

Full Length Research Paper

Detection of novel polymorphisms in the mitochondrial DNA D-Loop hypervariable region HVI from 400 healthy unrelated individuals from central and North-central Iraq

Ameera, Omran Hussein¹, Imad, Hadi Hameed^{1*} and Muhanned, Abdulhasan Kareem²

¹Department of Molecular Biology, Babylon University, Hilla City, Iraq.

²Babylon University, Centre of Environmental Research, Iraq.

Received 30 January, 2015; Accepted 25 March, 2015

The mitochondrial DNA (mtDNA) is a small circular genome located within the mitochondria in the cytoplasm of the cell and a smaller 1.2 kilobase pair fragment, called the control region (D-loop). The aims of this research were to study this region by using the Sanger sequencing technique and establish the degree of variation characteristic of a fragment. FTA® Technology (FTA™ paper DNA extraction) was utilized to extract DNA. PCR products were purified by EZ-10 spin column then sequenced and detected by using the ABI 3730xL DNA Analyzer. Novel polymorphisms discovered at positions 16037, 16075, 16104 and 16201 in future may be suitable sources for identification purpose.

Key words: D-loop, frequency, north-central Iraq, mitochondrial DNA, polymorphism.

INTRODUCTION

Mitochondrial DNA (mtDNA) is a useful genetic marker for answering evolutionary questions due to its high copy number, maternal mode of inheritance, and its high rate of evolution (Stoneking and Soodyall, 1996; Kraytsberg et al., 2004; Imad et al., 2015a). mtDNA is a small circular genome located within the mitochondria in the cytoplasm of the cell. The mitochondrial genome can be divided into two regions: a large coding region, which is responsible for the production of various biological molecules involved in the process of energy production in the cell, and a smaller 1.2 kilobase pair fragment, called the control region. It is found to be highly polymorphic and

harbors three hypervariable regions (HV), HV1, HV2 and HV3 (Helgason et al., 2003; Ingman and Gyllensten, 2003; Ukhee et al., 2005). mtDNA molecules contain 37 genes coding for 22 tRNAs, two rRNAs and 13 mRNAs (Brown et al., 1993; Giulietta et al., 2000; Young, 2009; Imad et al., 2014a). In modern population genetics research, studies based on mtDNA and Y-chromosome DNA are an excellent way of illustrating population structure while tracing uni-parental inheritance and ancestry since mtDNA is maternally inherited while the Y-chromosome is paternally inherited.

mtDNA is therefore inherited from generation to

*Corresponding author. E-mail: imad_dna@yahoo.com. Tel: 009647716150716.

Table 1. Primer sequence, region amplified, fragment size (bp) and PCR product length for HVI.

Primer name	Primer sequence (5' - 3') (Forward; F, Reverse; R)	Region amplified	Fragment size (bp)	PCR Product length
HVI-1	F: 5'- TCGTACATTACTGCCAGCCA -3' R: 5'- CATGGGGAGGGGGTTTTGAT -3'	16094-16113 16196-16177	20 20	103
HVI-2	F: 5'- AAACCCCTCCCATGCTTA -3' R: 5'- AGGGTGGGTAGGTTTGTGG -3'	16181-16200 16297-16278	20 20	117
HVI-3	F: 5'-TCACCCATCAACAACCGCTA -3' R: 5'- TAAGCATGGGGAGGGGGTTT -3'	16068-16087 16200-16181	20 20	133
HVI-4	F: 5'-TTCGTACATTACTGCCAGCCA-3' R: 5'- GGGAGGGGGTTTTGATGTGG -3'	16093-16113 16192-16173	20 20	100
HVI-5	F: 5'-TCATGGGGAAGCAGATTTGGG-3' R: 5'- TCATGGTGGCTGGCAGTAAT -3' R: 5'-GTCTGTGTGGAAGTGGCTGT-3'	16029-16049 16119-16100 277-257	21 20 21	91

