Ancient mitochondrial DNA analysis of human remains found in Korea

Hwan Young Lee[1], Ji-Eun Yoo[1], Myung Jin Park[1], Ukhee Chung[1], Kyoung-Jin Shin[1,2], Woo-Ick Yang[1], Sang-Ho Cho[1,2] and Chong-Youl Kim[1,2]

[1] Department of Forensic Medicine, College of Medicine, Yonsei University, Seoul, Korea

[2] Human Identification Research Institute, Yonsei University, Seoul, Korea

Introduction

In recent years, molecular studies have become widely employed to investigate old skeletal remains, because morphological indicators are much less precise than the genetic data potentially available by analysis of ancient DNA. Understanding genetic information within and between skeletal remains found in Korea will help us to understand the organization of polulations and the origin of human remains recovered. Therefore, we extracted DNA from bone samples of ancient skeletons excavated from several burial places in the Paleolithic, Neolithic and Bronze ages, and determined the mitochondrial DNA haplogroups.

Materials and Methods

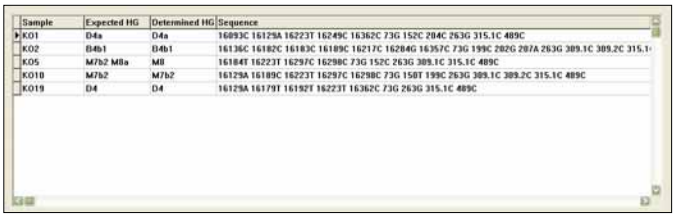
For DNA extraction from old skeletal remains, bone samples were pulverized by cryogenic grinding. From pulverized samples, DNA was extracted using QIAquick PCR purification kit after incubation in SDS/Proteinase K solution. To analyze mtDNA control region in highly degraded old skeletal remains, we performed PCR which was designed to produce amplicons with a length of 200 bp or less. Also, to determine mtDNA haplogroup exactly, we have selected 21 coding region SNPs and designed the 3 multiplex systems applying single base extension methods (see Table).

Haplo-group	SNP	Locus		PCR primer sequence	Length	tm (℃)	Amplicon size(bp)	primer conc.(uM)		SNaPshot primer sequence	Length	Length with T-tail	tm (℃)	primer conc.(uM)
PCR-Multiplex I														
M9	G/A	4491	F	cacctatcacaccccatcct	20	59.65	147	0.4	F	$(T)_{25}$GAATCCCCTGGCCCAACCC	19	45	64.4	0.2
			R	gatgagtgtgcctgcaaaga	20	59.99		0.4						
N9	G/A	5417	F	caccatcaccctccttaacc	22	59.26	142	0.4	F	$(T)_{25}$CATATCTAACAACGTAAAAATAAAATGAC A	30	65	56.6	0.1
			R	gatgagtgtggggaggaatg	20	60.33		0.4						
M11	G/A	7642	F	acatgcagcgcaagtaggtc	20	61.41	119	0.1	F	$(T)_{32}$GCTACTTCCCCTATCATAGAAGA	23	55	52.6	0.2
			R	gggcatacaggactaggaagc	21	60.11		0.1						
M10/A	T/C C/T	8793/8794	F	caacactaaaggacgaacctga	22	59.28	136	0.4	F(M10)	$(T)_{25}$CTAACCTCCTCGGACTCCTGCC	22	48	62.2	1.0
			R	gatggccatggctaggttta	20	59.92		0.4	R(A)	$(T)_{25}$GTGGTTGGTGTAAATGAGT	19	57	46.3	0.1
D5/M	A/G C/T	10397/10400	F	gccctaagtctggcctatga	20	59.29	133	0.8	F(D5)	$(T)_{5}$CTATGAGTGACTACAAAAAGGATTAGACTG	30	38	57.5	1.0
			R	tgtaaatgaggggcatttgg	20	60.69		0.8	R(M)	ATTCGTTTTGTTTAAACTATATACCAATTC	30	30	56.5	0.8
D4	C/T	14668	F	accccacaaaccccattact	20	60.34	138	0.8	F	$(T)_{11}$CCACACTCAACAGAAACAAAGCATA	25	36	59.2	0.5
			R	ttggggtcattggtgttctt	20	60.21		0.8						
PCR-Multiplex II														
F	C/T	3970	F	ggcttcaacatcgaatacgc	20	60.62	130	0.4	R	$(T)_{27}$GGTGTATTCGGCTATGAAGAATA	23	50	54.1	0.1
			R	caggggagagtgcgtcata	19	59.80		0.4						
G/D	A/G C/T	4833/4883	F	gtcccagaggttacccaagg	20	60.74	115	0.4	F(G)	$(T)_{9}$GTCCCAGAGGTTACCCAAGGC	21	29	60.7	0.5
			R	ggcttacgtttagtgagggaga	22	59.78		0.4	F(D)	$(T)_{4}$CACATGACAAAAACTAGCCCC	21	25	55.9	0.2
M8	C/A	7196	F	agaccaaacctacgccaaaa	20	59.61	122	0.4	F	$(T)_{15}$TAACTTTCTTCCCACAACACTTTCT	25	40	56.5	0.2
			R	catcggggtagtccgagtaa	20	59.95		0.4						
B	9bp del	8281-8289d	F	agggcccgtatttaccctat	20	58.74	132	0.8	R	$(T)_{37}$GTGGGCTCTAGAGGGGGT	18	55	55.1	0.2
			R	tttagttggggcatttcactg	21	59.98		0.8						
M7	T/C	9824	F	ggcatctacggctcaacatt	20	60.10	101	0.4	R	$(T)_{9}$GAAAGTTGAGCCAATAATGACGTG	24	33	58.7	1.0
			R	attagttggcggatgaagca	20	60.61		0.4						
R	C/T	12705	F	tgttcgttacatggtccatca	21	59.83	154	0.5	F	$(T)_{12}$AAACATTAATCAGTTCTTCAAATATCTACT CAT	33	45	58.0	0.2
			R	tctcagccgatgaacagttg	20	59.98		0.5						
PCR-Multiplex III														
D4	G/A	3010	F	gggataacagcgcaatccta	20	60.06	117	0.4	R	$(T)_{5}$TTTAATAGCGGCTGCACCAT	20	25	57.1	0.05
			R	tcgttgaacaaacgaacctt	20	58.26		0.4						
D4a	T/C	14979	F	gccatgcactactcaccaga	20	59.86	118	0.4	R	$(T)_{7}$CATTGGCGTGAAGGTAGCGGATG	23	30	66.3	0.2
			R	agaatattgaggcgccattg	22	60.06		0.4						
D4b	G/A	8020	F	cgactacggcggactaatct	20	59.36	125	0.4	R	$(T)_{13}$TTATACGAATGGGGGCTTCAAT	22	35	58.4	0.2
			R	ttatacgaatgggggcttca	20	60.28		0.4						
D4e	C/T	11215	F	ccacacttatccccaccttg	20	60.22	121	0.2	F	$(T)_{17}$CGCAGGCACATACTTCCTATTCTA	24	41	58.5	0.2
			R	agtgcgatgagtagggggaag	20	59.31		0.2						
D4g	G/A	8701	F	cgactaatcaccacccaaca	20	59.41	138	0.1	F	$(T)_{18}$CTAATCAAACTAACCTCAAAACAAATGATA	30	48	57.7	0.5
			R	tccgaggaggttagttgtgg	20	60.10		0.1						
D4j	G/A	11696	F	cagccattctcatccaaacc	20	60.46	113	0.6	R	$(T)_{29}$CGTGGGCGATTATGAGAATGA	21	50	59.5	0.2
			R	gcgttcgtagtttgagtttgc	19	59.94		0.6						

