Korean J Leg Med 2013;37:183-190
http://dx.doi.org/10.7580/kjlm.2013.37.4.183

원 저

한국인에서 높은 변이율을 갖는 조절 영역 단일염기다형성 분석을 통한 빠르고 간편한 미토콘드리아 DNA 검색법

김은혜[1] · 신경진[1] · 김혜연[2]
박수정[2] · 양우익[1] · 이환영[1]

[1]연세대학교 의과대학 법의학과 및 BK21 연세의과학사업단
[2]대검찰청 DNA수사담당관실

접 수 : 2013년 10월 23일
수 정 : 2013년 11월 1일
게재승인 : 2013년 11월 25일

본 연구과제는 2011년도 대검찰청의 '영상녹화 등 과학수사 연구개발비(1333-302-260-00)'의 지원을 받아 수행되었습니다.

책임저자 : 이환영
(120-752) 서울시 서대문구 연세로 50
연세대학교 의과대학 법의학과
전화 : +82-2-2228-2482
FAX : +82-2-362-0860
E-mail : hylee192@yuhs.ac

Rapid and Simple Screening of Mitochondrial DNA in Koreans by the Analysis of Highly Variable Control Region SNPs

Eun Hye Kim[1], Kyoung-Jin Shin[1], Hye Yeon Kim[2], Su Jeong Park[2], Woo Ick Yang[1], Hwan Young Lee[1]

[1]Department of Forensic Medicine and Brain Korea 21 Project for Medical Science, Yonsei University College of Medicine, Seoul, Korea
[2]DNA Forensic Division, Supreme Prosecutors' Office, Seoul, Korea

Human mitochondrial DNA (mtDNA) is generally used to identify highly degraded forensic samples, particularly when the extracted DNA is not sufficient for nuclear DNA analysis. However, direct sequencing, the most widely used mtDNA analysis method, is laborious and time-consuming, and precludes the simultaneous analysis of many samples. Here, we describe a rapid and simple screening method for mtDNA analysis in Koreans using single base extension (SBE) methods. Sixteen highly polymorphic mtDNA SNPs from the control region were selected, and a multiplex SBE system was constructed to analyze them. Because the developed system consists of two duplex PCRs, which produce small amplicons with fewer than 270 bp, it works well with highly degraded samples such as old skeletal remains. Using this multiplex SBE system, 145 different haplotypes were expected to be observed from 593 unrelated Koreans. Seventy-three haplotypes were expected to be observed only once, and the most frequent haplotype was expected to occur 80 times. Since the mean number of pairwise differences was estimated to be 4.55, the developed system could be useful to exclude samples that do not match evidence and reference samples. Therefore, the multiplex SBE system used in this study will be a useful tool to analyze many samples simultaneously and to efficiently screen out non-matching mtDNA sequences in forensic casework.

Key Words : Mitochondria DNA, Control region, Single nucleotide polymorphism, Single base extension, Koreans

서 론

미토콘드리아 DNA (mtDNA)는 세포소기관인 미토콘드리아 안에 존재하는 약 16 kb 크기의 원형 DNA로 어머니로부터 자손에게 유전되는 특성이 있어 모계 혈통의 추적을 위한 계통 연구에 널리 사용되어 왔다.[1)] 일반적으로 mtDNA 연구는 집단 내에서 높은 변이율을 보이는 약 1.1 kb 길이의 조절 영역(control region) 염기서열을 분석 대상으로 하며, 법과학 영역에서는 mtDNA가 세포 내 여러 copy로 존재하는 특성을 이용하여 분해된 시료나 환경에 노출되어 핵 DNA 분석이 어려운 손상된 시료의 개인식별에 주로 사용한다.

대부분의 법과학 실험실에서 이루어지는 mtDNA 염기서열 분석은 Sanger 방법에 따라 수행되는데,[2)] 이 분석법은 주의가

pISSN 1225-0589
eISSN 2287-2078
ⓒCopyright 2013 by the Korean Society for Legal Medicine
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

요구되는 기술로 분석에 많은 시간이 소요되고 시료의 비교분석을 위하여 연구자가 분석결과를 일일이 확인 및 검토해야만 하는 번거로움이 있어 동시에 다수의 시료를 신속하게 분석하는 데 어려움이 있다. 이러한 문제점을 해결하기 위하여 다수의 mtDNA single nucleotide polymorphism (mtDNA SNP)을 동시에 고속으로 분석하는 array 기반의 분석법이 최근 개발되었고, 임상 분야에서는 이를 진단에 효율적으로 이용하고 있다.[3] 그러나 이러한 고효율의 분석은 많은 양의 온전한 DNA를 필요로 하기 때문에 주로 오래되고 분해된 시료를 분석 대상으로 하는 법과학 영역에는 적용하기 어려운 한계점을 가진다. 따라서 법과학 영역에서는 미량의 분해된 시료에서 다수의 SNP를 한 번에 효과적으로 분석할 수 있는 single base extension (SBE) 방법에 기초한 multiplex SNaPshot mini-sequencing을 이용한 여러 mtDNA SNP 분석 시스템이 개발되었다. 일반적으로 이러한 분석 시스템은 집단에서 변이율이 높은 몇몇 mtDNA SNP를 우선으로 분석하여, 증거물과 대조시료 간에 일치하는지 여부를 먼저 검색하고 서로 일치하지 않는 시료를 향후 분석 대상에서 제외함으로써 분석의 효율을 높이고자 하는 데 목적이 있다. 그러나 현재까지 개발된 시스템은 대부분 유럽인과 아프리카인에서 높은 변이율을 보이는 mtDNA SNP 유전자형 결정을 위한 것으로,[4-6] 한국인의 개인식별에는 효과적으로 사용될 수 없다. 국내에서도 multiplex SNaPshot mini-sequencing 방법을 이용하여 mtDNA SNP를 분석한 연구가 이전에 보고된 적이 있으나,[7] 이는 코딩영역(coding region) SNP만을 대상으로 하였고mtDNA haplogroup을 정확하게 결정하는데 주된 목적을 두어 개인식별력이 떨어지는 단점이 있다.

