

# DNA typing을 통한 개인, 가족확인검사이론

책임연구원 윤경목, 선임연구원 박영민 / 다카라코리아 한국유전자검사센터

개인간 존재하는 DNA 다형성(polymorphism)의 분석을 통한 개인 식별 및 이를 이용한 증거물확인, 가족·혈연관계의 확인(forensic DNA typing)은 좁게는 법의학분야로부터 넓게는 인류학, 고고학에 이르는 다양한 범위에서 그 중요성을 높여가고 있다. 특히 1985년 Alec Jeffreys의 multi-locus RFLP probes 보고 이후 진행된 기술적 진보는 DNA typing의 결과가 법정에서 일반적으로 사용될 수 있을 정도의 높은 정확성을 가능케 했다.

## DNA typing에 사용되는 marker

DNA marker의 분석법 확립 이전에도 지문분석 및 ABO blood type에 의한 개인식별이나 1세대간 가족확인이 제한적으로나마 사용되었다. 그러나 여러 가지 DNA marker의 개발과 이를 이용한 정확한 개인식별이 가능해지며 현대의 개인식별, 증거물확인, 가족확인은 통상 DNA marker의 분석 즉, forensic DNA typing을 의미하게 되었다.

현재 흔히 사용되는 DNA marker들은 분석방법, 적용검체, 적용사례, 식별능력 등에서 표 1과 같은 차이점을 가지고 있다.

DNA marker	분석방법	적용검체	주 적용사례	식별능력
VNTR	RFLP-probing	대량의 검체	1세대간 가족확인	상대적으로 높다
STR	multiplex PCR	소량의 검체	1세대간 가족확인, 증거물확인	상대적으로 높다
Y-chromosome	multiplex PCR	소량의 검체	부계혈연의 확인, 남성에 한해 증거물확인	비교적 낮다
mtDNA	PCR -sequencing	극미량의 검체, 유골 등	모계혈연의 확인, 증거물확인	비교적 낮다

표1. 주로 사용되는 4가지 DNA typing marker

적용 가능한 검체의 범위, 사례 등의 차이로 case sensitive하게 최적의 marker를 선택하여 분석한다.

### · VNTR (variable number of tandem repeat)

가장 고전적인 DNA typing용 marker라고 할 수 있다. STR과 구조적으로 유사하지만 많은 대립형질을 갖고 있어서 식별능력면에서는 뛰어나다. 그러나 자동화가 쉽지 않다는 점과 일정량 이상의 혈액과 같이 비교적 많은 양의 검체가 필요하다는 단점이 있다.

### · STR (short tandem repeat)

VNTR과 구조적으로 유사하나 대립형질이 많지 않아 식별능력면에서는 떨어진다. 그러나 10여개 이상의 STR을 함께 자동분석할 수 있는 시스템

이 일반화되었고, PCR을 통한 DNA 증폭으로 모근, 타액, 혈흔, 정액반과 같이 비교적 소량의 검체를 분석에 사용할 수 있다는 장점으로 현재 가장 일반적인 DNA typing용 marker로 사용되고 있다.

### · Y-chromosome

남성만이 갖는 Y-chromosome상의 STR을 분석하므로, 피해자와 가해자의 체액이 섞여있는 강간현장의 증거물과 용의자의 비교, 수 세대에 걸친 부계혈통의 확인 등에 사용이 가능하다. 미국 3대 대통령 토머스 제퍼슨에게 흑인 사생아가 있었다는 사실을 200여년이 흐른 후에 밝혀내는데 사용된 것으로 유명하다. 여성을 대상으로 검사가 불가능하다는 단점이 있다.

### · mtDNA (mitochondrial DNA)

극미량의 검체를 활용한 DNA typing이 가능하다는 독보적인 장점을 갖는다. 부패된 사체, 작은 피부조각, 지문, 모근이 없는 머리카락 등에서 검사가 가능한 것으로 알려져 있다. 또한 모계를 통해 유전되어 모계혈통의 확인에 사용되기도 한다.

