

Improvement of the Discrimination Capacity through the Expansion of Y Chromosomal STR Markers

Dong Gyu Lee^{1,2,*}, So Eun Lee^{2,**}, Ji Hwan Park^{1,*}, Si-Keun Lim^{2,†,***} and Ju Yeon Jung^{3,†,*}

¹Forensic DNA Division, National Forensic Service, Wonju-si, Gangwon-do 26460, Korea

²Department of Forensic Sciences, Sungkyunkwan University, Suwon-si, Gyeonggi-do 16419, Korea

³Forensic DNA Section, National Forensic Service Jeju Branch, Jeju-si, Jeju-do 63309, Korea

Y chromosomal short tandem repeat (Y-STR) markers have been developed continuously to complement forensic DNA analyses and population genetic studies. Initially, we collected data from previously reported Korean population Y-STR haplotype studies on 1133 individuals. We then conducted a marker expansion analysis using a dataset from the Y-STR Haplotype Reference Database (YHRD), covering up to 29 Y-STRs, referred to as Ymax. Additionally, we examined the impact of rapidly mutating (RM) Y-STRs included in this expanded marker set on the discrimination capacity. We observed that marker expansions both with (0.9896), and without (0.9510), RM Y-STR improved the discrimination capacity. Subsequently, we focused on 16 individuals belonging to seven distinct groups sharing identical haplotypes. These particular haplotypes had been previously identified among 476 unrelated males using 23 Y-STR markers from the PowerPlex[®] Y23 System. We expanded the marker panel up to Ymax to explore how discrimination improved with an expansion of Y-STR markers for these 16 individuals. Among the expanded markers, DYS627, which had high discriminatory power, had a high mutation rate (1.10×10^{-2}) and high gene diversity (0.83). In contrast, DYS387S1 displayed high gene diversity (0.95) but a relatively low mutation rate (2.80×10^{-3}). We propose that these findings will be valuable in the selection of suitable Y-STR markers, depending on the objectives of forensic analyses. Additionally, the presence of frequently observed Y-haplotypes in Korean population will facilitate statistical interpretation in Y-STR DNA profiling.

Key Words: Y chromosomal short tandem repeat (Y-STR), Haplotype, Marker expansion, Gene diversity, Mutation rate, Discrimination capacity

서 론

Y 염색체 짧은연쇄반복(Y chromosomal short tandem repeat, 이하 Y-STR)은 범죄현장 증거물 분석, 부계 혈통 분석 등에 널리 활용되고 있다. 특히 성범죄 사건의 혼합

시료 분석에서 소량의 남성 DNA 형이 여성 DNA 형에 가리어져(masking effect) 검출되지 않는 경우 유용하게 활용되는 분석법이다(Lee et al., 2015; Cho et al., 2016). 또한 Y-STR 마커들은 Y 염색체 상에서 재조합이 거의 일어나지 않는 부위에 속하여 돌연변이가 발생하지 않을 경우 세대와 무관하게 동일한 하플로타입(haplotype)이 유전된

Received: October 4, 2023 / Revised: October 25, 2023 / Accepted: November 1, 2023

*Researcher, **Graduate student, ***Professor.

†Corresponding author: Ju Yeon Jung. Forensic DNA Section, National Forensic Service Jeju Branch, 221, Cheomdan-ro, Jeju-si, Jeju-do 63309, Korea. Tel: +82-64-795-3820, Fax: +82-64-795-3809, e-mail: jyy7@korea.kr

†Corresponding author: Si-Keun Lim. Department of Forensic Sciences, Sungkyunkwan University, 2066, Seobu-ro, Jangan-gu, Suwon-si, Gyeonggi-do 16419, Korea.

Tel: +82-31-290-5670, e-mail: sikeun.lim@skku.edu

©The Korean Society for Biomedical Laboratory Sciences. All rights reserved.

©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.

다(Yang et al., 2012). 따라서 동일부계 여부를 분석할 때 활용되며(Lee et al., 2014), 분석된 하플로그룹(haplogroup)은 인종과 인구 집단의 이동추정 등에 사용될 수 있다(Zhang et al., 2012). 이러한 Y-STR 프로필을 이용한 법과학적 해석의 정확성과 신뢰성을 향상시키기 위해서 통계적 계산이 중요하다. 혈연 관계가 없는 해당 인구 집단에서 특정 하플로타입이 얼마나 관찰되는지 빈도 조사와 세대간의 돌연변이율 측정 등 통계적 연구가 필요하다(Park et al., 2007; Yang et al., 2012).

한국인 집단을 대상으로 한 Y-STR 하플로타입과 돌연변이율 연구가 지속적으로 보고되었다(Park et al., 2005; Lee et al., 2007; Yoo et al., 2014; Oh et al., 2015; Jung et al., 2016). 최근 보고된 선행 연구에서도(Lee et al., 2023), 620쌍의 부-자 시료를 이용한 돌연변이 연구와 서로 연관 없는 476명을 대상으로 PowerPlex[®] Y23 System (Promega, Madison, WI, USA)에 포함된 23개의 Y-STR 마커를 기준으로 하플로타입 분석이 수행되었다. 그 결과 서로 연관이 없는 대상 내에서 두 명 이상이 동일한 하플로타입을 가지는 사례들이 발견되었다. 이 사례를 바탕으로, 본 연구에서는 동일한 하플로타입을 가진 16명을 대상으로 Y-STR 마커 수를 증가시킴(마커 확장)으로써 식별력이 어떻게 향상되는지 분석하였다. 더불어, 이전에 보고된 한국인 집단 Y-STR 하플로타입 연구 데이터로부터 최대 1,133명을 대상으로 마커 확장에 의한 식별력의 변화를 분석하였고, 한국인에게 많이 나타나는 하플로타입을 제시하였다.

본 연구에서 확장되는 마커에 포함된 rapidly mutating (RM) Y-STR이 식별력 변화에 어떠한 영향을 미치는지 분석하였다. PowerPlex[®] Y23 System (Promega)에 포함된 DYS570과 DYS576, 그리고 Yfiler[™] Plus PCR Amplification Kit (Applied Biosystems, Foster City, CA)에 포함된 5개의 마커(DYS449, DYS518, DYS627, DYF387S1)는 돌연변이율이 높은 RM Y-STR 마커로 알려져 있다(Ralf et al., 2020). RM Y-STR은 높은 돌연변이율 때문에 분석 목적에 따라 사용하는데 장단점이 있다(Henry et al., 2015). 현장증거물 분석에서의 적용에 있어서 Y-STR의 한계는 범죄현장에서 검출된 Y-STR DNA 형이 피의자와 동일하더라도, 이는 피의자의 부계 친족과도 동일할 가능성이 있다는 것이다. 이때 RM Y-STR을 사용하게 된다면 밀접하게 연관된 남성 간의 식별도 가능할 수 있다(Henry et al., 2015; Turrina et al., 2016; Ralf et al., 2020). 반면, 불상변사자 및 대량재난희생자 신원확인과 같이 부계 친족 간의 하플로

타입 비교를 통해 생물학적 가족 관계를 분석하는 경우, RM Y-STR을 포함하는 Y-STR 하플로타입 분석은 우도비가 낮아지거나 친족임에도 불구하고 잘못된 배제가 발생할 수 있어 사용에 적절하지 않을 수 있다(Henry et al., 2015; Ambers et al., 2018).

