



Humán polimorf mikroszatellita (short tandem repeat) lokuszok igazságügyi genetikai vizsgálata magyar populációkban

DOKTORI ÉRTEKEZÉS

Készítette:

Füredi Sándor

ELTE TTK Biológia Doktori Iskola

Iskolavezető: dr. Erdei Anna

Klasszikus és molekuláris genetika Doktori Program

Programvezető: dr. Orosz László

Témavezető: dr. Lontainé dr. Santora Zsófia
igazságügyi szerológus, intézetigazgató
a biológiai tudományok kandidátusa

ORFK Bűnügyi Szakértői és Kutatóintézet

2003

TARTALOMJEGYZÉK

1. BEVEZETÉS	4
1.1. IRODALMI ÉS TÖRTÉNETI ÁTTEKINTÉS	5
1.1.1. <i>Az igazságügyi genetikai vizsgálatok alapkérdései</i>	5
1.1.1.1. Az igazságügyi genetikai vizsgálatok statisztikai interpretációjának alapjai	7
1.1.2. <i>Az igazságügyi genetikai vizsgálatok rövid történeti áttekintése</i>	8
1.1.3. <i>A mikroszatellita (STR) lokuszok általános jellemzése</i>	9
1.1.4. <i>Az igazságügyi célokból analizált STR lokuszok és allélok nevezéktana</i>	11
1.1.5. <i>Az Y-kromoszómás STR markerek analízise az igazságügyi genetikában</i>	12
1.2. CÉLKITŰZÉSEK	14
1.2.1. <i>Humán STR markerek vizsgálatának laboratóriumi bevezetése, populációs minták összeállítása</i>	14
1.2.2. <i>Az alkalmazott automatikus genotipizáló rendszerek allélméret-meghatározó pontosságának hitelesítése</i>	17
1.2.3. <i>A populációs felmérések során detektált ritka, mikrovariáns allélok jellemzése</i>	17
1.2.4. <i>A vizsgált STR polimorfizmusok géndiverzitásának és személyazonosító hatékonyságának megállapítása a magyar népességben</i>	17
1.2.5. <i>A vizsgált populációk Hardy–Weinberg és linkage egyensúlyi tesztelése a vizsgált autoszómás STR lokuszokon</i>	18
1.2.6. <i>A felmérésbe bevont magyarországi populációk genetikai strukturáltságának megállapítása</i> 18	
1.2.7. <i>A vizsgált magyarországi populációk genetikai varianciájának hatásai az igazságügyi statisztikai interpretációra</i>	19
1.2.8. <i>A biológiai nyomot hátrahagyó személy etnikumának valószínűsítési lehetőségei: az etnikai eredet becslése (EEO)</i>	19
2. ANYAGOK ÉS MÓDSZEREK.....	20
2.1. A DNS KINYERÉSE A BIOLÓGIAI MINTÁKBÓL	20
2.2. A DNS-KONCENTRÁCIÓ MEGHATÁROZÁSA	20
2.3. AZ STR LOKUSZOK PCR-SOKSZOROZÁSA	21
2.4. A PCR HATÉKONYSÁGÁNAK ELLENŐRZÉSE	21
2.5. STR GENOTIPIZÁLÁS DNS-FRAGMENTANALÍZISSSEL.....	22
2.5.1. <i>Egy lokuszos (szimplex) genotipizálás a VWA és TH01 lokuszokon</i>	22
2.5.2. <i>Három lokuszos (triplex) genotipizálás a VWA, FES, HPRTB, TH01, TPOX, CSF1PO lokuszokon</i>	23
2.5.3. <i>Négy lokuszos (kvadruplex) genotipizálás VWA, TH01, TPOX, CSF1PO, LPL, F13B, FES, F13A01 lokuszokon</i>	23
2.5.3.1. A genotipizálás allélméret-meghatározás pontosságának (allele sizing accuracy, precision) mérése.24	
2.5.4. <i>Több lokuszos (multiplex) genotipizálás a CODIS STR és Amelogenin lokuszokon</i>	24
2.5.5. <i>Három lokuszos (Triplex I, Triplex II) és két lokuszos (Duplex) genotipizálás (haplotipizálás) az Y-STR lokuszokon</i>	25
2.5.5.1. A haplotipizálás allélméret-meghatározás precizitásának (allele sizing precision) mérése.....	25
2.6. A RITKA, VARIÁNS STR ALLÉLOK SZEKVENCIA ANALÍZISE	26
2.7. POPULÁCIÓ- ÉS STATISZTIKAI GENETIKAI ANALÍZISEK	27
2.6.1. <i>Populáción belüli elemzések</i>	27
2.6.1.1. Populációgenetikai alapértékek: Nei-féle géndiverzitás (h), a polimorfizmus információ tartalom (PIC), az igazságügyi megkülönböztető erő (power of discrimination, PD), az egyezési valószínűség (probability of match, pM), szülősi kizáró erő (power of exclusion, PE), minimális allélgyakoriság (p _{min})	27
2.6.1.2. A Hardy-Weinberg egyensúly (HWE) tesztelése.....	29
2.6.1.3. A linkage disequilibrium (LD) tesztelése.....	29
2.6.1.4. Tesztsorozat, kísérleti jellegű hibaráta, Bonferroni eljárás	30
2.6.2. <i>Populációk összehasonlító elemzése igazságügyi genetikai aspektusból</i>	31
2.6.2.1. Populáció-párok homogenitás-vizsgálata G-statisztikával.....	31
2.6.2.2. F-statisztika és AMOVA.....	31
2.6.2.3. Filogenetikai fa konstrukció	32
2.6.2.4. Az átlagolt fenotípuskülönbség-értékek [mean phenotype difference values (MVs)] meghatározása 33	
2.6.2.5. Logaritmikus sörét- (scatter) és egy személyű (within-person) Tippett-diagramok	33
2.6.2.7. Az etnikai eredet becslése (EEO).....	35

3	EREDMÉNYEK ÉS ÉRTÉKELÉSÜK	37
3.1.	AZ STR LOKUSZOK GENOTIPIZÁLÁSÁNAK BEMUTATÁSA	37
3.2.	A GENOTIPIZÁLÁS (FÉL)AUTOMATIKUS ALLÉLMÉRET-MEGHATÁROZÁSÁNAK PONTOSSÁGA (ALLELE SIZING ACCURACY, PRECISION).....	38
3.3.	RITKA ÉS MIKROVARIÁNS ALLÉLOK.....	42
3.3.1.	<i>A D7S820 lokusz egy 8.-1-es, egy 9.+1-es és egy 10-es mikrovariáns alléljának elemzése...</i>	42
3.3.2.	<i>A D13S317 lokusz egy 12-es és egy ritka 7-es alléljának elemzése</i>	44
3.3.3.	<i>A HPRTB lokusz egy 12.-2-es és egy 12.-6-os mikrovariáns alléljának elemzése.....</i>	45
3.3.4.	<i>Az Y-STR lokuszok ritka alléljainak elemzése.....</i>	46
3.4.	AZ ÖT POPULÁCIÓS MINTA IGAZSÁGÜGYI GENETIKAI ANALÍZISÉNEK POPULÁCIÓGENETIKAI ÉS POPULÁCIÓSTATISZTIKAI EREDMÉNYEI	51
3.4.1.	<i>Autoszómás lokuszok és az X-kapcsolt HPRTB lokusz vizsgálatának eredményei</i>	52
3.4.1.1.	A genotipizálás eredményei, konkordancia-vizsgálat, allélgyakorisági értékek és a lokuszok polimorfizmusának foka.....	52
3.4.1.2.	A Hardy-Weinberg és kapcsoltsági (linkage) egyensúlyi vizsgálatok eredményei.....	56
3.4.1.3.	Az elemzett populációk genetikai struktúrája	59
3.4.1.4.	A magyarországi populációs adatbázisok igazságügyi statisztikai célú alkalmazhatóságának elemzése 64	
3.4.1.5.	A biológiai nyomot hátrahagyó személy etnikumának valószínűsítési lehetőségei: az etnikai eredet becslése (EEO).....	75
3.4.2.	<i>Y-kromoszómás STR lokuszok vizsgálatának eredményei.....</i>	78
3.4.2.1.	Az Y-haplotipizálás eredményei.....	78
3.4.2.2.	A populációs felméréseink során detektált Y-haplotípusok kiértékelése	79
3.4.2.3.	Az Y-haplotípusok populációstatisztikai összehasonlításának eredményei	85
3.5.	ESETTANULMÁNY	88
4.	ÚJ MEGÁLLAPÍTÁSOK, A KÖVETKEZTETÉSEK LEVONÁSA.....	89
5.	KÖSZÖNETNYILVÁNÍTÁS.....	101
6.	IRODALOMJEGYZÉK	103
7.	FÜGGELÉK.....	114
7.1.	ÁBRÁK	115
7.2.	TÁBLÁZATOK	130
7.3.	RÖVIDÍTÉSEK JEGYZÉKE	137
8.	ÖSSZEFOGLALÁS.....	139
9.	SUMMARY.....	140

1. BEVEZETÉS

Az igazságügyi genetika, mint különálló tudományterület, nagykorú lett. Ez a nagykorúság nemcsak abból fakad, hogy törvényszéki eljárásban genetikai vizsgálatot először 18 évvel ezelőtt, 1985-ben alkalmaztak Angliában (Jeffreys és mtsai, 1985a, 1985b), hanem abból is, hogy mára már a világ szinte minden országában, Norvégiától Új-Zélandig, az USA-tól Japánig az igazságszolgáltatás szerves részét képezi (Brinkmann és Carracedo, 2003). A világ számos állama, az európai országok többsége az ujjlenyomat adatbázisokhoz hasonló, ún. nemzeti bűnügyi DNS-nyilvántartásokat állított föl, ahol a bűncselekmények helyszínéről származó, be nem azonosított biológiai nyomok DNS-információja mellett akár több százezer bűnelkövető DNS-mintázatát is őrzik (Interpol DNA Unit, 2003).

Az igazságügyi genetika, mint alkalmazott tudományág segítséget nyújthat bűncselekmények felderítésében, származás-megállapítások során (pl. apasági tesztekben), valamint ismeretlen személyazonosságú holttestek, emberi maradványok beazonosításában. A molekuláris biológiai technológia és az informatika szédítő fejlődésével azonban már napjainkban lehetővé vált, vagy a közeljövőben lehetővé válik a törvényszéki genetikai módszereknek az előbbieken említetteknél szélesebb körű felhasználása. Az egyik ilyen terület a személyazonosság DNS-sel történő megállapítása, a másik, ezzel összefüggésben lévő lehetőség az adott személy bizonyos tulajdonságainak (nem, kor, szemszín, testfelépítés, egészségi állapot stb.) a személy DNS-mintájából történő kiolvasása. Ez utóbbi két terület azért is olyan izgalmas és egyben potenciális veszélyekkel terhelt, mivel a vizsgálatra alkalmas DNS-minta származhat a célszemély akár egyetlen sejtjéből is. Csak szemléltetésképpen, tüszentéssel több ezer sejtet juttatunk a környezetünkbe, cigarettacsikkünkön, de még akár ujjlenyomatunkban is található a vizsgálatokhoz elegendő mennyiségű DNS (Oorschot és Jones, 1997). Amint az ismeretterjesztés és a média jóvoltából a társadalom széles rétege értesül ezekről a lehetőségekről és kockázatokról, valószínű, hogy a kutatók, jogalkotók, politikusok, vallási, gazdasági, és civil szervezetek között hasonlóan éles hangnemű, etikát és jogot érintő viták várhatóak, mint amelyet napjainkban az ember klónozása körül tapasztalunk (National Commission of the Future of DNA Evidence, 2001).

1.1. Irodalmi és történeti áttekintés

1.1.1. Az igazságügyi genetikai vizsgálatok alapkérdései

Az igazságügyi genetikai vizsgálatok fő feladata annak a megállapítása, hogy a törvényszéki eljárásban kitüntetett szerepet kapó biológiai nyom (ebben az értelemben „nyomnak” tekinthetjük a származás-megállapítási perekben szereplő utódot) köthető-e egy adott személyhez, vagy az élőlények adott fajához, egyedéhez. Ebből az okfejtésből világosan következik, hogy a biológiai nyom személy- vagy egyedazonosítási célokra történő felhasználása csak akkor lehet eredményes, ha a nyomból kiolvasható olyan tulajdonság, amely csak egyetlen személyre vagy egyedre jellemző. Emellett ennek a tulajdonságnak időben és térben viszonylag állandónak, módosulatlanak kell maradnia, hogy a szakértői összehasonlításokat – amelyek sokszor a kérdéses cselekmény után csak évekkel, évtizedekkel később kerülnek elvégzésre – érdemben meg lehessen tenni.

A törvényszéki genetika, illetve az elődjének tekinthető vércsoport-szerológia több mint száz éves fejlődése során három, egymástól teljesen szét nem választható alapkérdésre keresi a választ:

1. Az élő anyag mely összetevője bír információtartalommal a személyazonosítás szempontjából?
2. Ez az összetevő megváltozik-e, és ha igen, milyen mértékben az élet során, a halál után, illetve az élőlényből való kikerülés vagy átörökítés esetén (Parsons és Weedn, 1997)?
3. Ha összehasonlításra kerül sor, a megállapított típusok egyezése milyen mértékben bizonyítja a biológiai nyom adott személytől való származását (Evet és Weir, 1998)?

Az első kérdésre kezdetben, az 1900-as évek elejétől, a vércsoportok felfedezésével az orvos-biológia, szerológia, az 1960-as évektől, az enzimek és fehérjék elemzésével a biokémia, majd az 1980-as évektől, a genetikai, biotechnológiai ismeretek ugrásszerű gyarapodásával a genetika, molekuláris biológia kereste a válaszokat az igazságügy számára. A második kérdést, speciális jellegénél fogva elsősorban törvényszéki orvosok, kriminalisztikai szakértők, antropológusok firtatják, míg a harmadik kérdésre elsősorban populációs és statisztikai genetikai elemzésekkel lehet megfelelni. Az elmúlt évtizedekben

az ember örökítő anyagára, a DNS-re vonatkozó ismeretek tudtak leginkább megfelelni a törvényszéki eljárás által támasztott, egyre magasabb szintű követelmény-rendszernek. Így érthető, hogy napjainkban a biológiai nyomok igazságügyi célú személyazonosításában – a klasszikus ujjlenyomat-elemzéssel karöltve – a DNS-mintázat vizsgálata az egyeduralkodó.

A földi fajok és egyedek sokféleségének egyik fő meghatározó oka, hogy a természetben megtalálható szinte minden, az öröklődésért felelős nukleinsav-lánc mutációs változást és ennek eredményeképpen polimorfizmust hordoz. A polimorf szakaszok mérete egyetlen nukleotidtól (SNP) többszázezer nukleotidig terjedhet (Strachan és Read, 1999). Egy személy, egyed vagy biológiai nyom DNS-mintájából különböző lokuszokon megállapított allélok együttesét DNS-mintázatnak, genotípusnak vagy DNS-profilnak nevezzük. A DNS-mintázat felfogható személyi számunk genetikai megfelelőjeként. Azonban a személyi számmal ellentétben a DNS-profil nem feltétlenül egyedi, azaz nem 100%-ig biztos, hogy a Földön két személy nem rendelkezhet ugyanazzal a DNS-profillal.

A biológiai nyom DNS-profil egyediségének kérdését alapvetően három tényező befolyásolja. Az egyik, hogy az élőlények egy fajba tartozó egyedei szaporodási közösségükben kisebb-nagyobb rokonságban állnak egymással, tehát genetikai állományuk egy része közös. A DNS-profil alapú individualizáció másik korlátozó tényezője az a körülmény, hogy a genetikai lokuszok polimorfizmusának mértéke populációról populációra és földrajzi területenként is változhat (Cavalli-Sforza és mt sai, 1994). Az individualizációt korlátozó harmadik tényező egy elsősorban kriminalisztikai jellegzetesség: a bűncselekmény elkövetésekor a biológiai nyom több személy vagy egyed biológiai anyagának egymáshoz kerüléséből, vagy keveredéséből is létrejöhet.

Mindezen megfontolások alapján az igazságügyi genetikus szakértők – eltérően pl. az ujjlenyomat-szakértők által alkalmazott gyakorlattól – jelenleg nem kategorikus, hanem valószínűsítő véleményt adnak meg a személyazonosítás vagy származás-megállapítás során.

Két biológiai minta közös eredetének megállapítása a mintákból nyert DNS-profilok összehasonlításából áll. Ha a két DNS-profil minden egyes lokuszon megegyező, vagy legalábbis átfedő mintázatot mutat, akkor valószínűsíthető a közös eredetük. A vérrokonság genetikai vizsgálata során azt elemezzük, hogy a két feltételezett rokon személy DNS-profilja között van-e alléltípus egyezés, és ha igen, akkor ez megfelel-e az öröklődés törvényeinek.

1.1.1.1. Az igazságügyi genetikai vizsgálatok statisztikai interpretációjának alapjai

A DNS-bizonyíték statisztikai kiértékelésének a gyakorlatban legáltalánosabban elfogadott módszere az ún. Bayes-elvű hipotézistesztesztelés (Evetts és Weir, 1998). A valószínűségszámítás Bayes-tételének esély formájú leírása szerint a nyomozó hatóság ill. a bíróság által az igazságügyi szakértőnek feltett alapkérdést: pl. *Mekkora a valószínűsége annak, hogy a helyszínen talált vérnyom XY-tól származott?* mindig legalább egy másik, ún. ellenhipotézis valószínűséghez kell viszonyítani: pl. *Mekkora a valószínűsége annak, hogy a vérnyom az adott népességből véletlenszerűen kiválasztott másik személytől származott?* Az első kérdés általában a vád feltevésére, a második kérdés pedig általában a védelem hipotézisére vonatkozik.

Az előzőekben említett ún. *posterior* valószínűség további két valószínűségi értékre bontható. A két valószínűségi érték közül az ún. likelihood tag vonatkozik a genetikai tesztesztelésre. A likelihood azt mutatja meg, hogy mekkora a DNS-vizsgálat eredményének valószínűsége a vád ill. a védelem feltevésének (hipotézis és ellenhipotézis) teljesülése esetén. Például a vád hipotézisére: *Mekkora a valószínűsége annak, hogy éppen a vérnyomból kimutatott DNS-profil határozzuk meg akkor, ha feltesszük, hogy a vérnyom XY-tól származott?* Fontos azt hangsúlyozni, hogy az ellenhipotézis likelihood értékének számítása csak akkor korrekt, ha a statisztikai kalkulációhoz felhasznált populációs allél- vagy profilgyakorisági adatbázis típusát nem a terhelt személy etnikai hovatartozása, hanem a nyomozati információ (pl. a bűncselekmény elkövetésének helye, a sértett ill. a szemtanúk vallomása) határozza meg.

A hipotézis és ellenhipotézis teljesülésének feltételezésével számított valószínűségek hányadosát, a valószínűségi hányadost (LR) adják meg a szakértők az igazságügyi DNS-vizsgálatok statisztikai kiértékelése során. A vérrokonság (pl. apaság) genetikai tesztesztelésében a valószínűségi hányadost rokonsági (apasági) indexnek is hívják. Amennyiben a biológiai nyom csak egy személy genetikai anyagát tartalmazza, akkor a valószínűségi hányados – legegyszerűbb formájában – a DNS-profil egyezés valószínűségének reciprokával egyezik meg. A DNS-profil egyezési valószínűség azt mutatja meg, hogy mekkora a várható valószínűsége annak, hogy a populációban a biológiai nyom DNS-profilját a gyanúsítottól kívül még egyszer megfigyeljük.

1.1.2. Az igazságügyi genetikai vizsgálatok rövid történeti áttekintése

A polimeráz láncreakció (PCR) felfedezésének évében, 1985-ben került sor a DNS-vizsgálatok első törvényszéki alkalmazására. Alec Jeffreys és munkacsoportja (1995a,b) a kromoszóma azon szakaszait vizsgálta, amelyek nagy számban tartalmaztak rövid ismétlődő szekvenciákat, mivel ezek az ún. miniszatelliták lényegesen polimorfabbnak bizonyultak az eddig vizsgált markerekhez képest. Az RFLP-t (Restriction Fragment Length Polymorphism) mutató miniszatelliták (más néven VNTR-ek: Variable Number of Tandem Repeats) alkalmazása a '80-as évek végére terjedt el a külföldi igazságügyi vizsgálatokban (Butler, 2001).

A PCR-sokszorozás igazságügyi alkalmazásáról először Higuchi és munkatársai (1988) számoltak be, amikor egyetlen hajszálból kiindulva a HLA-DQ α -lokusz polimorf régióját genotipizálták. 1990-ben megjelent az első, igazságügyi alkalmazásra hitelesített kereskedelmi PCR-kit.

A mikroszatelliták, SSR-ek (Simple Sequence Repeats), vagy elterjedtebb nevén STR-ek (Short Tandem Repeat) igazságügyben való alkalmazhatóságát a PCR-technika gyors elterjedése tette lehetővé. 1991-ben leírták az STR markerek fluoreszcens analízisét.

1992-ben elindult az igazságügyi DNS-vizsgálat Magyarországon.

1993-ban a kereskedelmi forgalomban elérhetővé vált az első STR kit, kifejlesztették a genetikai nem meghatározását az Amelogenin lokusz PCR-fragmens analízisével. 1994-ben megjelent az első multiplex PCR-kit.

1995-ben az O. J. Simpson-ügy ráterelte a nyilvánosság figyelmét az igazságügyi DNS-vizsgálatokra. Nagy-Britanniában felállították a bűnügyi STR profilnyilvántartást. 1995-96-ban jelentek meg az első olyan tudományos közlemények, amelyek magyar igazságügyi DNS-vizsgálati eredményekről, köztük STR analízisekről számoltak be (Woller J, 1995; Woller és mtsai, 1995; Füredi és mtsai, 1995; Csete és mtsai, 1996).

1997-ben az USA Szövetségi Nyomozó Hivatala (FBI) definiálta a 13 STR törzslokuszt. Ebben az évben mutatták be az igazságügyi gyakorlatban vizsgálható Y-kromoszómás STR-ek első marker-készletét (Kayser at al., 1997a,b).

1998-ban az USA szövetségi szinten felállította bűnügyi DNS-profil nyilvántartó rendszerét (Combined DNA Index System, CODIS). Az évtized végére az STR vizsgálatokon alapuló igazságügyi genetikai szakértés Magyarországon elterjedtebbé vált (Kozma és mtsai, 1998; Lászik és mtsai, 1999)

Az ezredfordulóra a kereskedelmi multiplex STR kitek igazságügyi alkalmazása széleskörben elterjedt. Felállították az európai online Y-STR haplotípus referencia adatbázist. 2000-ben Magyarországon hatályba lépett a bűnügyi nyilvántartások rendszerét, köztük a DNS-profil nyilvántartást szabályozó törvény.

1.1.3. A mikroszatellita (STR) lokuszok általános jellemzése

Az STR-ek olyan DNS-régiók, ahol 2-6 bázispárból álló szakaszok ismétlődnek tandem módon egymásután. Az STR lokuszok ismétlődéseinek száma a személyek között nagy változatosságot mutat, és ez a polimorf tulajdonság alkalmassá teszi őket genetikai személyazonosítási célokra (Butler, 2001). Mivel általában az STR lokusz különböző alléljai méretükben is különböznek egymástól, ezért az STR polimorfizmusokat a hossz-polimorfizmusok közé sorolják.

A polimorf mikroszatelliták ezreit azonosították a humán genomban. Átlagosan minden 10.000. nukleotidnál megfigyelhető egy STR (Edwards és mtsai,1991). Új STR markereket általában DNS-szekvencia adatbázisokban lefolytatott kereséssel (Weber és May, 1989) vagy molekuláris biológiai izolációs módszerek révén fedeznek fel (Edwards és mtsai, 1991).

Az STR repeat szekvenciákat repeat egységük hosszának megfelelően nevezik el: a di-, tri-, tetra-, penta-, hexa-repeat egységek 2, 3, 4, 5, 6 nukleotidból állnak. Elméletileg 4, 16, 64, 256, 1024, és 4096 lehetséges repeat-motívum lehet a mono-, di-, tri-, tetra-, penta- és hexanukleotid-repeatek esetében. Mivel azonban a mikroszatellitákban az ismétlődő egységek tandem módon helyezkednek el, ezért bizonyos motívumok egymásnak gyakorlatilag megfelelnek.

Az STR szekvenciák nem csak repeat számuk, hanem a repeatek struktúrájában is variabilitást mutatnak. Az egyszerű repeatek azonos hosszúságú és szekvenciájú egységekből állnak. Az összetett repeatek két vagy több különböző egyszerű, egymással szomszédos repeatet tartalmaznak. A komplex repeatekben a különböző, változó hosszúságú repeat-egységeket változó méretű közbeiktatott szekvencia szakítja meg. A komplex hipervariábilis repeatek szekvenciájukban és méretükben is eltérő, nem konszenzus repeatek sokaságát tartalmazzák.

Nem minden STR allél mérete felel meg a repeat-egységük egész számú többszörösének. Vannak olyan, ún. mikrovariáns allélok is, amelyek a repeat- vagy a

repeateket határoló régióban egy vagy több nukleotidot ért megváltozást hordoznak, s emiatt a mikrovariáns allél mérete a konszenzus allél méretével nem egyezik meg.

Az STR markerek igazságügyi célú alkalmazhatósága – magas polimorfizmus fokuk mellett – PCR-sokszorozhatóságukban rejlik. A PCR-rel olyan módszer került a bűnügyi nyomozók és szakértők kezébe, amelynek segítségével addig használhatatlannak és lényegtelennek tűnő biológiai anyagmaradványokat is meg lehetett „szólaltatni” az igazság kiderítése céljából. A PCR legnagyobb előnye, hogy a genom általunk vizsgálni kívánt részéből nagyon gyorsan (egy-két óra alatt) akár több millió másolatot is készíthetünk. Másik igen kedvező tulajdonsága, hogy egy reakció során egyidejűleg több DNS-lokuszt is felsokszorozható (ez az ún. multiplex PCR), valamint a felsokszorozandó DNS-szakasz lerövidítésével igen leromlott minőségű és töredezett DNS-minták is azonosíthatóvá válnak. A multiplex PCR e tulajdonságainak köszönhetően akár néhány pikogramnyi DNS, azaz egyetlen sejt genetikai állománya is elegendő lehet a személyazonosítás számára. A minden tekintetből megbízható és a nagyon szigorú minőségbiztosítási követelményeknek is eleget tevő rutinszintű sorozatvizsgálatokhoz jelenleg azonban ennél két-három nagyságrenddel több, kb. 1 nanogramm örökítő anyag (kb. 100-200 sejt DNS-tartalma) szükséges.

Az STR lokuszok, kis méretüknél (100-500 bp) fogva az összes ismert hossz-polimorfizmusnál megbízhatóbban sokszorozhatók fel. Mindemellett az STR allélok DNS-fragmensanalízissel az esetek döntő többségében egyértelműen elkülöníthetők egymástól a genetikai analízis során. Az STR lokuszok multiplex PCR-alapú szimultán vizsgálata lehetőséget teremtett a genetikai személyazonosítás gyors (akár néhány órán belüli) kivitelezésére (Kimpton és mtsai, 1996). Mindezek miatt az igazságügyi genetikusok az 1990-es évek folyamán a mikroszatellitákat választották a személyazonosítás fő genetikai markeréül.

Doktori értekezésem írása időpontjáig igazságügyi célokra közel száz polimorf mikroszatellitát jellemeztek és hitelesítettek (Ruitberg és mtsai, 2001; Butler és Reader, 2003). Ezen STR-ek esetében lokuszonként mintegy 6-30 allélváltozat figyelhető meg a humán népességben. A polimorfizmusok egy szűkebb köre, 16 marker a jelenlegi nemzeti bűnügyi DNS-profil nyilvántartások alapját képezi. A Nemzetközi Rendőrség (Interpol), az Európai Igazságügyi Szakértői Intézetek Hálózata (ENFSI) valamint az USA Szövetségi Nyomozó Iroda (FBI) által vizsgálatra ajánlott markerek is ebből a körből kerülnek ki.

1.1.4. Az igazságügyi célokból analizált STR lokuszok és allélok nevezéktana

Az igazságügyi genetikai vizsgálatokban analizált STR lokuszok és alléljaik nevezéktanát a Nemzetközi Igazságügyi Genetikai Társaság (ISFG, korábban ISFH), több ajánlásban is összegezte (Bär és mtsai, 1997; Olaisen és mtsai, 1998; Gill és mtsai, 2001). A legfontosabb szabályok a következők:

- A szekvenciát mindig 5'-3' irányban kell olvasni, és a tipizáláshoz használatos DNS szál az kell hogy legyen, amelyet az irodalomban eredetileg leírtak, illetve a nyilvános adatbázisba először bejegyeztek.
- Fontos, hogy a nevezéktan amennyire lehet, egységesítve legyen a D#S# rendszer szerint. Mivel számos nem standard megnevezés használatos világszerte, ezeknek a megváltoztatására nincs mód, de az új lokuszok felfedezőinek már a D#S# rendszerű elnevezés tanácsolt annak érdekében, hogy elősegítsék az egységesítés folyamatát. Ugyanígy az ezirányú publikációkat kiadó szakmai folyóiratok szerkesztőinek is nagy részük lehet a standardizálás előrehaladásában.

Az allélikus nevezéktanra vonatkozó ajánlások fő pontjai:

- Az elnevezés a teljes ismétlődési egységek (repeatek) száma szerint történik.
- Amennyiben a szekvencia egy nem teljes repeatet is tartalmaz [mikrovariánsok, köztes (inter-) allélok], úgy a teljes egységek száma után egy ponttal elválasztva kell feltüntetni a hiányos repeat meglévő bázisainak számát.
- A következetesség érdekében a nevezéktani ajánlások szerint az újonnan felfedezett STR-ek alléljait a repeat-régióban található variáns és nem-variáns ismétlődési egységek együttes számának alapján kell elnevezni.
- A nagyon bonyolult STR-ek esetében, amelyek többféle és eltérő méretű repeatekből állnak, az elnevezés meglehetősen nehézkes. Ilyenkor az allélok hosszúsága között kell kapcsolatot keresni a már elnevezetteket véve alapul. Alternatív nevezéktanok természetesen létezhetnek, de az új STR-eknél már itt is ajánlott az egységesített szabályok alkalmazása.
- Időnként megjelenhetnek köztes allélok az ún. flanking (határoló) régiók egyszerű inszerció vagy deléciós mutációjának köszönhetően is. Leírtak ilyen típusú mutációt pl. a DYS385 lokusznál (Füredi és mtsai, 1999), ahol az eredeti primereket használva *.1 és *.2 köztes allélok keletkeznek, amelyeket a fent

említett esetektől való megkülönböztetés végett + illetve - jellel ajánlatos jelölni. A + jel inszercióra, a - pedig delécióra utal. A flanking régiók mutációinak zavaró hatását kiküszöbölő megoldás a mutációt az amplikonon kívül hagyó alternatív primerek használatában rejlik, habár ezzel a populációban esetleg meglévő, egy vagy kevés nukleotidot érintő (ún. SNP: single nucleotide polymorphism) polimorfizmus detektálásától esünk el (Schneider és mtsai, 1998).