## STR marker

현재 가장 보편적으로 사용되고 있는 STR marker의 이용을 중심으로 개인확인, 증거물확인, 가족확인에 대해 소개를 하려한다.

STR은 그 명칭이 의미하는 것처럼, 2~7개 정도의 염기로 구성된 단위체(unit)가 염색체의 특정지역에 적게는 몇 차례에서 많게는 몇 십 차례까지 반복하여 나오는 것을 의미한다. DNA typing에 사용되는 STR marker로는 흔히 4염기 반복의 것을 선택하고 있다. 그림 1.은 15번 염색체에 위치한 FESFPS라는 STR marker의 모식도이다.

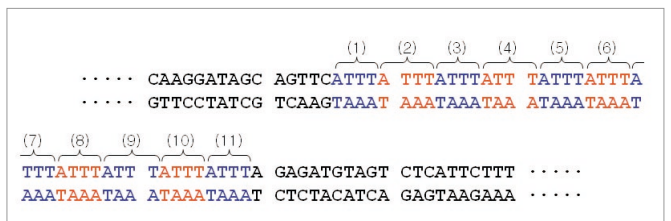


그림 1. FESFPS의 모식도  
ATTTA라는 4염기 단위체를 11회 반복으로 갖고 있다.

각 개인은 하나의 STR marker에 반복횟수로 표현되는 각 2개의 대립유전자(allele)를 갖게 되는데, 경우에 따라서 그 두 대립유전자는 동일한 반복

횡수를 가질 수도 있으며(homozygote), 서로 다른 반복횡수를 가질 수도 있다(heterozygote). 또한 한 개인이 갖는 대립유전자는 Mendelian 유전 법칙을 따라 부친과 모친으로부터 각 하나씩을 물려받은 것이다.

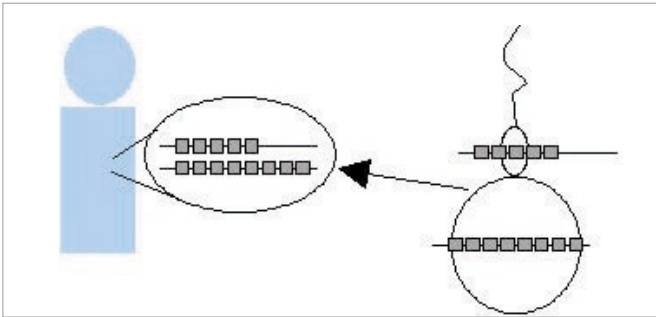


그림 2. STR의 유전 모식도  
부친의 두 대립유전자 중 allele 5, 모친의 두 대립유전자 중 allele 8의 생식세포가 만나 대립유전자 8, 5를 갖는 자손이 만들어졌다.

### PCR을 통한 STR marker의 분석

그림 1.에서와 같이 개인에 따라 혹은 한 개인도 부친과 모친으로부터 유전받은 2배수 염색체에 따라 단위체의 반복수인 대립유전자가 다를 수 있지만, 해당 반복부위 주변의 flanking region은 동일하며, 이 곳을 primer binding site로 하는 PCR(polymerase chain reaction)이 가능하다. PCR 이후 서로 다른 대립유전자는 다른 크기의 증폭산물을 내기 때문에 그림 3.에서와 같이 PAGE 등을 통해 대립유전자의 판정이 가능하다.