따라서 본 연구는 선행 연구를 포함한 한국인 집단 내 Y-STR 마커의 집단유전학적 수치(gene diversity 및 돌연변이율)와 한국인에게 많이 관찰되는 Y-STR 하플로타입 제시를 통해, 법과학적 분석 목적에 따른 마커의 적절한 선별에 도움을 주고, Y-STR의 통계적 기초자료로서 DNA 감식의 해석 정확성과 신뢰성을 높이는 데 도움을 줄 것이다.

재료 및 방법

한국인 집단 하플로타입 수집

본 연구에서는 2007년 Lee 등 (Lee et al., 2007), 2015년 Oh 등 (Oh et al., 2015)에 의해 보고된 351명의 Y-STR 하플로타입, 2016년 Jung 등 (Jung et al., 2016)에 의해 분석된 306명과 2023년 Lee 등 (Lee et al., 2023)에 의해 분석된 476명의 Y-STR 하플로타입을 수집하여 총 1,133개의 Y-STR 하플로타입을 확보하여 Y-STR 분석 마커의 확장에 따라 식별력이 얼마나 향상되었는지 분석하였다.

분석 카테고리 설정

마커 확장은 대표적인 웹 기반 Y-STR 하플로타입 데이터베이스 Y-STR haplotype reference database (YHRD; <https://yhrd.org>)에 존재하는 데이터세트를 기준으로 분석하였다. YHRD의 데이터세트는 Y-STR 상용화 키트의 변화에 따라 키트를 구성하는 마커 수가 증가함을 반영하여 Y-STR 키트 및 마커 수를 기준으로 구성되어 있다. 최소 9개의 마커로 구성된 minimal 하플로타입부터 PowerPlex[®] Y System (Promega, Madison, WI, USA)의 12개 마커로 구성된 Y12; AmpFLSTR[™] Yfiler[™] PCR Amplification Kit (Applied Biosystems, Foster City, CA)의 Y17; PowerPlex[®] Y23 System (PPY23, Promega)의 Y23; Yfiler[™] Plus PCR Amplification Kit (Yfiler Plus, Applied Biosystems)의 Y27; 대표적인 상용화 키트의 마커를 통합한 29개 마커로 구성된 Ymax까지 검색이 가능하다. 이를 참고하여 Table 1과 같이 카테고리(minimal, Y12, Y17, Y21, Y21(RM), Y22, Y23, Y27(1), Y27(2), Ymax)를 분류하고, 최소 9개의 마커(minimal)부터 최대 29개의 마커(Ymax)까지 분석을 수행

Table 1. Expansion of Y-STRs with and without RM Y-STR

Locus	Expansion with RM Y-STR					Expansion without RM Y-STR				
	Y21(RM) ¹⁾	Y23	Y27(1)	Y27(2)	Ymax	Minimal	Y12	Y17	Y21 ²⁾	Y22 ³⁾
		Powerplex Y23	Argus Y-28 QS	Yfiler Plus			Powerplex Y	Yfiler		
	n=1133	n=1133	n=673	n=673	n=673	n=1133	n=1133	n=1133	n=1133	n=673
DYS19	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
DYS389 I	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
DYS389 II	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
DYS390	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
DYS391	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
DYS392	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
DYS393	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
DYS385a/b*	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
DYS437	✓	✓	✓	✓	✓	-	✓	✓	✓	✓
DYS438	✓	✓	✓	✓	✓	-	✓	✓	✓	✓
DYS439	✓	✓	✓	✓	✓	-	✓	✓	✓	✓
DYS448	✓	✓	✓	✓	✓	-	-	✓	✓	✓
DYS456	✓	✓	✓	✓	✓	-	-	✓	✓	✓
DYS458	✓	✓	✓	✓	✓	-	-	✓	✓	✓
DYS635	✓	✓	✓	✓	✓	-	-	✓	✓	✓
YGATAH4	✓	✓	✓	✓	✓	-	-	✓	✓	✓
DYS481	✓	✓	✓	✓	✓	-	-	-	✓	✓
DYS533	✓	✓	✓	✓	✓	-	-	-	✓	✓
DYS570†	✓	✓	✓	✓	✓	-	-	-	-	-
DYS576†	✓	✓	✓	✓	✓	-	-	-	-	-
DYS549	-	✓	✓	-	✓	-	-	-	✓	✓
DYS643	-	✓	✓	-	✓	-	-	-	✓	✓
DYS449†	-	-	✓	✓	✓	-	-	-	-	-
DYS460	-	-	✓	✓	✓	-	-	-	-	✓
DYS518†	-	-	✓	✓	✓	-	-	-	-	-
DYS627†	-	-	✓	✓	✓	-	-	-	-	-
DYF387S1a/b*†	-	-	-	✓	✓	-	-	-	-	-

¹⁾ Common loci of Y23, Y27(1) and Y27(2)

²⁾ Loci except RM Y-STR in PPY23 kit

³⁾ Loci except RM Y-STR in QSY-28 kit

*Duplicated loci are counted twice as they provide two alleles

†RM Y-STR

하였다. Y27(1)은 Investigator[®] Argus Y-28 QS (QSY-28, Qiagen, Hilden, Germany)의 27개 마커를 의미하며, Y27(2)는 Yfiler Plus의 27개 마커를 의미한다. QSY-28은 5개 (DYS570, DYS576, DYS449, DYS518, DYS627)의 RM Y-STRs를 포함하고, Yfiler Plus는 앞서 언급한 5개의 RM Y-STRs에 DYF387S1을 추가로 포함하는 대신 non RM

Y-STR인 DYS549와 DYS643이 포함되지 않는 특징을 가진다. Y21(RM)은 Y23, Y27(1), Y27(2) 키트의 공통된 마커 21개를 의미한다. Y21은 PPY23에 포함된 21개의 non RM Y-STRs을 의미하고, Y22는 QSY-28에 포함된 22개의 non RM Y-STRs을 의미한다.

카테고리별 분석 시료

Table 1에서 각 카테고리별로 분석 시료의 수를 제시하였다. RM Y-STR을 포함한 마커 확장 분석에서 Y23까지는 취합한 총 1,133개의 하플로타입을 분석하였다. 이후 Ymax까지 분석은 2007년 Lee 등 (Lee et al., 2007), 2015년 Oh 등 (Oh et al., 2015)에 의해 보고된 351명의 하플로타입, 2016년 Jung 등 (Jung et al., 2016)에 의해 분석된 306명의 하플로타입과 2023년 Lee 등 (Lee et al., 2023)에 의해 분석된 16명의 하플로타입 데이터를 수집하여 총 673명의 하플로타입을 분석하였다. 2023년 Lee 등 (Lee et al., 2023)에 의해 분석된 16명의 하플로타입 데이터는 2명 이상이

동일한 Y23 하플로타입을 공유한 시료만을 선별한 것이다.

