

## 한국인에서의 Y 하플로그룹의 분포 양상 및 다른 민족과의 비교

장애화<sup>1</sup> · 이해영<sup>1</sup> · 서승범<sup>1,a</sup>  
이효정<sup>3</sup> · 김홍현<sup>1</sup> · 조소희<sup>1</sup>  
류성희<sup>2</sup> · 김기하<sup>2</sup> · 이재원<sup>3</sup>  
이승덕<sup>1,2</sup>

<sup>1</sup>서울대학교 의과대학 법의학교실

<sup>2</sup>서울대학교 의학연구원 법의학연  
구소

<sup>3</sup>고려대학교 정경대학 통계학과

접 수 : 2012년 4월 9일  
게재승인 : 2012년 5월 10일

이 논문은 2011년도 정부(교육과학기술  
부)의 재원으로 한국연구재단 바이오 · 의  
료기술개발사업의 지원을 받아 수행된 연  
구임(No. 2011-0027729).

본 연구는 질병관리본부 국립중앙인체자  
원은행에서 제공한 인체자원을 이용하여  
수행되었음(KOBB-2009-02).

— 위 논문 작성에 있어 장애화와 이해영이  
상당한 부분을 공동으로 작업하였음 (공동  
제1저자).

— 서승범의 현재 소속은 'Institute of  
Applied Genetics, Department of  
Forensic and Investigative Genetics,  
University of North Texas Health Science  
Center, Texas, USA' 임.

책임저자 : 이승덕  
(110-799) 서울시 종로구 대학로 103  
서울대학교 의과대학 법의학교실  
전화 : (02) 740-8359  
FAX : (02) 764-8340  
E-mail : sdlee@snu.ac.kr

### Y Haplogroup Distribution in Korean and Other Populations

Ai Hua Zhang<sup>1</sup>, Hye Young Lee<sup>1</sup>, Seung Bum Seo<sup>1,a</sup>, Hyo Jung Lee<sup>3</sup>,  
Hong Xuan Jin<sup>1</sup>, So Hee Cho<sup>1</sup>, Sung Hee Lyoo<sup>2</sup>, Ki Ha Kim<sup>2</sup>, Jae Won Lee<sup>3</sup>,  
Soong Deok Lee<sup>1,2</sup>

<sup>1</sup>Department of Forensic Medicine, <sup>2</sup>Institute of Forensic Medicine, Seoul National University  
College of Medicine, Seoul, Korea

<sup>3</sup>Department of Statistics, College of Political Science and Economics, Korea University,  
Seoul, Korea

With globalization, international exchange has increased. Accordingly, the necessity for individual identification using genetic polymorphism has also increased. Paternal lineages are distributed differently, and different distribution patterns can be used to predict ancestry. We studied the distribution pattern of different paternal lineages in Korea and compared them with other populations.

All 30 SNPs on the Y chromosome were selected for paternal lineage confirmation. Loci that could subclassify haplogroup O, the most frequent in the East Asian population, were added. After multiplex amplification for the target loci, SBE reactions were set up for each SNP site. One hundred Korean men as well as 60 Chinese, 60 Japanese, 19 African-American, 48 Caucasian, and 47 Mexican American were tested and compared. Five Y haplogroups [C (C3), D (D2), NO, O, Q (Q1a1)] were found in Koreans, with haplogroup O being the most frequent. Haplogroup O sub-classified into O\* (24%), O1 (6%), O2b (39%), O3a3c (4%), O3a3c1 (13%), and O3a3b(1%). This distribution pattern was similar to that of Chinese or Japanese, but minor differences were noted. With Fst, the Korean and Japanese patterns were close (0.01757) when using 6 SNPs. There were significant differences between Koreans and African Americans, Caucasians and Mexican Americans, and they were easily discernible without requiring haplogroup O sub-classification.

Sub-classification of haplogroup O is likely to be useful for East Asia group comparisons. Additional studies in populations from different areas of China or Japan or studies of mtDNA or autosomes may enhance the discriminatory power of genetic polymorphism in different Asian populations.

**Key words :** Y chromosome, SNP, paternal lineage, Korean

### 서 론

각국의 문호개방에 따라 상주 체류하거나 관광의 목적으로  
다른 나라를 오고가는 사람들의 범위는 점차 확대되는 추세이

다. 한국도 예외가 아니다. 몇 년 전만 해도 우리나라를 단일민  
족 국가라고 부르기도 하였지만, 현재는 110만 명이 넘는 외국  
인이 거주하고 있는 다민족국가로 변화되어 가고 있다. 타국과  
의 교류가 증가함에 따라 외국인 범죄도 증가하였다. 2009년  
법무부에서 제출한 국정감사보고 통계에 따르면 2004년

7,173건이었던 외국인 범죄가 2008년에는 약 4.7배로 증가된 34,108건에 이르러 범죄인원은 2007년에 비해 188.7%나 높아졌다고 한다.<sup>1)</sup> 또한 최근 들어 테러나 무기매매, 인신매매 등 초국가적인 범죄의 증가도 관심을 기울여야만 하는 대상이다.

이와 같은 상황에서는 범죄현장에서 얻은 증거물이 외국인에서 유래되었을 가능성을 아는 것은 수사 초기 매우 유용하다. 하지만 현 시점에서 우리나라 수사 당국은 이를 판단하기 어렵다. 모든 국가에서 자국민에 대해 범죄자를 대상으로 하는 DNA 데이터베이스를 구축하고 있지도 않고, 일부 관련 자료들을 가지고 있는 경우라도 우리나라에서 그들 자료에 자유로이 접근할 수도 없다.

STR을 기반으로 하는 고전적 개인식별의 개념에서는 위와 같은 경우 문제 해결이 쉽지 않다. STR에 비해 소위 SNP(SNP, Single Nucleotide Polymorphism)는 새로운 활용성을 제시하여 준다. 먼저, SNP를 활용하여 표현형에 대한 정보를 수집할 수 있다. FSS (Forensic Science Service, UK)는 모색 유전자 I에 관한 SNP 분석 기술을 개발하여 시료에 대한 붉은 색 모발을 확인할 수 있다고 하였고,<sup>2)</sup> 네덜란드의 Susan 등<sup>3-5)</sup>은 홍채색을 예측할 수 있는 민감도 높은 IrisPlex를 개발하였다. 한편 SNP를 이용하여 민족 기원을 밝히는 연구결과들이 최근 많이 보고되고 있다.<sup>6-8)</sup> 2003년 DNA Print라는 회사가 SNP를 이용하여 연쇄간사건에 수사정보를 제공한 사례는 널리 알려지기도 하였다.<sup>9)</sup> 이외 SNP를 증폭할 때, 증폭산물의 크기를 100 bp 보다 작게 조절하여 유골이나 부패된 시체 등과 같이 DNA가 절편화되어 검사가 어려운 경우에도 결과를 얻을 수 있고, STR 유전자에서 나타나는 stutter 등 결과 해석에 영향을 미칠 수 있는 현상이 나타나지 않으므로 결과를 쉽게 해석할 수 있다는 장점이 있다.

특히, Y-SNP는 남성을 결정하는 Y 염색체에 존재하는 SNP로써 pseudoautosomal region을 제외한 대부분에서 교차에 의한 재조합이 일어나지 않고 Y-STR보다 변이율이 낮으며, 아버지에서부터 아들에게 부계 유전되기 때문에 부계 유전자 검사, 인류 진화 연구나 인류 집단 이주 패턴을 추적하는 유용한 방법으로 사용되고 있다. 하지만, 같은 부계끼리는 모두 동일한 유전자형을 갖기 때문에 이것 또한 단점이기도 하다. 2008년 YCC (The Y Chromosome Consortium)에서 발표된 Y염색체 계통도는 A에서 T까지 주요 분기군(clade)을 포함하고, 600여개의 biallelic 마커들을 포함한 311개의 서로 다른 하플로그룹을 발표하였다. 이러한 서로 다른 Y-SNP 하플로그룹은 지리적인 기원을 예측하는데 도움을 줄 수 있다. 예를 들어 범죄현장에서 얻은 증거물의 지리적 기원에 관한 정보를 통하여 용의자의 가능성의 범위를 줄일 수 있어 범죄 수사에서 유용하게 쓰일 수 있다.

