

GENETIKA ÉS POPULÁCIÓGENETIKA

Genetikai polimorfizmusok



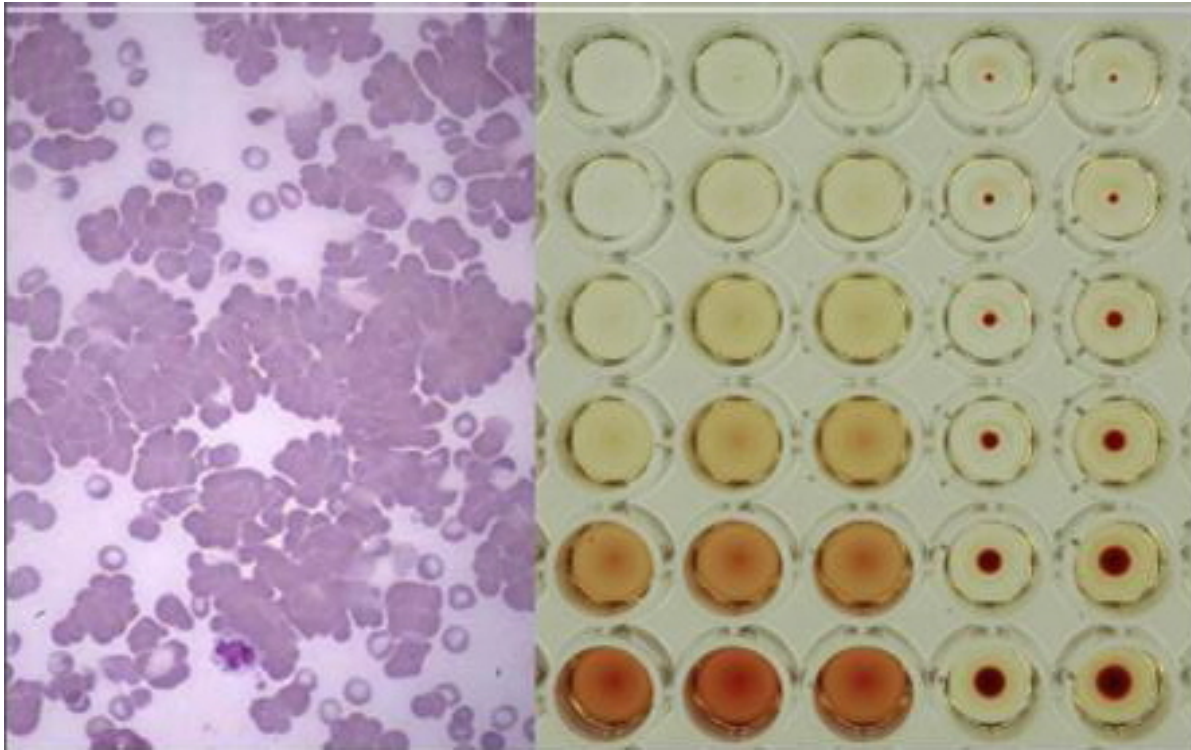
ELTE TTK Genetikai Tanszék

Recesszív gén hőmérséklet szenzitív expressziója

- TYR gén ► tirozináz enzim ► melanin szintézis (sötét szín)
 - Defektív tirozináz ► funkcióvesztés normál testhőmérsékleten
 - sötét színárnyalat csak az alacsonyabb testhőmérsékletű helyeken



Első genetikai marker: ABO vércsoport-rendszer



Landsteiner, 1900



FELIX BERNSTEIN (1933)

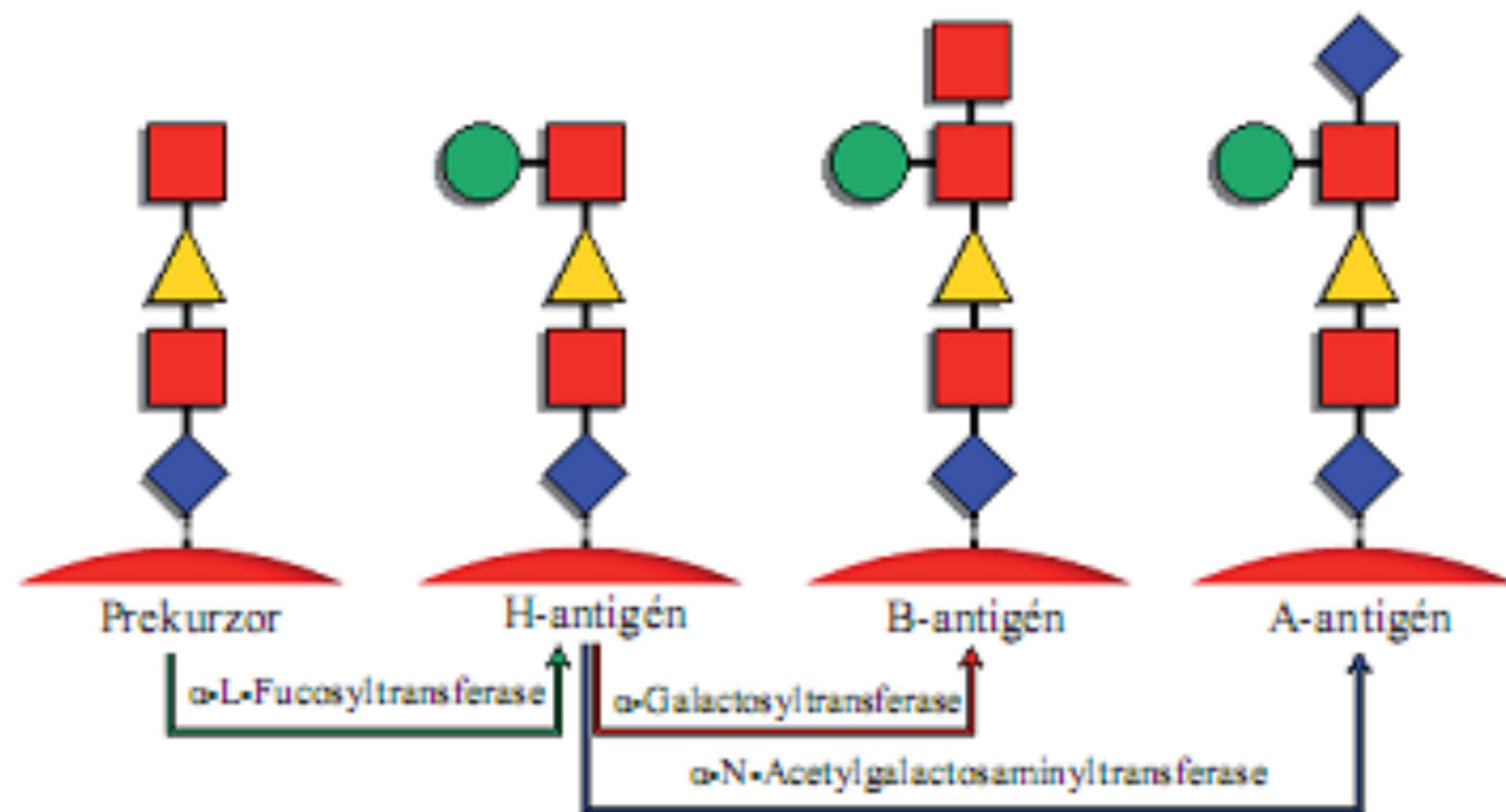
Two hypotheses of blood group inheritance

Group	VON DUNGERN and HIRZFELD		BERNSTEIN		Observed proportion
	Genotype	Expected proportion	Genotype	Expected proportion	
O	$aa\ bb$	$p_a^2\ p_b^2$	OO	p_O^2	0.294
A	$A-\ bb$	$(1 - p_a^2)p_b^2$	AA, OA	$p_A^2 + 2p_O p_A$	0.422
B	$aa\ B-$	$p_a^2(1 - p_b^2)$	BB, OB	$p_B^2 + 2p_O p_B$	0.206
AB	$A-\ B-$	$(1 - p_a^2)(1 - p_b^2)$	AB	$2p_A p_B$	0.078
Total		1		1	1.000

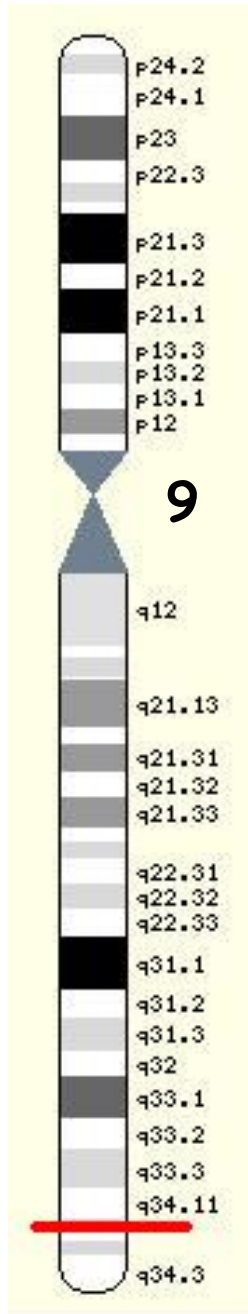
The expected proportions assume Hardy-Weinberg ratios and linkage equilibrium. The observed proportions are from 502 Japanese (BERNSTEIN 1925).

- L-Fucose
- D-Galactose
- ◆ N-Acetylgalactosamine
- ▲ N-Acetylglucosamine

AB0 antigének kialakulása



Various Alleles at the ABO Locus



Exon Number	6				7												
	2	2	4	5	6	6	6	7	7	7	8	8	8	8	9	1	1
Nucleotide Position	6	9	6	2	4	5	8	0	7	9	0	0	2	7	3	0	0
A alleles																	
A101	G	A	C	C	T	C	G	G	C	C	G	G	G	G	G	C	C
A102	*	*	T	*	*	*	*	*	*	*	*	*	*	*	*	*	*
A201	*	*	T	*	*	*	*	*	*	*	*	*	*	*	*	*	Δ
A301	*	*	*	*	*	*	*	*	*	*	*	*	*	*	A	*	*
Ax01	*	*	*	*	A	*	*	*	*	*	*	*	*	*	*	*	*
<i>cis</i> -AB01	*	*	T	*	*	*	*	*	*	*	*	C	*	*	*	*	*
B alleles																	
B101	*	G	*	G	*	T	*	A	*	A	*	C	*	*	A	*	*
B301	*	G	*	G	*	T	*	A	*	A	*	C	*	*	A	T	*
B(A)01	*	G	*	G	*	*	*	*	*	A	*	C	*	*	A	*	*
O alleles																	
O01	Δ	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
O02	Δ	G	*	*	A	*	A	*	T	*	*	*	A	*	*	*	*
O03	*	G	*	G	*	*	*	*	*	*	A	*	*	*	*	*	*
Possible Amino Acid Change	Frameshift	No change	P156L	R176G	F216I	No change	No change	G235S	No change	L266M	G268R	G268A	V277M	D291N	No change	R352W	Frameshift

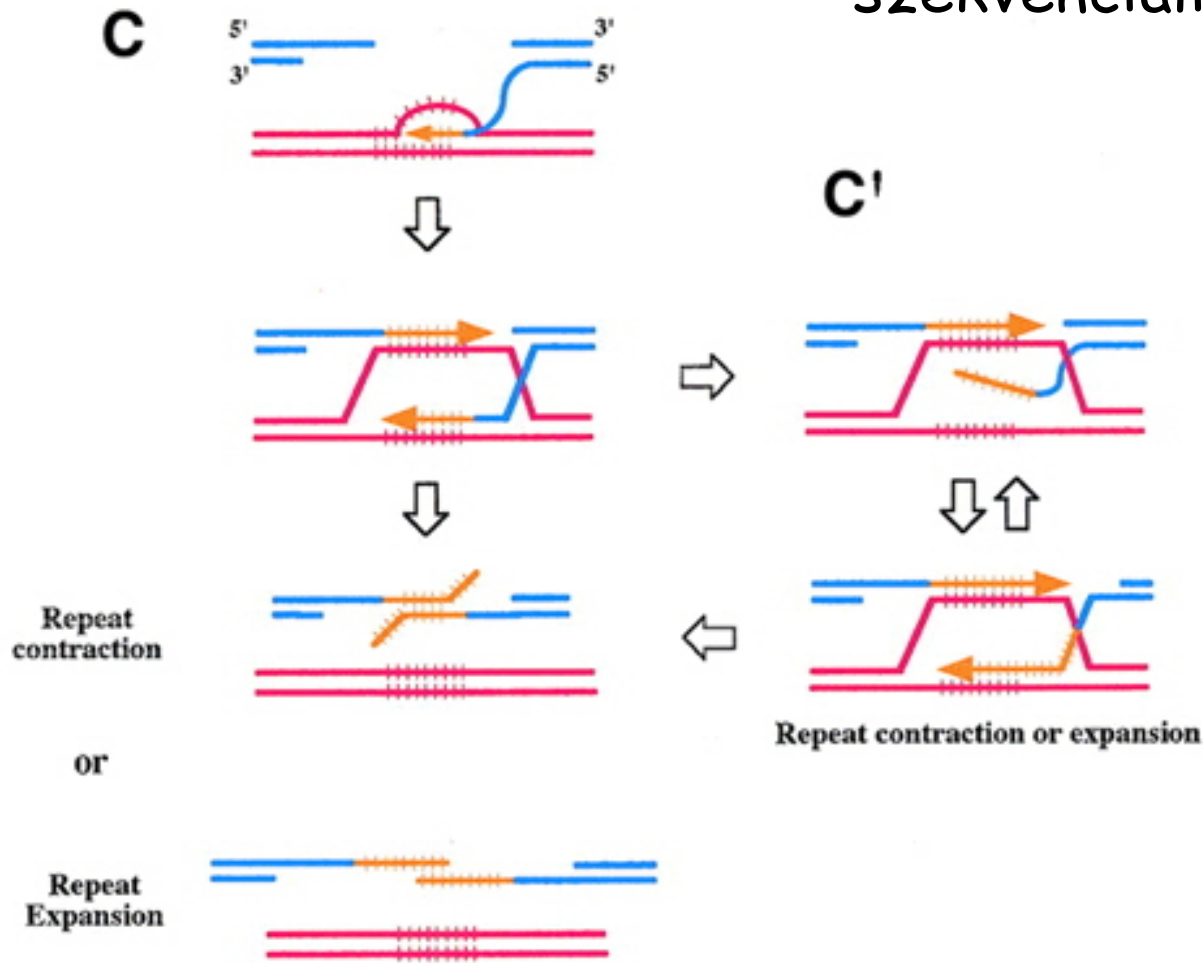
A humán genom projekt eredményei

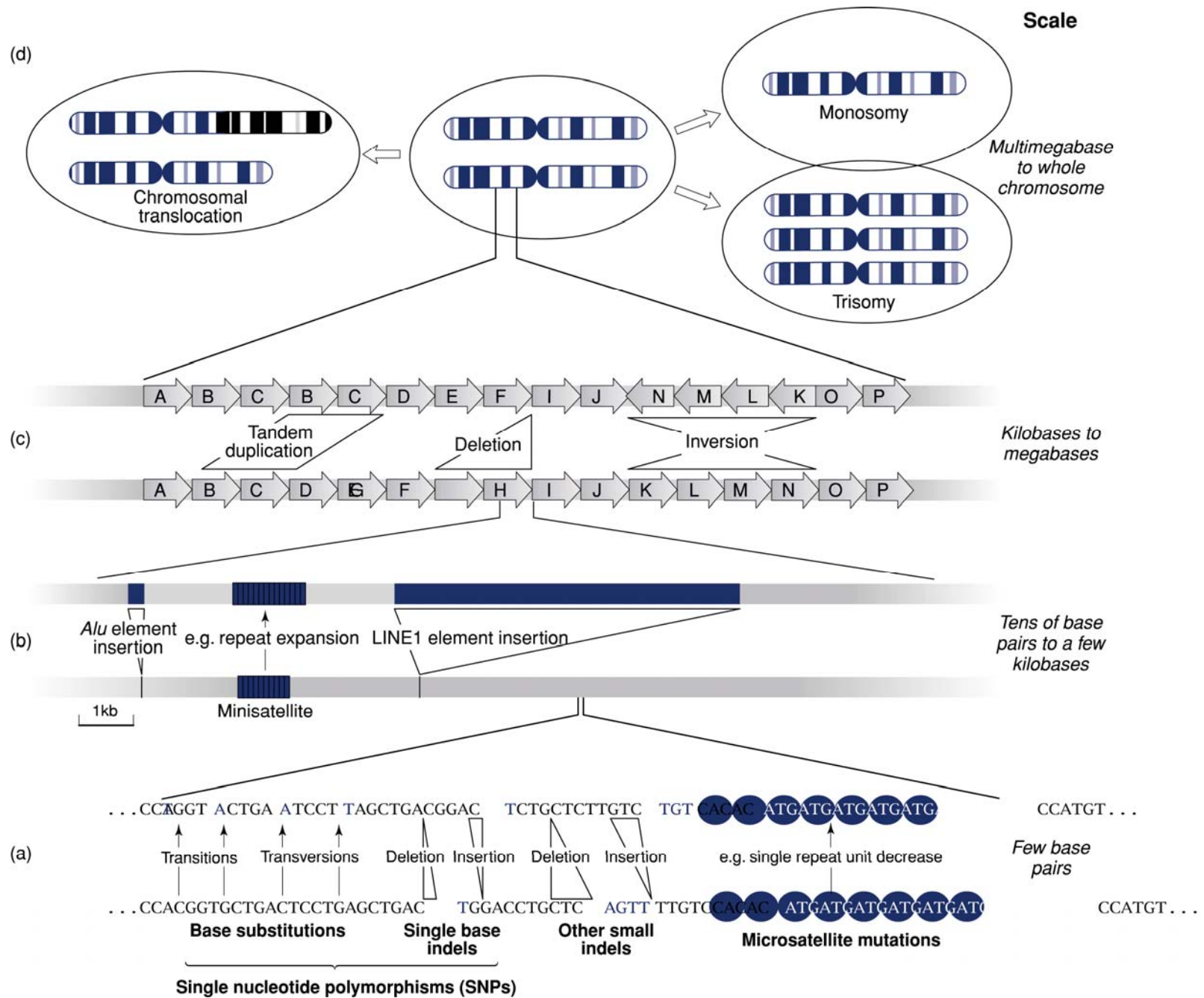
- Előzetes publikálás 2001-ben (Science, Nature)
- Az eddig leírt legnagyobb méretű teljes genom
- Az eukarióta genomra jellemző szerkezeti és szerveződési tulajdonságok (modellszervezetek)
- Meglepően alacsony a fehérje kódoló gének száma (~22000)
- RNS gének növekvő száma (tRNA, snRNA, miRNA, stb.)
- Orphan gének: nincs homológja, humán gének ~ 1/3-a
- Protein kódoló szekvenciák aránya: < 1 %
- **Repetitív szekvenciák túlnyomó aránya!**

REKOMBINÁCIÓ

Polimorfizmus hajtóereje:

- pontmutáció
- szekvenciális átrendeződés



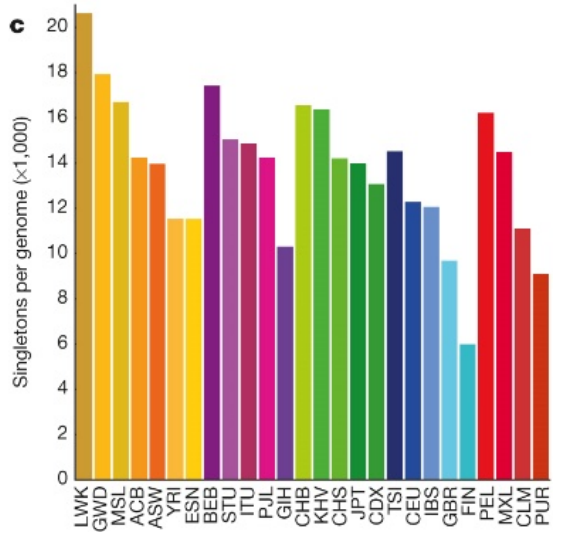
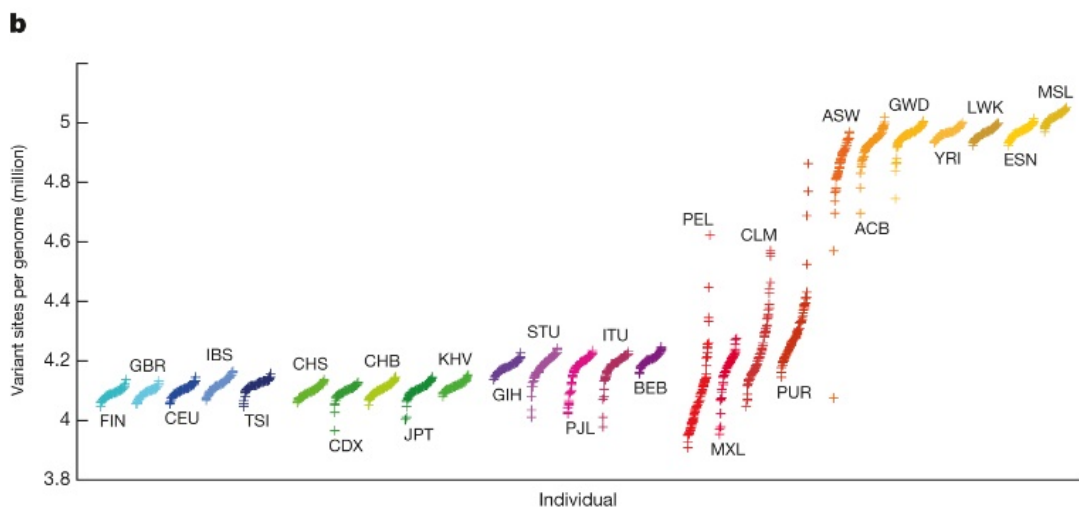
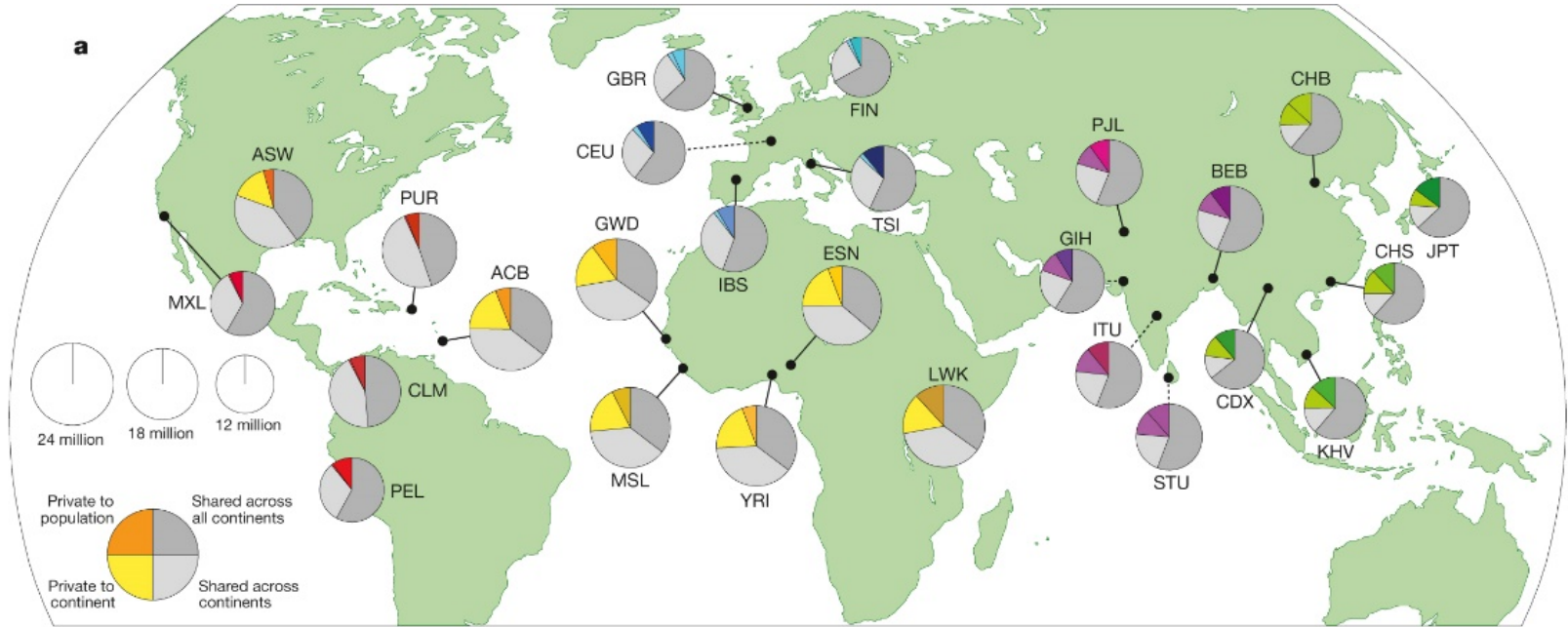


A global reference for human genetic variation

The 1000 Genomes Project Consortium*

The 1000 Genomes Project set out to provide a comprehensive description of common human genetic variation by applying whole-genome sequencing to a diverse set of individuals from multiple populations. Here we report completion of the project, having reconstructed the genomes of 2,504 individuals from 26 populations using a combination of low-coverage whole-genome sequencing, deep exome sequencing, and dense microarray genotyping. We characterized a broad spectrum of genetic variation, in total over 88 million variants (84.7 million single nucleotide polymorphisms (SNPs), 3.6 million short insertions/deletions (indels), and 60,000 structural variants), all phased onto high-quality haplotypes. This resource includes >99% of SNP variants with a frequency of >1% for a variety of ancestries. We describe the distribution of genetic variation across the global sample, and discuss the implications for common disease studies.

Population sampling



A global reference for human genetic variation

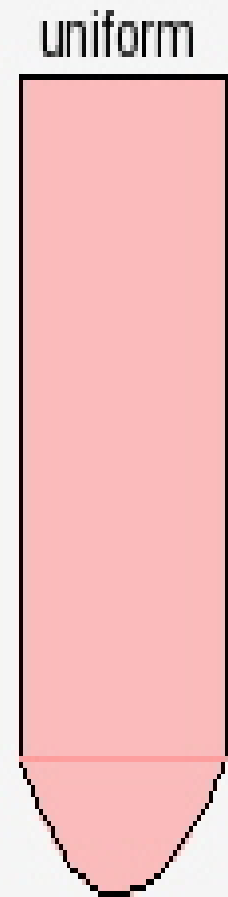
The 1000 Genomes Project Consortium*

Table 1 | Median autosomal variant sites per genome

	AFR		AMR		EAS		EUR		SAS	
Samples	661		347		504		503		489	
Mean coverage	8.2		7.6		7.7		7.4		8.0	
	Var. sites	Singletons	Var. sites	Singletons	Var. sites	Singletons	Var. sites	Singletons	Var. sites	Singletons
SNPs	4.31M	14.5k	3.64M	12.0k	3.55M	14.8k	3.53M	11.4k	3.60M	14.4k
Indels	625k	-	557k	-	546k	-	546k	-	556k	-
Large deletions	1.1k	5	949	5	940	7	939	5	947	5
CNVs	170	1	153	1	158	1	157	1	165	1
MEI (Alu)	1.03k	0	845	0	899	1	919	0	889	0
MEI (L1)	138	0	118	0	130	0	123	0	123	0
MEI (SVA)	52	0	44	0	56	0	53	0	44	0
MEI (MT)	5	0	5	0	4	0	4	0	4	0
Inversions	12	0	9	0	10	0	9	0	11	0
Nonsynon	12.2k	139	10.4k	121	10.2k	144	10.2k	116	10.3k	144
Synon	13.8k	78	11.4k	67	11.2k	79	11.2k	59	11.4k	78
Intron	2.06M	7.33k	1.72M	6.12k	1.68M	7.39k	1.68M	5.68k	1.72M	7.20k
UTR	37.2k	168	30.8k	136	30.0k	169	30.0k	129	30.7k	168
Promoter	102k	430	84.3k	332	81.6k	425	82.2k	336	84.0k	430
Insulator	70.9k	248	59.0k	199	57.7k	252	57.7k	189	59.1k	243
Enhancer	354k	1.32k	295k	1.05k	289k	1.34k	288k	1.02k	295k	1.31k
TFBSs	927	4	759	3	748	4	749	3	765	3
Filtered LoF	182	4	152	3	153	4	149	3	151	3
HGMD-DM	20	0	18	0	16	1	18	2	16	0
GWAS	2.00k	0	2.07k	0	1.99k	0	2.08k	0	2.06k	0
ClinVar	28	0	30	1	24	0	29	1	27	1

See Supplementary Table 1 for continental population groupings. CNVs, copy-number variants; HGMD-DM, Human Gene Mutation Database disease mutations; k, thousand; LoF, loss-of-function; M, million; MEI, mobile element insertions.