마커 확장에 따른 식별력 변화 분석

데이터세트의 마커가 확장될수록 haplotype diversity (HD)와 discrimination capacity (DC) 값의 변화를 관찰하였다. Nei가 제시한 방법에 따라 $HD = n(1 - \sum p_i^2) / (n-1)$ 으로 계산하였으며, 여기서 n은 population size이고 pi는 haplotype frequency를 나타낸다. 또한 $DC = h/n$ 으로 계산하였으며, 여기서 h는 different haplotypes의 총 수이고, n은 분석 개체의 총 수를 나타낸다(Nei, 1987).

Lee et al. (Lee et al., 2023) 선행 연구에서 서로 연관 없

Table 2. The forensic parameters evaluated for each haplotype set

No. of identical haplotype ¹⁾	Expansion with RM Y-STRs					Expansion without RM Y-STRs				
	Y21(RM)	Y23	Y27(1)	Y27(2)	Ymax	Minimal	Y12	Y17	Y21	Y22
	n=1133	n=1133	n=673	n=673	n=673	n=1133	n=1133	n=1133	n=1133	n=673
1 (unique haplotype)	1067	1083	656	659	660	591	734	973	1042	620
2	22	18	7	5	5	89	80	45	31	17
3	1	2	1		1	28	22	8	3	1
4	1			1		9	12	3	1	
5						7	1			
6	1					4	2		1	1
7						2	1			
8		1				1	2			
9	1					1				
10						1	1		1	1
11						3	1	2		
12								1		
14						1				
18							1			
19						1	1			
20						1				
22						1				
27							1			
36						1				
Different haplotype	1093	1104	664	665	666	741	859	1032	1079	640
HMP ²⁾	0.0010	0.0010	0.0015	0.0015	0.0015	0.0040	0.0026	0.0013	0.0010	0.0018
HD ³⁾	0.9999	0.9999	1.0000	1.0000	1.0000	0.9969	0.9983	0.9996	0.9998	0.9996
DC ⁴⁾	0.9647	0.9744	0.9866	0.9881	0.9896	0.6540	0.7582	0.9109	0.9523	0.9510

¹⁾ The count of individuals with an identical haplotype

²⁾ Haplotype match probability

³⁾ Haplotype diversity

⁴⁾ Discrimination capacity

는 남성 476명 중 2명 이상이 동일한 Y23 하플로타입을 가지는 16명의 시료를 선별하였고, Yfiler Plus를 사용한 추가적인 DNA 증폭을 시행하여 Ymax까지 마커 확장 분석을 시행하였다. 증폭 산물은 3,500 xL Genetic Analyzers (Applied Biosystems)를 사용하여 모세관 전기영동(Capillary Electrophoresis, CE)을 시행하였고, GeneMapper ID-X software v.1.4 (Applied Biosystems)를 사용하여 대립유전자를 확인하였다. 16명에 대한 시료의 수집과 이용은 국립과학 수사연구원의 IRB 승인(IRB 906-210118-BR-007-01)을 받아 수행되었다. 또한 29개의 마커에 대한 gene diversity

(GD)와 돌연변이율을 조사하여 제시하였다.

더불어, 한국인에게 많이 관찰되는 하플로타입을 분석 카테고리의 기준으로 제시하였으며, 각 하플로타입이 관찰된 개체의 수를 제시하였다.

결 과

마커 확장에 따른 식별력 변화 분석

본 연구는 마커가 확장되면서 변화하는 HD 값과 DC 값을 관찰하기 위해 YHRD 데이터세트와 선행 연구를

Table 3. Marker expansion from 23 Y-STRs to 29 Y-STRs

Classification	Group No.	1	2	3	4	5	6	7	
								7(1)	7(2)
	DYS19	15	16	16	15	16	15	15	
	DYS389 I	14	14	14	12	14	12	12	
	DYS389 II	29	29	29	29	29	29	29	
	DYS390	23	23	23	23	23	24	23	
	DYS391	10	10	10	10	10	10	10	
	DYS392	11	13	13	12	13	12	12	
	DYS393	14	13	13	12	13	12	12	
	DYS385a/b	11, 16	10, 18	10, 19	11, 19	10, 18	12, 12	11, 19	
	DYS437	14	14	14	15	14	15	15	
	DYS438	10	13	13	10	13	10	10	
	DYS439	12	12	12	11	12	11	11	
	DYS448	21	18	18	19	18	19	19	
	DYS456	15	15	15	16	15	15	16	
	DYS458	17	17	17	18	17	19	18	
	DYS635	24	21	21	21	21	20	21	
	Y GATA H4	11	11	11	11	11	12	11	
	DYS481	24	23	23	20	24	22	20	
	DYS533	12	11	11	11	11	10	11	
	DYS570*	16	18	18	18	16	18	18	
	DYS576*	16	17	17	20	18	20	19	
	DYS549	12	12	12	14	12	12	14	
Y23	DYS643	9	13	13	11	13	11	11	
	DYS449*	30¹⁾	29	31	32	31	31	32	32
	DYS460	10	10	11	11	11	11	11	11
	DYS518*	36	36	40	42	42	42	37	37
	DYS627*	21	20	21	21	21	22	24	24
Ymax	DYF387S1a/b*	36, 36	36, 37	35, 40	35, 40	39, 39	35, 40	39, 40	38, 39

¹⁾Differences in alleles are bolded and underlined.

*RM Y-STR

Unidentified groups are marked

참고하였고, PPY23과 Yfiler Plus 키트에 포함된 최대 마커 수인 29개까지 분석하였다. 또한 RM Y-STR 마커가 포함된 확장과 배제된 마커 확장의 식별력 변화도 함께 분석하였다. Table 2에서 제시된 것과 같이, RM Y-STR을 포함한 마커가 확장되면서 동일한 하플로타입을 공유하는 개체 수가 점차 감소하였으며, HD 값과 DC 값은 점차 증가하였다. Y23에서 Y27(1) 또는 Y27(2) 확장은 HD 값을 0.9999에서 1.0000으로 증가시켰으며, 특히 non RM Y-STR인 DYS549, DYS643 대신 RM Y-STR인 DYF387S1 마커를 추가한 Y27(2) (0.9881)가 Y27(1) (0.9866)에 비해 DC

값을 좀 더 증가시켰다. Ymax에서는 HD 값과 DC 값 모두 가장 높은 값을 보였다.