우리나라의 부계연구는 한국인에서 SNP를 기반으로 하는 하플로그룹 지점과 관련된 일부 연구결과를 확인할 수는 있지만,<sup>11)</sup> STR을 기반으로 한 연구에 비해 논문의 수가 많지는 않

고, 일부는 발표된 지 시간이 좀 지나 하플로그룹 구분이 세분화되어 있지 않았다.<sup>12)</sup>

본 연구에서는 계통발생학적 연구를 위해 혹은 현장 시료들이 어느 나라 민족에서 유래하였는지 관련 정보를 얻기 위해서는 한국인에서 다양한 부계 하플로그룹들이 어떻게 분포하는지를 확인할 필요가 있고, 이를 위해 국제적 기준의 하플로그룹을 판별할 수 있는 체계를 구성하고, 이를 활용하여 우리나라에서의 세분화된 Y 하플로그룹 분포 상황을 확인하며, 이를 주위 다른 민족과 비교 분석하여 보고자 한다.

## 재료 및 방법

### 1. 시료

본 연구는 다른 사업 목적으로 본 교실에서 보유하고 있었던 혈연관계가 없는 한국인 남자 100명을 대상으로 하였다. 민족 간 차이를 비교하기 위한 대상으로 사용한 중국인 남성 60명, 일본인 남성 60명 시료는 국립보건의원에서 공식적인 절차를 통해 분양 받았고(KOBB-2009-02), 이외 아프리카 아메리카인(19명), 코카시안(48명), 멕시코계 미국인(47명) [미국 국적의 백인, 흑인, 히스패닉] 시료들은 Coriell Cell Repositories, USA에서 공식적으로 구입하여 사용하였다. 본 연구에 대한 사항은 서울대학교 의과대학 및 서울대병원 IRB 심의 (IRB No: C-0808-056-254)를 거쳤다. 시료들에 대해 Hoefer DyNA Quant 300 (Hoefer Pharmacia Biotech, San Francisco, CA, USA)을 이용하여 DNA 농도를 측정하고, 일괄적으로 멸균수를 이용하여 1 ng/ul로 희석하여 사용하였다.

### 2. 대상 Y-SNP의 선정

2008년 YCC에 의해 발표된 하플로그룹수에 근거하여 주요 20개 하플로그룹(A-T)을 구분할 수 있도록 대상 Y-SNP를 선정하였다. 동아시아에서 흔히 관찰되는 하플로그룹을 좀 더 세부적으로 구분할 수 있도록 대상 Y-SNP를 추가하였다. 한국인에서 빈번하다고 알려진 O (M175) 그룹, C (M130), D (M174) 그룹 들을 구분할 수 있도록 대상 SNP를 선정하였다. 특히 제일 높은 빈도를 보이는 O 그룹을 O1 (M119), O2 (O2a-M95, O2b-M176), O3 나아가 O3 (O3a3b-M7), O3 (O3a3c-M134) 등으로 세분화할 수 있도록 Y-SNP 선정에 주의를 기울였다. 이렇게 선정된 유전자의 수는 모두 30개였다 (Fig. 1).

### 3. 시발체 설계

30쌍의 시발체는 모두 A, B C, D, E-5개 세트의 조합으로 나누어 다중증폭체계를 구성하였다. 계통 데이터베이스

(dbSNP or BLAST option)에서 대상 SNP 염기서열을 확인하고, Primer 3.0 Program V. 0.2 (<http://www-genome.wi.mit.edu/>)를 이용하여 시발체를 설계하였으며,<sup>13)</sup> Henegariu 등<sup>14)</sup>과 Markoulatos 등<sup>15)</sup>이 권유하는 사항을 따르면서 다중증폭이 적절하게 진행될 수 있도록 고안하였다. 분해된 시료에도 적용이 가능하게 하기 위하여 모든 유전자의 증폭산물을 200 bp보다 짧게 고안하였다(Table 1). 다중중합효소연쇄반응-E 방법은 O그룹 빈도가 높은 것으로 확인된 한국인 및 기타 동아시아인 2개 민족에 대하여서만 실행하였다.

#### 4. 대상 SNP 증폭 및 SBE (Single Base Extension)를 이용한 유전자형 확인

각 유전자에 대해 고안된 시발체를 이용하여 먼저 단일 대상 SNP의 증폭여부를 확인하고, 증폭 산물에 대해 BigDye®

Terminator V1.1 Cycle sequencing Kit (Applied Biosystems, Foster City, CA, USA) 및 자동염기서열분석기를 사용하여 염기서열을 확인하였다. 이후 이들 시발체를 조합하여 여러 유전자들을 한꺼번에 증폭하는 다중증폭 반응조건을 완성하였다(Table 1). 30쌍의 시발체를 5개 다중증폭으로 조합하였으며 최종 증폭조건은 다음과 같다. PCR반응의 총 볼륨은 25  $\mu$ l로써 반응에는 2.5 ng DNA, 10x\*PCR buffer 2.5  $\mu$ l, 25 mM  $MgCl_2$  2.0~2.5  $\mu$ l, 250 uM dNTP 4  $\mu$ l, 5 uM 시발체 0.36~7.2  $\mu$ l, 5U AmpliTag Gold DNA polymerase (Applied Biosystems, Foster City, CA, USA) 0.4  $\mu$ l와 20 mg/ml BSA 0.4  $\mu$ l를 첨부하였다. 반응조건은 94℃에서 5분 변성, 증폭순환은 95℃에서 30초 → 58℃에서 30초 → 65℃에서 30초로 33 cycles를 반응하였으며 최종 연장은 65℃에서 7분 반응하였다. 반응한 15  $\mu$ l 증폭산물에 5 Units의 SAP (USB, Cleveland, OH, USA)와 2 Unit의 Exo I (Fermentas, Glen Burnie, MD,

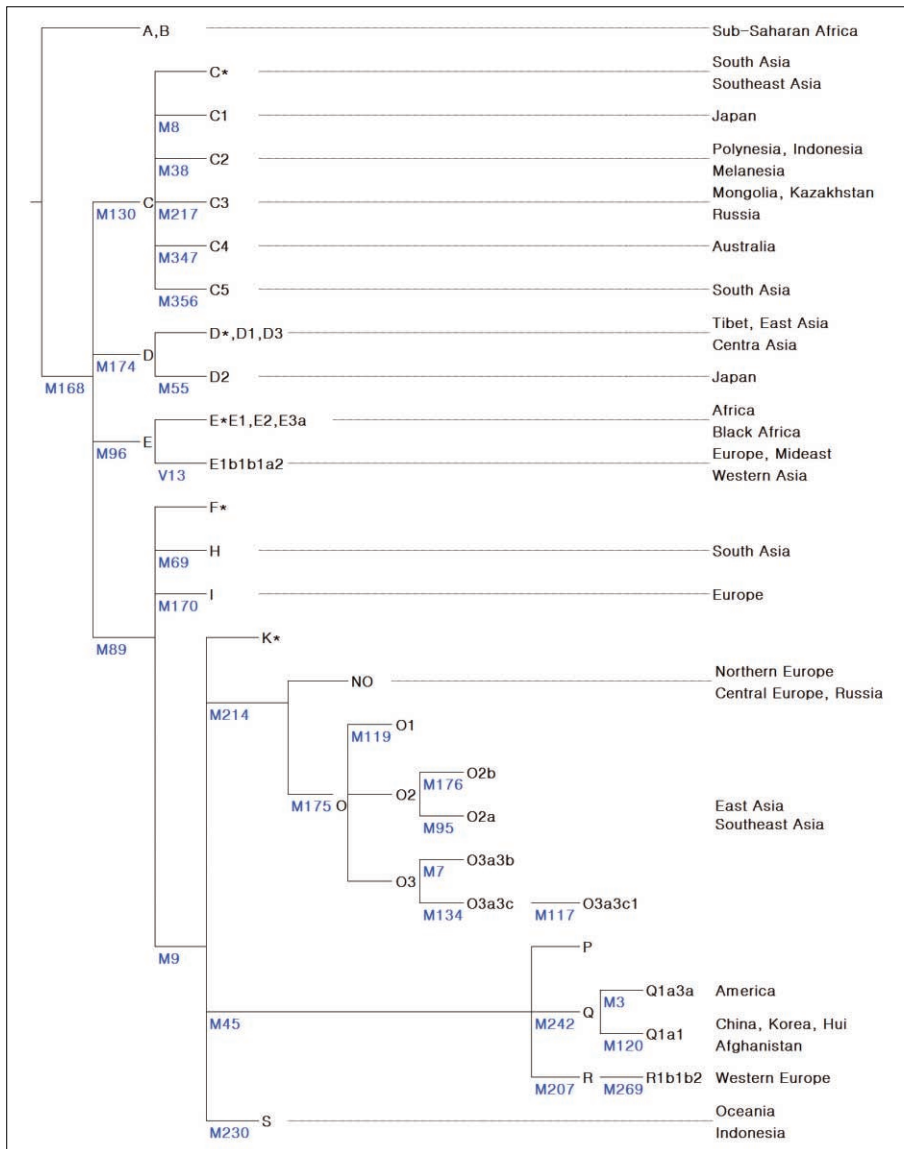


Fig. 1. Selected 30 Y-SNPs with phylogenetic tree used in this study for Y-haplogrouping. Several SNPs were selected to sub-classify haplogroup O.