- A duplikálódott haploid lokuszok esetében, mint pl. az Y-kromoszómás DYS385 lokusznál, ahol a megfigyelt fragmensek természetesen nem tekinthetők egy lokusz termékeinek, az allélokot együtt kell kezelni, és a nevezéktan ajánlása értelmében az allélok nevét kötőjellel elválasztva kell megnevezni: pl. DYS385 11-14.

1.1.5. Az Y-kromoszómás STR markerek analízise az igazságügyi genetikában

Az 1990-es években az igazságügyi azonosítási technikák fejlődésében a mikroszatelliták skálájának bővülése és vizsgálatuk módszereinek tökéletesítése dominált leginkább (Butler, 2001). Noha számos új autoszómás STR került törvényszéki vizsgálatokban kipróbálásra, s több közülük alkalmasnak bizonyult a peres ügyekben való rutinszerű használatra is (Urquhart és mtsai, 1994; Hammond és mtsai, 1994), az Y kromoszómás STR rendszerek fejlődése sokáig nagyon lassú ütemben haladt.

Az első humán Y-kromoszómás STR-t (Y27H39, későbbi nevén DYS19) 1992-ben írták le. (Roewer és mtsai, 1992). A '90-es évek elején csak a DYS19 volt az egyetlen Y-STR marker, amelyet igazságügyi alkalmazásokban teszteltek (Roewer és Eplen, 1992). Az 1990-es évek második felében felgyorsult a fejlődés az Y-kromoszómás mikroszatelliták kutatásának terén (Santos és mtsai, 1996; Santos és Tyler-Smith, 1996; Kayser 1997a,b). Napjainkban az igazságügyi genetikai célokra alkalmazott Y-STR markerek száma megközelítette az igazságügyben rutinszerűen vizsgált autoszómás STR markerek számát (Corach és mtsai, 2001).

Az erőszakos bűncselekmények 95-99%-át férfiak követik el. Az Y-kromoszóma a humán genom egyetlen, kizárólag férfiakra jellemző szakasza. Az Y-kromoszómás STR-ek vizsgálatával a férfi tettes genetikai ujjlenyomata közvetlenül kiolvasható olyan kevert jellegű biológiai mintákból is, ahol a női eredetű sejtek túlsúlya miatt a hagyományos

autoszómás STR analízis nem kecsegtet a bizonyítás szempontjából informatív eredménnyel (Jobling és mtsai, 1997).

A szigorúan apai öröklődésű Y-kromoszóma kb. 95%-a nem rekombinálódik a meiózis során, a férfiak Y-kromoszómájában megfigyelt eltéréseket kizárólag mutációk okozzák (Quintana-Murci és mtsai, 2001). Ezen tulajdonságainál fogva az Y-kromoszómás STR-ek elemzése különösen hasznos lehet olyan törvényszéki alkalmazásokban, ahol a férfiak közötti rokonsági kapcsolat megléte a kérdés (Kayser és Sajantila, 2001).

Az autoszómás és Y-kromoszómás STR-ek igazságügyi genetikai analízisében egyéb eltérések is vannak. Autoszómás rendszert tekintve az egyetlen személy, aki elméletileg ugyanazzal a DNS-profillal rendelkezhet, mint a gyanúsított, az csak a gyanúsított egyetértő ikertestvére lehet. Az Y-kromoszómás rendszernél azonban a gyanúsított minden férfiágú rokona: fia, unokája, apja, nagypapa, fiútestvére, apai nagybátyja stb. is ugyanazokat az allélokat hordozza. Ez az információ a bíróságon sok esetben a terhelt vád alól való felmentésére adhat okot, még ha az Y-kromoszómás markerek alapján egyezés látszik is az elkövető mintájával. Ugyanakkor bizonyos esetekben ez a rokonok közti allél-egyezés előnyös is lehet. Apasági pereknél ugyanis amennyiben a feltételezett apa nincs meg – elhunyt, vagy ismeretlen helyen tartózkodik –, úgy a fent említett rokonai közül bárkinek a DNS-mintáján elvégezhető a vizsgálat, hisz az alléljaik – a ritka mutációs eseményt leszámítva – teljesen megegyeznek (Jobling és mtsai, 1997).

A rekombináció hiánya miatt a vizsgált Y-kromoszómás lokuszokon az egyes allélok nem véletlenszerű kombinációban határoznak meg egy adott DNS-profilt. A kapcsolt allélok specifikus kombinációját haplotípusnak nevezzük (Mathias és mtsai, 1994; Jobling és mtsai, 1997). Y-kromoszómás populációs és evolúciós felmérések esetén informatívabb a haplotípus-gyakoriságokkal számolni az autoszómás rendszereknél használatos allélgyakoriságokkal szemben (Spurdle és Jankins, 1992; Kayser és mtsai, 1997a,b).

1994-ben elindult egy nemzetközi program, amely az európai populációk Y-kromoszómás genetikai struktúrátságát hivatott felmérni. Ennek eredményeit az európai haplotípus populációs adatbázis tartalmazza, az ún. Eu-YHRD (European Y Haplotypes Reference Database: <http://ystr.charite.de>, Roewer és mtsai, 2001). Az adatbázis jelenleg az európai népességben előforduló Y-STR haplotípusok legnagyobb gyűjteménye. A 2003. augusztus 15-i állapot szerint az Eu-YHRD 84 populációs minta 12.904 férfianak 9-

lokuszos, ún. minimál haplotípusát tartalmazza. Hasonló adatbázisok azóta az ázsiai és az USA populációkra is felállítottak.

1.2. Célkitűzések

Mikroszatellita markerek vizsgálatán alapuló igazságügyi genetikai szakértés 1994-ig nem létezett Magyarországon. STR polimorfizmusok bevonásával a magyar népesség genetikai összetételét 1994-ig nem vizsgálták igazságügyi aspektusból. A magyar népességre vonatkozóan nem álltak rendelkezésre STR allél- és profilgyakorisági adatbázisok.

Munkám alapvető célja a humán STR polimorfizmusok genetikai analízisének a magyarországi kriminalisztikai gyakorlatba történő bevezetése, hitelesítése, valamint olyan magyar populációs referencia adatbázis létrehozása, amelynek segítségével az STR lokuszok vizsgálatán alapuló DNS-eredmények bizonyító ereje populációstatisztikailag megbecsülhető a magyarországi törvényszéki eljárások során.

1.2.1. Humán STR markerek vizsgálatának laboratóriumi bevezetése, populációs minták összeállítása

Az igazságügyi genetikai mikroszatellita analízist olyan STR polimorfizmusok vizsgálatával szándékoztam elkezdni és lefolytatni, amelyek igazságügyi célú alkalmazhatóságát korábban letesztelték (Lygo és mtsai, 1994; Micka és mtsai, 1996; Pestoni és mtsai, 1995). Így egyrészt megalapozottan remélhettem, hogy a magyar igazságügyi szakértői gyakorlatba történő bevetésük is sikeres lesz. Másrészt azzal kecsegtetett, hogy eredményeim összehasonlíthatók lesznek más munkacsoportok eredményeivel. Ez nemcsak tudományos, hanem szakértői szempontból is rendkívül fontos, mivel a szakértői laboratóriumok közötti kompatibilitás a szakértői munka hitelének egyik zálogát jelenti.

Célul tűztem ki, hogy az STR polimorfizmusok vizsgálatát autoszómás és nemi kromoszómákra nézve is elvégezzem, ismerve az öröklődésükből és rekombinációs képességükből fakadó különbségeket.

17 autoszómás, egy X- és kilenc Y-kromoszómás lokusz elemzését kívántam megvalósítani. A kiválasztás szempontjai között szerepeltek: a lokuszok polimorfizmusának foka, a lokuszok igazságügyi hitelesítettsége, és a genetikai analízis

kivitelezhetősége. Céljaim között szerepelt, hogy olyan lokuszokat válasszak ki, amelyek kromoszómás lokalizációjuk alapján gyakorlatilag nem kapcsoltak.

Az általam kiválasztott STR lokuszok a kromoszómák sorrendjében a következők voltak (zárójelben kromoszómás lokalizációjuk):

A, autoszómás lokuszok: *F13B* (1q31-q32.1), *TPOX* (2p23-2pter), *D3S1358* (3p21), *FGA* (4q28), *D5S818* (5q21-q31), *CSFIPO* (5q33.3-34), *F13A01* (6p24-p25), *D7S820* (7q), *LPL* (8p22), *D8S1179* (8q24.1-24.2), *TH01* (11p15-15.5), *VWA* (12p12-pter), *D13S317* (13q22-q31), *FES/FPS* (15q25-qter), *D16S539* (16q22-q24), *D18S51* (18q21.3), *D21S11* (21p11.1),

B, X-kapcsolt lokusz: *HPRTB* (Xq26),

C, Y-kapcsolt lokuszok: *DYS19*, *DYS389-I*, *DYS389-II*, *DYS90*, *DYS391*, *DYS392*, *DYS393*, *DYS385*

A *DYS392* lokusz kivételével mindegyik lokusz tetramer, a *DYS392* trimer. A lokuszok viszonylag nagy száma és a doktori értekezés terjedelmi korlátai miatt a lokuszok részletes jellemzésétől eltekintek. A lokuszok igazságügyi genetikai szempontú, részletes leírását az STRBase honlap tartalmazza (Butler és Reader, 2003).

Az STR markerek igazságügyi hitelesítésének részét képezik a populációs felmérések. Allél- és profilgyakorisági adatbázisok nélkül az STR vizsgálatok törvényszéki bizonyító erejének statisztikai interpretációja nem lehetséges. A nemzetközi ajánlások szerint globális vizsgálatok mellett szükség van regionális felmérésekre is. A felmérésekbe bevont populációs mintáknak lehetőség szerint minél jobban reprezentálnia kell a népesség genetikai összetételét, különös tekintettel azokra a populációkra, akik szeparáltabb közösségeket alkotnak (NRC, 1996; Gill és mtsai, 2001).

Céлом egy olyan magyar populációs referencia adatbázis létrehozása volt, amelynek segítségével az STR lokuszok vizsgálatán alapuló DNS-eredmények bizonyító ereje populációstatisztikailag megbecsülhető a magyarországi törvényszéki eljárások során. Czeizel és munkatársai (1991) nyomán a referencia populációs mintát a budapesti kevert (heterogén) populáció véletlenszerűen kiválasztott személyeiből kívántam összeállítani.

Korábbi tanulmányok (Czeizel és mtsai, 1991; Guglielmino és Béres, 1996) felhívták a figyelmünket arra, hogy több magyarországi népcsoport genetikailag szeparált. Felvetődött tehát az a kérdés, hogy egy felállítandó budapesti referencia populációs adatbázis milyen feltételekkel és korlátokkal alkalmazható a magyarországi népcsoportokra. Természetesen a legnagyobb eltéréseket olyan magyarországi populációk

esetében vártam, akik a korábbi populációgenetikai vizsgálatok alapján a legnagyobb genetikai távolságot mutatták a többi magyar népcsoporttól.

A populációs minták gyűjtésében Intézetünk nem vett részt. A mintagyűjtés a mintaadó személyek beleegyezésével történt, az ORFK Bűnügyi Szakértői és Kutatóintézete a mintákat név és egyéb személyazonosító adat nélkül, csak a nem feltüntetésével kapta meg az értekezésem 5. *Köszönetnyilvánítás* c. fejezetében felsorolt, kollaboráló munkacsoportoktól.

A populációs minták a következők voltak (a mintagyűjtés időrendi sorrendjében):

1. 1993-ban gyűjtött budapesti, általános jellegűnek tekinthető (etnikai hovatartozás alapján nem szelektált) populációs minta (223 személy, jelölése BuCa). Ezt a budapesti mintát referencia populációs mintánknak választottuk,
2. 1996-97-ben gyűjtött Baranya megyei roma minta (216 személy, jelölése BaRo). A minták gyűjtése és feldolgozása két szakaszban történt. Az első szakaszban 135 személy mintáját gyűjtötték össze, majd ezt követően még 81 minta begyűjtésére és analizésére került sor,
3. 2000-ben gyűjtött kelet-magyarországi (Debrecen régió) roma minta (110 személy, jelölése DeRo),
4. 2000-ben gyűjtött budapesti askenázi zsidó minta (178 személy, jelölése BuAs),
5. 2000-ben gyűjtött kiskunsági (fülöpszállási) minta (37 személy, jelölése Kisk).

Célkitűzésem volt a vizsgálni kívánt STR markerekre vonatkozóan allél- és profilgyakorisági táblázatok létrehozása a felmérésembe bevont populációs mintákra.

Korábbi vizsgálatok alapján, amelyek a klasszikus vércsoportokra és fémérjemarkerekre irányultak, a hazai roma és askenázi zsidó populáció mutatta a legnagyobb genetikai távolságot a többi magyar népcsoporttól (Czeizel és mtsai, 1991; Guglielmino és Béres, 1996). A roma és zsidó populációs minták STR lokuszok analizésén alapuló vizsgálatával egyrészt ezt a kérdést a DNS-markerek szintjén is elemezni szerettem volna. Másrészt, ha az egyes populációs allél- és profilgyakorisági adatbázisok között az STR lokuszok esetében is szignifikáns különbségeket detektálunk, akkor meg kívántam állapítani azt, hogy ez a különbség lényeges hatással van-e az igazságügyi DNS-vizsgálatok statisztikai interpretációjára. A kiskun (fülöpszállási) populációs minta felmérésünkbe való bevonása lehetőséget teremtett arra, hogy egy olyan magyar etnikum populációgenetikai felmérését is elvégezhessük, akik a korábbi vizsgálatok alapján a

budapesti kevert népességtől a roma és a zsidó populációhoz képest kisebb genetikai távolságot mutattak (Czeizel és mtsai, 1991; Guglielmino és Béres, 1996).

1.2.2. Az alkalmazott automatikus genotipizáló rendszerek allélméret-meghatározó pontosságának hitelesítése

Az igazságügyi genetikai vizsgáló-rendszerek eredményeinek megbízhatóságát a vizsgálatokat kivitelező laboratóriumoknak biztosítaniuk kell (NRC, 1996). E hitelesítés keretében, le kívántam tesztelni a genetikai analízisünk döntő többségében alkalmazott automatizált elektroforetikus rendszerek (ALF és ABI Prism 310 szekvenátorok) allélméret-meghatározási pontosságát és precizitását. Céljaim közé tartozott a genotipizálási rendszerünk pontosságának növelése is.

1.2.3. A populációs felmérések során detektált ritka, mikrovariáns allélok jellemzése

Az igazságügyi genetikai vizsgálatok alapkövetelménye a genotípus pontos leírása. Az STR markerek esetében ez referencia-allélok (allél-létra) segítségével történik. Ha az ismeretlen allél nem felel meg egyik referencia-allélnak sem, a tipizálandó allélt – lehetőleg szekvenálással – jellemezni kell (NRC, 1996; Bar és mtsai, 1997)

Célul tűztem ki, hogy a populációs felméréseim során detektált ritkának számító vagy az észlelés időpontjában éppenséggel még le nem írt mikrovariáns allélok szekvencia-struktúráját jellemzem.

1.2.4. A vizsgált STR polimorfizmusok géndiverzitásának és személyazonosító hatékonyságának megállapítása a magyar népességben

Az allél- és STR-profil gyakorisági adatbázisokat tesztelni kell arra nézve, hogy felhasználhatók-e a DNS-vizsgálat bizonyító erejének korrekt megbecsüléséhez. Az STR profilegyezési valószínűség kalkulációjához számos paramétert és körülményt meg kell ismerni, amelyek közül a legfontosabbak: a vizsgált STR lokuszok polimorfizmusának

foka, az allélok lokuszokon belüli és lokuszok közötti lehetséges asszociációja, ill. a teljes genetikai variancia megoszlása személyeken belül, a személyek között a populáción belül és a populációk között (Evett és Weir, 1998).

Populációgenetikai elemzéseim első lépéseként célul tűztem ki a vizsgálataimba bevont STR markerek polimorfizmus-fokának meghatározását a magyarországi populációs minták esetében. Meg kívántam állapítani, hogy a markereken mért géndiverzitás kellően magas-e a hatékony személyazonosításra nézve Magyarországon.

1.2.5. A vizsgált populációk Hardy–Weinberg és linkage egyensúlyi tesztelése a vizsgált autoszómás STR lokuszokon

Az igazságügyi DNS-vizsgálatok statisztikai interpretációjának fő eleme a DNS-profil egyezési valószínűség meghatározása. Az egyezési valószínűség kalkulációjának alapmodelljét az szolgáltatja, hogy feltételezzük hogy a DNS-profilokat alkotó allélok között nincs számottevő asszociáció (Evett és Weir, 1998). Az elvégzendő Hardy–Weinberg és linkage egyensúlyi teszteléssel populációnként választ kívántam adni arra a kérdésre, hogy van-e jelentős mértékű allélikus asszociáció a vizsgált autoszómás STR-ek alléljai között a lokuszokon belül és a lokuszok között.

1.2.6. A felmérésbe bevont magyarországi populációk genetikai strukturáltságának megállapítása

Az igazságügyi genetikai gyakorlatban is széleskörűen alkalmazott populációstatisztikai tesztek [G-statisztika, F-statisztika, molekuláris varianciaanalízis (AMOVA)] segítségével, jellemezni kívántam a felmért magyarországi populációk genetikai strukturáltságát.

Tisztázni kívántam, hogy a referenciaként deklarált budapesti populáció és a többi magyarországi populáció között mekkora a genetikai variancia az analizált STR lokuszon.

Céлом volt annak megállapítása is, hogy a magyarországi populációk és más, külföldi populációk között mekkora a genetikai variabilitás.

1.2.7. A vizsgált magyarországi populációk genetikai variációjának hatásai az igazságügyi statisztikai interpretációra

Korábbi irodalmi adatok alapján (Czeizel és mtsai, 1991, Guglielmino és Béres, 1996) feltételeztem, hogy a magyar népesség szubstrukturáltsági foka igazságügyi genetikai aspektusból is jelentős lehet. Ennek alapján elemezni kívántam a vizsgált populációk autoszómás STR lokuszokon mért genetikai variációjának az igazságügyi statisztikai interpretációra gyakorolt hatását.

1.2.8. A biológiai nyomot hátrahagyó személy etnikumának valószínűsítési lehetőségei: az etnikai eredet becslése (EEO)

A populációs elemzéseinkhez kiválasztott öt magyarországi populáció STR allél- és profilgyakorisági eloszlásában, a populáció-párok esetében statisztikailag szignifikáns különbségeket mutattunk ki. A budapesti referencia populációtól legnagyobb mértékben a baranyai roma populáció különbözött.

2001-ben Michael Klitschar, a Grazi (jelenleg Hallei) Egyetem igazságügyi genetikusa megkeresett, hogy csatlakozzam ahhoz a kutatási projektjéhez, amely a biológiai nyomot hátrahagyó személy etnikai eredetének megbecsülésére (EEO) irányult. A dr. Klitscharral történt konzultációk alapján úgy gondoltuk, hogy az EEO megbízhatóságának tesztelését érdemes a budapesti és baranyai populáció-párra is kiterjeszteni. A közös projekt fő célkitűzése az volt, hogy az egyes populációkat alkotó személyek etnikai származását autoszómás STR profiljuk alapján milyen sikerességgel lehet helyesen megbecsülni.

2. ANYAGOK ÉS MÓDSZEREK

Munkánk során autoszómás, X- és Y-kromoszómás mikroszatellita (STR) lokuszokon végeztünk felmérést öt magyarországi populációban. A vizsgálat lépései röviden a következők voltak:

- A DNS kinyerése a biológiai mintákból: folyékony vérből, gézen szárított vér- ill. nyálfoltból
- A DNS-koncentráció meghatározása
- Az STR lokuszok PCR-amplifikációja
- A PCR-amplifikálás hatékonyságának ellenőrzése elektroforézissel
- STR genotipizálás DNS-fragmensanalízissel
- A félautomatikus/automatikus allélméret-meghatározás pontosságának mérése
- A ritka és mikrovariáns STR allélok szekvencia analízise
- Populáció- és statisztikai genetikai analízisek

2.1. A DNS kinyerése a biológiai mintákból

Nem hemolizált folyékony vérré a Kanter és munkatársai (1986) által kidolgozott protokollt alkalmaztuk, amely a limfociták sejtmagjainak szukróz ozmózisos sokkal történő feldúsításán alapul.

Hemolizált folyékony vérré, vér- és nyálfoltra a Comey és munkatársai (1994) által kidolgozott protokollt alkalmaztuk, amelynek fő lépései: proteináz K sejtlyízis, fenol-kloroform-izoamilalkohol szerves extrakció, Centricon-Amicon (Millipore) ultraszűrés, puffercsere és DNS-koncentrálás.

2.2. A DNS-koncentráció meghatározása

Az igazságügyi genetikai vizsgálatokban PCR-amplifikálás körülményeit úgy optimalizálják, hogy kis mennyiségű (0,5-2 ng) DNS-templát elegendő legyen a sikeres genotipizáláshoz. Azonban ez az amplifikálási rendszer – különösen több lokuszos (multiplex) vizsgálat esetén – meglehetősen érzékeny a túl sok DNS-templát bevitelére.

Az optimális DNS-mennyiség alkalmazása érdekében célszerű a DNS-extraktum koncentrációját meghatározni.

A minták DNS-tartalmának szemikvantitatív meghatározására agaróz gélelektroforézist és etídium-bromid/SYBR Gold fluorofór festést alkalmaztunk standard protokoll szerint.

A minták DNS-tartalmának kvantitatív meghatározására a Quantiblot Human DNA Quantification Kit-et (Applied Biosystems) alkalmaztuk. A szondázott humán DNS-t kolorimetriás vagy chemilumineszcenciás módszerrel tesszük láthatóvá.

2.3. Az STR lokuszok PCR-sokszorozása

A VWA és TH01 lokuszok egy lokuszos (szimplex) PCR-sokszorozását Wiegand és munkatársai (1993) valamint Möller és munkatársai (1994) módszere alapján végeztük.

A VWA, TH01, CSF1PO, FES, TPOX, HPRTB lokuszok három lokuszos (triplex) PCR-sokszorozását a GenePrint STR Multiplex System CSF1PO-TPOX-TH01/HPRTB-FESFPS-vWF kitekkel (Promega) végeztük.

A VWA, TH01, TPOX, CSF1PO, LPL, F13B, FES, F13A01 lokuszok négy lokuszos (kvadruplex) PCR-sokszorozását a GenePrint Fluorescent Quadruplex STR System CSF1PO-TPOX-TH01-vWA/ F13A01-FESFPS-F13B-LPL kitekkel (Promega) végeztük.

A CODIS STR lokuszok és az Amelogenin (genetikai nem meghatározását szolgáló) lokusz multiplex PCR-sokszorozását az AmpF/STR Profiler Plus/Cofiler kitekkel (Applied Biosystems) végeztük.

A DYS19, DYS389-I/II (Triplex I) és a DYS390, DYS391, DYS393 (Triplex II) Y-STR lokuszok három lokuszos (triplex) PCR-sokszorozását Kayser és munkatársai (1997) módszere alapján végeztük. A DYS392 és DYS385 lokuszok két lokuszos (duplex) PCR-sokszorozását az általunk kidolgozott módszerrel végeztük (Füredi és mtsai, 1999).

2.4. A PCR hatékonyságának ellenőrzése

Az STR lokuszok genotipizálása – hasonlóan a PCR-amplifikáláshoz – érzékeny a bevitt DNS mennyiségére. A további analízisek előtt célszerű a PCR kitermelését ellenőrizni,

hogy megbecsülhessük azt, hogy a genotipizálás, szekvenálás stb. optimális kivitelezéséhez a PCR-ezett mintából mekkora mennyiséget kell használni a PCR-termék ideális mennyiségének eléréséhez. A PCR-amplifikálás hatékonyságának ellenőrzése horizontális, natív (nem denaturáló), nagy felbontású, diszkontinuus poliakrilamid gélelektroforézissel (nhd-PAGE) és ezüstoffestéses vizualizációval történt (Budowle és mtsai, 1991).

2.5. STR genotipizálás DNS-fragmensanalízissel

Az STR lokuszok PCR-termékeinek fragmens-analízise alapvetően három részből áll. Először a PCR során felszaporított DNS-szakaszok méret szerinti elválasztására kerül sor elektroforézis alkalmazásával (elektroforetikus elválasztás). Mivel az STR lokuszok a hossz-polimorf markerekhez tartoznak, így a futtatás során a DNS-szakaszok méretére kapott adat az allél típusára vonatkozó információnak felel meg. Ennek meghatározásához azonban az elválasztott PCR-termék detektálására van szükség, ez a genotipizálás második része. A harmadik rész a kapott eredmények kiértékeléséből, azaz az egyes mintákhoz tartozó allélok típusának meghatározásából (genotipizálás) áll. Az alléltípus meghatározáshoz referencia-allélokból összeállított „koktél”, ún. allél-létrát használunk.

2.5.1. Egy lokuszos (szimplex) genotipizálás a VWA és TH01 lokuszokon

A genotipizálást Wiegand és munkatársai (1993) valamint Möller és munkatársai (1994) módszere alapján végeztük:

- Elektroforézis: nhd-PAGE
- Detektálás: ezüstoffestés
- Genotipizálás: az ismeretlen minta vizuális összehasonlítása a szomszédos sávban elválasztott allél-létrával („manuális” tipizálás). A tipizáláshoz szükséges allél-létrákat ismert genotípusú személyek DNS-mintájának ko-amplifikálásával valamint a PCR-termékek re-amplifikálásával hoztuk létre.

2.5.2. Három lokuszos (triplex) genotipizálás a VWA, FES, HPRTB, TH01, TPOX, CSF1PO lokuszokon

A genotipizálást Martin és munkatársai (1995) módszere alapján végeztük:

- Elektroforézis: vertikális, denaturáló, kontinuus PAGE
- Detektálás: ezüsfestés
- Genotipizálás: az ismeretlen minta vizuális összehasonlítása a szomszédos sávban elválasztott allél-létrával („manuális” tipizálás). A tipizáláshoz szükséges allél-létrákat a PCR kitek biztosították.

2.5.3. Négy lokuszos (kvadruplex) genotipizálás VWA, TH01, TPOX, CSF1PO, LPL, F13B, FES, F13A01 lokuszokon

A genotipizálást Moscetti és munkatársai módszere (1995) és az általunk kidolgozott eljárás alapján végeztük (Füredi és mtsai, 1997):

- Elektroforézis: vertikális, denaturáló, kontinuus PAGE ALF DNS szekvenátor készüléken (Amersham Pharmacia)
- Detektálás: fluoreszcens primerek, lézer gerjesztett fluoreszcencia, 1 fluorofór: FITC
- Genotipizálás: félautomatikus/automatikus módon ALF Fragment Manager v.1.2/Fragment Analyzer v.1.02 szoftverek (Amersham Pharmacia) segítségével. A szoftver az STR allélok méretét a gél bizonyos sávjaiban megfuttatott ún. külső méretstandardok segítségével határozza meg (allele sizing). A külső méretstandardok egyben allél-létrák is. A gél-sávok között esetlegesen létrejövő DNS-mozgékonyaság különbség („mobility shift”) kompenzálására minden mintához ún. belső méretstandardot (2-2 ismert nagyságú DNS-fragment) is hozzákevertünk. A Fragment Analyzer szoftver az ismert méretű STR allélokhoz a referencia allél-létra segítségével típusot rendel (allele calling). A tipizáláshoz szükséges allél-létrákat a PCR kitek biztosították.

2.5.3.1. A genotipizálás allélméret-meghatározás pontosságának (allele sizing accuracy, precision) mérése

Az ALF-rendszer számítógépes allélméret-meghatározás pontosságát Decorte és Cassiman (1996) módszere alapján mind az allél-létrákra, mind a mintákra egyaránt leteszteltük. Az egyes allélokra vonatkozó százalékos pontosság értékét a következő formulával számoltuk:

$$\text{pontosság(\%)} = (1 - |\text{számított méret} - \text{várt méret}| / \text{várt méret}) \times 100\%$$

A pontossági értéken kívül, ugyancsak az egyes allélokra vonatkozóan meghatároztuk a számított allélméret szórását (SD) és az allélikus méret-tartományt (size range, window). Ez utóbbi adatok a számítógépes allél-méret meghatározás precizitását (precision) fejezik ki (Lazaruk és mtsai, 1998). A precizitási érték formulája:

$$\text{precizitás(\%)} = (1 - \text{SD} / \text{várt méret}) \times 100\%$$

2.5.4. Több lokuszos (multiplex) genotipizálás a CODIS STR és Amelogenin lokuszokon

A genotipizálást az AmpF/STR Profiler Plus/Cofiler kitek (Applied Biosystems) felhasználói kézikönyve szerint végeztük azzal a különbséggel, hogy belső méretstandardként nem GeneScan-500 ROX-ot (Applied Biosystems), hanem CXR60-400-at (Promega) használtunk:

- Elektroforézis: denaturáló kapilláris gélelektroforézis ABI Prism 310 genetikai analizáló készüléken (Applied Biosystems)
- Detektálás: fluoreszcens primerek, lézer gerjesztett fluoreszcencia, 4 fluorofór: FAM, JOE, NED, ROX
- Genotipizálás: félautomatikus/automatikus módon GeneScan Analysis és Genotyper szoftverek (Applied Biosystems) segítségével. A GeneScan szoftver az STR allélok méretét a mintákhoz (és az allél-létrához) kevert ún. belső méretstandardok segítségével határozza meg (allele sizing). Ugyancsak a belső méretstandard szolgál az egyes injektálások között esetlegesen létrejövő DNS-mozgékonyosság különbség („mobility shift”) kompenzálására is. A Genotyper szoftver a referencia allél-létra és a minta méretének összehasonlítása útján tipizálja a PCR-termékeket (allele calling). A tipizáláshoz szükséges allél-létrákat a PCR kitek biztosították.