다음으로, 돌연변이를 고려하여 RM Y-STR을 배제한 확장 분석을 수행하여 HD와 DC 값에 미치는 영향을 관찰하였다. Table 2에서 제시된 것과 같이, minimal에서 HD 값은 0.9969로 높은 값을 보였지만 DC 값 0.6540으로 다소 낮았다. Non RM Y-STR 마커의 확장도 HD 값과 DC 값의 점진적인 향상에 기여하였으며, Y17부터 DC 값이

마커 수가 동일한 데이터세트 간 비교 분석을 통해

Table 4. Gene diversity (GD) and mutation rate for each locus

Locus	Korean population		Worldwide population	
	GD (Ranking)	Mutation rate (Ranking)	GD (Ranking)	Mutation rate (Ranking)
DYS19	0.70 ^[1] (13)	3.03×10^{-3} ^[1] (11)	0.70 ^[3] (11)	2.20×10^{-3} ^[4] (20)
DYS389 I	0.68 ^[1] (14)	3.03×10^{-3} ^[1] (11)	0.63 ^[3] (16)	2.43×10^{-3} ^[4] (18)
DYS389 II	0.72 ^[1] (12)	4.04×10^{-3} ^[1] (9)	0.63 ^[3] (16)	4.99×10^{-3} ^[4] (8)
DYS390	0.66 ^[1] (16)	1.01×10^{-3} ^[1] (20)	0.75 ^[3] (6)	2.00×10^{-3} ^[4] (21)
DYS391	0.28 ^[1] (27)	1.01×10^{-3} ^[1] (20)	0.52 ^[3] (22)	2.26×10^{-3} ^[4] (19)
DYS392	0.66 ^[1] (16)	0.00×10^{-3} ^[1] (24)	0.66 ^[3] (15)	0.44×10^{-3} ^[4] (26)
DYS393	0.63 ^[1] (19)	2.02×10^{-3} ^[1] (18)	0.53 ^[3] (21)	1.27×10^{-3} ^[4] (22)
DYS385	<u>0.96</u> ^[1] (1)	<u>2.02</u> $\times 10^{-3}$ ^[1] (18)	0.92 ^[3] (1)	2.78×10^{-3} ^[4] (15)
DYS437	0.41 ^[1] (26)	1.01×10^{-3} ^[1] (20)	0.62 ^[3] (18)	1.19×10^{-3} ^[4] (23)
DYS438	0.62 ^[1] (20)	0.00×10^{-3} ^[1] (24)	0.71 ^[3] (9)	0.39×10^{-3} ^[4] (27)
DYS439	0.61 ^[1] (21)	3.03×10^{-3} ^[1] (11)	0.69 ^[3] (12)	4.46×10^{-3} ^[4] (10)
DYS448	<u>0.75</u> ^[1] (11)	<u>0.00</u> $\times 10^{-3}$ ^[1] (24)	0.69 ^[3] (12)	1.19×10^{-3} ^[4] (23)
DYS456	0.51 ^[1] (24)	3.03×10^{-3} ^[1] (11)	0.71 ^[3] (9)	4.06×10^{-3} ^[4] (12)
DYS458	0.78 ^[1] (8)	7.08×10^{-3} ^[1] (4)	0.79 ^[3] (4)	6.52×10^{-3} ^[4] (6)
DYS635	0.67 ^[1] (15)	7.08×10^{-3} ^[1] (4)	0.74 ^[3] (7)	4.36×10^{-3} ^[4] (11)
Y GATA H4	0.59 ^[1] (22)	1.01×10^{-3} ^[1] (20)	0.61 ^[3] (20)	2.55×10^{-3} ^[4] (17)
DYS481	0.78 ^[1] (8)	4.07×10^{-3} ^[1] (8)	0.83 ^[3] (2)	3.75×10^{-3} ^[4] (13)
DYS533	0.50 ^[1] (25)	2.03×10^{-3} ^[1] (17)	0.62 ^[3] (18)	2.73×10^{-3} ^[4] (16)
DYS570*	0.79 ^[1] (6)	6.10×10^{-3} ^[1] (6)	0.80 ^[3] (3)	8.01×10^{-3} ^[4] (5)
DYS576*	0.78 ^[1] (8)	1.12×10^{-2} ^[1] (2)	0.79 ^[3] (4)	1.16×10^{-2} ^[4] (2)
DYS549	0.58 ^[1] (23)	3.05×10^{-3} ^[1] (10)	0.67 ^[3] (14)	3.03×10^{-3} ^[4] (14)
DYS643	<u>0.79</u> ^[1] (6)	<u>0.00</u> $\times 10^{-3}$ ^[1] (24)	0.72 ^[3] (8)	1.00×10^{-3} ^[4] (25)
DYS449*	0.86 ^[2] (3)	1.93×10^{-2} ^[2] (1)	–	9.64×10^{-3} ^[4] (4)
DYS460	0.65 ^[2] (18)	2.80×10^{-3} ^[2] (15)	–	4.55×10^{-3} ^[4] (9)
DYS518*	0.86 ^[2] (3)	5.50×10^{-3} ^[2] (7)	–	1.06×10^{-2} ^[4] (3)
DYS627*	0.83 ^[2] (5)	1.10×10^{-2} ^[2] (3)	–	1.39×10^{-2} ^[4] (1)
DYF387S1*	<u>0.95</u> ^[2] (2)	<u>2.80</u> $\times 10^{-3}$ ^[2] (15)	–	6.47×10^{-3} ^[4] (7)

¹⁾ Markers with high gene diversity and low mutation rate are bolded and underlined

*RM Y-STR

^[1] (Lee et al., 2023), ^[2] (Oh et al., 2015), ^[3] (Purps et al., 2014), ^[4] YHRD was referenced

Table 5. Haplotypes and number of individuals according to Y-STRs expansion (Continued)