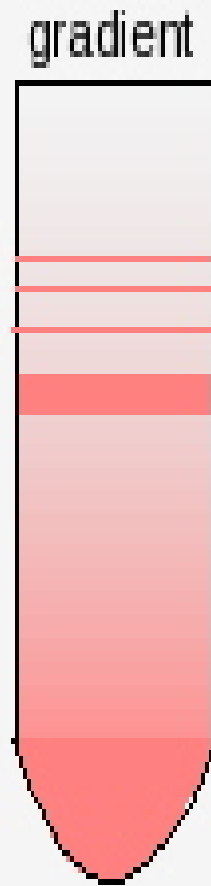
Satellite DNA



centrifugation



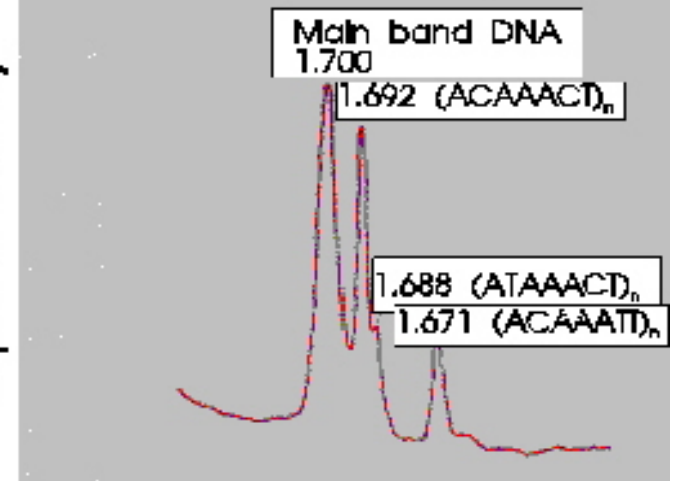
A large black arrow pointing from the uniform test tube to the gradient test tube, indicating the process of centrifugation.



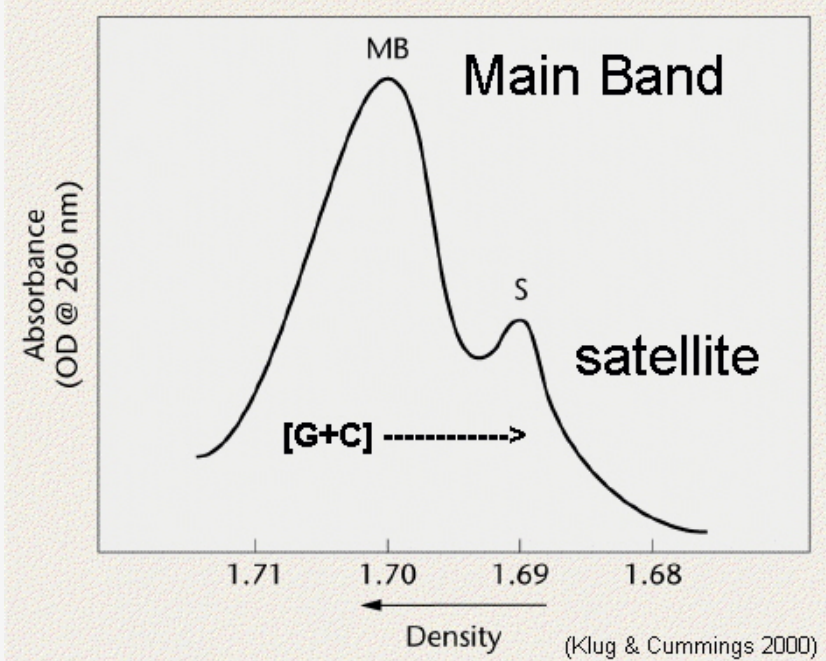
satellite bands

main band

Optical density

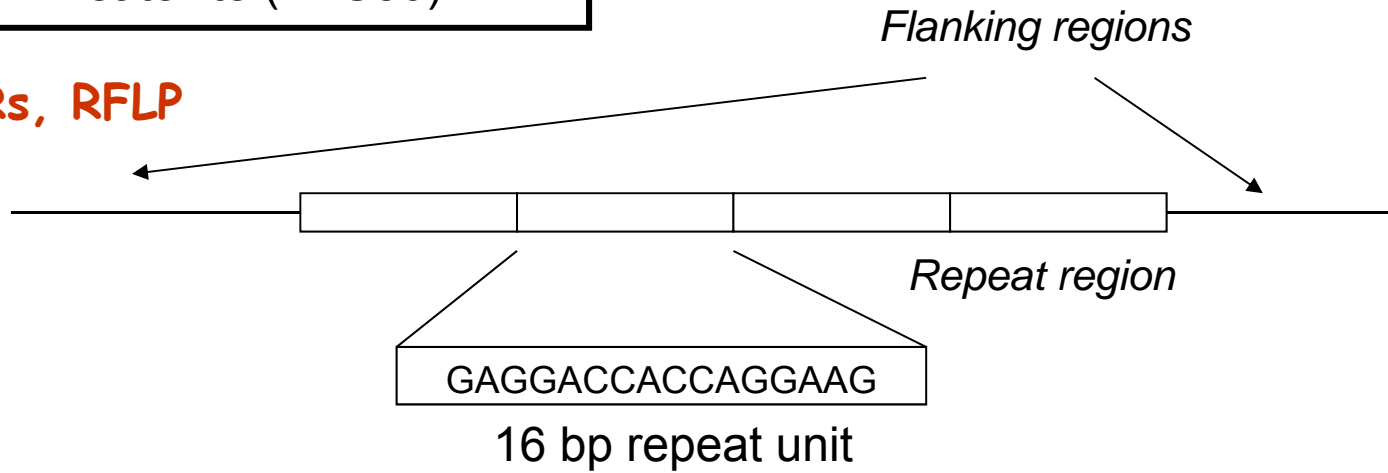


← Bouyant density in CsCl



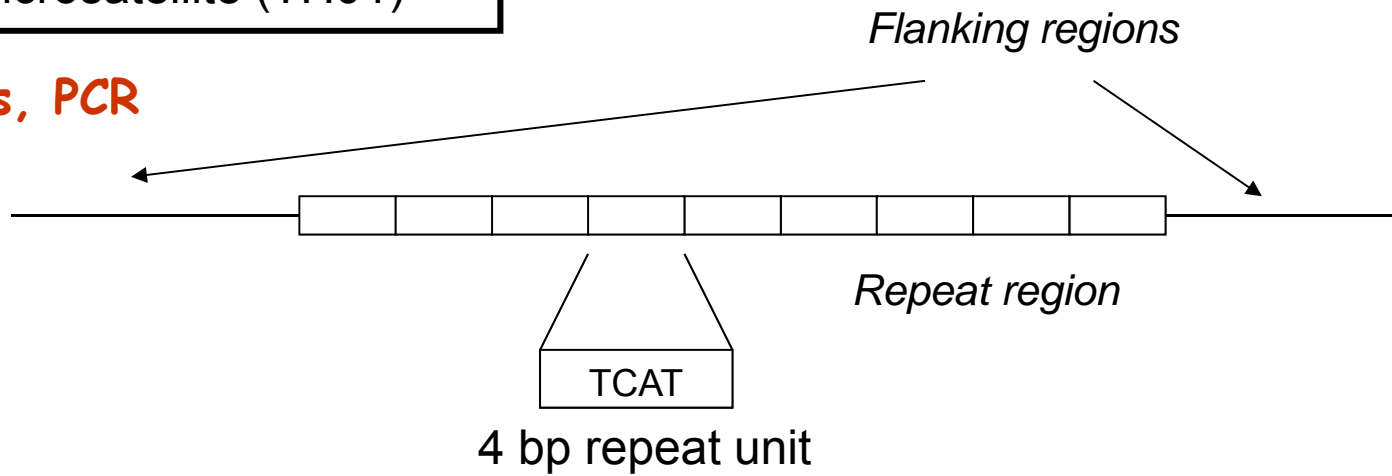
Minisatellite (D1S80)

VNTRs, RFLP



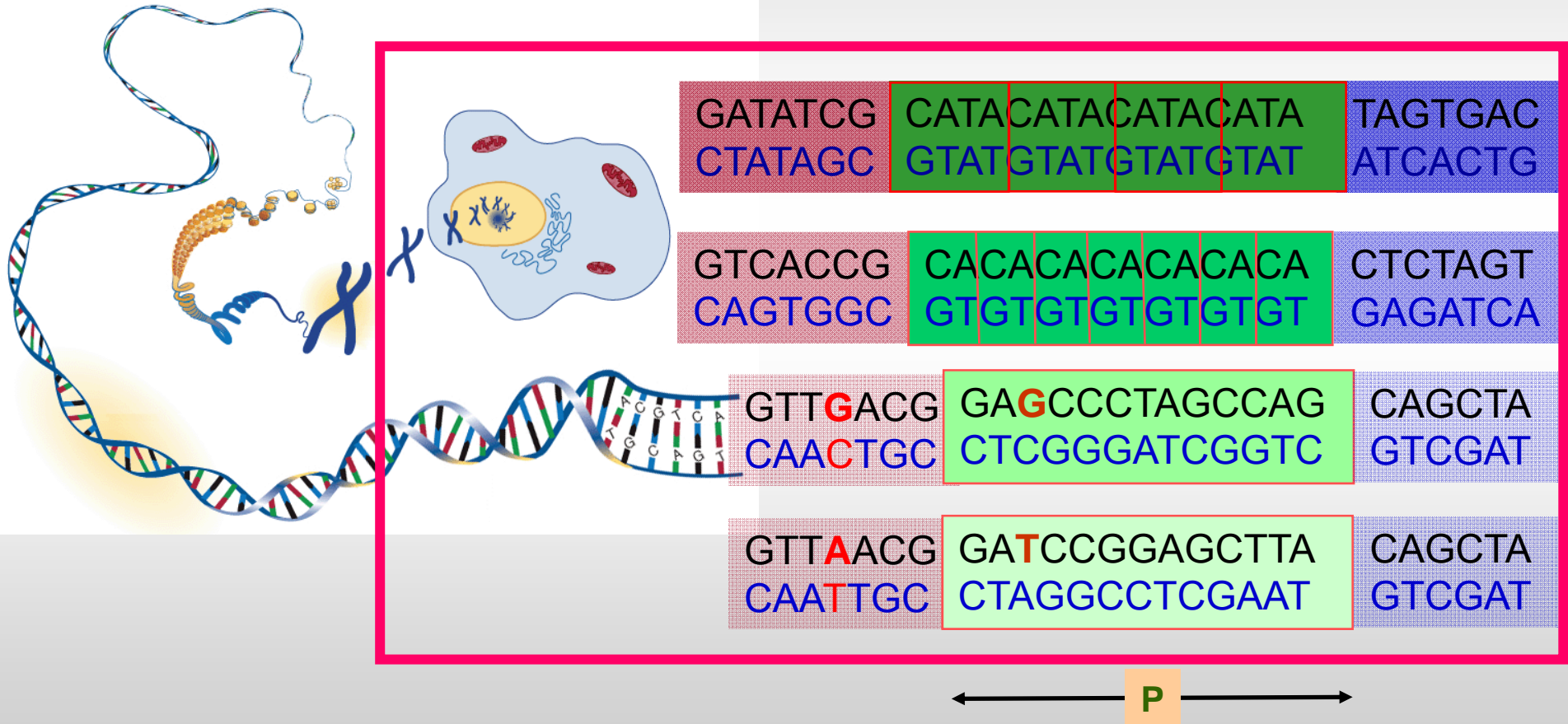
Microsatellite (TH01)

STRs, PCR



DNS polimorfizmusok

polimorfizmus - célrégió a PCR vizsgálatokban



P : polimorf célszakasz (marker, lokusz, allél)

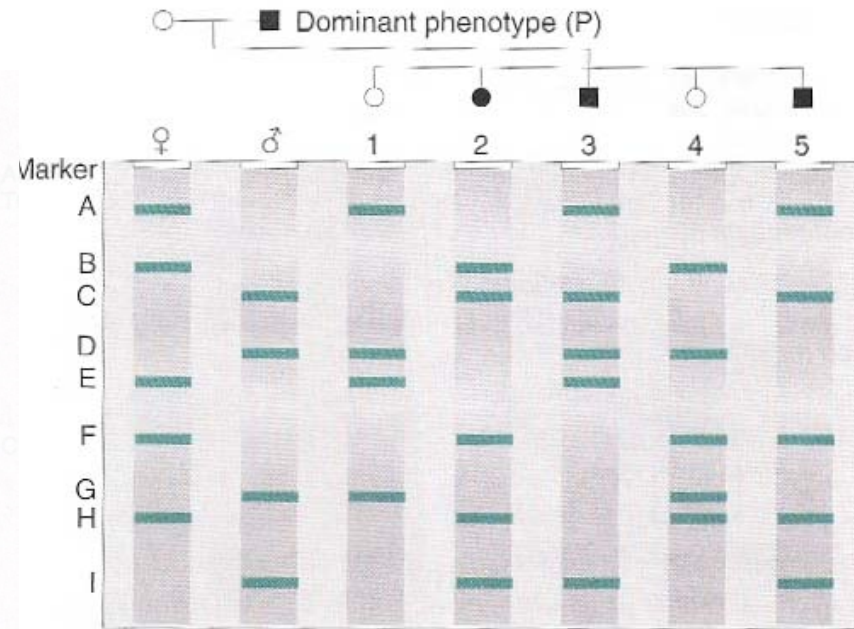


VNTR assay markerek: RFLP analitika

1985 - Sir Alec Jeffreys



1 2 3 Suspects 4 5 6 7

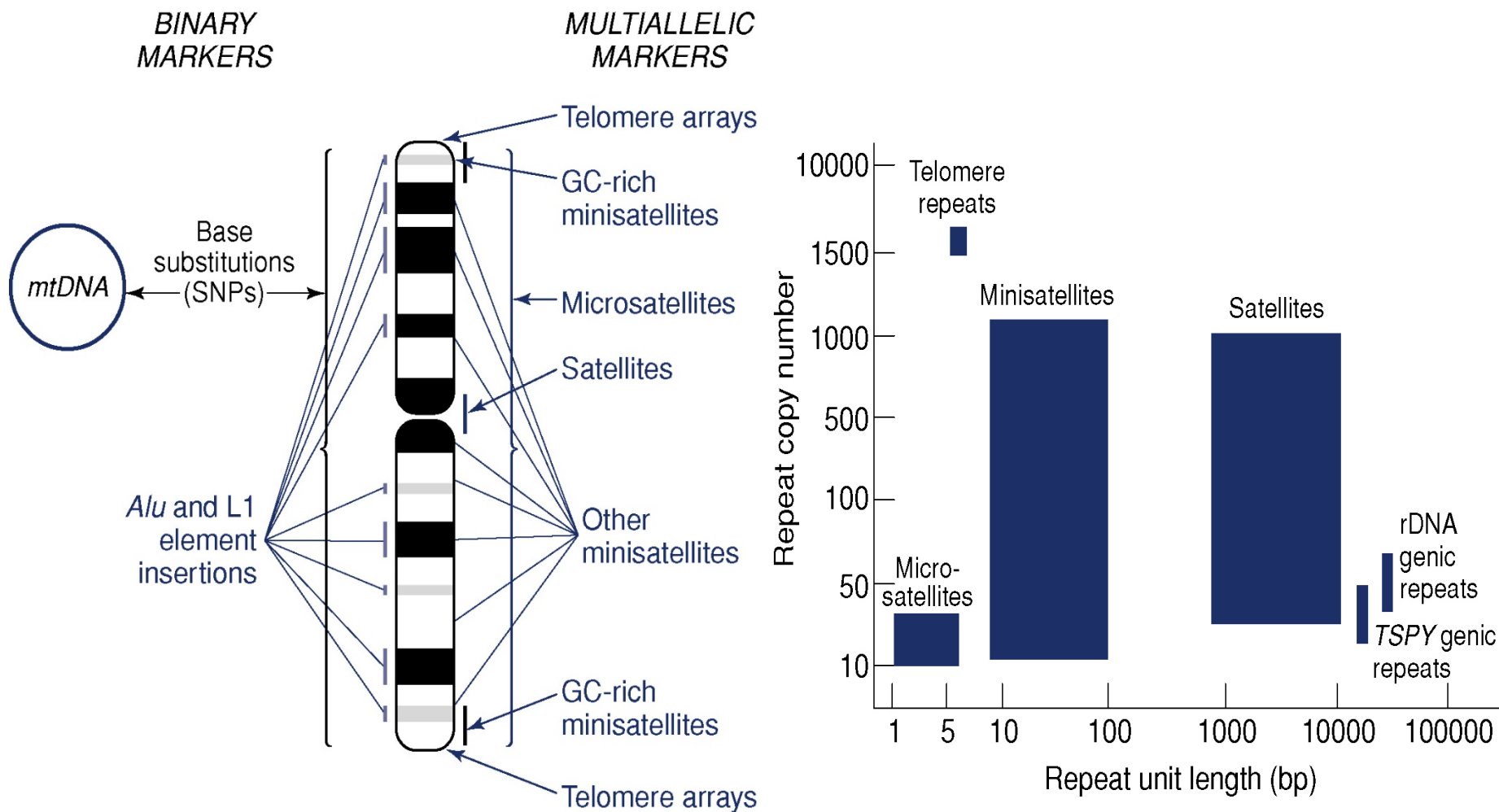


ANALYSIS EXAMPLES

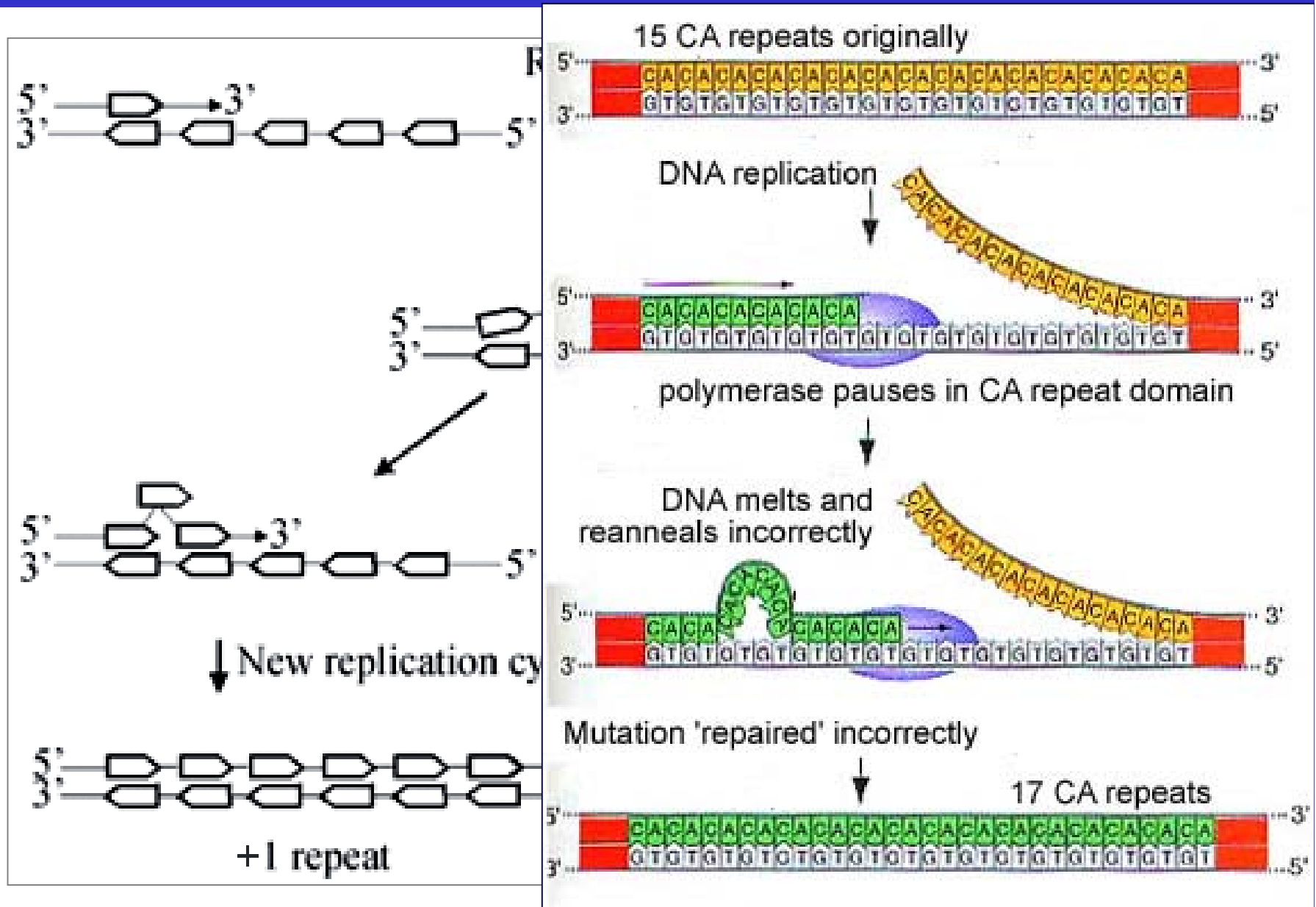
- F and H Always inherited together — linked?
- A and B In progeny, always *either* A or B — "allelic"?
- A and D Four combinations; A and D, A, D, or neither — unlinked?
- F, H, and E Always *either* F and H or E — closely linked in trans?
- Allele P Possibly linked to I and C.

Genetikai térképezés

Polimorf markerek eloszlása a genomban

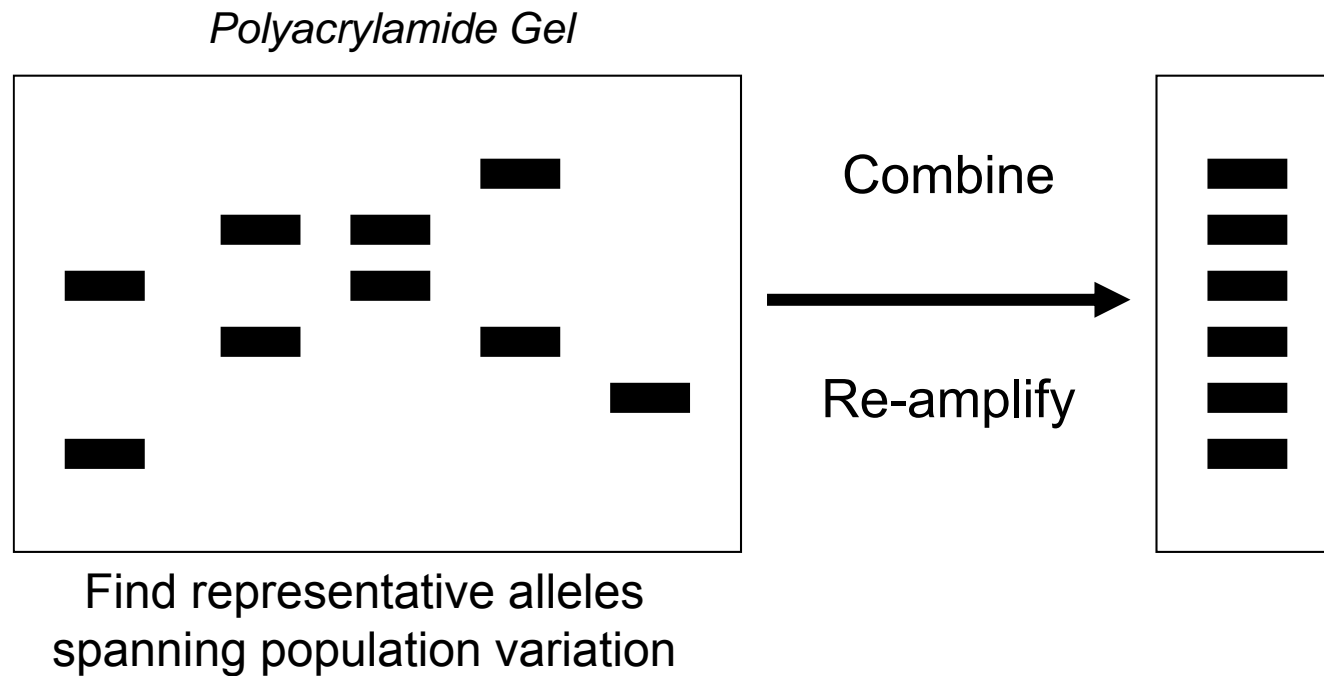


Mikroszatellita evolúció

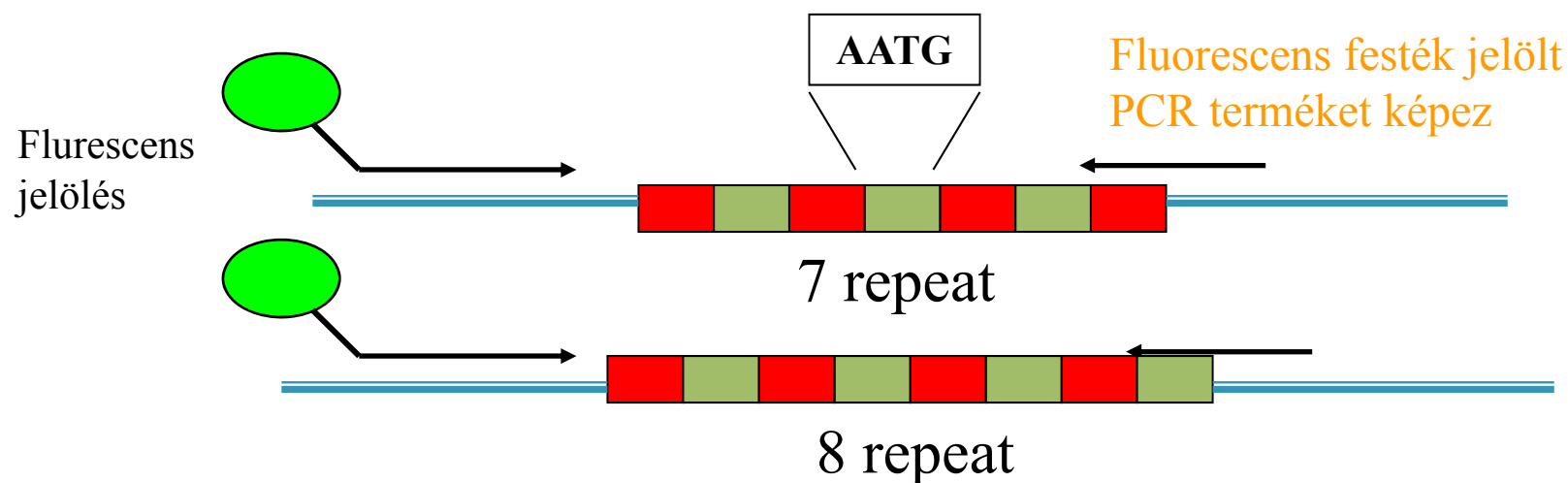


STR allélek genotipizálása: multiallélek

Separate PCR products from various samples amplified with primers targeted to a particular STR locus



Mikroszatellita - STR - markerek (Short Tandem Repeat)



a repeat régió variábilis az egyes minták között míg a flanking régió ahol a PCR primerek tapadnak állandó

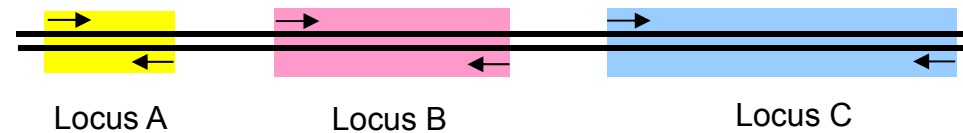
Homozigóta = a két allél megegyezik

Heterozigóta = allélek különbözők és egymástól elválaszthatók

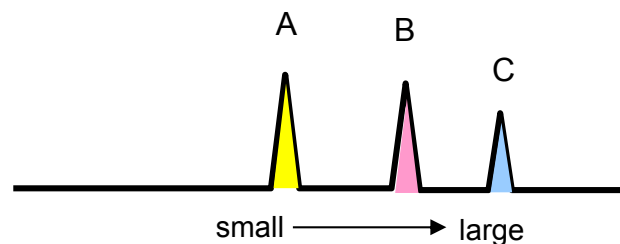
A primerek kötési helye meghatározza a PCR termék méretét!