2.5.5. Három lokuszos (Triplex I, Triplex II) és két lokuszos (Duplex) genotipizálás (haplotipizálás) az Y-STR lokuszokon

A haplotipizálás alkalmazott protokollját Füredi és munkatársai (1999) tanulmánya foglalja össze:

- Elektroforézis: ua. mint a 3.5.4. fejezetben
- Detektálás: fluoreszcens primerek, lézer gerjesztett fluoreszcencia, 2 ill. 3 fluorofór: FAM, ROX (Triplex I/II); FAM, TET, TAMRA (Duplex)
- Genotipizálás: ua. mint a 3.5.4. fejezetben azzal a különbséggel, hogy a tipizáláshoz szükséges allél-létrákat ismert haplotípusú személyek DNS-mintájának ko-amplifikálásával valamint a PCR-termékek re-amplifikálásával hoztuk létre. A minták haplotipizálása során a Nemzetközi Igazságügyi Genetikai Társaság (ISFG) DNS Bizottságának valamint az Eu-YHRD munkacsoport ajánlásait követtük (Kayser és mtsai, 1997; Roewer és mtsai, 2001; Corach és mtsai, 2001; Gill és mtsai, 2001).

2.5.5.1. A haplotipizálás allélméret-meghatározás precizitásának (allele sizing precision) mérése

Az ABI-rendszer számítógépes allélméret-meghatározás precizitását hat lokusz esetében (DYS19, DYS389-I/II, DYS390, DYS391, DYS393) az egyes allélokra vonatkozóan leteszteltük. A populációs mintákon, a CXR60-400 (Promega) belső méretstandard alkalmazásával az egyes allélokra vonatkozóan meghatároztuk a számított allélméret szórását (SD) és az allélikus méret-tartományt (size range, window). Az allél-létrák esetében ezt az elemzést a GeneScan-500 belső méretstandard alkalmazásával is elvégeztük (Füredi és mtsai, 1999). Az allél-létrák esetében az adatokat 50-50 egymást követő injektálásból nyertük. Az allél-létrák a következő allélokat tartalmazták: DYS19: 13-17, DYS389-I: 12--14, DYS389-II: 25-29, DYS390: 18-26, DYS391: 9-12, DYS393: 11-15.

2.6. A ritka, variáns STR allélok szekvencia analízise

Az alkalmazott nagy felbontású genotipizáló rendszereknek köszönhetően felméréseink során számos ritka, variáns STR allélt figyeltünk meg, amelyek nem voltak megtalálhatók az alkalmazott referencia allél-létrákban. E variáns allélok PCR-amplifikációját, nhd-PAGE gélelektroforézisét, gélből történő tisztítását és re-amplifikációját Intézetünkben végeztük standard protokoll szerint. A tisztított PCR-termékek szekvenálását két másik munkacsoport hajtotta végre:

- Pályi Anikó munkacsoportja (MTA Szegedi Biológiai Központ Szekvenáló Laboratórium) végezte a DYS385 lokusz 17.-1-es, 17.2-es és 18.-1-es alléljainak ciklusszekvenálását ABI Prism Dye Terminator Cycle Sequencing FS AmpliTaq Kit (Applied Biosystems) alkalmazásával (Füredi és mtsai, 1999)
- Antoon Vandenberghe munkacsoportja (Lyoni Claude Bernard Egyetem Humán Molekuláris Genetikai Laboratórium) végezte a HPRTB lokusz 12.-2-es és 12.-6-es alléljainak szekvenálását (Vandenberghe és mtsai, 2000); a D7S820 lokusz 8.-1-es, 9.+1-es és 10-es alléljainak (Egyed és mtsai, 2000) valamint a D13S317 lokusz 7-es és 12-es alléljainak ciklusszekvenálását ABI Prism d-Rhodamine Dye Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems) alkalmazásával (Boutrand és mtsai, 2001)

2.7. Populáció- és statisztikai genetikai analízisek

2.6.1. Populáción belüli elemzések

2.6.1.1. Populációgenetikai alapértékek: Nei-féle géndiverzitás (h), a polimorfizmus információ tartalom (PIC), az igazságügyi megkülönböztető erő (power of discrimination, PD), az egyezési valószínűség (probability of match, pM), szülősi kizáró erő (power of exclusion, PE), minimális allélgyakoriság (p_{min})

A genetikai polimorfizmusok populációgenetikai jellemzéséhez az igazságügyi genetika számos egyszerű mérőszámot használ. Az alábbiakban ismertetem az értekezésemben használt jellemzők formuláit.

A következő képlet segítségével mindegyik populációs mintán meghatároztuk a Nei-féle géndiverzitási értéket (h) az autoszómás lokuszokra, az X-kapcsolt HPRTB lokusz esetén a nőkre [ezekben az esetben h megegyezik a heterozigóták elfogulatlanul (unbiased) becsült arányával (H_{exp})] valamint a 9 Y-STR lokuszos minimál haplotípusokra vonatkoztatva:

$$h = n(1 - \sum x_i^2)/(n - 1), \quad \text{több lokuszra kombinálva} \quad h_{komb} = 1 - \prod (1 - h_i),$$

ahol n a kromoszómák száma, x_i az i -edik allél vagy haplotípus megfigyelt gyakorisága (Nei, 1973).

A Nei-féle géndiverzitás standard hibája $SE = \sqrt{2h(1-h)/n}$

A Nei-féle géndiverzitáshoz hasonló érték az igazságügyi megkülönböztető erő (power of discrimination, PD), amely az igazságügyi genetikában azt fejezi ki, hogy két személyt a népességből véletlenszerűen kiválasztva átlagosan mekkora a valószínűsége annak, hogy a két személy a vizsgált lokuszon eltérő genotípust hordoz. Az egyezési valószínűség (probability of match, pM) az 1-ből kivont PD értékkel egyezik meg.

$$PD = 1 - \sum x_i^2, \quad \text{több lokuszra kombinálva} \quad PD_{komb} = 1 - \prod (1 - PD_i),$$

$$pM = 1 - PD, \quad \text{több lokuszra kombinálva} \quad pM_{komb} = \prod pM_i,$$

ahol x_i az i -edik genotípus vagy haplotípus megfigyelt gyakorisága (Jones, 1972).

A p_M számítási módszere konkrét genotípus esetén a Bayes-elvű hipotézistesztelés igazságügyi statisztikai genetikában történő alkalmazásával jelentősen finomodott és bonyolultabbá vált (Balding és Nichols, 1994).

A *szülősegi kizáró erő* (*power of exclusion, PE; mean exclusion chance, MEC*) az igazságügyi genetikában azt fejezi ki, hogy egy személyt véletlenszerűen kiválasztva a népességből átlagosan mekkora a valószínűsége annak, hogy a személy egy szülő-gyermek relációban, mint másik lehetséges szülő kizárható:

$$PE = 2\left(\sum p_i^2\right)^2 - \sum p_i^4, \text{ több lokuszra kombinálva } PE_{\text{komb}} = 1 - \prod (1 - PE_i),$$

ahol p_i az i -edik allél megfigyelt gyakorisága (Ohno és mtsai, 1982).

A PD és PE értékeket az Y-STR markereken kívül szintén minden lokuszra kiszámítottuk. A PD és a PE az Y-STR lokuszok esetében megegyezik egymással; helyettük a Nei-féle géndiverzitást határoztuk meg.

A *polimorfizmus információs tartalmat (PIC)* (Botstein és mtsai, 1980) az Y-STR markereken kívül szintén minden lokuszra kiszámítottuk:

$$PIC = 1 - \sum p_i^2 - \sum_{i=1}^n \sum_{j=i+1}^n 2p_i^2 p_j^2$$

Az igazságügyi statisztikai számítások során a *minimális allélgyakorisági érték* (p_{\min}) használata lehetővé teszi, hogy populációs adatbázisokban a ritkán előforduló alléltípusok hiányos megfigyeléseit ellensúlyozhassuk (Budowle és mtsai, 1996). A minimális allélgyakoriság azt a küszöbértéket mutatja meg, amely fölött a megfigyelt allélok gyakorisága bizonyos (általában 95%-os) konfidenciával megbízhatónak tekinthető. A meg nem figyelt allélok gyakoriságát úgy lehet tekinteni, hogy az a p_{\min} -nel egyenlő. Az STR lokuszokra ajánlott minimális allélgyakoriság statisztikai és populációgenetikai elveken alapul, és a lokusz heterozigotásának növekedésével értéke nő, a populációs minta méretének növekedésével pedig csökken (Chakraborty, 1992):

$$p_{\min} = 1 - \left(1 - CF^{\frac{1}{c}}\right)^{\frac{1}{2N}},$$

ahol CF a választott konfidencia-szint, c a gyakran előforduló alléltípusok száma, N pedig a személyek száma.

A gyakran előforduló alléltípusok számát (c) a megfigyelt heterozigotásából (H_{obs}) számíthatjuk ki a következők szerint:

$$T = H_{obs} / (1 - H_{obs}), K = \sum_{i=0}^{2N-1} \frac{T}{T+i}, A = \sum_{i=1}^{2Nr} \frac{1}{i}, B = \sum_{i=1}^{2Nr} \frac{1}{2N-i}, P = AT - BT^2$$

$$c = K - P,$$

ahol r a ritka allél meghatározott gyakoriságát jelöli (általában $r \leq 0,01$).

2.6.1.2. A Hardy-Weinberg egyensúly (HWE) tesztelése

Az allélok lokuszon belüli függetlenségének vizsgálatát Hardy-Weinberg tesztelésnek hívjuk. A Hardy-Weinberg tesztelésnek számos változatát kidolgozták (Weir 1996), amelyek közül a legismertebbek: *a*, Nei-féle homozigotizációs teszt (Nei és Roychoudhury, 1974); *b*, hagyományos „goodness-of-fit” χ^2 -statisztika; *c*, valószínűségi hányados (likelihood ratio) teszt; *d*, exact teszt; *e*, log-lineáris modellek. Az igazságügyi genetikai szakirodalom útmutatásai és az általunk elérhető statisztikai szoftverek alapján a HWE tesztelésére az *a*, *b*, és *d* módszert választottuk. HWE tesztelésünk alapmódszerét a Fisher-féle exact tesztnek Guo és Thompson (1992) által továbbfejlesztett permutációs változata jelentette, amelyet Schneider, Roessli és Excoffier (2000) tett hatékonyabbá és implementált *Arlequin* nevű populációgenetikai szoftverükbe. Az *Arlequin* segítségével az exact tesztet az Y-STR lokuszok kivételével minden markeren elvégeztük. Az exact teszt alkalmazott paraméterei a következők voltak: Markov-lánc lépés-száma: 1.000.000, dememorizáció lépés-száma: 1000. A Nei-féle homozigotizációs tesztet saját készítésű Excel-táblázat segítségével szintén minden nem Y-kapcsolt STR markeren végrehajtottuk.

2.6.1.3. A linkage disequilibrium (LD) tesztelése

Az allélok lokuszok közötti függetlenségének vizsgálatát linkage disequilibrium (LD) tesztelésnek hívjuk. Az allélok lokuszok közötti lehetséges asszociációját az igazságügyi genetikai szakirodalom útmutatásai és az általunk elérhető statisztikai szoftverek alapján három módszerrel vizsgáltuk. Populációs elemzéseink kezdeti szakaszában Karlin és munkatársai (1981) által kidolgozott „inter-class” korrelációs kritériumot (ICC), valamint a Chakraborty-féle (1984) s_k^2 kritériumot alkalmaztuk. Az ICC két lokusz közötti, míg az s_k^2 kritérium az összes lokusz közötti asszociációt teszteli. Az s_k^2 kritérium azt vizsgálja, hogy a populációs mintában a függetlenség feltételezése esetén a heterozigóta lokuszok

számának megfigyelt varianciája beleesik–e a konfidencia tartományába. Az előbbieken említett két tesztet dr. Ranajit Chakraborty (Texasi Egyetem Humán-genetikai Központ) *DNATYPE* szoftvere segítségével végeztük el, amelyet nagylelkűen rendelkezésünkre bocsátott.

A két-locuszos LD tesztelésünk alapszoftverül az előző fejezetben említett *Arlequin* programot választottuk. Segítségével minden autoszómás lokuszpárban végrehajtottuk az analízist. Az *Arlequin* szoftver permutációs valószínűségi hányados (likelihood ratio) tesztet alkalmaz olyan genotípusos adatokra, ahol a gaméta-fázis nem ismert. A függetlenségi nullhipotézis teljesülése esetén az adatok valószínűségének kiszámításakor a haplotípus-gyakoriságot az allélgyakoriságok egyszerű szorzataként kapja meg. A függetlenségi nullhipotézis elvetése esetén a program az adatok valószínűségének kiszámításakor a haplotípusok gyakoriságát az „Expectation-Maximization” algoritmus (Slatkin és Excoffier, 1996) alkalmazásával kalkulálja ki. Az LD tesztelés beállításai a következők voltak: permutációk száma: 1000, kezdeti feltételek száma: 10.

2.6.1.4. Tesztsorozat, kísérleti jellegű hibaráta, Bonferroni eljárás

Amennyiben a HWE és az LD tesztelés azt célozza meg, hogy az egyensúlyt a lokuszok összességén vizsgálja, akkor a lokuszok ugyanazon hipotézisen elvégzett tesztsorozatban vesznek részt. Ebben az esetben a tesztsorozat szignifikancia-szintje (az ún. *kísérleti jellegű hibaráta*) annak a valószínűsége, hogy a tesztek közül legalább egy a nullhipotézis elvetését okozza, ha a nullhipotézis igaz. A kísérleti jellegű hibaráta (α') a következőképpen fejezhető ki:

$$\alpha' = 1 - (1 - \alpha)^L \approx L\alpha,$$

ahol α az egyedi teszt szignifikancia-szintje és L a tesztek száma.

A fenti procedúrát *Bonferroni eljárásnak* hívjuk (Weir, 1996). Annak érdekében, hogy elkerüljük a hipotézis hamis elvetését, minden egyedi tesztelést szorosabbá kell tenni a fenti képlet átalakított formája alapján:

$$\alpha = 1 - (1 - \alpha')^{\frac{1}{L}} \approx \alpha' / L$$

Pl. 100 tesztből álló tesztsorozat esetén az $\alpha' = 0,05$ kísérleti hibarátához $\alpha = 0,0005$ egyedi szignifikancia-szintet kell alkalmazni.

2.6.2. *Populációk összehasonlító elemzése igazságügyi genetikai aspektusból*

2.6.2.1. *Populáció-párok homogenitás-vizsgálata G-statisztikával*

A teszthez G. Carmody (Carleton University, Ottawa, Canada) által nagylelkűen rendelkezésünkre bocsátott *RXC* szoftvert használtuk. Az *RXC* program két utas, R-sor \times C-oszlop kontingencia táblázatok statisztikai szignifikanciáját teszteli. A szoftver a megfigyeléseket tartalmazó (eredeti) kontingencia táblázatból véletlenszerűen – az alapbeállítás szerint – 1000 új táblázatot replikál úgy, hogy az új táblázatok marginális (az egyes sorok és az egyes oszlopok) összegei megegyezzenek az eredeti táblázat hasonló összegeivel. A program minden egyes táblázatra kiszámítja a G-statisztikát, majd végeredményként kiírja a random táblázatok azon hányadát (P), ahol a G-statisztika értéke meghaladta az eredeti táblázat hasonló értékét. Ha a P érték alacsony, akkor számolni kell azzal, hogy az eredeti táblázat sorai és oszlopai egymástól nem függetlenek; az adatok nem mutatnak elegendő homogenitást ahhoz, hogy összevonhassuk őket. A populációgenetika nyelvén ez azt jelenti, hogy adott szignifikancia-szint mellett (általában ez 5%) a populációk között szignifikáns különbséget találtunk. A szoftver egyik legfőbb erénye, hogy a kis számú megfigyeléseket tartalmazó vagy éppen üres cellákat (nincs megfigyelés) is képes tolerálni.

2.6.2.2. *F-statisztika és AMOVA*

Az egyes populációk között a genetikai profilok interpopulációs variabilitásának megbecsléséhez hagyományos Wright-féle F-statisztika analízist és molekuláris varianciaanalízist (AMOVA) használtunk (Wright, 1951; Weir és Cockerham, 1984; Excoffier és mt sai, 1992; Michalakis és Excoffier, 1996). A teljes populációs mintára vonatkozó Wright-féle F-statisztika értékeket [F_{IS} , F_{IT} , F_{ST} (θ)] a *Genetic Data Analysis (GDA v.1.1)* szoftver segítségével (Lewis és Zaykin, 2001) az egyes lokuszokra nézve számítottuk ki. A populáció-párookra vonatkozó F_{ST} értékeket és azok statisztikai

szignifikanciáját az *Arlequin* szoftverrel határoztuk meg az alábbiakban ismertetett szimuláció segítségével.

Az AMOVA elemzése során ún. Φ_{ST} értékeket kapunk, amely a Wright-féle F_{ST} értékhez hasonló módon azt fejezi ki, hogy a populációkban meglévő totális molekuláris variancia mekkora része tulajdonítható a populációk közötti varianciának. Mivel az AMOVA nem képes a megkettőződött lokuszok kezelésére, így az analízisnél a DYS385 adatokat nem vettük figyelembe. Az AMOVA elemzéshez az *Arlequin* szoftvert használtuk. Az AMOVA a genotípusok között meglévő azon molekuláris távolságon alapul, amely egyenlő az egyes genotípusokat egymástól elválasztó mutációk minimális számával. A távolságszámításnál a mikroszatelliták lépésenkénti mutációs (Stepwise Mutation) modelljét vettük figyelembe (Ota és Kimura, 1973; Valdes és mtsai, 1993). A megfelelő távolság-mátrix kiszámításához az STR allélok repeat-szám különbségeinek négyzetét használtuk (Michalakis és Excoffier, 1996). A Φ_{ST} értékekkel jellemzett populációk közötti variabilitás statisztikai szignifikanciáját szimulációval teszteltük. A szimulációt végző szoftver az adott populáció-párban megfigyelt genotípusokat véletlenszerűen újra szétosztja a két populáció között; a szoftver a beállítás szerint összesen 1000 ilyen replikációt hajt végre. A szimulációval kapott P érték annak a gyakoriságát adja meg, hogy a replikációk hányad részében haladta meg a szimulációs Φ_{ST} érték az aktuális Φ_{ST} értéket. Ebben az esetben a $\{\Phi_{ST} = 0\}$ nullhipotézist teszteltük a $\{\Phi_{ST} > 0\}$ hipotézis ellenében. Ha a kapott $P < 0,05$, akkor $\alpha = 0,05$ szignifikancia-szint mellett a nullhipotézist elvethetjük, azaz molekuláris varianciaanalízissel szignifikáns különbséget találtunk a két populáció között. Az előbbieket alapján minél nagyobb a Φ_{ST} érték, annál nagyobb a populációk közötti molekuláris távolság. A Φ -statisztikát populáció-párokra lokuszonként, valamint a teljes populációs mintára is elvégeztük.

2.6.2.3. Filogenetikai fa konstrukció

A populációk közötti genetikai távolság kiértékelésére „neighbor-joining” módszerrel, a *PHYLP* v3.5c szoftver (Felsenstein J., 1989) segítségével „unrooted” filogenetikai fát konstruáltunk. A távolságmátrixhoz a Nei-féle standard genetikai távolságot (D_S) (Nei, 1972) használtuk. A távolságmátrixot a *MICROSAT* v.1.5d szoftver segítségével számítottuk ki (*MICROSAT*: <http://hpgl.stanford.edu/projects/microsat/programs/>). A bootstrap értékeket 1000 replikációból kaptuk meg.

2.6.2.4. Az átlagolt fenotípuskülönbség-értékek [mean phenotype difference values (MVs)] meghatározása

A vizsgálatot Meyer és munkatársai (1995) módszere alapján négy lokuszon (VWA, TH01, TPOX, CSF1PO) a BuCa-BaRo populáció-párra végeztük el. Mindkét populációs mintából véletlenszerűen kiválasztottunk 100-100 személyt (1-1 almintá), és az egyes személyekre kiszámítottuk az egy-, két- (VWA-TH01), három- (VWA-TH01-TPOX) és négy- (VWA-TH01-TPOX-CSF1PO) lokuszos DNS-profilokon a várt genotípusgyakorisági értékeket. Az egyes alminták közötti összehasonlítás két részre (BaRo→BuCa és BuCa→BaRo) bomlott attól függően, hogy a genotípus-gyakoriságot melyik almintából származó személy DNS-profiljára kalkuláltuk ki. Az egy személyhez tartozó, egymásnak megfelelő profil-gyakorisági értékeket egymással elosztottuk, és az almintára átlagoltuk. Az így kapott értékeket *átlagolt fenotípuskülönbség-értékeknek [mean phenotype difference values (MVs)]* nevezzük. Pl. ennek alapján a BaRo→BuCa összehasonlítás azt jelentette, hogy meghatároztuk 100 BaRo személy DNS-profiljának ($G_{BaRo,i}$) gyakoriságát a BaRo ill. a BuCa allélgyakorisági adatbázis (A_{BaRo} , A_{BuCa}) segítségével. A BaRo→BuCa irányban az MV-értéket a következő számítás révén kaptuk meg:

$$MV_{BaRo \rightarrow BuCa} = \frac{\sum_{i=1}^{100} \frac{P(G_{BaRo,i} | A_{BaRo})}{P(G_{BaRo,i} | A_{BuCa})}}{100}$$

Az ellentétes irányú összehasonlítás során a BuCa személyek profiljait vizsgáltuk hasonló módon. Az elemzést a HWE és linkage egyensúly feltételezésével, a minimális allélgyakorisági értékek felhasználásával végeztük el.

2.6.2.5. Logaritmikus sörét- (scatter) és egy személyű (within-person) Tippett-diagramok

A teljes populációs mintánk genetikai adatainak felhasználásával az elemzett populációs adatbázisok igazságügyi genetikai célú alkalmazásának hatásait logaritmikus sörét- (scatter) diagramokkal és az ún. egy személyű (within-person) Tippett-diagramokkal vizsgáltuk. Ezekkel az analízisekkel azt a bűnügyi helyzetet szimuláljuk, amikor a tettes és a gyanúsított egy személy, azaz a bűncselekmény során keletkezett biológiai nyom a

gyanúsítottól származott. Az analízishez az öt populációs adatbázis DNS-profiljait egyesítettük. Valamennyi DNS-profilra kiszámítottuk a kriminalisztikai cél- (target)-profil egyezés ("találat") valószínűségi hányadosát (LR). A valószínűségi hányados ebben az esetben az egyezési valószínűség (pM) reciproka. A pM értéket – a HWE és kapcsoltsági egyensúly feltételezésével, a termék szabálynak megfelelően, a minimális allélgyakoriság alkalmazásával – a Balding (1995) által kidolgozott, Bayes-elvű módszerrel kalkuláltuk ki. A Balding-módszer lényege, hogy a pM kiszámításánál a cél-genotípust kétszer hozzáadjuk ahhoz a populációs mintához, amelyből az allélgyakoriságokat vesszük. Így az allélgyakoriságok *posterior* értéke minden cél-profil esetében a következőképpen számítható:

$$p_{i,cél} = \frac{x_i + a}{n + 4},$$

ahol x_i a cél-genotípus vizsgált alléljának megfigyelt száma a populációs mintában, n a kromoszómák száma, $a = 4$ vagy 2 attól függően, hogy a cél-genotípus homozigóta vagy heterozigóta.

. A sörét-diagramon minden egyes cél-profilnak megfelelő LR értéket tízes alapú logaritmikus skálán ábrázoltunk. A diagramok tengelyeit annak az allélgyakorisági adatbázisnak a jelzésével láttuk el, ahonnan a tengelyhez tartozó LR értéket számítottuk (Budowle és mtsai, 1994; Monson és Budowle, 1998).

A cél-profilok valószínűségi hányados értékeinek eloszlását az egy személyű Tippett-diagramokon hasonlítottuk össze. A Tippett-diagramon azon cél-profilok részarányát tüntettük fel egy adott pontban, amelyek log LR értékei meghaladják a pont x-tengely vetületénél feltüntetett értéket (Evett és mtsai, 1996; Evett és Weir, 1998). Hasonló módon ábrázoltuk az LR értékek populáció-párokra vonatkozó különbségeinek eloszlását is. A sörét- és Tippett-diagramot a budapesti allélgyakorisági adatbázis (BuCa) $F_{ST}(\theta)$ és F_{IS} F-statisztikával korrigált valószínűségi hányados értékeivel is megrajzoltuk. A korrekció során a pM értékeket az USA Nemzeti Kutatótanácsának (NRC) 1996. évi, Bayes-elven alapuló ajánlásának (NRC, 1996) Ayres és Overall (1999) által javasolt kibővítésével módosítottuk. A szubpopulációs teória alapján a korrekció az egyezési valószínűség két F-statisztikával [$F_{ST}(\theta)$ és F_{IS}] történő módosítását jelenti az alábbiak szerint:

$$P(A_i A_i | A_i A_i) = \frac{\theta + (1-\theta)p_{i,cél}}{F_{IS} + (1-F_{IS})[\theta + (1-\theta)p_{i,cél}]} \left[F_{IS}^2 + 2F_{IS}(1-F_{IS}) \frac{2\theta + (1-\theta)p_{i,cél}}{1+\theta} + (1-F_{IS})^2 \frac{[2\theta + (1-\theta)p_{i,cél}][3\theta + (1-\theta)p_{i,cél}]}{(1+\theta)(1+2\theta)} \right]$$

$$P(A_i A_j | A_i A_j) = \frac{2[\theta + (1-\theta)p_{i,cél}][\theta + (1-\theta)p_{j,cél}]}{(1+\theta)(1+2\theta)},$$

ahol $P(A_i A_i | A_i A_i)$ az egyezési valószínűség értéke homozigóta genotípus esetén;

$P(A_i A_j | A_i A_j)$ az egyezési valószínűség értéke heterozigóta genotípus esetén;

$p_{i,cél}, p_{j,cél}$ a cél-profilhoz tartozó allélgyakoriságok.

Megfigyelhetjük, hogy $\theta = 0$ és $F_{IS} = 0$ esetén visszakapjuk a HWE alapképleteket

$$[P(A_i A_i | A_i A_i) = p_{i,cél}^2, P(A_i A_j | A_i A_j) = 2p_{i,cél}p_{j,cél}].$$

A korrekció esetén a Tippet-diagramokon csak azokat a cél-profilokat ábráztuk, ahol a BuCa adatbázisból számolt LR értékek meghaladták a többi adatbázisból kalkulált valószínűségi hányadosokat.

2.6.2.7. Az etnikai eredet becslése (EEO)

Bűncselekmény során biológiai nyomot hátrahagyó személy etnikai hovatartozásának megbecsülésére (Estimating of Ethnic Origin, EEO), illetve az EEO procedúra megbízhatóságának ellenőrzésére Michael Klintschar és munkacsoportjával (Hallei és Grazi Egyetemek Igazságügyi Orvostani Intézetei) karöltve a következő módszert alkalmaztuk (Klintschar és mtsai, 2003):

A vizsgálathoz nyolc populációs adatbázist használtunk fel, amelyeket három csoportba osztottunk:

- | | | |
|--------------|----|---|
| I. csoport | 1, | BuCa, (Budapest), 223 személy, 17 lokusz |
| | 2, | BaRo, (Baranya), 135 személy, 17 lokusz |
| II. csoport: | 3, | osztrák (Styria), 500 személy, 19 lokusz |
| | 4, | egyiptomi (Kairó), 140 személy, 19 lokusz |
| III. csoport | 5, | amerikai kaukázusi (New York), 125 személy, 13 lokusz |
| | 6, | afro-amerikai (New York), 127 személy, 13 lokusz |
| | 7, | amerikai ázsiai (New York), 126 személy, 13 lokusz |
| | 8, | amerikai spanyol ajkú (New York), 127 személy, 13 lokusz. |

Az EEO procedúra szimulációját a csoportokon belül végeztük el. A szimuláció során először egy személyt kiemeltünk a saját (referencia-) populációjából, és a kiemelt személy hiányában az allélgyakorisági értékeket újra kiszámítottuk a referencia-populációra nézve.

Ezzel a lépéssel kiküszöbölhető, hogy az adott személyre végzett profilgyakorisági számítás elfogult legyen a saját populációjára nézve (Brenner, 1998). A szimulációban a kiemelt személy (akinek etnikai hovatartozása ismert) DNS-profilja felel meg az ismeretlen etnikai eredetű biológiai nyom DNS-profiljának. Második lépésben a kiválasztott személy profilgyakorisági értékét – a HWE és kapcsoltsági egyensúly feltételezésével, a termék szabálynak megfelelően, a minimális allélgyakoriság alkalmazásával – kiszámítottuk a saját, módosított ill. a többi, módosítatlan populációs minta allélgyakorisági adatbázisának felhasználásával. Ha a kiemelt személy profilja a saját populációjában bizonyult a leggyakoribbnak, akkor az etnikai hovatartozás becslését (ti. azt, hogy a DNS-profil ebből a populációból származott) korrektnek fogadtuk el. Ha a kiemelt személy profilja nem a saját populációjában lett a leggyakoribb, az EEO-t hamisnak tekintettük. Végezetül a kiválasztott személyt visszahelyeztük saját populációjába és a fenti procedúrát a maradék személyekkel is elvégeztük. Ennek alapján az elvégzett EEO-k száma megegyezik a referencia-populáció egyedszámával. Ebben az összefüggésben az etnikai származás becslésének megbízhatóságát a korrektnek tekintett EEO-k aránya fejezi ki.

Az EEO megbízhatóságának tesztelése során arra is kíváncsiak voltunk, hogy a DNS-profilokat alkotó lokuszok számától függően hogyan változik a korrekt becslések aránya. Ennek érdekében a fenti szimulációt 1, 3, 5, 7, 9, 11, 13, 15, 17 és 19 lokusz bevonásával külön-külön is elvégeztük.