generation through the maternal line. Mechanisms for this include simple dilution (an egg contains 100,000 to 1,000,000 mtDNA molecules, whereas a sperm contains only 100 to 1000), degradation of sperm mtDNA in the fertilized egg, and at least in a few organisms, failure of sperm mtDNA to enter the egg. Most mitochondria are present at the base of the sperm's tail, which is used for propelling the sperm cells; sometimes the tail is lost during fertilization. Also, unlike nuclear DNA, where there is a shuffling of the chromosomes at every generation, the mitochondrial DNA does not recombine with any other DNA type and remains intact from generation to generation (Brown, 2000; Brown, 2002; Dobbs et al., 2002; Imad et al., 2014b). mtDNA from all mammalian is very similar, with the order and position of the genes being the same.

Genetic analyses in population studies of the mitochondrial genome can be done either by sequencing the mtDNA or through the use of restriction fragment length polymorphisms (RFLPs) (Guntheroth et al., 1986; Pastore, 1994). RFLPs utilize restriction enzymes that can recognize the presence or absence of specific polymorphic DNA regions, and cut sites in the coding region of the mtDNA. These polymorphisms allow scientists to compare mtDNA from crime scenes to mtDNA from given individuals to ascertain whether the tested individuals are within the maternal line (or another coincidentally matching maternal line) of people who could have been the source of the trace evidence.

MATERIALS AND METHODS

Sample collection

Population sample collected from 400 healthy unrelated volunteer

donors, recruited from central and north-central Iraq, where the samples were collected from donors after they have been displaced to the center of Iraq because of the bad security situation in the provinces at north of the center stricken.

mtDNA extraction and amplification

DNA was extracted from all dried blood samples on FTA cards following the manufacturer's procedure as described in Whatman FTA Protocol BD01 except that the Whatman FTA purification reagent was modified to half the volume reference (Imad et al., 2014c). A 1.2 mm diameter disc was punched from each FTA card with a puncher. The discs were transferred to new Eppendorf tubes and washed three times in 100 µl Whatman FTA purification reagent. Each time, the disc was incubated for 5 min at room temperature with moderate manual mixing and the reagent was discarded between washing steps. The discs were then washed twice in 200 µl TE buffer (10 mM Tris-HCl, 0.1 mM EDTA, pH 8.0), the buffer was discarded and the discs were left to dry at room temperature for 1 h. Amplification of HVI region was carried out using five sets of primers (Table 1). A portion of a noncoding region encompassing positions from 16024 to 16365 for HVI was amplified in accordance with the Anderson reference sequence (Anderson et al., 1981). 20 µL of Master Mix was added into PCR tube. Change the pipette tip and add 20 µL of Primer Mix into PCR tube. Add 10 µL of extracting DNA into the PCR tube after changing the pipette tip again. Allow all the liquid settles at the bottom of the tube, and not elsewhere. Check the volume in the PCR tube using the PCR tube with 50 µL in it. The location of the tube on the grid after putting the mixture in the thermal cycler 95°C hold for 10 min, 30 cycles of 94°C for 30 s, 52.5°C for 30 s, 65°C for 1 min, 72°C hold for 10 min, 4°C hold ∞ infinity which is the cycling protocol for amplification of mtDNA PCR.

Purification, cycle sequencing and sequence analysis of mitochondrial DNA

mtDNA PCR products were purified by EZ10-spin column DNA clean up kit 100 prep. The DNA sequencing of the PCR products

was done using the BigDye TM Terminator and utilizing POP-7 polymer (Applied Biosystems). The separation of the cycle sequencing products was carried out. Detection was by using the ABI 3730xL DNA Analyzer, cap array size of 96 and cap array length of 50. The reference sequence described by Anderson et al. (1981) was compared to the data observed. Within the mtDNA coding region, sequencing results were studied from a consensus sequence derived from multiple sequence results. Data were analysed by Sequencher™ and aligned with the Anderson sequence (Anderson et al., 1981) using sequence Navigator software.

Statistical analysis

Genetic diversity for the analyzed DNA fragment was calculated according to the formula:

$$h = (1 - \sum x_i^2) / (n - 1)$$

Where, n is sample size and x_i is the frequency of i-th mtDNA type (Gu, 2001).