Classification	Haplotype	DYS 19	DYS 389I	DYS 389II	DYS 390	DYS 391	DYS 392	DYS 393	DYS 385	DYS 437	DYS 438	DYS 439	DYS 448	DYS 456	DYS 458	DYS 635	DYS GATA H4	DYS 481	DYS 533	DYS 570	DYS 576	DYS 549	DYS 643	DYS 449	DYS 518	DYS 627	DYS 387S1	DYS 387S1	No.
	9	15	12	29	23	10	12	13	12,17	14	10	11	19	17	16	19	13	21	11	16	19	12	11					2	
	10	15	12	29	24	10	12	12	12,12	15	10	11	19	15	19	20	12	22	10	18	20	12	11					2	
	11	15	14	29	23	10	11	14	11,16	14	10	12	21	15	17	24	11	24	12	16	16	12	9					2	
	12	15	14	29	23	10	11	15	11,17	14	10	12	21	15	16	21	11	24	12	16	17	12	9					2	
	13	15	14	30	23	10	11	15	11,19	14	10	10	21	15	15	21	11	26	12	16	17	12	8					2	
	14	15	14	31	23	10	13	13	10,19	14	13	12	18	16	19	20	12	22	11	19	17	12	12					2	
	15	16	13	29	23	10	11	14	12,19	14	10	12	22	15	18	21	11	24	12	16	18	14	9					2	
	16	16	13	29	23	11	11	14	12,18	14	10	11	22	15	18	21	11	25	12	16	19	13	9					2	
	17	16	14	29	23	10	13	13	10,18	14	13	12	18	15	17	21	11	24	11	16	18	12	13					2	
	18	16	14	29	23	10	13	13	10,18	14	13	12	18	15	18	21	11	23	11	16	17	12	12					2	
	19	16	14	29	23	10	13	13	10,19	14	13	12	18	15	17	21	11	23	11	17	17	12	13					2	
	20	16	14	29	23	10	13	13	10,19	14	13	12	18	15	17	21	11	23	11	18	17	13	13					2	
	21	16	14	29	23	10	13	13	10,19	14	13	12	18	15	18	21	11	23	11	18	17	12	13					2	
	1	15	12	29	23	10	12	12	11,19	15	10	11	19	16	18	21	11	20	11	18	19	14	11	32	11	37	23		3
	2	14	13	30	25	11	13	13	13,18	15	11	12	20	15	18	21	11	23	11	19	18	14	12	41	10	34	21		2
	3	15	12	29	23	10	12	12	11,19	15	10	11	19	16	18	21	11	20	11	18	19	14	11	32	11	37	24		2
	4	15	12	29	23	10	12	12	11,19	15	10	11	19	16	18	21	11	20	11	18	20	14	11	32	11	37	24		2
	5	15	12	29	23	10	12	12	14,16	15	10	12	19	17	18	21	12	22	11	17	18	13	11	35	10	37	21		2
	6	15	12	29	23	10	12	13	12,17	14	10	11	19	17	16	19	13	21	11	16	19	12	11	34	11	42	22		2
	7	15	12	29	24	10	12	12	12,12	15	10	11	19	15	19	20	12	22	10	18	20	12	11	34	10	35	21		2
	8	16	14	29	23	10	13	13	10,18	14	13	12	18	15	17	21	11	24	11	16	18	12	13	31	11	39	22		2
	1	15	12	29	23	10	12	12	11,19	15	10	11	19	16	18	21	11	20	11	18	19	14	11	32	11	37	23	38,40	4
	2	14	13	30	25	11	13	13	13,18	15	11	12	20	15	18	21	11	23	11	19	18	14	10	34	10	34	21	37,40	2
	3	15	12	29	23	10	12	12	11,19	15	10	11	19	16	18	21	11	20	11	18	19	14	11	32	11	37	24	38,40	2
	4	15	12	29	23	10	12	12	14,16	15	10	12	19	17	18	21	12	22	11	17	18	13	11	35	10	37	21	37,39	2
	5	15	12	29	24	10	12	12	12,12	15	10	11	19	15	19	20	12	22	10	18	20	12	11	34	10	35	21	37,38	2
	6	16	14	29	23	10	13	13	10,18	14	13	12	18	15	17	21	11	24	11	16	18	12	13	31	11	39	22	35,40	2
	1	15	12	29	23	10	12	12	11,19	15	10	11	19	16	18	21	11	20	11	18	19	14	11	32	11	37	23	38,40	3
	2	14	13	30	25	11	13	13	13,18	15	11	12	20	15	18	21	11	23	11	19	18	14	12	41	10	34	21	37,40	2
	3	15	12	29	23	10	12	12	11,19	15	10	11	19	16	18	21	11	20	11	18	19	14	11	32	11	37	24	38,40	2
	4	15	12	29	23	10	12	12	14,16	15	10	12	19	17	18	21	12	22	11	17	18	13	11	35	10	37	21	37,39	2
	5	15	12	29	24	10	12	12	12,12	15	10	11	19	15	19	20	12	22	10	18	20	12	11	34	10	35	21	37,38	2
	6	16	14	29	23	10	13	13	10,18	14	13	12	18	15	17	21	11	24	11	16	18	12	13	31	11	39	22	35,40	2
	1	15	12	29	23	10	12	12	11,19	15	10	11	19	16	18	21	11	20	11	18	19	14	11	32	11	37	23	38,40	3
	2	14	13	30	25	11	13	13	13,18	15	11	12	20	15	18	21	11	23	11	19	18	14	12	41	10	34	21	37,40	2
	3	15	12	29	23	10	12	12	11,19	15	10	11	19	16	18	21	11	20	11	18	19	14	11	32	11	37	24	38,40	2
	4	15	12	29	23	10	12	12	14,16	15	10	12	19	17	18	21	12	22	11	17	18	13	11	35	10	37	21	37,39	2
	5	15	12	29	24	10	12	12	12,12	15	10	11	19	15	19	20	12	22	10	18	20	12	11	34	10	35	21	37,38	2
	6	16	14	29	23	10	13	13	10,18	14	13	12	18	15	17	21	11	24	11	16	18	12	13	31	11	39	22	35,40	2
	1	15	12	29	23	10	12	12	11,19	15	10	11	19	16	18	21	11	20	11	18	19	14	11	32	11	37	23	38,40	3
	2	14	13	30	25	11	13	13	13,18	15	11	12	20	15	18	21	11	23	11	19	18	14	12	41	10	34	21	37,40	2
	3	15	12	29	23	10	12	12	11,19	15	10	11	19	16	18	21	11	20	11	18	19	14	11	32	11	37	24	38,40	2
	4	15	12	29	23	10	12	12	14,16	15	10	12	19	17	18	21	12	22	11	17	18	13	11	35	10	37	21	37,39	2
	5	15	12	29	24	10	12	12	12,12	15	10	11	19	15	19	20	12	22	10	18	20	12	11	34	10	35	21	37,38	2
	6	16	14	29	23	10	13	13	10,18	14	13	12	18	15	17	21	11	24	11	16	18	12	13	31	11	39	22	35,40	2
	1	15	12	29	23	10	12	12	11,19	15	10	11	19	16	18	21	11	20	11	18	19	14	11	32	11	37	23	38,40	3
	2	14	13	30	25	11	13	13	13,18	15	11	12	20	15	18	21	11	23	11	19	18	14	12	41	10	34	21	37,40	2
	3	15	12	29	23	10	12	12	11,19	15	10	11	19	16	18	21	11	20	11	18	19	14	11	32	11	37	24	38,40	2
	4	15	12	29	23	10	12	12	14,16	15	10	12	19	17	18	21	12	22	11	17	18	13	11	35	10	37	21	37,39	2
	5	15	12	29	24	10	12	12	12,12	15	10	11	19	15	19														

RM Y-STR 유무로 인한 DC 변화 및 DC 향상에 기여하는 마커의 특징에 대해 알아보았다. 마커 수가 동일한 Y27(1)과 Y27(2)를 비교해보았을 때 non RM Y-STR인 DYS549, DYS643을 포함한 Y27(1)보다 RM Y-STR인 DYF387S1을 포함한 Y27(2)의 DC 값이 더 높았다(Y27(1): 0.9866, Y27(2): 0.9881). 또한 Y21과 Y21(RM)을 비교해보았을 때 non RM Y-STR인 DYS549, DYS643을 포함한 Y21보다 RM Y-STR인 DYS570, DYS576을 포함한 Y21(RM)의 DC 값이 높았다(Y21: 0.9523, Y21(RM): 0.9647). Y21에서 Y22로 non RM Y-STR인 DYS460 확장은 DC 향상에 영향을 미치지 못하였지만, Y21(RM)에서 Y23 또는 Y27(2)에서 Ymax로 non RM Y-STR인 DYS549와 DYS643의 동시 확장은 DC 향상에 영향을 미쳤다.