3 EREDMÉNYEK ÉS ÉRTÉKELÉSÜK

Az értekezésemben tárgyalt magyarországi populációs minták mikroszatellita-alapú genetikai feltérképezését hat, egymásra épülő szakaszban hajtottuk végre (1. táblázat). Mindegyik vizsgálati szakasz az adott STR rendszer elemzésének laboratóriumi protokollunkba történő bevezetését, hitelesítését és esetleges továbbfejlesztését jelentette. Ezzel párhuzamosan az adott időszakban rendelkezésünkre álló populációs mintákon végrehajtottuk a genetikai analízist. Az eredményeink publikálására az egyes vizsgálati egységek lezárása után került sor.

1. táblázat Az értekezésben tárgyalt STR rendszerek és populációs minták vizsgálatának szakaszai

Időszak	STR-rendszer						Populációs minta			Referencia	
	VT	HFV/CTT	CTTV	FFFL	P+	Cof	Y	BuCa	BaRo		DeRo/BuAs/Kisk
1994-95	+							+			Füredi et al., 1995
1995-96		+						+			Füredi et al., 1996
1996-97			*	+				+	*		Füredi et al., 1997a, b
1997-98				+					+		Füredi et al., 1998
1998-2000					+		+	+	+		Füredi et al., 1999 Egyed et al., 2000
2000-03						*	+	*	*	* +	kézirat folyamatban

VT = VWA és TH01 szimplex

HFV = HPRTB-FES-VWA triplex

CTT = CSF1PO-TPOX-TH01 triplex

CTTV = CSF1PO-TPOX-TH01-VWA kvadruplex

FFFL = F13A01-FES-F13B-LPL kvadruplex

P+ = D3S1358-VWA-FGA-Amel-D8S1179-D21S11-D18S51-D5S818-D13S317-D7S820 multiplex

Cof = D3S1358-D16S539-Amel-TH01-TPOX-CSF1PO-D7S820 multiplex

Y = DYS19-DYS389-I/II (Triplex I), DYS393-DYS390-DYS391 (Triplex II), DYS392-DYS385 (Duplex)

Az eredmények bemutatása során a szövegben, a táblázatokban és ábrákon használt, esetlegesen ki nem fejtett rövidítések jelentése a *Függelék* végén található meg.

3.1. Az STR lokuszok genotipizálásának bemutatása

Ezt a rövid alfejezetet célszerűnek tartottam értekezésem *Eredmények és értékelésük* részének elejére beilleszteni, mivel a tudományos publikációk terjedelmi korlátai miatt saját közleményeinkbe nem kerülhettek bele. Mivel az igazságügyi genetikai szakértői

munka alapfeltétele a korrekt genotipizálás, ezért annak jelentőségét nem lehet eléggé hangsúlyozni.

Az STR lokuszok igazságügyi célú genotipizálása során a tipizálandó minta alléljait referencia-allélokkal hasonlítjuk össze. A referencia-allélokot allél-létrába gyűjtve, szimultán alkalmazzuk. A *Függelék* (F.) F.1-7. ábráin az értekezésemben tárgyalt genotipizálási rendszerek képét allél-létrával és a velük együtt futtatott, vizsgálni kívánt mintával mutatom be. Az allélok mellett azok típusát is feltüntettem. Az ábrákon az allélok mérete alulról fölfelé ill. balról jobbra növekszik. Az ábrák felirataiban hivatkozok azokra a publikációinkra, amelyeknél a bemutatott genotipizálási módszert alkalmazzuk.

3.2. A genotipizálás (fél)automatikus allélméret-meghatározásának pontossága (allele sizing accuracy, precision)

Az általunk alkalmazott genotipizálási rendszerek (fél)automatikus allélméret-meghatározásának pontosságát (allele sizing accuracy, precision) az ALF és ABI szekvenátorokon egyaránt leteszteltük. A méréseket az ALF rendszer esetében 8 lokuszon (VWA, TH01, TPOX, CSF1PO, LPL, F13B, FES, F13A01), az ABI rendszer esetében pedig 6 lokuszon (DYS19, DYS389-I/II, DYS390, DYS391, DYS393) végeztük el.

Az ALF szekvenátoron vizsgált 135 baranyai roma (BaRo) személy VWA, TH01, TPOX, CSF1PO lokuszának 928 allélikus fragmensére mért adatokat a 2. táblázatban mutatom be.

2. táblázat A számított allélméret szórása, maximális mérettartománya, pontossága és precizitása^a (Füredi és mtsai, 1997b alapján)

Lokusz	Allél	SD (nt)	Tartomány (nt)	Pontosság% (accuracy)	Precizitás% (precision)	N
VWA	14	0,07	0,3	99,95	99,95	21
	15	0,05	0,2	99,98	99,96	17
	16	0,06	0,3	99,96	99,96	75
	17	0,07	0,3	99,97	99,96	55
	18	0,07	0,3	99,97	99,96	52
	19	0,04	0,1	99,99	99,97	5
TH01	6	0,06	0,2	99,98	99,97	55
	7	0,06	0,2	99,98	99,97	23
	8	0,05	0,2	99,98	99,97	53
	9	0,06	0,2	99,98	99,97	46
	9.3	0,06	0,3	99,97	99,97	61
TPOX	8	0,08	0,4	99,98	99,97	96
	9	0,08	0,2	99,98	99,97	16
	10	0,05	0,2	99,99	99,98	39
	11	0,06	0,2	99,98	99,98	77
CSF1PO	8	0,08	0,2	99,97	99,97	5
	9	0,10	0,2	99,98	99,97	3
	10	0,10	0,4	99,97	99,97	68
	11	0,09	0,4	99,97	99,97	65
	12	0,11	0,6	99,96	99,97	79
	13	0,10	0,3	99,97	99,97	11

^a Az adatok legalább 3 megfigyelést tartalmaznak
N = PCR-fragmensek száma
nt = nukleotid

A számított allélméret szórása 0,04-0,11 nukleotid (nt) közé, míg a maximális allélikus mérettartomány 0,1-0,6 nt közé esett. A maximális allélikus mérettartomány (más néven allél-ablak) azt fejezi ki, hogy az adott alléltípus esetén az allélok mért nagysága között mekkora volt a legnagyobb különbség. A 2. táblázatból kiolvasható, hogy a pontosság és a precizitás a vizsgált teljes mérettartományon (113-327 nt) belül alig változik (átlaguk mindkét esetben 99,97%), míg a számított allélméret szórása az allélok méretének növekedésével kis mértékben nő.

A számítógépes allélméret-meghatározás pontosságát nemcsak a baranyai roma populációs mintán, hanem az allél-létrán is leteszteltük. Öt darab, csak allél-létrát tartalmazó gél egymást követő futtása és kiértékelése révén összesen 4428 allélikus fragmens méretét határoztuk meg. Ebben az esetben a méret-meghatározás átlagos pontossága 99,98%-osnak, átlagos precizitása 99,97%-osnak bizonyult.

A 2. táblázatban szereplő értékekhez teljesen hasonló eredményeket kaptunk az LPL, F13B, FES, F13A01 lokuszok alléljaira is (Füredi és mtsai, 1998; az adatok nincsenek bemutatva).

Az ALF rendszeren az allélméret-meghatározás pontosságára irányuló vizsgálataink mindegyik analizált STR rendszer esetében kielégítő eredményre vezettek. Két db belső méretstandard alkalmazásával a pontossági és precizitási értékek átlaga elérte a 99,97-99,98%-ot. A számított allélméretes maximális tartománya, az allél-ablak nem érte el az 1 nt-ot, egy alléltípus kivételével a teljes mérési tartományon belül 0,5 nt alatt maradt. A számított allél-méretes szórása 0,15 nt-nál kisebbnek adódott. A kapott eredmények alapján az általunk kis mértékben továbbfejlesztett ALF fragmens-analizáló rendszer alkalmas STR lokuszok 1 nukleotid pontossággal történő genotipizálására az igazságügyi genetikai szakértői gyakorlatunkban.

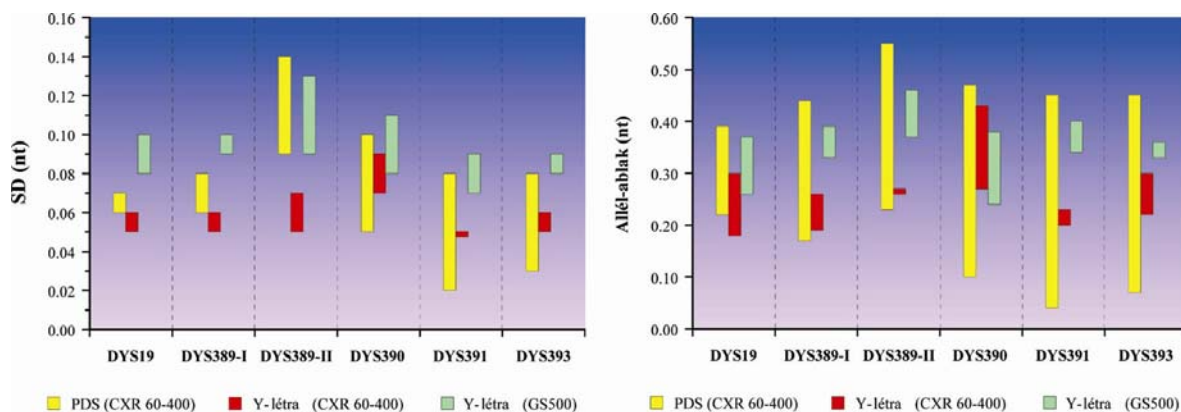
Az ABI Prism 310 szekvenátoron végzett precizitás-vizsgálatunk eredménye a 3. táblázatban látható. Az 1. ábrán a táblázat értékeit grafikus formában is bemutatjuk.

3. táblázat A populációs adatbázis-mintákra (PDS) és az Y-STR allél-létrákra számított allélméret szórás-tartománya és maximális mérettartománya (allél-ablak) (Füredi és mtsai, 1999)

Lokusz	PDS (CXR60–400)		Y-létra (CXR60–400)		Y-létra (GS500)	
	SD (nt)	Allél-ablak (nt)	SD (nt)	Allél-ablak (nt)	SD (nt)	Allél-ablak (nt)
DYS19	0,06-0,07	0,22-0,39	0,05-0,06	0,18-0,30	0,08-0,10	0,26-0,37
DYS389-I	0,06-0,08	0,17-0,44	0,05-0,06	0,19-0,26	0,09-0,10	0,33-0,39
DYS389-II	0,09-0,14	0,23-0,55	0,05-0,07	0,26-0,27	0,09-0,13	0,37-0,46
DYS390	0,05-0,10	0,10-0,47	0,07-0,09	0,27-0,43	0,08-0,11	0,24-0,38
DYS391	0,02-0,08	0,04-0,45	0,05	0,20-0,23	0,07-0,09	0,34-0,40
DYS393	0,03-0,08	0,07-0,45	0,05-0,06	0,22-0,30	0,08-0,09	0,33-0,36

nt = nukleotid

1. ábra Az 1. táblázat grafikus megjelenítése



A vizsgálatban a BuCa és BaRo populációs mintákban lévő 194 férfi hat Y-STR lokuszának valamint a lokuszokhoz tartozó allél-létrák allélméret-meghatározását teszteltük a precizitás szempontjából. Az allél-létrák esetében a vizsgálatot a CXR60-400 belső méretstandard mellett a GS500 méretstandardra is elvégeztük. Mindegyik precizitási kísérletben az egyes alléltípusok számított allélméretének szórása 0,15 nt-nál kisebbnek adódott. A szórás-értékek a DYS389-II és a DYS390 lokusz kivételével 0,1 nt-on belül mozogtak. A mért allél-ablakok szélső értéke csak a DYS389-II lokusz esetében haladta meg a 0,5 nt-ot. Az allél-létrák esetében a 3. táblázat grafikus formájából (1. ábra) könnyen megállapítható, hogy a precizitást szemléltető szórások és allél-ablakok maximális értékei szinte minden lokuszon jelentősen kisebbek lettek, ha az allél-létrákhoz GS500 helyett CXR60-400 belső méretstandardot kevertünk. Az egyetlen kivételt a DYS390 lokusz jelentette.

Az ABI Prism 310 kapilláris elektroforetikus rendszeren az allélméret-meghatározás pontosságára irányuló vizsgálatunk mindegyik analizált STR rendszer esetében kielégítő eredményre vezettek. Két különböző belső méretstandard összehasonlítása során egyértelmű allélméret meghatározási pontosság-növekedést detektáltunk abban az esetben, amikor egy több fragmensből álló méretstandardot (CXR60-400) alkalmaztunk. Az általunk kapott precizitási értékek gyakorlatilag megegyeztek egy másik igazságügyi genetikai munkacsoport velünk párhuzamosan elvégzett teszteredményeivel (Lazaruk és mtsai, 1998). A számított allélméret maximuma, az allél-ablak nem érte el az 1 nt-ot, egy lokusz kivételével a teljes mérési tartományon belül 0,5 nt alatt maradt. A számított allél-méret szórása 0,15 nt-nál kisebbnek adódott. A kapott eredmények alapján az általunk kis mértékben továbbfejlesztett ABI Prism 310 fragmens-analizáló

rendszer alkalmas STR lokuszok 1 nukleotid pontossággal történő genotipizálására az igazságügyi genetikai szakértői gyakorlatunkban.

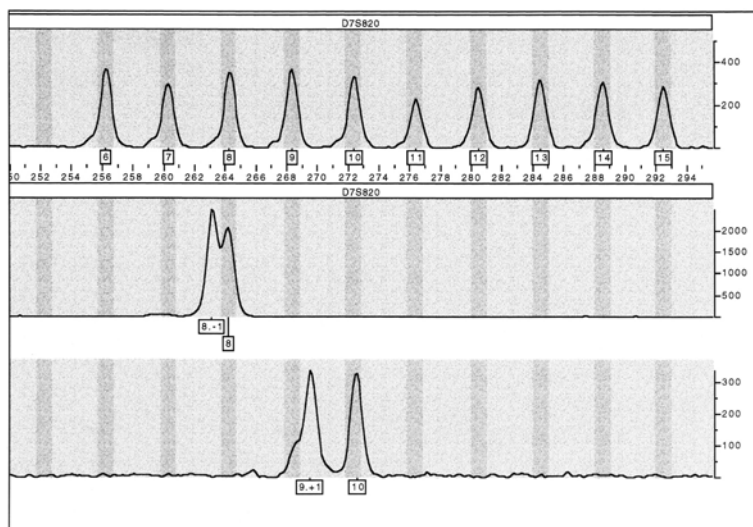
3.3. Ritka és mikrovariáns allélok

Az STR markerekkel végzett kriminalisztikai DNS-vizsgálataink során több lokuszon olyan allélok is detektáltak, amelyek a hazai és a külföldi populációs felmérések eredményei alapján igen ritkának számítanak vagy az észlelés időpontjában éppenséggel még nem voltak ismertek. Négy lokusz (D7S820, D13S317, HPRTB, DYS385) esetében ezen allélok közös vonása az volt, hogy a lokusz allél-létrájából hiányoztak. Az ötödik, DYS19 markernél egy DeRo személy mintájában két-allélos genotípust figyeltünk meg. A ritka allélok döntő többségét a mikrovariáns (nem konszenzus), ún. interallélok tették ki. Elnevezésük elsősorban elektroforetikus eredetű, abból fakad, hogy elektroforetikus mobilitásuk általában eltér a többi, az allél-létrát alkotó és konszenzusnak tekintett allél mobilitásától. Az általunk alkalmazott genotipizálási rendszerek nagy elválasztóképességének köszönhetően ezek az allélok a konszenzus alléloktól megkülönböztethetők voltak annak ellenére, hogy a közöttük lévő távolság sokszor csak 1 nt-nyi volt.

3.3.1. A D7S820 lokusz egy 8.-1-es, egy 9.+1-es és egy 10-es mikrovariáns alléljának elemzése

Kriminalisztikai vizsgálataink során a D7S820 lokuszon két személyben találtunk mikrovariáns allélt. A budapesti referencia populációs minta (BuCa) 146-os sorszámú férfi mintájából 9var/10-es genotípust határoztunk meg. Egy bűneset kapcsán 8var/8-as genotípust figyeltünk meg egy jugoszláv állampolgárságú, albán nemzetiségű férfi mintájában (Egyed és mtsai, 2000). A D7S820 lokusz mikrovariáns alléljainak elektroforetikus képe a 2. ábrán látható.

2. ábra A D7S820 lokuszon megfigyelt mikrovariáns allélok elektroferogramja



Mindkét genotípus PCR-termékeit pGEM-T-ben (Promega) klónoztuk és mindegyik allél legalább 2-2 klónját megszekvenáltuk. A mikrovariáns allélokat szekvencia struktúrájuk alapján 8.-1-es és 9.+1-es allélnak neveztük el (3. ábra).

3. ábra Négy D7S820 allél fragmensmérete és szekvencia struktúrája (Egyed és mtsai, 2000)

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

Referencia-szekvenciaként a GenBank G08616 szekvenciáját használtuk, amely 12 GATA repeatet tartalmaz. A 9.+1-es allél a repeat régióban 9 teljes repeatet hordozott és egy T addícióval rendelkezett a 3' flanking régió (T)₉-es blokkjában. A 8.-1-es allél 8 teljes GATA repeattal bírt a repeat régióban, míg a mikrovariáns allél 3' flanking régiójának előbbieken említett T-sorozatából egy T hiányzott. A 9var/10-es minta 10-es alléja is szekvencia-variációt mutatott a referencia-szekvenciához képest, mivel egy T→A transzverzió volt detektálható a T-blokk végén (3. ábra). A (T)₉-es blokkban megfigyelt variáns szekvenciák azt sejtették, hogy ez a szakasz – magasabb mutációs rátája miatt – a D7S820 lokusz polimorf része. A szekvenálás időpontjában, 1999. tavaszán a *BLAST*

program (Altschul és mtsai, 1997) segítségével elvégzett szekvenciahomológia-keresés során 100%-os egyezést találtunk a humán DJ0649P17 PAC klónnal (GenBank AC004848). A klón a D7S820 lokusz teljes szekvenciáját tartalmazta. A D7S820 lokusz kromoszómás helyét korábban csak a 7. kromoszóma hosszabbik karjára (7q) lokalizálták (Bouffard és mtsai, 1997). Mivel a homológiát mutató klónt a 7q11.23-q21 kromoszóma-szakaszra térképezték, ezért így lehetőség nyílt a D7S820 lokusz kromoszómás helyének pontosabb meghatározására

3.3.2. A D13S317 lokusz egy 12-es és egy ritka 7-es alléljának elemzése

Egy bűncselekményben végzett igazságügyi szakértői munkánk során egy ázsiai származású fivér-pár mintájában D13S317 7/12-es genotípust találtunk (Boutrand és mtsai, 2001). Mivel a 7-es allélt addig nem figyeltünk meg és az allél az allél-létrából is hiányzott, a ritka allél jellemzéséhez az allél megszekvenálását határoztuk el. Mivel a D13S317 lokusz PCR-sokszorozásához használt Profiler AmpF/STR Kit gyártója titkosan kezelte a kitben alkalmazott primer szekvenciákat, ezért a szekvenáláshoz a publikált primer szekvenciákat (GDB G09017) kíséreltük meg felhasználni. Ez a 12-es allélok esetében sikerre is vezetett, amelynek eredményeképpen a 12-es allélok szekvenciája a referencia-szekvenciával megegyezőnek bizonyult. Azonban ugyanezzel a primer-párral a 7-es allélok PCR-amplifikálása nem sikerült. Az amplifikálást egy másik gyártótól beszerzett PCR-kittel megismételve (Gamma STR Multiplex Kit, Promega) ugyancsak nem kaptunk 7-es allélra jellemző PCR-terméket. A sikeresen amplifikált 12-es allélok fragmens-mérete a Profiler kit használata esetén 224 bp-nak, a GDB primerek esetén 189 bp-nak, míg a Gamma kit esetén 185 bp-nak adódott. Ezek az értékek azt sugallták számunkra, hogy a GDB primerek és a Gamma kitben alkalmazott primerek nagyban átfedtek egymással. Mindezek alapján azt a hipotézist állítottuk fel, hogy a 7-es allélok kieséséért (drop-out) a primer-szekvencia a felelős.

Hipotézisünk bizonyítása érdekében a forward primertől felfelé és a reverz primertől lefelé primereket terveztünk, amelynek eredményeképpen a PCR-ek során megkaptuk a 7-es allélt is (4. ábra). A 4. ábra 12-es allélján aláhúzással jelöltük a PCR-amplifikálásra használt GDB primereket. A klónozáshoz és szekvenáláshoz használt primerek nukleotidjait kisbetűvel tüntettük fel. A 7-es allél egy igen érdekes variánsnak bizonyult, mivel szekvenciája 2 helyen is módosult. Egyrészt a normál 12-es allél GDB

reverzprimer-kötő részében lévő egyik TGTC nukleotid-négyes kiesett, másrészt egy ATCA-négyes beékelődött a reverz primer szekvencia elé (a deletálódott és inszertálódott szekvencia-részleteket a 4. ábrán vastag betűtípussal szedtük). Nagy valószínűséggel a reverzprimer-kötő szekvenciának e megfigyelt drasztikus módosulása eredményezhette a 7-es allél PCR-sokszorozásának sikertelenségét és kiesését.

4. ábra A D13S317 lokusz vizsgált 12-es és 7-es alléljainak szekvencia struktúrája (Boutrand és mtsai, 2001)

D13S317 allele 12

5'-gggttgctggacatggtatc**ACAGAAGTCTGGGATGTGGA**---N82---(ATCT)₁₂ATCAATC(ATCT)₃**TTCTGTCTGTCTTTTTGGGC**---N36---
gaccaacaattcaagctctc-3'

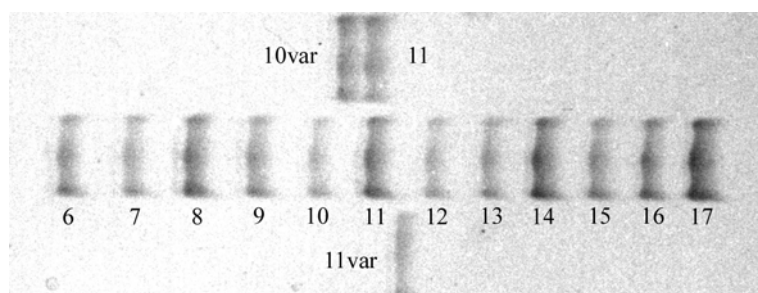
D13S317 allele 7 (variant)

5'-gggttgctggacatggtatc**ACAGAAGTCTGGGATGTGGA**---N82---(ATCT)₇**ATCAATCAATC(ATCT)₃TTCTGTCTTTTTGGGC**---N36---
gaccaacaattcaagctctc-3'

3.3.3. A *HPRTB* lokusz egy 12.-2-es és egy 12.-6-os mikrovariáns alléljának elemzése

A BuCa populációs minta 11-es jelzésű, férfitől származó mintájában variáns 11-es, 79-es sorszámú női mintájában 10var/11-es fenotípust észleltünk (Füredi és mtsai, 1996) (5. ábra). (A variáns allélokot a referált publikációnkban M betűvel jelöltük.)

5. ábra A *HPRTB* lokusz vizsgált 10var/11-es és 11var fenotípusának gélképe



A szekvencia-elemzés eredménye azt mutatta, hogy a 11var allél 12 teljes repeatból áll, és a 3' flanking régióban egy AG deléciót hordoz (6. ábra, az AG deléciót az ábrán vastag nagybetűvel jelöltük). A szekvencia-struktúra alapján az allél jelölése: 12.-2. A 10var allél a 12.-2-eshez hasonlóan 12 teljes repeatból állt, azonban az allél 5' flanking

régiójában lévő két szomszédos CTATCT hexamer közül az egyik kiesett. A 6. ábrán ezt a két hexamert zárójelek között, vastagított nagybetűvel jelöltük. A hatos deléció miatt az allélt 12.–6-os jelöléssel láttuk el. Mertens és munkacsoportja (1999) a flamand populációban szintén megfigyelt a HPRTB lokuszon olyan mikrovaráns allélokot, ahol a deléciók a flanking régiókat érintette. A 6. ábrán megfigyelhető, hogy a két CTATCT hexamer magában foglalja a HPRTB lokusz TATC repeat motívumát (az ábrán aláhúzással jeleztük). A HPRTB lokusz ezen szakasza a lokusz tökéletlen repeat régiójának egy részét képezi, mivel háromszor tartalmazza a TATC repeat motívumot, amelyeket egy TC nukleotid-kettős kétszer, egy TGTC-négyes pedig egyszer megszakít. Mertens és munkacsoportja (1999) korábban rámutatott, hogy ezt a tökéletlen régiót (TN)₁₃ blokkal is jelölhetjük, ahol N = A, C vagy G, és minden második pozícióban egy T foglal helyet. A tökéletlen repeat régió valószínűleg destabilizálja a polimeráz komplexet replikáció közben. A 3' flanking régióban ehhez hasonló tökéletlen repeatekből álló szakasz nem figyelhető meg. Ezek és a korábbi eredmények azt sugallják, hogy az allélt érintő mutációk preferenciálisan az 5' repeat régióját érintik, amelynek következtében új allélok születhetnek.

6. ábra A HPRTB lokusz 12-es alléljának szekvenciája. A számozás a GenBank M26434 felső szálának megfelelő. A primer szekvenciákat vastag kisbetűvel szedtük. A lokusz repeat régiója a szögletes zárójelben található (Vandenberghe és mtsai, 2000)

```
22 798-ctc tccagaatag ttagatgtag gtataccact ttgatgttga cactagtta cctagaactt
atetttgta aatctgtctc tatttccate tetgtcteca ttttgtct(C TATCT)(CTATC T)gtctatctc
[tatctatcta tctatctate tatctateta tctatctate tatctatc]ta aagcaaatte atgcccttc
cctatttatt gaatcgagac catagacAGg ggtgagagaa agaatttgge aggaatgggg atgtgtatta
tctgtggcat-23 080
```

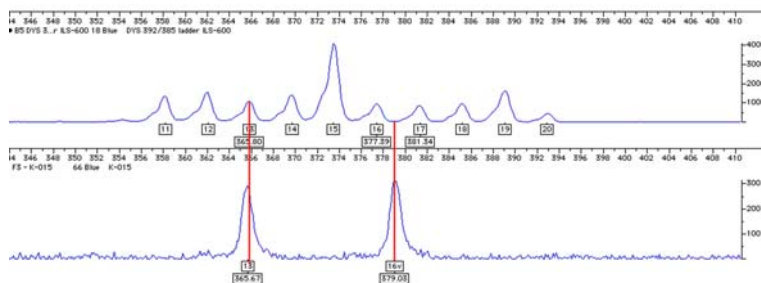
3.3.4. Az Y-STR lokuszok ritka alléljainak elemzése

Az értekezésemben tárgyalt öt populációban összesen 11 mikrovariáns allélt figyeltünk meg, amelyek mindegyike a DYS385 lokuszhoz tartozott. A budapesti referencia (BuCa) és a baranyai roma (BaRo) populációs mintákban észlelt mikrovariánsokat az 1999. évi publikációnkban elemeztük (Füredi és mtsai, 1999). DYS385 *.2-es interallélt három

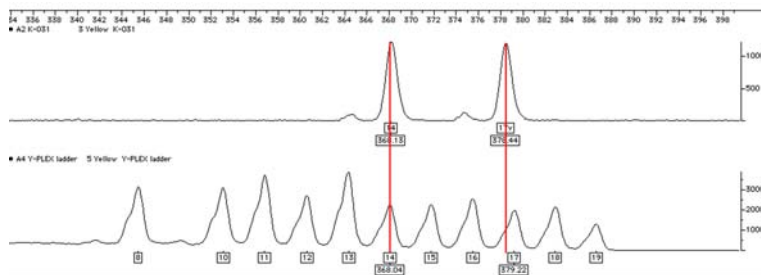
személyben találtunk: a 60-as jelzésű BuCa, a 15-ös jelzésű Kisk és a 142-es jelzésű DeRo férfiban. DYS385 *-1-es interallélt a következő nyolc személyben tipizáltunk: BuCa_003, BuCa_026, BuCa_113, BuCa_142, BaRo_079, BuAs_095, BuAs_178, Kisk_31. Az interallélok helyes elnevezésének bizonyítása szekvencia meghatározást igényel, aminek egy része értekezésem írása idején még folyamatban van. A mikrovariáns allélokon kívül egy északkelet-magyarországi roma személyben (DeRo_133), a DYS19 lokuszon két allélt figyeltünk meg.

Az interallélok és a két-allélos DYS19 lokusz elektroferogramja a 7-9. ábrákon látható.

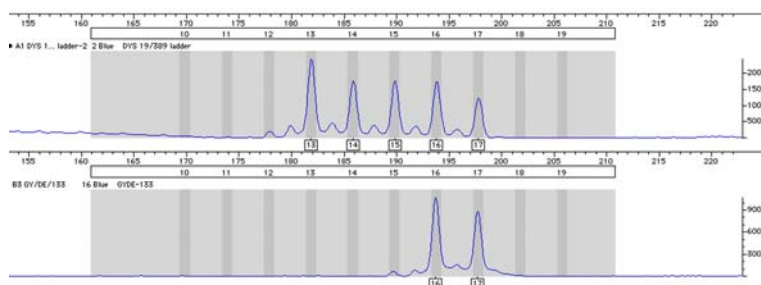
7. ábra Egy DYS385 *.2 interallél elektroferogramja



8. ábra Egy DYS385 *-1 interallél elektroferogramja



9. ábra Egy két allélos DYS19 lokusz elektroferogramja



A DYS19 lokuszon dupla allélt okozó módosulás a korábbi megfigyelések alapján nem túl ritka (kb. 0,4% gyakoriságú, Kayser, és mtsai, 1997). A jelenség a lokusz duplikációjával, majd az egyik allél repeatszám mutációjával magyarázható. Igazságügyi genetikai jelentőséggel bír, mivel téves következtetés vonható le belőle a biológiai nyomok eredetére vonatkozóan: a biológiai nyomot min. két férfitől származtathatjuk – tévesen. Ugyanakkor felismerése esetén a DNS-vizsgálat bizonyító erejét növelheti.

