Rom J Leg Med [21] 125-136 [2013]
DOI: 10.4323/rjlm.2013.125

The effectiveness of sequence variants of MTCOI and MTCYB besides entire D-Loop for haplotyping in eight population groups living in Taiwan

Yen-Ching Chen[1], James Chun-I Lee[1,2], Chun-Yen Lin[2], Tsang-Ming Ko[3], Yu-Hsuan Huang[2], Hsiang-Yi Yin[1], Li-Hui Tseng[4], Hsiao-Lin Hwa[1,4,5,]*

Abstract: Objectives: Analysis of human mitochondrial DNA (mtDNA) polymorphisms of hypervariable regions has become an important tool in forensic casework. This study evaluates the usefulness of the coding region polymorphisms in addition to hypervariable regions of different populations to increase the value of mtDNA polymorphic data in forensic genetics.

Methods: The sequence polymorphisms and haplotype frequencies of MTCOI, MTCYB, and the entire D-Loop of mtDNA, including hypervariable regions 1,2, and 3 of 432 individuals of eight population groups living in Taiwan, including Taiwanese Han, indigenous Taiwanese of Taiwan Island, Tao of Orchid Island, mainland Chinese, Filipinos, Thais, Vietnamese, and Caucasians were analyzed.

Results: Sequence positions with high heterogeneity were observed in D-Loop, MTCOI as well as MTCYB. Among the 432 subjects, 317 different haplotypes were observed for a combination of MTCOI, MTCYB, and entire D-Loop sequences, and 268 haplotypes were noted only once. The forensic parameters were calculated for different combinations of mtDNA regions. In most of these populations, sequence variants of MTCOI and MTCYB could further discriminate the haplotypes, besides the polymorphisms of hypervariable regions 1 and 2. The effectiveness of sequence variants of MTCOI and MTCYB for further haplotype differentiation varies in different population groups.

Conclusions: In addition to the polymorphisms of mtDNA hypervariable regions 1 and 2, the sequence variants of MTCOI and MTCYB are helpful for individual identification with varied effectiveness in different population groups.

Key Words: D-loop, hypervariable region 3, MTCOI, MTCYB, random match probability.

The analysis of mitochondrial DNA (mtDNA) sequences has become a powerful tool in forensic casework, especially in cases with a small amount of samples, severely degraded samples, or matrilineal determination [1, 2].

The sequence of the mtDNA hypervariable region 1 (HV1, nt16024-16365) and hypervariable regions 2 (HV2, nt73-340) is commonly used for forensic genetic purposes and many databases of mtDNA hypervariable region 1 (HV1) and hypervariable region 2 (HV2) have been established [3-5].

Besides HV1 and HV2, the entire D-loop,

1) Department and Graduate Institute of Forensic Medicine, College of Medicine, National Taiwan University
** Corresponding author: Associate Professor Hsiao-Lin Hwa, Department and Graduate Institute of Forensic Medicine, College of Medicine, National Taiwan University, No. 1, Jen-I Road Section 1, 100, Taipei, Taiwan, E-mail: hwahl013@ntu.edu.tw*
2) Institute of Forensic Medicine, Ministry of Justice, Taiwan, R.O.C.
3) Genephile Bioscience Laboratory, Taipei, Taiwan
4) Department of Medical Genetics, National Taiwan University Hospital
5) Department of Obstetrics and Gynecology, National Taiwan University Hospital and College of Medicine, National Taiwan University

including the region between nt 16366 and 72, called the 7S DNA region (7S-SP), and the region between 341 and 576, call the hypervariable region 3 (HV3) extended region (HV3ex), has been described to provide additional information for haplotype determination [6-8].

The sequence variants of MTCYB and MTCOI have also been suggested to be markers of individual identification in different population groups [9-12]. Moreover, in forensic casework, the significance of mtDNA matching requires comparison with a large mtDNA sequence database in order to determine the relative rarity of a particular case. Because of the haploid maternal inheritance of mtDNA, and that most mtDNA types are very rare, extending the mtDNA typing database can augment the forensic power of mtDNA evidence [13].

Because Taiwan is an island located at the junction the Western Pacific and the South China Sea, the heterogeneous population of Taiwan comprises two major groups: a Han population (about 22,680,000, 97.4%) and an indigenous population (about 530,000, 2.3%) [14].

The Taiwanese Han population includes the descendants of individuals who migrated from southeastern mainland China from the 1600s to the early 1900s, and a large scale migration from all areas of China around 1949. The indigenous population consists of 13 tribes living mainly in the mountainous region and east coast of Taiwan proper, and a tribe (Tao) living on Orchid Island [15]. There are also different population groups from other countries throughout the world living or working in Taiwan presently (about 880,000) [16].

The aim of this study was to present sequence variants data based on the entire mtDNA D-loop (including HV1, HV2, HV3ex, and 7S-SP), MTCOI, and MTCYB of different population groups living in Taiwan. The forensic genetic parameters were calculated according to polymorphisms of different combinations of regions of the D-loop, MTCOI and MTCYB, in order to evaluate the efficiency of various combinations. A multi-dimensional scaling (MDS) plot of these population groups based on the genetic distances between each group is also demonstrated.

MATERIALS AND METHODS

This retrospective study was approved by the Ethics Committee of the Institution in accordance with the World Medical Association Declaration of Helsinki. A total of 432 DNA samples from apparently healthy and unrelated individuals living in Taiwan were analyzed. These samples were obtained from 111 Taiwanese Han (TWH), 120 indigenous Taiwanese of Taiwan Island (TWI), 46 Orchid Islanders (TAO), 26 mainland Chinese(CHI), 31 Filipinos (FIL), 35 Thais (THA), 36 Vietnamese (VIE), and 27 Caucasians (CAU). The Caucasian subjects were individuals with European, Near Eastern or South Asian ancestry from the United States of America (8), United Kingdom (2), France (3), Brazil (3), Germany (2), Australia (2), New Zealand (2), India (1), Peru (1), the Netherlands (1), Nicaragua (1), and Syria (1).

The blood samples and buccal swab samples were obtained from volunteer donors between 1993 and 2005. Standard methods of phenol-chloroform isoamyl alcohol extraction were used for DNA extraction from peripheral whole blood samples, and the QIAamp DNA Mini kit (Qiagene, Valencia, CA, USA) was used for DNA extraction from buccal cells.

For each DNA sample, sequencing of the entire D-Loop (nt 16024-576), MTCOI (nt 5904-7445), and MTCYB (nt 14747-15887) was performed. The primer pairs for PCR amplification of MTCOI, MTCYB, and D-loop are listed in Table S1. Some primer pairs have been described previously and others were designed by the authors [8,12,17-19]. The primers were numbered according to the revised Cambridge reference human mtDNA sequence (rCRS) [20,21]. The PCR amplification and cycle sequencing of the PCR products of MTCOI, MTCYB and the entire D-Loop have been described in our previous reports [8,11,12].

PCR amplification of MTCOI, MTCYB, and the entire D-Loop was performed with the reaction mixture presented in Table S2.Briefly, PCR amplification for MTCOI was carried out in 50μl of reaction mixture, which contained 25ng genomic DNA, 10μM dNTP, reaction buffer (10mM Tris-HCl, pH 8.3, 2.5mM MgCl2, 50mM KCl, 0.1% gelatin), 2.5 units of GeneTaq DNA polymerase (Genemark, Taipei, Taiwan) and 10μM each of the primers (L5798, H7493). PCR amplification for MTCYB was performed in 50μl of reaction mixture, containing 25ng genomic DNA, 10μM dNTP, reaction buffer (10mM Tris-HCl, pH 8.3, 2.5mM MgCl2, 50mM KCl, 0.1% gelatin), 2.5 units of Gene Taq DNA polymerase (Genemark), and 10μM each of the primers (L14724, H15915). PCR amplification for D-Loop was carried out in 50μl of reaction mixture, with 12.5ng genomic DNA, 10μM dNTP, reaction buffer (20mM Tris-HCl, pH 8.0, 1.5mM MgCl2, 0.1mM EDTA, 1mM DTT, stabilizers and 50% glycerol), 0.5 unit of Super-therm polymerase (Bertec Enterprise, Taipei, Taiwan) and 10μM each of the primers (L15969, H638).

