FGC11>> ZS9061	H19	35.67
	J1a>> P58>> L860	H25	84.47
	J1a>> P58>> S4924 >> L818	H11	73.54
	J1a>> P58>> S4924 >> L818	H15	99.8
	J1a>> P58>> S4924 >> L818	H24	94.13
	J1a>> P58>> S4924 >> L818	H27	75.9
	J1a2 > ZS6591	H3	37.88
	J1a2 > ZS6591	H14	23.76
	J1a2 > ZS6591	H17	95.11
	J1a2 > ZS6591	H21	41.9
	J1a2 > ZS6591	H22	26.22
	J1a2a2 PF7264> P56	H8	51.11
	J1a2a2 PF7264> PF7263	H4	76.82
	J1a2a2 PF7264> PF7263	H5	94.55
	J1a2a2 PF7264> PF7263	H6	37.76
	J1a2a2 PF7264> PF7263	H7	78.51
	J1a2a2 PF7264> PF7263	H9	98.29
	J1a2a2 PF7264> PF7263	H10	65.92
	J1a2a2 PF7264> PF7263	H12	50.59
	J1a2a2 PF7264> PF7263	H18	33.25
	J1a2a2 PF7264> PF7263	H20	66.94
	J1a3 Z1828> BY69J1a3 Z1828> BY69	H26	47.19
	J1a3 Z1828> Z1842> Z18436> CTS1460> BY70	H16	36.22
	J1a3 Z1828> Z1842> ZS3089	H1	58.43
	J1a3 Z1828> Z1842> ZS3089	H2	14.48
	J1b F4306>> F1614>> Z2223	H13	30.45
	J2a1 M319> S18579	H33	29.06
	J2a1 PF5191>> PF5177	H50	32.02
	J2a1 PF5191>> S19231	H38	28.73
	J2a1 Z6063>> Y28259	H59	59.11
	J2a1 Z6063>> Y28259	H61	36.96
	J2a1 Z6063>> Y28259	H62	37.09
	J2a1 Z7700> A5360	H45	78.38
	J2a1 Z7700> A5360	H53	65.44
	J2a1 Z7700> A5360	H54	47.92

Table 4 Continue...

mutation at each generation. So, this study depended on results of Y-DNA haplotypes by aid of web site Haplogroup Predictor (https://www.nevgen.org/) to determine Iraqi haplogroup. Table 4 shows the main and sub-haplogroups of haplotypes with their maximum Probability from web site Haplogroup Predictor.

The sixty two haplotypes sorted in six main haplogroups (E, G, I, J, Q and R). haplogroup J was the biggest one, it share 72% of total haplogroups (Fig. 3) with 45 haplotypes distributed in seven sub-haplogroups). Which be note all the Hashemities sub-population (H1-H27) employs in sub-haplogroup J1, while people of haplogroup J2 belongs to commons sub-populations.

The second biggest haplogroup was E which share 16% of total haplogroups. It has 10 haplotypes branched to three sub-haplogroups, all haplotypes persons of this haplogroup was commons sub-populations.

The the ridhaplo group is R, its share 6% of all haplo groups with 4 haplo types distributed in two sub-haplo groups. All haplo group's peoples was commons sub-populations.

Finally, the lowest reminds haplo groups was G, I and Q each of one share 2% of total haplogroups with 1 haplotypes distributed in one sub-haplogroups. And also each peoples was belongs to commons sub-populations.

Phylogenetic network of Iraqi haplogroups

The phylogenetic network of haplogroups has been designed and drawn automatically by help of program NETWORK 5.0.1.1 ©2005. Fig. 4 show the phylogenetic network of 62 haplotypes according to its haplogroups and by sorting factor of program the haplogroup J1 (haplotype H1-H27) of hashemites sub-population selected as recognized by green node. Its show highly interaction and complex association in the core of network in compare with other haplogroups and it share some links with neighbor trunk of network.

Table 4 Continue...

	J2a1 Z7700 > A5360	H55	69.74
	J2a1 Z7700 > A5360	H58	72.84
	J2a1 Z7700 > A5360	H60	42.37
	J2a1 Z7700 > Z44288	H57	88.03
	J2a2 PF5008	H39	49.52
	J2a2 PF5008 >> Y31569	H43	73.2
	J2a2 PF5008 >> Z39726	H48	46.04
	J2a2 PF5008 >> Z39726	H56	50.23
	J2b1 M205 >> CTS1969	H52	48.52
Q	Q M346 >> L330	H44	96.71
R	R1a YP1051	H41	37.89
	R1a YP1051	H47	78.91
	R1a Z93 > Z94 > L657 > Y5	H35	42.77
	R1b V88	H42	22.63

In compare these results with other studies in Iraq and Arabic area it showed a high converge and agree with them. In study of geographical and social structure in the paternal lineages of Saudi Arabia for Khubrani *et al.*, (2018), found that 71% of its population belong to haplogroup j and the other was A, B, E, G, H, L, Q and R haplogroups.