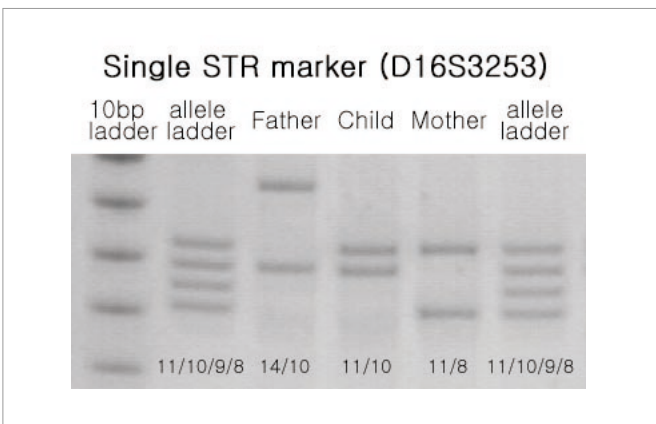


그림 3. 부-모-자로 구성된 3인 가족의 STR marker(D16S3253) 확인 결과 각자가 가지고 있는 allele은 증폭산물의 크기에 따라 전기영동을 통해 분리 및 관찰이 가능하다. 8% PAGE 후 silver stain.

또한 이와 같이 PCR을 통한 STR marker의 typing은 VNTR marker의 분석과 비교해 소량의 검체를 분석에 적용할 수 있다는 특징이 된다.

### Multiplex STRs

한 개인이 하나의 STR marker에서 갖는 대립유전자는 통상적으로 10여 개의 가능한 대립유전자 중 선택된 2개이므로, 하나의 STR marker만 기준으로 보았을 때 아무런 혈연적 연관성 없이 우연히 동일한 대립유전자 즉, 동일 DNA type을 갖는 사람의 수는 무척 많을 것으로 예상된다. 실제로 한국인 100명당 D5S818이라는 STR marker에서 11/10의 대립유전자를 갖는 경우는 15.5명, FESFPS라는 STR marker에서 11/11의 대립유전자를 갖는 경우는 23.4명 수준으로 그 빈도가 매우 높은 것으로 조사되고 있다.

이러한 STR marker별 10여개의 많지 않은 대립유전자의 수는 VNTR marker에 비교 열세를 보이는 STR marker의 가장 큰 약점이라 할 수 있다. 하지만 표.2와 같이 다소 식별능력이 떨어지는 복수의 STR marker를 한꺼번에 분석하고 여기서 나오는 확률을 곱하기 법칙(product rule)을 적용하여 계산하면 기하급수적으로 높은 식별능력을 확보할 수 있다.

검사대상자	STR marker	allele 1(p)	allele 2(q)	발생빈도 (frequency)
○○○씨	CSF1PO	13	12	0,056
	D5S818	13	13	0,015
	D7S820	12	8	0,052
	D13S317	12	8	0,089
	TH01	7	7	0,061
	TPOX	12	8	0,025
	vWA	18	18	0,033
	D3S1358	18	17	0,033
	D8S1179	15	11	0,024
	D16S3253	9	9	0,004
	D18S51	16	13	0,032
	D21S11	32,2	30	0,057
	FGA	23	21	0,063
	PentaD	12	11	0,050
	PentaE	12	10	0,006
	D12S391	20	17	0,043
	D14S608	12	11	0,066
결합 발생빈도 (Combined freq.)				2,19X10 <sup>-15</sup>

표 2. 17개 STR marker의 분석을 통한 한 개인의 DNA typing 결과 개개 STR marker 분석의 식별능력은 높지 않으나, 다수 marker의 결과를 통한 전체적인 식별능력은 전 인류 중에 한 개인만의 독특한 타입의 수준으로 분석이 가능하다.

하지만 다수의 STR marker를 독립적으로 PCR 후 각각의 product를 PAGE와 같은 manual 방식으로 분석하는 것은 많은 노동력과 시간을 필요로 한다. 이를 극복하기 위해 최근에는 10여개 이상의 STR marker를 하나의 tube에서 multi-dye로 labeling된 primer set를 이용해 PCR수행하고 이를 한번에 분석할 수 있는 하드웨어 및 소프트웨어 시스템이 개발되어 상용화되어있는 상태이다.

현재 본 센터에서는 96명에 대한 15개의 STR marker와 amelogenin marker(성별확인 DNA marker)를 한번에 확인할 수 있는 시스템을 활용하고 있다.