Autoclaved deionized H2O served as a negative control, and reference female DNA 9947A (Applied Biosystems, Foster City, CA, USA) was used as a positive control, respectively. Amplification of MTCOI, MTCYB, and D-Loop was conducted in an ABI 2720 thermal cycler (Applied Biosystems) with the conditions described in Table S2: Cycle sequencing of PCR products of MTCOI, MTCYB, and D-Loop was conducted in an ABI 2700 or 9700 thermal cycler (Applied Biosystems) with the primers and conditions shown in Table S3. Sequencing was performed using the Big Dye Terminator v3.1 Cycle Sequencing Kit (Applied

Figure 1. The frequencies of each varied position compared to rCRS.

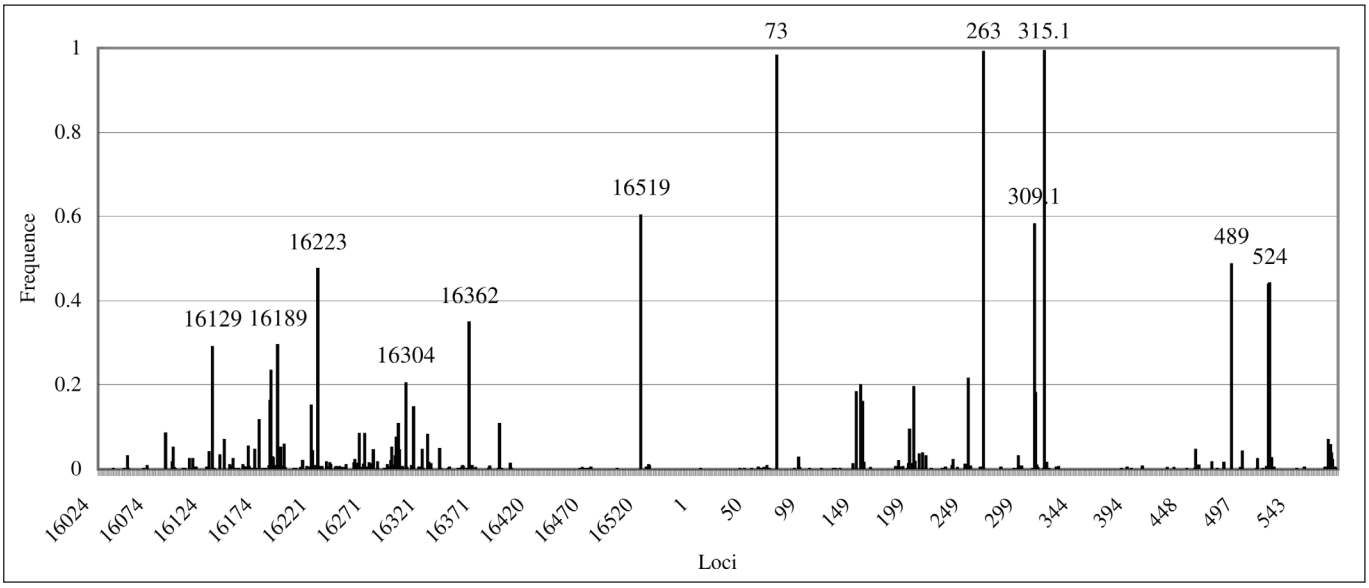

Figure 1 - A. Based on the entire D-loop.

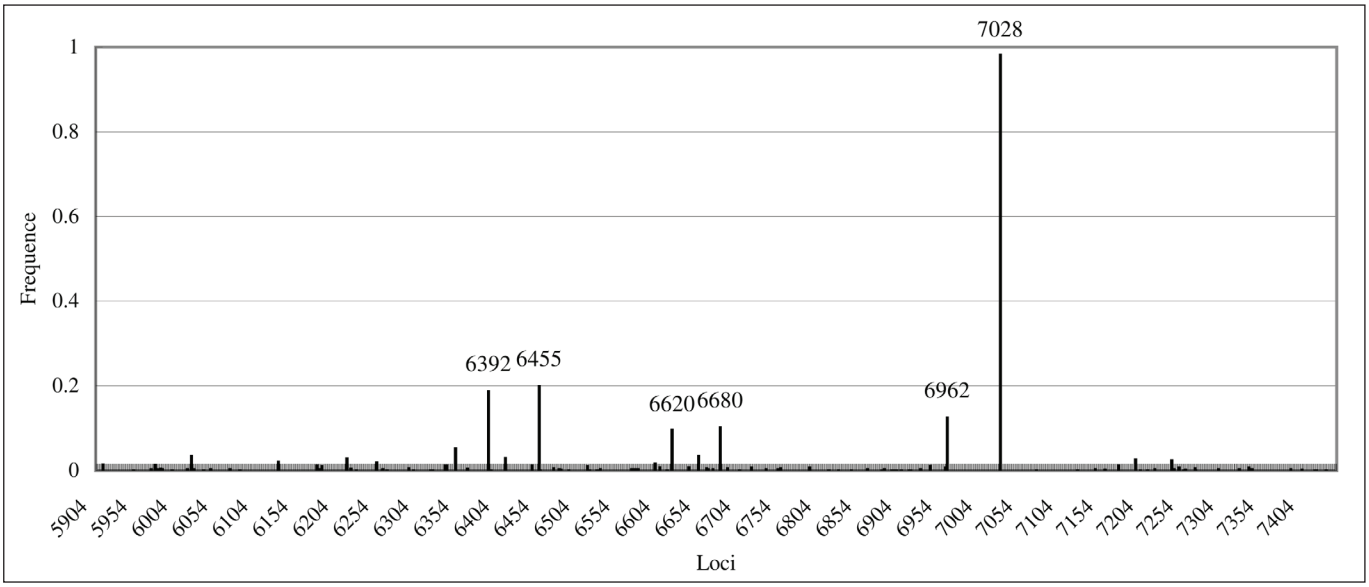

Figure 1 - B. Base on MTCOI.