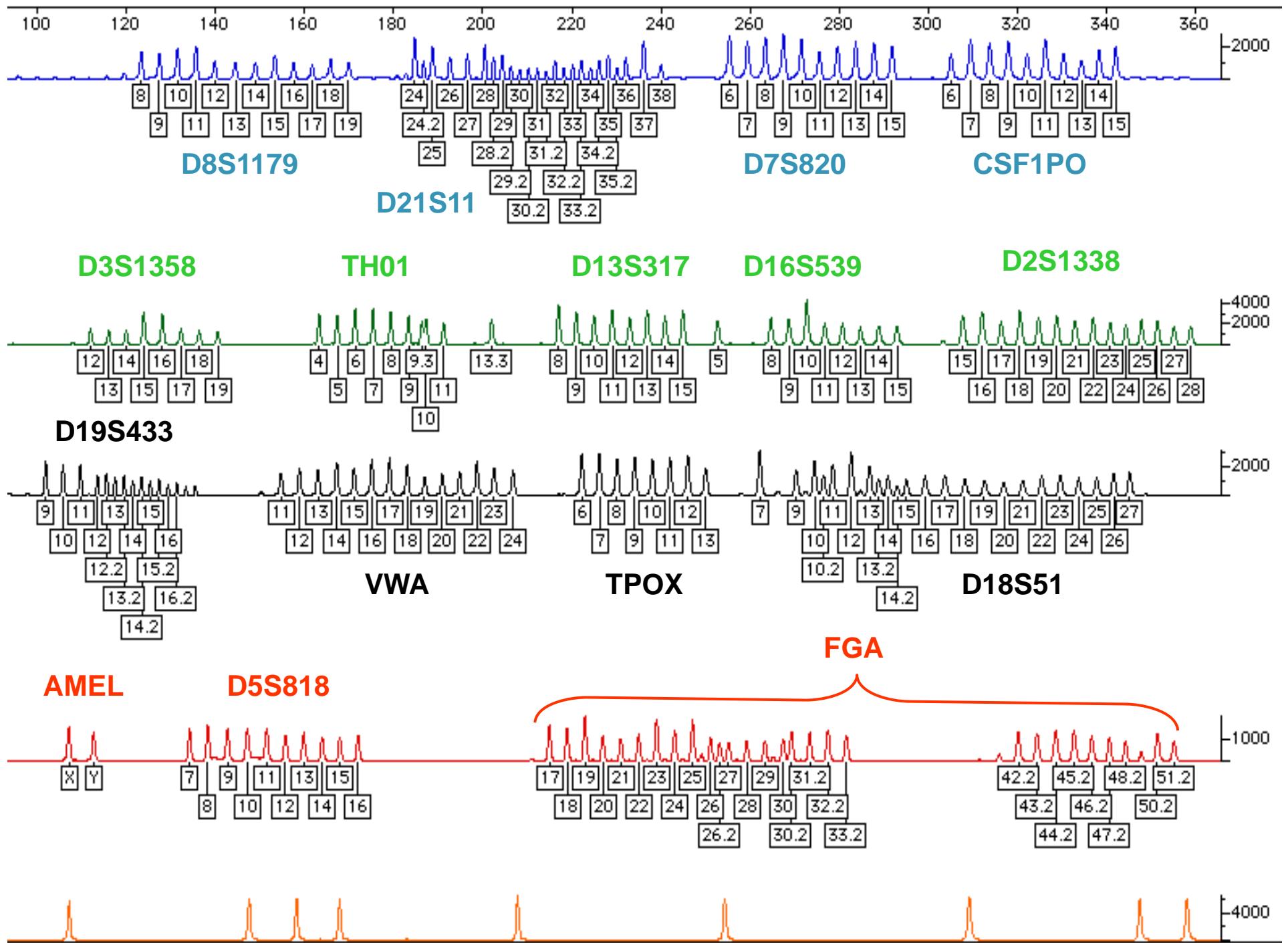
Multiplex - PCR

(A) Három lokusz szimultán amplifikálása egy reakcióban

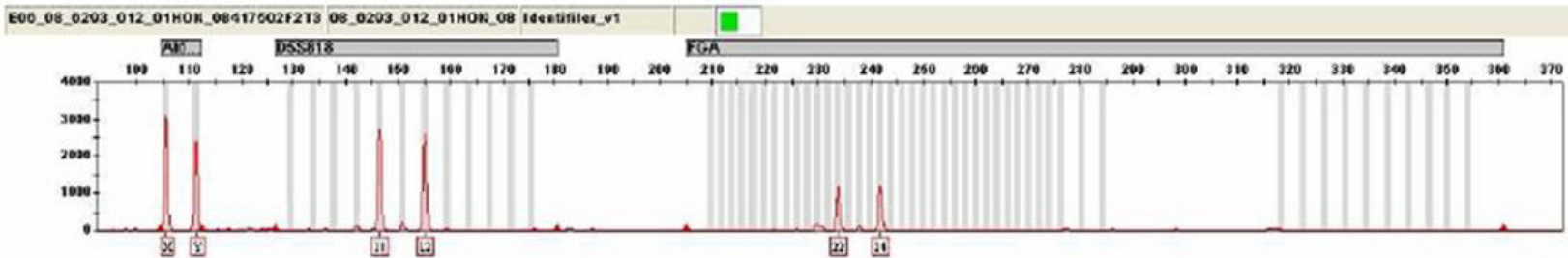
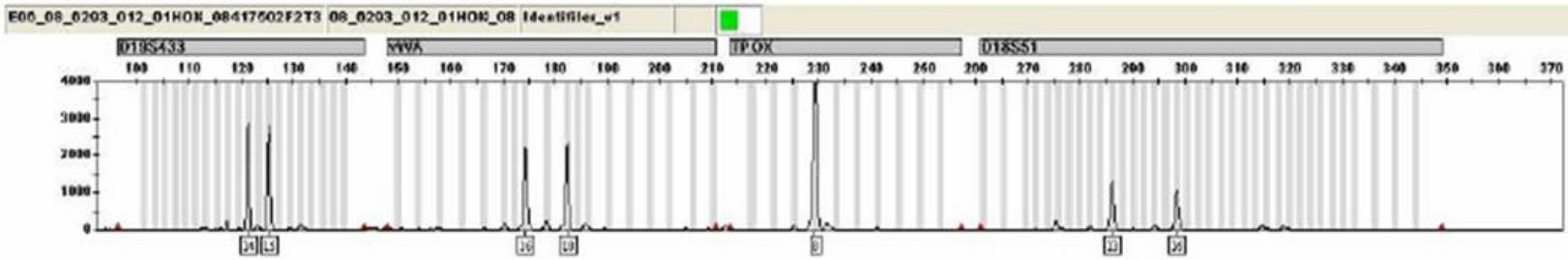
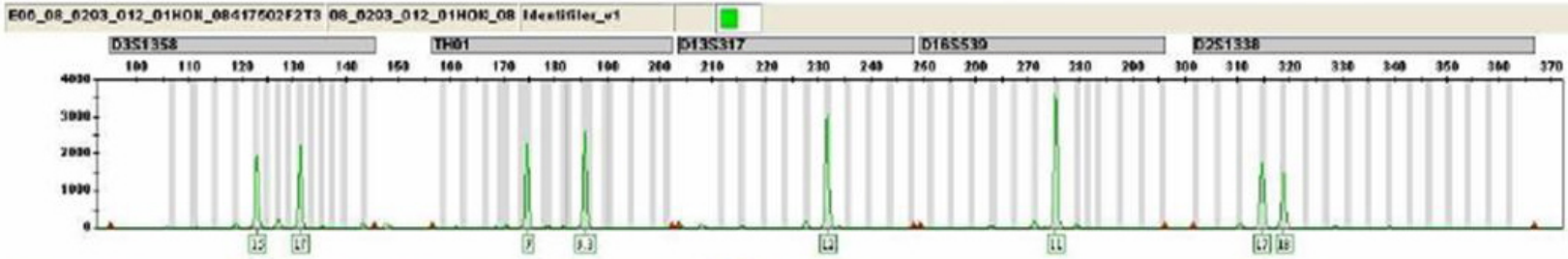
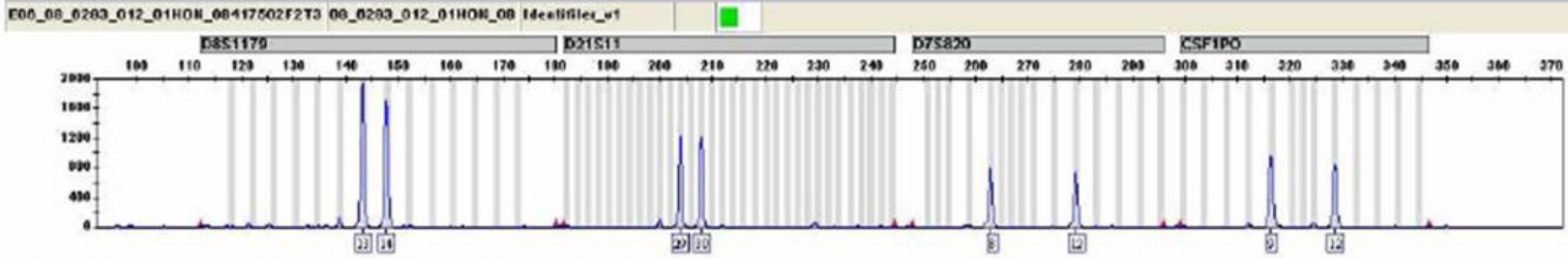


(B) PCR termékek elválasztása fragmens méret alapján





15 autoszómás lókuszt ...



Genotípus gyakoriságok statisztikai interpretálása

- **Allél-gyakorisági adatok** generálása a vizsgált populációban gyűjtött minták sorozatából (általában 100-200 nem rokon egyed / populáció autoszóma STR lókuszok esetén)
- **Allél-gyakorisági értékek meghatározása** minden lókuszon és az összes megfigyelt allélra
- Allél-gyakorisági értékek felhasználásával **a kérdéses DNS-profil előfordulási gyakoriságának kiszámolása**
 - Homozigóta (p^2), Heterozigóta ($2pq$)
 - Product rule: független lókuszok genotípus gyakoriságának multiplikálhatósága $PM = (P1)(P2)\dots(Pn)$

Bayes-féle hipotézis tesztelés

■ Vád hipotézise (H_p):

- Pl.: A helyszínen talált vércsoma XY-tól származott

■ Védelem ellenhipotézise (H_d):

- Pl.: A vércsoma az adott népességből véletlenszerűen kiválasztott másik, NN személytől származott

■ Bayes-formula (valószínűségi tesztelés)

$$\frac{\Pr(H_p|I)}{\Pr(H_d|I)} \times \frac{\Pr(E|H_p, I)}{\Pr(E|H_d, I)} = \frac{\Pr(H_p|E, I)}{\Pr(H_d|E, I)}$$

a priori
valószínűségi
hányados
(PrPR)

valószínűségi
hányados
(genetikai bizonyíték)

a posteriori
valószínűségi
hányados
(PoPR)

Pr: valószínűség

E: bizonyíték

I: információ

H: hipotézis

p: ügyész

d: ügyvéd

A valószínűségi hányados

- Ha a biológiai nyom egy személytől származott, és a laboratóriumi hiba kizárt, $\Pr(E | H_p, I) = 1$.

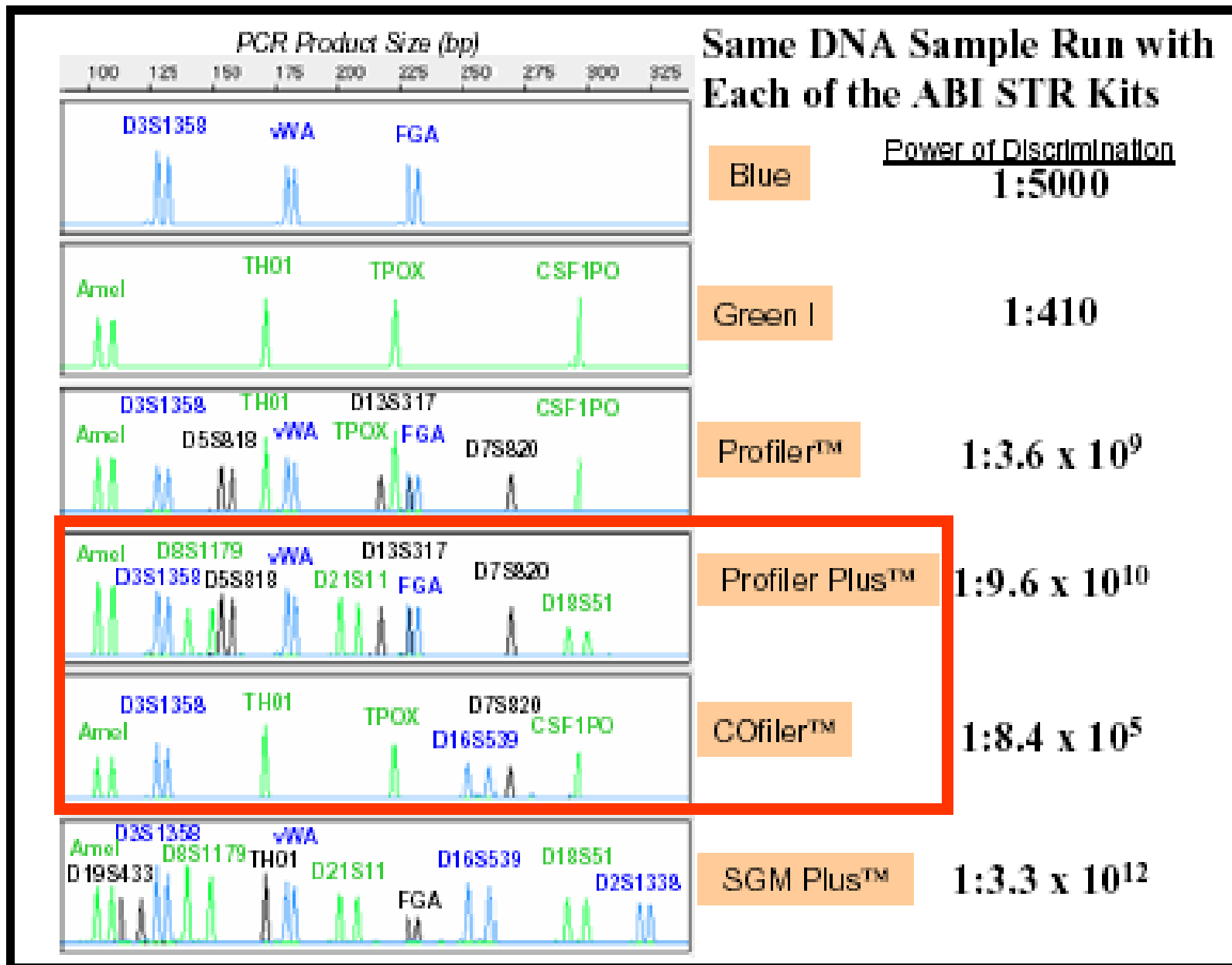
$$LR = \frac{\Pr(E | H_p, I)}{\Pr(E | H_d, I)} = \frac{1}{\Pr(G_c | G_s, H_d, I)} \approx \frac{1}{\Pr(G_c | H_d, I)}$$

Egyezési valószínűség

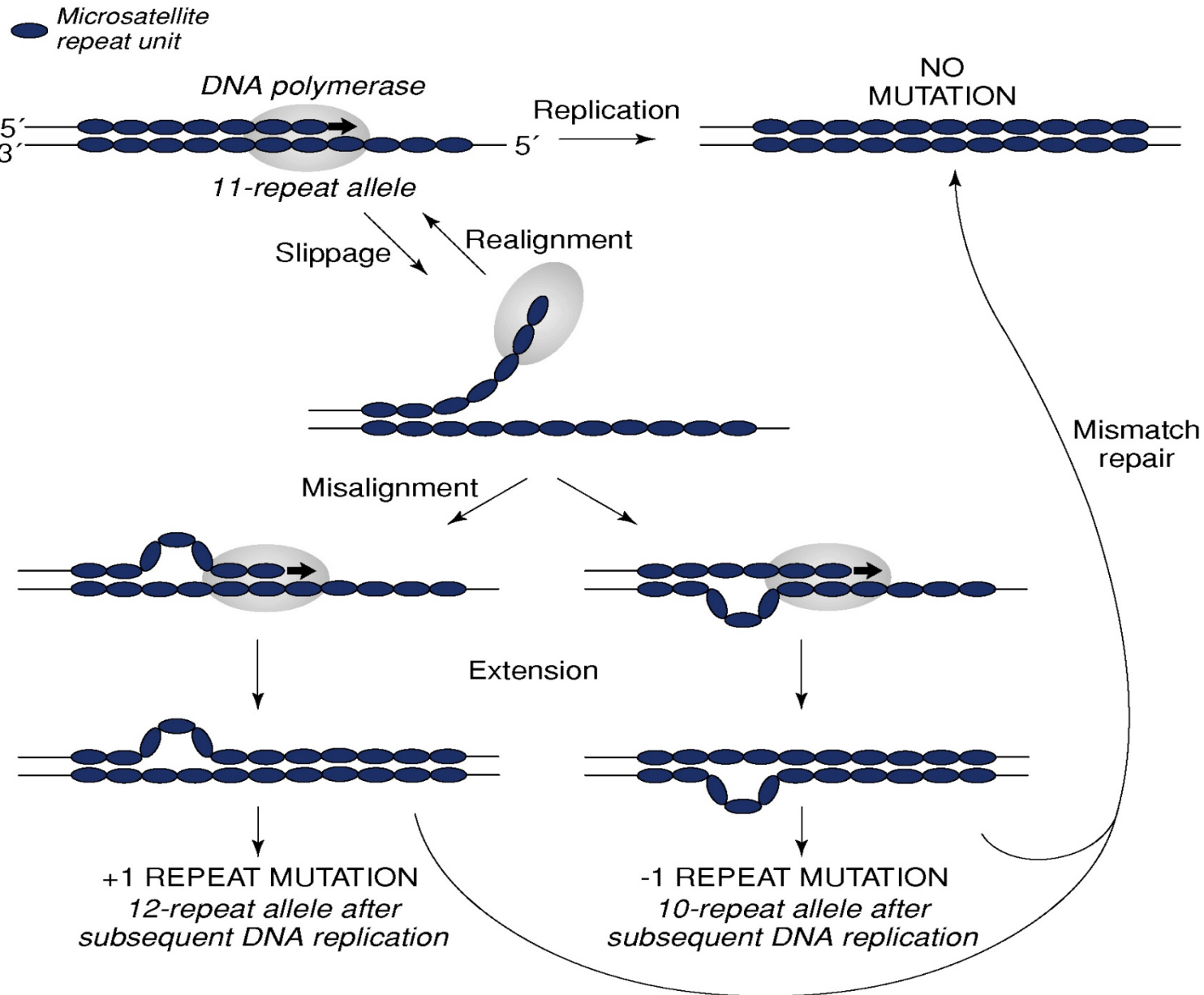
Profilgyakoriság

Pr: valószínűség
E: bizonyíték
G: genetikai profil
I: információ
H: hipotézis
p: ügyész
d: ügyvéd
c: helyszín
s: gyanúsított

- A valószínűségi hányados – legegyszerűbb formájában – a genetikai-profil előfordulási gyakoriságának reciprokával egyezik meg.



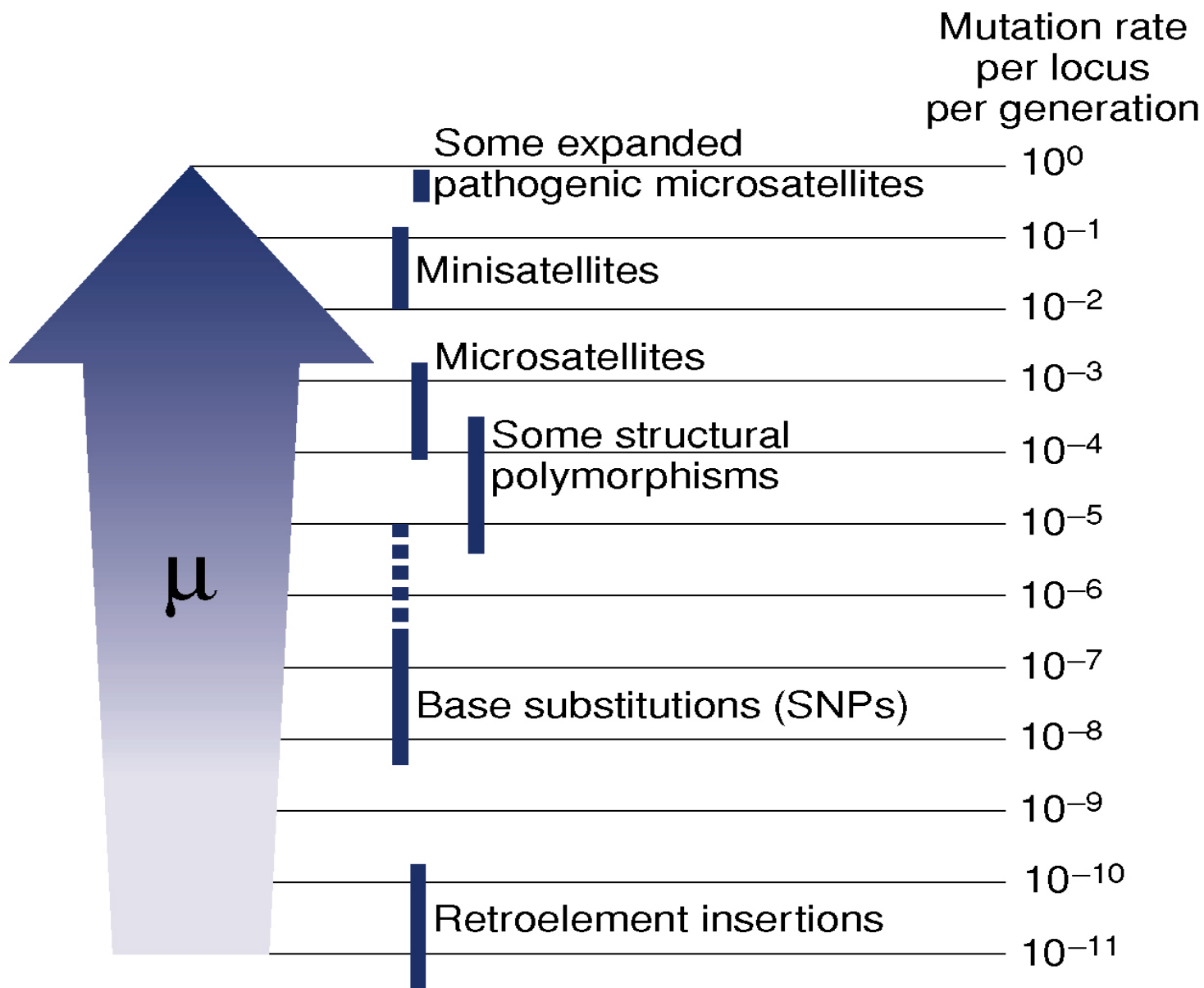
„Replikációs slippage” - Mikroszatellita mutáció



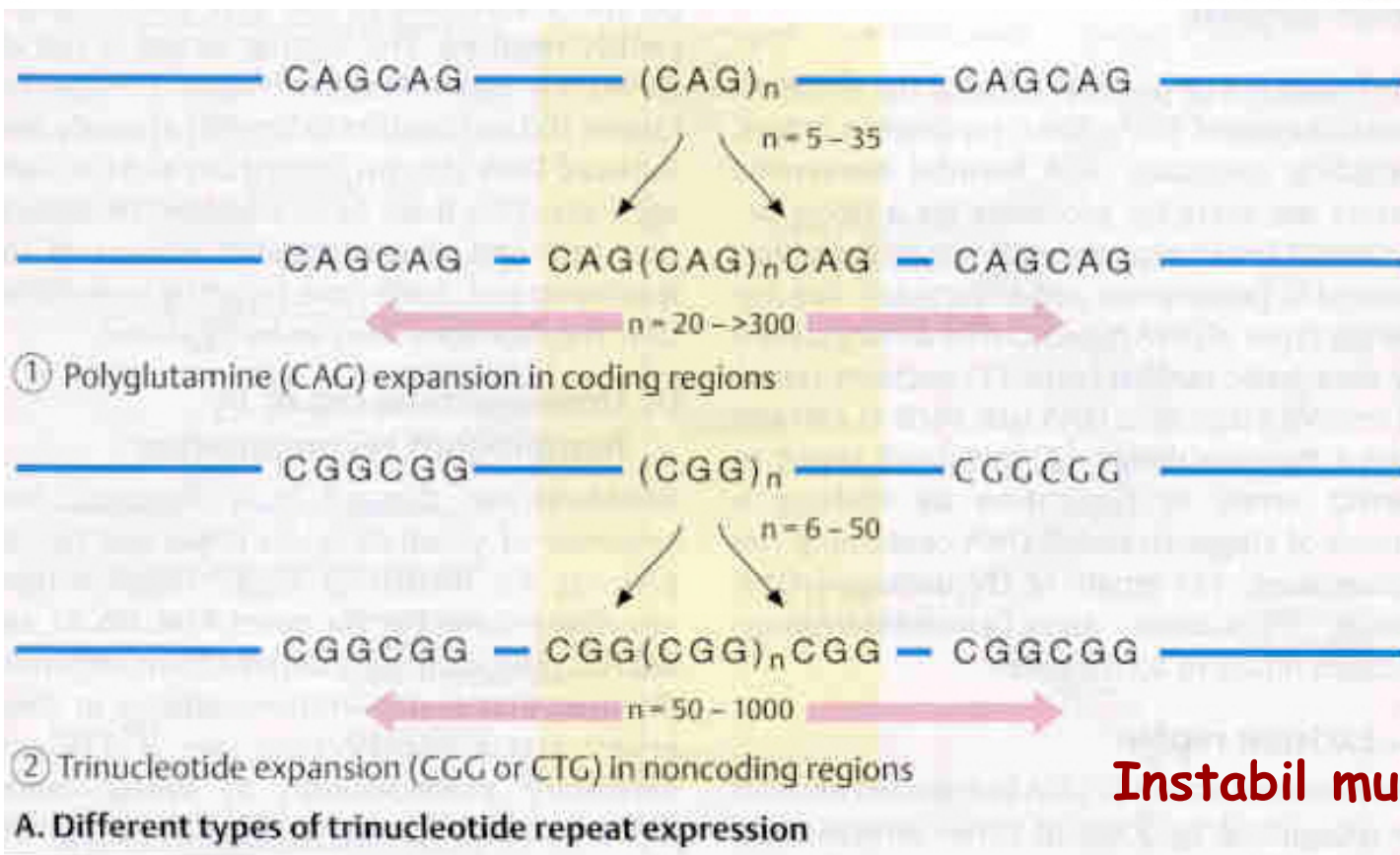
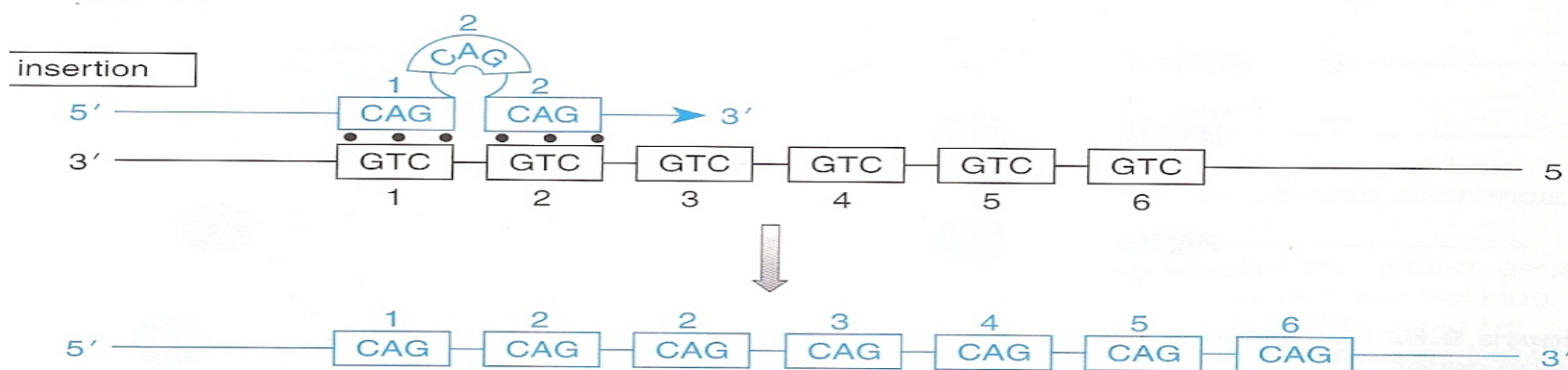
STR System	Maternal Meioses (%)	Paternal Meioses (%)	Number from either	Total Number of Mutations	Mutation Rate
CSF1PO	95/304,307 (0.03)	982/643,118 (0.15)	410	1,487/947,425	0.16%
FGA	205/408,230 (0.05)	2,210/692,776 (0.32)	710	3,125/1,101,006	0.28%
TH01	31/327,172 (0.009)	41/452,382 (0.009)	28	100/779,554	0.01%
TPOX	18/400,061 (0.004)	54/457,420 (0.012)	28	100/857,481	0.01%
VWA	184/564,398 (0.03)	1,482/873,547 (0.17)	814	2,480/1,437,945	0.17%
D3S1358	60/405,452 (0.015)	713/558,836 (0.13)	379	1,152/964,288	0.12%
D5S818	111/451,736 (0.025)	763/655,603 (0.12)	385	1,259/1,107,339	0.11%
D7S820	59/440,562 (0.013)	745/644,743 (0.12)	285	1,089/1,085,305	0.10%
D8S1179	96/409,869 (0.02)	779/489,968 (0.16)	364	1,239/899,837	0.14%
D13S317	192/482,136 (0.04)	881/621,146 (0.14)	485	1,558/1,103,282	0.14%
D16S539	129/467,774 (0.03)	540/494,465 (0.11)	372	1,041/962,239	0.11%
D18S51	186/296,244 (0.06)	1,094/494,098 (0.22)	466	1,746/790,342	0.22%
D21S11	464/435,388 (0.11)	772/526,708 (0.15)	580	1,816/962,096	0.19%
Penta D	12/18,701 (0.06)	21/22,501 (0.09)	24	57/41,202	0.14%
Penta E	29/44,311 (0.065)	75/55,719 (0.135)	59	163/100,030	0.16%
D2S1338	15/72,830 (0.021)	157/152,310 (0.10)	90	262/225,140	0.12%
D19S433	38/70,001 (0.05)	78/103,489 (0.075)	71	187/173,490	0.11%
SE33 (ACTBP2)	0/330 (<0.30)	330/51,610 (0.64)	None reported	330/51,940	0.64%

STR lókuszok mutációs rátája: 10^{-3} - 10^{-4} / meiózis

Polimorf szekvenciák mutációs rátája (μ)



Trinukleotid repeat expanzió I.