In other study on Iraqi population for Al-Zahery *et al.*, 2011, on marsh peoples. She found that about 84.6% of this population have J haplogroup and share other haplogoups with different rate such as E, R, L, G and Q haplogroups.

So, many studies prove that the origin of Arab belong to J haplogroup in the first matter with rate up to 60% and then E and R haplogroups, with low rate of other rare haplogroups. (Abu-Amero *et al.*, 2009); (Mohammad *et al.*, 2009).

Acknowledgements

Conclusion

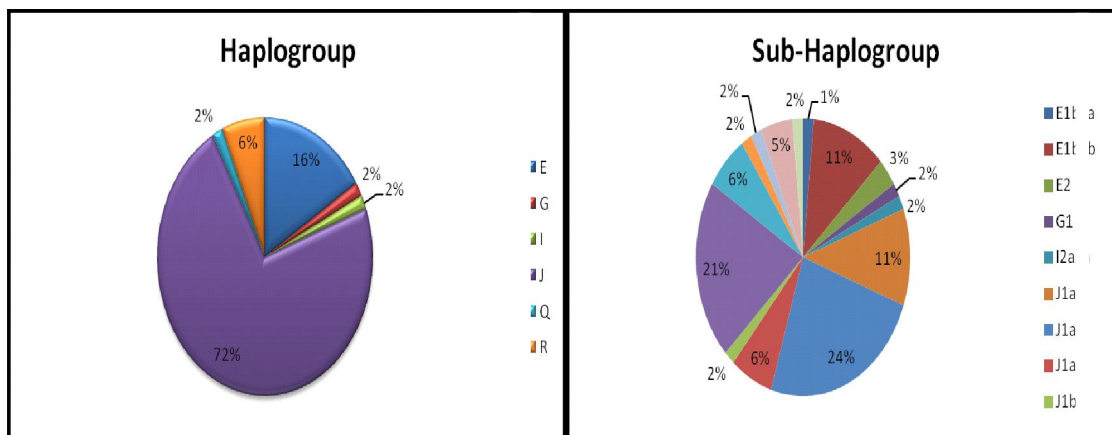In the study of comparison Iraqi population with (YHRD) its matched

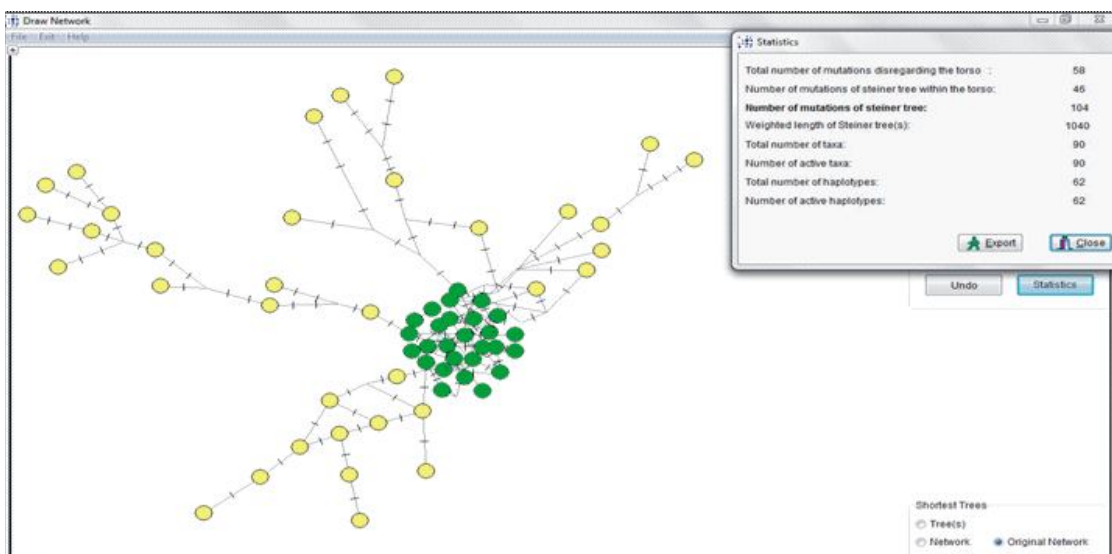**Fig. 3:** The main Haplogroups and sub-Haplogroups rating share of Iraqi population.

Fig. 4: The phylogenetic network of Iraqi haplogroups.

with 44 Haplotypes. Analysis of AMOVA and MDS plots of Iraq and Arabia population show Kuwait was more closed to Iraqi populationas well as Japan in compare with global populations. The haplogroupstudy show allHashemites sub-population was belongs to sub-haplogroup J2.

References

Abbai, R., V. K. Singh, V.V. Nachimuthu, P. Sinha, R. Selvaraj, A.K. Vipparla and A. Kumar (2019). Haplotype analysis of key genes governing grain yield and quality traits across 3K RG panel reveals scope for the development of tailor made rice with enhanced genetic gains. *Plant biotechnology journal*, **17(8):** 1612-1622.

