

## Mutation Cases in the Korean Population using 23 Autosomal STR Loci Analysis

Jeongyong Kim<sup>1,§,\*</sup>, Hyojeong Kim<sup>1,§,\*</sup>, Ja Hyun Lee<sup>1,§,\*</sup>, Hyo Sook Kim<sup>1,\*</sup> and Eungsoo Kim<sup>2,†,\*</sup>

<sup>1</sup>National Forensic Service, Forensic DNA Division, Wonju 26460, Korea

<sup>2</sup>National Forensic Service Seoul Institute, DNA Analysis Division, Seoul 08036, Korea

Short Tandem Repeats (STR) analysis which characterized by genetic polymorphism has been widely used in the forensic genetic fields. Unfortunately, mutation occurred in various STR loci could make it difficult to interpret STR data. Thus, the mutation rate of STR loci plays an important role for the data interpretation in human identification and paternity test. To verify the mutation of the STR loci in the Korean population, 545 trio sets (father, mother, and child) were analyzed with two commercial STR kits that include the 23 autosomal STR loci (D1S1656, TPOX, D2S441, D2S1338, D3S1358, FGA, D5S818, CSF1PO, D7S820, D8S1179, D10S1248, TH01, D12S391, VWA D13S317, D16S539, D18S51, D19S433, D21S11, D22S1045, SE33, Penta E and Penta D). As a result, 36 mutations were observed in 14 STR loci. The types of mutation were also classified by the increase or decrease of the alleles. The overall mutation rate was  $1.4 \times 10^{-3}$ , and the paternal mutation rate was four times higher than that of the maternal. This study will provide more detailed criterion for human identification by the mutation rate of STR loci in the Korean population.

**Key Words:** Autosomal DNA, Short Tandem Repeats (STR), Mutation, Korean population

DNA에는 과변이 반복 염기서열(hypervariable repetitive sequences)인 다형성 부위가 존재한다. 이러한 생물학적 특징을 통해 개인마다 고유한 DNA 정보를 가지며, 해당 부위를 구성하는 유전자 단위는 유전자좌 또는 유전자 마커라고 불리고 있다. 인간의 DNA는 polymerase chain reaction (PCR) 기법 기반 유전자 마커를 이용한 다양한 DNA 분석법들이 법과학 분야에 활용되고 있다. 오늘날 개인 식별 및 신원확인 과정에서 short tandem repeat (STR) 분석 기법이 주로 사용되고 있으며 2~7 base pair (bp)의 반복 단위로 존재하는 microsatellite를 말한다. STR은 사람마다 반복횟수가 다르고 독립적으로 유전되어 높은 다형성을 가지며, 다중 증폭 기법을 이용하여 많은 유전자 마커를 한 번에 분석 가능하다는 특징이 있다(Butler, 2007).

이러한 특징들은 STR 유전자의 대립형질을 비교해 개인 DNA 프로파일 사이의 식별력을 강화시켜, 감정물의 개인 식별에 이용하거나, 혈연관계의 부모, 자녀 사이에 발생되는 유전적 돌연변이를 판단할 수 있다(Kim et al., 2016).

국내에서는 DNA를 이용하여 강력사건 현장 감정물 분석 시 용의자 DNA profile과 1대 1로 대조하여 감정하거나 비교가 불가능한 DNA profile에 대해서는 DNA 신원확인 정보의 이용 및 보호에 관한 법률(법률 제16866호)에 의해 데이터베이스를 운영하며 범죄사건 해결에 활용하고 있으며, 불상 변사자의 경우, 불상 변사자의 DNA profile을 1촌 관계의 직계 가족과 비교하여 신원확인을 진행하고 있다.

현재는 DNA 감정에 20 Combined DNA Index System

Received: April 9, 2021 / Revised: June 22, 2021 / Accepted: June 25, 2021

\* Researcher.

§ Authors contributed equally.

† Corresponding author: Eungsoo Kim. National Forensic Service Seoul Institute, DNA Analysis Division, Seoul 08036, Korea.