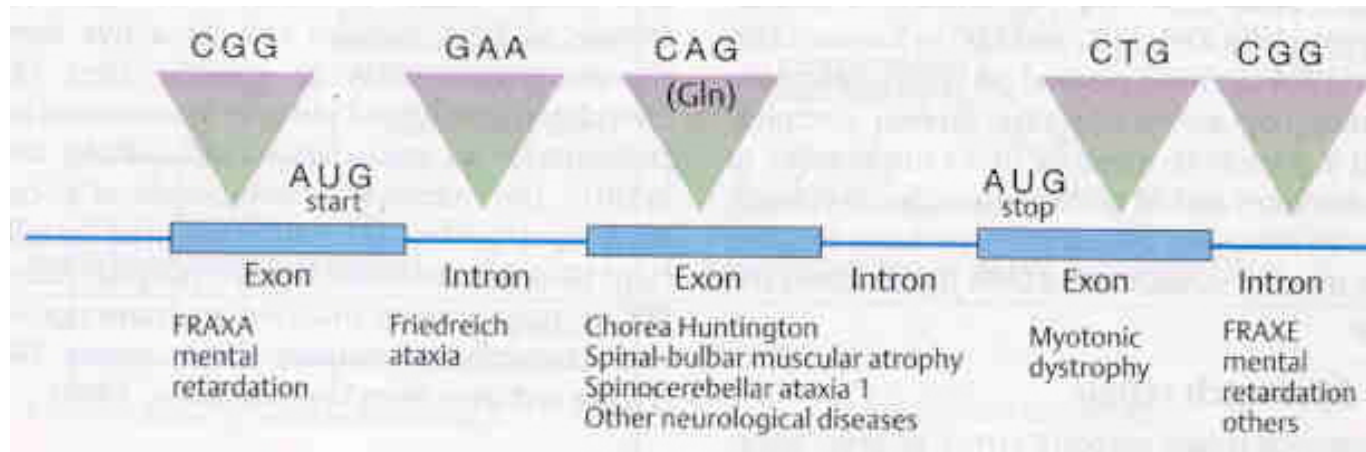


Type I

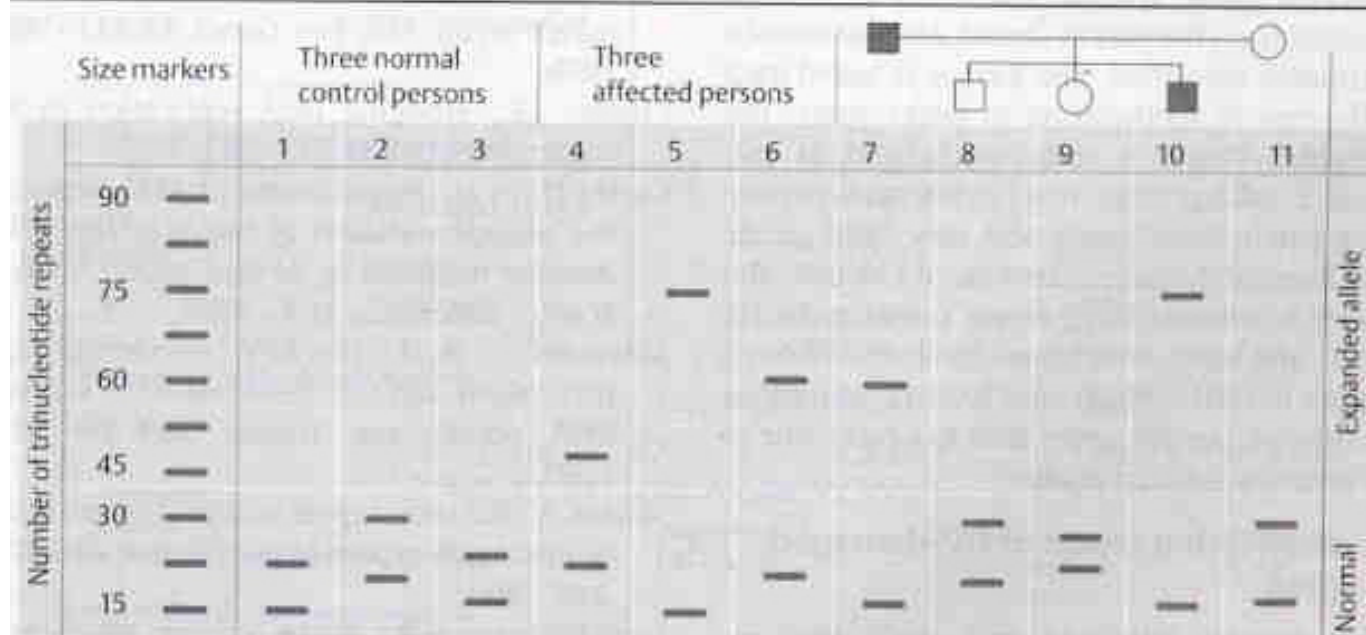
Type II

Instabil mutáció!

Trinukleotid repeat expanzió II.



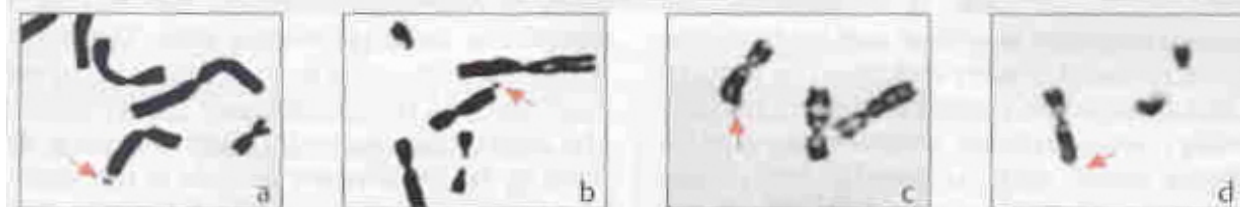
B. Unstable trinucleotide repeats in different diseases



C. Principle of laboratory diagnosis of unstable trinucleotide repeats leading to expansion

Repeat expanziók okozta defektusok

Disease (Examples)	Gene	Frequency	Tri-nucleotide	Normal Number	Mutant Allele	Chromosome
Huntington disease	<i>HD</i>	1:10 000	(CAG) _n	0–26	36–121	4p16.3
Fragile X syndrome	<i>FMR1</i>	1:5 000	(CGG) _n	6–50	52–500	Xq27.3
Myotonic dystrophy	<i>DMPK</i>	1:8 000	(CTG) _n	5–37	50–500	19q13.2
Spinal-bulbar muscular atrophy (Kennedy)	<i>SBMA</i>	<1:50 000	(CAG) _n	11–31	36–65	Xq11-12



Fragilis X
 Huntington disease
 Myotonic dystrophy
 Friedrich ataxia
 stb.

Mikroszatellita pontmutációk

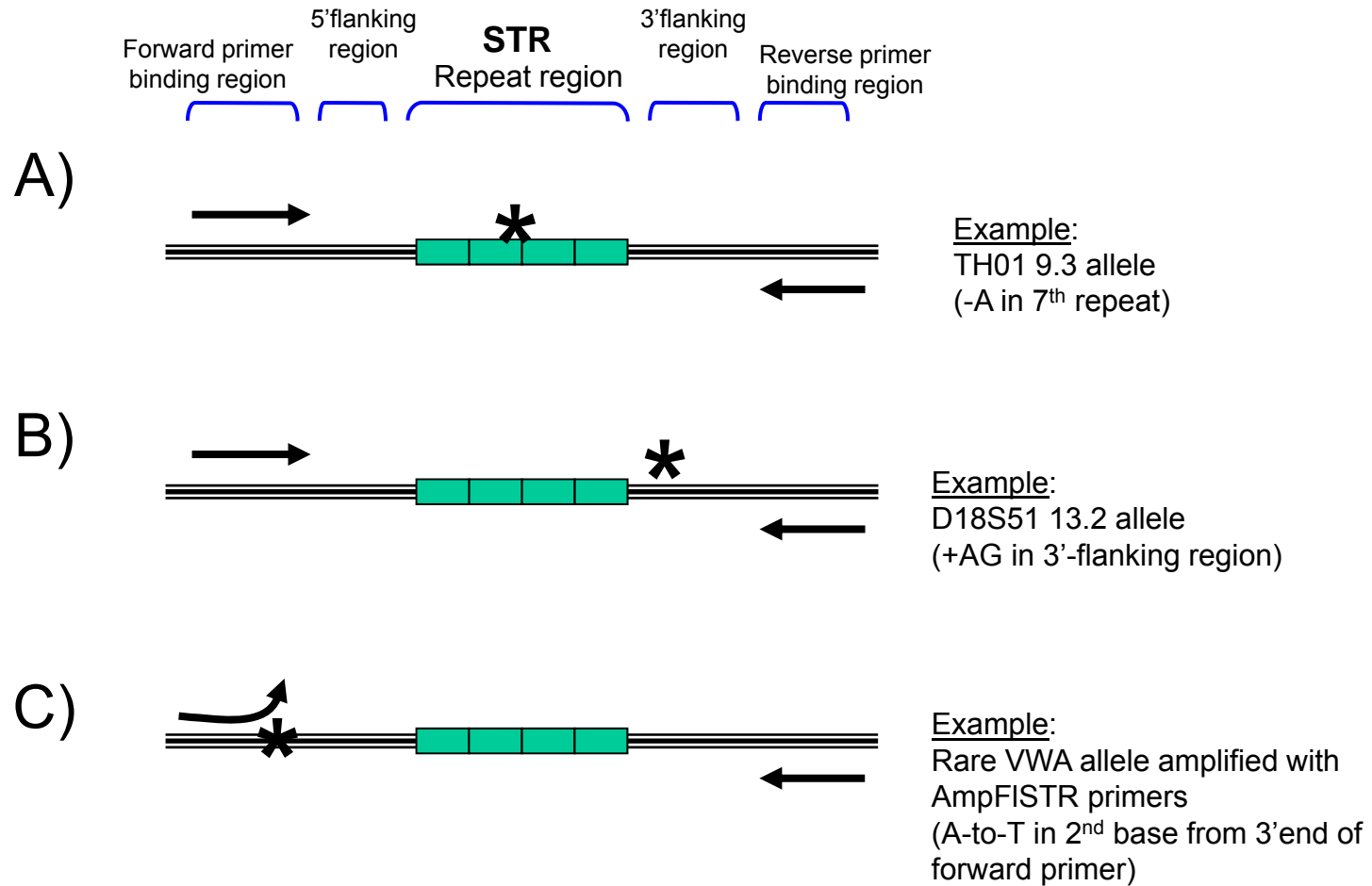
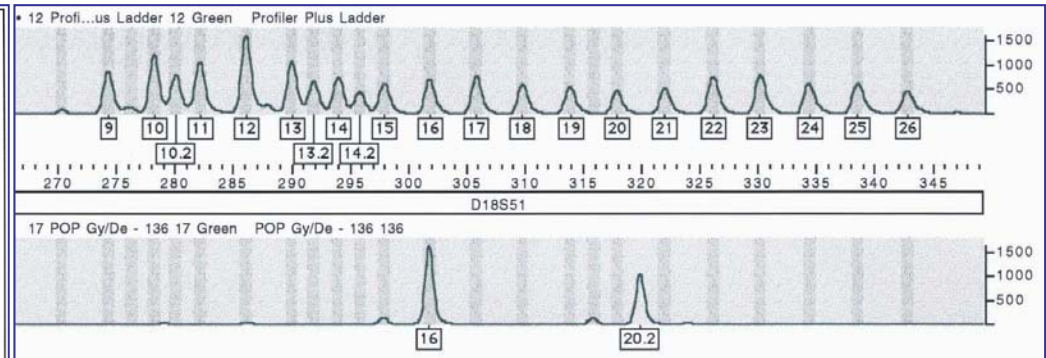
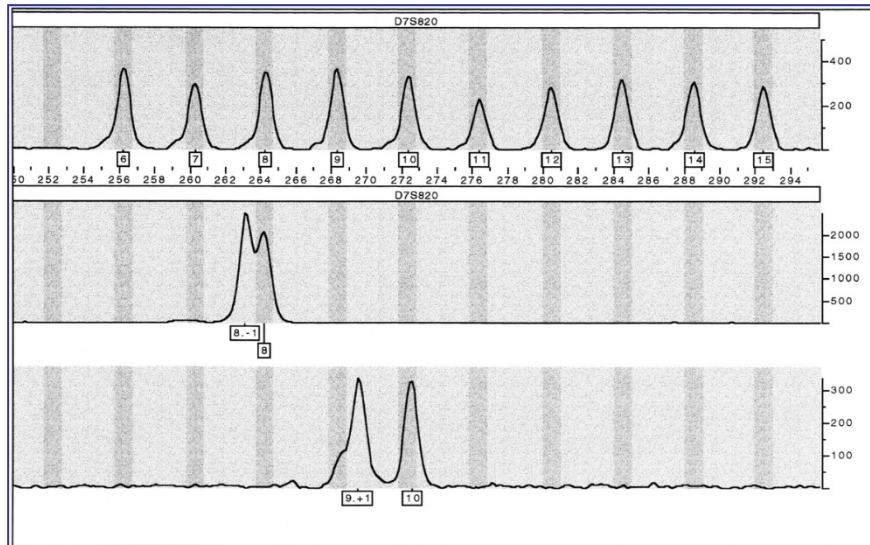


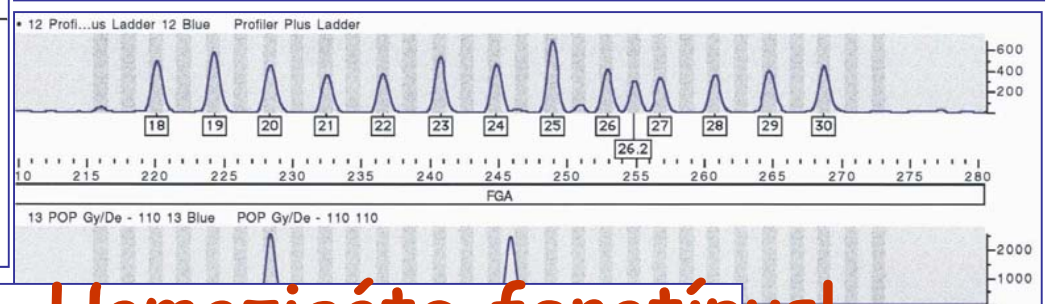
Figure 6.8, J.M. Butler (2005) *Forensic DNA Typing*, 2nd Edition © 2005 Elsevier Academic Press

Variáns mikroszatellita allélek: Null-allélek



Allél	Fragmenthossz	5' flanking régió	Repeat régió	3' flanking régió
16	299 bp	██████████	- (AGAA) ₁₆ -	ΔΔAG AGAGAG - ██████████
20.2	317 bp	██████████	- (AGAA) ₂₁ -	AG AGAGAG - ██████████
15*	295 bp	██████████	- (ATAG) ₁₅ -	AAAG AGAGAG - ██████████

Allél	Fragmens méret	5' Flanking régió	Repeat régió	3' Flanking régió
8.-1	205 bp	██████████	-(GATA) ₈ -	██████████ - (T) ₈ ATCT -
9.+1	211 bp	██████████	-(GATA) ₉ -	██████████ - (T) ₁₀ ATCT -
10	214 bp	██████████	-(GATA) ₁₀ -	██████████ - (T) ₈ AATCT -
12 (ref.)	222 bp	██████████	-(GATA) ₁₂ -	██████████ - (T) ₉ ATCT -
		24 bp		13 bp 124 bp



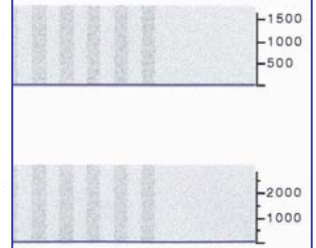
Homozigóta fenotípus!

D13S317 allele 12

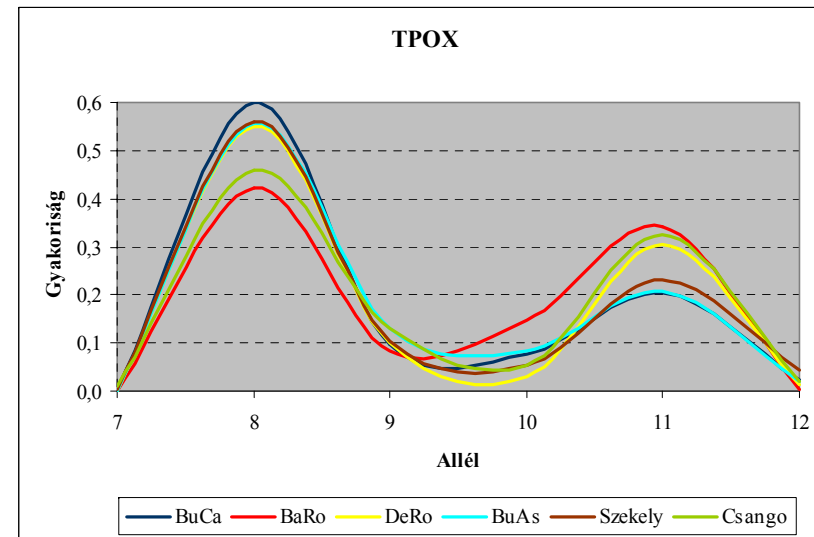
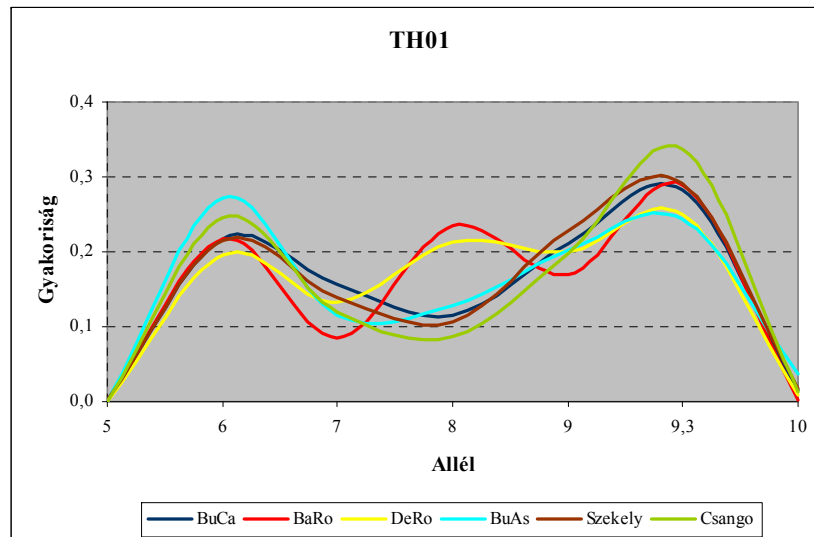
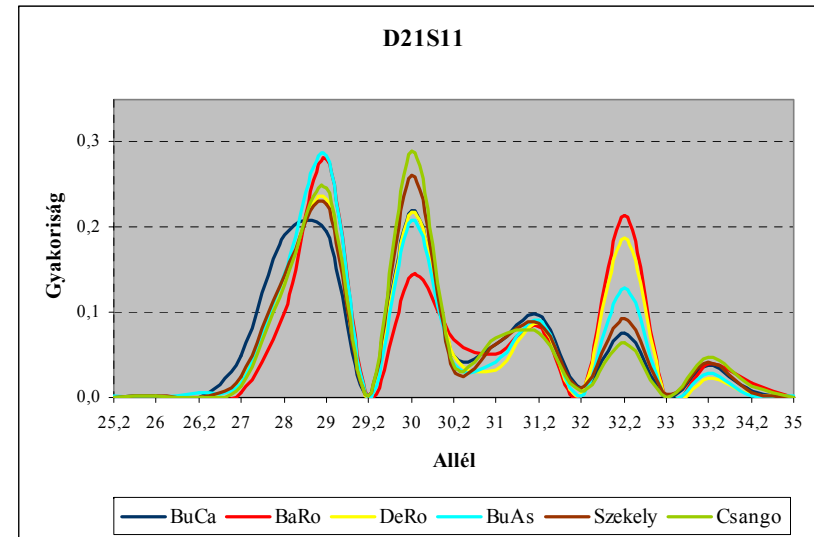
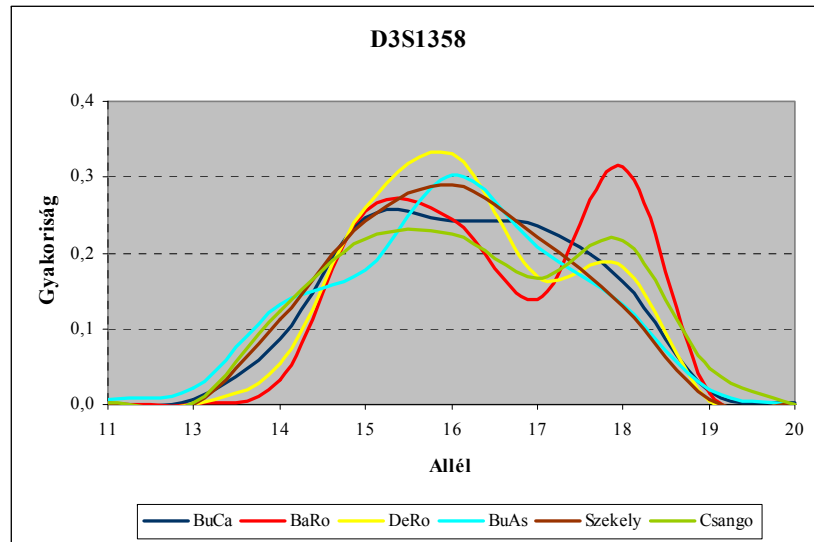
5'-gggttgctggacatggtatcACAGAAGTCTGGGATGTGGA---N82---(ATCT)₁₂ATCAATC(ATCT)₃TTCTGTCTTTTTGGGC---N36---
gaccaacaattcaagctctc-3'

D13S317 allele 7 (variant)

5'-gggttgctggacatggtatcACAGAAGTCTGGGATGTGGA---N82---(ATCT)₇ATCAATCAATC(ATCT)₃TTCTGTCTTTTTGGGC---N36---
gaccaacaattcaagctctc-3'



Mikroszatellita allélgyakoriság diagramok



Genetic Structure (AMOVA)

a,

F_{ST}	P	BuCa															BaRo															DeRo															BuAs														
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
BuCa		[Pattern]															[Pattern]															[Pattern]															[Pattern]														
BaRo		[Pattern]															[Pattern]															[Pattern]															[Pattern]														
DeRo		[Pattern]															[Pattern]															[Pattern]															[Pattern]														
BuAs		[Pattern]															[Pattern]															[Pattern]															[Pattern]														

b,

Φ_{ST}	P	BuCa															BaRo															DeRo															BuAs														
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
BuCa		[Pattern]															[Pattern]															[Pattern]															[Pattern]														
BaRo		[Pattern]															[Pattern]															[Pattern]															[Pattern]														
DeRo		[Pattern]															[Pattern]															[Pattern]															[Pattern]														
BuAs		[Pattern]															[Pattern]															[Pattern]															[Pattern]														

Jelölések: ■ $(F_{ST}, \Phi_{ST}) > 0,02; P < 0,1$ ◐ $0,02 > (F_{ST}, \Phi_{ST}) > 0,01; 0,01 < P < 0,05$ □ $(F_{ST}, \Phi_{ST}) < 0,01; P > 0,05$

F_{ST}	P	BuCa																	Szekely																	Csango																
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
BuCa		[Pattern]																	[Pattern]																	[Pattern]																
Szekely		[Pattern]																	[Pattern]																	[Pattern]																
Csango		[Pattern]																	[Pattern]																	[Pattern]																

b,

Φ_{ST}	P	BuCa																	Szekely																	Csango																
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
BuCa		[Pattern]																	[Pattern]																	[Pattern]																
Szekely		[Pattern]																	[Pattern]																	[Pattern]																
Csango		[Pattern]																	[Pattern]																	[Pattern]																

Jelölések: ■ $(F_{ST}, \Phi_{ST}) > 0,02; P < 0,1$ ◐ $0,02 > (F_{ST}, \Phi_{ST}) > 0,01; 0,01 < P < 0,05$ □ $(F_{ST}, \Phi_{ST}) < 0,01; P > 0,05$

Fajtaazonosítás?.....

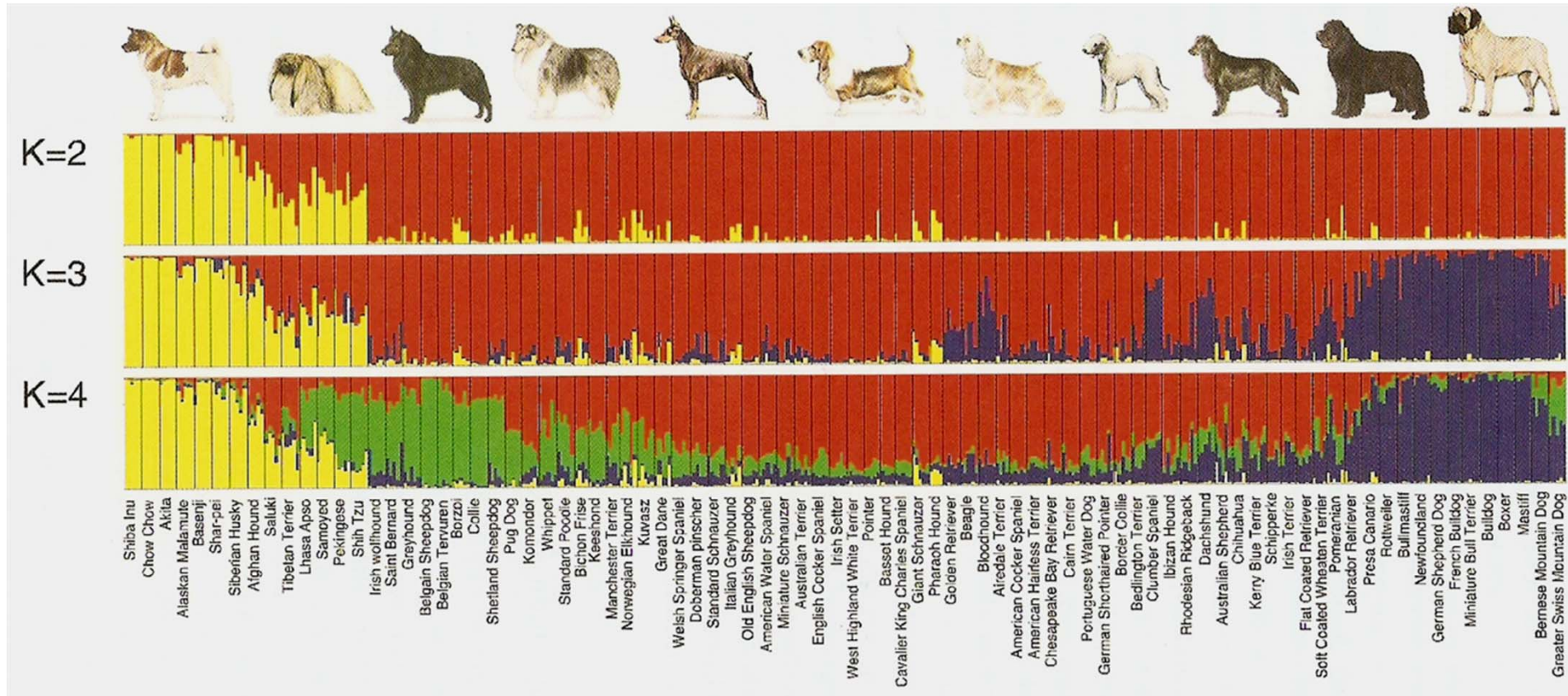
Populációs vizsgálatok!!!