따라서 본 연구에서는 mtDNA를 이용한 한국인의 개인식별에서 불일치 시료 배제를 위한 1차 검색에 효과적으로 사용할 수 있는 빠르고 간편한 mtDNA 분석법을 개발하고자 하였다. 이를 위하여 한국인 집단에서 높은 변이율을 보이는 mtDNA control region SNP를 선정하고, 이를 효과적으로 분석할 수 있는 multiplex SNaPshot mini-sequencing system을 개발하였다.

재료 및 방법

1. 시료

mtDNA control region의 염기서열이 밝혀진 593명의 한국인 시료[8] 중 일부를 취하여 사용하였으며, 이때 DNA는 구강상피세포 또는 혈액으로부터 QIAamp DNA Mini Kit (QIAgen, Hilden Germany)를 이용하여 추출하였다. 법과학 분야에서의 효용성을 평가하기 위한 시료로는 6.25 전쟁 전사자 4명의 유해로부터 Lee 등[9]의 방법에 따라 추출한 DNA와 북유럽인 7명, 인도-파키스탄인 7명, 아프리카계 미국인 4명, 중동인 9명, 일본인 10명, 멕시코인 5명, 동남아시아인 5명, 남아메리카인 7명, Ashkenazi 유대인 9명, 중국인 9명을 포함하는 총 72명의 DNA 시료(Coriell Cell Repositories, Camden, NJ, USA)를 사용하였다.

2. 분석대상 mtDNA SNP 선정

한국인 593명의 mtDNA control region 염기서열을 분석한 Lee 등의 자료[8]를 바탕으로 한국인에서 10% 이상의 변이율을 보이는 mtDNA SNP를 control region의 hypervariable region (HV) I 과 HV II 영역 염기서열로부터 선정하였다. 이때 길이 다형성을 보이는 시료에서 나타나는 변이인 309.1C, 315.1C는 분석 대상에서 제외하였다.

3. mtDNA control region 증폭을 위한 PCR system 구축

mtDNA control region 증폭을 위한 PCR system은 Lee 등의 연구[10]에서 구축된 midi-primer set를 바탕으로 이 중 2개의 primer를 수정 설계하여 구축하였다(Table 1). 사용된 primer는 3´ end가 결합하는 부위에 한국인에서 1% 이상의 변이율을 가지는 mtDNA SNP가 위치하지 않도록 설계되었고, 여러 primer를 동시에 효과적으로 증폭할 수 있도록 T_m 값

Table 1. Final Concentrations of Primer Sets for Each PCR System*

System	Amplicon	Primer	Sequences (5' →3')	Amplicon size (bp)	Conc. (uM)
Duplex I	P11	F15989	CCCAAAGCTAAGATTCTAAT	267	0.8
		R16255	CTTTGGAGTTGCAGTTGATG		
	P21	F015	CACCCTATTAACCACTCACG	247	0.4
		R261	GCTGTGCAGACATTCAATTGTT		
Duplex II	P12	F16159-15m	$(tat)_5$CATAAAAACCCAATCCACAT	267	0.3
		R16410m	GAGGATGGTGGTCAAGGGA		
	P22	F155m	TATTTATCGCACCTACGTTCA	235	1.0
		R389	CTGGTTAGGCTGGTGTTAGG		

*Monoplex PCR using F15989-R389 produces a 970 bp amplicon ; Duplex PCR using F15989-R16410m and F015-R389 produces two amplicons with 422 bp and 375 bp sizes.

http://dx.doi.org/10.7580/kjlm.2013.37.4.183

을 일정하게 조정하였다. 구축된 PCR system은 각각 270 bp 이하의 크기를 갖는 두 개의 PCR 증폭 산물을 포함하는 두 세트의 duplex (duplex I & II)로 구성된 PCR system을 기본으로 하였고, 분석 대상 DNA의 분해 상태에 따라 더욱 큰 PCR 증폭 산물을 생성할 수 있는 primer set를 선택하여 사용할 수 있도록 하였다. 또한, 각 PCR set의 primer에 poly T-tail을 붙여 증폭 산물 간 20 bp 이상의 크기 차이가 나도록 primer를 설계하여 PCR 반응 후 생성물을 polyacrylamide gel electrophoresis (PAGE)에서 쉽게 확인할 수 있도록 하였다.