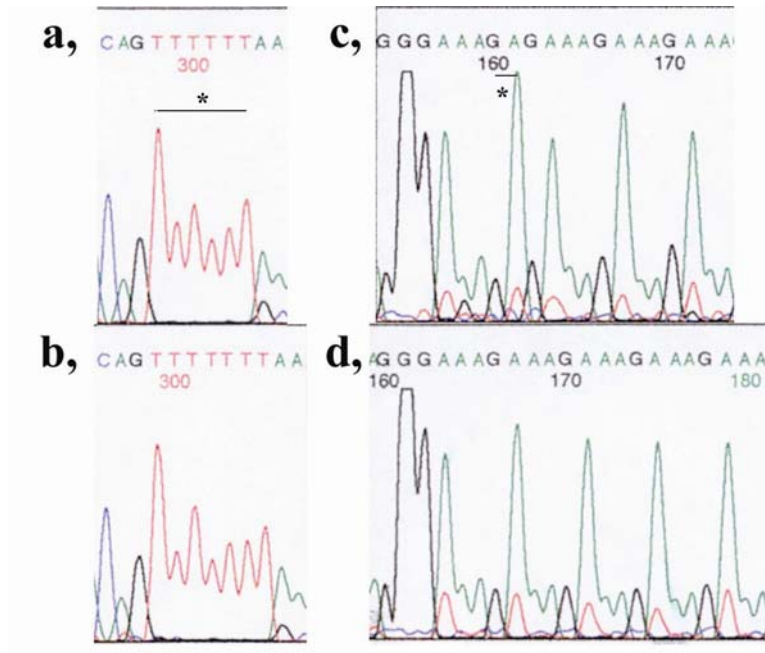
Az összes DYS385 mikrovariánst olyan férfiak mintájában detektáltuk, ahol a duplikálódott lokusz két allélja elkülöníthető volt egymástól. Ugyan a BaRo, BuAs, Kisk populációs mintákban megfigyelt DYS385 mikrovariánsokat az értekezésem írása időpontjáig még nem szekvenáltuk meg, azonban korábbi felmérésünk (Füredi és mtsai, 1999) alapján valószínűsíthető a még nem elemzett DYS385 *.2 és a DYS385 *.–1 interallélok szekvenciája. A már megszekvenált, összesen hat allél három alléltípust képvisel, amelyek struktúrája a 10. ábrán látható.

10. ábra A DYS385 mikrovariáns allélok fragmens-mérete és struktúrája (Füredi és mtsai, 1999)

Alléljelölés	Fragment-hossz (bp)	5' Flanking régió	Repeat régió	3' Flanking régió
17.-1	392	██████████ 0	-(GAAA) ₁₇ -	██████████ - G TTTTTTA- ██████████ 324
18.-1	396	██████████ 0	-(GAAA) ₁₈ -	██████████ - G TTTTTTA- ██████████ 328
17.2	395	██████████	-(GAAA) ₁ -(GA) ₁ -(GAAA) ₁₆ -	██████████

A 11a-b. ábrán bemutatott szekvencia-elemzés eredménye alapján a DYS385 *.–1 interallélok 3' flanking régiójában deléció történt: pl. a 17.–1-es interallél esetében a 324. bázisnál kezdődő hét egymás utáni T helyett csupán hat található meg. Ugyanez a helyzet a 18.–1-es interallél esetében is azzal a különbséggel, hogy lévén ez egy repeattal nagyobb méretű, itt a 328. bázisnál kezdődik a hat T. A 10. ábrán a (T)₆ blokk kezdőpontját a mikrovariáns allélokhoz az EMBL Z93950 referencia-szekvenciával való összehasonlítása alapján adtuk meg. A DYS385 *.2 interallélokban a repeat régióban történt a mutáció, a második (GAAA) ismétlődési egység két A-t érintő deléciójával (11c-d. ábra).

11. ábra A DYS385 mikrovariáns allélok szekvenálása. A delécióit érintő szakaszokat az ábrán *-gal jelöltük
a, *-1 mikrovariáns allél (T)₆ blokkja; **b**, referencia-allél (T)₇ blokkja; **c**, *.2 mikrovariáns allél repeat régiójának kezdeti szakasza; **d**, referencia-allél repeat régiójának kezdeti szakasza



Már a BuCa és BaRo populációs minták vizsgálata során észrevettük, hogy a *-1 mikrovariánssal rendelkező öt BuCa vagy BaRo férfi haplotípusa öt lokuszon (DYS389-I, DYS390, DYS391, DYS392, DYS393) megegyezik egymással (Füredi és mtsai, 1999). Mindezek mellett az öt személy közül három BuCa férfi 7-monolokuszos haplotípusa azonosnak bizonyult (4. táblázat). Érdekesnek ígérkezett tehát annak elemzése, hogy a BuAs és Kisk mintákban talált további három DYS385 *-1 mikrovariáns személy vajon milyen haplotípusú. A haplotipizálás eredménye azt mutatta, hogy mind a nyolc férfi haplotípusa négy lokuszon (DYS389-I, DYS390, DYS391, DYS392) teljesen azonos (4. táblázat). Ahol a többi lokusz esetében az egyes allélok közötti különbség legfeljebb három repeat egységnyinek adódott. A nyolc személy mikrovariáns Y-STR haplotípusának az előzőekben leírt nagyfokú hasonlóságának egyik magyarázata lehet az, hogy ezen Y-kromoszómák eredete közös és a haplotípusok szétválása nem túl régen történt meg.

Populációs felmérésünk öt populációra történő kibővítésével a DeRo és Kisk mintában két újabb DYS385 *.2-es mikrovariáns allélt detektáltunk. A három *.2-es mikrovariánst hordozó férfi közül két személy (BuCa, Kisk) haplotípusa a mikrovariáns

allélokon kívül teljes allélikus egyezést mutatott (4. táblázat), amely szintén ezen Y-kromoszómák közös eredetére utal.

4. táblázat DYS385 interallélokot hordozó személyek haplotípusának összehasonlítása. A konszenzus allélok hátterét piros színnel jelöltük

ID	DYS19	DYS389-I	DYS389-II	DYS390	DYS391	DYS392	DYS393	DYS385	DYS385
BuCa_026	14	12	28	24	10	11	12	14	17.-1
Kisk_031	14	12	28	24	10	11	12	14	17.-1?
BaRo_079	15	12	27	24	10	11	12	13	18.-1
BuCa_142	15	12	28	24	10	11	12	12	17.-1
BuCa_003	15	12	28	24	10	11	12	13	17.-1
BuCa_113	15	12	28	24	10	11	12	14	17.-1
BuAs_178	15	12	29	24	10	11	12	15	17.-1?
BuAs_095	15	12	29	24	10	11	13	15	17.-1?
DeRo_142	14	12	28	24	11	13	13	11	13.2?
BuCa_060	16	12	28	24	11	11	14	13	17.2
Kisk_015	16	12	28	24	11	11	14	13	16.2?

A kérdőjellel jelölt interallélok típusa szekvenálással még nincs megerősítve

Annak érdekében, hogy az egymással nagyfokú hasonlóságot mutató haplotípusok nem túl régi elválásának hipotézisét más megközelítéssel is alátámasszuk, az érintett Y-kromoszómák SNP-vizsgálatát tűztük ki célul. A DNS-minták Y-SNP-analízisét az Oxfordi Egyetem Biokémiai Tanszékén Chris Tyler-Smith munkacsoportja végezte el. Az analízis eredménye az 5. táblázatban látható. Mivel az SNP-vizsgálatokhoz rendelkezésre álló DNS-minta koncentrációja meglehetősen csekély volt, ezért első lépcsőben Tyler-Smith munkacsoportja három mintán kezdte az elemzést a Paracchini és munkatársai (2002) által kidolgozott primer extenzió/MALDI TOF tömegspektrometriás eljárással. A vizsgált minták és SNP-k körét a minták DNS-koncentrációjától függően terjesztették ki. Az elégtelen mennyiségű DNS-minta miatt a BuAs_095 férfi SNP-analízise nem járt eredménnyel. A többi hét DYS385 *-1 mikrovariáns allélt hordozó Y-kromoszóma az SNP-típusuk alapján egy haplocsoportba, az Underhill-féle (Underhill és mtsai, 2000) 60-as csoportba volt sorolható. Ettől a csoporttól egyértelműen elkülönült az SNP-vizsgálatba bevont két *-2-es mikrovariáns Y-kromoszóma haplocsoportja, amely mindkét személy esetében az 52-esnek bizonyult. Az SNP-eredmények tehát alátámasztották a mikrovariáns Y-STR haplotípusok közös eredetére vonatkozó hipotézisünket, mivel mindegyik, haplotípusuk szerint egy csoportba sorolt Y-kromoszóma egyben egy haplocsoportba tartozónak bizonyult SNP-típusuk alapján.

5. táblázat Az Y-STR haplotípusuk alapján egymással nagyfokú hasonlóságot mutató, DYS385 mikrovariáns allélokot hordozó személyek Y-kromoszómájának SNP vizsgálati eredménye (C. Tyler-Smith és S. Paracchini, személyes közlés)

	mult3 ^a		mult10 ^a	mult6 ^a						mult8 ^a			Hcs	
	2 ^b	35 ^b	173 ^b	170 ^b	172 ^b	52 ^b	62 ^b	26 ^b	161 ^b	21 ^b	12 ^b	102 ^b		99 ^b
BuCa_026	-	-	-	-	-	-	-	-	-	-	1	1	0	60
Kisk_031	-	-	-	-	-	-	-	-	-	-	1	1	0	60
BaRo_079	-	-	-	-	-	-	-	-	-	-	1	1	0	60
BuCa_142	0	0	0	0	1	0	0	0	0	0	1	1	0	60
BuCa_003	0	0	0	0	1	0	0	0	0	0	1	1	0	60
BuCa_113	-	-	-	-	-	-	-	-	-	-	1	1	0	60
BuAs_178	-	-	-	-	-	-	-	-	-	-	1	1	0	60
BuAs_095											NA	NA		NA
BuCa_060	-	-	-	1	0	0	0	0	0	0	-	-	-	52
Kisk_015	0	0	-	1	0	0	0	0	0	0	0	0	0	52

^a Az SNP multiplex jelzése

^b Az Underhill-féle marker (M) száma (Underhill et al., 2000)

- = nincs vizsgálat

NA = eredménytelen vizsgálat

Hcs = haplocsoport

3.4. Az öt populációs minta igazságügyi genetikai analízisének populációgenetikai és populációstatisztikai eredményei

Az igazságügyi genetikai célú populációs felmérésünk során öt populációs mintában összesen 754 személy STR-alapú genotipizálását végeztük el. A genetikai analízishez használt STR lokuszok körét az évek során úgy bővítettük ki, hogy mindegyik személy legalább 13 lokuszon, az ún. CODIS törzs-markereken meg legyen tipizálva. A populációs mintákban lévő, összesen 299 férfi kilenc Y-STR lokuszra épülő haplotipizálását is végrehajtottuk. Mindezek mellett a vizsgálatok egy korai szakaszában négy másik mikroszatellita (LPL, F13B, FES, F13A01) elemzését is elvégeztük a BuCa populációs minta egészen és a BaRo mintának a HPRTB-vizsgálat kivitelezése idejében rendelkezésre álló részén. Felmérésünk ugyancsak korai szakaszában, a HFV triplex PCR-kitnek köszönhetően lehetőségünk nyílt a BuCa populációs mintánkat egy X-kromoszómához kapcsolt markerrel, a HPRTB lokusszal is analizálni.

Az értekezésemben tárgyalt öt magyarországi populációs minta igazságügyi aspektusú populációgenetikai és populációstatisztikai vizsgálatának eredményeit két alfejezetben mutatom be. A 3.4.1. alfejezet tárgyalja az autoszómás kromoszómákhoz és

az X-kromoszómához kapcsolt STR lokuszokat, a 3.4.2 alfejezet pedig az Y-kromoszóma nem rekombinálódó részén lévő mikroszatellitákkal foglalkozik.

A statisztikai tesztek eredményeit bemutató összesítő táblázatokban a szignifikáns eltérést ($0,01 \leq P < 0,05$) világosabb (szürke, narancssárga, világoszöld) színnel, a nagyon szignifikáns eltérést ($0,01 > P$) sötétebb (fekete, vörös, sötétzöld) színnel jelöltem

3.4.1. Autoszómás lokuszok és az X-kapcsolt HPRTB lokusz vizsgálatának eredményei

3.4.1.1. A genotipizálás eredményei, konkordancia-vizsgálat, allélgyakorisági értékek és a lokuszok polimorfizmusának foka

A doktori értekezés terjedelmi korlátai miatt az öt populációs minta 754 személyének autoszómás STR profilját – a BuCa minta esetében a HPRTB lokusszal történt kibővítéssel – az értekezésem nyomtatott formájában nem közlöm. A DNS-profilokat tartalmazó öt táblázat az értekezésemhez csatolt CD-rom *Autoszómás STR profilok.xls* nevű Excel fájlban található meg. A táblázatban minden személy esetében közlöm a vizsgált személy genetikai nemét, amely az Amelogenin lokusz PCR-fragmens analízisén alapult. A genetikai szex és a fenotípusos nem között eltérést nem tapasztaltunk.

Az öt populációs mintában a mikroszatellita lokuszok polimorf repeat régiójának száma alapján, a mikrovariánsokkal együtt összesen 49 alléltípust különböztettünk meg. A legtöbb alléltípust, 19 db-ot az FGA lokuszon detektáltuk, a legkevesebbet, ötöt a TPOX markeren észleltünk. Az allélok repeat száma 4-35 között mozgott. [Az F13A01 lokusz 3.2-es allélja valójában négy teljes polimorf repeat-et tartalmaz, de a flanking régióban bekövetkezett 2 bp-t érintő deléciónak miatt az elektroforézis során, mint mikrovariáns megkülönböztethető a 4-es alléltól (Puers és mtsai, 1994). Az allél jellemzésekor a kutatók 3.2-es típusjelzést adtak ennek az allélnak és ezt a jelölést a későbbiekben is megtartották.] A megfigyelt alléltípusok 39%-a (19 db) interallél volt. Interallél alatt ebben az esetben azokat az allélvariánsokat értjük, amelyek elektroforetikus mobilitása különbözött az egész számú repeat-egységet tartalmazó allélok mozgékonyaságától. A legtöbb interallélt, 8

db-ot a D21S11 markernél figyeltük meg. Interallél(oka)t találtunk még a következő hat lokuszon: FGA, F13A01, D7S820, TH01, D18S51, HPR1B.

Értekezésem Eredmények és értékelésük fejezetének elején bemutatott 1. táblázatból kiolvasható, hogy a multiplex PCR-kitek egymást követő laboratóriumi alkalmazásával és a kitekkel végzett populációs analízissel lehetőség nyílt arra, hogy öt autoszómás STR lokuszt különböző – legalább két – mikroszatellita genotipizáló rendszerrel is, ugyanazon személyeken jellemezhesünk (Egyed és mtsai, 2000). E kisebb fajta konkordancia-vizsgálatunk felépítése a 6. táblázatban látható. A konkordancia-elemzés során alkalmazott multiplex PCR-kitek felhasználói kézikönyvében leírt (PCR-sokszorozott) lokusz-méretük alapján egyértelmű volt, hogy a HFV-FFFL kit-pár kivételével az összehasonlításra kerülő STR tipizálási rendszerekben szekvenciájukban legalább részben eltérő primerek szerepeltek. A konkordancia-vizsgálatunk eredményeképpen nem észleltünk eltérő genotípust egyik személy esetében sem, tehát a genetikai analízis végeredménye nem függött az alkalmazott genotipizálási protokolltól.

6. táblázat A BuCa és BaRo populációs minta STR tipizálásának konkordancia-vizsgálata (A táblázat rövidítéseinek jelentését lásd az 1. táblázatban ill. a Rövidítések jegyzékében)

	BuCa (N = 233)	BaRo (N = 135)
VWA	V(s)-HFV-P+ ^a	CTTV-P+ ^a
TH01	T(s)-CTT-Cof ^a	CTTV-Cof ^a
TPOX	CTT-Cof ^a	CTTV-Cof ^a
CSFIPO	CTT-Cof ^a	CTTV-Cof ^a
FES	HFV-FFFL	-

^a (Részben) eltérő primerek

V(s), T(s) = VWA ill. TH01 szimplex

A populációs minták vizsgált lokuszain megfigyelt allélgyakorisági értékeket, és a markerek populációgenetikai alapértékeit a *Függelék* F.1-4. táblázatai mutatják be. A könnyebb vizuális összehasonlíthatóság érdekében az egyes lokuszokra vonatkozó allélgyakoriságokat diagramok formájában is feltüntettem (F.8-25. ábrák) Az allélfrekvencia-értékeket és populációgenetikai alapadatokat a CD-rom *allélgyakoriság.xls* Excel-fájl tartalmazza digitálisan szerkeszthető formában.

Felmérésünk során a leggyakoribb allélnak a TPOX 8 bizonyult, amelynek frekvenciája a budapesti referencia mintában csaknem elérte a 60%-ot (F.9. ábra) (Füredi

és mtsai, 1996). E magas érték miatt a TPOX lokusz polimorfizmusának foka várhatóan a legkisebb a lokuszok között. A populációs felmérésünk alapján az allélok a legkisebb maximális frekvencia-értékkel a DeRo személyek FGA lokuszán oszlottak meg, ahol a leggyakoribb, 22-es allél megfigyelt gyakorisága alig 16% volt (F.11. ábra, F.1. táblázat). Az öt populációs mintát együtt vizsgálva a lokuszok közül ugyancsak az FGA marker mutatta a legkisebb maximális allélgyakorisági értéket (27,0%). Ebből a szempontból az FGA lokuszt szorosán követte a D18S51 mikroszatellita a maga 27,7%-os értékével. Ezek az eredmények azt sejtetik, hogy a magyar népességben a DNS-alapú személyazonosításhoz a vizsgált 18 STR marker közül az FGA és a D18S51 lokusz bír a legnagyobb igazságügyi megkülönböztető ill. individualizáló erővel.

Az öt populációs minta allélgyakoriság-eloszlása szemmel láthatóan kisebb-nagyobb mértékben különbözött egymástól. Roma-specifikus eloszlási görbét nem figyeltünk meg, kivéve talán csak a D5S818 és a TH01 lokuszok valamint a D21S11 lokusz 32.2-es allélja esetében. RXC kontingencia táblák komputerezált G-statisztikája alapján a BuCa mintán felmért X-kapcsolt HPRTB lokusz allélgyakoriság-eloszlása nem mutatott szignifikáns különbséget ($P = 0,95$) a két nem között (Füredi és mtsai, 1996).

Az értekezésemben vizsgált 17 autoszómás és egy X-kapcsolt lokusz minimális allélgyakoriságát és polimorfizmusának fokát kifejező jellemzők szélső értékeit a 7. táblázatban foglaltam össze. A szélsőértékek mellett a táblázat azt is bemutatja, hogy az adott legkisebbnek vagy legnagyobbban bizonyult értéket mely populációs minta mely lokuszán figyeltük meg

7. táblázat Az öt populációs mintán analizált, 18 STR lokuszra vonatkozó minimális allélgyakoriságok és polimorfizmus-fokok szélsőértékei

		Minimum	Maximum
PD	Érték	0,7590	0,9740
	Populáció	DeRo	BuCa
	Lokusz	TPOX	D18S51
PE	Érték	0,3325	0,7640
	Populáció	DeRo	BuCa
	Lokusz	TPOX	D18S51
PIC	Érték	0,5286	0,8722
	Populáció	DeRo	BuCa
	Lokusz	TPOX	D18S51
H_{obs}	Érték	0,6054	0,9000
	Populáció	BuCa	DeRo
	Lokusz	TPOX	D18S51
H_{exp}	Érték	0,5880	0,8856
	Populáció	BuCa	BuCa
	Lokusz	TPOX	D18S51
P_{min}	Érték	0,0107	0,0752
	Populáció	BuCa	BuCa
	Lokusz	TPOX	D18S51

A 7. táblázatból egyértelműen kitűnik, hogy minden esetben a legkisebb értéket a TPOX, a legnagyobb értéket pedig a D18S51 lokusznál tapasztaltuk. (Mivel az X-komoszóma a férfiakban haploid, ezért az összehasonlításnál a BuCa férfiakra számított adatokat nem vettük figyelembe.) Az igazságügyi személyazonosítási hatékonyságot (PD, PE, PIC) kifejező értékek esetében a minimumot mindig a DeRo mintában, a maximum értéket pedig mindig a BuCa mintában kaptuk. A minimális allélfrekvencia 1,1-7,5% között mozgott. A viszonylag magas, 7%-os érték a kiskun adatbázis kis mintaszámából fakadt. Ha a kiskun mintától eltekintünk, akkor a populációs adatbázisunk allélgyakorisági értékei 1-2% fölött 95%-os konfidenciával megbízhatónak bizonyultak.

Az STR markerek igazságügyi személyazonosítási hatékonyságát az egyes multiplex PCR-alapú genetikai analizáló rendszerekre összesítve is kifejeztük (Woller és mtsai, 1997; Füredi és mtsai, 1997a, b, 1998; Egyed és mtsai, 2000). Természetesen az individualizáció hatékonysága a szimultán vizsgált STR lokuszok számával arányosan növekedett. A hatékonyság növekedését jól illusztrálja a következő PD adatsor, amelyet a BuCa populációs mintára számítottunk ki: VWA szimplex (1 lokusz): 92,6%, CTTV-kvadruplex (4 lokusz): 99,98%, Profiler Plus multiplex (9 lokusz): 0,99999999999904. A

rendelkezésünkre álló öt magyarországi populációs minta széleskörű igazságügyi genetikai STR analízise után kíváncsiak voltunk arra, hogy van-e lényeges különbség az individualizáció várt hatékonyságában az egyes populációs minták között. Ennek érdekében a pM és PE értékeket minden populációs minta esetében a 13 CODIS törzslokuszra kombináltuk. A kapott eredményeket a 8. táblázat mutatja be. A táblázatban látható, hogy a 13 STR lokusból álló CODIS-rendszerrel az igazságügyi személyazonosítás hatékonysága mindegyik populációban rendkívül magasnak és egymással közel azonos nagyságrendűnek bizonyult.

8. táblázat Az öt populációs mintán analizált, 13 CODIS STR törzslokuszra vonatkozó kombinált PD és PE értékek

	pM_{komb}	PE_{komb}
BuCa	2×10^{-15}	0,999994
BaRo	2×10^{-14}	0,999983
DeRo	9×10^{-15}	0,999991
BuAs	6×10^{-15}	0,999991
Kisk	1×10^{-13}	0,999986

3.4.1.2. A Hardy-Weinberg és kapcsoltsági (linkage) egyensúlyi vizsgálatok eredményei

Az allélok lokuszon belül és lokuszok között esetlegesen meglévő asszociációjának detektálására a teljes lokusz-készleten az *Arlequin* szoftverrel statisztikai vizsgálatokat végeztünk. A lokuszok közötti asszociáció esetében a lokuszpárokat elemeztük. Az *Arlequin* programmal kapott eredményeket a *Függelék F.5.* táblázatában összesítettük. A tesztelésbe bevont allélok száma (n) az egyes lokuszokra vonatkozóan az F.1-4. táblázatokban olvasható le. Az F.5. táblázat a vizsgált 17 autoszómás STR marker allélikus asszociációjára irányuló analízis eredményeit mutatja be. A mindkét teszt esetében azt ábrázoltuk, hogy a mért valószínűségi értékek kisebbek-e a felállított 5%-os ill. 1%-os szignifikancia-szintnél. A HPRTB lokuszra vonatkozó HWE-tesztelés eredményét a táblázat nem tartalmazza. A BuCa populációs mintát alkotó nők HRTB-genotípusainak megfigyelt gyakorisága az exact teszt alapján megfelelt a HWE alapján várt értékeknek ($P = 0.390$, Füredi és mtsai, 1996).

Az Arlequin szoftverbe épített Guo és Thompson- (1992) féle exact teszt hat lokuszon a HW egyensúlytól statisztikailag szignifikáns eltérést mutatott ki a két roma populációs minta esetében (Füredi és mtsai, 1998; Egyed és mtsai, 2000; nem publikált eredmények). Az érintett lokuszok a következők voltak:

- a BaRo minta esetében: TPOX, D3S1358, CSF1PO, D7S820, LPL
- a DeRo minta esetében: D13S317

Mind a hat lokuszon a HWE-től való eltérés nem volt jelentősen szignifikáns ($0,01 \leq P < 0,05$).

A VWA, TH01, TPOX és CSF1PO markerek esetében a HWE-tesztelést korábban a BaRo populációs minta egy kisebb almintáján, 135 személy bevonásával végeztük el (Füredi és mtsai, 1997b). Míg az almintán végrehajtott teszt eredménye mind a négy lokusz esetében kevéssel bár, de meghaladta az 5%-os szignifikancia-szintet, addig a teljes (206 személyből álló) minta analízise szignifikáns eltérést adott a TPOX és a CSF1PO lokuszra. A HWE nullhipotézis tesztelt valószínűségének a populációs minta méretének növelésével együtt járó csökkenését a populáció heterogén volta (Wahlund-hatás) és/vagy a mintavételi bizonytalanság (sampling effect) okozhatta (Evett és Weir, 1998).

A két roma populációs minta exact teszttel történő HWE elemzését a Bonferroni-eljárás (Weir, 1996) alapján, az $\alpha' = 0,05$ kísérleti hibaráttával is elvégeztük. Ebben az esetben a BaRo mintára (17 teszt) a tesztsorozat egyedi szignifikancia-szintje $\alpha = 0,0029$ volt, míg a DeRo mintára (13 teszt) az $\alpha = 0,0038$ értéket kapott. Az exact teszt során a TPOX, D3S1358, CSF1PO, D7S820, LPL, D13S317 lokuszok mindegyikére $P > 0,01$ valószínűségi értéket kaptunk, ami meghaladta a tesztsorozatok egyedi szignifikancia-szintjeit. Ennek alapján egyik roma mintára sem vethetjük a HWE hipotézist, ha a HW-egyensúlyt a populációk teljes genomi szintjén vizsgáljuk.

Populációs felmérésünk korai szakaszában az *Arlequin* szoftver EM-algortmuson alapuló valószínűségi hányados tesztjével egy lehetséges allélikus asszociációt detektáltunk a BaRo mintában a FES és az F13A01 lokusz között (Füredi és mtsai, 1998). Későbbiekben a kapcsoltsági egyensúly EM-alapú analízisét a teljes autoszómás lokuszkészletre is kiterjesztettük. Az X-kapcsolt HPRTB lokusz LD-analízise az *Arlequin* szoftverrel nem volt lehetséges. Tesztjeinkben lokuszpárokat vizsgáltunk. Az elemzett lokuszok számának megfelelően a BuCa és BaRo mintán 136 tesztet, a többi három adatbázison pedig 78 tesztet végeztünk el (F.5. táblázat). A szignifikáns eltérést mutató tesztek aránya az egyes populációs mintákban a következőképpen alakult: BuCa = 9/136

(6,6%), BaRo = 30/136 (22,1%), DeRo = 10/78 (12,8%), BuAs = 6/78 (7,7%), Kisk = 9/78 (11,5%). A legkisebb értéket a budapesti referencia populációs mintában, a legnagyobb arányt a baranyai roma mintában figyeltünk meg. A szignifikáns eltérések aránya csak a BuCa és a BuAs adatbázisban maradt 10% alatt. Az összes szignifikáns eltéréshez hasonlóan a nagyon szignifikáns deviációk aránya is a BuCa mintában bizonyult a legmagasabbnak: 19/136 (14,0%).

A Bonferroni-eljárással az LD tesztsorozatok egyedi szignifikancia-szintjét a kísérleti hibaráta 5%-on való tartása érdekében a BuCa, BaRo minta (136-136 teszt) esetében $\alpha = 0,0004$ értékre, a maradék három adatbázis esetében (78-78 teszt) $\alpha = 0,0006$ értékre korrigáltuk. A BuCa, DeRo, BuAs, Kisk populációs mintában a valószínűségi értékek minden tesztben meghaladták az egyedi szignifikancia-szintet. A baranyai roma populáció esetében azonban hét olyan lokuszpár is volt, ahol az egyedi szignifikancia-szintnél kisebb valószínűséget mértünk. Ennek alapján megalapozottnak tűnik az a feltételezés, hogy a baranyai roma populációs mintában genom szinten jelentős linkage disequilibrium van.

A baranyai roma LD-eredmények kiértékelése során azonban figyelembe kell venni a következőket. A kapcsoltsági egyensúly EM-algoritmuson alapuló, genotípusos adatokra vonatkozó valószínűségi hányados tesztelése abból a feltételezésből indul ki, hogy a gamétafázisukban nem ismert genotípusok a HW-egyensúlynak megfelelően oszlanak el a populációban (Schneider és mtsai, 2000). Ennek alapján az LD-teszt akkor is elvetheti az allélok lokuszok közötti függetlenségének hipotézisét, ha a populáció a vizsgált lokuszokon jelentősen eltér HW-egyensúlytól (Excoffier és Slatkin, 1998). A BaRo mintában az exact teszt alapján öt lokuszon figyeltünk meg jelentős deviációt a Hardy-Weinberg egyensúlytól. A BaRo adatbázis LD-analízisében kapott 30 szignifikáns eltérés közül 13 deviációt olyan lokuszok között észleltünk, amelyek közül legalább az egyik az öt HW disequilibriumot mutató lokusz közül került ki. Ezekben az esetekben elképzelhető, hogy a lokuszokon belül megfigyelt allélikus asszociáció a statisztika torzításával hozzájárult a LE hipotézis téves elvetéséhez.

Az STR-alapú populációs felméréseink kezdeti szakaszában az LE meglétét az „interclass” korrelációs analízissel valamint a Chakraborty-féle (1984) s_k^2 kritériummal is teszteltük (Füredi és mtsai, 1996, 1997a, b). A vizsgálatokat a BuCa populáción genotipizált nyolc lokuszra (VWA, TH01, TPOX, CSF1PO, LPL, FES, F13B, F13A01) és a BaRo mintán vizsgált négy lokuszra (VWA, TH01, TPOX, CSF1PO) végeztük el. Az elemzés eredményeképpen egyik teszt sem mutatott szignifikáns allélikus asszociációt

sem a lokuszpárokra, sem a teljes lokuszkészletre vonatkozóan. Ez az eredmény részben ellentétes az fentiekben leírt, EM-alapú LD-tesztelés során kapott eredményekkel, mivel abban az esetben a BuCa minta TH01-CSF1PO lokuszpárjában és a BaRo minta TH01-TPOX, VWA-TPOX, TH01-CSF1PO lokuszpárjaiban a teszt elvetette a kapcsoltsági egyensúlyi hipotézist (F.5. táblázat). Az alkalmazott statisztikai tesztek eredményei között tapasztalt eltérést a TPOX és CSF1PO lokuszok esetében az EM-alapú LD-analízis előbbiekben vázolt HWE érzékenysége is okozhatta, de a kérdés tisztázásához további elemzésekre van szükség.