The probability of two randomly selected individuals from a population having identical mtDNA types is:

$$(P = \sum x_i^2)$$

Where, p frequencies of the observed Haplotypes (Jones, 1972).

RESULTS AND DISCUSSION

The basic aim of this work was to assess the degree of variation characterizing a selected segment of the non-coding region of mtDNA. The study enabled identification of 117 different haplotypes and 30 polymorphic nucleotide positions (Table 2). Within these 30 variations, the most frequent variant (H1) was consistent with the Anderson sequence, transitions between T and C, transitions between A and G and only and transversions. This fact is consistent with abundant literature data revealing significant domination of transitions over transversions (Brown et al., 1982; Yang and Yoder, 1999; Imad et al., 2015b).

Eight polymorphic positions (16037, 16041, 16067, 16075, 16104, 16144, 16201 and 16234) showed transversions (Table 2) and 4 are novel (Table 3); the number of analyzed markers was been increased to compensate for the increasing number of profiles in the databases in order to minimize accidental matches between unrelated individuals. Progression of new technology is therefore very slow and the use of SNPs has sometimes met a reluctant reception.

Genetic diversity for the analysed DNA fragment was calculated according to the formula: $D = 1 - \sum p^2$ and recorded 95.4%. The calculated value of the genetic diversity should be understood as high in the context of noncoding function of the analysed DNA fragment. The

relatively high gene diversity and a relatively low random match probability were observed in this study.

Comparative analysis of our results with previously published Iraqi data revealed significant differences in varying patterns (Nadia et al., 1999). This observation supports the hypothesis that different SNP-type polymorphisms can be strongly associated with a given population. Haplotypes detected in this study group have been compared with other global populations: German (n = 200) (Lutz et al., 1998), US Caucasian (n = 604), Africa (n = 111), Malaysia (n = 195) (Budowle et al., 1999) and India (n = 98) (Mountain et al., 1995) (Table 4).

Walsh et al. (1991) and Tang (2002) showed that the polymorphisms of mtDNA coding area are less than that of mtDNA control region. Therefore, more efficient polymorphic sites should be used to provide an improved discrimination power for forensic mtDNA testing (Imad et al., 2014). As forensic markers, they should be phenotypically neutral to avoid landing investigators into serious situations of medical genetic privacy and ethnics, especially for mtDNA coding area whose mutation is often correlated with an increased risk of some disease. With the whole mtDNA sequences being researched, we are optimistic that the polymorphisms sites within mtDNA coding area will be useful in combination with SNPs in the control region in order to increase the discrimination power of mtDNA (David et al., 2013; Imad et al., 2014).

Conclusion

Sequence analysis of the noncoding region of mtDNA (HVI) conducted on a population of 400 unrelated individuals enabled identification of 117 different haplotypes. The novel polymorphic positions 16037, 16075, 16104 and 16201 described in future may be suitable sources for identification purpose.

Conflict of interests

The author(s) did not declare any conflict of interest.

ACKNOWLEDGEMENTS

We sincerely wish to thank Dr. Issam for providing us the opportunity to work on this project. We are also indebted to you for choosing the project, your enthusiasm for helping us through the various analysis stages, and for providing helpful criticism and feedback throughout the writing process and your patience and guidance as we progress to put this project together. The authors would like to thank Dr. Khalifa from the Institution of medico-legal for all time put to discuss the project together. The authors would also like to thank Zainab Al-Habubi from the Department of Biology for her guidance and help in the laboratory work.

Table 3. Contd.