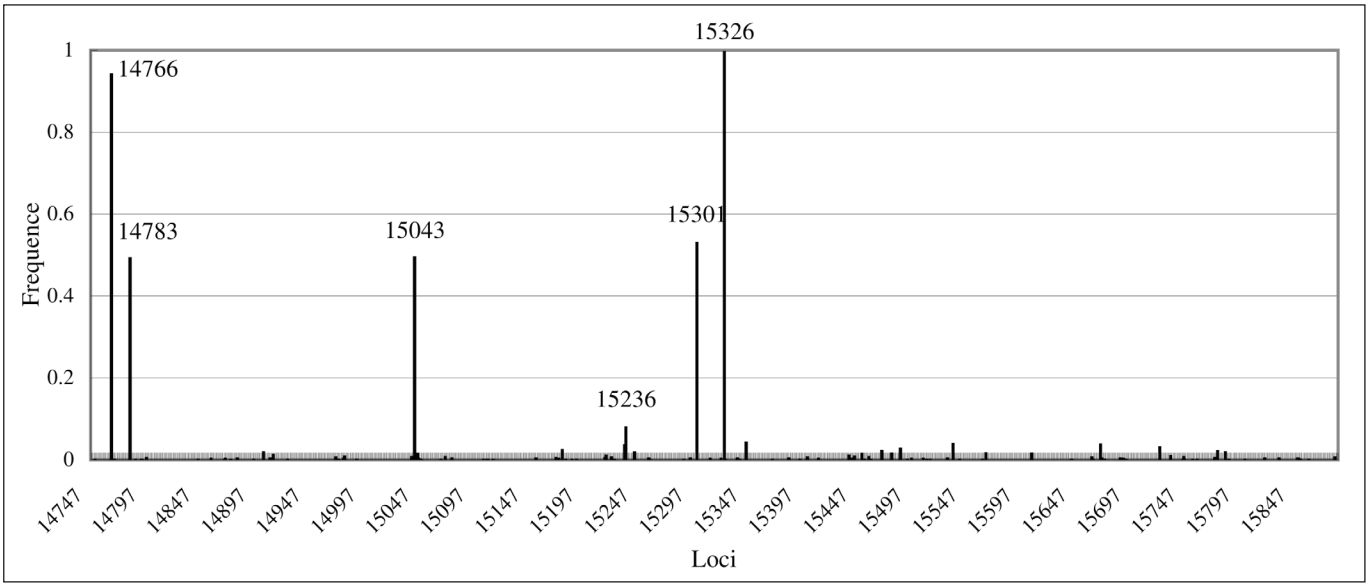

Figure 1 - C. Base on MTCYB.

Table 1. Random match probalility of various combinations of HV1 and HV2, or the entire D-loop (HV1+7S-SP+HV2+HV3ex), MTCOI, and MTCYB of eight population groups.

	TWH	TWI	TAO	CHI	FIL	THA	VIE	CAU	Combined
HV1+HV2	1.05	3.38	20.32	4.14	4.68	3.67	3.55	3.98	0.79
D-loop	1.05	2.50	20.32	4.14	4.68	3.67	3.09	3.98	0.67
MTCOI	9.93	9.35	23.91	15.98	10.72	16.08	10.49	27.57	8.28
MTCYB	9.00	8.69	19.28	26.33	11.76	19.51	17.44	11.93	7.09
HV1+HV2+MTCOI	1.05	2.96	19.57	3.85	4.47	3.67	3.24	3.70	0.69
D-loop+MTCOI	1.05	2.29	19.57	3.85	4.47	3.67	3.09	3.70	0.61
HV1+HV2+MTCYB	1.01	3.15	17.96	3.85	4.47	3.67	3.24	3.70	0.69
D-loop+MTCYB	1.01	2.35	17.96	3.85	4.47	3.67	3.09	3.70	0.61
HV1+HV2+MTCOI+MTCYB	1.01	2.93	17.20	3.85	4.27	3.67	3.09	3.70	0.63
D-loop+MTCOI+MTCYB	1.01	2.26	17.20	3.85	4.27	3.67	3.09	3.70	0.56

Biosystems). The DNA sequence was detected with an ABI 3730 DNA sequencer (Applied Biosystems). A portion of the MTCYB, MTCOI, and D-loop sequence data of these subjects has been reported in our previous studies [8,11,12].

Sequences were aligned using the DNAMAN for Windows version 5.2.10 (Lynnon BioSoft, Quebec, Canada). All of the sequences were compared with the rCRS using BioEdit version 7.0.0 (Ibis Therapeutics, Carlsbad, CS, USA), and genotyped according to the guidelines [22]. Cytosine insertions at positions 16193, 309, and 573 were ignored [23]. The frequencies of variants among different population groups were calculated using the Forensic DNA Data Analysis System (FORDDAS) v8.1 Build 0115 (http://140.112.138.124/forddas). Haplogroups were classified using HaploGrep (http://haplogrep.uibk.ac.at/) [24].

Characteristic parameters for population, including gene diversity, nucleotide diversity, and random match probability (RMP), and average number of nucleotide differences were calculated based on data combinations of different regions of the entire D-loop, MTCOI, and MTCYB. Gene diversity was estimated according to the formula presented by Nei [25]. The genetic relationship among different populations was evaluated based on the description by Excoffier et al [26-27]. Statistical analyses were performed using Arlequin v3.11 software [28] and tested for statistical significance using 10,000 randomizations. The pair-wise comparisons between population groups using Fst values and the Reynolds' genetic distance based on the combination of MTCOI, MTCYB, and the entire D-loop (nt16024-576) [29] were explored by Analysis of Molecular Variance (AMOVA) implemented in Arlequin 3.11 (http://cmpg.unibe.ch/software/arlequin3). A difference with a p-value <0.05 was taken as statistically significant. Based on Reynolds' genetic distance, MDS plot analysis was performed using the Statistical Package for the Social Sciences (SPSS) v16.0 (SPSS. Inc., Chicago, USA).

RESULTS

The 1542bp (nt5904-7445) fragment of the MTCOI, the 1141bp (nt14747-15887) fragment of the MTCYB, the entire D-Loop region (nt16024-576, about 1122bp), and the different combinations of these sequences as a haplotype were analyzed for forensic parameters.

There were 486 polymorphic sites that varied from the rCRS, with 506 kinds of variants identified. The average variant positions of the entire D-Loop, MTCOI, and MTCYB from the rCRS were 11.3 (3 to 24 positions), 2.3 (0 to 8 positions), and 4.3 (1 to 10 positions), respectively. By combining the entire D-loop, MTCOI, and MTCYB sequence as a haplotype, the analyzed samples presented a total of 317 haplotypes, 268 of which were noted only once. Figure 1 presents the frequencies of each varied position compared to rCRS. The most polymorphic region was HV1, with 133 (38.89%, 133/342) variable positions, followed by HV2 with 60 (22.39%, 60/268) variable sites. The sequence of MTCYB was more polymorphic than that of MTCOI. Only 7 samples of the Caucasian group carried the same sequence as the rCRS at nt5904-7445 (MTCOI) in this study. None of these 432 samples were the same as the rCRS at nt14747-15887 (MTCYB), nor at nt16024-576 (entire D-Loop).

Table 1 summarizes the RMP of the different combinations of MTCOI, MTCYB, HV1 and HV2 or the entire D-loop in these population groups. The sequence polymorphisms of HV3ex and 7S-SP improved the RMP based on HV1 and HV2 in TWI, and VIE. In addition to the combination of HV1 and HV2, the sequence variants of MTCOI or MTCYB could further discriminate the haplotypes of subjects of all population groups except THA. The sequence polymorphisms of MTCYB showed better further discriminative effectiveness than that of MTCOI in THW and TAO, whereas the sequence variants of MTCOI revealed better further discriminative

effectiveness than that of MTCYB in TWI. In the TWH, CHI, THA, and CAU groups, the RMPs of sequence variants of combinations of HV1, HV2, and MTCYB were the same as that of combinations of D-loop, MTCOI, and MTCYB. As a whole group, the RMP of these 432 subjects decreased as the combined regions expanded, and the RMP was 0.56 with a combination of the entire D-loop, MTCOI, and MTCYB. Table S4 presents the nucleotide diversity, gene diversity, and average number of nucleotide differences of various combinations of MTCOI, MTCYB and D-loop regions in these population groups.

Phylogenetic haplogroups based on variations in the complete D-loop, MTCOI, and MTCYB sequences of these eight population groups are shown in supplementary Table S5. The MDS plot of these eight population groups constructed on the basis of Reynolds' genetic distances and according to the sequences of the entire D-loop, MTCOI and MTCYB is illustrated in Figure 2.