PCR은 2.0 ㎕의 10X Gold ST*R buffer (Promega, Madison, WI, USA), 1.0 unit의 AmpliTaq Gold DNA polymerase (Applied Biosystems, Foster City, CA, USA), primer와 0.1 ng의 template DNA를 포함하는 총 20 ㎕의 반응액을 95℃에서 11분간 처리한 후 94℃ 20초, 55℃ 1분, 72℃ 30초의 조건으로 총 33회 중합한 후, 72℃에서 7분간 처리하였다.

4. 16개 control region SNP 분석을 위한 multiplex SBE 구축 및 수행

Multiplex SBE system은 duplex I의 PCR 증폭 대상 영역에 존재하는 SNP를 분석하는 multiplex SBE I과 duplex II의 PCR 증폭 대상 영역에 존재하는 SNP를 분석하는 multiplex SBE II로 구성되도록 구축하였다. 기본적으로 SBE primer는 BatchPrimer3 프로그램(http://probes.pw.usda.gov/batchprimer)을 이용하여 설계하였으며, 각 SBE primer는 5´ end에 poly T-tail을 붙여 electropherogram에서 SNP의 유전자형이 revised Cambridge Reference Sequence (rCRS) 염기서열 순서대로 결정될 수 있도록 하였다. 또한, 이때 증폭 산물 간에는 5 nucleotide 이상의 차이가 나도록 SBE primer를 설계하여 유전자형 결정을 위한 결과 자료 분석이 쉽도록 하였다.

SBE 반응의 효율을 높이기 위하여, 설계된 SBE primer가 결합하는 부위에 분석대상인 다른 SNP이 위치할 경우 이를 반영한 degenerate primer를 제작하여 기존의 primer를 대치하

Table 2. SBE Primer Sequences and Final Concentrations of Multiplex SBE System

SBE system	Primer	Sequences (5' →3')	Conc. (uM)
I	16129F21-21	CAGCCACCATGAATATTGTAC	0.13
	16129-03CF20-21†	tAGCCACCATGAATATTGCAC	0.10
	16129-18TF21-21†	CAGTCACCATGAATATTGTAC	0.50
	16172F25-30	$(t)_5$CCACCTGTAGTACATAAAAACCCAA	0.50
	16182-10YF26-39‡	$(t)_{13}$GTACATAAAAACCCAAYCCACATCAA	2.00
	16183-11YF28m-46‡	$(t)_{17}$aAGTACATAAAAACCCAAYCCACATCAAM	3.20
	16189R20-50	$(t)_{29}$aACTTGCTTGTAAGCATGGGG	0.25
	146F23-58	$(t)_{35}$AGTATCTGTCTTTGATTCCTGCC	0.20
	150F21-65	$(t)_{44}$TGTCTTTGATTCCTGCCTCAT	0.30
	150-04CF19-65	$(t)_{46}$TCTTTGATTCCTGCCCCAT	0.45
	152R24-72	$(t)_{48}$TTGAACGTAGGTGCGATAAATAAT	0.35
II	16217F23-23	ATGCTTACAAGCAAGTACAGCAA	0.40
	16223-06YF22-32‡	$(t)_{10}$CAAGCAAGTACAGCAAYCAACC	0.60
	16261R24-36	$(t)_{12}$TTTGTTGGTATCCTAGTGGGTGAG	0.20
	16261+05AR24-36*	$(t)_{12}$TTTGTTGGTATCCTAGTGGATGAG	0.70
	16261+13TR24-36*	$(t)_{12}$TTTGTTGGTATTCTAGTGGGTGAG	0.60
	16261+17AR24-36*	$(t)_{12}$TTTGTTGATATCCTAGTGGGTGAG	0.40
	16311+08YR26-45‡	$(t)_{19}$GCTATGTACGGTAAATGGYTTTATGT	0.80
	16311+14GR26-45*	$(t)_{19}$GCTATGTACGGTGAATGGCTTTATGT	0.60
	16319R25-52	$(t)_{26}$aGTAATGTGCTATGTACGGTAAATGG	0.80
	16319+06GR25-52*	$(t)_{26}$aGTAATGTGCTATGTACGGTGAATGG	1.00
	16362R17-59	$(t)_{42}$GGGGGTCATCCATGGGG	0.35
	199F27-67	$(t)_{40}$TTCAATATTACAGGCGAACATACTTAC	0.20
	199-04CF27-67*	$(t)_{40}$TTCAATATTACAGGCGAACATACCTAC	0.10
	199-05TF27-67*	$(t)_{40}$TTCAATATTACAGGCGAACATATTTAC	0.50
	199-10GF27-67*	$(t)_{40}$TTCAATATTACAGGCGAGCATACTTAC	0.12
	249R18-73	$(t)_{55}$AAGCGGCTGTGCAGACAT	0.15

*Additional SBE primers for samples with 3% or more polymorphic sites

†Additional SBE primers for 16111 or 16126 mutation with 2% polymorphism frequency

‡Degenerate primers were used to amplify samples which have target SNPs in another SNP's primer binding site (Degeneracy Y indicated including nucleotide C or T, Degeneracy M indicated including nucleotide A or C)

였다. 또한, SBE primer 결합 부위에 3% 이상의 변이율을 나타내는 mtDNA SNP가 존재하는 경우(16266, 16274, 16278, 16325, 189, 194, 195), 이들 변이를 sequence에 포함하는 추가적인 SBE primer를 제작하여 primer mix에 포함했다. 16111, 16126 위치의 SNP는 한국인에서 2%대의 변이율을 보이지만 SBE 반응의 효율을 높이기 위하여 추가적인 primer를 제작하여 primer mix에 포함했다(Table 2).