3.4.1.3. Az elemzett populációk genetikai struktúrája

Az értekezésemben tárgyalt öt populációs minta genetikai homogenitását ill. tagoltságát három különböző statisztikával igyekeztünk meghatározni. A G-statisztika és a Wright-féle F-statisztika fenotípusos szinten, az allélgyakoriságok felhasználásával képez különböző értékeket és valószínűségeket. Az AMOVA (molekuláris varianciaanalízis) a Wright-féle F-statisztika molekuláris megfelelője (Excoffier és mtai, 1992). Az AMOVA a mikroszatelliták esetén, a stepwise mutációs modell (Ota és Kimura, 1973; Valdes és mtai, 1993) alkalmazásával az egyes allélok közötti távolságot molekuláris szinten, az STR lokuszt alkotó polimorf repeatek száma alapján határozza meg.

Korábbi publikációinkban bemutattuk a budapesti referencia populáción és a baranyai roma mintán elvégzett összehasonlító vizsgálatok eredményeit a D16S539 lokusz kivételével az összes, értekezésemben tárgyalt autoszómás lokuszra (Füredi és mtai, 1996, 1997b, 1998; Egyed és mtai, 2000). Az elemzések további három populációs adatbázissal való kibővítésének összesített eredményét a *Függelék* F.6. táblázata tartalmazza. Az eredményeket lokuszpárokon belül az egyes lokuszokra vonatkozóan tüntettem fel. A G-statisztika esetében csak azt ábrázoltam, hogy a mért valószínűségi értékek kisebbek-e a felállított 5%-os ill. 1%-os szignifikancia-szintnél. Az F-statisztika és AMOVA esetében a közös leszármazási együttható [$F_{ST}(\theta)$] ill. a Φ_{ST} értékének nagyságrendjét is feltüntettem a *b*, és *c*, táblázatok alsó háromszögeiben.

Az F.6. táblázatból az első pillantásra megállapítható, hogy egymáshoz viszonyítva az öt populációs minta genetikai összetételében, a lokuszokon mért allélgyakoriságok és allél-struktúrák eloszlása tekintetében szignifikáns különbségek vannak. Mindhárom tesztben a legtöbb szignifikáns eltérést a BuCa-BaRo és a BaRo-BuAs populáció-páros

adta. A G- és F-statisztika alapján figyelemre méltó volt a roma populációk egymáshoz viszonyított nagyfokú heterogenitása. A roma populációknak az allélgyakoriságok szintjén tapasztalt heterogenitását az allélok molekuláris szerkezetén alapuló AMOVA csak csekély mértékben támasztotta alá, mivel a két populációban megállapított Φ_{ST} mindössze három lokuszon jelzett szignifikáns eltérést. A populáció-párok közül a legkisebb genetikai tagozódást a BuCa-Kisk minta mutatta, mivel valamennyi teszt csak két lokuszon jelzett jelentős deviációt. A kelet-magyarországi roma populáció és a budapesti referencia-populációk allélgyakoriság-eloszlásban sok, összesen nyolc lokuszon tapasztaltunk inhomogenitást. Ennél kisebb genetikai tagolódást észleltünk a két minta között, amikor a közös leszármazási együtthatókat elemeztük. A két budapesti populációs adatbázis között szintén deviációt mutattunk ki, de ez az eltérés tesztenként legfeljebb három vagy négy lokuszra korlátozódott.

A populációk összehasonlítása során megvizsgáltuk azt is, hogy mely lokuszok bírnak a legkevesebb és a legtöbb varianciával a populáció-párok esetében. Mivel ebből a szempontból nagyon sok egymáshoz igen hasonló értéket kaptunk, ezért csak pár jellemző vonást ill. szélsőértéket emeltünk ki. A vizsgált 17 autoszómás STR között nem volt egy olyan marker sem, amely legalább egy tesztben ne adott volna szignifikáns eltérést. Azon 13 lokusz közül, amelyet mindegyik populációs mintán elemeztünk, a legkisebb változatosságot, a TH01 lokuszon tapasztaltunk. A TH01 marker analízise során az összesen lehetséges 10-10 db párosításból 7, 8 ill. 10 alkalommal nem észleltünk szignifikáns eltérést. A legnagyobb varianciát a 13 CODIS STR törzs-lokuszt közül a D21S11 lokuszon detektáltuk, amely 8, 8 és 7 db tesztben mutatott deviációt.

A populációk páronkénti összehasonlítását a 13 CODIS lokusz együttes bevonásával is elvégeztük. A 9. táblázat alsó háromszöge tartalmazza az F_{ST} - és Φ_{ST} -alapú genetikai távolságmátrixot, míg a táblázat felső háromszögében Φ -statisztika szignifikanciáját jelöltük. Az F-statisztika valószínűségi értékei az *Arlequin* szoftverrel ebben az esetben nem voltak meghatározhatók. Az AMOVA vizsgálat nem mutatott jelentős különbséget a BuCa-Kisk párban, valamint a roma populációk között. A többi párosításban azonban a molekuláris varianciaanalízis szignifikáns különbséget jelzett. Három populáció-párban, nevezetesen a BuCa-BaRo, BaRo-BuAs, BaRo-Kisk mintapárban a 13 lokuszra vonatkoztatott F-statisztikai értékek meghaladták a 0,02 szintet. A közelmúltban az USA ill. az európai népességben elvégzett igazságügyi populációs felmérések eredményei alapján a 0,02 szint meglehetősen magasnak számít (Budowle és mtsai, 2001; Budowle és

Chakraborty, 2001). Mivel a felmérések során nagyobb F_{ST} -értéket elsősorban izolált ill. eltérő genetikai eredetű népcsoportokban detektáltak, ezért megfigyeléseinket szintén a BuCa, BaRo, BuAs, Kisk populáció eltérő genetikai múltja ill. izoláltsága magyarázhatja.

9. táblázat Az öt vizsgált magyarországi populációs minta páronkénti összehasonlítása a 13 CODIS STR törzs-lokusz bevonásával

<i>P</i> Érték	BuCa		BaRo		DeRo		BuAs		Kisk	
	F_{ST}	Φ_{ST}	F_{ST}	Φ_{ST}	F_{ST}	Φ_{ST}	F_{ST}	Φ_{ST}	F_{ST}	Φ_{ST}
BuCa			-	$< 10^{-4}$	-	0,011	-	$< 10^{-4}$	-	0,118
BaRo	0,022	0,015			-	0,157	-	$< 10^{-4}$	-	$< 10^{-4}$
DeRo	0,004	0,004	0,012	0,001			-	$< 10^{-4}$	-	0,040
BuAs	0,002	0,008	0,025	0,027	0,006	0,016			-	0,021
Kisk	0,004	0,004	0,022	0,017	0,011	0,008	0,010	0,009		

Annak érdekében, hogy a az öt vizsgált populációs adatbázis további, igazságügyi aspektusú összehasonlítását elvégezhessük (sörét- és Tippett-diagramok, lásd később), meghatároztuk a teljes öt adatbázis közös leszármazási (coancestry) és a populáción belüli beltenyésztési (inbreeding) együtthatóját [F_{ST} (θ) és F_{IS}] a 13 CODIS STR lokuszon. A megállapított F-statisztika értékek a 10. táblázatban láthatóak.

10. táblázat Az öt vizsgált magyarországi populációs minta 13 CODIS STR törzs-lokuszra kalkulált F_{ST} (θ) és F_{IS} értékei

	TPOX	D3S1358	FGA	D5S818	CSF1PO	D7S820	D8S1179	TH01	VWA	D13S317	D16S539	D18S51	D21S11
F_{ST}	0,021	0,014	0,012	0,010	0,003	0,033	0,008	0,006	0,008	0,016	0,010	0,017	0,012
F_{IS}	$0 \geq$	0,003	0,002	$0 \geq$	$0 \geq$	0,062	0,024	0,007	$0 \geq$	0,040	0,012	0,021	0,008

A 10. táblázat F_{ST} -értékei alapján az öt vizsgált populációs mintában a 13 CODIS STR marker közül a D7S820 lokusz mutatta a legnagyobb variabilitást a populációk között. Ugyancsak magas variációt kaptunk a TPOX lokusz esetében. A CSF1PO mikroszatellita mutatta a legkisebb változékonyságot az adatbázisok között. Ez utóbbi megfigyelés összhangban van korábbi felmérések CSF1PO lokuszra vonatkozó eredményével (Deka és mtsai, 1995; Budowle és mtsai, 2001; Budowle és Chakraborty, 2001). A CSF1PO mikroszatellita az 5. kromoszóma rövid karján lévő növekedési faktor és növekedésifaktor-receptor génklaszter részét képezi. E génklaszter a humán populációkban szelekciós nyomás alatt állhat. Az alacsony variabilitás annak a bizonyítéka lehet, hogy a CSF1PO lokuszon a szelekció kiegyensúlyozó (stabilizáló)

hatású (Deka és mtsai, 1995). A 10. táblázat F_{IS} adatai arra engednek következtetni, hogy a populációkon belüli beltenyésztés foka nem teljesen elhanyagolható. A magasabb F_{IS} értéket elsősorban az egyéneknek a populáción belüli nem véletlenszerű párválasztása okozhatta (Evett és Weir, 1998).

Az értekezésemben tárgyalt populációk allélgyakoriság-eloszlásának összehasonlítására irányuló vizsgálataink utolsó lépésében a budapesti referencia mintánk (BuCa) allélgyakorisági értékeit más külföldi, kaukázusi referencia populáció hasonló adatával vetettük össze (Füredi és mtsai, 1996; Egyed és mtsai 2000). Mivel a publikált eredményeink öt és nyolc lokuszra valamint négy külföldi populációs mintára vonatkoznak, ezért ez az összevetés korántsem nevezhető teljesnek. Ugyanakkor úgy gondoltuk, hogy e vizsgálat konklúziója is segít alátámasztani a budapesti referencia adatbázisunk igazságügyi alkalmazásának hitelességét. A publikált eredményeket táblázatos formában gyűjtöttük össze (11. táblázat). Egy szlovén, egy olasz és két spanyol populációs mintával való összehasonlításban a budapesti referencia populáció csak az olasz mintától mutatott szignifikáns eltérést. A megfigyelt négy deviáció közül csak a D8S1179 markeren detektáltunk statisztikailag nagyon jelentős eltérést. A budapesti referencia populációs mintánk allélgyakoriság-eloszlása tehát nagyfokú hasonlóságot mutatott más európai népcsoportokkal. Ezek az eredmények egybevágtak azzal a tudományos nézettel, amely szerint a kevert magyar népesség az indoeurópai népekkel közeli genetikai rokonságban áll (Czeizel és mtsai, 1991; Guglielmino és Béres, 1996; Guglielmino és mtsai, 2000).

11. táblázat A budapesti referencia populációs adatbázis (BuCa) és négy külföldi kaukázusi populációs minta G-statisztikán alapuló összehasonlító elemzése. A táblázatban szereplő értékek valószínűségi (P) értékek

	TPOX	D3S1358	FGA	D5S818	CSF1PO	D7S820	D8S1179	TH01	VWA	D13S317	FES	D18S51	D21S11
szlovén ^a	-	NS	NS	NS	-	NS	NS	NS	NS	NS	-	NS	NS
olasz ^b	-	NS	0,034	NS	-	NS	$< 10^{-3}$	NS	NS	NS	-	0,010	0,015
spanyol 1 ^c	-	NS	NS	NS	-	NS	NS	NS	NS	NS	-	NS	NS
spanyol 2 ^d	NS	-	-	-	NS	-	-	NS	NS	-	NS	-	-

NS = statisztikailag nem szignifikáns

- = nincs vizsgálat

^a Zupanic et al., 1998

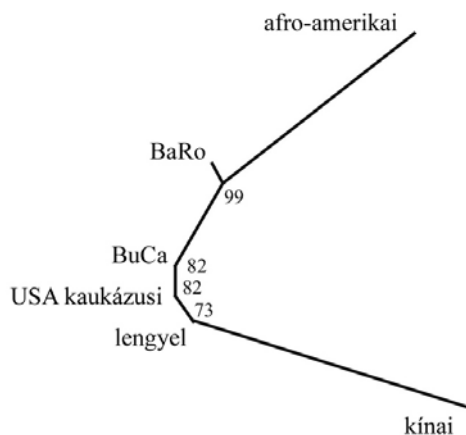
^b Garofano et al., 1998

^c Entrala et al., 1998

^d Martín et al., 1995

A budapesti referencia és a baranyai roma populációs mintánk három nagy népességgel, a kaukázusi, az afro-amerikai, és az ázsiai embercsoporttal való lehetséges genetikai kapcsolatát is elemeztük (Füredi és mtsai, 1998). A vizsgálatba bevont hazai és külföldi mintákon nyolc STR lokusz (VWA, TH01, TPOX, CSF1PO, LPL, F13B, FES, F13A01) allélgyakorisági értékének felhasználásával a Nei-féle genetikai távolságon alapuló távolság-mátrixot szerkesztettünk. A kapott távolságmátrix segítségével, a *PHYLIP* szoftver alkalmazásával gyökértelen szomszéd-kapcsolású (unrooted neighbor-joining) filogenetikai fát szerkesztettünk (12. ábra). A fán az USA kaukázusi (Schumm és mtsai, 1997) és a lengyel populáció (Miścicka-Śliwka és mtsai, 1998) a budapesti referencia populációval együtt egy csoportot alkotott. Ehhez a csoporthoz a baranyai roma populáció távolabbról kapcsolódott. Az afro-amerikai (Schumm és mtsai, 1997) és a kínai populáció (Lee és mtsai, 1997) hosszú ágak mentén élesen különvált a kaukázusi csoporttól. A filogenetikai eredményeink azt mutatták, hogy a vizsgált kaukázusi populáció-csoport és a baranyai roma populáció között viszonylag távolabbi a genetikai rokonság.

12. ábra A Nei-féle genetikai távolságon alapuló „unrooted neighbor-joining” fa. Az elágazási pontoknál feltüntetett számok 1000 replikáció bootstrap értékét jelölik (Füredi és mtsai, 1998)



- afro-amerikai és USA kaukázusi minta: Schumm és mtsai, 1997
- lengyel minta: Miścicka-Śliwka és mtsai, 1998
- kínai minta: Lee és mtsai, 1997

3.4.1.4. A magyarországi populációs adatbázisok igazságügyi statisztikai célú alkalmazhatóságának elemzése

Az igazságügyi genetikai szakértői vizsgálatok egyik legsarkalatosabb kérdése, hogy a tettes által hátrahagyott biológiai nyom mekkora bizonyító erővel bír a vádlott bűnösségére nézve. Ha a biológiai nyom és a vádlott DNS-profilja különbözik, akkor a szakértő általában könnyű helyzetben van – a nyom nem a vádlottól származik. A biológiai anyagmaradvány és a vádlott DNS-profiljának teljes egyezése azonban az értekezésem *Bevezetés* fejezetében vázoltak alapján sokkal komplexebb helyzetet teremt. A genetikai bizonyíték Bayes-elven nyugvó statisztikai interpretációjához meg kell becsülni azt, hogy mekkora a valószínűsége annak, hogy a tettes biológiai nyomának DNS-profilja és a védelem szerint ártatlan vádlott genotípusa között detektált azonosság csupán a véletlennek köszönhető. Az egyezési valószínűség kiszámításához meg kell becsülni a kérdéses DNS-profil népességben való előfordulási gyakoriságát (Evett és Weir, 1998). Ebben a lépésben az egyik kulcskérdést az jelenti, hogy a DNS-profil frekvenciáját és így az egyezési valószínűség értékét milyen mértékben befolyásolja az, hogy a statisztikai kiértékeléshez melyik allélgyakorisági adatbázist használjuk.

Az alábbiakban leírt analízisekkel azt a bűnügyi helyzetet szimuláltuk, amikor a tettes és a gyanúsított egy személy, azaz a bűncselekmény során keletkezett biológiai nyom a gyanúsítottól származott (ún. egyszemélyű, „within person” vizsgálatok).

A mikroszatellita-alapú populációs felméréseink korai szakaszában a budapesti referencia és a baranyai roma adatbázist elemeztük az előbbieken vázolt igazságügyi statisztikai szempontból (Füredi és mtsai, 1997b, 1998).

Az átlagolt fenotípus-különbségi értékek (MV) analízise során azt vizsgáltuk, hogy a budapesti referencia adatbázis alkalmazása milyen hatással volt a baranyai roma személyek STR-profil frekvenciájára és vice versa. Itt az *A* populációból kiválasztott profil azt esetet szimulálta, amikor a bűncselekmény során keletkezett biológiai nyom az *A* népességből származó személytől eredt. Az MV-tesztelést négy STR lokuszra külön-külön, illetve a maradék három lokusz egymás utáni folyamatos bevonásával is elvégeztük (12. táblázat)

12. táblázat A BuCa és BaRo populációs minták között mért átagolt fenotípus különbségi értékek. A populációk összehasonlítása két irányból történt. **a**, egy-egy STR marker; **b**, 2-4 marker együtt vizsgálva (Füredi és mtsai, 1998)

a,

Populációk	VWA	TH01	TPOX	CSF1PO
BaRo→BuCa	1,5	1,4	1,5	1,4
BuCa→BaRo	1,5	1,4	1,8	1,2

b.,

Populációk	2 lokusz (VWA, TH01)	3 lokusz (TPOX)	4 lokusz (CSF1PO)
BaRo→BuCa	2,1	3,2	4,6
BuCa→BaRo	2,1	2,8	3,3

Amint a 12. táblázatban látható, minden összehasonlítás során annak az allélgyakorisági adatbázisnak az alkalmazásakor kaptunk átlagosan nagyobb fenotípusgyakoriságot, amelyikből a vizsgált DNS-profil származott. Az eltérés mértéke a lokuszok fokozatos bevonásával kismértékben bár, de növekedett. Ezek az eredmények összhangban álltak más populációs mintákon elvégzett vizsgálatok eredményeivel (Meyer és mtsai, 1995), amelyekből azt a következtetést vontuk le, hogy a magyar népesség genetikai struktúrája az igazságügyi DNS-elemzés bizonyító erejének statisztikai interpretációját befolyásolhatja. Ráadásul, ha nem a tettes alpopulációjából választjuk ki az allélgyakorisági értékeket, akkor átlagosan gyakrabban tévedünk a gyanúsított kárára tévedhetünk, mivel az esetek többségében kisebb DNS-profil gyakoriság miatt az egyezési valószínűség értéke is kisebb lesz. Ez pedig ellentétes a demokratikus jogrendszerek ártatlanság vélelme elvével.

A populációs allélgyakorisági adatbázis típusának a négy lokuszos STR profilgyakorisági értékre kifejtett hatását egy másik vizsgálati rendszerben is elemeztük. Annak érdekében, hogy a bűnügyi szimulációba bevont DNS-profilok száma minél több legyen, a populációs mintákat alkotó személyek DNS-profiljainak egyesítésével egy nagyobb DNS-profil készletet (poolt) hoztunk létre. A profil-készlet kezdetben 358 db STR genotípusból állt, mivel csak a budapesti referencia minta és a baranyai roma populációs minta 135 főből álló almintája állt rendelkezésünkre. 2000-2003. között további három mintabevonásával, és a vizsgált STR markerek számának növelésével sikerült egy 754 személyből álló 13 STR lokuszos DNS-profil készletet létrehozni. A

egyesített profil-állomány vizsgálata során az egyezési valószínűségeket annak reciprokával, a valószínűségi hányadalossal (LR) fejeztük ki.

Már a 358 személyt tartalmazó profil-készlet négy lokuszon történő analízise is azt mutatta, hogy a különböző adatbázisokból számított profilgyakorisági értékek között akár majdnem két nagyságrend különbség is lehet (Füredi és mtsai, 1997b, 1998). A VWA, TH01, TPOX, CSF1PO lokuszok alléljaiból álló DNS-profil készlet esetében pl. ha az 1 : 1 milliónál nagyobb ($\log LR < 6$) egyezési valószínűséget mutató profilokat szemléljük, akkor e profilok 5,3%-a mutatott legalább egy nagyságrendnyi eltérést (Füredi és mtsai, 1997b). Ehhez hasonló eredményt kaptunk, amikor másik négy (LPL, F13B, FES, F13A01) STR lokusszal ismételtük meg ugyanezt az analízist (Füredi és mtsai, 1998). Figyelembe véve Magyarország lakosainak számát, az előbbieket alapján kijelenthetjük, hogy az igazságügyi statisztikai interpretáció számos esetben méltánytalanul elfogult profilgyakoriságot eredményezhet akkor, ha az egyik populáció allélgyakorisági adatait a másik populációra alkalmazzuk.

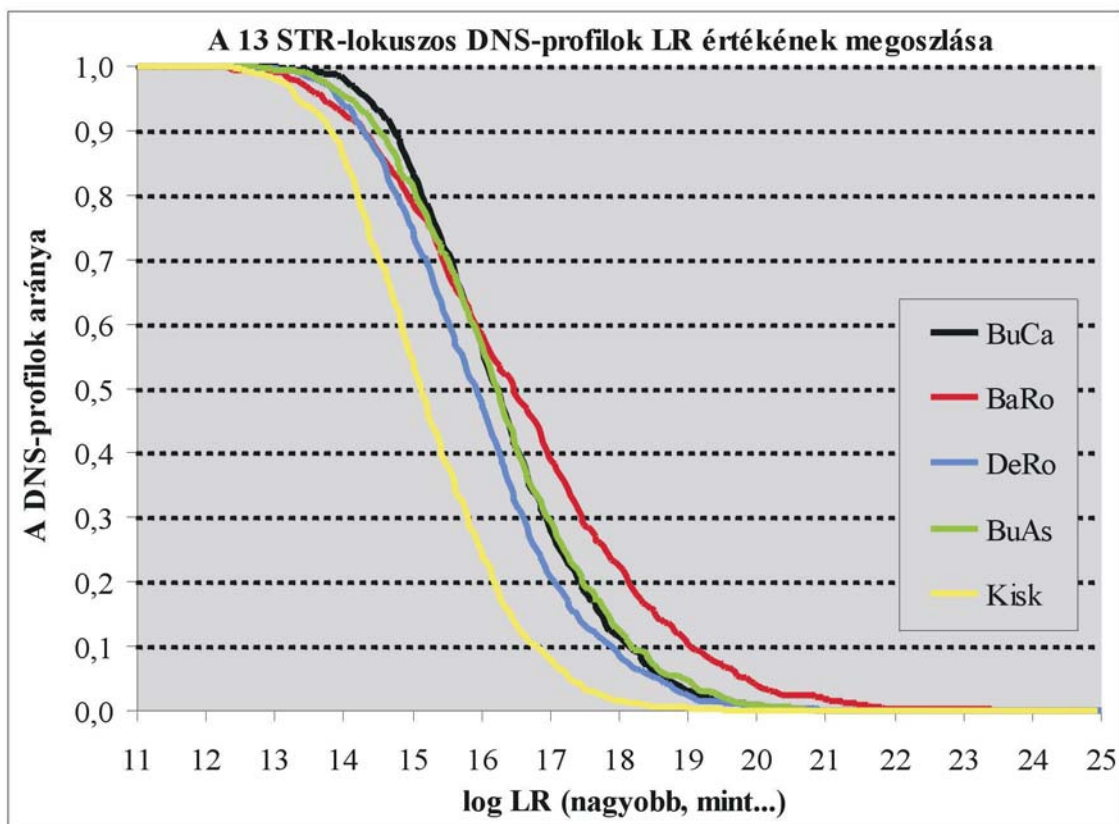
A genetikai személyazonosítási technológia rendkívül gyors fejlődése lehetőséget teremtett arra, hogy Intézetünkben is alkalmazhassunk olyan multiplex PCR-alapú STR vizsgálati rendszereket, amelyek segítségével 6-9 STR lokusz genetikai analízise szimultán elvégezhető. A multiplex PCR-kitek használatával megnövekedett valószínűségi hányados érték lehetővé tette a rutin igazságügyi szakértői vizsgálat bizonyító erejének fokozását. A nagy igazságügyi személyazonosító erővel rendelkező STR rendszerek bevezetésével kíváncsiak voltunk arra, hogy a populációs adatbázis-választás fentiekben kifejtett hatása további három magyarországi populációs minta, és összesen 13 STR lokusz (a CODIS törzs-lokuszek) bevonásával is érvényesül-e.

Az összehasonlító analízis eredményét, az egyes populáció-párookra vonatkozó valószínűségi hányados értékeket a *Függelék* F.26a-j. ábrái tartalmazzák. A valószínűségi hányados számítását ebben az esetben a Balding (1995) által kidolgozott, Bayes-elvű módszerre alapoztuk, amely itt nem tartalmazott olyan korrekciót, amely figyelembe venné az elemzett populációk genetikai strukturáltságát. A F.26. ábra BuCa-BaRo párra vonatkozó diagramja (a) azt mutatta, hogy a populációs adatbázis között négy STR markeren megfigyelt különbség 13 lokusz alkalmazásával tovább nőtt: a különbség sok profil esetében meghaladta a két nagyságrendet. (A diagramokon az 1 nagyságrend különbséget a szaggatott vonalak, a 2 nagyságrend különbséget a pontozott vonalak jelzik.) Ehhez hasonló képet mutatott a baranyai roma adatbázis akkor is, ha azt a többi három (DeRo, BuAs, Kisk) populációval vetettük össze (F26.e-g. diagramok). Itt érdemes külön

is megemlíteni a két roma populációt, amely az egymással való összehasonlításban jelentős szórást mutatott, amely a magyarországi roma népesség jelentősebb tagoltságára utalhat. A diagramok vizuális kiértékelésben a 10 populáció-pár közül a budapesti referencia és a budapesti askenázi minta között észleltük a legkisebb mértékű szóródást. A kiskun minták esetében észrevettük, az LR értékek az átlótól elhajoltak a vizsgált populáció-pár másik tagjának adatbázisa irányába. A jelenség arra utalhat, hogy az LR-számításhoz alkalmazott Balding-féle módszer kis adatbázis esetén igazságügyi értelemben túlságosan konzervatív, azaz túlságosan felülbecsülheti a biológiai nyom és a vádlott DNS-profiljának a véletlenül alapuló egyezési valószínűségét.

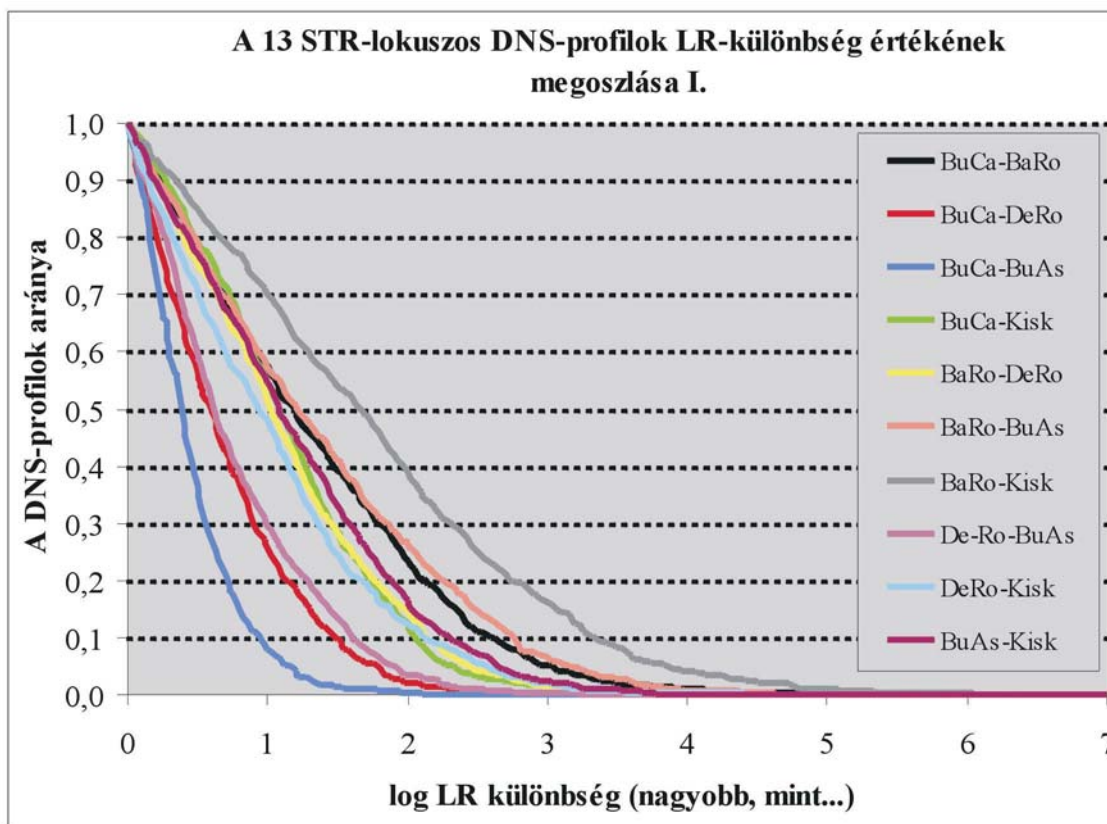
A sötét-diagramokon ábrázolt LR-érték megoszlás pontosabb elemzése érdekében a sörét-analízis során kapott LR-eredményeket az ún. egyszemélyű Tippett-diagramon is ábrázoltuk. A Tippett-diagramon azon profilok részarányát tüntettük fel egy adott pontban, amelyek logaritmus LR értékei meghaladták a pont x-tengely vetületénél feltüntetett értéket. A Balding-módszer által számított, korrigálatlan LR-értékek Tippett-diagramját a 13. ábra mutatja be az öt vizsgált populációs mintára nézve. Az ábráról leolvasható, hogy valamennyi DNS-profil valószínűségi hányados értéke meghaladta az $LR = 10^{11}$ szintet, függetlenül attól, hogy melyik populációs adatbázis allélgyakorisági értékével számoltunk. Az $LR = 10^{11}$ szint (1 a százmilliárdhoz) a magyar népességben már rendkívül nagy igazságügyi bizonyító erővel rendelkezik. A BuCa, DeRo és BuAs görbék lefutása egymáshoz hasonló volt. A baranyai roma mintához tartozó görbe meredeksége a többi populációhoz viszonyítva kisebbnek adódott, ami a BaRo minta LR-értékeinek nagyobb szórásából következik. Valószínűsíthető, hogy az LR-értékek nagyobb szórását a baranyai populációnak és a többi populációnak – az allélgyakoriság-eloszlás szintjén tapasztalt –elkülönülése okozta. A kiskun görbe lefutása a BuCa, DeRo, BuAs triótól abban különbözött, hogy a görbe az előbbieken tárgyalt konzervatív becslés miatt balra csúszott.