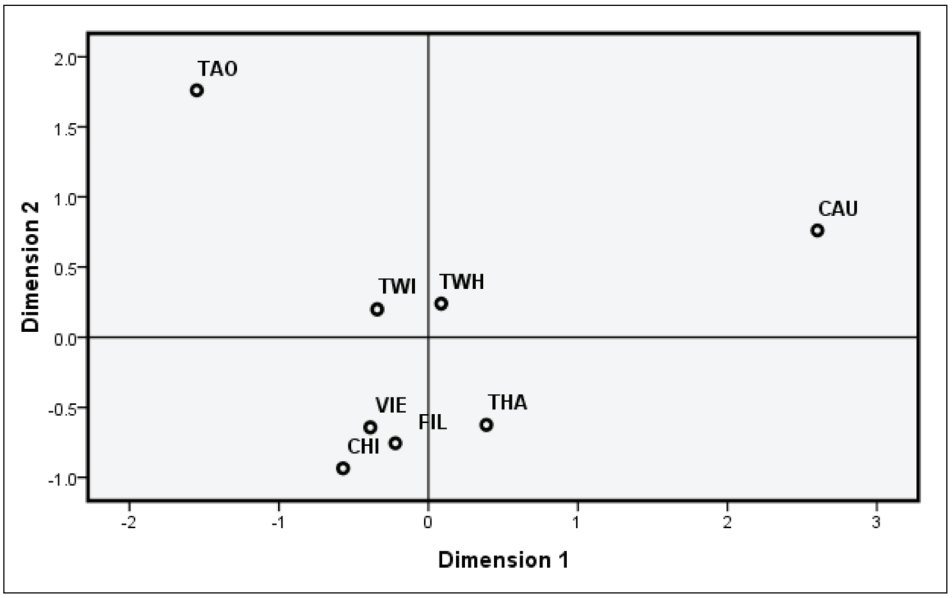

Figure 2. The multi-dimensional scaling plot based on Reynolds' genetic distance calculated according to the combination of sequence of the entire D-loop, MTCOI, and MTCYB.

DISCUSSION

Analysis of sequence polymorphisms of mtDNA HV1 and HV2 regions have been commonly used for forensic identification of individuals. Additional polymorphisms are helpful for further differentiation of haplotypes, because some haplotypes defined by HV1 and HV2 may be shared by many people in a population. The data of the frequency of a particular mDNA haplotype in a specific population is important for the usefulness of mitochondrial markers in forensic casework. Assessment of sequence variations of the entire D-loop and mtDNA coding regions can be useful for both individual characterization and further determination of mtDNA haplogroup lineages of Asian populations [8,9,12,30,31]. In this study, in addition to polymorphisms of HV1 and HV2, the sequence variants of HV3ex and 7S-SP improved RMP in the TWI and VIE groups only. The sequence polymorphisms of HV3ex and 7S-SP seem to be slightly helpful for further discrimination of haplotypes besides HV1 and HV2, and the effectiveness of discrimination may vary in different population groups.

In this study, we combined the sequences of the entire D-loop, MTCOI, and MTCYB in a total of 432 unrelated individuals from eight population groups living in Taiwan, and evaluated the efficiency of different combinations of regions applied in forensic genetics. In most of the population groups, the RMPs of a combination of HV1, HV2, and MTCYB were among the lowest or the second lowest, so sequence analysis of HV1, HV2 and MTCYB may be as effective as sequence analysis of the entire D-loop, MTCYB and MTCOI in individual identification. This is consistent with the finding that the MTCYB was more polymorphic than MTCOI, and the discrimination power of MTCYB was higher than that of the MTCOI in most of the population groups [12]. However, in specific population groups, such as TWI, sequence analysis of MTCOI can be more helpful than that of MTCYB, besides analysis of the entire D-loop. The efficiency of haplotype differentiation of polymorphisms of MTCOI, and MTCYB may also vary in different population groups.

In the TAO group, polymorphisms of both MTCOI and MTCYB were helpful for further haplotype differentiation in addition to HV1 and HV2. In an isolated small population with high RMP such as the TAO, further data on sequence variants of mtDNA coding regions, including MTCOI and MTCYB, would be valuable for individual identification. The relatively high RMP in the Tao group alerts us to interpret conservatively the results of mtDNA sequence analyses of forensic casework involving this population. More sequence polymorphism loci should be analyzed for a person from a small population group, and a haplotype match within a specific small group may not be conclusive [32].

The most frequently observed macrohaplogroups classified in TWH was M, followed by the B macrohaplogroup. The most common macrohaplogroups identified in TWI are M, B and E, and the most common in TAO are B and M. The haplogroup distribution of these

Table S1. The amplification and sequencing primers for mtDNA MTCOI, MTCYB and D-loop.

Primer	Sequences(5'→3')	
L5798*†	AAG CTG CTT CTT CGA ATT TGC	MTCOI
L6512†	TAT CTC TCC CAG TCC TAG CT	MTCOI
H6649†	GGA GAT TAT TCC GAA GCC TG	MTCOI
L6668†	CAG GCT TCG GAA TAA TCT CC	MTCOI
H7493*†	AAA AGT CAT GGA GGC CAT GG	MTCOI
L14724*†	CGA AGC TTG ATA TGA AAA ACC ATC GTT G	MTCYB
H15149†	TAA CTG TAG CCC CTC AGA ATG ATA TTT GTC CTC A	MTCYB
L15182†	TGA GGA CAA ATA TCA TTC TGA GGG GCT ACA GTT A	MTCYB
H15915*†	AAC TGC AGT CAT CTC CGG TTT ACA AGA C	MTCYB
L15969*†	CCA AGG ACA AAT CAG AGA AAA AGT C	D-loop
L15997†	CAC CAT TAG CAC CCA AAG CT	D-loop
L16262†	CTC ACC CAC TAG GAT ACC AAC	D-loop
H16401†	TGA TTT CAC GGA GGA TGG TG	D-loop
L29†	GGT CTA TCA CCC TAT TAA CCA C	D-loop
H408†	CTG TTA AAA GTG CAT ACC GCC A	D-loop
H638*†	GGA CCA AAC CTA TTT GTT TAT GGG	D-loop

* For PCR amplification
† For DNA sequencing

Table S2. The reaction mixture, primers, and the reaction conditions of PCR amplification of MTCOI, MTCYB, and the entire D-loop.

	MTCOI	MTCYB	D-loop
Reaction mixture (µl)	50	50	50
genonic DNA (ng)	25	25	12.5
10x reaction buffer	5	5	5
dNTP (10µM)	1	1	1
DNA polyperase (unit)	GenTaq DNA polymerase 2.5	GenTaq DNA polymerase 2.5	Super-therm polymerase 0.5
primers 10µM each (µl)	L5798 & H7493 1	L14724 & H15915 1	L15969 & H638 1
Reaction condition initial denaturation further exertion	95°C 5min 95°C 30sec 62°C 30sec 30cycle 72°C 1min 72°C 10min 4°C ∞	95°C 5min 95°C 30sec 55°C 30sec 35cycle 72°C 1min 72°C 10min 4°C ∞	95°C 5min 95°C 30sec 62°C 30sec 40cycle 72°C 1min 72°C 7min 4°C ∞

Table S3. The primers and the reaction conditions of cycle sequencing of PCR products of MTCOI, MTCYB and D-loop.