Results and Discussions

From mtDNA control region mutation motif, we could assigned 5 mtDNA haplotypes into mtDNA haplogroup, D4a, B4b1, M8, M7b2 and D4, respectively (Fig.1). Also we could confirm the results using 3 multiplex single base extension (SNaPshot) reactions (Fig.2). In this way, we could complement coding region information to control region mutation motifs and also unequivocally assign each ancient individual to a mitochondrial haplogroup.

Fig 1.



Sample	Expected HG	Determined HG	Sequence
K01	D4a	D4a	16093C 16129A 16223T 16249C 16362C 73G 152C 204C 263G 315.1C 489C
K02	B4b1	B4b1	16136C 16182C 16183C 16189C 16217C 16284G 16357C 73G 199C 202G 207A 263G 309.1C 309.2C 315.1
K05	M7b2 M8a	M8	16184T 16223T 16297C 16298C 73G 152C 263G 309.1C 315.1C 489C
K018	M7b2	M7b2	16129A 16189C 16223T 16297C 16298C 73G 150T 199C 263G 309.1C 309.2C 315.1C 489C
K019	D4	D4	16129A 16179T 16192T 16223T 16362C 73G 263G 315.1C 489C

Fig 2.

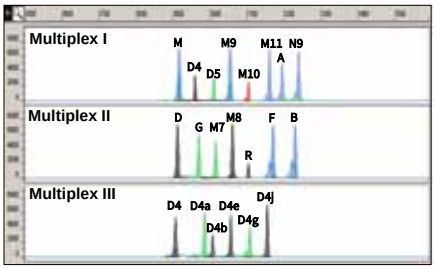


Fig 1. mtDNA haplogroup determination by control region mutation motifs.

Fig 2. mtDNA haplogroup determination by coding region SNP score. The result is representative of 5 bone samples.

Summary

We extracted DNA from bone samples of ancient skeletons excavated from several burial places in the Paleolithic, Neolithic and Bronze ages. To establish a high quality mtDNA database and phlygenetic haplogroups for ancient Koreans, we analyzed mtDNA control region sequence and coding region SNPs. We first assigned each mtDNA into an appropriate haplogroup according to the important control region mutation motifs. However, quite a few haplogroup-diagnostic SNPs are located in mtDNA coding region, and in some cases, scoring of coding region SNPs is required for exact haplogroup determination. Accordingly we have selected 21 coding region SNPs and designed the 3 multiplex systems applying single base extension methods. Using 2 multiplex systems, we allocated each ancient mtDNA into one of 15 haplogroups: M, D, G, D4, D5, M7, M8, M9, M10, M11, R, F, B, A and N9. Using the other multiplex, we further determined D4 subhaplogroups; D4a, D4b D4e, D4g and D4j, since D4 haplotypes occurred most frequently in Koreans. In this way, we could complement coding region information to control region mutation motifs and also unequivocally assign each ancient individual to a mitochondrial haplogroup.