13. ábra A 13 STR lokuszos DNS-profilok LR-értékének megoszlása az öt vizsgált populációs mintában. Az LR-értékeken F-statisztika korrekciót nem végeztünk



A Tippett-analízist az öt populáció páronkénti összehasonlítására is elvégeztük, amelynek eredménye a 14. ábrán látható. Ahogy az előbbieken bemutatott eredmények alapján sejteni lehetett, a legnagyobb arányú profilgyakorisági különbségeket, a legelnyújtottabb görbét a BaRo-Kisk párra kaptuk. Figyelemre méltó, hogy e populáció-páros allégyakorisági adatainak alkalmazása esetén a DNS-profilok 16%-ában kaptunk legalább három nagyságrend különbséget. E két adatbázis extrém helyzetét az is jól szemléltette, hogy ebben az összehasonlításban találtunk olyan profilokat, amelyeknek valószínűségi hányados érték különbsége meghaladta a hat nagyságrendet. Ez az eredmény azt jelenti, hogy az adott profil frekvenciája milliószor nagyobb (vagy kisebbnek) bizonyult attól függően, hogy melyik populáció adatbázisát használtuk. A legmeredekebb különbség-eloszlási görbét (a legkisebb különbséget) a két budapesti populáció összevetése során észleltük.

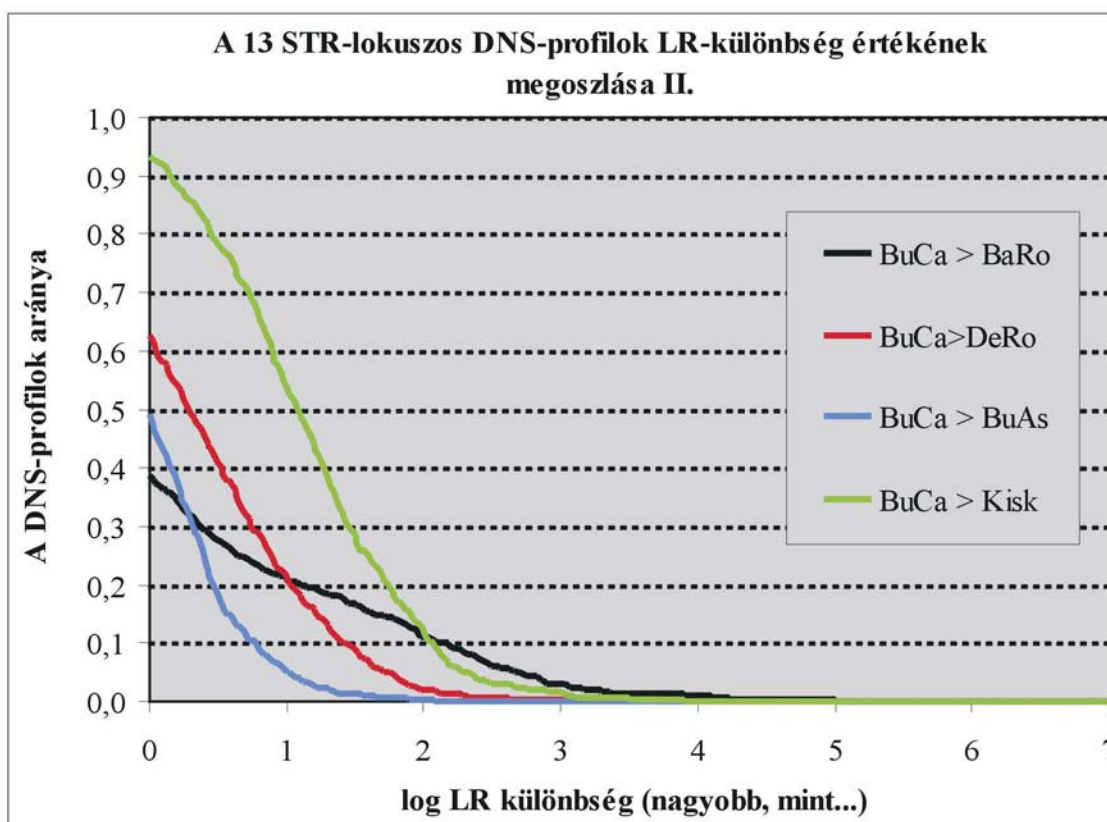
14. ábra A 13 STR lokuszos DNS-profilok LR-érték különbségének megoszlása a vizsgált populáció-párokban. Az LR-értékeken F-statisztika korrekciót nem végeztünk



A magyarországi populációs mintákon elvégzett igazságügyi genetikai kutatásaink egyik fő célkitűzése az volt, hogy megállapítsuk, a referencia adatbázisnak választott budapesti populációs minta allélgyakorisági adatbázisa milyen feltételekkel használható fel igazságügyi statisztikai célokra a magyar népességben. E kérdés megválaszolása érdekében olyan Tippett-diagramot szerkesztettünk, ahol csak azokat az LR értékeket vontuk be az analízisbe, ahol a BuCa allélgyakorisági adatbázisból számolt LR értékek meghaladták a többi adatbázisból kalkulált valószínűségi hányadosokat. Ezzel azt az igazságügyi helyzetet modelleztük, amikor a biológiai nyom bizonyító erejének megbecsülésekor, az LR-érték kalkulációjában azért tévedünk a terhelt személy kárára, mert az elkövető nem abból a populációból származott, amelyikből az allélgyakorisági adatbázist generáltuk. Ha e modellt a 14. ábra adatsorán érvényesítettük, a 15. ábrához jutottunk. A 15. ábra alapján megállapíthattuk, hogy a négy lehetséges párosítás közül kettőben, a baranyai roma és a kiskun adatbázissal való összehasonlításban a profilok 11-12%-a mutatott min. két nagyságrend LR különbséget. Ez az alkalmazott modellünk

alapján azt jelenti, hogyha a tettes e két populáció valamelyikéből származik, akkor az LR-érték megbecsülésekor várhatóan az esetek kb. 10%-ában tévednénk legalább két nagyságrendet a vádlott kárára. Ez az arány igazságügyi kontextusban jelentősnek mondható (Evett és Weir, 1998).

15. ábra A 13 STR lokuszos DNS-profilok LR-érték különbségének megoszlása a BuCa mintát tartalmazó populáció-párokban. Az LR-értékeken F-statisztika korrekciót nem végeztünk. Csak azokat a különbség-értékeket tüntettük fel, ahol: $LR_{BuCa} > LR_{egyéb\ adatbázis}$



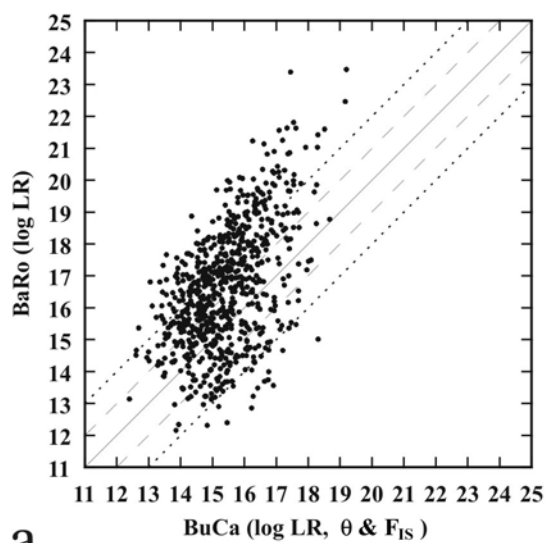
A biológia nyom, mint DNS-bizonyíték igazságügyi bizonyító erejének megbecsüléséhez a Bayes-elv alapján számíthatjuk ki a profil egyezési valószínűséget. A profil egyezési valószínűség azt mutatja meg, hogy mekkora annak az esélye, hogy a gyanúsítotton kívül még egy olyan személyt találunk a népességben, akinek a DNS-profilja teljes egyezést mutat a biológiai nyom genotípusával. Az 1990-es évek elejéig az egyezési valószínűség számításakor a szakértők azzal a feltételezéssel éltek, hogy a profilok függetlenek egymástól. E feltételezés szerint a második, egyezést mutató profil megfigyelésének várt valószínűsége nem függ az első profil (ti. a biológiai nyom genotípusa) megfigyelésének valószínűségével. Az egyezési valószínűség kalkulációja a

Hardy-Weinberg alapképletnek megfelelően történt. Balding és Nichols (1994) az 1990-es évek közepén rámutatott arra, hogy a bűncselekmények felderítése során a büntett elkövetésével meggyanúsított személyek gyakran abba az alpopulációba tartoznak, amelyből maga a tettes is származik. Ez azt is maga után vonja, hogy ebben a „gyanúsított” alpopulációban a biológiai nyom genotípusa gyakoribb lehet, mint a teljes népességben. A nagyobb profilgyakoriságot az alpopulációt alkotó személyek között potenciálisan meglévő genetikai korreláció okozhatja. Az előbbieken elmondottakra tekintettel Balding és Nichols az egyezési valószínűség számítására egy alternatív módszert javasolt, amely figyelembe veszi, hogy a tettes és a többi lehetséges gyanúsított egy vagy több, nem túlságosan régen élt közös őssel rendelkezhet. A genetikai korreláció hatását kifejezendő, a kutatók a profil egyezési valószínűséget megadó képleteikbe a közös leszármazási együtthatót (F_{ST}) építették be. 1996-ban az USA Nemzeti Kutatótanácsa is javasolta (NRC, 1996) a számítási módszert, amelynek hatására ez a megközelítés a szakértői társadalomban elfogadottá vált (Evetts és Weir, 1998). Ayres és Overall (1999) azonban felhívta a figyelmet arra, hogy Balding és Nichols formulái azon a feltételezésen alapultak, hogy az alpopuláción belül a párválasztás véletlenszerű. Ez a feltétel azonban sokszor nem teljesül migráns populációkban, ill. olyan népességekben, ahol a házasságkötések tradicionális vagy vallási alapon nyugszanak (Ayres és Overall, 1999). Ha a gyanúsítottak alpopulációjában pl. vérrokonság következtében magasabb fokú a beltenyésztés, akkor a homozigóta genotípusok aránya nagyobb a HW-egyensúlyi értékhez képest. Ennek megfelelően a második homozigóta genotípus megfigyelésének esélye több, mint azt a véletlenszerű párválasztás alapján várnánk. Ennek alapján Ayres és Overall (1999) tanulmányukban a profil egyezési valószínűség számítására egy olyan formulát javasolt, amely az F_{ST} mellett az F_{IS} (populáción belüli beletenyésztési együttható) értékét is tartalmazta.

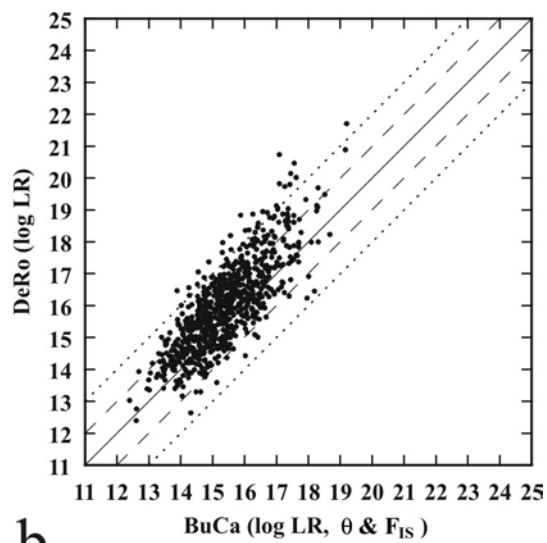
Az értekezésemben tárgyalt öt magyarországi populációs minta populációgenetikai vizsgálata során több olyan STR markert is találtunk, ahol az F_{IS} kalkulált értéke számottevő volt. Így a budapesti kevert jellegű referencia populációnk további igazságügyi statisztikai jellemzéséhez az Ayres és Overall által javasolt egyezési valószínűségi formulát alkalmaztuk. A BuCa minta további igazságügyi jellemzésében a budapesti referencia populációra úgy tekintettünk, mint a magyarországi népességet reprezentáló populációs mintára. A többi négy populációs minta a tettesek és potenciális gyanúsítottak egy-egy alpopulációját hivatott képviselni. A valószínűségi hányadosok sörét- és Tippett-diagramokon alapuló analízisével azt vizsgáltuk, hogy az LR-értékek és

különbségeinek eloszlása hogyan módosul, ha a profil egyezési valószínűség számítását a BuCa adatbázis esetében az Ayres és Overall által javasolt módon korrigáltuk. A kapott eredményeket a fentiekben leírt módon sörét- és Tippett-diagramon ábrázoltuk (16-17. ábrák). A 16. ábra sörét diagramjain megfigyelhető volt, hogy az LR értékek szóródásában – az F-statisztikákkal nem korrigált LR-értékeket bemutató diagramokhoz képest – elhajlás következett be. Az elhajlás mindig a BuCa adatbázissal összehasonlításra kerülő másik adatbázis tengelye felé történt. A 17. ábra alapján megállapíthattuk, hogy mind a négy lehetséges párosításban a profilok legfeljebb 2,5%-a mutatott min. két nagyságrend LR különbséget. Ez az alkalmazott alpopulációs modellünk alapján azt jelenti, hogyha az igazságügyi statisztikai interpretáció során a referencia budapesti allélgyakorisági adatbázissal történik a kalkuláció, akkor várhatóan az esetek legfeljebb 2-3%-ában tévednénk min. két nagyságrendet a terhelt kárára. Figyelembe véve azonban az LR értékek nagyságrendjét és az LR különbség-eloszlási görbék gyors lecsengését, kijelenthetjük, hogy a detektált különbség igazságügyi szempontból elhanyagolható.

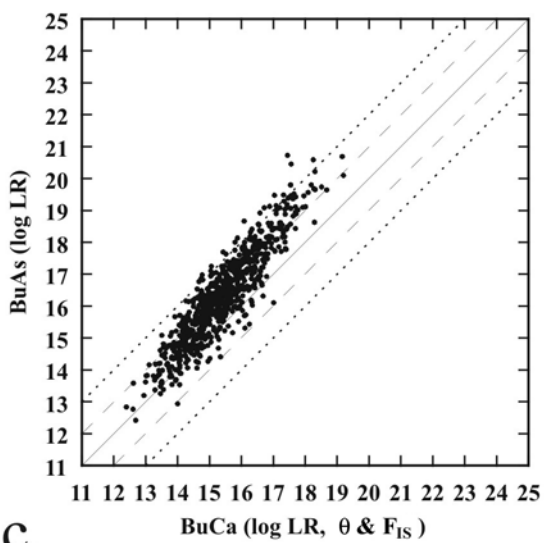
16. ábra A 13 STR lokuszos DNS-profilok sörét-diagramjai a BuCa mintát tartalmazó populáció-párokra. A BuCa adatbázisból számított LR-értékeken F-statisztika korrekciót végeztünk



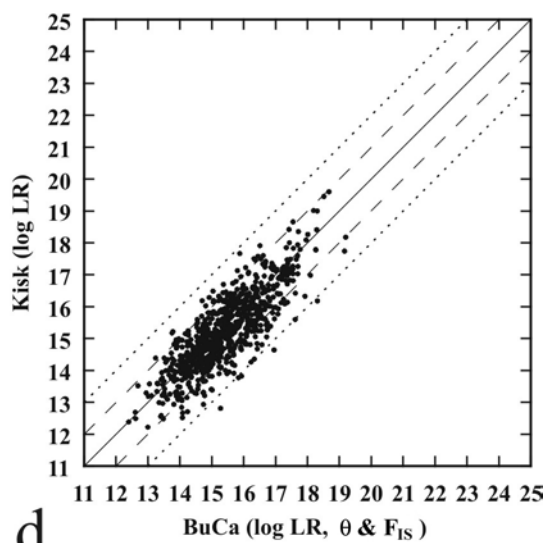
a,



b,

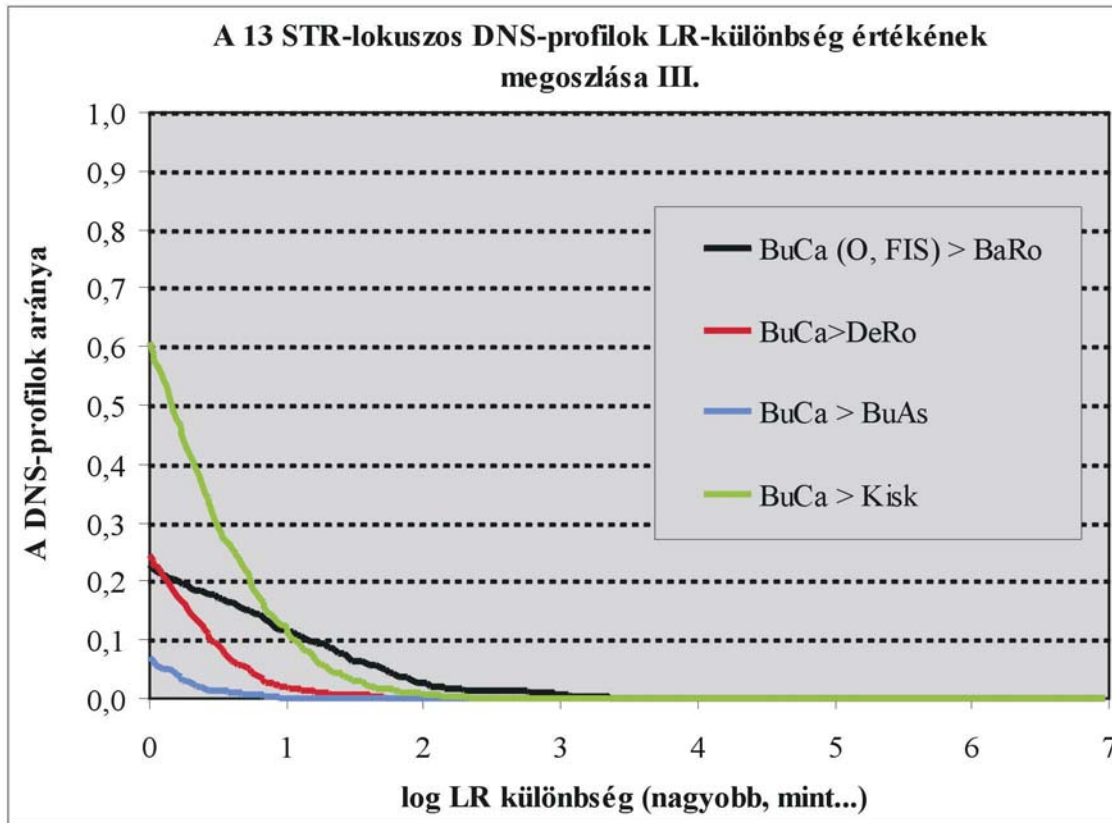


c,



d,

17. ábra A 13 STR lokuszos DNS-profilok LR-érték különbségének megoszlása a BuCa mintát tartalmazó populáció-párokban. Az LR-értékeken F-statisztika korrekciót végeztünk. Csak azokat a különbség-értékeket tüntettük fel, ahol: $LR_{BuCa} > LR_{egyéb\ adatbázis}$



A budapesti referencia populációs minta és a többi négy magyarországi populációs adatbázis igazságügyi statisztikai összehasonlításának eredménye alátámasztotta azt a feltevésünket, hogy a budapesti referencia minta alkalmas lehet a Magyarországon elkövetett bűncselekmények DNS-vizsgálati eredményének korrekt statisztikai kiértékelésére. A budapesti referencia adatbázison alapuló, Ayres és Overall- (1999) féle egyezési valószínűség számítási képlet bizonyult a legmegfelelőbbnek. A genetikai korrelációt kifejező F-statisztika értékekkel korrigált kalkulációval elérhető, hogy a statisztikai becslések ne a terhelt kárára történjenek, ugyanakkor a kiszámított valószínűségi hányados ne legyen indokolatlanul kedvező a terhelt számára.

3.4.1.5. A biológiai nyomot hátrahagyó személy etnikumának valószínűsítési lehetőségei: az etnikai eredet becslése (EEO)

Napjainkban több ezer DNS-polimorfizmus ismeretes. Valamennyi polimorfizmus kisebb-nagyobb mértékben alkalmas lehet annak megbecslésére, hogy a bűncselekmény elkövetője melyik populációcsoportba tartozott. A kérdés inkább úgy vetődik fel, hogy mely lokuszok bírnak a legnagyobb információtartalommal az etnikai származás megbecsléséhez (EEO), illetve ezen lokuszok analízise kivitelezhető-e az igazságügyi laboratóriumok számára. Nyilvánvalóan olyan markerek kiválasztása a legelőnyösebb, amelyek vizsgálatához a laboratóriumoknak nem kell teljesen új vizsgálati technológiát beállítani és hitelesíteni. A humán populációkban polimorfizmust mutató STR markereknek egy szűk köre, 25-30 lokusz genetikai analízise terjedt el az igazságügyi genetikai szakértői munka során széles körben (Ruitberg és mtsai, 2001; STRbase: <http://www.cstl.nist.gov/biotech/strbase>). Az igazságügyben elemzett lokuszok egy részét (pl. az USA CODIS törzs-lokuszeit) a bűnügyi DNS-adatbázisokban is elemzik. A nemzeti bűnügyi DNS-nyilvántartások szaporodásával világszerte sok laboratórium rendelkezik olyan technológiai háttérrel, amely lehetővé teszi biológiai minták 10-20 STR markerre irányuló tömeges tipizálását.

Shriver és munkatársai (1997) korábban elvégeztek egy széleskörű elemzést, amelyben nagy számú DNS-polimorfizmus EEO potenciálját mérték fel az USA kaukázusi, afro-amerikai és spanyolajkú népességeiben. A vizsgálatba bevont lokuszok némelyike nagyobb populációs megkülönböztető erővel bírt, mint amekkora a korábbi populációs felmérések alapján az értekezésemben tárgyalt STR lokuszoktól várható volt. Azonban ezeket az EEO szempontjából informatív lokuszokat az igazságügyi genetikai laboratóriumokban rutinszerűen nem vizsgálják. Jelenleg hasonló megfontolások érvényesek az egynukleotid polimorfizmusokra (SNP-k) és számos Y-kromoszómás markerre, amelyek a legígéretesebbnek tűnnek az etnikai származás tisztázása szempontjából (Brooks AJ, 1999; Kayser és mtsai, 2001)

Számos olyan tanulmány is született az elmúlt években, amelyek az EEO megvalósíthatóságát a DNS-profil gyakoriságok valószínűségi hányadosa alapján tesztelte (Evett és mtsai, 1992; Meyer és mtsai, 1995; Brenner, 1998; Smouse és Chevillon, 1998). E megközelítések fő hiányossága, hogy az eredményekből a helyes és téves becslések várt aránya, így a procedúra megbízhatósága nem volt kiolvasható.

Az általunk analizált autoszómás STR markerek etnikai származás megállapító képességét nyolc populáció három csoportjában mértük fel (Klitschar és mtsai, 2003). A nyolc populációt a következő szempontok szerint választottuk ki:

I. csoport: a budapesti referencia és a baranyai roma populációs mintánk, amely olyan populációkat hivatott reprezentálni, ahol a népesség nagyobbik fele együtt él egy jelentős arányú, de viszonylag izolált kisebbséggel;

II. csoport: az osztrák kaukázusi ill. az egyiptomi populációs minta (Klitschar és mtsai, 1998) olyan európai népségeket képviselt, akik egymástól földrajzilag meglehetősen távol élnek;

III. csoport: kaukázusi, afro-amerikai, ázsiai és spanyol ajkú populációs minta (Mechthild Prinz, személyes közlés), az USA négy nagy népcsoportját reprezentálta.

A 13 CODIS STR lokusz-készlettel vizsgált EEO megbízhatóságát a helyes és téves becslések arányának meghatározásával mértük meg a három populáció-csoporton belül (13. táblázat). A EEO mért megbízhatósága az ismeretlen személy valódi etnikai eredetétől, a eredeti populációk feltételezett számától és populációk genetikai különbözőségétől függött.

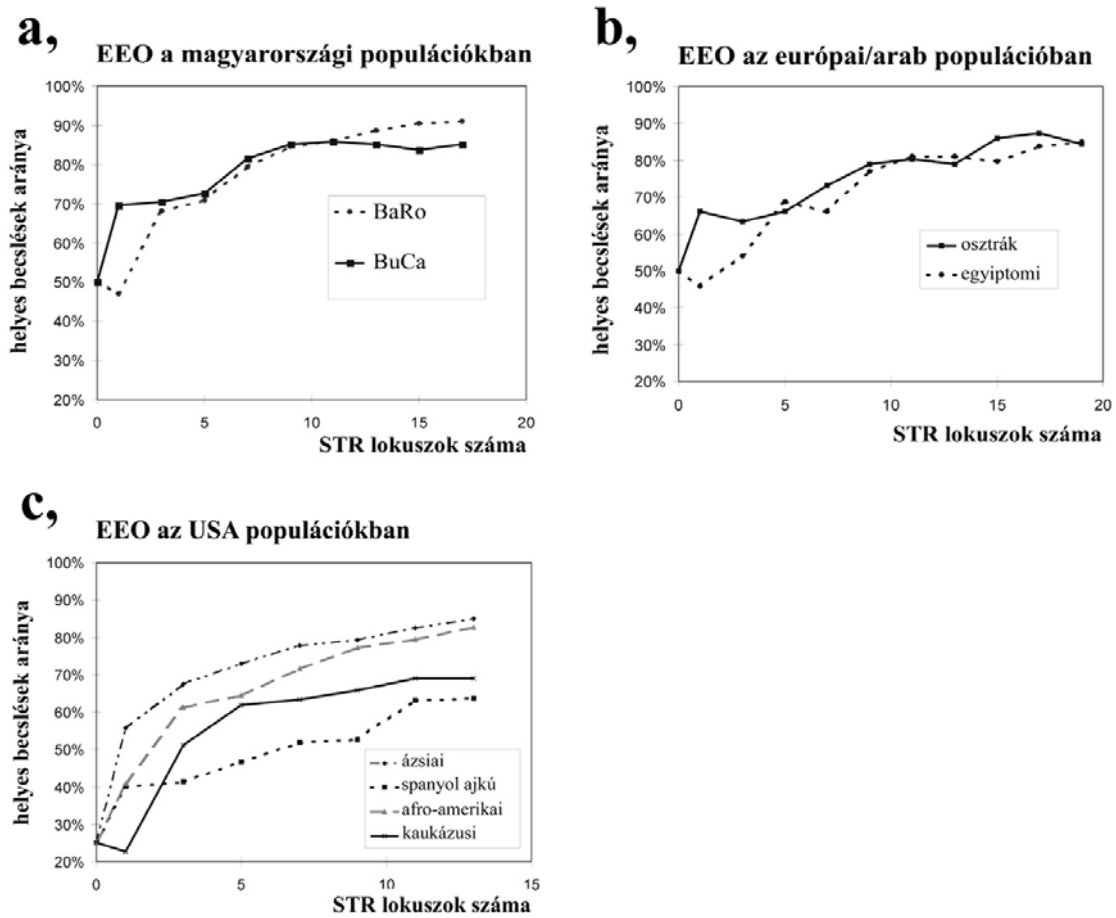
A magyarországi populáció-csoportban a személyek közel 90%-ára az EEO helyes eredményt adott. Az EEO az osztrák/egyiptomi párra már csak 80%-ban volt megfelelő. Az USA populáció-csoportban az etnikai származás helyes becslése *a priori* kisebb sikerrel kecsegtetett az első két csoporthoz képest, mivel a lehetséges eredeti populációk száma kettőről négyre emelkedett. Ehhez képest az ázsiai és az afro-amerikai populációra kapott 80% körüli sikeres azonosítás meglehetősen magas értéknek számít. A kaukázusi és a spanyol ajkú populációs eredet beazonosítása lényegesen kisebb sikerrel járt. Az EEO megbízhatóságának csökkenése a kaukázusi és a spanyol ajkú populáció nyilvánvaló genetikai kapcsolatára utal. Az EEO-vizsgálat során a helytelenül beazonosított kaukázusi származású személyek túlnyomó része spanyol ajkúnak lett titulálva és vice versa. Mindezek mellett a „spanyol ajkú” terminus különbözőféleképpen értelmezhető, így a megfelelő populációs adatbázis felállítása a spanyol ajkú populációra nehézkes. Az eredmények tükrében megbízhatóbbnak tűnt egy olyan EEO vizsgálati szempont felállítása, ahol a kaukázusi vagy spanyol ajkú hovatartozást nem kívántuk tisztázni. Ekkor a két populációt alkotó személyek 93%-a sikeresen elkülöníthető volt az ázsiai és afro-amerikai populációtól.

13. táblázat A helyes eredményt adó etnikaieredet-becslések aránya 13 STR lokusz vizsgálata esetén (Klitschar és mtsai, 2003)

I. magyarországi populációk	BaRo	88.8%
	BuCa	85.2%
II. európai/arab populáció	osztrák	78.9%
	egyipomi	81.1%
III. USA populációk	ázsiai	84.9%
	afro-amerikai	82.7%
	kaukázusi	69.1%
	spanyol ajkú	63.8%

A helyes etnikaieredet-becslések aránya a vizsgált lokuszok számával nőtt (18a-c. ábra). Habár már a 8 STR markert tartalmazó profil is meglehetősen magas információtartalommal rendelkezett az EEO szempontjából, azonban optimális eredményt 15-17 lokusz bevonásával kaptunk. Az EEO sikerességi görbék 15 lokusz után ellaposodtak, ami azt jelezte számunkra, hogy további lokuszok alkalmazása már csak csekély kiegészítő információval járulhat hozzá a kérdéses személy etnikumának megállapításához. Ha a 18. ábrán látható görbéket extrapoláljuk, akkor megjósolható, hogy rendkívül valószínűtlen az EEO sikerességének 90-95% fölé emelése. Az EEO sikerességi görbék 100% alatti telítődését a vizsgálatba bevont személyek egy részének közös vagy idegen eredete magyarázhatja. Az eredményeink arra utalhatnak, hogy a modern, globalizált társadalmakban teljesen izolált, érintetlen etnikai közösségek gyakorlatilag nem léteznek.