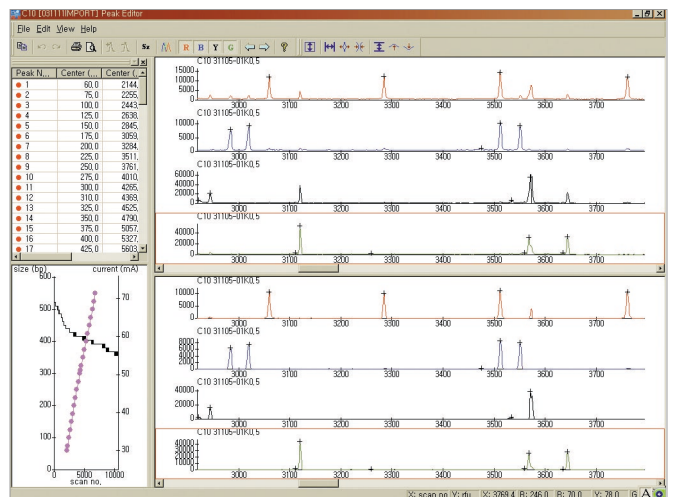


그림 4. Multiplex STR 분석  
복수의 fluorescent dye로 labeling된 primer set를 이용한 multiplex PCR와 capillary DNA analyzer를 이용한 분석

하지만 DNA typing에 사용될 수 있는 STR marker의 선택에 있어서, 현재 수없이 많이 존재하는 marker들 중에 연구자마다 서로 다른 것을 선택해 사용하게 된다면, 표준화된 DB를 확보하거나 연구자간 정보공유 및 활용의 측면에서 부정적인 것이다. 이에 따라 몇몇 국가기관 및 국제기관이 표준화된 STR marker 시스템을 도입하여 운영하고 있다. (그림 5.) 현재 국내에서는 FBI의 13CODIS나 이를 일부 변경한 STR marker 시스템을 사용한 분석이 일반화되어있다.

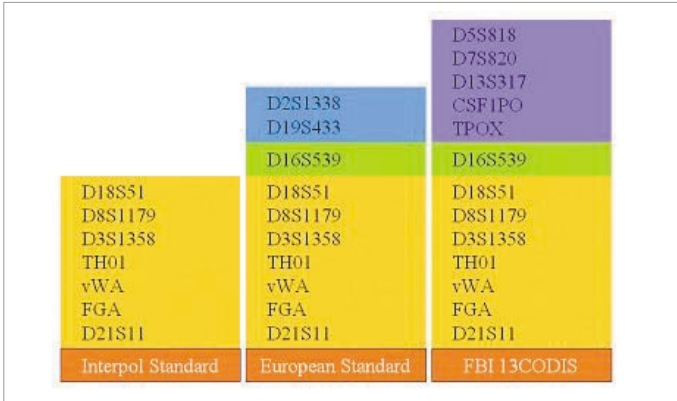


그림 5. 각 기관에서 정한 표준 STR marker 시스템

### 개인의 유전자타입 검사

검체의 유전자타입간 비교 없이 단순히 하나의 검체(증거물이나 개인이 될 수 있다)의 유전자타입의 검사결과는 통상 검사대상 STR marker 들에서의 대립유전자 타입, 대립유전자 타입의 발생빈도, 결합 발생빈도 등으로 제공된다. 이를 위해서는 피검자가 속한 population에 대한 기 작성된 각 STR marker별 대립유전자 frequency DB를 필요로 한다.