1. Ősi

2. Terelő

3. Vadász

4. Védő



- 85 fajta (n=414)
- 95 mikroszatellita marker (dinukleotid repeat)
- 99%-ban helyes besorolás

Mobilis Genetikai Elemek

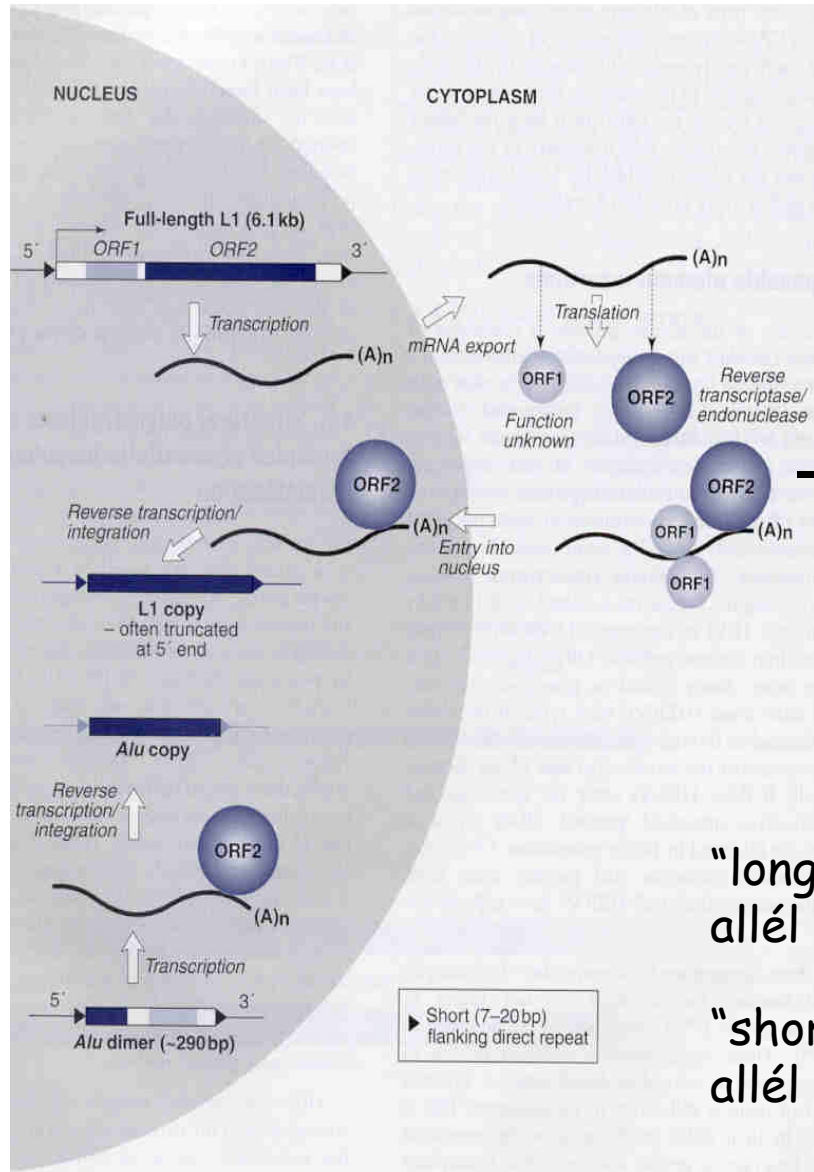
- Extra mennyiség a humán genomban: 100/gén (~ 45 %)
- Humán genom ~ 75 %-a lehet mobilis elem maradványa
- Mutagén hatások: pl. inzerció, nem homológ rekombináció,
→ negatív konzekvencia a gazdára nézve
- Retrotranszpozon: „copy-and-paste”, LINEs, SINEs, LTRs
- Transzpozonok: „cut-and-paste”

TABLE 2.2: CLASSES OF DISPERSED REPEATS IN THE HUMAN GENOME.

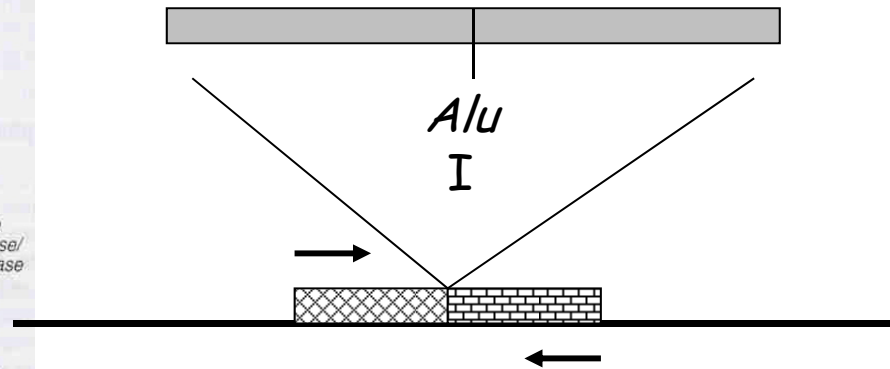
Class	Copy no. per haploid genome	Fraction of genome	Autonomous transposition or retrotransposition?	Length
LINEs	850 000	21%	Yes	Up to 6–8 kb
SINEs	1 500 000	13%	No	Up to 100–300 bp
Retrovirus-like elements	450 000	8%	Complete copies, yes	6–11 kb (1.5–3 kb)
DNA transposon copies	300 000	3%	Complete copies, yes	2–3 kb (80–3000 bp)

Values given in parentheses are lengths of incomplete elements, incapable of autonomous transposition (see Section 3.4). Adapted from Lander *et al.* (2001).

Mobilis elemek: biallélikus hossz-polimorfizmus



Human *Alu* Repeat (~300 bp)



Kétféle alléltípus

"long" (+)
allél



400
bp

"short" (-)
allél



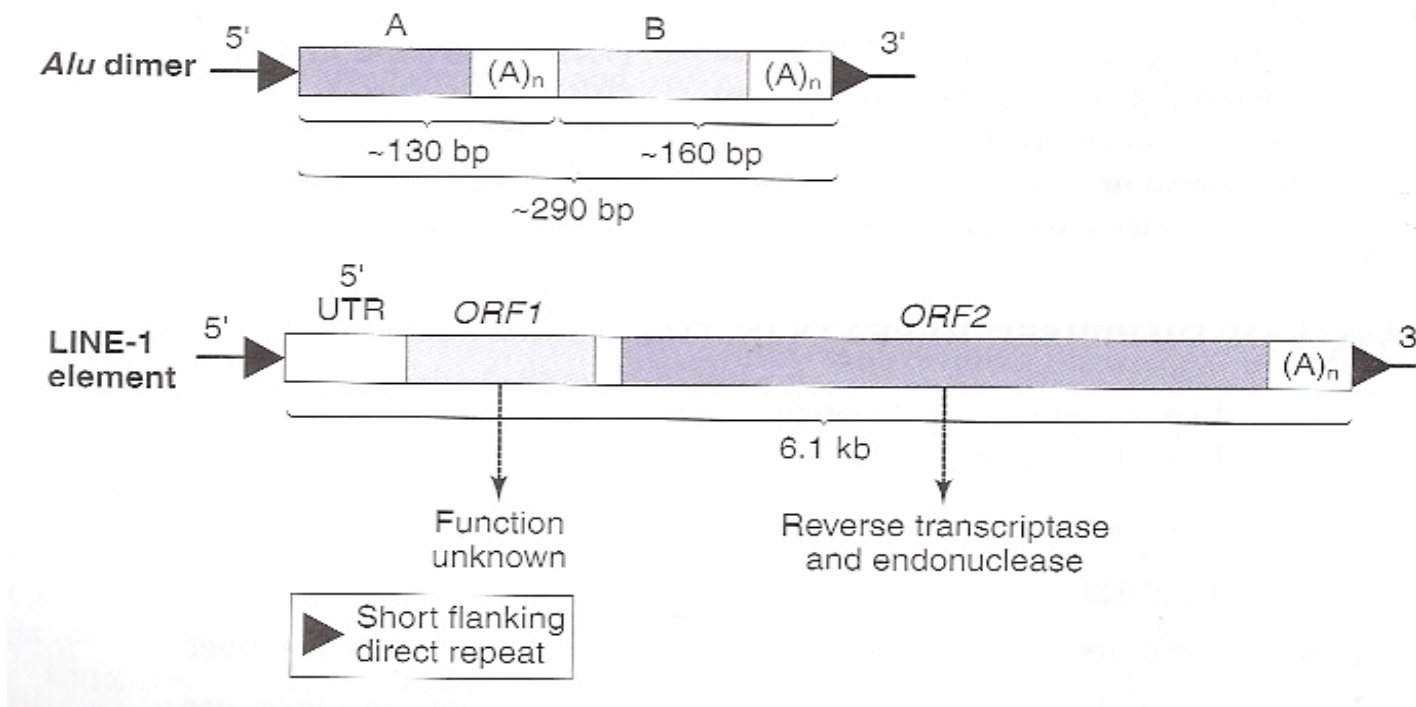
100
bp

Long Interspersed Nuclear Elements - LINEs

- LINEs v. *Kpn*: emberi genom 20 %-a, 870.000 kópia
- kb. 100 LINEs működőképes retrotranszpozonként
- ~ 6 kb, belső 5' promóter, 2 ORF (RNS-kötő fehérje, endonukleáz + reverz transzkriptáz), poly(A)-farok,
- Target-primed reverz transzkripció: TT | AAAA - target
- Hanyag másolás
(transzkripció „read-through”, „dead-on-arrival”, nagyobb szekvenciárszek átrendeződése, egyéb nem autonóm szekvenciákhoz való kötődés)
- Önmaguktól nem tudnak a genomból kivágódni (deléció)
- ősi és relatíve új szekvenciák (pl. LINE-1: 5 MYs)

Short Interspersed Nuclear Elements - SINEs

- SINEs v. *Alu*: 1.500.000 kópia, 70 % *AluI*, 300 bp,
- Főemlős specifikus, *Alu I*: AGCT, polimorfizmusok,
- Nem kódoló szekvencia, önállóan nem mobilizálódik
- *Alu* - LINE-1 retrotranszpozíció, 0.05 /genom / generáció



Egynukleotid polimorfizmusok (SNPs)

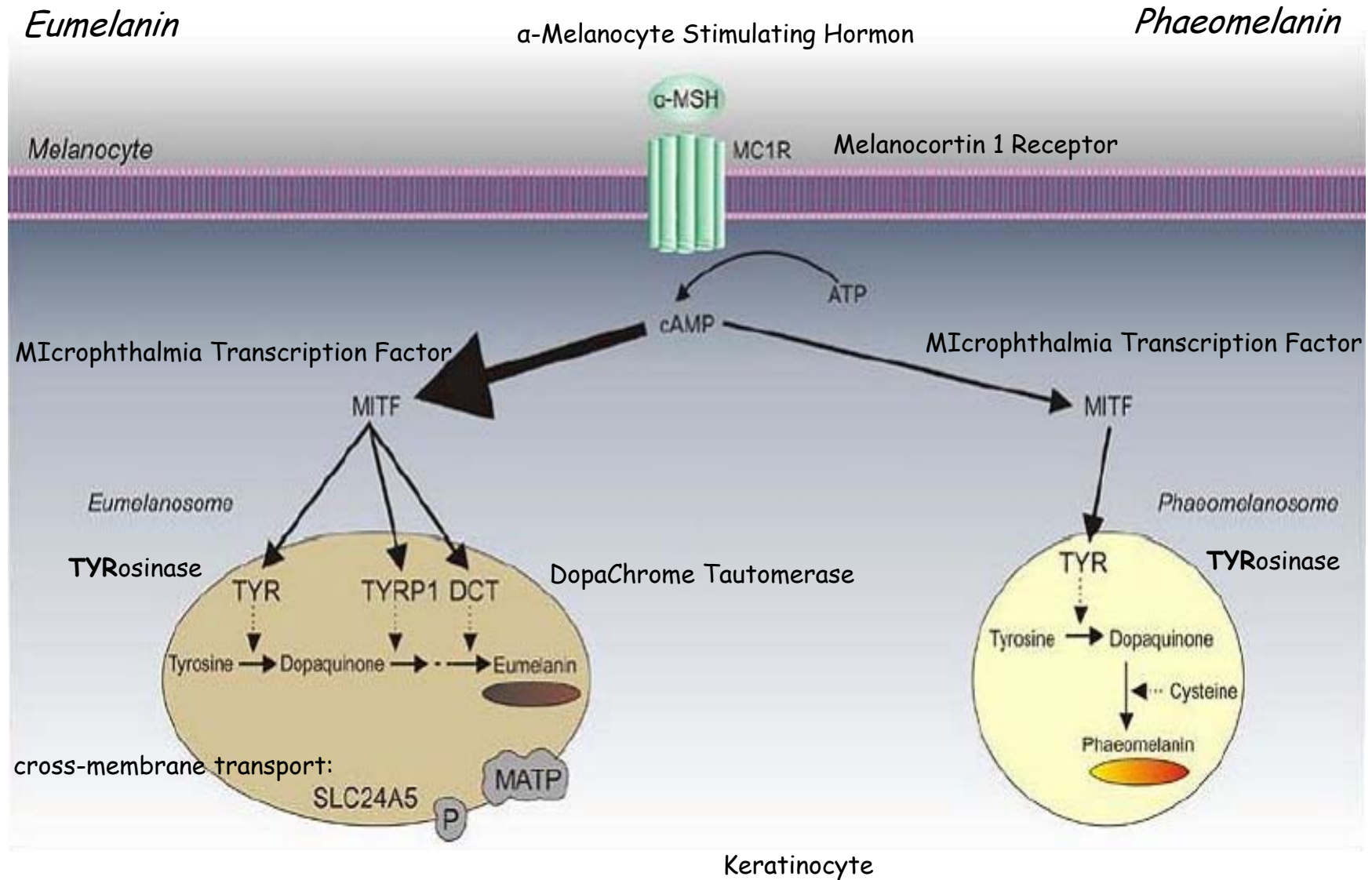
- Biallélikus markerek (6 lehetőség)
 - (A / G, C / T, A / T, C / G, T / G, A / C)
- Több millió SNP a genomban
 - kb. 500-1000 bázispáronként
 - pontmutációk génekben és/vagy regulátor régiókban
- Fenotípus kapcsolatok
 - pigmentáció, testalkat, ...
- Leszármazási vonalak (Y-SNP's)
- Diagnosztika
 - multifaktoriális poligénes betegségek

SNP markerek a humán genomban

TABLE 12.2 Categories of SNP Markers (See Budowle & van Daal 2008, Butler et al. 2008).

Category	Characteristics	Examples
Identity SNPs Individual Identification SNPs (IISNPs)	SNPs that collectively give very low probabilities of two individuals having the same multi-locus genotype	FSS 21plex (Dixon et al. 2005) SNPforID 52plex (Sanchez et al. 2006) Kidd group SNPs (Pakstis et al. 2010)
Lineage SNPs Lineage Informative SNPs (LISNPs)	Sets of tightly linked SNPs that function as multi-allelic markers that can serve to identify relatives with higher probabilities than simple bi-allelic SNPs	mtDNA coding region SNPs (Coble et al. 2004) Japanese Y-SNPs (Mizuno et al. 2010) Haplotype blocks (Ge et al. 2010)
Ancestry SNPs Ancestry Informative SNPs (AISNPs)	SNPs that collectively give a high probability of an individual's ancestry being from one part of the world or being derived from two or more areas of the world	SNPforID 34plex (Phillips et al. 2007b) 24 SNPs (Lao et al. 2010) FSS YSNPs (Wetton et al. 2005)
Phenotype SNPs Phenotype Informative SNPs (PISNPs)	SNPs that provide a high probability that the individual has particular phenotypes, such as a particular skin color, hair color, eye color, etc.	Red hair (Grimes et al. 2001) "Golden" gene pigmentation (Lamason et al. 2005) IrisPlex eye color (Walsh et al. 2010)

A human melanogenesis



A bőrpigmentáció kialakulásáért felelős gének

Principal skin pigmentation candidate genes

Locus	Chromosome	Protein	Mut phenotype	Function
Melanosome proteins				
TYR	11q14-11q21	Tyrosinase	OCA1	Oxidation of tyrosine
TYRP1	9p23	Gp75, TRYP1	OCA3	DHICA-oxidase, TYR stabilisation
DCT	13q32	DCT, TRYP2		Dopachrome tautomerase
OCA2	15q11.2-15q12	P-protein	OCA2 (eye)	pH of melanosome
SLC45A2	5p14.3-5q12.3	MATP, AIM-1	OCA4 (skin)	Melansome maturation
SLC24A5	15q21.1	Cation exchanger		Melanosome precursor
Signal proteins				
ASIP	20q11.2-20q12	Agouti signal protein		MC1R antagonist
MC1R	16q24.3	MSH receptor	Red hair (skin)	G-protein coupled receptor
POMC	16q24.3	MSH receptor	Red hair	MC1R antagonist
OA1	Xp22.3	OA1 protein	OA1	G-protein coupled receptor
MITF	3p12.3-3p14.1	MITF	Waardenburg	Transcription factor
Proteins involved in melanosome transport or uptake by keratinocytes				
MYO5A	15q21	Myosin Va	Griscelli	Motor protein
RAB27A	15q15-15q21.1	Rab27a	Griscelli	RAS family protein
HPS1	10q23.1-10q23.3	HPS1	Hermansky-Pudlak	Organelle biogenesis and size
HPS6	10q24.32	HPS6	Hermansky-Pudlak	Organelle biogenesis

ACTH: adrenocorticotrophin hormone; DCT: dopachrome tautomerase; DHICA: 5,6-dihydroxyindole-2-carboxylic acid; MATP: membrane-associated transporter protein; MC1R: melanocortin-1 receptor; MITF: microphthalmia-associated transcription factor; MSH: melanocyte stimulating hormone; OCA: oculocutaneous albinism; POMC: pro-opiomelanocortin; TYRP1: tyrosinase-related protein 1.

MC1R gén mutációk

Mutations in the MC1R gene, their penetrance and functional significance (where known)

Mutation	Type	Designation	Penetrance (odds ratio)	Functional significance	References (for functional significance and penetrance)
R151C	Mis-sense	R	63.3	Altered cellular location	[16,26]
R160W	Mis-sense	R	63.3	Altered cellular location	[16,26]
D294H	Mis-sense	R	63.3	Impaired G coupling ability	[26,27]
D84E	Mis-sense	R	63.3	Altered cellular location	[16,26]
I155T	Mis-sense	Lack of statistical data—strong familial association		Altered cellular location	[16,26]
V92M	Mis-sense	r	5.1	Reduced α -MSH binding	[26,28,29]
V60L	Mis-sense	r	5.1		[26]
R163Q	Mis-sense	r	5.1	Slightly reduced α -MSH binding	[26,29]
R142H	Mis-sense	Lack of statistical data—strong familial association			[26]

- MC1R allélváltozatok különböző aktivitással rendelkeznek.
- 317 AS, 7 transzmembrán domén,
- SNPs, RHC fenotípus,
- neandervölgyi pigmentáció
- genetikai tesztek, fenotípus predikció

SNPs - pigmentációs gének

- ASIP (aguti): 3'UTR 8818A - MSH antagonista - phaeomelanin termelés
- MATP: melanoszóma pH reguláció, 374Leu allél - sötét szín, albinizmus
- SLC24A5: „arany” gén, zebrafish, Ala111Thr allél, világos árnyalat, europid rasszban fixált, szelekciós nyomás?
- OCA2: albinizmus gén, 305 Arg/Trp, Afrika / Európa

Gene	Location	Protein	Reference SNP ID (rs#) ^a	Alleles	Variation type
<i>MC1R</i>	16q24.3	MC1R: melanocortin 1 receptor	rs1805007 rs1805008	C/T C/T	ns coding, c.451C>T, p.R151C ns coding, c.478C>T, p.R160W
<i>HERC2</i>	15q13	Unknown	rs12913832	A/G	Non-coding, intron 86
<i>OCA2</i>	15q11.2-15q12	P-protein: NA ⁺ /H ⁺ antiporter or glutamate transporter	rs7495174 rs6497268 or rs4778241 rs11855019 or rs4778138 rs1545397	T/C G/T T/C G/A	Non-coding, intron 1 Non-coding intronic
<i>SLC45A2</i>	5p13.3	MATP: membrane-associated transporter protein	rs16891982	C/G	ns coding, c.1122C>G, p.F374L
<i>SLC24A5</i>	15q21.1	SLC24A5 (or NCKX5): solute carrier family 24, member 5; potassium-dependent sodium- calcium ion exchanger	rs1426654	G/A	ns coding, p.A111T
<i>DCT</i>	13q32	DCT or TYRP2/TRP-2: dopachrome tautomerase or tyrosinase-related protein-2	rs2031526	G/A	Non-coding, intronic

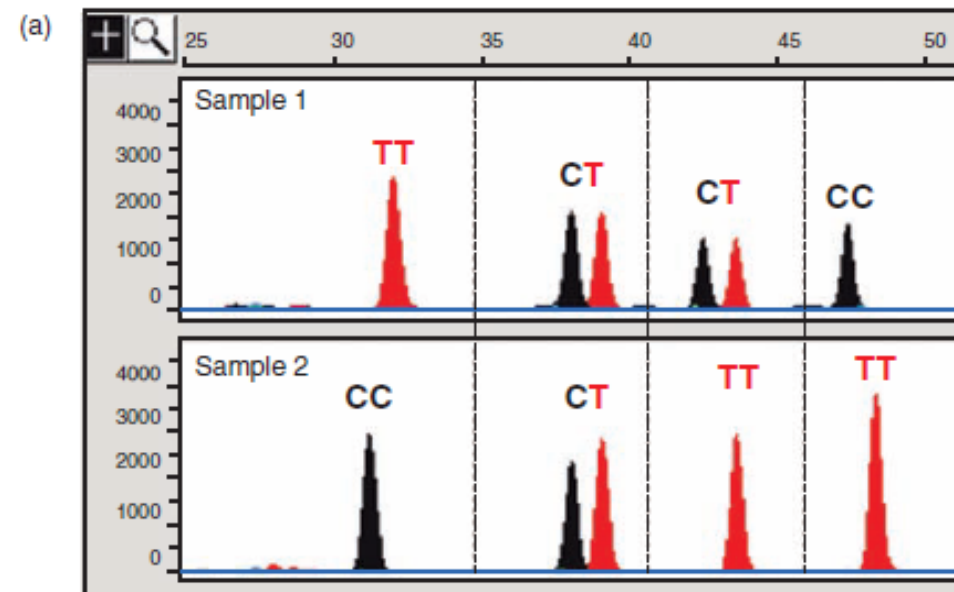
ns non-synonymous

^a Reference SNP ID refer to the reference sequence identifier given to the SNP in the dbSNP database

SNaPshot: A Primer Extension Assay Capable of Multiplex Analysis

Minisequencing
(SNaPshot assay)

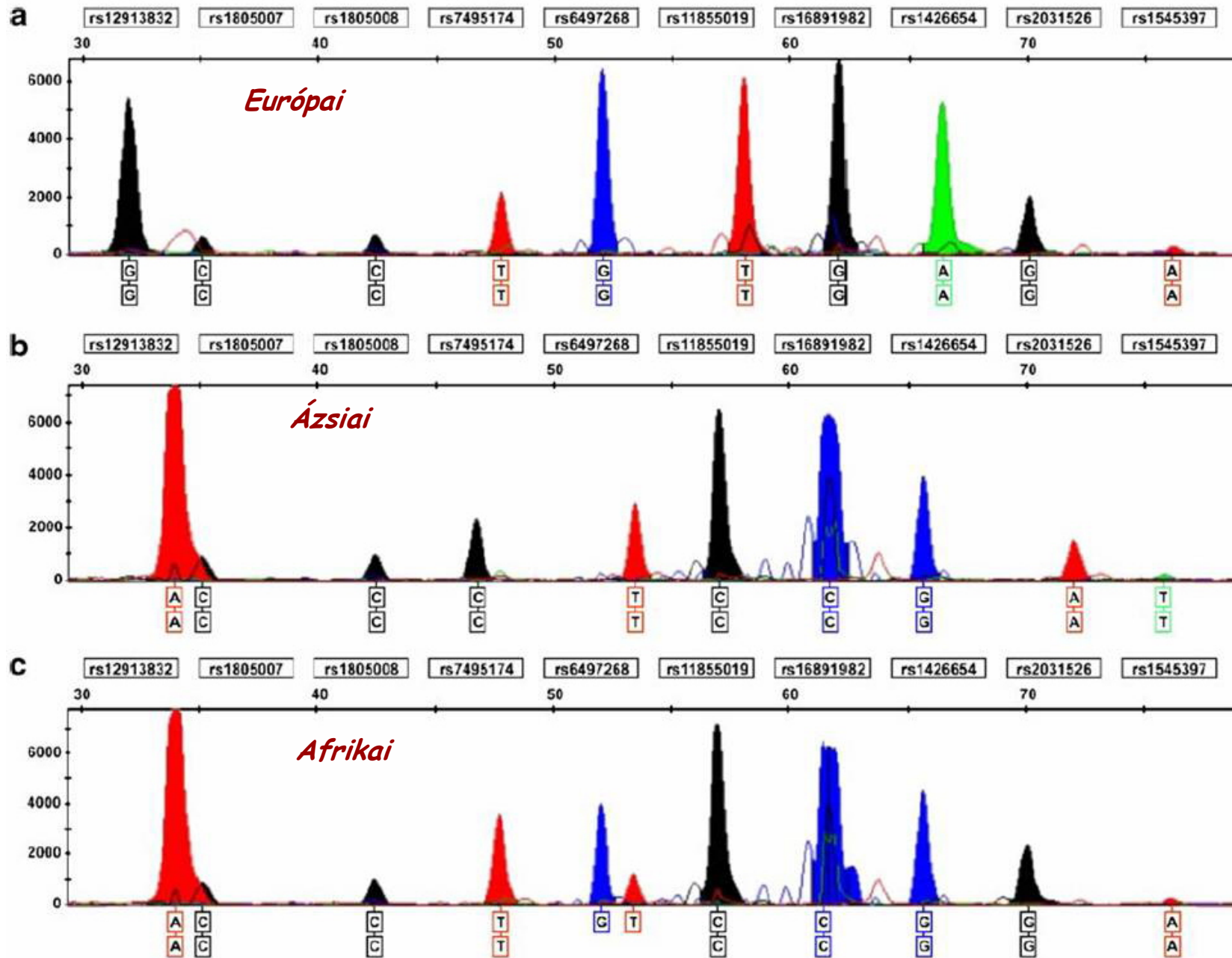
Allele-specific primer extension across the SNP site with fluorescently labeled ddNTPs; mobility modifying tails can be added to the 5'-end of each primer in order to spatially separate them during electrophoresis.



- (b) (TTTTT)-**primer1** (chromosome 20)-**ddT/ddT**
 (TTTTT)-(TTTTT)-**primer2** (chromosome 6)-**ddC/ddT**
 (TTTTT)-(TTTTT)-(TTTTT)-**primer3** (chromosome 14)-**ddC/ddT**
 (TTTTT)-(TTTTT)-(TTTTT)-(TTTTT)-**primer4** (chromosome 1)-**ddC/ddC**

FIGURE 12.2 Allele-specific primer extension results using four autosomal SNP markers on two different samples (a). SNP loci are from separate chromosomes (1, 6, 14, and 20) and therefore unlinked. Electrophoretic resolution of the SNP primer extension products occurs due to poly(T) tails that are 5 nucleotides different from one another (b).