Tel: +82-2-2600-4850, Fax: +82-2-2600-4889, e-mail: sophist@korea.kr

©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.

(CODIS) STR 좌위를 기준으로 분석되고 있으며, 불상 변사자 또는 실종아동 등의 신원확인 분석 식별력을 높이기 위해 추가 좌위들(Penta E, Penta D, SE33)을 사용하고 있다. DNA profile을 직접 비교하는 범죄사건 감정물의 DNA 분석과 달리 불상 변사자와 가족을 비교하는 신원확인 DNA 분석의 경우, STR 좌위에서 발생하는 돌연변이 때문에 어려움이 있다. STR의 돌연변이율이  $10^{-3} \sim 10^{-4}$  정도로 높기 때문에 좌위를 추가하여 분석하면, 돌연변이로 인해 부모와 자녀 사이 공유하는 대립유전자를 찾지 못하는 경우도 많아질 수 있다(Yang et al., 2013). 따라서 STR 기법을 이용한 신원확인 감정의 법과학적 확률계산을 위해 좌위별 돌연변이율이 참고되어야 한다(Mardini et al., 2013). STR 유전자에서의 대립유전자 분포와 빈도는 집단마다 다르게 발생된다(Budowle et al., 2001). 같은 좌위에서도 어떤 민족을 대상으로 연구가 진행되는지에 따라 다른 빈도가 나타날 수 있기 때문에, 집단이나 민족에 따른 좌위별 연구가 진행되어야 할 필요성이 있다. 이미 중국을 포함한 여러 나라에서는(브라질, 멕시코, 아제르바이잔, 터키 등) 특정 인구에서 발생하는 STR 좌위별 돌연변이 연구를 활발히 진행하고 있다(Aşıcıoğlu et al., 2004;ardini et al., 2013; Lopez Gonzalez et al., 2019; Mustafayev et al., 2019). 하지만 우리나라의 경우 좌위별 돌연변이율에 대한 연구가 미진하였고(Han et al., 2000), 좌위 확장에 따른 추가 좌위에 대한 돌연변이 연구가 필요한 실정이다(Oh et al., 2015; Kim et al., 2016). 따라서 본 연구에서는 혈연관계의 가족 시료를 대상으로 부모와 자녀 사이에서 발생할 수 있는 돌연변이를 관찰하여 DNA 분석을 통한 신원확인 감정의 신뢰성을 높이고자 하였다.

부모와 자녀 사이에 발생하는 돌연변이를 파악하기 위해, 2020년 3월부터 7월까지 일촌관계가 확인된 가계를 대상으로 총 303가계의 1,151명의 시료를 수집하였다. 한 가계의 구성원은 최소 3명에서 최대 6명이었으며, 부, 모, 자녀를 하나의 세트로 총 545개의 trio 세트가 만들어졌다. 검사 시료로는 멸균된 면봉으로 공여자의 구강 상피 세포를 채취하여 이용하였으며, 인체 유래물의 연구 진행에 대하여 국립과학수사연구원 생명윤리위원회(Institutional Review Board, IRB)로부터 승인을 받아 수행하였다(IRB number: 906-200214-BR-011-03). Genomic DNA (gDNA)는 DNA IQ™ System (Promega, Madison, WI, USA)을 이용하여 사용자 매뉴얼에 따라 추출하였다. 추출된 gDNA는 Quantifiler™ Trio DNA Quantification Kit (Thermo Fisher Scientific, Waltham, MA, USA)를 사용하여 7500 Real-Time