마커 확장 분석

본 연구자의 선행 연구(Lee et al., 2023)에서 한국인 집단의 서로 연관 없는 남성 476명을 대상으로 Y23 하플로타입 분석을 시행하였고, Table 3와 같이 16명이 7그룹으로 동일한 하플로타입을 공유했다. 하플로타입이 동일하였던 16명을 대상으로 Table 3와 같이 Ymax까지 마커 확장 분석을 시행하였고, 확장된 5개의 마커인 DYS449, DYS460, DYS518, DYS627 및 DYF387S1a/b (multi-locus marker, 이하 DYF387S1)에 대해 식별력이 향상되는 정도를 분석하였으며, 각 마커별 gene diversity (GD) 및 돌연변이율을 조사하여 제시하였다(Table 4).

Table 3에서 확장된 5개의 마커 중 16개체 식별에 대한 식별력은 DYS627(4그룹 식별)가 가장 높았고, DYF387S1(3그룹 식별), DYS449(2그룹 식별), DYS518(1그룹 식별), DYS460(식별 불가)의 순서로 관찰되었다. 1그룹 이상을 식별한 DYS627, DYF387S1, DYS449 및 DYS518 마커는 모두 RM Y-STR 마커로, 1그룹도 식별하지 못한 DYS460 마커는 non RM Y-STR 마커로 알려져 있다. 반면, 4그룹(그룹 5, 6, 7(1), 7(2))의 8명은 마커를 29 Y-STRs까지 확장했음에도 대립유전자 차이가 발생하지 않았다(Table 3).

선행 연구(Oh et al., 2015)에서 확장된 5개의 마커에 대한 GD 값을 조사해본 결과, 한국인 집단 내 GD 값은 DYF387S1 (0.95), DYS518 (0.86), DYS449 (0.86), DYS627 (0.83), DYS460 (0.65) 순으로, RM Y-STR 마커는 모두 GD 값 0.8 이상의 높은 다형성을 보였다(Table 4). 이 마커들의 돌연변이율은 DYS449 (1.93×10^{-2}), DYS627 (1.10×10^{-2}), DYS518 (5.50×10^{-3}), DYF387S1 (2.80×10^{-3}), DYS460 (2.80×10^{-3}) 순으로 나타났다(Table 4).

한국인 남성의 하플로타입 특징

한국인 남성 집단에서 분석 카테고리의 마커가 확장될수록 관찰되는 하플로타입의 종류(그룹)와 빈도(개체) 수는 Table 5와 같았다. Y21(RM)에서는 26종류의 하플로타입이 관찰되었으며 최대 9개체가 하플로타입이 동일하였다. 마커가 확장될수록 관찰되는 하플로타입의 종류와 빈도 수가 줄어들었고, 최대 확장한 Ymax에서는 6종류의 하플로타입이 관찰되었으며 최대 3개체가 하플로타입이 동일하였다.

고 찰

Y-STR 분석은 상용화 키트의 지속적인 개발로 인해 마커의 수가 증가하고, 이로 인해 식별력이 계속해서 향상되어 왔다. 한국인 남성 306명을 대상으로 한 선행 연구에서 23 Y-STRs 분석 기준으로는 유일한(unique) 하플로타입이 296개였지만(HD: 0.9999; DC: 0.9837), 27 Y-STRs 분석 기준으로는 유일한 하플로타입이 304개로 증가하여(HD: 1.0000; DC: 0.9935), 마커 수가 증가할수록 식별력이 향상된 것을 확인하였다(Jung et al., 2016). 또 다른 한국인 남성 351명을 대상으로 한 선행 연구에서 23 Y-STRs 분석 기준으로는 유일한 하플로타입이 343개였지만(HD: 0.9999; DC: 0.9886), 27 Y-STRs 분석 기준으로는 전체 351개 하플로타입이 식별 가능하였다(Oh et al., 2015).

다양한 인구 집단에서 29 Y-STRs 하플로타입에 대한 연구가 지속적으로 발표되고 있다. 중국에서는 여러 하위 인종 집단을 대상으로 수행한 29 Y-STRs 하플로타입 연구 결과를 비롯하여(Zhou et al., 2018; Zhang et al., 2019; Li et al., 2020; Song et al., 2020), 자체 개발한 29개의 Y-STRs 마커를 다중 증폭하여 분석할 수 있는 Microreader™ 29Y Prime ID system (Song et al., 2020)과 DNA Typer™ Y29 PCR Amplification kit (Zhang et al., 2019)를 사용한 연구를 보고하였다. 그 중 상하이 중국 한족 집단 내 연관 없는 남성 843명을 대상으로 수행된 연구에서도, 23 Y-STRs (HD: 1.0000, DC: 0.9917)에서 29 Y-STRs (HD: 1.0000, DC: 0.9976)로 분석 마커 수가 증가함에 따라 식별력이 향상됨을 관찰하였다(Zhou et al., 2018).

그리하여 본 연구는 마커가 확장되면서 변화하는 HD 값과 DC 값을 관찰하기 위해 YHRD 데이터세트와 선행 연구를 참고하여 최대 마커 수인 29개까지 분석하였다. RM Y-STR을 포함한 마커 확장 분석에서 동일한 하플로

타입을 공유하는 개체 수가 점차 감소하였으며, HD 값과 DC 값은 점차 증가하여 높은 변별력을 얻을 수 있다. 또한 RM Y-STR을 배제한 마커 확장 분석에서도 HD 값과 DC 값이 점진적으로 향상했지만, Y21에서 Y22로 DYS460 마커 확장은 HD 값과 DC 값이 감소하였다. 이는 분석 시료 간의 차이(bias)가 원인일 가능성이 있다. 결과적으로 Y21에서 HD 값과 DC 값이 가장 높았으며, 이는 RM Y-STR 마커를 포함한 확장 분석에서 마커 수가 가장 작은 Y21(RM)의 HD 값과 DC 값보다는 낮았지만 HD 값은 0.99 이상, DC 값은 0.95 이상으로 분석하기에 신뢰성 있는 수치로 판단된다. 따라서 RM Y-STR을 배제한 확장 분석은 변이를 최소화하면서 변별력을 높일 수 있으므로 불상변사자 및 대량재난희생자 신원확인에서 동일부계의 잘못된 배제 가능성을 줄이는데 도움이 될 것으로 사료된다(Ambers et al., 2018). 상용화된 Y-STR 키트 내 마커의 구성 측면에서는, QSY-28 키트가 DYS460을 포함한 최대 22개의 non RM Y-STR을 동시에 분석할 수 있으므로 동일부계 감정에서 유용할 것으로 사료된다.