10 pigmentációs gén SNP genotipizálás (SNaPshot)



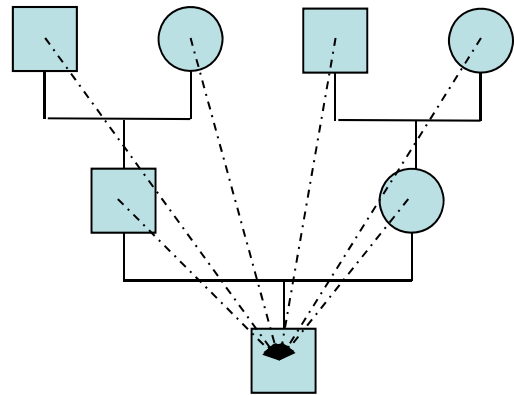
Sample	Self-reported pigmentary traits			rs12913832 HERC2	rs1805007 MC1R	rs1805008 MC1R	OCA2 diplotype ^a	rs16891982 SLC24A2	rs1426654 SLC24A5	rs2031526 DCT	rs1545397 OCA2	Inferred ancestry of individuals ^b		
	Eye color	Hair color	Skin color									European	Asian	African
E1	Blue	Red	Fair	<u>G/G</u>	C/C	C/T	<u>TGT/TGT</u>	G/G	A/A	G/G	A/A	0.963	0.012	0.024
E2	Green	Light brown	Fair	<u>G/G</u>	C/C	C/C	<u>TGT/TGT</u>	G/G	A/A	A/G	A/A	0.954	0.021	0.025
E3	Blue	Blond	Fair	<u>G/G</u>	C/C	C/C	<u>TGT/TGT</u>	G/G	A/A	A/G	A/A	0.954	0.024	0.022
E4	Blue	Blond	Fair	<u>G/G</u>	C/C	C/C	<u>TGT/TGT</u>	G/G	A/A	A/G	A/A	0.960	0.020	0.020
E5	Blue/gray	Auburn	Fair	<u>G/G</u>	C/T	C/C	<u>TGT/TGT</u>	G/G	A/A	G/G	A/A	0.961	0.013	0.026
E6	Green/gray	Light brown	Fair	<u>G/G</u>	C/C	C/C	<u>TGT/TGT</u>	C/G	A/A	G/G	A/A	0.787	0.038	0.175
E7	Green/hazel	Light brown	Fair	A/G	C/C	C/C	<u>TGT/TGT</u>	G/G	A/A	A/G	A/A	0.955	0.022	0.024
E8	Green/hazel	Dark brown	Fair	A/A	C/C	C/C	TGT/CTC	G/G	A/A	G/G	A/A	0.961	0.013	0.027
E9	Green/hazel	Dark brown	Fair	A/A	C/C	C/C	TTT/CTC	G/G	A/A	G/G	A/A	0.963	0.013	0.024
E10	Blue	Light brown	Fair	<u>G/G</u>	C/C	C/C	<u>TGT/TGT</u>	C/G	A/A	G/G	A/A	0.789	0.049	0.163
E11	Green	Auburn	Fair	<u>G/G</u>	C/T	C/C	TGT/TGC	G/G	A/A	G/G	A/A	0.958	0.014	0.028
E12	Blue/hazel	Light brown	Fair	A/G	C/C	C/C	TGT/TTT	G/G	A/A	G/G	A/A	0.962	0.012	0.026
E13	Blue/hazel	Light brown	Fair	A/G	C/C	C/C	TGT/TTT	G/G	A/A	G/G	A/A	0.965	0.013	0.022
E14	Green	Light brown	Fair	<u>G/G</u>	C/C	C/C	<u>TGT/TGT</u>	C/G	A/A	G/G	A/T	0.763	0.165	0.073
E15	Brown	Dark brown	Fair	A/G	C/C	C/C	<u>TGT/TGT</u>	G/G	A/A	A/G	A/A	0.957	0.022	0.021
E16	Brown	Dark brown	Fair	A/A	C/C	C/C	TGT/CTC	C/G	A/A	A/G	A/T	0.669	0.283	0.048
E17	Green/hazel	Dark brown	Medium	A/G	C/C	C/C	TGT/TTT	C/G	A/A	G/G	A/T	0.755	0.170	0.076
E18	Blue	Light brown	Fair	<u>G/G</u>	C/C	C/C	<u>TGT/TGT</u>	G/G	A/A	G/G	A/T	0.935	0.045	0.021
E19	Brown	Red	Fair	A/G	C/T	C/C	<u>TGT/TGT</u>	G/G	A/A	G/G	A/A	0.964	0.013	0.022
E20	Green	Light brown	Fair	<u>G/G</u>	C/C	C/C	<u>TGT/TGT</u>	C/G	A/A	G/G	A/A	0.792	0.047	0.161
E21	Green/gray	Blond	Fair	<u>G/G</u>	C/C	C/C	<u>TGT/TGT</u>	G/G	A/A	A/G	A/A	0.957	0.022	0.021
E22	Blue	Light brown	Fair	<u>G/G</u>	C/C	C/C	<u>TGT/TGT</u>	G/G	A/A	G/G	A/A	0.959	0.014	0.026
E23	Green/hazel	Light brown	Fair	A/G	C/C	C/C	TGT/TTT	G/G	A/A	A/G	A/A	0.957	0.020	0.022
E24	Green	Light brown	Fair	<u>G/G</u>	C/C	C/C	<u>TGT/TGT</u>	C/G	A/A	G/G	A/A	0.786	0.049	0.166
E25	Brown	Red	Fair	A/G	C/C	T/T	TGT/TGC	G/G	A/A	G/G	A/A	0.963	0.014	0.023
E26	Blue	Light brown	Fair	<u>G/G</u>	C/C	C/C	<u>TGT/TGT</u>	G/G	A/A	A/G	A/A	0.954	0.021	0.025
E27	Blue	Red	Fair	<u>G/G</u>	C/C	C/T	<u>TGT/TGT</u>	G/G	A/A	G/G	A/A	0.958	0.014	0.028
Af1	Brown	Black	Dark	A/A	C/C	C/C	TGC/TTC	C/C	G/G	A/G	A/A	0.028	0.094	0.878
Af2	Brown	Black	Dark	A/A	C/C	C/C	TGC/TTC	C/C	G/G	G/G	A/A	0.023	0.031	0.946
Af3	Brown	Black	Dark	A/A	C/C	C/C	TGC/TTC	C/C	A/G	G/G	A/A	0.164	0.041	0.795
As1	-	-	-	A/A	C/C	C/C	TTT/CTC	C/C	G/G	A/G	A/T	0.042	0.649	0.308
As2	-	-	-	A/A	C/C	C/C	CTC/CTC	C/C	G/G	A/G	T/T	0.020	0.921	0.060
As3	-	-	-	A/A	C/C	C/C	CTC/CTC	C/C	G/G	A/A	T/T	0.013	0.964	0.023
As4	-	-	-	A/G	C/C	C/C	TTT/CGC	C/C	A/G	A/A	A/T	0.212	0.708	0.080
As5	-	-	-	A/A	C/C	C/C	TTC/CGC	C/C	G/G	A/G	T/T	0.019	0.922	0.059
As6	-	-	-	A/A	C/C	C/C	CTC/CTC	C/G	G/G	A/A	T/T	0.119	0.858	0.023

E European modern sample, Af African modern sample, As Asian modern sample

^a OCA2 diplotype correspond to markers rs7495174/rs6497268/rs11855019. OCA2 diplotype and rs12913832 genotype predictive of blue eye color phenotype are underlined

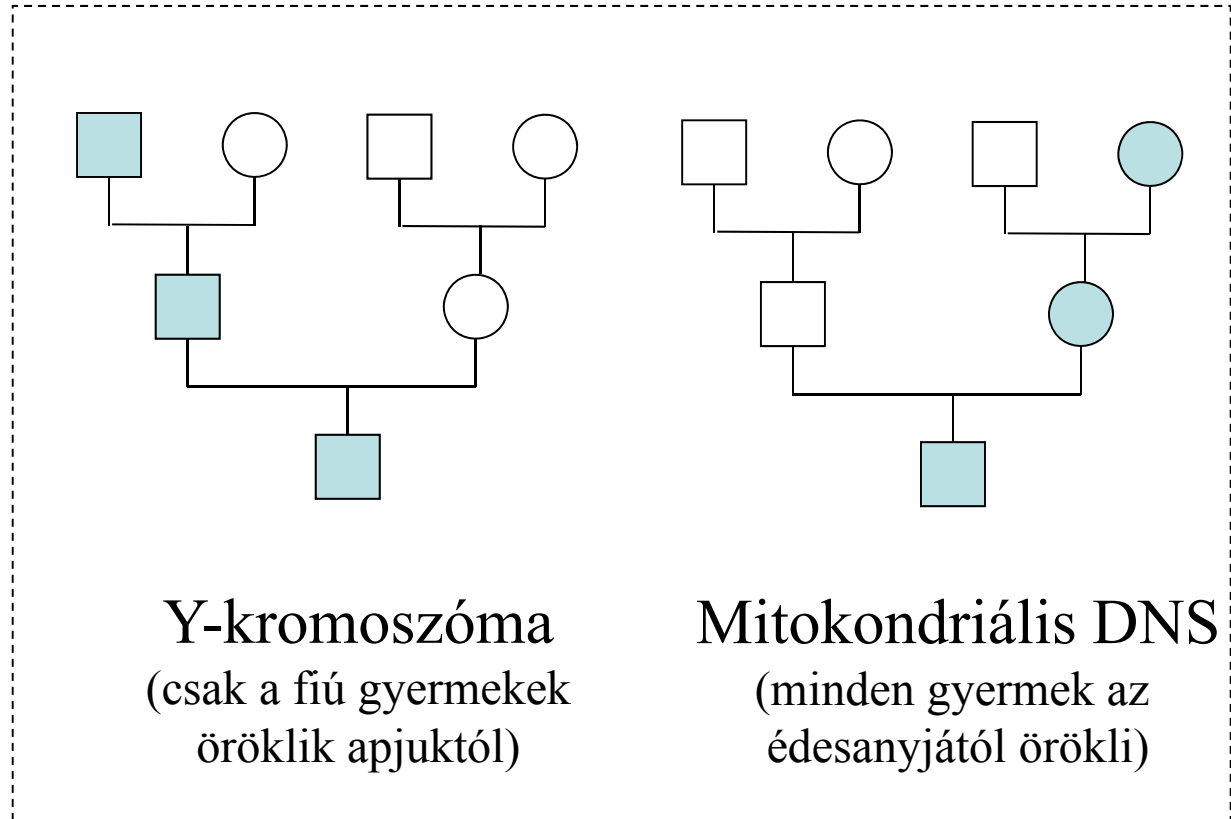
^b Probability of being from European/Asian/African population determined using the STRUCTURE program. The greatest probability, most likely estimate of ancestry, is indicated in bold

Eltérő öröklődési mintázatok



Autoszómás marker
(az összes felmenőntől örököljük részletekben)

Leszármazási markerek



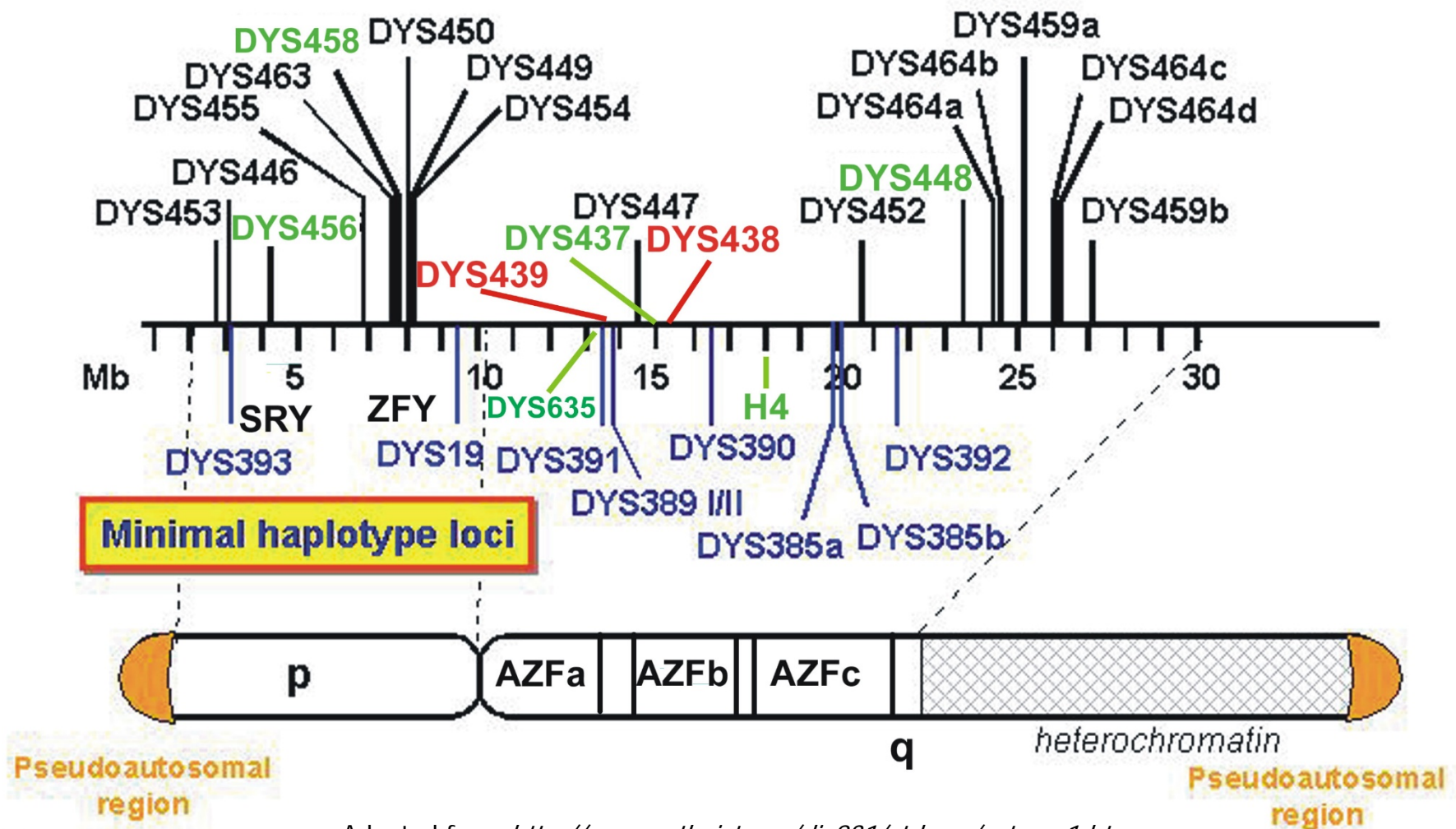
Y-kromoszóma
(csak a fiú gyermekek örökölik apjuktól)

Mitokondriális DNS
(minden gyermek az édesanyjától örököli)

Y STR Positions along Y Chromosome

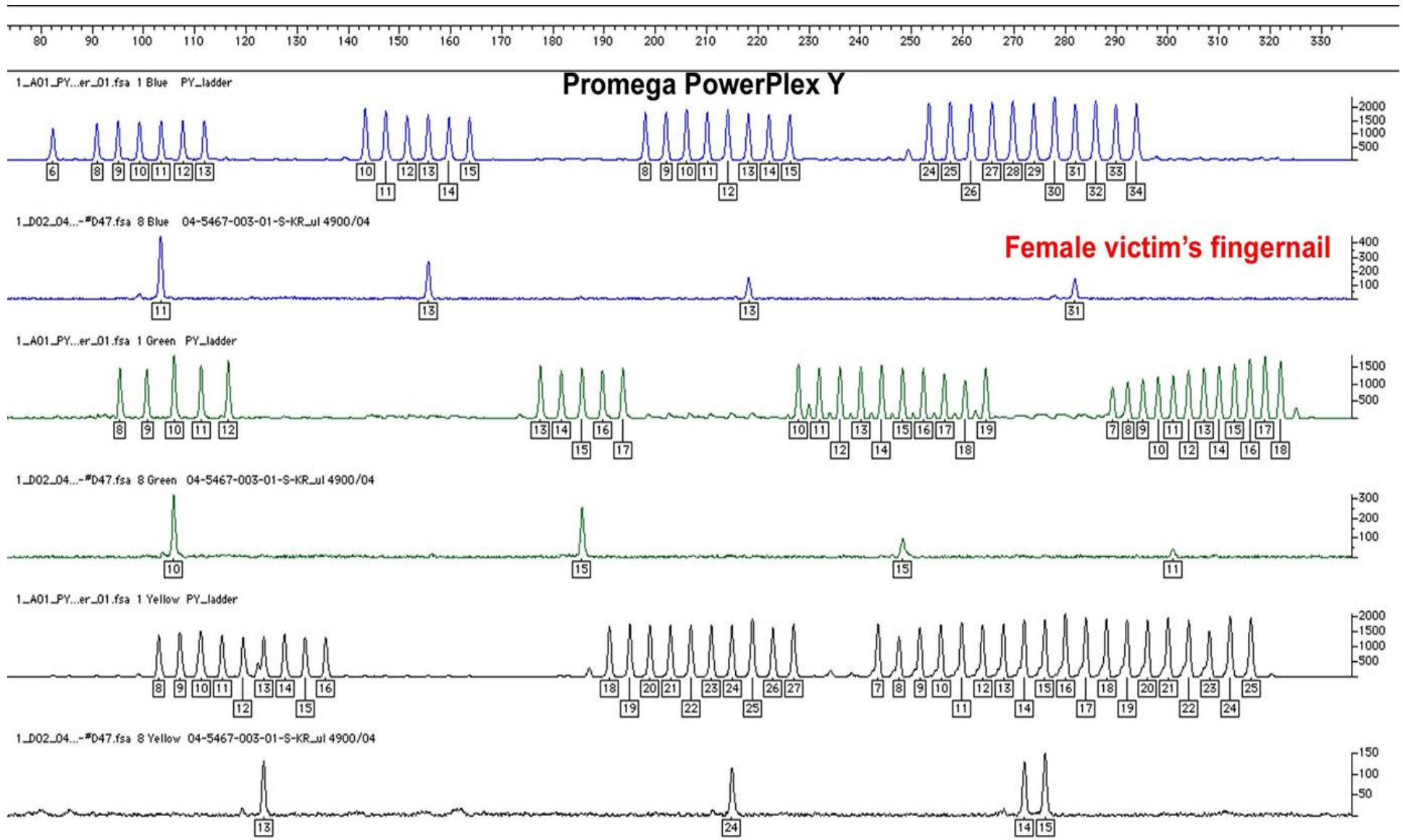
Extended **haplotype loci**

ABI AmpF/STR **Yfiler loci**

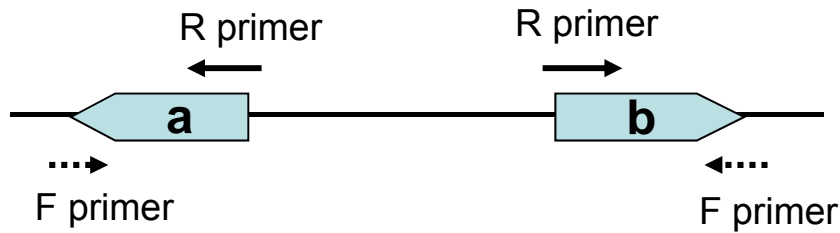


Adapted from <http://www.cstl.nist.gov/div831/strbase/ystrpos1.htm>

Y kromoszómás vizsgálat bűnügyi mintákon

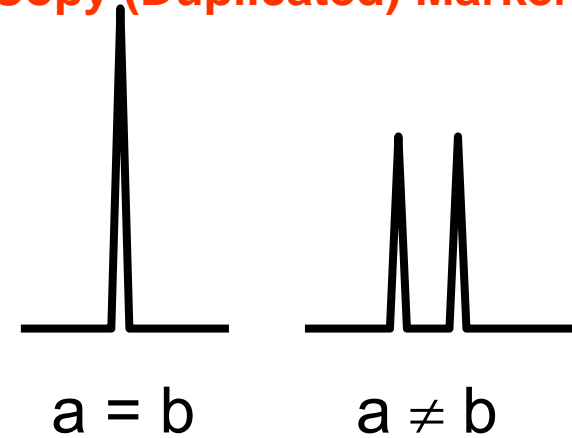


(A) DYS385 a/b

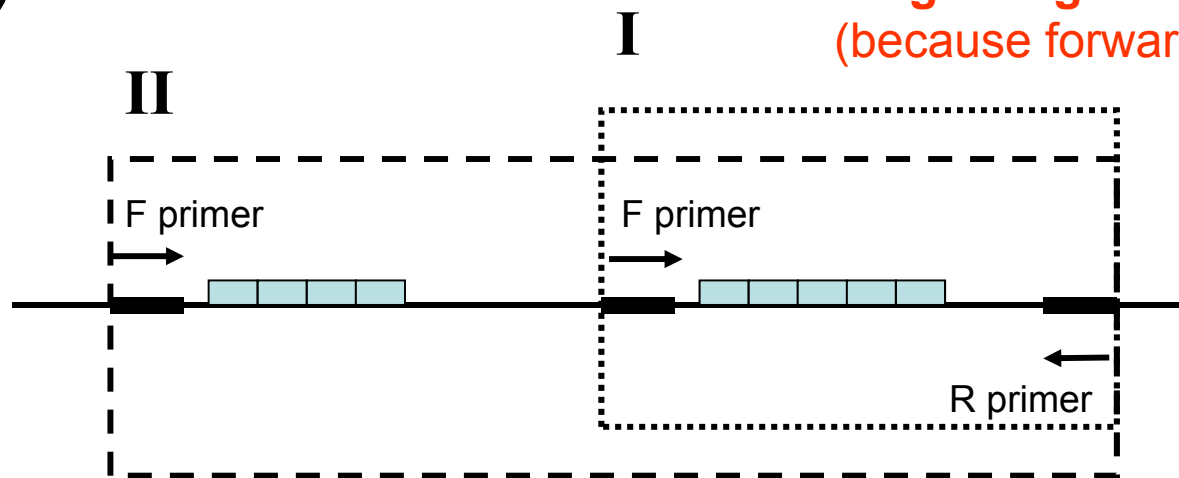


*Duplicated regions are
40,775 bp apart and facing
away from each other*

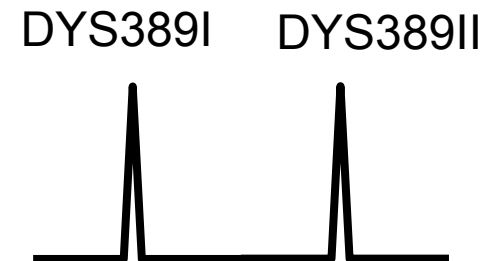
Multi-Copy (Duplicated) Marker



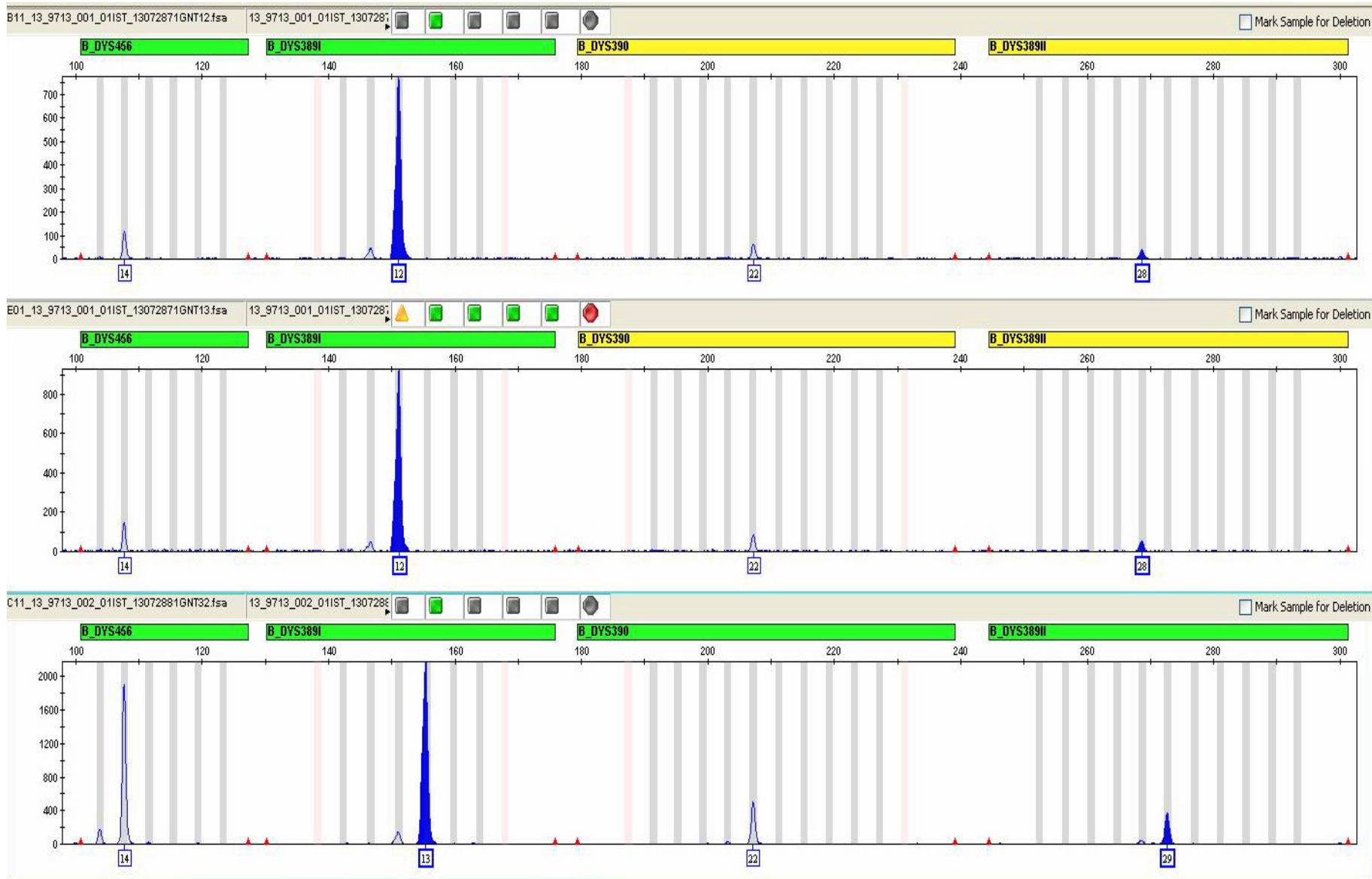
(B) DYS389 I/II



**Single Region but Two PCR Products
(because forward primers bind twice)**



DYS389 I-II mutation





R39: 101055 haplotypes

Search

Haplotypes

SNPs

Populations

Contributors

Contributions

Analyse

Research

Contribute

Meet

DYS19

16

DYS389I

13

DYS389II

31

DYS390

25

DYS391

12

DYS392

11

DYS393

13

DYS385

14.16

National database

Metapopulations

SNP

Search

Whole database

DYS438

10

DYS439

13

DYS437

15

DYS448

20

DYS456

15

DYS458

18

DYS635

23

YGATAH4

11

Please note: The database size will vary based on the loci you have entered.

- 7 loci haplotype (DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393): **101055 haplotypes**
- 9 loci haplotype (+ DYS385a/b): **99258 haplotypes**
- 11 loci haplotype (+ DYS438, DYS439): **72171 haplotypes**
- 12 loci haplotype (+ DYS437): **52628 haplotypes**
- 17 loci haplotype (+ DYS448, DYS456, DYS458, DYS635, YGATAH4): **40987 haplotypes**

Y-SNPs:

- 124 Y-SNP branches (defined by 134 Y-SNP markers)
- 9039 haplotypes with Y-SNP information



YHRD by Sascha Willuweit & Lutz Roewer is licensed under a Creative Commons Attribution-NonCommercial-Share Alike 3.0 Unported License.