PCR System (Thermo Fisher Scientific) 장비로 정량 분석을 수행하였다. 정량이 끝난 DNA는 다중 유전자 증폭(multiplex PCR) 키트인 GlobalFiler™ PCR Amplification Kit (Thermo Fisher Scientific, GF)와 PowerPlex® Fusion System (Promega, PPF)을 이용하여 ABI 9700 Thermal Cycler (Thermo Fisher Scientific)로 증폭하였다. 증폭과정은 각각의 제조사에서 권고하는 표준 매뉴얼에 따라 진행되었다. 증폭된 PCR 산물은 3,500 xL Genetic Analyzer (Thermo Fisher Scientific)를 사용하여 모세관 전기 영동을 수행하였으며, GeneMapper ID-X software, ver. 1.4 (Thermo Fisher Scientific)를 사용하여 최종 결과를 분석하였다. 분석된 결과를 토대로 부모의 STR 결과와 자녀의 STR 결과를 비교하여 돌연변이 여부를 판단하였다. 돌연변이가 발생된 혈연관계에 대한 추가 검증을 위해 상염색체 STR 분석과 mitochondrial DNA (mt DNA) 분석을 수행하였다. Paternal 돌연변이 확인을 위해 아들의 경우 PowerPlex Y23 System (Promega, Y23)으로, 딸의 경우 Investigator Argus X-12 QS Kit (Qiagen, Hilden, Germany, X-12)로 검사를 진행하였다. Maternal 돌연변이의 경우 아들, 딸 모두 mt DNA sequencing으로 추가적인 돌연변이가 있는지 확인하였다.

친자확인을 위한 Trio set 분석 결과, 303가계 중 33개의 가계(10.89%)에서 총 36개의 돌연변이가 관찰되었다. 돌연변이는 분석된 23개의 상염색체 STR 좌위 중 D1S1656, D2S441, D2S1338, FGA, CSF1PO, SE33, D7S820, vWA, D12S391, Penta E, D18S51, D21S11, Penta D, D22S1045의 14개 좌위에서 관찰되었다(Table 1). 특히 SE33 좌위에서 7개의 돌연변이가 관찰되어 가장 높은 돌연변이율(1.28%)을 나타냈고, 다음으로 Penta E 좌위와 D12S391 좌위가 높은 돌연변이율을 보였다(Fig. 1). 높은 돌연변이율을 보였던 해당 좌위들은 넓은 대립유전자 범위를 갖고 있거나 높은 가변성을 갖고 있기 때문에, 다른 좌위들에 비해 높은 돌연변이율이 관찰되는 것으로 판단된다(Butler et al., 2011; Butler, 2012; Butler and Hill, 2013; Liu et al., 2017). 발견된 돌연변이는 반복 서열의 횟수가 1회 증가하는 경우를 +1, 1회 감소하는 경우를 -1, 돌연변이 발생 대립유전자 확인이 어려워 증가/감소가 정확하지 않은 경우  $\pm 1$ 로 표기하였다(Kim et al., 2016). 총 36개의 돌연변이 중 17개가 부모의 대립유전자보다 증가하였고(47.22%), 12개는 부모의 대립유전자보다 감소하였으며(33.33%), 나머지 7개의 돌연변이는 부모의 어떤 대립유전자로부터 유래되었는지 확인이 불가능하여 증가 또는 감소를 판단할 수 없었다(19.44%). 대부분의 돌연변이는 한 가계에서 한 개의 돌연