특정 STR에서 대립유전자 p, q를 갖는 type의 빈도는 다음과 같이 계산된다.

$$P(p,q) = 2pq \quad \text{---- 식.1.}$$

해당 STR에서 대립유전자 p, p를 갖는 homozygote인 경우는 다음으로 계산 된다.

$$P(p,p) = p^2 \quad \text{---- 식.2.}$$

여기서 p, q는 해당 population에서 대립유전자 p, q가 갖는 각각의 발생 빈도를 의미한다. 또한 n개의 multiplex STR분석의 결과로 나타나는 한 개인 DNA typing결과와 발생빈도는 각 STR에서 확인된 빈도의 곱으로 구할 수 있다.

$$P(x) = \prod_{i=1}^n P_i \quad \text{---- 식.3.}$$

### 증거물확인 검사

범행 현장의 증거물-혐의자, 대형 참사 현장의 사체-생존당시 피해자의 혈액 등과 같은 특정 증거물-비교대상 개인간의 일치여부는, 분석 대상으로 선택한 STR marker들에서 양자간 대립유전자의 정확한 일치 여부를 보이는 것으로 확인할 수 있다. 단, 이 경우에도 무작위로 추출된 한 개인이 증거물의 대립유전자와 우연하게 일치할 수 있는 가능성을 전혀 배제할 수 없으므로, 이에 대한 검토가 뒤따르게 된다.

베이즈의 정리를 통해 DNA type의 일치 증거가 있을 때, 해당 개인이 증거물의 실제 출처일 확률은 다음과 같이 정리될 수 있다. 단, 편의를 위해 범피수사 경우의 표현(증거물, 혐의자, 유죄, 무죄 등)을 쓰기로 한다.

$$P(G|E) = \frac{P(E|G) \times P(G)}{P(E|G) \times P(G) + P(E|\sim G) \times P(\sim G)} \quad \text{---- 식.4.}$$

- E : 두 검체의 DNA type이 일치함
- G : 혐의자가 유죄임
- ~G : 혐의자가 무죄임
- P(G|E) : 유전자기반의 증거(E)가 있을 때 혐의자가 유죄(G)일 확률
- P(E|G) : 혐의자가 유죄(G)일 때 유전자기반의 증거가 확인(E) 될 확률 = 1

식.4.를 다시 정리하면 다음과 같다.

$$\frac{P(G|E)}{P(\sim G|E)} = \frac{P(E|G)}{P(E|\sim G)} \times \frac{P(G)}{P(\sim G)} \quad \text{---- 식.5.}$$

이러한 증거물-개인의 확인에 있어서 가능성비(LR, likelihood ratio)의 개념이 사용되는데, 식.5.의 P(E|G)/P(E~G)에 해당되는 LR는 다음과 같이 구할 수 있다.

$$LR = 1 / P(x) \quad \text{---- 식.6.}$$

만약에 P(x)가 0.0001이라서 LR이 10,000의 값을 갖는다면, '두 검체의 DNA typing 결과가 동일했을 경우, 증거물이 실제로 혐의자로부터 유래했을 가능성이 그렇지 않을 가능성의 10,000배' 라는 의미를 갖게 된다. 간혹 범피수사에서 이것을 식.5.의 P(G|E)/P(~G|E)에 '해당하는 DNA type이 일치하는 경우, 피의자가 범인일 확률의 비' 로 판단하는 오류(Prosecutor's fallacy)를 범하기도 하는데, 이는 P(G)/P(~G)로 표현되는 전기상대확률(prior odds)에 대한 고려를 누락했기 때문이다. 식.5.를 다시 표현하면 식.7.과 같다.

$$\text{posterior odds} = LR \times \text{prior odds} \quad \text{---- 식.7.}$$

범피수사에서 전기상대확률은 목격자, 혐의자의 유사성과 유무, 알리바이, 거짓말 탐지기의 결과 등의 정량화하기 어려운 정황들로 구성된다. 만약 사건의 혐의 없이 전적으로 유전자 DB로부터 혐의자를 찾은 경우, National Research Council은 해당 유전자 DB의 크기의 역수인 1/N을 전기상대확률로 정의할 것을 권장하기도 한다.