Supported by



Endorsed by



R39: 101055 haplotypes

Contributors
Contributions

Analyse

Research

Contribute

Meet

DYS19	DYS359I	DYS359II	DYS390	DYS391	DYS392	DYS393	DYS385
16	13	31	25	12	11	13	14.16
DYS435	DYS439	DYS437	DYS445	DYS456	DYS455	DYS635	YGATAH4
10	13	15	20	15	18	23	11

National database | Metapopulations | SNP

Whole database

Search Reset

Matches grouped by Metapopulations

Matches grouped by Continents

Matches grouped by Haplogroups

Frequency surveying estimates

African - Afro-American

Frequency estimates with given haplotype not included in the database: Mean: 3.366×10^{-4} , Mode: 2.843×10^{-4}

Frequency estimates with given haplotype included in the database: Mean: 3.889×10^{-4} , Mode: 3.366×10^{-4}

Afro-Asiatic - Semitic

Frequency estimates with given haplotype not included in the database: Mean: 4.267×10^{-4} , Mode: 4.064×10^{-4}

Frequency estimates with given haplotype included in the database: Mean: 4.47×10^{-4} , Mode: 4.267×10^{-4}

East Asian - Japanese

Frequency estimates with given haplotype not included in the database: Mean: 5.41×10^{-4} , Mode: 4.677×10^{-4}

Frequency estimates with given haplotype included in the database: Mean: 6.143×10^{-4} , Mode: 5.41×10^{-4}

East Asian - Korean

Frequency estimates with given haplotype not included in the database: Mean: 1.786×10^{-4} , Mode: 1.395×10^{-4}

Frequency estimates with given haplotype included in the database: Mean: 2.177×10^{-4} , Mode: 1.786×10^{-4}

East Asian - Sino-Tibetan - Chinese

Frequency estimates with given haplotype not included in the database: Mean: 6.028×10^{-5} , Mode: 3.951×10^{-5}

Frequency estimates with given haplotype included in the database: Mean: 8.104×10^{-5} , Mode: 6.028×10^{-5}

Eurasian - Altaic

Frequency estimates with given haplotype not included in the database: Mean: 5.634×10^{-4} , Mode: 4.953×10^{-4}

Frequency estimates with given haplotype included in the database: Mean: 6.315×10^{-4} , Mode: 5.634×10^{-4}

Eurasian - European - Eastern European

Frequency estimates with given haplotype not included in the database: Mean: 7.657×10^{-5} , Mode: 3.381×10^{-5}

Frequency estimates with given haplotype included in the database: Mean: 1.193×10^{-4} , Mode: 7.658×10^{-5}

Eurasian - European - South-Eastern European

Frequency estimates with given haplotype not included in the database: Mean: 3.772×10^{-4} , Mode: 2.878×10^{-4}

Frequency estimates with given haplotype included in the database: Mean: 4.667×10^{-4} , Mode: 3.772×10^{-4}

Eurasian - European - Western European

Frequency estimates with given haplotype not included in the database: Mean: 2.693×10^{-5} , Mode: 1.444×10^{-5}

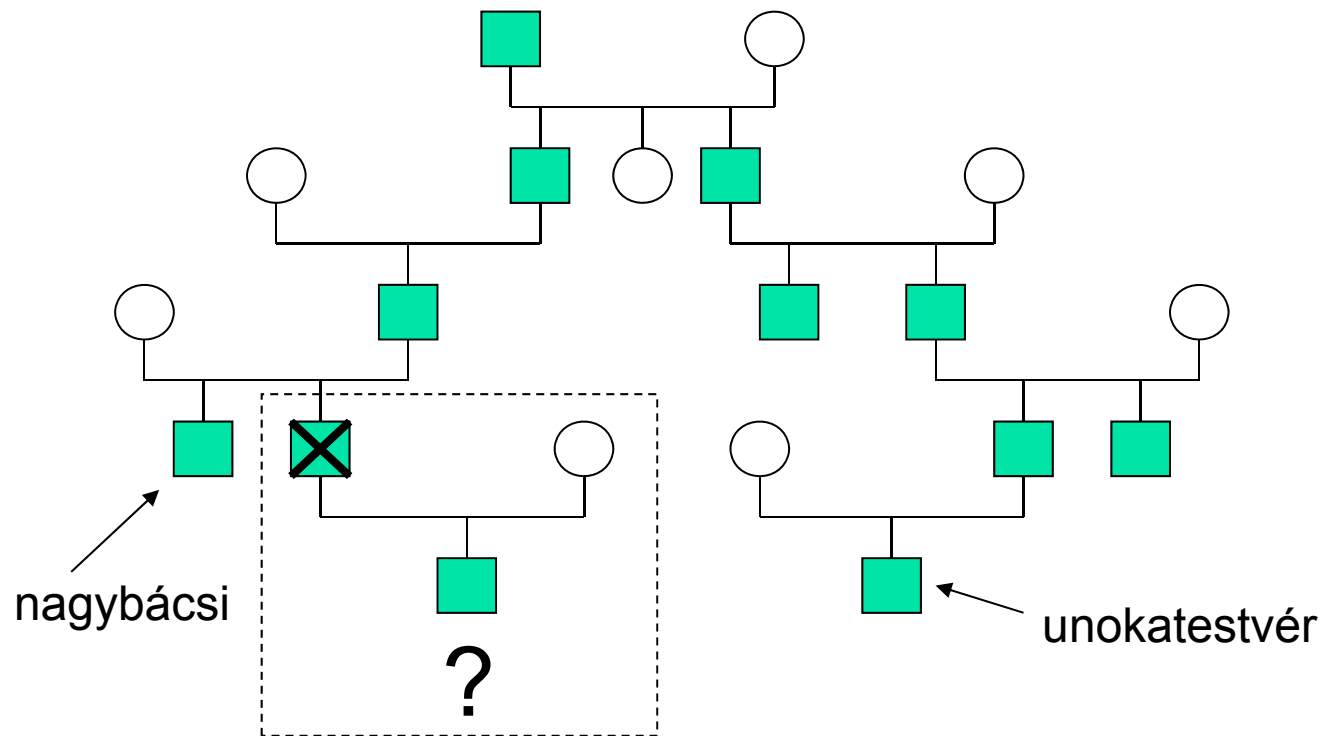
Frequency estimates with given haplotype included in the database: Mean: 3.941×10^{-5} , Mode: 2.693×10^{-5}

Counting Method

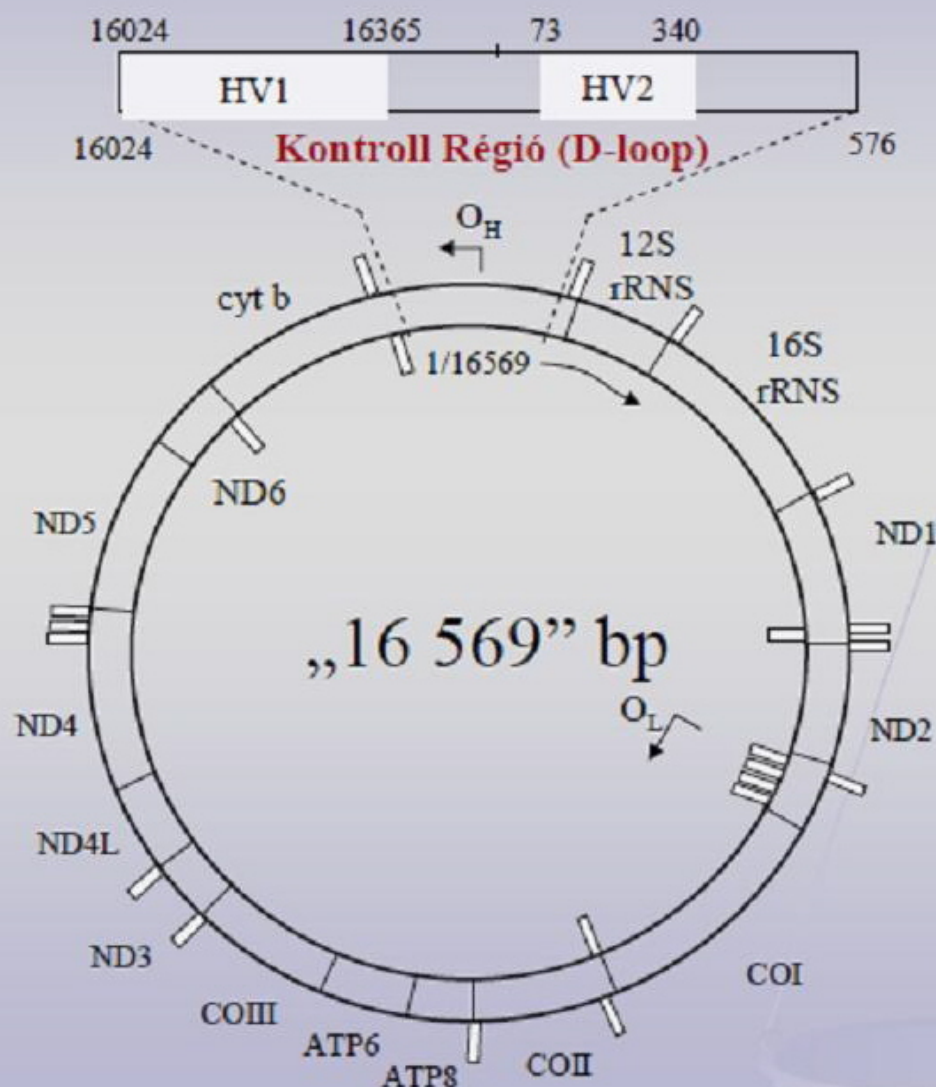
95% confidence interval

$$1-(0.05)^{1/N} \approx 3/N$$

Y kromoszóma vizsgálat során a férfi ági fiú rokonok mindegyike potenciális mintaadó



A kontroll régió

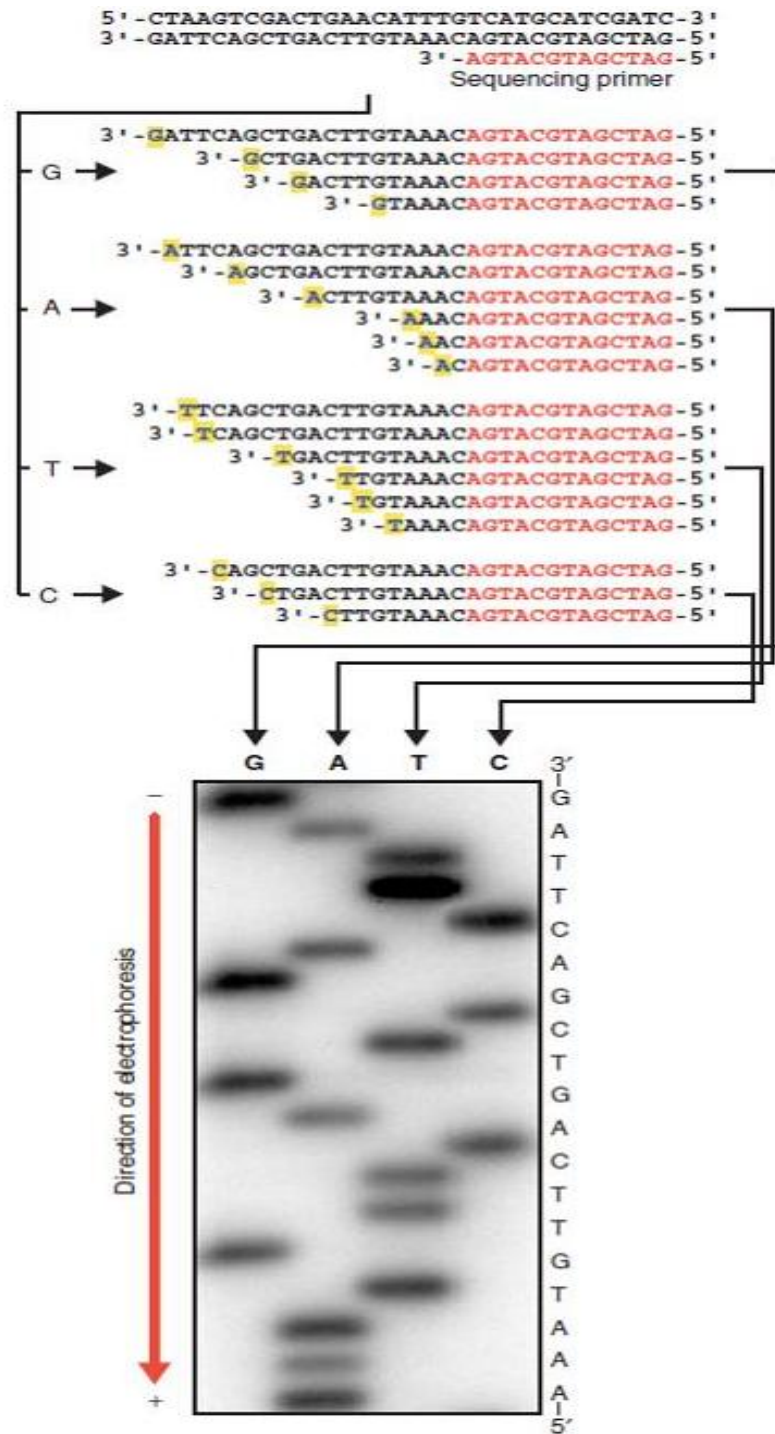


■ Kontroll Régió (CR) – D(isplacement)-loop

- ~1120 bp
- replikációs origó
- hipervariábilis szakaszok
 - HV1, HV2 és HV3
- alacsony szelekciós nyomás
- a variációk forrása csak mutáció

■ Szekvencia-adatbázisok

- EMPOP



D-loop: anyai leszármazási vonal genetikai azonosítása

A vizsgált minták mtDNS szekvenciáját a referenciaként használt rCRS-hez illesztjük (pl., 16071-16140)

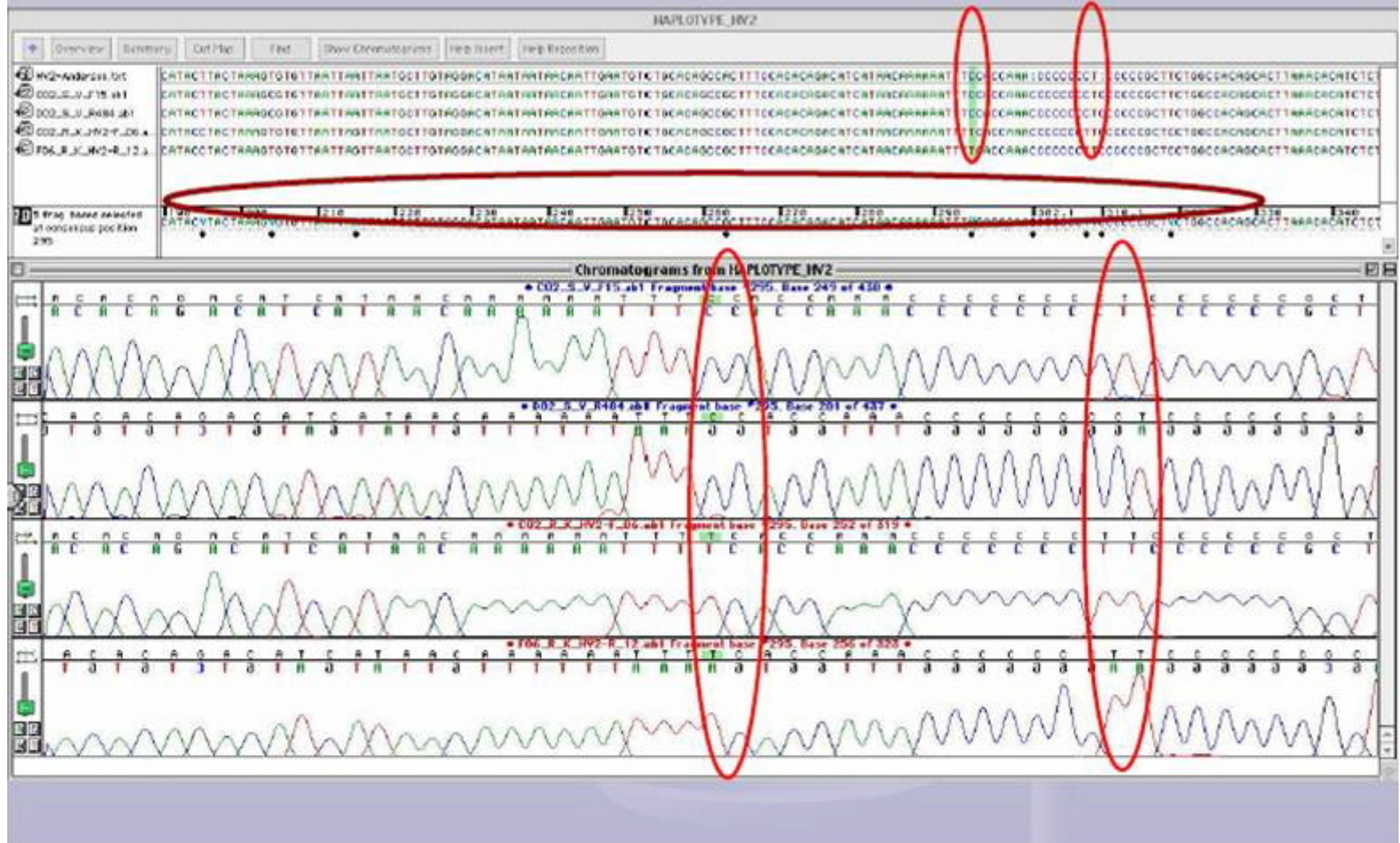
	16090	16100	16110	16120	16130	16140
rCRS	ACCGCTATGT	ATTCGGTACA	TTACTGCCAG	CCACCATGAA	TATTGTACGG	TACCATAAAT
Q	ACCGCTATGT	ATTCGGTACA	TTACTGCCAG	CCACCATGAA	TATTGTACAG	TACCATAAAT
K	ACCGCTATGT	ATTCGGTACA	TTACTGCCAG	CCACCATGAA	TATTGTACAG	TACCATAAAT

16093
16129

A referencia szekvenciától való eltérések pozícióját és nukleotidját jegyezzük le, ami megadja az adott minta mtDNS haplotípusát

<u>Bizonyíték</u>	<u>Összehasonlító</u>	
16093C	16093C	Bizonyíték: - Kizárható
16129A	16129A	- Inkonzkluzív
		- Nem zárható ki

A haplotípusok ...



Igazságügyi szempontok ...

A mtDNS egyedi azonosításra kevésbé alkalmas, mint a nuDNS

- Nem az egyént, hanem az anyai leszármazási haplotípust identifikálja

Ugyanakkor

- A diverzitás mértéke miatt nagy a valószínűsége annak, hogy két véletlenül kiválasztott személy különböző mtDNS haplotípussal rendelkezik
 - Populációtól függetlenül ez az érték általában 95 % valószínűség fölé van
 - Új minták esetén kb. 2/3 része a mintáknak új haplotípust mutat
 - Csak kevés ún. “közönséges” gyakori haplotípus van, melyek gyakorisága több mint 1%
 - Több ezer haplotípus figyelhető meg, emiatt a haplotípusok relatív gyakorisága alacsony

EMPOP: mtDNA szekvencia adatbázis

empop.org/modules/haplotype/

Version: 2.1, Release: 6

Welcome egyedb! [Logout](#)

EMPOP

- Home
- Contribute
- Help
- Imprint
- Contact
- Terms of use

▶ Query

▶ Tools

▶ Account

Input | Summary

Type

Sample Info

Query	Range	Profile
<input type="text"/>	<input type="text" value="16024-576"/>	<input type="text" value="16129A 93G 263G 315.1C"/>
<input type="text"/>	<input type="text"/>	<input type="text"/>
<input type="text"/>	<input type="text"/>	<input type="text"/>

Options

Match type: pattern literal

Number of differences displayed:

Disregard InDels in length variants at positions:
 16193 309 455 573

Source: Forensic (10869) Literature (5252)

endorsed by



www.empop.org

EMPOP

- Home
- Contribute
- Help
- Imprint
- Contact
- Terms of use

▶ Query

▶ Tools

▶ Account

Input **Summary**

Results Identification: d3d282a3fd917b56b101259843bb8a760b033543

Search execution date: 2012-02-16 22:15:47 UTC

Sample Info: not specified

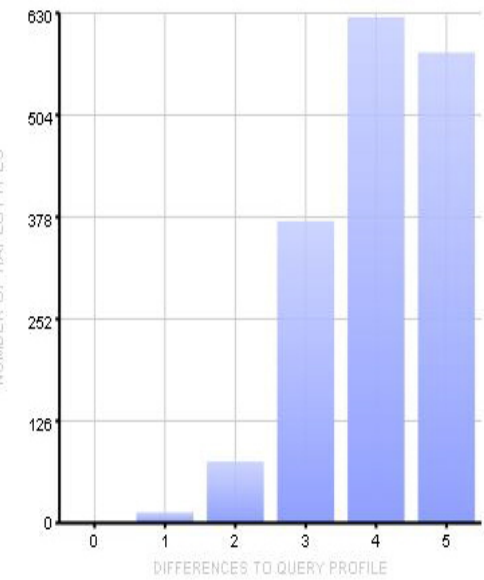
Type: string-based search: haplotype as differences to rCRS

Options: Match type: **pattern**
 Maximum differences displayed: 5
 Disregard InDels in length variants at positions: **16193 309 455 573**

Source: **Forensic data (9052/10869)**
Literature data (1789/5252)

Query: 16024-576 G16129A A93G A263G -315.1C

Geographic affiliation		Metapopulation	
All		All	
DIFFERENCES TO QUERY PROFILE	NUMBER OF HAPLOTYPES	CUMULATIVE NUMBER OF HAPLOTYPES	
0	0	0	
1	1	1	
2	76	77	
3	373	450	
4	625	1075	
5	581	1656	
6+	9185	10841	



Frequency estimates

P_{Uc} **0.000e+0** [2.710e-20 ; 3.542e-4]

P_{N+1} **9.223e-5** [1.628e-5 ; 5.224e-4]



Counting Method

CALCULATION OF mtDNA PROFILE FREQUENCY ESTIMATES USING THE COUNTING METHOD

In cases where an mtDNA profile is observed a particular number of times (X) in a database containing N profiles, its frequency (p) can be calculated as follows:

$$p = X/N$$

A 95% upper bound confidence interval can be placed on the profile's frequency using:

$$p + 1.96 \sqrt{\frac{(p)(1-p)}{N}}$$

In cases where the profile has not been observed in a database, the 95% upper bound on the confidence interval is

$$1 - \alpha^{1/N} = 1 - (0.05)^{1/N}$$

where $\alpha = 0.05$ is the confidence coefficient and N is the number of individuals in the database.

For example, the mtDNA type 16129A, 263G, 309d, 315.1C occurs twice in 1148 African-American profiles, twice in 1655 Caucasian profiles, and not at all in 686 Hispanic profiles when searched against the mtDNA Population Database (Monson et al. 2002). Using the equations above, calculations for the rarity of this profile in the respective sample sets are as follows:

For African-Americans: $p = 2/1148 + 1.96 [(2/1148)(1 - (2/1148))/1148]^{1/2} = 0.0017 + 0.002 = 0.004 = 0.40\%$

For Caucasians: $p = 2/1655 + 1.96 [(2/1655)(1 - (2/1655))/1655]^{1/2} = 0.0012 + 0.0017 = 0.0029 = 0.29\%$

For Hispanics: $1 - (0.05)^{1/686} = 1 - 0.9956 = 0.0044 = 0.44\%$

These calculations demonstrate that the statistical weight can be similar whether or not a match is found to a few previously observed samples in a database.

As shown in D.N.A. Box 13.2, the Clopper-Pearson method (Clopper & Pearson 1934) may also be used to provide a conservative estimate for the upper 95% confidence interval.

Sources:

- Clopper, C.J., & Pearson, E.S. (1934). *The use of confidence or fiducial limits illustrated in the case of the binomial*. *Biometrika*, 26, 404-413.
- Evett, I.W., & Weir, B.S. (1998). *Interpreting DNA Evidence*. Sunderland, MA: Sinauer Associates, Inc., p. 142.
- Monson, K.L., et al. (2002). *The mtDNA population database: an integrated software and database resource*. *Forensic Science Communications*, 4(2). Available at <<http://www2.fbi.gov/hq/lab/fsc/backissu/april2002/miller1.htm>>.
- Tully, G., et al. (2001). *Considerations by the European DNA profiling (EDNAP) group on the working practices, nomenclature and interpretation of mitochondrial DNA profiles*. *Forensic Science International*, 124, 83-91.

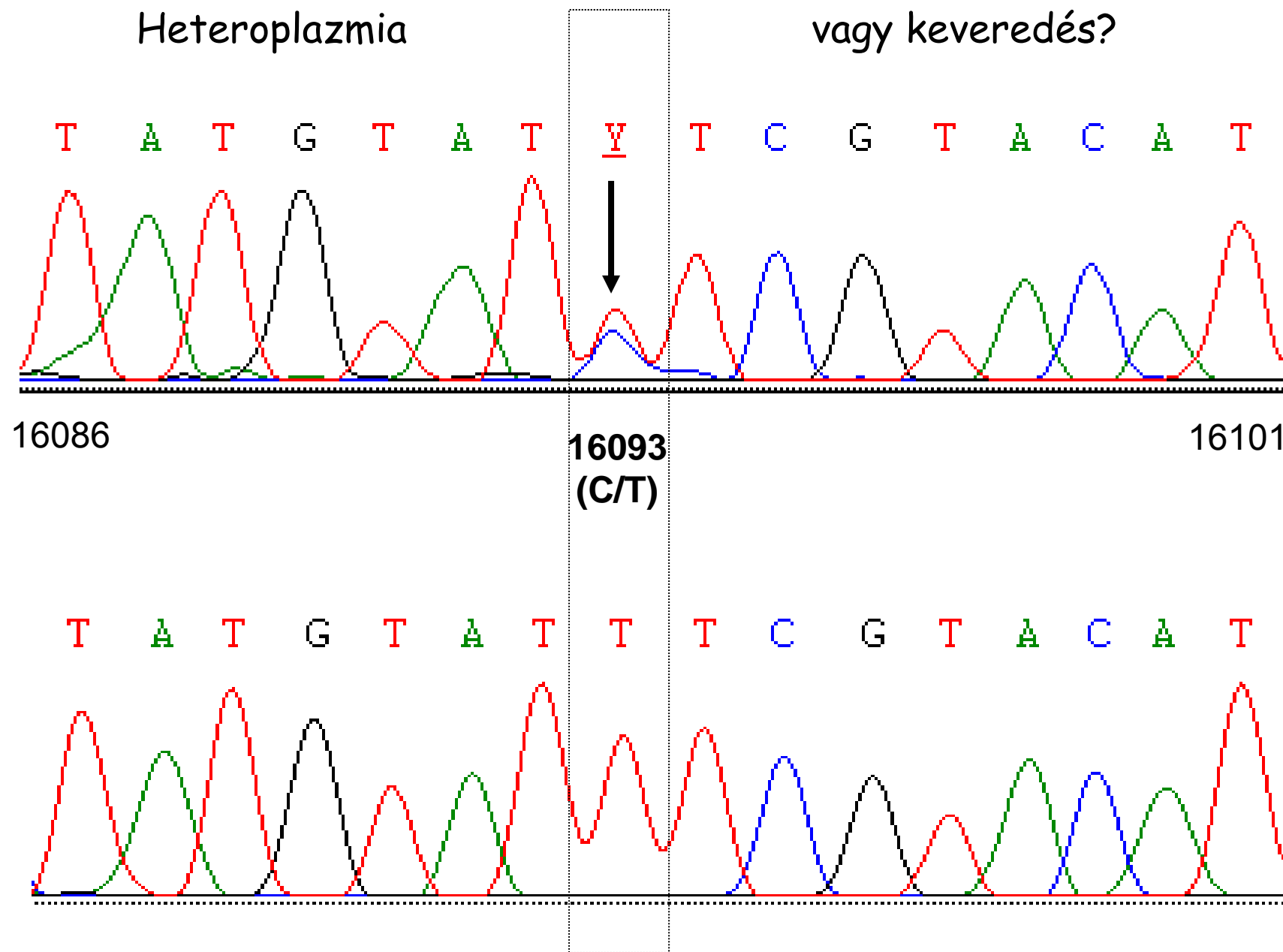
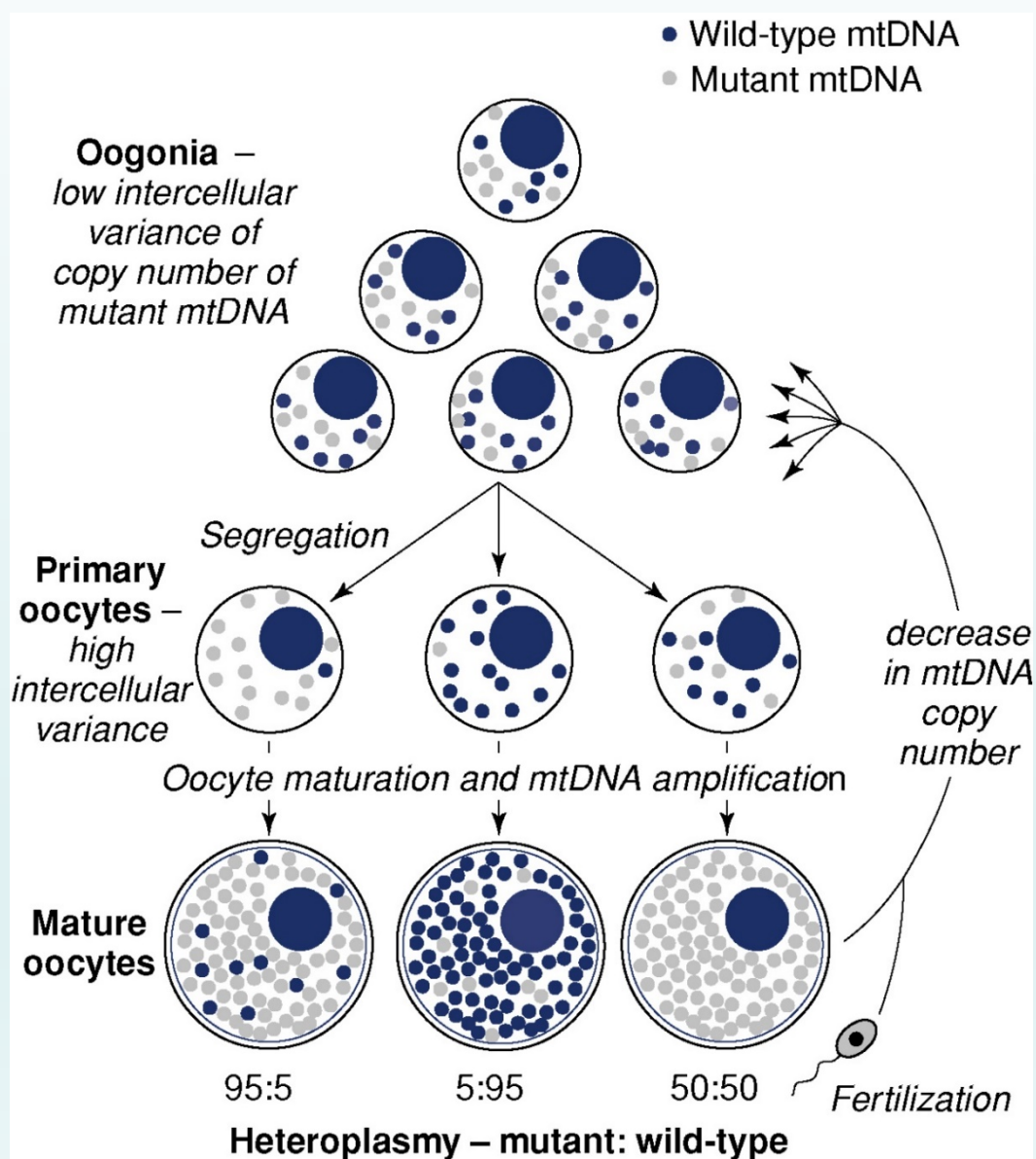