USA)를 첨부하여 37℃에서 1시간 배양 후 80℃에서 20분 동안 효소 불활성화를 실행하였다.

염기서열 변이 확인을 위한 SBE의 시발체 고안을 위해 Primo SNP 3.4 (<http://www.changbioscience.com/primo/primosnp.html>)를 사용하였고, SBE산물을 자동염기서열분석기를 이용하여 분석함을 고려하여 서로 다른 길이로 구성된 시발체가 형성되도록 시발체는 특이서열을 보합서열로 인간 게놈에 존재하지 않는 염기서열-AACTGACTAACTAGG TGCCACGTCGTGAAAGTCTGACAA와 Poly C-를 이용하여 길이 차이를 구성하였다(Table 2). SBE반응은 SNaPshot® Primer Focus™ Kit (Applied Biosystems, Foster City, CA, USA)를 사용하여 제작자가 권유하는 방법을 활용하여 진행하였고, 결과는 ABI Prism® 3130 genetic Analyzer (Applied Biosystems, Foster City, CA, USA)를 이용하여 분석하였다. (Fig. 2).

## 5. 여러 민족 시료에 대한 결과 확인 및 통계적 접근

연구에 사용된 시료들이 각 민족을 적절하게 대표할 수 있는지를 확인하기 위해 한국인 시료와 기타 동아시아 2개 민족 시료들에 대해 상염색체 검사를 시행하고, 각각의 유전자가 하디-와인버그평형 상태에 있는지를 확인하여 보았다. 나아가 Y 염색체 유전자가 편향되게 분포하는지 여부를 알아보기 위하여 Y-STR 유전자를 분석하고 대립유전자의 수가 적은 유전자좌에서 각 유전자형의 빈도를 기준에 발표한 데이터와 비교함으로써 본 연구에 활용된 시료의 민족 대표성을 확인하였다.<sup>16-18)</sup>

각 민족에서 나타나는 하플로그룹의 빈도를 비교하여 보았다. 한편 구성된 체계가 어느 만큼의 구분력을 가지고 있는지를 수치로써 표현하기 위하여 다범주 로짓모형을 이용한 로지스틱 회귀분석 방법을 이용하여 한국, 일본, 중국 민족 예측 확

**Table 1.** Primer Sequences for Amplification of Each 30 Y-SNPs. The Amplification was Done Through 5 Separate Multiplex System

Marker	SNP	PCR Primer Sequence (5' → 3' )		Conc ( $\mu$ M)	Size (bp)
		Forward	Reverse		
Multiplex-A					
M175	_/TTCTC	GATTTAAACTCTCTGAATCAGGCACAT	TTCTACTGATACCTTTGTTTCTGTTTCATTC	0.05	79
M170	A/C	CAGCTCTTATTAAGTTATGTTTTCATATTCTGTG	GTCCTCATTTTACAGTGAGACACAAC	0.05	119
M347	A/G	CCTGAATTGGGTCACTGCTTCCCA	ACATGGCTGCCCTTGGACGA	0.02	125
M9	C/G	CAGGACCCTGAAATACAGAACT	AGCGCTACCTTACTTACATAAC	0.04	128
M89	C/T	TGGATTGAGCTCTCTTCCTAAGGTTAT	CTGCTCAGGTACACACAGAGTATCA	0.05	135
M45	G/A	GAGAGAGGATATCAAAAATTGGCAGT	TGACAGTGGCACCAAAGGTC	0.04	138
M96	G/C	TGCCCTCTCACAGAGCACTT	CCACCCACTTTGTTGCTTTG	0.04	143
M174	T/C	CCGTCACAGCAAAAATGTA	AAAAGGAGAAGGACAAGACC	0.05	178
Mutiplex-B					
M207	A/G	GAAGTATCCCTGAAGAAGGAA	TCACTTCAACCTCTTGTTGG	0.05	120
M55	T/C	ATACAAATAGGTGGGGCAAG	CTGAATCCTAATGGCTGTTT	0.08	133
M269	T/C	CCTAGCCTCATTCTCTAAAA	TCTTTTGTGTGCCTTCTGAG	0.19	148
M214	T/C	AGGCTGATTTTGCTGCTGAT	GTGGTGAGCCATGGAAATG	0.11	156
M242	C/T	CTACGGCATAGAAAAGTTTGTG	GGGCTTTCAGCATAATACCT	0.03	160
M38	T/G	CCCTAGGGTATGGCAATGGT	TCTCCACAAAACAGCTGGAG	0.06	174
Multiplex-C					
M3	C/T	ATGTGGCCAAGTTTATCTG	AGGGCATCTTTCATTTTAGG	0.02	112
M130	C/T	TCTCCTCCCTTCTCTTCTGTA	CCAGCCTCTTATCTCTCTCTT	0.04	128
M356	C/G	CCTGCTCAGTTGTTGTCATTA	GGAACCAAGTGCTCCTTATAGT	0.15	158
M217	A/C	GGCCAGTATCTCCAAAATCC	AAAGCTGCTGTGGCTTTCAT	0.02	180
M8	G/T	TGCCTCAGTATCACAAACAC	AGACAAGTCCACCAAACAGATG	0.12	189
Multiplex-D					
V13	G/A	TGAGATAGCACAAACAGTGGA	AATCCCATCTCAATCCCTTA	0.08	91
M168	C/T	GCTTGGAGATAATTCTGGTG	ACATCTCTTACCCAAACTGC	0.06	100
M69	T/C	GTAGCCTGTTCAAATCCAAA	CCCTTAGCTCTCCTGTTTTT	0.16	110
M230	T/A	TACATGGCCATACTTAACCT	TCTTCATTGCTAGCTATTTT	0.63	131
M120	T/C	CACTGCTAATGATCCGTTTT	ATAGCAATGTAGGGGGGATTT	0.34	166
Multiplex-E					
M134	G/-	AGAATCATCAAACCCAGAAGG	GAGAATTCGCATTTACCACTGT	0.16	105
M7	C/G	GCATGTAATCATTCTCTCT	TTGTCCCTGCAGCCTTGTA	0.03	111
M119	A/C	TCCAAACCGCAGTGCTATGTGT	GATGTAGAAAAATGTTATGGGT	0.05	112
M95	C/T	AATAGTGTTGCACCTTCTTGG	TGGACTCTCTAGACCTACTG	0.06	115
M117	ATCT/-	TCCTATAGAAGCAAAGATAATG	AGAGCTTATTAGATGATAGAAAAAC	0.16	125
M176	C/T	TACAGGCCATGCACAGAGAG	CATTCTTGAGTGTGTGGGTTT	0.03	181

를 분석하였다.

## 결 과

### 1. 민족 시료의 SNP 유전자형 확인

구성한 체계를 한국인과 기타 민족 시료에 적용하여 전체 유전자형을 확인하였다. 전체 샘플에서 관측된 하플로그룹은 C1, C3, D2, E, E1b1b1a2, F\*, K\*, NO, O\*, O1, O2a, O2b, O3a3c, O3a3c1, O3a3b, Q1a1, Q1a3a, I, P, R, R1b1b2이다.

#### 가) 대립유전자 분포빈도

얻은 유전자형 데이터로부터 각 민족별로 나타내는 유전자형의 빈도는 Table 3과 같다.