### 가족확인 검사

가족확인 STR marker 타이핑의 경우, 가족관계가 부정되지 않기 위해서는 그림 3.에서와 같이 아버지와 어머니로부터 각각 하나씩의 STR을 받았음이 확인되어야 한다. 경우에 따라서는 아버지와 어머니 중 1인과의 대조만으로도 친부 혹은 친모확인 가족확인검사가 진행가능하다. 그림 3.의 예에서 아버지의 14, 10 allele 중 10을, 어머니의 11, 8 allele 중 11을 전달받았음을 확인할 수 있다. 식.4.에서와 같이 베이즈의 정리를 통해 상기의 유전자기반 증거가 있을 때, 추정부(alleged father)가 친부(biological father)일 확률은 다음과 같이 정리된다.

$$P(H|E) = \frac{P(E|H) \times P(H)}{P(E|H) \times P(H) + P(E|\sim H) \times P(\sim H)} \quad \text{---- 식.8.}$$

E : 추정부의 대립유전자 1개가 자의 대립유전자와 일치함  
 H : 추정부가 자의 친부임  
 P(E|~H) : 추정부가 친부가 아닐 때(~H), 두 사람이 하나의 대립유전자를 공유할 확률  
 P(E|H) : 추정부가 친부일 때(H) 둘 중 하나의 대립유전자를 전해줄 확률 = 0,5

가족확인의 경우, 전기 상대확률로 통상 0.5를 사용. 즉, P(H)=P(~H)=0.5 인 것을 고려해 식.8.을 다시 정리하면 다음과 같다.

$$P(H|E) = \frac{[P(E|H)/P(E|\sim H)] \times P(H)}{[P(E|H)/P(E|\sim H)] \times P(H) + P(\sim H)}$$

$$= \frac{[P(E|H)/P(E|\sim H)]}{[P(E|H)/P(E|\sim H)] + 1} \quad \text{--- 식.9.}$$

가족확인의 경우, 증거물확인에서의 가능성비(LR) 대신 P(E|H)/P(E|~H)로 정의되는 부계지수(PI, paternity index)가 중요한 의미를 갖게 된다. 즉, 추정부와 자의 1개 대립유전자가 일치할 때에 추정부가 친부일 가능성인 P(H/E)는 PI/(PI+1)로 계산이 가능하다. 또한 각각의 STR marker에서 확인된 PI는 product rule의 적용으로 결합친부지수(CPI, combined paternity index)로 계산될 수 있으므로 검사대상의 모든 STR marker에서의 증거를 통한 생물학적친부일 가능성은 CPI/(CPI+1)로 계산이 가능하다. 모친을 확인 하는 경우도 이와 동일한 방법이 사용된다. 표.3.은 본 센터에서 실제 분석한바 있는 父, 子 2인이 참여한 친부 확인의 예이다. 통상적으로 母의 검사 참여가 있을 경우, 친부가 아닌 추정부의 배제능력, 추정부가 친부인 경우에는 보다 높은 친부가능성을 확보할 수 있다.

STR marker	추정부		자		부계지수 (PI)
	allele 1	allele 2	allele 1	allele 2	
CSF1PO	13	12	12	12	1,31
D5S818	13	13	13	11	4,13
D7S820	12	8	12	12	2,28
D13S317	12	8	12	8	2,38
TH01	7	7	9	7	2,02
TPOX	12	8	8	8	1,05
vWA	18	18	18	17	2,75
D3S1358	18	17	17	15	1,36
D8S1179	15	11	15	15	3,55
D16S3253	9	9	11	9	7,82
D18S51	16	13	14	13	1,36
D21S11	32,2	30	32,2	30	3,96
FGA	23	21	22	21	1,77
PentaD	12	11	11	10	1,54
PentaE	12	10	20	12	2,21
D12S391	20	17	21	17	2,14
D14S608	12	11	12	11	2,96
결합부계지수 (CPI)					1,325,725
친부(생물학적 친부) 가능성					99,99992%

표 3. 17개 STR marker의 분석을 통한 추정부와 자의 DNA typing 결과 및 친부 가능성 각 STR marker에서 추정부와 자의 1개 allele이 일치함이 확인되었으며, 각 STR marker에서의 PI, CPI 및 친부 가능성을 구했다.