Figure 10.9, J.M. Butler (2005) *Forensic DNA Typing*, 2nd Edition © 2005 Elsevier Academic Press

Mitokondriális DNS mutációk: heteroplazmia



Deléció: mitokondriális eredetű betegségek (anyai)

- MITOMAP

Szubsztitúció: ált. neutrális

Mutációs ráta: non-uniform

$10^{-6} - 10^{-7}$ / bp / generation

Mutációs hot-spotok

Citoplazmikus szegregáció:

„bottleneck”
(palacknyak effektus)

Heteroplazmia - homoplazmia:

szövet- és módszer specifikus

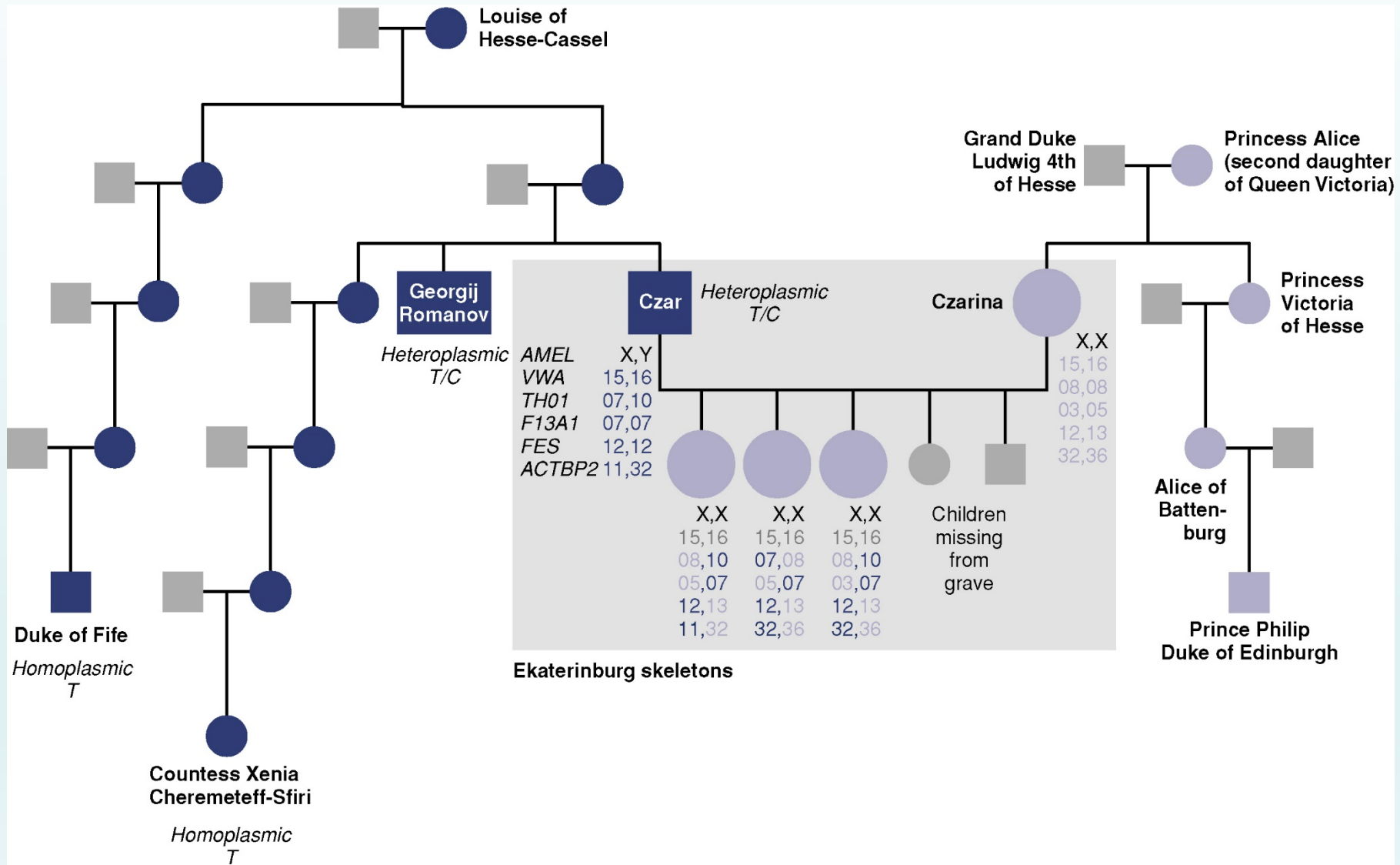
Identification of the remains of the Romanov family by DNA analysis

Peter Gill¹, Pavel L. Ivanov², Colin Kimpton¹, Romelle Piercy¹, Nicola Benson¹, Gillian Tully¹, Ian Evett¹, Erika Hagelberg³ & Kevin Sullivan¹

Nine skeletons found in a shallow grave in Ekaterinburg, Russia, in July 1991, were tentatively identified by Russian forensic authorities as the remains of the last Tsar, Tsarina, three of their five children, the Royal Physician and three servants. We have performed DNA based sex testing and short tandem repeat (STR) analysis and confirm that a family group was present in the grave. Analysis of mitochondrial (mt) DNA reveals an exact sequence match between the putative Tsarina and the three children with a living maternal relative. Amplified mtDNA extracted from the remains of the putative Tsarina has been cloned to demonstrate heteroplasmy at a single base within the mtDNA control region. One of these sequences matches two living maternal relatives of the Tsar. We conclude that the DNA evidence supports the hypothesis that the remains are those of the Romanov family.



A Romanovok maradványainak azonosítása molekuláris genetikai módszerrel



Mitokondriális DNS: PCR és szekvenálási stratégia

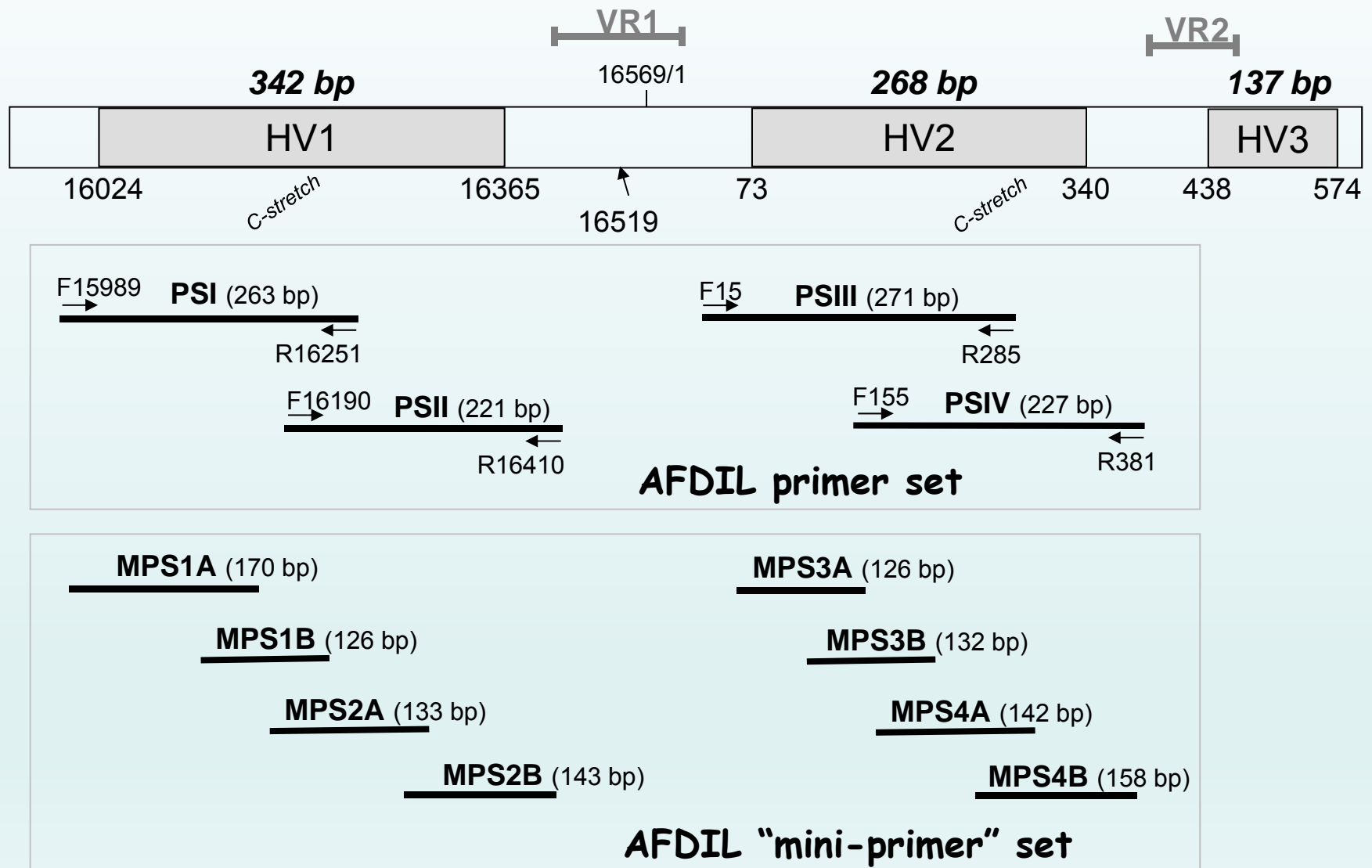
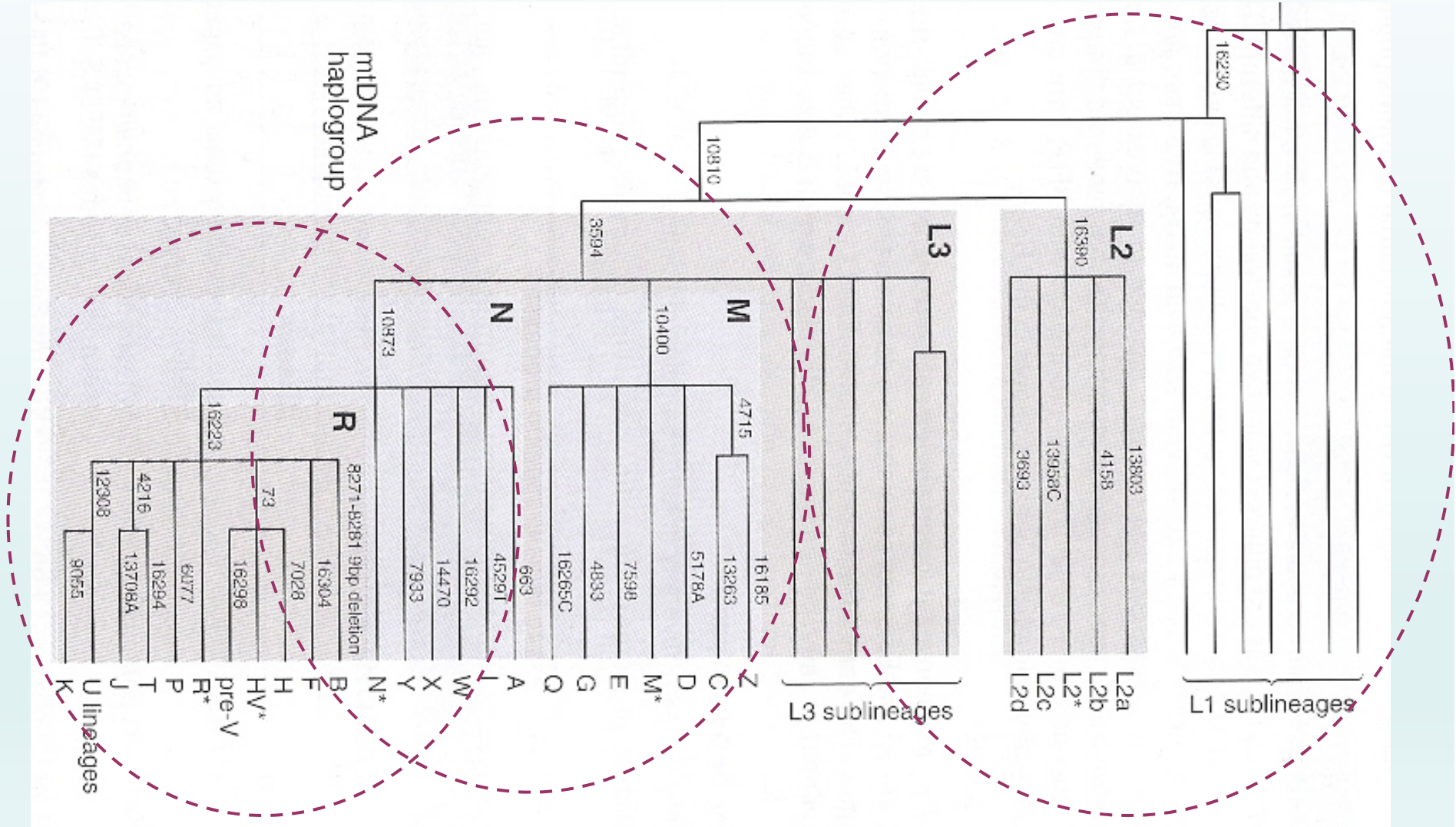
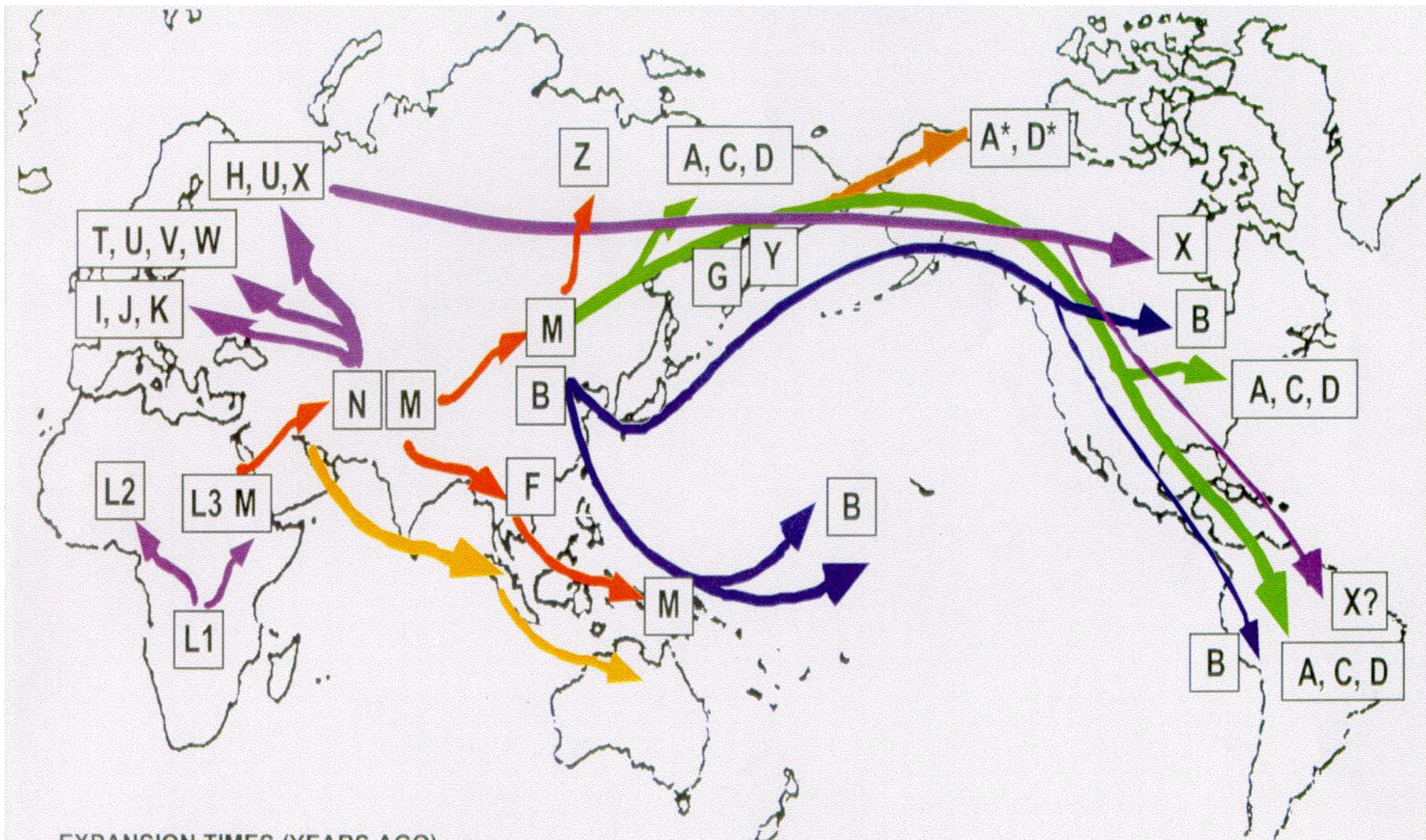


Figure 10.3, J.M. Butler (2005) *Forensic DNA Typing*, 2nd Edition © 2005 Elsevier Academic Press

A fő mitokondriális DNS leszármazási vonalak



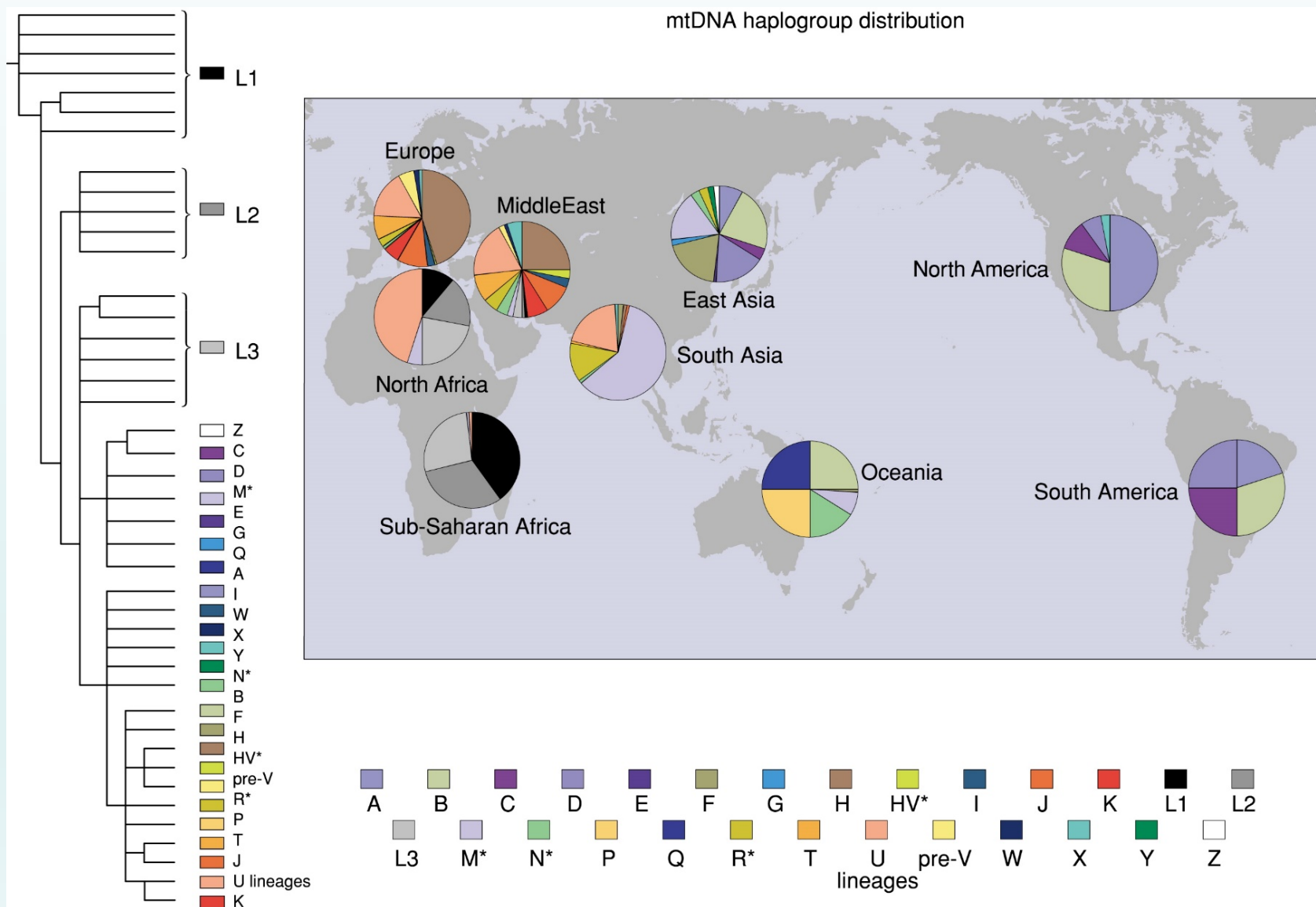


EXPANSION TIMES (YEARS AGO)

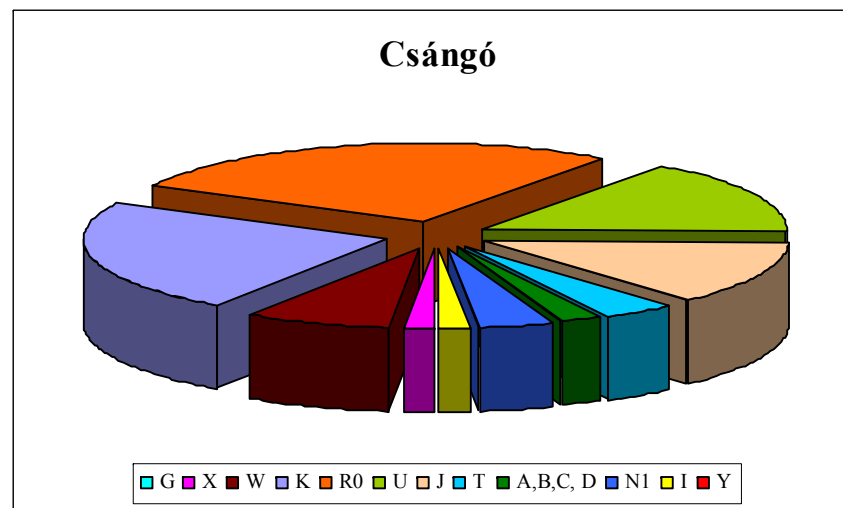
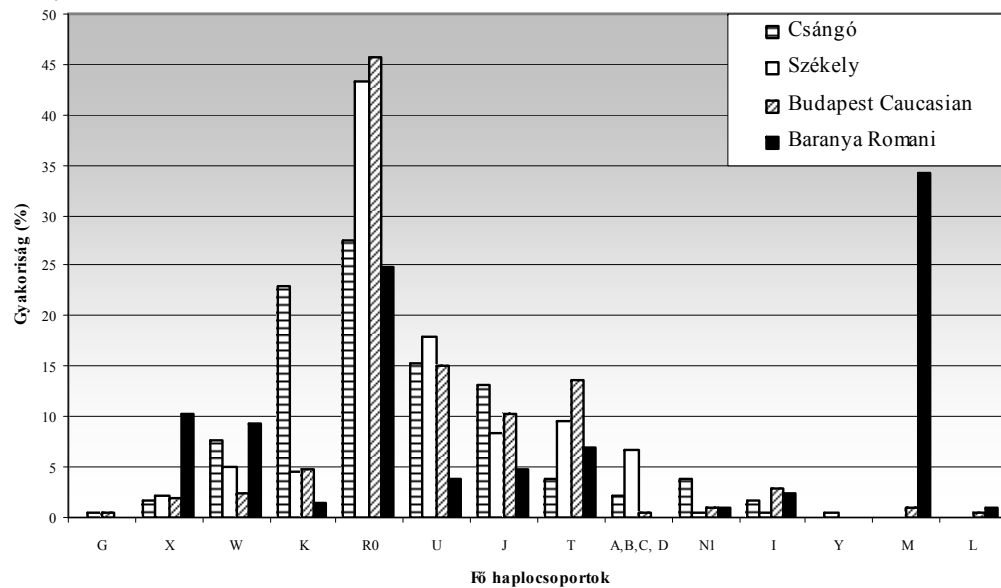
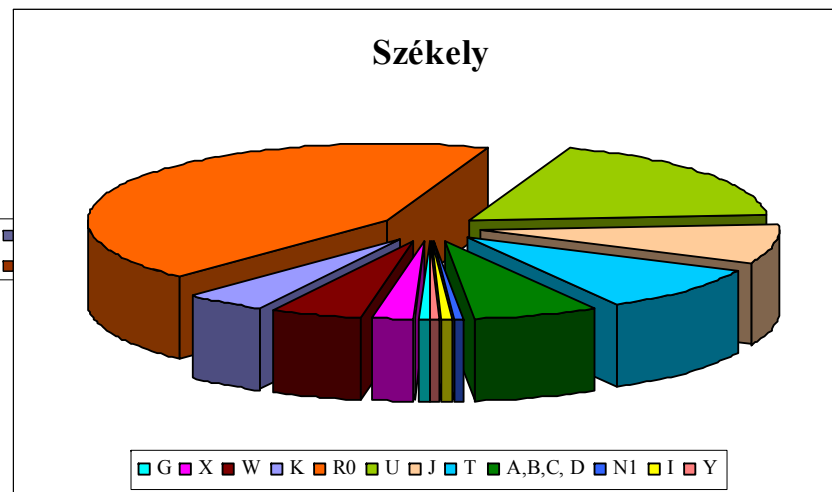
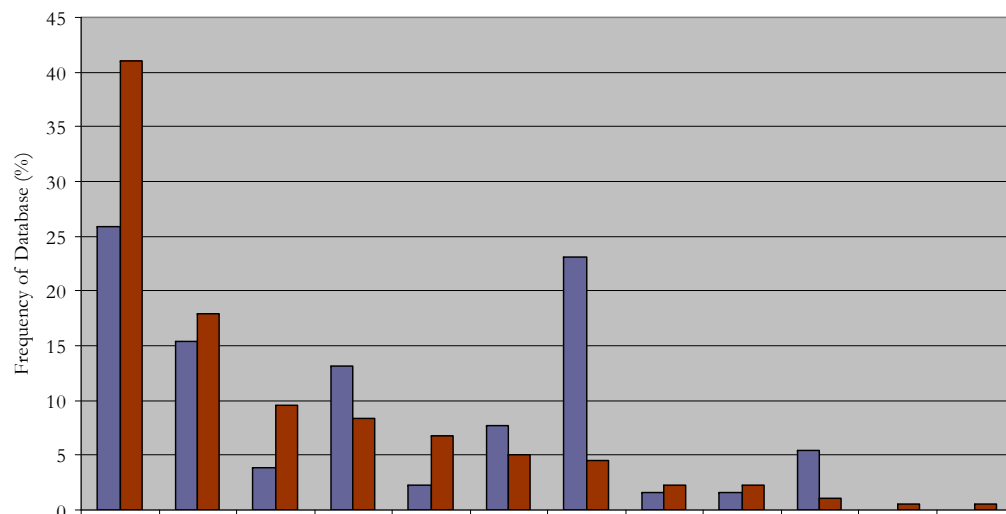
Africa	120,000- 150,000
Out of Africa	55,000- 75,000
Asia	40,000- 70,000
Australia/PNG	40,000- 60,000
Europe	35,000- 50,000
Americas	15,000- 35,000
Na-Dene/Esk/Aleuts	8,000- 10,000

Family Tree DNA
mtDNA Migrations Map

A fő mitokondriális DNS haplocsoportok eloszlása



Mitokondriális DNS haplocsoport eloszlások



A fő mitokondriális DNS haplocsoportok eloszlása a nyugat-eurázsiai kontinensen