16162	Transition	A-G	Presence	0.0275	2.75
16163	Transition	A-G	Presence	0.0125	1.25
16171	Transition	A-G	Presence	0.03	3.00
16179	Transition	C-T	Presence	0.0475	4.75
16183	Transition	A-G	Presence	0.025	2.50
16201	Transition	C-T	Presence	0.0425	4.25
16201	Transversion	C-G	New*		
16209	Transition	T-C	Presence	0.0275	2.75
16217	Transition	T-C	Presence	0.0425	4.25
16224	Transition	T-C	Presence	0.035	3.50
16234	Transition	T-C	Presence	0.0305	3.10
16247	Transition	T-C	Presence	0.035	3.50
Genetic diversity* $D = 1 - \sum p^2 = 0.954 = 95.4\%$					

New*: new polymorphic positions; Genetic diversity*, Genetic diversity for the analysed DNA fragment was calculated according to the formula: $D = 1 - \sum p^2$.

Table 4. Comparisons of the characteristics across D-loop region in different human population groups.

Population	Iraq ¹	India ²	Malaysia ³	Africa ⁴	German ⁵	US Caucasian ⁶
Sample size	400	98	195	111	200	604
No. of variant sites	30	83	149	97	153	233
A→G	104	233	473	323	330	1112
G→A	155	66	81	78	55	219
T→C	110	145	461	382	308	1007
C→T	115	117	321	486	199	688
A→T	2	1	2	0	4	2
A→C	0	23	81	15	5	47
G→T	0	0	0	18	0	1
G→C	0	0	3	0	1	6
C→A	0	0	30	17	11	12
C→G	9	4	1	6	19	6
T→A	2	7	5	0	1	0
T→G	0	0	3	0	0	0
Insertion	0	168	322	140	291	983
Deletion	0	0	28	6	6	14

Note: % of transitions and transversions were calculated as number of observations divided by total substitutions. ¹This study; ²Mountain et al. (1995); ³Budowle et al. (1999); ⁴Budowle et al. (1999); ⁵Lutz et al. (1998); ⁶Budowle et al. (1999).