SBE 반응을 위한 PCR 증폭 산물의 정제는 ExoSAP-IT® (USB, Cleveland, OH, USA) 1 ㎕를 PCR 증폭 산물 5 ㎕에 더해 37℃에서 45분간 반응시킨 후, 80℃에서 15분간 처리하여 수행하였다. 정제한 duplex I의 PCR 증폭 산물 1 ㎕를 취하여 ABI PRISM® SNaPshot™ Multiplex Kit (Applied Biosystems)와 multiplex SBE I primer set를 이용하여 SBE 반응을 수행하였고, 정제한 duplex II의 PCR 증폭 산물 1 ㎕를 취하여 ABI PRISM® SNaPshot™ Multiplex Kit와 multiplex SBE II primer set를 이용하여 multiplex SBE II를 수행하였다. SBE 반응 산물은 ABI PRISM 310 Genetic Analyzer (Applied Biosystems)에서 모세관 전기영동을 수행한 후, GeneScan software 3.7 (Applied Biosystems)을 이용하여 확인하였다. 또한, GeneMapper™ ID Software Version 3.1 (Applied Biosystems)을 이용하여 ladder panel을 구축함으로써 각 SNP 위치에서 유전자형을 쉽게 확인할 수 있도록 하였다.

5. 다양한 시료를 이용한 multiplex SBE system의 검증

구축된 multiplex SBE system의 법과학 분야에서의 효용성을 검증하고자 60년 이상 오래되어 백골화된 유해 4구로부터 추출한 DNA를 이용하여 multiplex SBE 반응을 수행하였다. 기술된 방법에 따라 반응을 수행하였으며, PCR 증폭에 한해 AmpliTaq Gold DNA polymerase를 1.5 unit으로 증가시켜 사용하였다. 얻어진 PCR 반응의 증폭 산물은 PAGE에서 확인하였으며, ExoSAP-IT®로 정제하고 ABI PRISM® SNaPshot™ Multiplex Kit과 SBE primer를 이용하여 multiplex SBE 반응을 수행하였다. 반응의 결과는 ABI PRISM 310 Genetic Analyzer, GeneScan software 3.7과 GeneMapper™ ID Software Version 3.1을 이용하여 확인하였다.

또한, 다양한 집단의 시료를 대상으로 multiplex SBE 반응을 수행하였다. mtDNA 증폭을 위하여 monoplex PCR을 수행하였으며(Table 1), 이때 PCR은 2.0 ㎕의 10X Gold ST*R buffer, 1.0 unit의 AmpliTaq Gold DNA polymerase, 0.6 ㎛M의 primer와 1 ng의 template DNA를 포함하는 총 20 ㎕의 반응액을 95℃에서 11분간 처리한 후 94℃ 20초, 55℃ 30초, 72℃ 1분의 조건으로 총 33회 중합한 후, 72℃에서 7분간 처리하였다. 이후 PCR 산물은 ExoSAP-IT®로 정제하여 multiplex SBE I과 multiplex SBE II로 반응을 진행하고 16개 mtDNA