#### 나) 한국인 하플로그룹 분포

한국인 100명 시료에서 모두 10개의 하플로그룹이 확인 되었다. 하플로그룹 O가 87%로써 제일 많은 분포를 나타냈다. 이 그룹에 대한 세부적인 하플로그룹 분포에서 O2b가 제일 높은 퍼센트를 차지하였다. 구체적인 분포로는 O\*-M175, O1-M119, O2b-M176, O3a3c-M134, O3a3c1-M117, O3a3b-M7이 전체 O그룹에서 각각 28%, 7%, 45%, 4%, 15%, 1%를 차지하였다(Table 4). 하플로그룹 O를 제외한 나머지 하플로그룹으로는 NO-M204, D2-M55, C3-M217, Q1a1-M120이 각각 2%, 2%, 8%, 1%로 분석되었다.

#### 다) 기타 민족 하플로그룹 분포

##### (1) 다른 2개 동아시아 민족의 하플로그룹 분포

일본인, 중국인 역시 한국인과 유사하게 하플로그룹 O가 65%, 78%로 제일 높은 빈도를 보였다. 일본인은 세부 하플로그룹에서 O2b가 제일 높은 빈도를 나타냈으며, 중국인은

**Table 2.** SBE Primers for Genotyping Each SNP. Genotyping was Done Through Two Separate SBE Reactions, One for Macro-haplogrouping and the Other for O Sub-haplogrouping

SBE Primer-A			
Marker	Size	Primer Sequence (5' → 3')	Conc (μM)
M175	28	GTCTGACAAACATGCCTTCTCACTTCTC	0.03
M356	28	CAAGGCTTGGTATAGATATACAGATTTT	0.02
M168	32	GAAAGTCTGACAAGTATGTGTTGGAGGTGAGT	0.02
M45	39	ACGTCGTGAAAGTCTGACAACCTCAGAAGGAGCTTTTTTCG	0.03
M347	44	GTGCCACGTCGTGAAAGTCTGACAATGAGGTTTTATGGACTTC	0.05
V13	49	TAGGTGCCACGTCGTGAAAGTCTGACAAATTTGGCTCAAACCTCCCTTG	0.03
M120	50	TAGGTGCCACGTCGTGAAAGTCTGACAAATTTCCCTTAAAAACAGCATG	0.03
M89	56	ACTAACTAGGTGCCACGTCGTGAAAGTCTGACAAAACCTCAGGCAAAGTGAGAGAT	0.02
M130	56	GACTAACTAGGTGCCACGTCGTGAAAGTCTGACAAGGCAATAAACCTTGGAATTC	0.06
M214	62	CCCAACTGACTAACTAGGTGCCACGTCGTGAAGTCTGACAAGACACTGTCTGAAAACAAC	0.05
M269	62	CCAAGTACTAACTAGGTGCCACGTCGTGAAAGTCTGACAATGATCAGGGTTTGGTTAAT	0.16
M230	68	(C) <sub>2</sub> -AACTGACTAACTAGGTGCCACGTCGTGAAAGTATGACAACCTTATTTTTATTATCACATTATA	0.13
M9	68	(C) <sub>9</sub> -AACTGACTAACTAGGTGCCACGTCGTGAAAGTCTGACAACGGCCTAAGATGGTTGAAT	0.04
M55	74	(C) <sub>15</sub> -AACTGACTAACTAGGTGCCACGTCGTGAAAGTCTGACAACCTGGATGACTGATGAAAT	0.03
M207	74	(C) <sub>14</sub> -AACTGACTAACTAGGTGCCACGTCGTGAAAGTCTGACAAAAGGTATTGTTATTCTCTTT	0.08
M38	80	(C) <sub>19</sub> -AACTGACTAACTAGGTGCCACGTCGTGAAAGTCTGACAAGGGAATGCTTACTGAATAAAA	0.09
M170	80	(C) <sub>21</sub> -AACTGACTAACTAGGTGCCACGTCGTGAAAGTCTGACAACAACCCACACTGAAAAAAA	0.05
M217	84	(C) <sub>25</sub> -AACTGACTAACTAGGTGCCACGTCGTGAAAGTCTGACAATGAAAAAGTTGGGTGACAC	0.02
M8	84	(C) <sub>25</sub> -AACTGACTAACTAGGTGCCACGTCGTGAAAGTCTGACAAATAACTTGGACTGGGTCA	0.16
M96	88	(C) <sub>27</sub> -AACTGACTAACTAGGTGCCACGTCGTGAAAGTCTGACAAGAAAAGAGGTCTCTCATAATA	0.09
M174	92	(C) <sub>32</sub> -AACTGACTAACTAGGTGCCACGTCGTGAAAGTCTGACAAGAATACCTTCTGGAGTGCCC	0.03
M242	96	(C) <sub>37</sub> -AACTGACTAACTAGGTGCCACGTCGTGAAAGTCTGACAAAAAGGTGACCAAGGTGCT	0.06
M69	100	(C) <sub>41</sub> -AACTGACTAACTAGGTGCCACGTCGTGAAAGTCTGACAAGCTGTTTACACTCTGAAA	0.14
M3	104	(C) <sub>45</sub> -AACTGACTAACTAGGTGCCACGTCGTGAAAGTCTGACAAGGGTCACCTCTGGGACTGA	0.12
SBE Primer-B			
Marker	Size	Primer Sequence	Conc (μM)
M7	25	CAATAGTTGAGTTACTGTTCTTCTT	0.10
M176	32	TGAAAGTCTGACAAGTTGTCCAGTTGCACCTC	0.11
M117	39	ACGTCGTGAAAGTCTGACAAACCAAAGGAATGCACATCT	0.20
M95	44	TGCCACGTCGTGAAAGTCTGACAAGAAAGACTACCATATTAGTG	0.49
M119	80	(C) <sub>21</sub> -AACTGACTAACTAGGTGCCACGTCGTGAAAGTCTGACAATCCAATTCAGCATACAGGC	0.24
M134	92	(C) <sub>33</sub> -AACTGACTAACTAGGTGCCACGTCGTGAAAGTCTGACAAAAAGGCCAGGAAAGTAT	0.24

O\*M175가 40%로 제일 많은 퍼센트를 차지하였다.

## (2) 서양인 하플로그룹 분포

114명의 서양인 3개 민족, 아프리카 아메리카인, 코카시안, 멕시코계 미국인에 대한 하플로그룹을 분석 하였을 때 모두 9 개의 하플로그룹이 관측 되었다. 이 3개 민족에서 얻은 하플로그룹은 한국인을 포함한 3개 동아시아 민족과 완전히 다른 하플로그룹이었다(Table 3).

## 2. 유전학적 통계수치를 이용한 민족간 비교

연구에 사용된 시료들에 대한 상염색체 검사 결과 시료들은

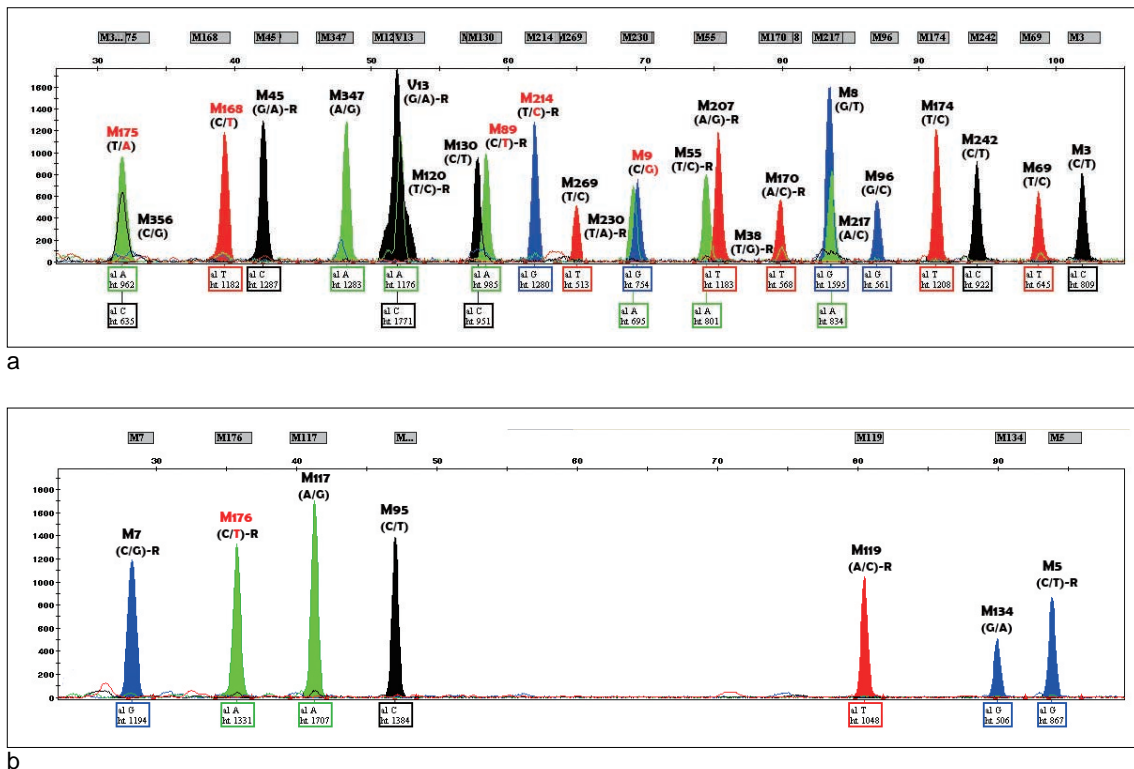
하디-와인버그평형 상태에서 벗어나지 않았다(구체적인 자료는 제시하지 않음).