**Table 1.** List of mutations found as a result of 545 Trio analysis

Mutation type	Family No.	Loci	Father	Mother	Child	Type
1 family-1 mutation	3-d	D7S820	12~12	11~13	11~13	±1
	6-e	Penta E	11~19	11~16	11~ <b>20</b>	+1
	21-d	D22S1045	15~15	11~15	11~ <b>17</b>	+2
	29-d	SE33	26.2~28.2	18~21	21~ <b>27.2</b>	±1
	44-c	CSF1PO	12~12	10~12	10~ <b>13</b>	+1
	63-c	D12S391	21~21	18~18	18~ <b>20</b>	-1
	81-d	D12S391	22~24	19~19	19~ <b>25</b>	+1
	94-c	SE33	28.2~30.2	18~22	<b>21</b> ~30.2	-1
	103-d	D18S51	19~22	17~22	<b>18</b> ~22	±1
	109-c	D12S391	21~23	19~22	19~ <b>20</b>	-1
	117-c	D18S51	13~20	13~19	19~ <b>19</b>	-1
	124-c	D7S820	11~12	8~11	8~ <b>13</b>	+1
	136-c	SE33	21~29.2	21~24.2	21~ <b>22</b>	+1
	138-d	SE33	19~26.2	23.2~30.2	<b>20</b> ~23.2	+1
	170-c	SE33	19~19	18~21	18~ <b>20</b>	+1
	172-d	Penta E	11~24	16~20	20~ <b>23</b>	-1
	181-c	SE33	17~25.2	19~23.2	23.2~ <b>26.2</b>	+1
	195-e	D21S11	29~31	31.2~32	31.2~ <b>33</b>	+2
	221-c	Penta D	9~11	9~14	9~ <b>12</b>	+1
	225-e	CSF1PO	9~13	11~13	13~ <b>14</b>	+1
	228-c	FGA	24~26	20~21	20~ <b>25</b>	±1
	233-c	D21S11	29~29	31~32.2	<b>30</b> ~32.2	+1
	234-c	D2S1338	24~25	18~19	19~ <b>26</b>	+1
	236-d	vWA	16~18	14~20	16~ <b>19</b>	-1
	238-d	Penta D	9~12	9~9	9~ <b>13</b>	+1
	243-f	D22S1045	16~17	11~16	11~ <b>15</b>	-1
	267-d	D1S1656	13~18	13~15	<b>12</b> ~18	-1
	287-c	CSF1PO	10~12	10~12	10~ <b>11</b>	-1
	289-c	SE33	17~26	25.2~30.2	<b>25</b> ~25.2	-1
	304-d	Penta E	11~16	13~20	16~ <b>21</b>	+1
1 person-2 mutation	104-c	D1S1656	13~16	14~15	<b>14</b> ~ <b>15</b>	±1
		D12S391	19~21	19~22	<b>20</b> ~22	±1
1 person-2 mutation	231-c	D2S441	11~13	11~11	11~ <b>12</b>	±1
		Penta E	16~19	11~16	<b>17</b> ~19	+1
1 family-2 mutation	223-c	FGA	23~25	23~25	23~ <b>24</b>	-1
	223-d	Penta E	10~22	5~22	10~ <b>21</b>	-1

Bold letters indicate allele mutation or allele with high potential for mutation. c: first child of each family, d: second child of each family, e: third child of each family, f: fourth child of each family, +2: 2-step mutation

변이가 부모의 대립유전자 보다 1 반복(1-step) 증가 또는 감소를 하였으나(77.78%), 한 사람이 2개의 돌연변이를 갖는 경우(104c, 231c)도 있었으며, 한 가계의 경우 2개 좌위

의 돌연변이가 2명의 자녀에서 각각 관찰되는 경우(223c, 223d)도 있었다. 또한 부모의 대립유전자보다 반복횟수가 2회 증가(2-step)하는 경우(21d, 195e, Fig. 2)도 관찰되었다.