	MTCOI	MTCYB	D-loop
Primers	L5798 L6512 H6649 L6668 H7493	L14724 H15149 L15182 H15915	L15969 L15997 L16262 H16401 L29 H408 H638
Reaction condition	96°C 3min30sec 95°C 30sec 50°C 30sec 25cycle 60°C 4min 4°C ∞	96°C 3min30sec 95°C 30sec 50°C 30sec 25cycle 60°C 4min 4°C ∞	96°C 3min30sec 96°C 10sec 50°C 5sec 25cycle 60°C 4min 4°C ∞

Table S4. The nucleotide diversity, gene diversity, and average number of nucleotide differences of varied combinations of HV1 and HV2 or the entire D-loop (HV1+7S-SP+HV2+HV3ex) of eight population groups.

	ND±SD	GD±SD	ANND
TWH			
HV1+HV2	0.015923 ± 0.008139	0.9985 ± 0.0015	9.1
D-loop	0.010379 ± 0.005246	0.9985 ± 0.0015	11.7
MTCOI	0.001625 ± 0.000978	0.9089 ± 0.0209	2.3
MTCYB	0.002883 ± 0.001655	0.9183 ± 0.0171	4.3
HV1+HV2+MTCOI	0.005684 ± 0.002865	0.9985 ± 0.0015	11.4
D-loop+MTCOI	0.005312 ± 0.002658	0.9985 ± 0.0015	14.0
HV1+HV2+MTCYB	0.007428 ± 0.003732	0.9989 ± 0.0014	13.4
D-loop+MTCYB	0.006598 ± 0.003293	0.9989 ± 0.0014	16.0
HV1+HV2+MTCOI+MTCYB	0.004713 ± 0.002348	0.9989 ± 0.0014	15.7
D-loop+MTCOI+MTCYB	0.004583 ± 0.002272	0.9989 ± 0.0014	18.3
TWI			
HV1+HV2	0.014268 ± 0.007353	0.9744 ± 0.0058	8.6
D-loop	0.009698 ± 0.004918	0.9832 ± 0.0035	11.1
MTCOI	0.001871 ± 0.001097	0.9141 ± 0.0099	2.6
MTCYB	0.002772 ± 0.001600	0.9207 ± 0.0108	4.4
HV1+HV2+MTCOI	0.005395 ± 0.002726	0.9786 ± 0.0053	11.2
D-loop+MTCOI	0.005168 ± 0.002587	0.9853 ± 0.0033	13.7
HV1+HV2+MTCYB	0.006786 ± 0.003424	0.9766 ± 0.0053	13.0
D-loop+MTCYB	0.006204 ± 0.003103	0.9847 ± 0.0033	15.4
HV1+HV2+MTCOI+MTCYB	0.004486 ± 0.002238	0.9789 ± 0.0053	15.6
D-loop+MTCOI+MTCYB	0.004449 ± 0.002207	0.9856 ± 0.0033	18.1
TAO			
HV1+HV2	0.016702 ± 0.008624	0.8145 ± 0.0305	10.7
D-loop	0.010041 ± 0.005153	0.8145 ± 0.0305	13.3
MTCOI	0.001190 ± 0.000774	0.7778 ± 0.0293	2.2
MTCYB	0.002593 ± 0.001534	0.8251 ± 0.0259	4.6
HV1+HV2+MTCOI	0.005593 ± 0.002860	0.8222 ± 0.0315	12.9
D-loop+MTCOI	0.004918 ± 0.002503	0.8222 ± 0.0315	15.5
HV1+HV2+MTCYB	0.007511 ± 0.003822	0.8386 ± 0.0283	15.3
D-loop+MTCYB	0.006284 ± 0.003185	0.8386 ± 0.0283	18.0
HV1+HV2+MTCOI+MTCYB	0.004553 ± 0.002302	0.8464 ± 0.0289	17.5
D-loop+MTCOI+MTCYB	0.004221 ± 0.002128	0.8464 ± 0.0289	20.2
CHI			
HV1+HV2	0.012348 ± 0.006636	0.9969 ± 0.0117	8.1
D-loop	0.00792 ± 0.004203	0.9969 ± 0.0117	10.6
MTCOI	0.001447 ± 0.000918	0.8738 ± 0.0567	2.2
MTCYB	0.002100 ± 0.001314	0.7662 ± 0.0775	4.3
HV1+HV2+MTCOI	0.004542 ± 0.002394	1.0000 ± 0.0107	10.3
D-loop+MTCOI	0.004174 ± 0.002183	1.0000 ± 0.0107	12.8
HV1+HV2+MTCYB	0.005672 ± 0.002987	1.0000 ± 0.0107	12.4
D-loop+MTCYB	0.004984 ± 0.002604	1.0000 ± 0.0107	14.9
HV1+HV2+MTCOI+MTCYB	0.003695 ± 0.001922	1.0000 ± 0.0107	14.6
D-loop+MTCOI+MTCYB	0.003552 ± 0.001838	1.0000 ± 0.0107	17.1
FIL			
HV1+HV2	0.012882 ± 0.006853	0.9849 ± 0.0124	8.0
D-loop	0.008564 ± 0.004489	0.9849 ± 0.0124	11.3
MTCOI	0.001448 ± 0.000913	0.9226 ± 0.0262	2.2
MTCYB	0.003282 ± 0.001894	0.9118 ± 0.0281	4.9
HV1+HV2+MTCOI	0.004693 ± 0.002452	0.9871 ± 0.0122	10.3
D-loop+MTCOI	0.004445 ± 0.002300	0.9871 ± 0.0122	13.5
HV1+HV2+MTCYB	0.006629 ± 0.003434	0.9871 ± 0.0122	13.0
D-loop+MTCYB	0.005900 ± 0.003034	0.9871 ± 0.0122	16.3
HV1+HV2+MTCOI+MTCYB	0.004204 ± 0.002158	0.9892 ± 0.0119	15.2
D-loop+MTCOI+MTCYB	0.004096 ± 0.002091	0.9892 ± 0.0119	18.5
THA			
HV1+HV2	0.014671 ± 0.007698	0.9916 ± 0.0097	8.8
D-loop	0.009281 ± 0.004821	0.9916 ± 0.0097	11.3
MTCOI	0.001687 ± 0.001029	0.8639 ± 0.0427	2.5
MTCYB	0.002539 ± 0.001519	0.8286 ± 0.0521	4.0
HV1+HV2+MTCOI	0.005373 ± 0.002773	0.9916 ± 0.0097	11.3
D-loop+MTCOI	0.004886 ± 0.002506	0.9916 ± 0.0097	13.9
HV1+HV2+MTCYB	0.006768 ± 0.003488	0.9916 ± 0.0097	12.8
D-loop+MTCYB	0.005880 ± 0.003012	0.9916 ± 0.0097	15.3
HV1+HV2+MTCOI+MTCYB	0.004390 ± 0.002240	0.9916 ± 0.0097	15.4
D-loop+MTCOI+MTCYB	0.004182 ± 0.002124	0.9916 ± 0.0097	17.9
VIE			
HV1+HV2	0.014804 ± 0.007756	0.9921 ± 0.0084	9.1
D-loop	0.009531 ± 0.004938	0.9968 ± 0.0075	11.3
MTCOI	0.001921 ± 0.001144	0.9206 ± 0.0282	2.8
MTCYB	0.002556 ± 0.001526	0.8492 ± 0.0467	4.3
HV1+HV2+MTCOI	0.005578 ± 0.002871	0.9952 ± 0.0078	11.9
D-loop+MTCOI	0.005127 ± 0.002621	0.9968 ± 0.0075	14.1
HV1+HV2+MTCYB	0.006825 ± 0.003513	0.9952 ± 0.0078	13.4
D-loop+MTCYB	0.006013 ± 0.003074	0.9968 ± 0.0075	15.6
HV1+HV2+MTCOI+MTCYB	0.004531 ± 0.002306	0.9968 ± 0.0075	16.2
D-loop+MTCOI+MTCYB	0.004355 ± 0.002207	0.9968 ± 0.0075	18.4
CAU			
HV1+HV2	0.015593 ±0.008220	0.9972 ±0.0111	7.9
D-loop	0.009886 ±0.005162	0.9972 ±0.0111	9.9
MTCOI	0.001213 ±0.000798	0.7521 ±0.0692	1.5
MTCYB	0.002644 ±0.001585	0.9145 ±0.0334	3.1
HV1+HV2+MTCOI	0.005299 ±0.002763	1.0000 ±0.0101	9.4
D-loop+MTCOI	0.004868 ±0.002520	1.0000 ±0.0101	11.4
HV1+HV2+MTCYB	0.007162 ±0.003714	1.0000 ±0.0101	11.0
D-loop+MTCYB	0.006235 ±0.003214	1.0000 ±0.0101	13.0
HV1+HV2+MTCOI+MTCYB	0.004379 ±0.002255	1.0000 ±0.0101	12.5
D-loop+MTCOI+MTCYB	0.004201 ±0.002153	1.0000 ±0.0101	14.5
Combined			
HV1+HV2	0.015706 ±0.007981	0.9944 ±0.0009	8.8
D-loop	0.010150 ±0.005102	0.9956 ±0.0007	11.3
MTCOI	0.001692 ±0.001004	0.9194 ±0.0089	2.3
MTCYB	0.002875 ±0.001640	0.9313 ±0.0069	4.3
HV1+HV2+MTCOI	0.005652 ±0.002831	0.9954 ±0.0008	11.1
D-loop+MTCOI	0.005254 ±0.002613	0.9962 ±0.0007	13.6
HV1+HV2+MTCYB	0.007333 ±0.003662	0.9954 ±0.0008	13.0
D-loop+MTCYB	0.006482 ±0.003216	0.9962 ±0.0007	15.6
HV1+HV2+MTCOI+MTCYB	0.004693 ±0.002323	0.9960 ±0.0007	15.3
D-loop+MTCOI+MTCYB	0.004538 ±0.002236	0.9967 ±0.0006	17.9