**맺음말**

국가적인 차원의 강력범죄자 유전자 DB 구축, 미야 유전자 DB 구축 등은 개인정보의 국가기관 운용에 따르는 사생활 및 인권침해라는 의견과 함께, 범죄수사의 신뢰도 향상, 미야찾기 활성화 등 공익적 목적이 크다는

의견으로 찬반양론이 팽팽한 상태에서 준비 중이거나 진행되고 있다. 반면 몇몇 바이오벤처를 중심으로 DNA typing을 이용한 가족확인 등의 서비스가 실시되고 있다. 하지만 바이오벤처들의 공신력에 대한 일반인들의 의구심이 끊임없이 제기되고 있다. 해당 업체가 DNA typing에 대한 이론적인 배경이 뒷받침 되어있는지, 한국인을 대표할 수 있는 자체의 유전자 DB를 확보하고 있는지, 검사에 적용하고 있는 DNA marker의 선정은 합리적으로 이루어지고 있는지, 개인의 신상 및 유전정보 관리가 완벽한 보안의 틀 안에서 관리가 되는지, 회사의 영업활동이 건전하고 공익에 기여하고 있는지 등의 검토가 이루어져야할 것이다.

표.4.에서는 의뢰인이 DNA typing업체의 선정 시 검토해야할 몇 가지 사항과 함께 이와 관련된 현재 본 센터에서 적용하고 있는 기준을 나열하고 있다.

검토내용	한국유전자검사의 기준
이론적 배경을 갖추었는가?	전문교육을 받은 석박사급 연구원에 의한 통계학적인 검토가 이루어지고 있음.
한국인을 대표하는 유전자DB를 확보하였는가?	수백 명 이상의 한국인을 토대로 구성된 독자적 STR DB와 mtDNA DB 등을 확보하고 있다.
STR marker의 선택은 합리적인가?	FBI 13CODIS의 모든 STR marker와 한국인에게 적합하다 판단되는 자체의 추가 STR marker를 적용하여 FBI의 검사결과보다 신뢰도 높은 결과 도출가능.
실험자의 능력과 실험법은 검증되었는가?	FBI AABB(미국혈액은행협회)에서 마련한 기준에 부합되는 인력, 장비 및 검사방법을 통해 검사를 진행. 주기적으로 자체검증 프로그램 실시.
개인 신상 및 유전정보 관리는 완벽한가?	모든 검체는 익명으로 검사 진행 됨. 온라인을 통한 개인 신상, 유전정보 관리 금하고 있음.
유전자검사의 결과가 공익을 위해 사용되는가?	입양기관의 입양아 친부모확인 작업지원, 2-3세대 해외교포의 국적회복업무 지원, 장기미야 확인업무 지원 등의 공익사업 진행.

표 4. 검사기관 선정에 있어서의 검토사항 및 한국유전자검사센터의 기준

국가 및 민간주도의 DNA typing 모두는 몇몇 가지 면에서 국민들의 우려를 사고 있음이 사실이나, 다른 방법으로는 확보할 수 없는 수준의 정확한 검사결과로 사회에 기여하고 있다는 부분이 간과되어서는 안 될 것이다. Forensic DNA typing이 국민의 신뢰를 얻고 사회에 이바지하기 위해서는 끊임없는 연구개발과 정확한 결과를 도출하기 위한 스스로의 노력 이 경주되어야 할 것이다.

## 유전자 검사의 모든 것!

- 한우판별검사
- 동물성 성분검사
- PES 돈육검사
- 가족확인 유전자검사



**Takara**  
한국 유전 자 검사 센터

TEL 031-739-3360 / Fax 031-739-3361  
<http://kgac.takara.co.kr>