Y23에서 RM Y-STR (DYS627, DYF387S1, DYS449, DYS-518)이 포함된 확장 분석은 식별력을 향상시킬 수 있고, 현장증거물 분석에 매우 유용할 것으로 사료된다. 하지만 PYP23 키트와 Yfiler Plus 키트를 교차분석 해야 Ymax 까지 분석할 수 있으므로, 현장증거물 시료의 양적 한계를 고려하여 상황에 맞는 선택과 효율적인 분석이 필요하다. 중국에서 이러한 식별력을 고려하여 자체 개발한 Microreader™ 29Y Prime ID system 키트의 경우, 29개의 마커를 다중 증폭하여 한 번에 분석함으로써 노동력과 비용적 측면에서 효율성을 높였다. 또한, Scientific Working Group on DNA Analysis Methods (SWGDM)의 "Validation Guidelines for DNA Analysis Methods (2016)"에 따라 그 유효성(validation)이 검증되어 신뢰성 있는 결과를 얻었다(Zhao et al., 2020). 29개 마커에 대한 분석 결과를 한 번에 확보한 후 분석 목적에 따라 RM Y-STR의 이용 여부를 달리함으로써 선별적 활용이 가능할 것으로 사료된다.

선행 연구에서 총 22개 인구 집단으로 구성된 아시아 인구 그룹에서는 3,458개의 23 Y-STRs 하플로타입 중 165명이 동일한 하플로타입을 공유하였고, 총 129개 인구 집단으로 구성된 전세계 인구 그룹에서는 18,237개의 23 Y-STRs 하플로타입 중 1,393명이 동일한 하플로타입을 공유하였다(Purps et al., 2014). 한국인 집단(Lee et al., 2023)에서는 서로 연관 없는 남성 476명의 23 Y-STRs 하플로타입 중 16명이 7그룹으로 동일한 하플로타입을 공유하

였다. 이처럼 크기가 큰 인구 집단이나 특정 인구 집단에서 23 Y-STRs 하플로타입만으로는 서로 연관 없는 남성 간의 식별에 한계가 있다. 따라서 한국인 집단에서 하플로타입이 동일하였던 16명을 대상으로 Ymax까지 마커 확장 분석을 시행하였고, 7그룹으로 동일한 하플로그룹을 공유한 16개체 식별에 대한 식별력, GD 값 및 돌연변이율을 마커 내에서 순위로 비교 분석해보면, DYS627은 16개체의 식별력(1위)과 돌연변이율(2위)이 GD 값(4위)에 비해 높았고, DYF387S1은 16개체의 식별력(2위)과 GD 값(1위)이 돌연변이율(4위)에 비해 높았다. 이는 향후 개인 식별을 목적으로 Y-STR 마커 선정 시, GD 값과 돌연변이율의 종합적인 고려가 필요할 수 있음을 보여준다. 더불어, 세계 인구 집단과 비교하여 한국인 집단에서 DYS635 및 DYS449 마커의 돌연변이율이 높은 것으로 분석되었는데(Table 4), 이는 한국인에 초점을 맞춘 Y-STR 키트 자체 개발 시 고려될 필요가 있다.

한국인 집단 내 선행 연구(Oh et al., 2015)에서 DYF387S1은 RM Y-STR 마커임에도 비교적 낮은 돌연변이율에 속했고, YHRD에 등록된 전세계 인구 집단에서도 non RM Y-STR인 DYS458 (6.52×10^{-3})의 돌연변이율과 유사하였다(Table 4). DYF387S1과 유사하게 multi-locus Y-STR 마커인 DYS385도 한국인 인구 집단에서 23 Y-STRs 마커 중 가장 높은 GD (0.96) 값을 갖는 반면, 돌연변이율(2.02×10^{-3})은 13번째 순위로 비교적 낮은 편이었고, 전세계 인구 집단에서도 높은 GD (0.92) 값과 낮은 돌연변이율(2.78×10^{-3}) 양상을 보였다(Table 4). DYS448과 DYS643 또한 한국인 인구 집단에서 GD 값이 각각 0.75와 0.79로 비교적 높은 값을 보였지만 돌연변이율은 발생하지 않았으며, 전세계 인구 집단에서도 GD 값이 각각 0.69와 0.72로 비교적 높은 값을 보였지만 돌연변이율은 각각 1.19×10^{-3} 와 1.00×10^{-3} 로 비교적 낮은 편이었다(Table 4). 본 연구자의 선행 연구(Lee et al., 2023)에서도 높은 다형성(높은 GD 값)을 가지더라도 돌연변이율은 매우 낮은 마커가 존재하였고, 이와 같이 GD 값과 돌연변이율은 비례한다고 볼 수 없었다. GD 값은 높지만 돌연변이율이 낮은 마커들은 동일부계 친족 감정에 유용한 마커일 수 있으며, 분석 목적에 따라 GD 값과 돌연변이율의 종합적인 고려를 통해 마커를 선정하는 것이 중요하다.

본 연구는 한국인 남성 집단에서 분석 카테고리의 마커가 확장될수록 관찰되는 하플로타입의 종류(그룹)와 빈도(개체) 수를 제시하였고, 마커가 확장될수록 하플로타입의 종류와 빈도 수가 줄어드는 결과를 관찰하였다.

더불어, 27개의 동일한 마커 수를 가진 Y27(1)과 Y27(2) 하플로타입을 비교한 결과, 마커 수는 27개로 동일하지만 포함하고 있는 마커의 RM 여부(유무)에 따라(non RM Y-STR: DYS549, DYS643, RM Y-STR: DYF387S1) 하플로타입 양상이 다르게 관찰되었다. Non RM Y-STR이 포함되었을 때는 가장 많이 관찰되는 하플로타입의 빈도 수가 줄었고, RM Y-STR이 포함되었을 때는 하플로타입의 종류의 수가 감소하였다. 따라서 식별 목적에 따라 적절한 키트 선정 및 non RM Y-STR 또는 RM Y-STR의 선별적 확장이 가능하다.

선행 연구를 포함하여 한국인 집단에서 각 분석 카테고리별 가장 많은 수의 개체가 공유하고 있는 하플로타입을 분석하여 제시하였는데, 이는 범죄현장에서 검출된 Y-STR 프로필이 관련이 없는 남성 간에 우연히 일치할 확률을 제시하는데 도움을 줄 수 있을 것으로 사료되며, 특히 부분 프로필이 검출되었을 때 Y-STR 결과의 식별력을 고려하는데 도움이 될 것으로 사료된다.

또한, 본 연구를 통해 법과학적 분석 목적에 따른 확장 마커 선별과 상용화 키트 선정의 필요성에 대한 근거 자료를 제시하고자 하였다. 현장증거물에서 Y-STR 하플로타입의 식별력을 높이기 위해, PPY23과 Yfiler Plus를 교차 분석하여 GD 값이 높은 5개의 RM Y-STR (DYS627, DYF387S1, DYS449, DYS518)가 활용될 수 있다. 하지만 현장증거물 시료의 양적 한계와 키트의 성능이 같이 고려되어야 하며, 식별력이 높은 Y-STR 마커들을 한 번에 다중 증폭할 수 있는 시스템의 대한민국 자체 개발이 필요하다. 더불어, 식별력이 높은 마커는 한국인 집단 내 GD 값과 돌연변이율을 종합적으로 판단하여 선정해야 한다. 불상변사자 및 대량재난희생자 신원확인을 위한 친족 관계 내 동일부계 분석에서는 돌연변이로 인한 잘못된 배제를 최소화해야 하므로 RM Y-STR는 제외될 필요가 있다.