REFERENCES

- Anderson S, Bankier AT, Barrell BG (1981). Sequence and organization of the human mitochondrial genome. *Nature* 290:457-465.
- Brown JR, Bechenbach AT, Smith MJ (1993). Intraspecific DNA sequence variation of the mitochondrial control region of white sturgeon (*Acipenser transmontanus*). *Mol. Biol. Evol.* 10:326-341.
- Brown MD (2000). The role of MtDNA background in disease expression: Anew primary LHON mutation associated with Western Eurasian haplogroup. *J. Hum. Genet.* 110(2):130-138.
- Brown MD (2002). Functional analysis of lymphoplast and cybrid mitochondria containing the 3460, 11778, or 14484 Lebersheredity optic neuropathy mitochondrial DNA mutation. *J. Biol. Chem.* 275(51):39831-39836.
- Brown WM, Prager EM, Wang A (1982). Mitochondrial DNA sequences of primates: tempo and mode of evolution. *J. Mol. Evol.* 18:225-239.
- Budowle B, Wilson MR, DiZinno JA (1999). Mitochondrial DNA regions HV1 and HVII population data. *Forensic Sci. Int.* 103:23-35.
- David HW, Jonathan K, Arthur JE (2013). Validation of the PLEX-IDTM mass spectrometry mitochondrial DNA assay. *Int. J. Legal Med.* 127(2):277-286.
- Dobbs LJ, Madigan MN, Carter AB (2002). Use of FTA gene guard filter paper for the storage and transportation of tumor cells for molecular testing. *Arch. Pathol. Lab Med.* 126 (1):56-63.
- Giulietta DB, Ivane SN, Michele S (2000). Mitochondrial DNA sequences in prehistoric human remains from the Alps. *Eur. J. Hum. Genet.* 8:669-677.
- Gu MB (2001). Polymorphism study of the mitochondrial DNA control region in Chinese Han population group. *Chin. J. Forensic Med.* 16(1):6-9.
- Guntheroth WG, Lohmann R, Spiers PS (1986). Risk of sudden infant death syndrome in subsequent sibling. *J. Pediatr.* 116(4):520-524.
- Helgason A, Hrafnkelsson B, Gulcher JR (2003). A populationwide coalescent analysis of Icelandic matrilineal and patrilineal genealogies: evidence for a faster evolutionary rate of mtDNA lineages than Y chromosomes. *Am. J. Hum. Genet.* 72:1370-1388.
- Imad H, Mohammed A, Cheah Y (2014a). Genetic variation of twenty autosomal STR loci and evaluate the importance of these loci for forensic genetic purposes. *Afr. J. Biotechnol.* 13:1-9.
- Imad H, Muhanned A, Aamera J, Cheah Y (2014c). Analysis of eleven Y-chromosomal STR markers in middle and South of Iraq. *Afr. J. Biotechnol.* 13(38):3860-3871.
- Imad HH, Ameer IA, Mohammed AJ (2014b). Haplotypes and variable position detection in the mitochondrial DNA coding region encompassing nucleotide positions 10,716-11,184. *Mitochon. DNA.* 1-6.
- Imad HH, Mohammed AJ, Muhanned AK (2015a). Forensic analysis of mitochondrial DNA hypervariable region HVII (encompassing nucleotide positions 37 to 340) and HVIII (encompassing nucleotide positions 438-574) and evaluate the importance of these variable positions for forensic genetic purposes. *Afr. J. Biotechnol.* 14(5):365-375.
- Imad HH, Muhanned AK, Rafid HH (2015b). X-chromosome short tandem repeat, advantages and typing technology review. *African J. of Biotechnology.* 14(7): 535-541.
- Ingman M, Gyllensten U (2003). Mitochondrial genome variation and evolutionary history of Australian and New Guinean Aborigines. *Genome Res.* 13(7):1600-1606.
- Jones DA (1972). Blood samples: probability of discrimination. *J. Forensic Sci. Soc.* 12:355-358.
- Kraytsberg Y, Schwartz M, Brown TA (2004). Recombination of human mitochondrial DNA. *Science* 304(56731):28-981.
- Lutz S, Weisser H-J, Heizmann J (1998). Location and frequency of polymorphic positions in the mtDNA control region of individuals from Germany". *Int. J. Legal Med.* 111:67-77.
- Mohammed A, Imad H. (2013) Autosomal STR: From locus information to next generation sequencing technology. *Res. J. Biotechnol.* 8(10):92-105.
- Mohammed AJ, Imad HH, Muhanned AK (2015). Detection of New Variant "Off-ladder" at the (D12S391, D19S433 and D1S1656 loci) and Tri-allelic Pattern at the D16S539 Locus in a 21 Locus Autosomal Short Tandem Repeat Database of 400 Iraqi Individuals. *Afr. J. of Biotechnol.* 14(5): 375-399.
- Mountain JL, Hebert JM, Bhattacharyya S (1995). Demographic history of India and mtDNA-sequence diversity. *Am. J. Hum. Genet.* 56:979-992.
- Nadia A, Maria P, Vincenza B (2011). In search of the genetic footprints of Sumerians: a survey of Y-chromosome and mtDNA variation in the Marsh Arabs of Iraq. *BMC Evolut. Biol.* 11: 288.
- Pastore GM (1994). Leigh syndrome and hypertrophic cardiomyopathy in an infant with a mitochondrial DNA point mutation (T8993). *Am. J. Med. Genet.* 50: 265-71.
- Stoneking M, Soodyall H (1996). Human Evolution and the Mitochondrial Genome. *Curr. Opin. Genet. Dev.* 6(6):731-736.
- Tang H (2002). Polymorphisms of mitochondrial DNA hypervariable region of Beijing Han population. *Chin. J. Forensic Med.* 17(1):32-33.
- Ukhee C, Hwan YL, Ji-Eun Y (2005). Mitochondrial DNA CA dinucleotide repeats in Koreans: the presence of length heteroplasmy. *Int. J. Legal Med.* 119:50-53.
- Walsh PS, Metzger DA, Higuchi R (1991). Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *Biotechniques* 10(4):506-513.
- Yang Z, Yoder AD (1999). Estimation of the transition/transversion rate bias and species sampling. *J. Mol. Evol.* 48:274-283.
- Young KL (2009). The Basques in the Genetic Landscape of Europe. PhD Dissertation. University of Kansas.