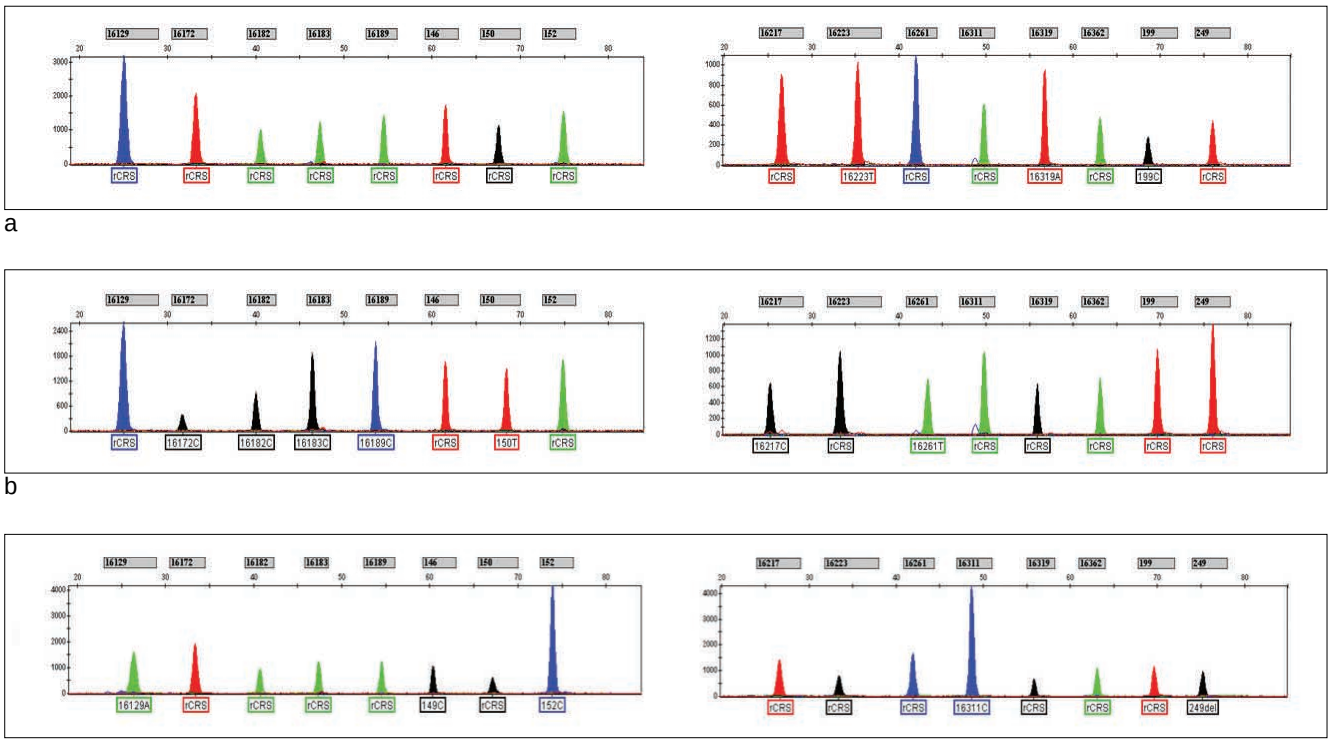

Fig. 1. Multiplex SBE genotyping of various Korean mtDNA samples (a-c). Results for multiplex SBE I were shown on the left, and the results for multiplex SBE II were shown on the right.

http://dx.doi.org/10.7580/kjlm.2013.37.4.183

SNP의 유전자형을 결정하였다.

6. 한국인 mtDNA 데이터베이스를 이용한 multiplex SBE system의 통계적 평가

선행 연구[7, 8]에서 보고된 한국인 593명의 mtDNA control region 염기서열과 coding region SNP 정보를 토대로 본 연구에서 구축한 multiplex SBE system의 효용성을 평가하였다. Nei,[11] Schoske 등[12]에 따라 haplotype diversity (HD)와 discrimination capacity (DC)를 구하였고, 통계 프로그램인 Arlequin, v. 3.5.1.2[13]를 이용하여 haplotypes의 mean number of pairwise differences (MNPD)를 구하였다. 또한, 선정한 16개 mtDNA SNP의 개인식별력을 비교해 보기 위하여 최근에 한국에서 보고된 coding region SNP 분석법[14]에 대하여서도 동일한 통계적 지표들을 구하여 함께 비교하였다.

결 과

1. 16개 control region SNP 분석을 위한 multiplex SBE system 구축

한국인에서 식별력이 높은 multiplex SBE system을 구축하기 위하여 한국인에서 10% 이상의 변이율을 보이는 16개 mtDNA SNP를 HV I과 HV II 영역으로부터 선정하였다 (16129, 16172, 16182, 16183, 16189, 16217, 16223, 16261 16311 16319 16362, 146, 150, 152, 199, 249). 이를 분석하기 위한 PCR system은 HV I과 HV II 영역의 앞쪽 절반 부분을 증폭하는 duplex I과 HV I과 HV II 영역의 뒤쪽 절반 부분을 증폭하는 duplex II로 이루어진 두 개의 duplex PCR로 설계되었다. 각각의 duplex PCR 증폭 산물은 multiplex SBE I과 multiplex SBE II를 이용하여 반응시킴으로써 각각 8개씩 총 16개의 SNP 값을 성공적으로 결정할 수 있도록 하였다.

다양한 유전자형을 갖는 다수의 한국인 시료를 이용하여 구축된 multiplex SBE system이 16개 mtDNA SNP에서 정확한 유전자형 분석 결과를 생성함을 확인할 수 있었다(Fig. 1). 또한, SBE primer가 결합하는 부위에 다른 분석대상 SNP가 위치하여 변이가 있는 경우에도 16개 mtDNA SNP의 유전자형이 성공적으로 결정되는 것을 확인할 수 있었다. 16172, 16182, 146, 16217, 16319 SNP이 16182, 16183, 150, 16223, 16311 SNP의 값을 결정하는 SBE primer 결합부위에 각각 존재하였다. 그리고 SBE primer 결합 부위에 2% 이상의 변이율을 보이는 SNP를 포함하는 시료를 대상으로 추가 실험을 수행하여 이들 모두에서 16개 mtDNA SNP의 유전자형이