ND: nucleotide diversity; SD: standard deviation; GD: gene diversity; RMP: random match probability; ANND: average number of nucleotide differences

Table S4. The nucleotide diversity, gene diversity, and average number of nucleotide differences of varied combinations of HV1 and HV2 or the entire D-loop (HV1+7S-SP+HV2+HV3ex) of eight population groups.

Haplogroups	TWH (111)	TWI (120)	TAO (46)	CHI (26)	FIL (31)	THA (35)	VIE (36)	CAU (27)	Sum
A2+C64T+T16111C!	0	0	0	0	0	0	0	1	1
A2w	0	0	0	0	0	0	0	1	1
A4	0	0	0	1	0	0	0	1	2
A4+A200G	0	0	0	1	0	0	0	0	1
A4e1	2	0	0	0	0	0	0	0	2
A5b	2	0	0	0	0	0	0	0	2
B2b	0	0	0	0	0	0	0	1	1
B2c	0	0	0	0	0	0	0	1	1
B4+C16261T	2	0	0	0	0	0	0	0	2
B4+C16261T+G16129A!	6	0	0	1	0	0	0	0	7
B4a1+T16311C!	0	1	0	0	0	0	0	0	1
B4a1a	1	0	0	0	0	0	0	0	1
B4a1a1a	0	0	0	0	0	0	0	1	1
B4a1a1a1	0	0	0	0	0	0	0	1	1
B4a2a	0	4	10	0	0	0	0	0	14
B4a4	1	0	0	0	0	0	0	0	1
B4b1a	0	0	0	0	0	0	1	0	1
B4b1a2	2	9	0	0	0	1	0	0	12
B4b1a3	1	0	0	0	0	0	0	0	1
B4b1b'c	2	0	0	0	0	0	0	0	2
B4c1b+A16335G	1	0	0	0	0	0	0	0	1
B4c1b2a	2	0	0	0	0	0	0	0	2
B4c1b2a2	0	0	11	0	0	0	0	0	11
B4c1c	1	0	0	0	0	0	0	0	1
B4c2	0	0	0	0	0	2	0	0	2
B4e	0	0	0	0	0	0	1	0	1
B4g	2	0	0	0	0	0	0	0	2
B5a	0	1	0	0	0	0	0	0	1
B5a1	3	0	0	0	0	0	0	0	3
B5a2a	0	10	0	0	0	0	0	0	10
B5b	1	0	0	0	0	0	0	0	1
B5b2a	1	0	0	0	0	0	0	0	1
C7	0	0	0	0	0	2	1	0	3
C7a1	0	0	0	1	0	0	0	0	1
C7a2	0	0	0	1	0	0	0	0	1
D4+T195C!	0	1	0	1	0	0	0	0	2
D4a	2	0	0	0	0	0	0	0	2
D4b1a	0	0	0	0	0	1	0	0	1
D4b1b	1	0	0	1	0	0	0	0	2
D4b1d	0	0	0	0	1	0	0	0	1
D4b2b	0	0	0	1	0	1	0	0	2
D4b2b5	1	0	0	0	0	0	0	0	1
D4e3	0	0	0	0	0	0	1	0	1
D4g2	1	0	0	1	0	0	0	0	2
D4j1a	1	0	0	0	0	0	0	0	1
D4j1a1	0	0	0	0	0	1	0	0	1
D4j3	0	0	0	1	0	0	0	0	1
D4j6	1	0	0	0	0	0	0	0	1
D5	1	0	0	0	0	0	0	0	1
D5a2a1	1	0	0	0	0	0	0	0	1
D5b	1	9	0	0	0	0	2	0	12
D5b1	2	0	0	0	0	0	0	0	2
D6c	0	0	0	0	0	0	1	0	1
E1a	2	4	0	0	0	0	0	0	6