가) Pairwise genetic distances-Fst값으로부터 확인된 차이  
한국인의 하플로그룹 분포 양상이 기타 민족과 비교하여 어느 만큼의 유전학적 차이가 있는지를 확인하기 위하여 유전학적 거리를 통하여 비교하였다. 동아시아 3개 민족은 기타 3개 민족과 현저한 차이를 보였다. 한국인은 다른 2개 민족-중국인, 일본인과의 Fst값이 각각 0.01155, 0.11127로써 가까운 수치를 보였으나 기타 3개 민족의 Fst값과 비교하여 많은 차이를 보였다. 한국인과 아프리카 아메리카인, 코카시안, 멕시코계 아메리카인과의 Fst값은 각각 0.68834, 0.55257, 0.57500이

**Table 3.** Observed Haplogroup Frequency in Each Population without Sub-classification of Haplogroup O. Haplogroups Observed in Korean were not the Same with that of with African-American, Caucasian and Mexican-American

	C1	C3	D2	NO	O	Q1a1	E	E1b1b1a2	F*	I	K*	P	Q1a3a	R	R1b1b2
Korean		0.08	0.02	0.02	0.87	0.01									
Chinese		0.10	0.02	0.10	0.78										
Japanese	0.03	0.08	0.24		0.65										
AF-AM							0.63							0.05	0.32
Caucasian							0.12	0.02	0.19	0.08		0.02		0.15	0.42
ME-AM							0.05	0.02	0.06	0.06	0.06		0.21	0.05	0.49

AF-AM: African-American ME-AM: Mexican-American



**Fig. 2.** Electropherograms of SBE reaction of two sets. (a). SBE-A set show five markers with derived state - M168, M89, M9, M214, M175. (b). SBE-B set show one marker with derived state - M176. The SNPs M168, M89, M9, M214, M175, M176 show the derived state, which leads to the assignment of the sample to haplogroup O2b with the final marker M176.



었다(Table 5). 전 대륙을 대상으로 구성된 24개 SNP를 이용하였을 경우에는 한국인은 중국인과 가까운 Fst값을 보여 주었으나 O그룹의 세부 하플로그룹들을 이용하여 분석한 결과는 중국인보다 일본인과 가까운 Fst값-0.01757(중국인-0.07269)을 보여 주었다(Table 6).

나) 한국인과 기타 동아시아 2개 민족에 대한 체계의 구분력  
확인된 하플로그룹들로부터 한국인은 서양인 3개 민족과는 구분이 가능함을 알 수 있다. 그러나 한국인과 기타 동아시아 2개 민족 사이에는 비슷한 하플로그룹 분포를 이루고 있어 동아시아 3개 민족 사이에는 구분이 잘 되지 않았다.

선정된 30개 Y-SNP를 통하여 동아시아 3개 민족에 대하여 어느 만큼의 식별력을 가지는지 수치로써 나타내기 위해, 여러 가지 통계방법을 시도한 결과 로지스틱 회귀분석(logistic regression)이 응용된 형태인, 결과변수가 다항일 때 적용 가능한 다범주 로짓모형을 이용한 로지스틱 회귀분석방법을 사용하였다. 한국인과 중국인, 일본인에서 확인된 12개 하플로그룹을 이용하여 먼저 세 민족을 구분함에 있어서 차이를 보여주는 하플로그룹들을 전진선택법(forward selection)으로 선택한 결과 하플로그룹 C1, D2, O2b가 선택되었다. 다음 3개 하플로그룹을 이용하여 한국인, 중국인, 일본인에 대한 예측 확률을 분석하였다. 동일한 민족으로 예측할 확률, 즉 다시 말하면 한국인을 한국인으로, 중국인을 중국인으로, 일본인을 일본인으로 정확하게 예측 할 확률은 36.7%였다. 한편 다른 민족으로 예측 할 확률, 다시 말하면 한국인을 일본인이나 중국인으로 예측 할 확률, 일본인을 중국인이나 한국인으로 예측 할 확률은 20.4%였다. 이외 동일 민족인지 다른 민족인지를 구분할 수 없는 확률은 42.9%였다(Table 7).

## 고 찰

각 SNP 유전자형을 분석하기 위하여 본 연구에서는 SNaPshot™ 방법을 이용하였다. 이 방법은 실험과정이 간편한 반면, 아래와 같은 결점을 보여 주었다. 첫째, 큰 길이의 시발체 합성이 필요하였다. 같은 형광염료가 붙은 증폭산물은 산물의 서로 다른 크기에 의하여 식별하게 된다. 그러므로 한 번에 많은 수의 SNP를 증폭하기 위하여서는 큰 길이의 시발체가 필요하였다. 길이가 길어짐에 따라 시발체간의 보합 가능성도 커지게 된다. 본 연구에서도 2개 시발체에서 이런 현상이 일어나 짧은 길이로 다시 디자인 하는 과정이 필요하였다. 한편 60 mer 이상의 시발체부터는 합성이 어려워지며 비용이 증가하는 동시에 합성된 양은 적게 된다. 그리고 고순도로 정제되지 않은 시발체는 n-1, n-2, n-3…… 등과 같은 연속적인 꼬리현상의 피크를 보였다. 이런 꼬리현상은 흔히 사용하는 정제방법인 HPLC (high performance liquid chromatography)를 사용하여도 피할 수 없었다. 이와 비교하여 Ultramers™ technology (IDT, Coralville, IA, USA)를 이용하여 합성된 시발체는 104 mer 크기에서도, Ultra PAGE정제를 실행하지 않아도 꼬리현상은 나타나지 않았다. 또한 Ultramers™ technology는 200 mer 길이의 시발체도 합성할 수 있어 더 많은 수의 SBE실험

**Table 6.** Fst was Calculated Through 6 Y-SNPs Using AMOVA. 6 Y-SNPs were M119, M95, M176, M134, M117, M7. Korean was More Closer to Japanese than to Chinese

	Korean	Chinese	Japanese
Korean	0.00000		
Chinese	0.07269	0.00000	
Japanese	0.01757	0.018502	0.00000

**Table 4.** Different Sub-haplogrouping for Haplogroup O in Three Northeast Asia Populations. In Korean, Sub-haplogroup O\*, O2b and O3a3c1 are the Major Sub-haplogroups

	O*	O1	O2a	O2b	O3a3c	O3a3c1	O3a3b
Korean	28	7	0	45	4	15	1
Chinese	40	2	2	17	4	34	0
Japanese	26	0	0	62	5	8	0

(Unit : %)

**Table 5.** Fst was Calculated Through 24 Y-SNPs Using AMOVA. Korean was the Most Closest to Chinese

	Korean	Chinese	Japanese	AF-AM	Caucasian	ME-AM
Korean	0.00000					
Chinese	0.01155	0.00000				
Japanese	0.11127	0.06391	0.00000			
AF-AM	0.68834	0.57547	0.47533	0.00000		
Caucasian	0.55257	0.44287	0.36214	0.16087	0.00000	
ME-AM	0.57500	0.46707	0.38588	0.22802	0.03722	0.00000

AF-AM: Afrian-American, ME-AM: Mexican-American

**Table 7.** Influential Haplogroups and Prediction Probabilities for Three Northeast Asia Population

		C1	D2	O2b
Odds ratio	Point estimate	<0.001	0.065	0.928
	95% Wald confidence limits	<0.001, 999.999	0.018, 0.230	0.539, 1.597
Predicted probabilities	Concordant(predicted same population)		36.7%	
	Discordant(predicted different population)		20.4%	
	Tied (unknown)		42.9%	
	Discrimination Capacity		0.582	

Using forward selection method, haplogroup C1, D2, O2b selected as influential haplogroups to predict each Northeast Asia population. According to logistic regression, the possibility to designate each Northeast Asia population concordancy was low using above there haplogroups.