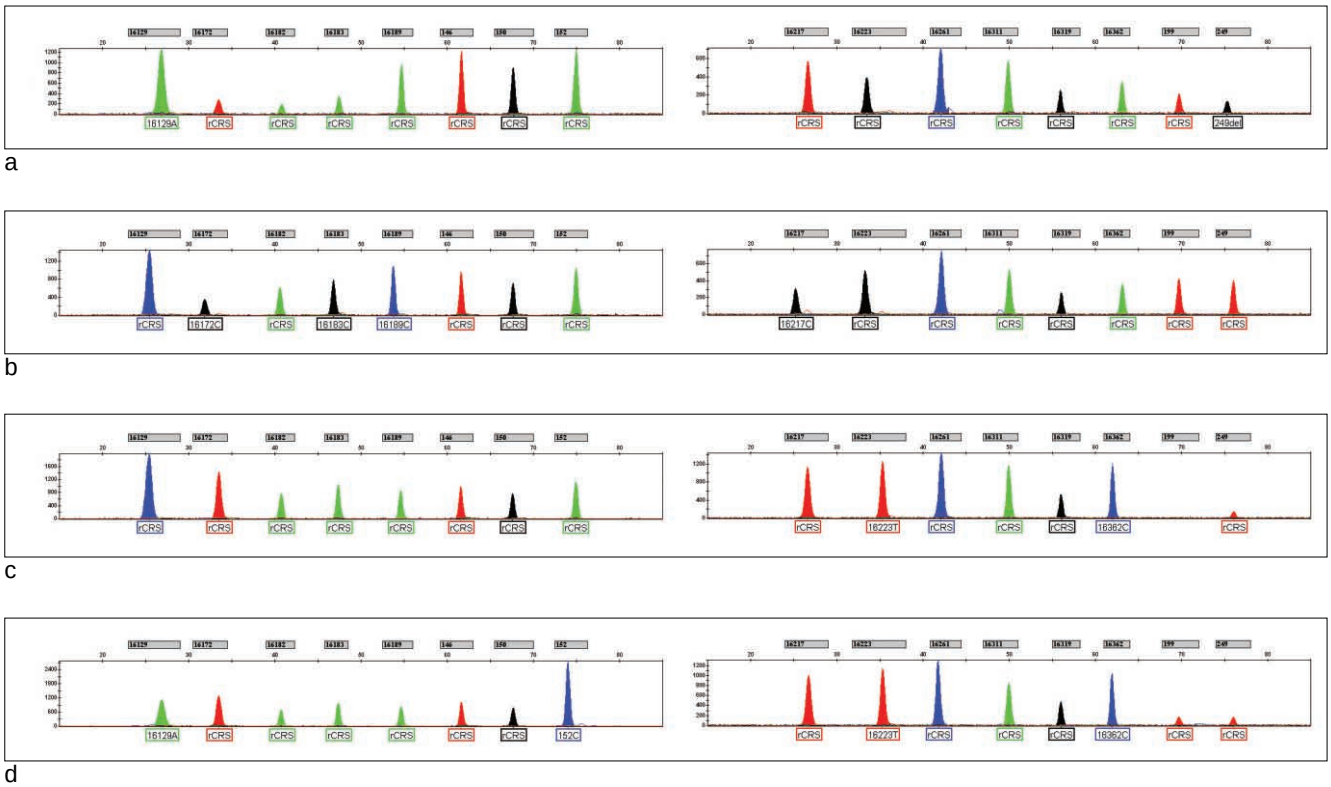

Fig. 2. Electropherograms of multiplex SBE using degraded DNA obtained from old skeletal remains (a-d). Results for multiplex SBE I were shown on the left, and the results for multiplex SBE II were shown on the right.

잘 결정되는 것을 확인하였다. 따라서 본 연구에서 구축된 multiplex SBE system은 대부분의 한국인 시료에서 16개 SNP의 유전자형을 성공적으로 결정할 수 있을 것으로 기대되었다.

2. 다양한 시료를 이용한 multiplex SBE system의 검증

두 개의 duplex로 구축된 PCR system은 PCR 증폭 산물의 크기가 270 bp 이하가 되도록 설계하였다. 따라서 이를 이용한 mtDNA control region SNP 분석이 오래되고 분해된 DNA에서도 효과적으로 이루어지는지를 알아보기 위하여 60년 이상 된 유해에서 추출한 DNA 시료를 이용하여 multiplex SBE 반응을 수행하여 유전자형을 결정하였다. 4개의 시료 중 3개에서 16개 control region SNP의 유전자형 결정이 모두 성공적으로 이루어지는 것을 확인할 수 있었다(Fig. 2). 다만, 시료 c의 경우 199 위치의 SNP 값을 결정할 수 없었는데, 이는 시료가 199 SBE primer의 3´ end 결합부위인 198 위치에 돌연변이(198T)를 갖고 있기 때문임을 direct sequencing 결과를 통해 알 수 있었다. 198T 돌연변이는 약 0.3%의 한국인에서 발견되는 매우 드문 변이임을 이전 조사된 염기서열 자료를 통해 확인할 수 있었다.[8]

한편 다양한 집단의 시료를 대상으로 16개 mtDNA SNP의 유전자형을 결정한 결과, 비교적 적은 수의 시료임에도 불구하고 일부 집단의 경우 동일한 유전자형을 갖는 시료가 많이 발견되는 것을 볼 수 있었다. 특히, 북유럽인으로 구성된 집단에서는 총 7명 중 6명이 rCRS와 동일한 유전자형을 갖는 것으로 밝혀졌고, 총 9명으로 구성된 중동인 집단의 경우에는 2개의 유전자형이 각각 4명과 2명에서 공유되는 것으로 나타났다. 이에 반해 동아시아인을 포함한 다른 나머지 집단들은 모두 2명 내지 1명에서만 발견되는 유전자형으로 구성된 것을 확인할 수 있었다. 각 집단의 DC는 북유럽인 29%, 인도파키스탄인 86%, 아프리카-아메리카인 100%, 중동 아시아인 56%, 일본인 80%, 멕시코인 83%, 동남아시아인 100%, 서아메리카인 86%, Ashkenazi 유대인 78%, 중국인 100%으로 나타났다.