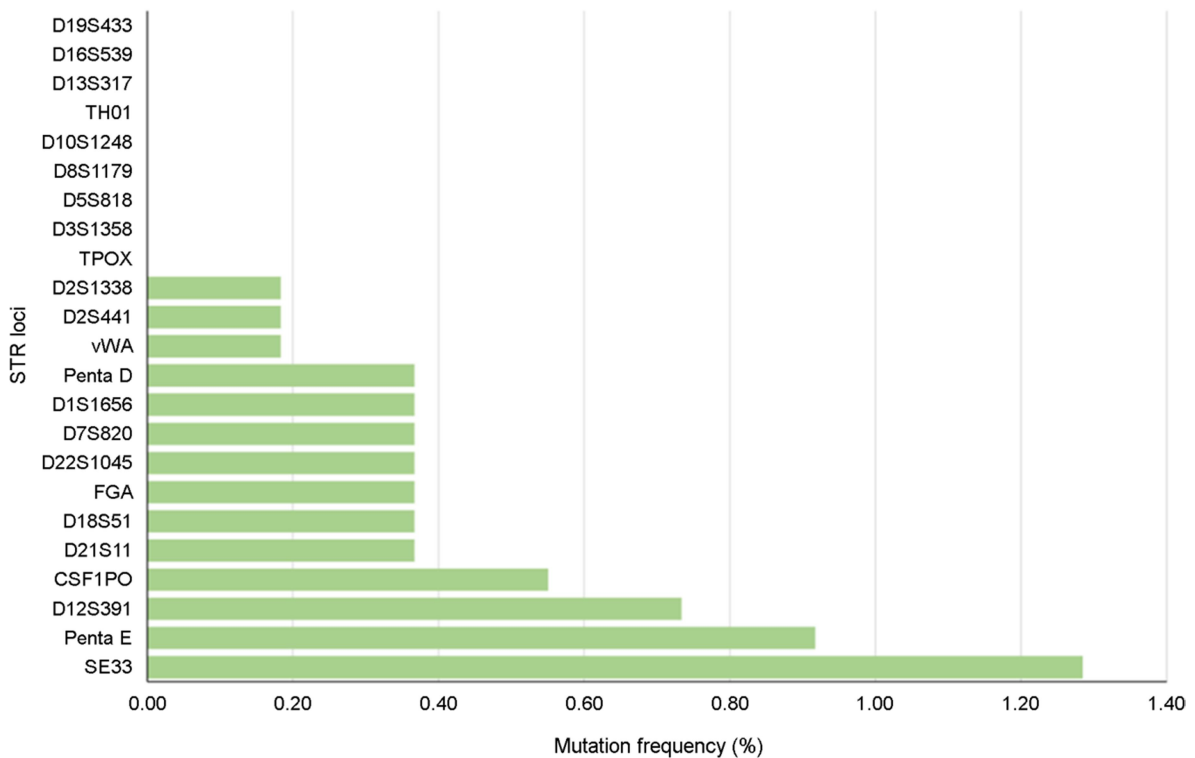


Fig. 1. Mutation frequency for each STR locus observed in the 545 trio sets.

본 연구 결과, 총 25,070개의 Meiosis 중 36개의 돌연변이가 관찰되어 돌연변이율(mutation rate)이  $1.4 \times 10^{-3}$ 로 계산되었다. 그 중 아버지로부터 유래된(paternal) 돌연변이는 25개가 관찰되어 돌연변이율은  $2.0 \times 10^{-3}$ 을 나타냈으며, 어머니로부터 유래된(maternal) 돌연변이는 6개가 관찰되어  $4.8 \times 10^{-4}$ 의 돌연변이율로 확인되었다. 그리고 나머지 5개의 돌연변이는 부모 중 누구로부터 유래되었는지 확실히 구분할 수 없었다. 예를 들어 223번 가족은 FGA의 좌위에서 부(23~25), 모(23~25), 자녀(23~24)의 대립유전자를 갖는다. 이때 자녀의 경우 부모의 대립유전자와 비교하여 돌연변이 여부를 알 수 있지만 실제 발생된 돌연변이가 부, 모 중 어디서부터 유래되었는지 알 수 없었다. Paternal 돌연변이는 maternal 돌연변이에 비해 약 4배 가량 높게 나타나는 것을 관찰할 수 있었는데 이는 Jin 등이 발표한 논문의 결과(5배) 보다 낮았지만 표본집단의 수가 상대적으로 부족하여 발생하는 차이로 보인다(Jin et al., 2016). 남성의 경우, 여성에 비해 돌연변이가 더 자주 발생하는 경향이 있는데, 이는 생식 세포 발생 시 세포 분열의 수와 유형의 차이가 원인으로 보고되었다(Brinkmann et al., 1998).