가능성을 보여준다. 둘째, 반응에 참여하지 못한 시발체에 대한 정제가 완전하지 못하다. 다중증합효소연쇄반응을 실행한 후의 증폭산물은 2개의 효소를 이용하여 정제과정을 진행한다. 하나는 Exo I(*Escherichia coli* exonuclease I: 3' → 5' 단 일가닥 핵산말단 가수분해효소)로써 반응에 참여하지 않은 잔여 단일가닥 시발체를 제거하며 다른 하나는 SAP (shrimp alkaline phosphatase: 유기산과 에스테르, 폴리인산을 가수분해하는 효소)로써 반응 중 주형에 결합되지 않은 dNTP를 제거한다. Exo I에 의한 잔여 시발체 제거가 완전하게 이루어지지 않으면 전기영동결과에서 20–30 bp 위치에서 비특이적 피크가 발생하여 결과분석이 어려웠다. 이런 영향을 피하기 위하여 본 연구에서는 제일 짧은 시발체의 길이를 25 mer 이상으로 디자인함으로써 유전자 실제 피크에 대한 판단이 가능하였다. 결과분석에 있어서 또 하나의 비특이적 피크로는 70 bp 위치에서 나타나는 녹색 형광피크이다. 이 피크는 SAP가 소화시키지 못한 ddATPs에 의하여 나타나는 피크로써 이를 완전히 제거하기는 어려웠다. 그리하여 본 연구에서는 여러 가지 방법을 결합하여 비특이적 피크의 높이를 최대한으로 줄임으로써 분석 기준인 200 RFU보다 낮게 나타나게 하려고 노력하였다. 우선 시발체의 디자인에서 70 bp 위치에서 비특이적 피크가 생성되지 않도록 하였으며 신선한 SAP를 사용함과 동시에 효소 불활성화과정의 온도를 75℃에서 80℃로 높였으며 반응시간은 15분에서 20분으로 늘였다.<sup>19)</sup> 모든 반응물과 증폭산물은 아이스박스를 사용하였고 반응산물은 가능한 빨리 분석단계까지 진행하도록 하였으며 부득이한 경우에는 냉동 보관하였다. 냉동보관시간도 24시간을 초과하지 않게끔 하였다. 셋째, 변이에 의한 각 유전자의 증폭효율이 서로 다르기에 증폭산물의 균형을 조절하기 어려웠다. 일부 SNP에서는 두 유전자형의 피크 비례가 2배, 3배나 되는 차이를 보이기도 하였다. 이에 대한 해결방법은 아직 보고되어 있지 않으며 본 연구에서도 해결하기 어려웠던 점이었다. 이 현상은 다중증폭에 의한 불균형적인 증폭효율에 의하여 산생하는 것으로써 TagMan® real-time PCR assays, Sequenom® iPLEX™ MALDI-TOF spectrometry 등 SNP 관련 분석방법에서도 나타난다고 한다.<sup>20)</sup> 다만 이와 같은 문제점들이 유전자형 판단에는 영향을 미

치지 않았다.

구성된 체계를 이용하여 한국인에 대한 하플로그룹 C, D, O의 세부 하플로그룹을 분석 할 수 있다. Ahn 등<sup>21)</sup>과 Jin 등<sup>22)</sup>의 연구에 의하면 한국인에서 주요한 하플로그룹 O의 세부 하플로그룹 가운데서도 O3가 제일 많은 분포를 차지하였고 본 연구 결과도 이와는 크게 다르지는 않다. 본 연구에서는 더 나아가 O3그룹에 대한 세부 하플로그룹인 O3a3c, O3a3c1, O3a3b에 대해서까지도 분포 양상을 분석할 수 있었다. 또한 15~20%를 차지하였던 하플로그룹 C, 2% 정도로 관찰된 하플로그룹 D에 대하여서도 세부 하플로그룹의 분포 분석이 가능하였다.

한국인에서 관찰된 하플로그룹으로는 O, C, D, NO, Q가 있다. 이 가운데 흔한 분포를 가지는 하플로그룹은 O와 C이다. 하플로그룹 O는 C에 비해 압도적으로 빈번하였고, 이러한 전반적인 분포는 최근 발표된 다른 연구결과와 크게 다르지는 않았다.<sup>11)</sup> 한편 새로운 SNP 검사에 따라 좀 더 세분화된 하플로그룹 구분이 가능할 수도 있겠다.<sup>23)</sup> 다만 세부적인 하플로그룹 비교에 있어 목적에 맞게 세분화의 정도에 대한 고민이 필요할 수도 있겠다. 한국인은 중국인보다 하플로그룹 NO의 분포가 적었다. 본 연구에서는 일본인에서 하플로그룹 NO가 관측되지 않았지만 일본인에서도 한국인과 비슷하게 2%정도로 관측되었다고 한다.<sup>24)</sup> 세분화하지 않은, 대략적인 하플로그룹 수준에서만 볼 때 한국인은 중국인과 비슷한 하플로그룹을 가지고 있다. 이는 전 대륙을 대상으로 한, 24개 SNP를 이용하여 얻은 Fst값에서도 같은 결과를 나타냈다.

세부적으로 하플로그룹을 분석하면 한국인에서 관측되었던 세부 하플로그룹으로는 C3, D2, O\*, O1, O2b, O3a3c, O3a3c1, O3a3b, Q1a1이 관찰되었다. 다른 지역 민족에 비해 한국인에서만 특이하게 나타나는 세부 하플로그룹은 보이지 않았다. 동일한 하플로그룹이라도 지역마다 가지는 세부 하플로그룹은 다르다. 본 연구에서 얻은 한국인에 대한 세부 하플로그룹 정보로부터 기타 지역과의 분포 차이를 비교할 수 있었다. 하플로그룹 C인 경우 일본인(10%미만)에서만 나타나는 세부 하플로그룹 C1은 한국인뿐만 아니라 중국인에서도 볼 수 없다고 하였다.<sup>24, 25)</sup> 대양주 주변 지역에서 일정한 분포를 가지



는 세부 하플로그룹 C2도 마찬가지로 한국인에서는 관찰되지 않았다. 한국인에서 확인된 C3그룹은 중국 서북으로부터 중앙아시아, 나아가 중앙아시아 북쪽 지역에서 25~84%의 분포를 가진다고 하였다.<sup>25)</sup>

하플로그룹 D의 경우, 한국인을 대상으로 진행된 기존의 연구에서는 한국인에서 DE-YAP이나 D-M174가 2%, 2.5%로 확인되었다. 본 연구에서도 2%로 분석되었지만 D그룹에 대한 세부 하플로그룹을 측정한 결과 D그룹의 세부 하플로그룹 가운데서 D2임이 확인되었다. 같은 D그룹은 티베트와 아시아 서북쪽 지역에서도 보이는데 주로는 D1그룹이나 D\*그룹이라고 하였다.<sup>26)</sup>