3. 한국인에서 16개 control region SNP 분석을 위한 multiplex SBE system의 통계적 평가

본 연구에서 구축된 multiplex SBE system은 한국인에서 높은 식별력을 가질 것으로 추정되므로, 이를 검증하기 위하여 한국인 593명의 mtDNA 자료[7]를 대상으로 통계 지표를 구하여 그 효용성을 평가하였다. Table 3에서와 같이 본 연구에서 구축한 multiplex SBE를 이용하여 16개의 mtDNA control region SNP으로 분석하였을 경우, 593개의 한국인 시료 중 145 종류의 서로 다른 유전자형을 갖는 haplotype을 얻을 수 있음을 추정할 수 있었다. 이 중 73개의 haplotype이 한 번씩 나타났고, 가장 빈번하게 보이는 하나의 haplotype은 80번 나타나는 것을 알 수 있었다. 그리고 HD와 DC는 각각 0.9691과 24.45%였다.

반면에 이전에 보고된 15개 coding region SNP 분석법[14]을 선택하였을 경우에는 혈연관계가 없는 593개의 한국인 시료 중 20개의 서로 다른 유전자형을 갖는 haplotype을 얻을 수 있었고, 이 중 5개 haplotype이 한 번씩, 가장 빈번하게 보이는 haplotype은 128번 나타나는 것을 알 수 있었다. 산출된 HD는 0.8855, DC는 3.37%였다.

또한, Arlequin, v. 3.5.1.2를 이용하여 얻은 MNPD로부터 임의의 두 시료를 선택하여 분석했을 때 차이를 보이는 nucleotide의 평균 개수를 추정해 볼 수 있는데, 16개 control region SNP를 분석하면 MNPD가 4.55로 나타나 평균적으로 약 5개 정도의 SNP 위치에서 차이를 보일 것을 예상해 볼 수 있었다. 반면, haplogroup을 결정하는 15개 coding region SNP분석법으로 분석하면 MNPD가 2.74로 평균적으로 약 3개 정도의 SNP 위치에서 차이를 보일 것을 예상할 수 있었다.

mtDNA control region 16개 SNP를 대상으로 하는 분석법과 haplogroup을 지정하는 15개 coding region SNP를 이용한 두 가지의 분석법을 동시에 이용하여 분석하였을 경우에는 174 종류의 서로 다른 유전자형을 갖는 haplotype을 얻을 수 있음을 확인할 수 있었다. 이 중 90개 haplotype이 한 번씩 나타났으며 가장 빈번하게 보이는 haplotype은 48번 나타나는 것을 알 수 있었다. Coding region 분석만으로 128명의 한국인에서 공통으로 나타났던 haplotype은 D4 haplogroup을 나타

Table 3. Statistical Evaluation of 16 Control Region SNPs in the Present Study and 15 Coding Region SNPs of the Previous Report using 593 Unrelated Koreans' mtDNA Sequences

	Number of haplotypes	Discrimination capacity (%)	Haplotype diversity	Mean number of pairwise differences
16 control region SNPs	145	24.45	0.9691	4.55
15 coding region SNPs*	20	3.37	0.8855	2.74
31 SNPs†	174	29.34	0.9836	7.62

*15 coding region SNPs used for East asian haplogroup determination[11]
†31 SNPs consisting of 16 control region SNPs and 15 coding region SNPs

http://dx.doi.org/10.7580/kjlm.2013.37.4.183

내는 haplotype으로 16개 control region SNP 분석에 의해서 29개의 서로 다른 haplotype으로 다시 나뉘는 것을 확인할 수 있었다. 총 31개 SNP 분석을 통한 HD와 DC는 각각 0.9836과 29.34%였으며, MNPD는 7.62로 임의로 선택한 두 개의 시료를 비교해 보면 평균적으로 약 8개 정도의 SNP 위치에서 차이를 보일 것으로 예상하였다.

고 찰

본 연구에서는 기존의 mtDNA control region 염기서열분석법에 비하여 법과학 실무에서 더욱 신속하고 간편하게 사용할 수 있는 mtDNA SNP 분석법을 개발하였다. 분석을 통하여 얻게 되는 결과는 16개 SNP (deletion 포함) 값으로 구성된 haplotype의 형태로 나타나게 되는데, 이는 sequencing 분석을 통하여 얻게 되는 약 1.1 kb 상당의 control region 염기서열을 rCRS와 비교 분석하여 차이를 보이는 위치와 변화된 nucleotide만을 표시하는 기존의 haplotype (e.g. 16223T 16362C 73G 263G 315.1C)과 표시방식이 다르지만, 전체 control region 염기서열을 분석하고 haplotype을 생성하는 어렵고 복잡한 과정을 1차 스크리닝 과정을 통해 대폭 줄이는 데 효과적으로 사용할 수 있으므로 실무에서 mtDNA 분석에 대한 부담을 상당 부분 해소시킬 수 있을 것으로 기대된다. 또한, 구축된 mtDNA SNP 분석법은 시료의 상태에 따라 PCR 증폭산물의 크기를 적절히 선택할 수 있도록 다양한 primer set를 제공하였고, 특히 270 bp 미만의 PCR 증폭산물 두 개로 구성된 두 세트의 duplex PCR system은 유해와 같이 매우 분해되고 오래된 시료에서도 효과적으로 16개 SNP의 분석이 가능함을 보여주었다. 그리고 PCR primer의 3´ end가 결합하는 부위에 한국인에서 1% 이상의 변이율을 가지는 mtDNA SNP가 위치하지 않도록 primer를 설계하였기 때문에 한국인에서 높은 분석 성공률을 보일 것으로 기대된다. 실제로 593명의 한국인 mtDNA control region 염기서열을 살펴보았을 때, PCR primer의 3´ end에 변이를 갖는 시료는 단 하나 발견되었기 때문에(16178C) 개발된 mtDNA SNP 분석법은 한국인에서 매우 효과적으로 사용될 수 있을 것이다. SBE primer의 3´ end에 변이가 있는 경우(593명 중 17명)는 그 primer가 결정하는 SNP 값이 null allele로 나타날 수 있는데 동일인에서 얻은 시료일 경우 서로 일치하는 결과를 나타낼 것이므로 증거물과 대조 시료를 비교하는 데에는 문제가 없을 것으로 추정된다.