본 연구는 29개의 Y-STR 마커를 포함한 Ymax까지 확장함으로써 한국인 집단 내 Y-STR 하플로타입의 식별력 변화 및 RM Y-STR이 미치는 영향을 분석하였고, 이를 통해 법과학적 분석 목적에 따른 마커의 적절한 선별에 도움을 줄 것으로 기대한다. 더불어, 한국인에게 많이 관찰되는 Y-STR 하플로타입을 제시하여 DNA 감식의 통계적 해석에 관한 근거자료로서 활용될 수 있을 것이다.

ACKNOWLEDGEMENT

This work was supported by the Korean Government and

the Forensic Research Program of the National Forensic Service (NFS) [grant number NFS2023DNA03].

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

REFERENCES

- Ambers A, Votrubova J, Vanek D, Sajantila A, Budowle B. Improved y-str typing for disaster victim identification, missing persons investigations, and historical human skeletal remains. *International Journal of Legal Medicine*. 2018. 132: 1545-1553.
- Cho NS, Choi DH, Kim YJ, Hong SB, Kim SH, Lee WH, Han MS. An expanded korean reference db of y-chromosome str haplotype for forensic application. *Journal of Scientific Criminal Investigation*. 2016. 10: 48-54.
- Henry J, Simon C, Linacre A. The benefits and limitations of expanded y-chromosome short tandem repeat (y-str) loci. *Forensic Science International: Genetics Supplement Series*. 2015. 5: e28-e30.
- Jung JY, Park J-H, Oh Y-L, Kwon H-S, Park H-C, Park K-H, Kim EH, Lee D-S, Lim S-K. Forensic genetic study of 29 y-strs in korean population. *Legal Medicine*. 2016. 23: 17-20.
- Lee DG, Kim SJ, Cho W-C, Cho Y, Park JH, Lee J, Jung JY. Analysis of mutation rates and haplotypes of 23 y-chromosomal strs in korean father-son pairs. *Forensic Science International: Genetics*. 2023. 65: 102875.
- Lee H, Hwang IK, Park M-S, Jeong K-S, Park S-H, Kang P-W, Cho NS, Choi D-H. A validation study to apply of powerplex®y23 kit in y-str genotyping of sexual assault samples. *Korean Journal of Scientific Criminal Investigation*. 2015. 9: 181-186.
- Lee HJ, Lee SD, Lee SH, Park SJ, Jeong SJ, Lee JW. Statistical evaluation of lineage markers in individual identification. *The Korean Journal of Laboratory Medicine*. 2014. 38: 39-47.
- Lee HY, Park MJ, Chung U, Lee HY, Yang WI, Cho S-H, Shin K-J. Haplotypes and mutation analysis of 22 y-chromosomal strs in korean father-son pairs. *International Journal of Legal Medicine*. 2007. 121: 128-135.
- Li M, Zhou W, Zhang Y, Huang L, Wang X, Wu J, Meng M, Wang H, Li C, Bian Y. Development and validation of a novel 29-plex y-str typing system for forensic application. *Forensic Science International: Genetics*. 2020. 44: 102169.

- Nei M. Molecular evolutionary genetics. 1987. Columbia university press.
- Oh YN, Lee HY, Lee EY, Kim EH, Yang WI, Shin K-J. Haplotype and mutation analysis for newly suggested y-strs in korean father-son pairs. *Forensic Science International: Genetics*. 2015. 15: 64-68.
- Park MJ, Lee HY, Kim NY, Sim JE, Yang WI, Cho SH, Shin KJ. Genetic characteristics of 22 y-str loci in koreans. *Korean J Leg Med*. 2007. 31: 162-170.
- Park MJ, Lee HY, Yoo J-E, Chung U, Lee SY, Shin K-J. Forensic evaluation and haplotypes of 19 y-chromosomal str loci in koreans. *Forensic Science International*. 2005. 152: 133-147.
- Purps J, Siegert S, Willuweit S, Nagy M, Alves C, Salazar R, Angustia SM, Santos LH, Anslinger K, Bayer B. A global analysis of y-chromosomal haplotype diversity for 23 str loci. *Forensic Science International: Genetics*. 2014. 12: 12-23.
- Ralf A, Lubach D, Kousouri N, Winkler C, Schulz I, Roewer L, Purps J, Lessig R, Krajewski P, Ploski R. Identification and characterization of novel rapidly mutating y-chromosomal short tandem repeat markers. *Human Mutation*. 2020. 41: 1680-1696.
- Song F, Xie M, Xie B, Wang S, Liao M, Luo H. Genetic diversity and phylogenetic analysis of 29 y-str loci in the tibetan population from sichuan province, southwest china. *International Journal of Legal Medicine*. 2020. 134: 513-516.
- Turrina S, Caratti S, Ferriani M, De Leo D. Are rapidly mutating y-short tandem repeats useful to resolve a lineage? Expanding mutability data on distant male relationships. *Transfusion*. 2016. 56: 533-538.
- Yang IS, Lee HY, Park MJ, Yang WI, Shin K-J. Web-based y-str database for haplotype frequency estimation and kinship index calculation. *Korean Journal of Legal Medicine*. 2012. 36: 45-55.
- Yoo SY, Han S, Hwang JH, Park MJ, Kwon JA, Kim NS, Park JH, Cho NS. Population genetics of 23 y-chromosomal str loci in south koreans. *Korean Journal of Scientific Criminal Investigation*. 2014. 8: 16-30.
- Zhang AH, Lee HY, Seo SB, Lee HJ, Jin HX, Cho SH, Lyoo SH, Kim KH, Lee JW, Lee SD. Y haplogroup distribution in korean and other populations. *Korean Journal of Legal Medicine*. 2012. 36: 34-44.
- Zhang J, Mo X, Shang L, Jin X, Chen D, Zhu H, Zhang Y, Kang B, Li W, Ye J. Genetic analysis of 29 y-str loci in han population from dongfang, southern china. *International Journal of Legal Medicine*. 2019. 133: 1033-1035.
- Zhao Y, Zhang J, He T, Chen C, Guo D, Jiao Z, Yang Z, Yu Z, Ma W, Zhang Q. Validation of the microreader™ 29y prime id system for forensic use. *Forensic Science International*. 2020. 309: 110147.
- Zhou Y, Shao C, Li L, Zhang Y, Liu B, Yang Q, Tang Q, Li S, Xie J. Genetic analysis of 29 y-str loci in the chinese han population from shanghai. *Forensic Science International: Genetics*. 2018. 32: e1-e4.

<https://doi.org/10.15616/BSL.2023.29.4.302>

Cite this article as: Lee DG, Lee SE, Park JH, Lim SK, Jung JY. Improvement of the Discrimination Capacity through the Expansion of Y Chromosomal STR Markers. *Biomedical Science Letters*. 2023. 29: 302-313.