Abu-Amero, K.K., A. Hellani, A.M. González, J.M. Larruga, V.M. Cabrera and P.A. Underhill (2009). Saudi Arabian Y-Chromosome diversity and its relationship with nearby regions. *B.M.C. genetics*, **10(1):** 59.

Al-Zahery, N., M. Pala, V. Battaglia, V. Grugni, M.A. Hamod, B.H. Kashani and O. Semino (2011). In search of the genetic footprints of Sumerians: a survey of Y-chromosome and mtDNA variation in the Marsh Arabs of Iraq. *BMC evolutionary biology*, **11(1):** 1-16.

Anderson, S., A.T. Bankier, B.G. Barrell, M.H. de Bruijn, A.R. Coulson, J. Drouin an P.H. Schreier (1981). Sequence and organization of the human mitochondrial genome. *Nature*, **290(5806):** 457-465.

Balloux, F. and N. Lugon Moulin (2002). The estimation of population differentiation with microsatellite markers. *Molecular ecology*, **11(2):** 155-165.

Excoffier, L., P.E. Smouse and J.M. Quattro (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131(2):** 479-491.

Han, Y., G. He, S. Gong, J. Chen, Z. Jiang and P. Chen (2019). Genetic diversity and haplotype analysis of Guizhou Miao identified with 19 X-chromosomal short tandem repeats. *International Journal of legal medicine*, **133(1):** 99-101.

Hartl, D.L. and A.G. Clark (1997). Principles of population genetics (Vol. 116). Sunderland, MA: Sinauer associates.

Immel, U.D., M. Krawczak, J. Udolph, A. Richter, H. Rodig, M. Kleiber and M. Klintschar (2006). Y-chromosomal STR haplotype analysis reveals surname-associated strata in the East-German population. *Eur. J. Hum. Genet.*, **14:** 577-582

Khubrani, Y.M., J.H. Wetton and M.A. Jobling (2018). Extensive geographical and social structure in the paternal lineages of Saudi Arabia revealed by analysis of 27 Y-STRs.

Forensic Science International: *Genetics*, **33:** 98-105.

Mohammad, T., Y. Xue, M. Evison and C. Tyler-Smith (2009). Genetic structure of nomadic Bedouin from Kuwait. *Heredity*, **103(5):** 425-433.

Mohapatra, B.K., K. Chauhan, P. Shrivastava, A. Sharma, S. Dagar and K. Kaitholia (2019). Haplotype data for 17 Y-STR loci in the population of Himachal Pradesh, India. *International Journal of legal medicine*, **133(5):** 1401-1402.

Robinson, D.H. and G.J. Lafleche (2000). Nucleic Acid Electrophoresis in Agarose Gels. In: Essential Molecular Biology: A Practical Approach, Volume 1, Brown, T.A. (Ed.). 1st Edn., Oxford University Press, New York, USA., ISBN-13:9780199636426, 89-119.

Roewer, L. (2019). Y-chromosome short tandem repeats in forensics-Sexing, profiling and matching male DNA. Wiley Interdisciplinary Reviews: *Forensic Science*, **1(4):** e1336.

Share, F.M. (2014). Evaluation of Y-Chromosome Short TandemRepeats (STR) Allele. Distribution in the Saudi population. M.Sc. Thesis. Department of Criminal Neighborhoods.College of Criminal Sciences. Naif Arab University forSecurity Sciences, KSA.

Stephens, J.C., J.A. Schneider, D.A. Tanguay, J. Choi, T. Acharya, S.E. Stanley and J. Duan (2001). Haplotype variation and linkage disequilibrium in 313 human genes. *Science*, **293(5529):** 489-493.

Tadmouri, G.O., K.S. Sastry and L. Chouchane (2015). Arab gene geography: From population diversities to personalized medical genomics. *Global Cardiology Science and Practice*, **2014(4):** 54.

Tareq, Z.M.A. (2017). Study of Y-Chromosome STR Markersin United Arab Emirates Population. M.Sc. Thesis. Department of Biology, College of Science, United ArabEmirates University, UAE.

Triki-Fendri, S., S. Alfadhli, I. Ayadi, N. Kharrat, H. Ayadi and A. Rebai (2010). Genetic structure of Kuwaiti population revealed by Y-STR diversity. *Annals of Human Biology*, **37(6):** 827-835.

Underhill, P.A. and T. Kivisild (2007). Use of Y-chromosome and mitochondrial DNA population structure in tracing human migrations. *Annu. Rev. Genet.*, **41:** 539-564.

Weir, B.S. and C.C. Cockerham (1984). Estimating F-statistics for the analysis of population structure. *Evolution,* 1358-1370.

Willuweit, S. and L. Roewer (2015). The new Y chromosome haplotype reference database. *Forensic Science International: Genetics*, **15:** 43-48.

Wright, S. (1978). Evolution and the Genetics of Populations. Vol. 4. Variability within and among Natural Populations. University of Chicago Press, Chicago.