또한, 혈연관계가 없는 593명의 한국인 mtDNA 자료를 이용한 통계분석 자료를 살펴보면 본 연구에서 개발된 multiplex SBE system의 16개의 control region SNP이 mtDNA haplogroup을 결정하는 15개의 coding region SNP보다 식별력이 상당히 높음을 알 수 있다. 임의의 두 시료를 대상으로 분석했을 때 평균적으로 차이를 보이는 SNP의 수도 5개로 서로 같은 시료임을 배제할 수 있는 2개 이상 위치에서의 차이를 보는 데도 문제가 없을 것으로 기대된다. 그러나 분석대상 SNP의 수를 증가시키면 식별력이 증가할 것을 기대해 볼 수 있으므로, 보다 많은 수의 SNP를 분석할 수 있는 system을 개발하거나 본 연구에서 비교한 두 가지 system을 동시에 사용하면 보다 좋은 분석 system을 구축할 수 있을 것으로 예상된다. 또한, haplogroup의 결정이 가능한 coding region SNP 분석법은 인종 간 구별의 가능성이 있으므로 앞으로 이러한 두 가지 system을 함께 사용하는 것을 고려해 볼 필요가 있을 것이다.

구축된 mtDNA SNP 분석법이 다른 집단에서도 한국인에서와 같이 효과적으로 사용될 수 있을지를 알아보기 위해서는 보다 많은 시료에 대한 평가가 필요할 것으로 생각되나, 개발된 mtDNA SNP 분석법은 한국인을 포함한 동아시아인에서 Sanger 방법에 기초한 염기서열분석에 앞서 증거물과 대조 시료 간에 서로 일치하지 않는 시료를 우선적으로 배제하기 위한 스크리닝의 목적에 효과적으로 적용할 수 있을 것이다. 특히 SBE 방법을 사용하기 때문에 오래되고 분해된 시료의 분석이 가능하고, 실험 및 결과의 해석이 빠른 시간 안에 이루어질 수 있기 때문에 다수의 시료를 한 번에 분석할 수 있게 함으로써 법과학 실무에 매우 유용할 것이다.

참 고 문 헌

1. Butler JM. Fundamentals of forensic DNA typing. 2nd ed. San Diego: Elsevier Academic Press; 2009.
2. Butler JM. Forensic DNA typing. Biology, technology and genetics of STR markers. 2nd ed. Burlington: Elsevier Academic Press; 2005.
3. Vallone PM, Jakupciak JP, Coble MD. Forensic application of the Affymetrix human mitochondrial resequencing array. Forensic Sci Int Genet 2007;1:196-8.
4. Brandstätter A, Salas A, Niederstätter H, et al. Dissection of mitochondrial superhaplogroup H using coding region SNPs. Electrophoresis 2006;27:2541-50.
5. Grignani P, Turchi C, Achilli A, et al. Multiplex mtDNA coding region SNP assays for molecular dissection of haplogroups U/K and J/T. Forensic Sci Int Genet 2009;4:21-5.
6. Köhnemann S, Hohoff C, Pfeiffer H. An economical mtDNA SNP assay detecting different mitochondrial haplogroups in identical HVR 1 samples of Caucasian ancestry. Mitochondrion 2009;9:370-5.
7. Lee HY, Yoo JE, Park MJ, et al. East Asian mtDNA haplogroup determination in Koreans: haplogroup-level coding region SNP analysis and subhaplogroup-level control region sequence analysis. Electrophoresis 2006;27:4408-18.
8. Lee HY, Yoo JE, Park MJ, et al. 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;120:5-14.
9. Lee HY, Park MJ, Kim NY, et al. Simple and highly effective DNA extraction methods from old skeletal remains using silica columns. Forensic Sci Int Genet 2010;4:275-80.
10. Lee HY, Kim NY, Park MJ, et al. DNA typing for the identification of old skeletal remains from Korean War victims. J Forensic Sci 2010;55:1422-9.
11. Nei M. Molecular Evolutionary Genetics. New York: Columbia University Press; 1987.
12. Schoske R, Vallone PM, Kline MC, et al. High-throughput Y-STR typing of U.S. populations with 27 regions of the Y chromosome using two multiplex PCR assays. Forensic Sci Int 2004;139:107-21.
13. Excoffier L, Lischer HE. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour 2010;10:564-7.
14. Lee HY, Yoon JA, Yang WI, et al. A one step multiplex PCR assay for rapid screening of East Asian mtDNA haplogroups on forensic samples. Leg Med (Tokyo) 2013;15:50-4.

 http://dx.doi.org/10.7580/kjlm.2013.37.4.183