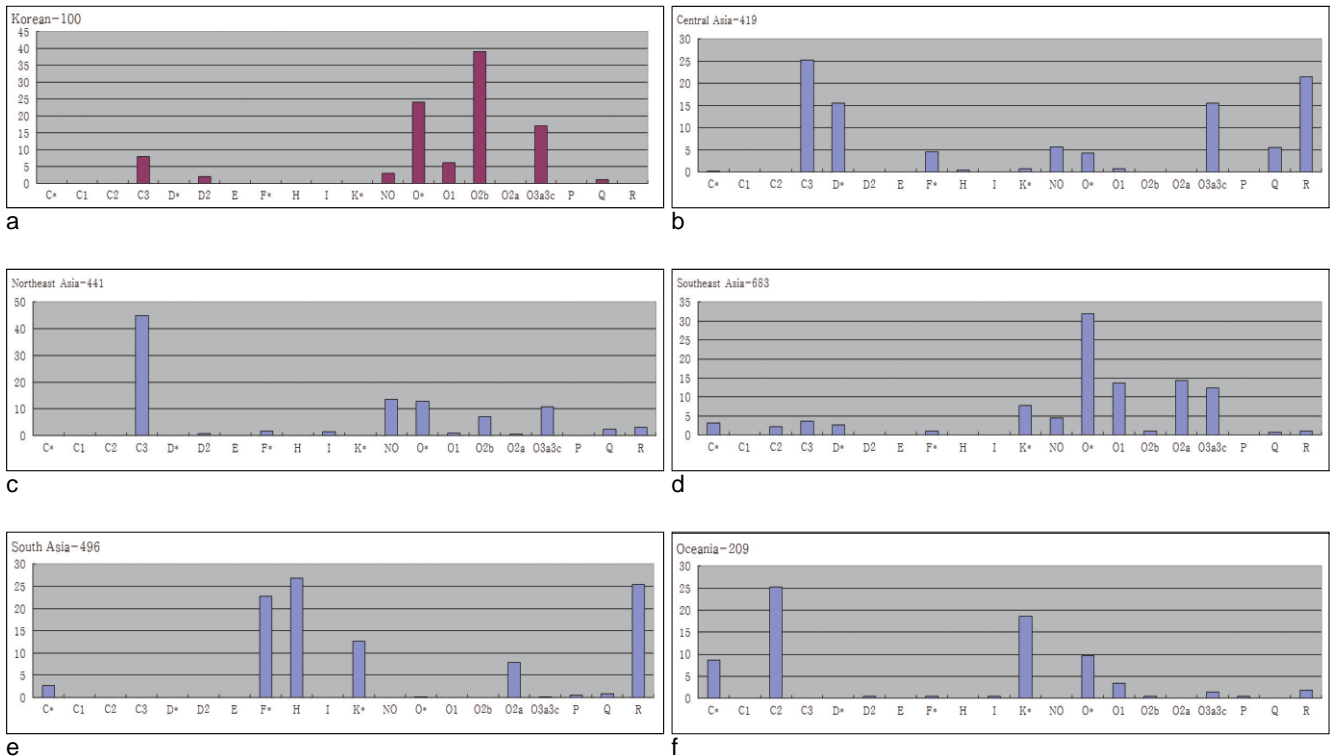
한국인에서만 특이하게 나타나는 세부 하플로그룹은 관찰되지 않았지만 기타 지역과 분포 차이는 존재하였다. 한국인에서 관측된 전체 세부 하플로그룹 가운데서 O2b가 제일 많은 분포를 가진다. 한국인과 일본인인 경우 O그룹의 세부 하플로그룹 가운데서 제일 많은 비율을 차지하는 하플로그룹이었다. 한국인에서의 분포는 27~40%로써 일본인(7~30%)에 비해 더 많은 분포를 가진다. O2b는 한국, 일본, 중국 동북에서 보이는 하플로그룹으로써 비슷한 서북 분포형 하플로그룹인 C, D와 다르게 중앙아시아 지역에서는 보이지 않는다. 또한 중국 남부와 오세아니아 지역에서는 극히 낮은 분포를 가지며, 그 사이에 있는 동남아시아-베트남과 인도네시아 지역에서는 일정한 빈

도를 가진다(Fig. 3).<sup>24)</sup>

O2b 다음으로 한국인과 중국인, 일본인 사이에서 다형성을 보이는 하플로그룹은 O1a-M119와 O2a-M95였다. 두 하플로그룹은 모두 남방계 마커로써 북에서 남으로 갈수록 분포가 증가한 추세를 가진다. 이 가운데서 한국인에서 보이지 않았던 하플로그룹은 O2a이다. 이는 한국인뿐만 아니라 일본인, 중국 북방 지역에서도 분석되지 않았지만 중국 남부와 동남아시아에서는 10~60%의 분포를 가진다. 중국인에서 적은 분포를 보이지만 높은 빈도를 가지는 동남아 민족과 구분하기에는 유용한 SNP이다.

세부 하플로그룹 O2b 다음으로 한국인에서 많은 분포를 가진 세부 하플로그룹은 O\*와 O3a3c1이다. 세부 하플로그룹 O3a3c1은 O3그룹의 하위 그룹인 O3a3c의 하위 그룹이다. O3a3c는 동북아시아보다 동남아시아에서 더 많은 분포를 가지는 하플로그룹이다. O3a3c의 세부 하플로그룹에 대한 자료가 부족하여 O3a3c1에 대한 지역 분포 분석은 어려웠다. 본 연구에서는 한국인 O3a3c1(13%)의 분포가 중국인(26.7%)보다 적었다.

하플로그룹 O의 3개 세부 하플로그룹 가운데서 O3를 정의하는 SNP는 M122이다. O3-M122는 동아시아와 동남아시아에서 비교적 높은 분포를 차지하는 그룹으로써 중국 북부와 남부에서 50%이상의 분포를 가지며 동남아시아에서도 일정한



**Fig. 3.** Comparison of our Koreans data with that of previously reported for other populations. Other population data were from published studies. Korean have different haplogroup distribution especially in sub-haplogroup C, D, O from Central Asia(D\*), Northeast Asia(C3, O2b) and Oceania population(C\*, C2, K\*).

a: our data b, c, d, e, f: previously reported data.

분포를 가지고 있다. 그러나 본 연구에서는 M122 SNP를 선택하지 않았다. 본 연구는 민족을 구분하려는 용도로 집단간의 분포 차이점을 나타내는 SNP를 선택하여야 한다. Shi 등<sup>28)</sup>의 연구 결과에 의하여 본 연구체계에서는 O3그룹의 세부 하플로그룹 가운데서 주요한 분포를 가지는 O3a3c-M134, O3a3c1-M117을 선택하였다. O3a3c1-M117은 O3a3c-M134의 하위 세부 하플로그룹으로써 중국인에서 O\* 다음으로 높은 빈도를 보였고 한국인과 일본인에서도 O\*과 O2b 다음으로 높은 빈도를 보였다.

O\*의 퍼센트는 하플로그룹 O의 세부 하플로그룹 가운데서 어떤 세부 하플로그룹으로 나누는가에 따라 각 연구결과마다 어느 정도의 차이를 보인다. 본 연구에서는 O3a3c, O3a3b를 제외하곤 나머지 O3의 세부 하플로그룹(M122, M324 등)과 O-M175 등 여러 하플로그룹이 O\*로 분류된다. 이러한 원인으로 O\*에 대한 민족간의 비교분석은 할 수 없지만 동일한 SNP를 이용한 하플로그룹 분석체계에서는 한국인은 중국인보다 낮은 분포를 가진다.

하플로그룹 O의 세부 하플로그룹 O2b1-47z는 사용할 수 없었다. 왜냐하면 O2b1-47z와 같은 SNP는 X 염색체(DXY5X)와 Y 염색체(DXYS5Y)에 모두 존재하여 부계에만 특이하게 증폭이 이루어지기 어렵기 때문이다.<sup>26, 27)</sup> 그리하여 본 연구에서는 대상 Y-SNP 마커 선정시 O2b1-47z를 제외하였으나, 다른 연구에서 O2b1-47z를 포함시킨 연구결과도 볼 수 있었다.<sup>11, 24)</sup>

하플로그룹 O의 세부 하플로그룹에서 한국인은 중국인보다 일본인과 비슷한 하플로그룹 분포 양상을 보였는데 Fst값을 비교하여도 같은 결과를 보였다. 이는 근대에 발생하였다는 한국인과 일본인의 혼합설로 설명이 가능하지 않을까 생각한 다.<sup>26)</sup>

한국인은 중국인, 일본인과 세부 하플로그룹 분포가 차이는 있었지만 전체적으로 볼 때 비슷한 분포 양상을 가진다. 30개 Y-SNP를 이용하여 한국인과 중국인, 일본인과의 구분 정도를 수치화 하였더니 구분력은 0.582로써 구분이 낮았다. 본 연구를 동남아시아 시료에까지 시행하지 못하여 구성 체계로 한국인과 동남아시아 민족간의 구분력을 정확히 계량할 수 없었지만 상염색체상의 변이를 조사한 보고를 참고하면 그 분석이 쉽지는 않을 것이라고 추측된다.<sup>29)</sup> 단순 Y-하플로그룹이나 mtDNA-하플로그룹, 상염색체를 이용하여 민족을 구분하기는 어려우나 여러 가지 요소를 함께 활용한다면 동아시아 북쪽 지역과 동남아시아 지역 민족사이의 구분은 가능할 것으로 예상된다.