Haplogroups	TWH (111)	TWI (120)	TAO (46)	CHI (26)	FIL (31)	THA (35)	VIE (36)	CAU (27)	Sum
E1a1a	4	15	0	0	0	0	0	0	19
E1b+C16261T	0	0	0	0	3	0	0	0	3
E2	0	0	0	0	1	0	0	0	1
E2a3	0	0	0	0	1	0	0	0	1
E2b1	0	6	3	0	0	0	0	0	9
F1+T16189C!	0	0	0	0	0	0	1	0	1
F1a	3	0	0	1	0	1	0	0	5
F1a1	4	0	0	0	0	1	3	0	8
F1a1a	3	0	0	1	0	3	4	0	11
F1a1d	0	0	4	0	0	1	0	0	5
F1a2	0	0	0	0	0	1	0	0	1
F1a3a	1	2	0	0	4	0	0	0	7
F1a4a	0	2	0	0	1	0	0	0	3
F1c1a	2	0	0	1	0	3	1	0	7
F1d	2	0	0	0	0	0	0	0	2
F1f	0	0	0	0	0	2	2	0	4
F2a	0	0	0	1	0	0	0	0	1
F2b	0	1	0	0	0	0	0	0	1
F2d	0	0	0	1	0	0	0	0	1
F3a1	0	0	0	1	0	0	1	0	2
F3b1	1	3	0	0	0	0	0	0	4
F4b	1	14	0	0	0	0	0	0	15
G1	1	0	0	0	0	0	0	0	1
G1a1	1	0	0	0	0	0	0	0	1
G2a	1	0	0	1	0	0	0	0	2
G2a1d2	0	0	0	0	0	1	0	0	1
H1+C16239T	0	0	0	0	0	0	0	1	1
H1+T152C!	0	0	0	0	0	0	0	1	1
H1+T16189C!	0	0	0	0	0	0	0	1	1
H1n6	0	0	0	0	0	0	0	1	1
H2a2	0	0	0	0	0	0	0	1	1
H4a1a1a	0	0	0	0	0	0	0	1	1
H5a1+T152C!	0	0	0	0	0	0	0	1	1
HV+T16311C!	0	0	0	0	0	0	0	1	1
I	0	0	0	0	0	0	0	1	1
K1a1b1a	0	0	0	0	0	0	0	1	1
K1a4c	0	0	0	0	0	0	0	1	1
L1c3b	0	0	0	0	0	0	0	1	1
M10a1	2	0	0	0	0	0	0	0	2
M10a1a	1	0	0	0	0	0	0	0	1
M10a2	1	0	0	0	0	0	0	0	1
M11a	0	0	0	0	0	0	1	0	1
M13’46’61+T16362C	0	4	0	0	0	0	0	0	4
M13a1b	0	0	0	0	0	1	0	0	1
M13a2	0	0	0	0	0	1	0	0	1
M17a	0	0	0	0	0	0	1	0	1
M20	1	0	0	0	0	0	0	0	1
M21a	0	0	0	0	1	0	0	0	1
M25	0	0	0	0	2	0	0	0	2
M36a	0	0	0	0	1	0	0	0	1
M37+T152C!	0	0	0	0	0	1	0	0	1
M49a’b	0	0	0	0	0	1	0	0	1
M7	0	1	0	0	0	0	0	0	1
M71	1	0	0	0	0	0	1	0	2
M71a1	1	0	0	0	0	0	0	0	1
M72	0	0	0	0	0	1	0	0	1

Haplogroups	TWH (111)	TWI (120)	TAO (46)	CHI (26)	FIL (31)	THA (35)	VIE (36)	CAU (27)	Sum
M73b	0	0	0	0	0	1	0	0	1
M74	0	0	0	0	0	0	1	0	1
M74a	1	0	0	0	0	0	0	0	1
M7b1	1	0	0	0	0	0	1	0	2
M7b1'2'4'5'6'7'8	2	0	0	1	0	0	1	0	4
M7b1'2'4'5'6'7'8+(C16192T)	6	0	0	2	0	0	3	0	11
M7b1'2'4'5'6'7'8+T16189C!	0	0	0	0	0	0	3	0	3
M7b2	1	0	0	0	0	0	0	0	1
M7b3	0	16	4	0	2	0	0	0	22
M7b4	1	0	0	0	0	0	0	0	1
M7b6	0	0	0	0	0	1	0	0	1
M7b8	1	0	0	0	0	0	0	0	1
M7c	1	0	0	1	0	0	1	0	3
M7c2b2	1	0	0	1	0	0	0	0	2
M7c3+T16295C!	0	0	0	0	1	0	0	0	1
M7c3a	0	2	0	0	0	0	0	0	2
M7c3b	1	0	0	1	0	0	1	0	3
M7c3c	1	5	0	0	5	0	0	0	11
M7c3c1	1	2	14	0	0	0	0	0	17
M7c'e'f	0	2	0	0	0	0	0	0	2
M7e	1	0	0	0	0	0	0	0	1
M8a2'3	3	0	0	0	0	1	0	0	4
M8a2a	1	0	0	0	0	0	0	0	1
M9	0	0	0	2	0	0	0	0	2
M9a1a	1	0	0	0	0	0	0	0	1
M9a1b1	0	0	0	0	0	1	0	0	1
M9a4a	0	0	0	0	0	1	0	0	1
M9a5	1	0	0	0	0	0	0	0	1
N	1	0	0	1	0	0	0	0	2
N22	0	0	0	0	1	0	0	0	1
N9a	1	0	0	0	0	0	1	0	2
N9a1	0	0	0	0	0	2	0	0	2
N9a10a	2	1	0	0	0	0	0	0	3
P9	0	0	0	0	1	0	0	0	1
R	0	1	0	0	0	0	0	0	1
R5a2b	0	0	0	0	0	1	0	0	1
R9b1a1a	1	0	0	0	0	0	0	0	1
R9b1a3	0	0	0	0	0	0	1	0	1
R9b1b	0	0	0	0	0	0	1	0	1
R9c	2	0	0	0	0	0	0	0	2
R9c1a	1	2	0	0	2	0	0	0	5
T1a	0	0	0	0	0	0	0	1	1
T1a1'3	0	0	0	0	0	0	0	1	1
T2b11	0	0	0	0	0	0	0	1	1
T2b7a1	0	0	0	0	0	0	0	1	1
T2f+T195C!	0	0	0	0	0	0	0	1	1
U2b	0	0	0	0	0	1	0	0	1
U3a	0	0	0	0	0	0	0	1	1
U4	0	0	0	0	0	0	0	1	1
U5a2a	0	0	0	0	0	0	0	1	1
Y1	1	0	0	0	0	0	0	0	1
Y1b	1	0	0	0	0	0	0	0	1
Y2	0	2	0	0	4	0	0	0	6
Z+T152C!	1	0	0	0	0	0	0	0	1
Total NO. of HG	111	120	46	26	31	35	36	27	432

HG: haplogroup

three population groups in this study is similar to that of previous reports [5,7,33,34]. Some specific positions in MTCOI (nt6221, 6776, 7028, and 7476,), and MTCYB (nt14766, 14798, 14905, 15607, and 15257)have been used for haplogrouping [35]. Therefore, the haplogroups of some subjects are different from those that depend on D-loop sequences only [8].

The MDS plot (Figure 2) based on the phylogenic analysis of a combination of sequence variants of the entire D-loop, MTCOI, and MTCYB of these eight population groups shows a matrilineal genetic substructure in this area. This MDS plot is similar to the MDS plots of our previous reports, and the analysis of a combination with more mtDNA loci may result in a MDS plot closer to the real population group substructure in this area [8,12].

The TWH and TWI are clustered together, which is different from the MDS plot based on Y-STR with TWH and CHI clustered together [36]. This result indicates that TWH may be a population with a matrilineal combination of both CHI and TWI. This is compatible with the migration history of the mainland Chinese Han men arriving Taiwan and the possible subsequent interbreeding with indigenous Taiwanese women of Taiwan Island. In the MDS plot, the outlying CAU population demonstrates the long genetic distance between Caucasian and Asian populations. The relatively long distances between the Tao and other Asian population groups represent the geographical and genetic isolation of this tribe on Orchid Island.

In conclusion, the performance of different combinations of mtDNA regions, including the entire D-loop, MTCOI, and MTCYB of eight population groups living in Taiwan, was presented in this study. In addition to HV1 and HV2, the sequence polymorphisms of MTCOI and MTCYB could be helpful in further haplotype differentiation, with varied effectiveness for individual identification in different population groups.

ACKNOWLEDGMENTS

We would like to thank Ms Pi-Mei Hsu, and Ms Shwu-Fang Li, for technical support on DNA extraction. This study was supported by a grant (No.101-1301-IFM(08)-06) from the Ministry of Justice, Taiwan R.O.C..