돌연변이가 발생된 혈연관계에 대한 추가 검증을 위해 성염색체 STR 분석과 mt DNA 분석을 수행하였다. 34개의 돌연변이 시료에서는 추가 돌연변이가 확인되지 않았으며, 2 step 돌연변이가 확인된 195번 가계의 DYS458 좌위에서 아버지(195a)는 16 allele이 아들(195e)은 17 allele이 관찰되었다. 부녀 간 돌연변이가 확인된 103번 가계의 경우, DXS10079 좌위에서 아버지(103a)는 19 allele이 딸(103d)은 18 allele이 관찰되었다. mt DNA 데이터 분석 결과에서는 추가적인 돌연변이가 발견되지 않았다. 따라서, 보다 신뢰성 있는 신원확인 데이터 제공을 위해서는 추가 분석 시험법인 X-STR, Y-STR 및 mt DNA 염기서열 분석 등에 대한 돌연변이 연구가 필요할 것으로 판단된다. 세대 간 돌연변이와 별도로 사용한 두 STR 키트(GF, PPF)의 결과가 상이한 10개의 시료를 확인할 수 있었으며 이는 부모와 자녀에서 동일하게 관찰되었다. 또한, 두 가계에서는 null allele이 부모에서 자녀에게로 유전됨을 확인하였다. 프라이머 접합 부위(primer binding site) 3' 말단의 돌연변이 발생 또는 삽입(insertion), 결실(deletion)은 침묵 대립유전자(silent allele) 발생을 유발하여 두 키트 간 다른 결과 또는 null allele의 원인이 되며(Butler and Hill,



**Fig. 2.** Electropherogram of trio sets with a 2-step mutation. (A) Mutation found at locus D22S1045 in family 21, (B) Mutation found at locus D21S11 in family 195.

2013), 경우에 따라 부모-자녀 사이의 친자관계 불일치를 일으킬 수 있다(Mizuno et al., 2008). 이를 명확히 확인하기 위해서는 돌연변이 좌위 분석을 위한 새로운 primer를 제작하거나 염기서열 분석을 수행하여야 한다(Cotton et al., 2000; Delamoye et al., 2004; Ricci et al., 2007; Lane, 2013).

본 연구에서는 혈연관계의 부모로부터 자녀에게 유전되는 대립유전자에 대해 23개의 상염색체 STR 좌위에서 발생하는 돌연변이를 총 303 가계의 1,151명의 시료를 분석하여 확인하였으며, 그 중 14개의 STR 좌위에서 돌연변이가 관찰되었고, 보다 정확한 신원확인 감정의 기준을

제공하기 위해서 STR 좌위들에 대한 돌연변이율 연구의 필요성을 제시하였다. 하지만 본 연구에 사용된 시료 수의 한계로 인해 돌연변이가 확인되지 않은 STR 좌위의 해석을 위해서는 보다 확장된 모집단을 통해 추가 연구 수행이 필요할 것으로 판단되며, 이를 통해 신원확인 감정의 신뢰성 있는 기준을 제공할 수 있을 것으로 기대된다.

#### ACKNOWLEDGEMENT

This work was supported by National Forensic Service (NFS2021DNA03), Ministry of the Interior and Safety, Republic of Korea.

## CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest.

## REFERENCES

- Aşıcıoğlu F, Oguz-Savran F, Ozbek U. Mutation rate at commonly used forensic STR loci: Paternity testing experience. *Dis Markers*. 2004. 20: 313-315.
- Brinkmann B, Klintschar M, Neuhuber F, Huhne J, Rolf B. Mutation Rate in Human Microsatellites: Influence of the Structure and Length of the Tandem Repeat. *Am J Hum Genet*. 1998. 62: 1408-1415.
- Budowle B, Shea B, Niezgodza S, Chakraborty R. CODIS STR Loci Data from 41 sample Populations\*. *J Forensic Sci*. 2001. 46: 453-489.
- Butler JM. Short tandem repeat typing technologies used in human identity testing. *Biotechniques*. 2007.
- Butler JM. *Advanced Topics in Forensic DNA Typing: Methodology*, *Advanced Topics in Forensic DNA Typing: Methodology*. 2012.
- Butler JM, (Becky) Hill CR, Kline MC, et al. SE33 variant alleles: Sequences and implications. *Forensic Sci Int Genet Suppl Ser*. 2011. 3.
- Butler JM, Hill CR. Biology and genetics of new autosomal STR loci useful for forensic DNA analysis. *Forensic DNA Anal. Curr Pract Emerg Technol*. 2013. 181-198.
- Cotton EA, Allsop RF, Guest JL, et al. Validation of the AMPFISTR® SGM Plus (TM) system for use in forensic case-work. *Forensic Sci Int*. 2000. 112: 151-161.
- Delamoye M, Duverneuil C, Riva K, et al. False homozygosities at various loci revealed by discrepancies between commercial kits: Implications for genetic databases. *Forensic Sci Int*. 2004. 143: 47-52.
- Han GR, Lee YW, Lee HL, et al. A Korean population study of the nine STR loci FGA, VWA, D3S1358, D18551, D21S11, D8S1179, D7S820, D13S317 and D5S818. *Int J Legal Med*. 2000. 114: 41-44.
- Jin B, Su Q, Luo H, et al. Mutational analysis of 33 autosomal short tandem repeat (STR) loci in southwest Chinese Han population based on trio parentage testing. *Forensic Sci Int Genet*. 2016. 23: 86-90.
- Kim MY, Cho S, Lyoo SH, et al. Looking Back at Our 5-Year Experience of Paternity Testing: A Summary. *Korean J Leg Med*. 2016. 40: 48-54.
- Lane AB. STR null alleles complicate parentage testing in South Africa. *South African Med J*. 2013. 103: 1004-1008.
- Liu QL, Chen YF, Huang XL, Liu KY, Zhao H, Lu DJ. Population data and mutation rates of 19 STR loci in seven provinces from China based on Goldeneye DNA ID System 20A. *Int J Legal Med* 2017. 131: 653-656.
- Lopez Gonzalez PN, Bautista-Gonzalez JE, Lopez-Gonzalez MJ, Sosa-Escalante JE, Gonzalez-Herrera L. The most frequent autosomal STRs involved in exclusion of paternity cases in a population from southeast, Mexico. *Forensic Sci Int Genet Suppl Ser*. 2019. 7: 465-467.
- Mardini AC, Rodenbusch R, Schumacher S, et al. Mutation rate estimates for 13 STR loci in a large population from Rio Grande do Sul, Southern Brazil. *Int J Legal Med*. 2013. 127: 45-47.
- Mizuno N, Kitayama T, Fujii K, et al. A D19S433 primer binding site mutation and the frequency in Japanese of the silent allele it causes. *J Forensic Sci*. 2008. 53: 1068-1073.
- Mustafayev NS, Mammadov ER, Mammadov AC. Mutation cases in the paternity tests using 15 autosomal STR markers. *J Life Sci Biomed*. 2019. 1: 5-17.
- Oh YN, Lee HY, Lee EY, et al. Haplotype and mutation analysis for newly suggested Y-STRs in Korean father-son pairs. *Forensic Sci Int Genet*. 2015. 15: 64-68.
- Ricci U, Melean G, Robino C, Genuardi M. A single mutation in the FGA locus responsible for false homozygosities and discrepancies between commercial kits in an unusual paternity test case. *J Forensic Sci*. 2007. 5: 393-396.
- Yang IS, Lee HY, Park SJ, Yang W-I, Park MJ. Analysis of Kinship Index Distributions in Koreans Using Simulated Autosomal STR Profiles. *Korean J Leg Med*. 2013. 57-65.

<https://doi.org/10.15616/BSL.2021.27.2.105>

**Cite this article as:** Kim J, Kim H, Lee JH, Kim HS, Kim E. Mutation Cases in the Korean Population using 23 Autosomal STR Loci Analysis. *Biomedical Science Letters*. 2021. 27: 105-110.