한국인을 기타 동아시아인, 동남아시아인과 구분하기는 어려웠지만 구성된 체계를 이용하여 유럽계-하플로그룹 민족과의 구분은 가능하다. 세계 5대륙을 대상으로 분포하는 하플로그룹을 보면 아시아, 아프리카와 유럽 및 아메리카, 오세아니아 등 기타 3개 부분으로 나눌 수 있다. 아시아 지역에서는 동

아시아와 동남아시아가 한 부분으로, 다른 한 부분은 유럽계-하플로그룹을 주요 분포로 가지는 중앙아시아, 서아시아, 남아시아로 분류할 수 있겠다. 즉 다시 말하면 하플로그룹 D2, O그룹에 의하여 동아시아와 동남아시아가 한 부분으로, 기타 아시아 지역과 대륙이 한 부분으로 분류가 될 수 있다. 본 연구에서는 한국인에서 보였던 하플로그룹을 기타 3개 서양인 민족에서는 볼 수 없었으며 중국인, 일본인에서도 마찬가지였다. 동아시아 3개 민족의 80-90%이상이 하플로그룹 D2, O그룹에 속하였으며 기타 서양인 3개 민족에서는 이런 하플로그룹이 전혀 분석되지 않았다.

법의학 영역에서 SNP는 STR에 비해 실험기법이나 연구영역이 늦게 출발하였지만 어쩌면 앞으로의 어느 시기에 STR보다 SNP를 일반적으로 분석하게 될지도 모른다고 학자들은 예측한다.<sup>30)</sup> 또한 변질된 시료에 대한 분석에 있어서 SNP는 유용하다. 30개 Y-SNP로 구성된 본체계는 민족을 구분함에 있어서 한국과 유럽계 하플로그룹을 가지는 민족과는 구분이 가능하나 중국, 일본 등 기타 동아시아 민족 및 동남아 민족과의 구분은 어렵다. 이를 위하여 다양한 지역을 대상으로 O\*, O1, O2a, O2b, O3a3c에 대한 세부 하플로그룹 자료가 필요하며 이런 부계 하플로그룹 자료는 mtDNA나 상염색체와 함께 시료의 민족을 추정함에 있어서 유용한 도구로 쓰일 것이다.

#### Acknowledgments

We are grateful to National Biobank of Korea for providing DNA samples. Two anonymous reviewers are acknowledged for their helpful comments on the manuscript.

#### 참 고 문 헌

1. Lim JT. Fundamental research for foreign crime countermeasure. IOM Migration Research & Training Centre Working Paper 2010;08.
2. Grimes EA, Noake PJ, Dixon L, Urquhart A. Sequence polymorphism in the human melanocortin 1 receptor gene as an indicator of the red hair phenotype. *Forensic Sci Int* 2001;122:124-9.
3. Walsh S, Liu F, Ballantyne KN, van Oven M, Lao O, Kayser M. IrisPlex: a sensitive DNA tool for accurate prediction of blue and brown eye colour in the absence of ancestry information. *Forensic Sci Int Genet* 2011;5:170-80.
4. Walsh S, Lindenbergh A, Zuniga SB, et al. Developmental validation of the IrisPlex system: determination of blue and brown iris colour for forensic intelligence. *Forensic Sci Int Genet* 2011;5:464-71.
5. Walsh S, Wollstein A, Liu F, et al. DNA-based eye colour prediction across Europe with the IrisPlex system. *Forensic Sci Int Genet* 2012;6:330-40.
6. Nassir R, Kosoy R, Tian C, et al. An ancestry informative

- marker set for determining continental origin: validation and extension using human genome diversity panels. *Bmc Genet* 2009;10.
7. Londin ER, Keller MA, Maista C, et al. CoAIMs: A Cost-Effective Panel of Ancestry Informative Markers for Determining Continental Origins. *Plos One* 2010;5: e13443.
  8. Phillips C, Salas A, Sanchez JJ, et al. Inferring ancestral origin using a single multiplex assay of ancestry-informative marker SNPs. *Forensic Sci Int-Gen* 2007;1:273-80.
  9. Butler JM. *Forensic DNA typing. Biology, technology and genetics of STR markers*. 2nd ed. Burlington: Elsevier Academic Press; 2005
  10. Butler J, Coble M, Vallone P. STRs vs. SNPs: thoughts on the future of forensic DNA testing. *Forensic Sci Med Pathol* 2007;3:200-5.
  11. Kim SH, Han MS, Kim W, Kim W. Y chromosome homogeneity in the Korean population. *Int J Legal Med* 2010;124:653-7.
  12. Shi M, Hou Y, Yan J, Bai R, Yu X. 6 Y-SNP Typing of China and Korean Samples Using Primer Extension and DHPLC. *J Forensic Sci* 2007;52:235-6.
  13. Sanchez JJ, Borsting C, Hallenberg C, Buchard A, Hernandez A, Morling N. Multiplex PCR and minisequencing of SNPs--a model with 35 Y chromosome SNPs. *Forensic Sci Int* 2003;137:74-84.
  14. Henegariu O, Heerema NA, Dlouhy SR, Vance GH, Vogt PH. Multiplex PCR: critical parameters and step-by-step protocol. *Biotechniques* 1997;23:504-11.
  15. Markoulatos P, Siafakas N, Moncany M. Multiplex polymerase chain reaction: a practical approach. *J Clin Lab Anal* 2002;16:47-51.
  16. Kim SH, Kim NY, Hong SB, et al. Genetic polymorphisms of 16 Y chromosomal STR loci in Korean population. *Forensic Sci Int Genet* 2008;2:9-10.
  17. Hashiyada M, Umetsu K, Yuasa I, et al. Population genetics of 17 Y-chromosomal STR loci in Japanese. *Forensic Sci Int Genet* 2008;2:69-70.
  18. Yan J, Tang H, Liu Y, et al. Genetic polymorphisms of 17 Y-STRs haplotypes in Chinese Han population residing in Shandong province of China. *Legal Med* 2007;9:196-202.
  19. Bender K. SNaPshot for pharmacogenetics by minisequencing. *Methods Mol Biol* 2005;297:243-52.
  20. Phillips C, Fang R, Ballard D, et al. Evaluation of the Genplex SNP typing system and a 49plex forensic marker panel. *Forensic Sci Int Genet* 2007;1:180-5.
  21. Ahn HJ, Kim JJ, Han MS, Hur BK. Typing of Y-SNPs in a Korean population using allele-specific hybridization. *Korean J Forensic Sci* 2005;1:15-19
  22. Jin HJ, Kwak KD, Hammer M, et al. Y-chromosomal DNA haplogroups and their implications for the dual origins of the Koreans. *Hum Genet* 2003;114:27-35.
  23. Jota MS, Lacerda DR, Sandoval JR, et al. A new subhaplogroup of native American Y-Chromosomes from the Andes. *Am J Phys Anthropol* 2011;146:553-9.
  24. Hammer MF, Karafet TM, Park H, et al. Dual origins of the Japanese: common ground for hunter-gatherer and farmer Y chromosomes. *J Hum Genet* 2006;51:47-58.
  25. Tajima A, Hayami M, Tokunaga K, et al. Genetic origins of the Ainu inferred from combined DNA analyses of maternal and paternal lineages. *J Hum Genet* 2004;49:187-93.
  26. Mizuno N, Kitayama T, Fujii K, et al. A forensic method for the simultaneous analysis of biallelic markers identifying Y chromosome haplogroups inferred as having originated in Asia and the Japanese archipelago. *Forensic Sci Int Genet* 2010;4:73-9.
  27. Chae S-H, Kim J-M, Kim I-C, et al. Identification of novel allele on the locus 47z (DXYS5) in the Korean population. *J Hum Genet* 2005;50:664-6.
  28. Shi H, Dong YL, Wen B, et al. Y-chromosome evidence of southern origin of the East Asian-specific haplogroup O3-M122. *Am J Hum Genet* 2005;77:408-19.
  29. Abdulla MA, Ahmed I, Assawamakin A, et al. Mapping human genetic diversity in Asia. *Science* 2009;326:1541-5.
  30. Xue Y, Tyler-Smith C. The hare and the tortoise: one small step for four SNPs, one giant leap for SNP-kind. *Forensic Sci Int Genet* 2010;4:59-61.