References

1. Giles RE, Blanc H, Cann HM, Wallace DC. Maternal inheritance of human mitochondrial DNA. Proc Natl Acad Sci U S A. 1980 Nov;77(11):6715-9.
2. Robin ED, Wong R. Mitochondrial DNA molecules and virtual number of mitochondria per cell in mammalian cells. J Cell Physiol. 1988 Sep;136(3):507-13.
3. Melton T, Wilson M, Batzer M, Stoneking M. Extent of heterogeneity in mitochondrial DNA of European populations. J Forensic Sci. 1997 May;42(3):437-46.
4. Pfeiffer H, Forster P, Ortmann C, Brinkmann B. The results of an mtDNA study of 1,200 inhabitants of a German village in comparison to other Caucasian databases and its relevance for forensic casework. Int J Legal Med. 2001 Feb;114(3):169-72.
5. Tsai LC, Lin CY, Lee JC, Chang JG, Linacre A, Goodwin W. Sequence polymorphism of mitochondrial D-loop DNA in the Taiwanese Han population. Forensic Sci Int. 2001 Jun;119(2):239-47.
6. Lee HY, Yoo JE, Park MJ, Chung U, Shin KJ. Mitochondrial DNA control region sequences in Koreans: identification of useful variable sites and phylogenetic analysis for mtDNA data quality control. Int J Legal Med. 2006 Jan;120(1):5-14.
7. Nur Haslindawaty AR, Panneerchelvam S, Edinur HA, Norazmi MN, Zafarina Z. Sequence polymorphisms of mtDNA HV1, HV2, and HV3 regions in the Malay population of Peninsular Malaysia. Int J Legal Med. 2010 Sep;124(5):415-26.
8. Hwa HL, Ko TM, Chen YC, Lin CY, Huang YH, Tseng LH, Su YN, Lee JCI. Sequence polymorphisms of mtDNA HV1, HV2 and HV3 regions in eight population groups living in Taiwan. Aust J Forensic Sci. 2012 Sep;44(3):243-52.
9. Lee SD, Lee YS, Lee JB. Polymorphism in the mitochondrial cytochrome b gene in Koreans. An additional marker for individual identification. Int J Legal Med. 2002 Apr;116(2):74-8.
10. Ray AM, Zuhlke KA, Levin AM, Douglas JA, Cooney KA, Petros JA. Sequence variation in the mitochondrial gene cytochrome c oxidase subunit I and prostate cancer in African American men. Prostate. 2009 Jun 15;69(9):956-60.
11. Hwa HL, Ko TM, Chen YC, Chang YY, Tseng LH, Su YN, Lee JCI. Study of the cytochrome b gene sequence in populations of Taiwan. J Forensic Sci. 2010 Jan;55(1):167-70.
12. Hwa HL, Lin CY, Ko TM, Yin HY, Tseng LH, Su YN, Lee JCI. Analysis of MTCOI and MTCYB sequence variations in eight population groups living in Taiwan. Rom J Leg Med. 2011 Sep;19(3):219-28.
13. Holland MM, Parsons TJ. Mitochondrial DNA sequence analysis - validation and use for forensic casework. Forensic Sci Rev. 1999 Jun;11(1):22-50.
14. Department of Statistics, Ministry of the Interior, Taiwan. [http://www.moi.gov.tw/]. Assessed on 02 April 2013.
15. Council of Indigenous Peoples, Executive Yuan, Taiwan.[http://www.apc.gov.tw/]. Assessed on 02 April 2013.
16. National Immigration Agency Taiwan. [http://www.immigration.gov.tw/]. Assessed on 02 April 2013.
17. Hsieh HM, Huang LH, Tsai LC, Liu CL, Kuo YC, Hsiao CT, Linacre A, Lee JCI. Species identification of Kachuga tecta using the cytochrome b gene. J Forensic Sci. 2006 Jan;51(1):52-6.
18. Kocher TD, Thomas WK, Meyer A, Edwards SV, Pääbo S, Villablanca FX, Wilson AC. Dynamics of mitochondrial DNA evolution in mammals: amplification and sequencing with conserved primers. Proc Natl Acad Sci U S A. 1989 Aug;86(16):6196-200.
19. Irwin DM, Kocher TD, Wilson AC. Evolution of the cytochrome b gene of mammals. J Mol Evol. 1991 Feb;32(2):128-44.
20. Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJ, Staden R, Young IG. Sequence and organization of the human mitochondrial genome. Nature. 1981 Apr 9;290(5806):457-65.
21. Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. Nat Genet. 1999 Oct;23(2):147.
22. Carracedo A, Bär W, Lincoln P, Mayr W, Morling N, Olaisen B, Schneider P, Budowle B, Brinkmann B, Gill P, Holland M, Tully G,

Wilson M. DNA commission of the international society for forensic genetics: guidelines for mitochondrial DNA typing. Forensic Sci Int. 2000 May;110(2):79-85.
23. Irwin JA, Ikramov A, Saunier J, Bodner M, Amory S, Röck A, O'Callaghan J, Nuritdinov A, Atakhodjaev S, Mukhamedov R, Parson W, Parsons TJ. The mtDNA composition of Uzbekistan: a microcosm of Central Asian patterns. Int J Legal Med. 2010 May;124(3):195-204.
24. Kloss-Brandstätter A, Pacher D, Schönherr S, Weissensteiner H, Binna R, Specht G, Kronengerg F. HaploGrep: a fast and reliable algorithm for automatic classification of mitochondrial DNA haplogroups. Hum Mutat. 2011 Jan;32(1):25-32.
25. Nei M. Molecular Evolutionary Genetics. New York: Columbia University Press; 1987.
26. Excoffier L, Smouse P, Quattro J. Analysis of molecular variance inferred from metric distances among DNA haplotype: application to human mitochondrial DNA restriction data. Genetics. Genetics. 1992 Jun;131(2):479-91.
27. Excoffier L, Smouse P. Using allele frequencies and geographic subdivision to reconstruct gene trees within a species: molecular variance parsimony. Genetics. 1994 Jan;136(1):343-59.
28. Excoffier L, Laval G, Schneider S. Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol Bioinform Online. 2007 Feb;1:47-50.
29. Reynolds J, Weir BS, Cockerham CC. Estimation of the coancestry coefficient: basis for a short-term genetic distance. Genetics. 1983 Nov;105(3):767-79.
30. Lutz-Bonengel S, Schmidt U, Schmitt T, Pollak S. Sequence polymorphisms within the human mitochondrial genes MTATP6, MTATP8 and MTND4. Int J Legal Med. 2003 Jun;117(3):133-42.
31. Nohira C, Maruyama S, Minaguchi K. Phylogenetic classification of Japanese mtDNA assisted by complete mitochondrial DNA sequences. Int J Legal Med. 2010 Jan;124(1):7-12.
32. Tillmar AO, Coble MD, Wallerström T, Holmlund G. Homogeneity in mitochondrial DNA control region sequences in Swedish subpopulations. Int J Legal Med. 2010 Mar;124(2):91-8.
33. Trejaut JA, Kivisild T, Loo JH, Lee CL, He CL, Hsu CJ, Lee ZY, Lin M. Traces of archaic mitochondrial lineages persist in Austronesian-speaking Formosan populations. PLoS Biol. 2005 Aug;3(8):e247.
34. Loo JH, Trejaut JA, Yen CJ, Chen ZS, Lee CL, Lin M. Genetic affinities between the Yami tribe people of Orchid Island and the Philippine Islanders of the Batanes archipelago. BMC Genet. 2011 Jan 31;12:21
35. Butler JM. Mitochondrial DNA analysis. In: Butler JM. Forensic DNA typing. 2nd ed. Burlington: Elsevier Academic Press; 2005. p. 241-98.
36. Hwa HL, Tseng LH, Ko TM, Chang YY, Yin HY, Su YN, Lee JC. Seventeen Y-chromosomal short tandem repeat haplotypes in seven groups of population living in Taiwan. Int J Legal Med. 2010 Jul;124(4):295-300.