18. ábra A három populáció-csoportban elvégzett etnikaieredet-becslés sikerességi rátája a vizsgált STR lokuszok számának függvényében (Klitschar és mtsai, 2003)



3.4.2. Y-kromoszómás STR lokuszok vizsgálatának eredményei

3.4.2.1. Az Y-haplotipizálás eredményei

Az értekezésemben tárgyalt öt magyarországi populációs minta igazságügyi genetikai analizését kilenc Y-kromoszómás mikroszatellita (STR) lokusz bevonásával is elvégeztük. Az Y-kromoszómára irányuló elemzéseink első lépcsőjében a budapesti referencia és a baranyai roma mintán végeztük el a felmérést (Füredi és mtsai, 1999). E két populációs minta az európai Y-STR haplotípus referencia adatbázis (Eu-YHRD) részét képezi (Roewer és mtsai, 2001). A 2000-2003. közötti időszakban Y-STR vizsgálatainkat a kelet-

magyarországi roma, a budapesti askenázi és a kiskun populációs mintára is kiterjesztettük.

A genetikai analízis során a 299 férfi mintáján meghatározott allélok, illetve lévén ezek kapcsoltak, azok specifikus kombinációi, a haplotípusok az értekezésemhez csatolt CD-rom *Y-STR-haplotípusok.xls* nevű Excel fájlában található meg populációs bontásban.

Korábban bemutattuk, hogy a mikrovariáns allélokat hordozó személyek haplotípusainak összehasonlítása során hasonlóságokat, egyezéseket figyeltünk meg. Ha feltételezzük azt, hogy a DYS385 *-1 interallélhoz vezető, egy nukleotidot érintő mutáció gyakorisága az SNP polimorfizmusok mutációs rátájával azonos nagyságrendű (kb. 10^{-7} /lokusz/generáció, Kayser és mtsai, 2000), akkor felállíthatjuk azt a hipotézist, hogy a kérdéses interallélokat hordozó személyeket között feltehetően közös leszármazás lehet. Ezt a feltételezést a haplotípusok négy lokuszon, a DYS389-I, a DYS390, a DYS391 és a DYS392 markereken való teljesen egyezése és az érintett kromoszómákon elvégzett SNP-analízis eredménye is alátámasztotta. Ahol különbséget találtunk az STR allélok között, az is csak max. három repeat egységnyi volt, s ez az eltérés magyarázható a közös eredetű Y-kromoszómák mikroszatellitáinak single step (lépésenkénti) mutációs modellen alapuló módosulásával. Az STR lokuszok repeatjeit érintő mutációs ráta az SNP-k mutációs gyakoriságához viszonyítva nagyságrendekkel nagyobb (kb. 3×10^{-3} /lokusz/generáció (Heyer és mtsai, 1997; Kayser és mtsai, 2000). A vizsgált, interallélt tartalmazó haplotípusok alléljainak általunk megfigyelt elkülönülése összhangban van az STR allélok lépésenként történő mutációs hipotézisével. A nyolc *-1 interallélt hordozó személy négy populációs mintánkból (BuCa, BaRo, BuAs, Kisk) származott, ami e populációk férfi tagjainak keveredésére (admixture) utalhat.

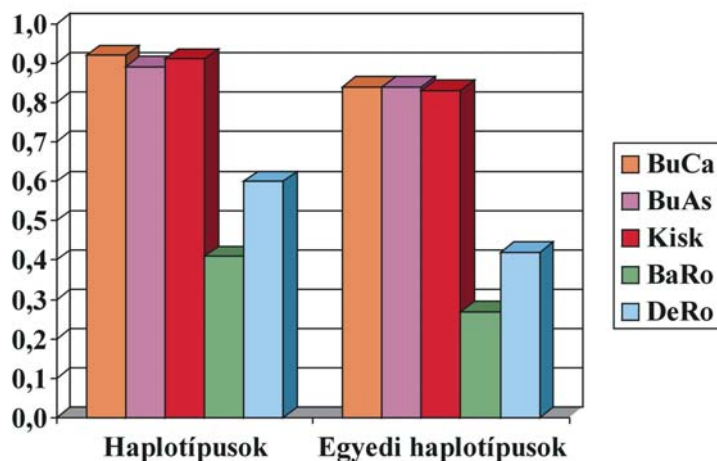
3.4.2.2. A populációs felméréseink során detektált Y-haplotípusok kiértékelése

A haplotípus eredmények kiértékelése során először a haplotípusok populációs megoszlását vizsgáltuk meg az öt populációs mintában. Minden populáció esetében megszámloltuk, hogy hány különböző haplotípus fordul elő, majd ezt elosztottuk a minták számával. Így az összehasonlíthatóság szempontjából is kedvező, százalékos formában kaptuk meg az eredményt, ami a 19. ábra bal oldalán látható populációkra bontva. A 299

férfiban összesen 192 különböző (64%) haplotípust találtunk. Ezeknek az Y kromoszómás STR lokuszoknak a kriminalisztikai gyakorlatban való felhasználhatóságának természetesen az kedvez, ha minél több a különböző haplotípusok aránya, hisz annál kevesebb személy hordozza ugyanazt az allélkombinációt, s ezzel csökken a valószínűsége az ugyanolyan haplotípust hordozó személyek előfordulásának. Az ábrán látható, hogy a számolások eredményeként adódott értékek a két roma népességben, főleg a BaRo mintában feltűnően alacsonyak a BuCa, a BuAs és a Kisk populációkhoz képest, melynek okai a romák más populációkkal való kismértékű keveredésében és a beltenyésztés magasabb fokában keresendő (Czeizel és mtsai, 1991).

A 19. ábra jobb oldalán az egyedi haplotípusok aránya látható populációkra bontva. Ehhez minden populációban megszámláltuk a csak egy személyben előforduló haplotípusokat, és a kapott értéket elosztottuk a minták számával. Ha a teljes populációs mintára is elvégeztük ezt a kalkulációt, akkor 299 férfiban összesen 156 egyszer előforduló, egyedi haplotípust találtunk (52%). Ezen a grafikonon is jól észrevehető, hogy a két roma népességben a többihez képest kiugróan alacsony az egyedi allélkombinációk száma, azaz egy adott haplotípust a többi populációhoz képest sokkal több egyén hordoz.

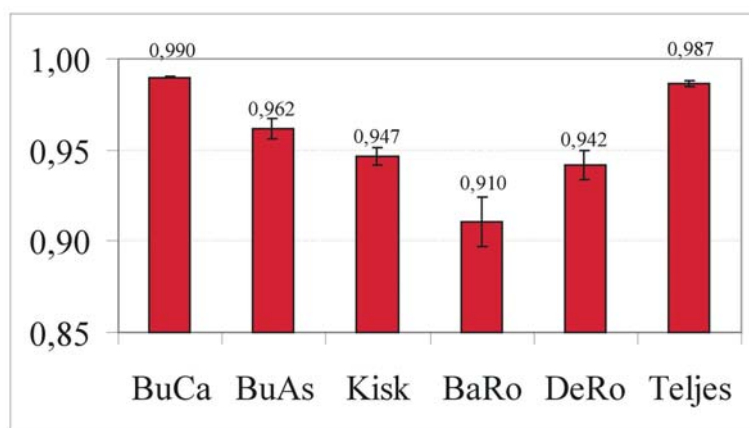
19. ábra A haplotípusok populációs megoszlása



A 20. ábrán a haplotípus-diverzitás látható populációkra bontva, melyet a Nei-féle géndiverzitás alapján számoltunk. Mint látható, a roma populációs minták diverzitásértékei lényegesen kisebbnek adódtak a többi populációhoz képest, a kiskun minta kivételével. Az összevont populációs mintára számolva a teljes haplotípus-diverzitás 0,987. Ez azt jelenti, hogy két személyt véletlenszerűen kiválasztva a népességből 98,7% a

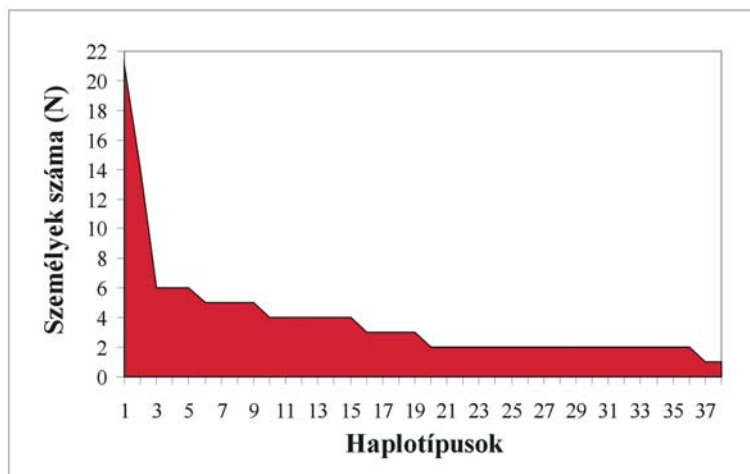
valószínűsége annak, hogy a két személy eltérő haplotípust hordoz az Y kromoszómáján, tehát az igazságügyi vizsgálatok során egymástól megkülönböztethető, feltéve, ha a két személy nincs férfiágú rokonságban egymással. A budapesti referencia populáció haplotípus-diverzitása hasonlóan bizonyult más kaukázusi populációk diverzitásához (Füredi és mtsai, 1999). A többi populációs mintánk esetében a kisebb géndiverzitást bizonyos haplotípusok közös eredete (elsősorban a roma populációknál) és/vagy a mintában lévő személyek viszonylag kis száma (elsősorban az askenázi és kiskun populációknál) okozhatta.

20. ábra Haplotípus-diverzitás és szórás-értékek a vizsgált öt populációra és az egyesített populációkra nézve



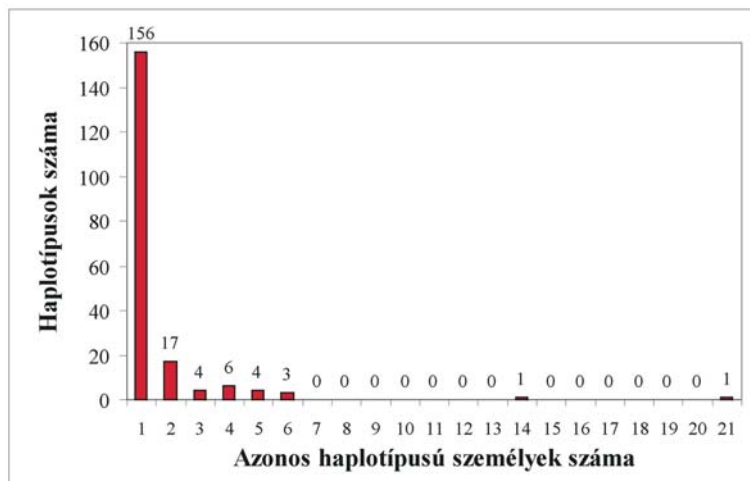
A 21. és a 22. ábrán a haplotípusgyakoriság-eloszlás látható az öt populációra együtt nézve. A 21. ábra vízszintes tengelyén a számok az előfordult haplotípusokat jelölik oly módon, hogy a leggyakoribb kapta az 1-es számot, a következő a 2-est, és így tovább. A függőleges tengelyről azon személyek száma olvasható le, akik az adott haplotípust hordozták. Igazságügyi szempontból az az előnyös, ha a görbe alacsonyan indul, és minél előbb lecseng, hisz annál kevesebb az olyan haplotípus, amelyet sok egyén hordoz. Esetünkben az egyedi haplotípusok magas arányának köszönhetően a görbe gyors esést mutatott, azonban az is megfigyelhető, hogy az ábrán 1-essel és 2-essel jelölt haplotípus kiugróan sok személyben (21 ill. 14) képviseltette magát. Ezek az ún. extra-domináns (uralkodó) haplotípusok egy kivételével a roma populációkhoz tartoztak.

21.ábra Haplotípusgyakoriság-eloszlás a teljes populációs mintán I.



A 22. ábra vízszintes tengelye az azonos haplotípusú személyek számát jelöli, a függőleges tengely pedig a haplotípusok számát. A 192-ből 156 olyan haplotípus volt, amelyet csupán egy egyén hordozott, és olyan haplotípust már nem találtunk, mely hét személyben fordult volna elő. Mindazonáltal találtunk két olyan allélkombinációt, amelyet 14 illetve 21 egyén mintája mutatott, ezek a 21. ábrán 1. és 2. számmal jelölve láthatók.

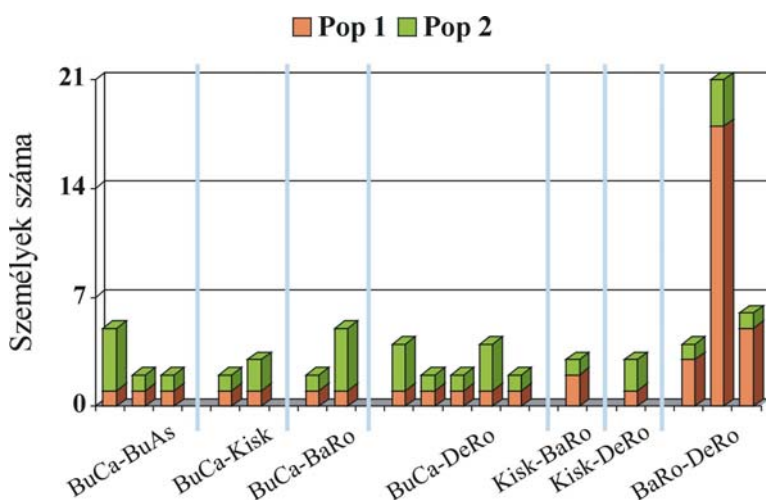
22.ábra Haplotípusgyakoriság-eloszlás a teljes populációs mintán II.



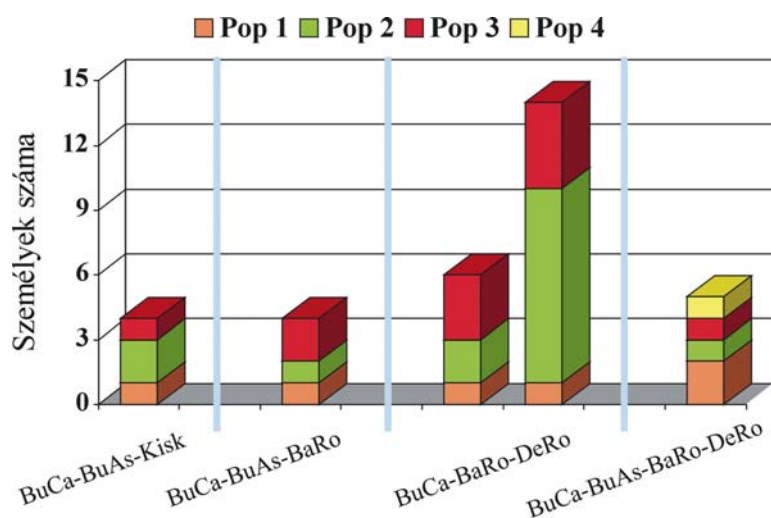
A többször előforduló haplotípusok közül 17 figyelhető meg 7 populáció-párban, 4 haplotípus 3 populáció-trióban, és 1 populáció-négyesben. A többször előforduló haplotípusok megoszlása a 23. és 24. ábrán látható. Az ábrákon az oszlopok egy-egy haplotípust jelölnek, a függőleges tengelyről pedig azon személyek száma olvasható le,

akik azt a haplotípust hordozták az adott két, három vagy négy populációban. Olyan haplotípust nem találtunk, mely mind az öt mintában előfordult volna. A 23. ábrán, amely a közös haplotípusok populáció-páros megoszlását ábrázolja, megfigyelhető, hogy a BuCa és a DeRo populációkban öt olyan haplotípus is volt, melyen ez a két populáció osztozott. Ezzel ellentétben a Kisk népesség a két roma mintával csak egy-egy haplotípuson osztozott. A BuAs mintának a BuCa mintán kívül nincs olyan párja, akivel ketten osztoznának egy haplotípuson. A 23. ábrán bemutatott eredmények alapján a két roma populáció igen sok egyéne hordozta a mindkét népességben megjelenő haplotípusokat; különösen feltűnő a már többször bemutatott 21 személy által képviselt allélkombináció. A 24. ábra a többször előforduló haplotípusok közül a populáció-triókban és populáció-négyesben megjelenőket ábrázolja. Populáció-négyesben csak egyetlen haplotípus képviseltette magát, és populáció-trióban is összesen csak négy, azaz igen magas fokú populációs specificitás figyelhető meg a haplotípusok szintjén. Ezt természetesen részben a relatíve kis mintaszám okozta. A 23. és 24. ábrából leolvasható, hogy az előbbieken említett két extra-domináns haplotípus mindkét roma populációs mintában előfordult, s ezeken kívül csupán egy BuCa férfiban volt meg az egyik ilyen haplotípus. Mivel azonban ebben a populációban a mintagyűjtés nem etnikai alapon történt, ezért a kérdéses személy lehetett roma is. Az előbbieken alapján e két uralkodó haplotípus roma-specifikus lehet. A roma populációkban megfigyelt alacsony diverzitási értékeket az uralkodó férfi leszármazási vonalak közös eredete magyarázhatja. Ebben a roma népesség viszonylagos zártsága, és ebből fakadóan a beltenyésztés magasabb foka játszhat szerepet.

23. ábra A közös haplotípusok megoszlása populáció-párokban



24. ábra Közös haplotípusok megoszlása populáció-triókban és populáció-négyesben



A roma-specifikusság kérdésének eldöntése érdekében az európai adatbázisból (Eu-YHRD) kikerestük azokat a haplotípusokat, amelyeken a két roma populáció vagy a két roma és a BuCa populáció osztozik (összesen 5). Megnéztük, hogy a túlnyomó többségében európai kaukázusi személyek tipizálásából nyert haplotípusok között ezek az allélkombinációk milyen gyakran fordultak elő, majd a kapott adatokat összevetettük a magyarországi előfordulási gyakoriságokkal (lekérdezés időpontja: 2003.08.15.). A kapott eredményeket a 14. táblázatban összegeztük.

14. táblázat A BaRo-DeRo ill. a BaRo-DeRo-BuCa populációk közös haplotípusainak detektált száma és aránya a magyar és az európai (Eu-YHRD) adatbázisban (lekérdezés időpontja: 2003.08.15.)

Haplotípus DYS19/389-I/II/390/391/392/393/385	Magyar adatbázis (299)	Európai adatbázis			
		teljes (12.904)	BuCa (116)	BaRo (78)	bulgáriai roma ^a (81)
15/13/31/25/11/11/13/11-14	21 (7,0%)	32 (0,2%)	0	18	3
15/14/30/22/10/11/12/15-17	14 (4,7%)	37 (0,3%)	1	9	15
15/14/30/22/10/11/12/15-16	6 (2,0%)	9 (0,1%)	0	5	1
14/14/31/23/10/11/12/13-16	6 (2,0%)	14 (0,1%)	1	2	0
14/12/28/22/10/11/13/14-14	4 (1,3%)	54 (0,4%)	0	3	7

^a Zaharova et al., 2001

A 14. táblázatban látható, hogy az összesített magyarországi populációs mintában mindegyik haplotípus sokkal gyakoribb volt a teljes európai adatbázishoz viszonyítva. Az Eu-YHRD adatbázis az értekezésem írása időpontjában két roma populációs mintát tartalmazott: a baranyai roma mintánkat és egy 81 személyből álló bulgáriai roma mintát

(Zaharova és mtsai, 2001). Az általunk vizsgált öt haplotípus közül négy detektálható volt a bulgáriai roma mintában is, a négy haplotípus közül három típust min. három bulgáriai roma személyben is megtaláltunk. Ezek az eredmények nagymértékben alátámasztják azt a feltételezést, hogy e haplotípusok nagymértékben köthetők a roma népcsoporthoz.

3.4.2.3. Az Y-haplotípusok populációstatisztikai összehasonlításának eredményei

A felmért öt populáció AMOVA alapú, páronkénti összehasonlításának eredményét a 15. táblázat mutatja be. Az ábra bal oldalán a Φ_{ST} értékek kerültek ábrázolásra. Az AMOVA eredménye azt mutatta, hogy az összesített populációra vonatkoztatva a teljes genetikai variancia nagy részét, 97,52%-át az egyes populációkon belüli variancia adta. Mindazonáltal jelentős különbség is adódott több populáció között. A roma populációk egymással összehasonlítva alacsony Φ_{ST} értéket mutattak, azonban magas értékeket kaptunk, ha a roma mintákat a Kisk és BuCa mintákkal teszteltük. A roma populációk egyedül az askenázi mintával nem adtak magas Φ_{ST} értéket. Figyelemre méltó az is, hogy a BuCa-Kisk párra számolt Φ_{ST} érték a legkisebb kategóriába esett.

A 15. táblázat jobb felső háromszögében a populáció-párookra számolt Φ_{ST} értékek statisztikai szignifikanciáját (P) ábrázoltuk. A szignifikancia-tesztelés teljes egészében megerősíti a Φ_{ST} értékekből levont következtetéseinket; pl. még $\alpha = 48\%$ -os szignifikancia-szint mellett sem találtunk különbséget a roma populációk között (ez volt a legmagasabb P érték). Ezzel ellentétben mindkét roma minta – az askenázi populáció kivételével – mindegyik populációtól jelentősen eltért. Az analízis alapján a Φ_{ST} értékekhez hasonlóan nem volt jelentős különbség a BuCa és BuAs minták között.

Érdekes, hogy az askenázi populációban nem volt észlelhető hasonló jelenség, mint a romáknál; a BuAs minta nem különbözött számottevően a többi populációtól, és nem találtunk uralkodó haplotípusokat sem. Ennek oka lehet a relatíve kis mintaszám (37 férfi), valamint az a megfigyelés, hogy e népesség tagjainak vallási/etnikai identitására vonatkozóan elsősorban az anya vallási/etnikai hovatartozása a mérvadó. Így olyan egyének is bekerülhettek a BuAs felmérésünkbe, akik bár férfiágról – ahonnan az Y kromoszómájukat kapták – nem zsidó származásúak, az anyai ág miatt azonban mégis ebbe az etnikai csoportba tartoztak. Egy, a közelmúltban elvégzett populációs felmérés megerősítette ezt a feltételezést. Thomas és munkacsoportja (2002) földrajzilag szeparált

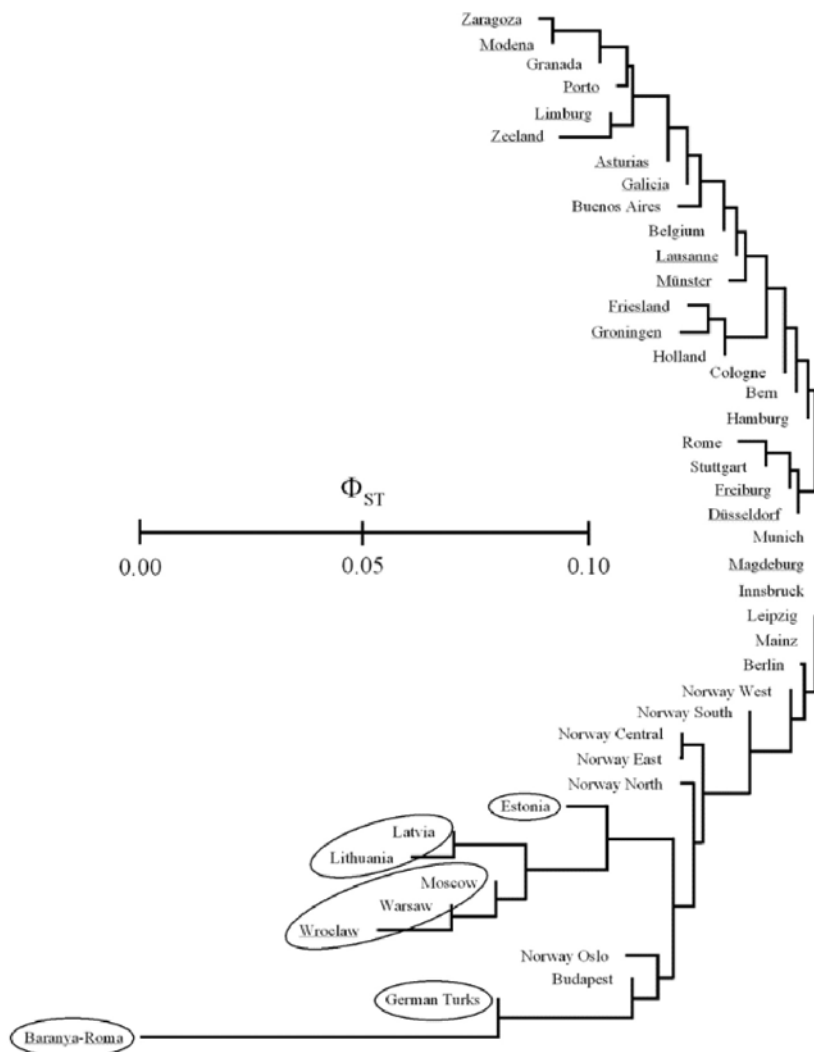
kilenc zsidó, nyolc nem-zsidó és egy izraeli palesztin populációt hasonlított össze mitokondriális (anyai) ágon, valamint Y-kromoszómás DNS-polimorfizmusok segítségével. A felmérés azt mutatta, hogy mitokondriális markerek alapján a zsidó populációk kevés számú női ősrre vezethetők vissza, az alapító hatás függetlennek bizonyult a különböző földrajzi területek tekintetében, és a környező népeiségekből történő genetikai módosító hatás korlátozott volt a női oldalon. Ezzel éles ellentétben állt az Y kromoszómás markerek alapján végzett felmérés, amely hasonlóságot mutatott a többi populációval, nem adódtak szignifikáns eltérések, és alapító hatásra sem utalt semmi az adatok alapján.

15. táblázat Populációk Y-haplotípusainak AMOVA alapú statisztikai összehasonlítása

	BuCa	BaRo	DeRo	BuAs	Kisk
BuCa		$P \leq 0,01$	$0,05 \geq P > 0,01$	$P > 0,05$	
BaRo	$\Phi_{ST} > 0,03$				
DeRo	$0,03 \geq \Phi > 0,02$				
BuAs	$0,02 \geq \Phi > 0,01$				
Kisk	$\Phi_{ST} \leq 0,01$				

Roewer és munkacsoportja (2001) 41 európai populáció AMOVA-alapú összehasonlítását végezte el a budapesti referencia és a baranyai roma populációnk Y-haplotípus adatainak bevonásával. A populáció-párokra meghatározott molekuláris genetikai távolságok (Φ_{ST}) alapján filogenetikai fát konstruáltak, amelyet a 25. ábra mutat be.

25. ábra 41 európai populáció Y-STR haplotipizálása alapján szerkesztett neighbor-joining filogenetikai fa. Az ellipszisek a statisztikailag elkülönülő populációcsoportokat jelölik (Roewer és mtsai, 2001)



Az ábrázolt filogenetikai fa alapján a balti, norvég és magyar referencia férfi populáció kis mértékben elkülönült a vizsgált európai populációk többi tagjától. Jelentősebb genetikai távolság volt megfigyelhető a Németországban élő törökök és a baranyai romák esetében.

A fent leírt eredményeink alapján elmondható, hogy az igazságügyi Y-STR azonosításban a szakértőknek a valószínűsítő vélemény elkészítéséhez a magyarországi népesség szubstrukturálódását számításba kell venniük, hiszen analízisünk során némely populáció között statisztikailag is jelentős különbség adódott. Szükség van tehát további lokális és globális felmérésekre és összehasonlításokra a szakértői vélemények legmegfelelőbb populációstatisztikai interpretálásához.

3.5. Esettanulmány

Az igazságügyi genetikai aspektusú populációs felmérések fontosságát hangsúlyozandó, az alábbiakban ismertetünk egy esettanulmányt, amely szerintünk kellőképpen rávilágít a populációs felmérések hiányából eredő tévedések lehetőségére.

2001-ben a rendőrség büntetőeljárást indított egy fiatalember ellen, akivel szemben alapos gyanú merült fel, hogy egy nő sérelmére nemi erőszakot követett el. A rendőrségi tanúkihallgatási jegyzőkönyv alapján, a sértett szerint az elkövető kreol bőrű, sötét hajszínű fiatalember volt. A rendőrség az elkövető DNS-alapú személyazonosítása érdekében kirendelte az ORFK Bűnügyi Szakértő és Kutatóintézet igazságügyi genetikus szakértőjét szakértői vizsgálat elvégzésére. Az összehasonlító DNS-vizsgálathoz rendelkezésre álló biológiai minta (1 db tárgylemezen fixált, festett, fedett hüvelykenet) rossz állapota és a rajta lévő spermiumok csekély mennyisége miatt az autoszómás STR lokuszokon végzett analízis nem vezetett eredményre, így Y-kromoszómás STR vizsgálatra került sor. Hat Y-kromoszómás lokuszon az analízis sikerrel járt és egyértelmű egyezést mutatott a bűnjelminta (az elkövető) és a gyanúsított haplotípusa között. Ez a haplotípus azonos volt a 14. táblázat 3. haplotípusának megfelelő allélkombinációjával.

A valószínűsítő vélemény kiadásához az igazságügyi genetikus szakértőnek az adott haplotípus előfordulási gyakorisági értékére volt szüksége. Ehhez először az európai adatbázist (Eu-YHRD) használta fel. A lekérdezés időpontjáig haplotipizált 9340 kaukázusi férfiból három hordozta a kérdéses haplotípust (0,032%) az adatbázisban. Amikor azonban a magyarországi populációs mintában is elvégezte ugyanezt a keresést, hatalmas különbség mutatkozott. A vizsgált 299 magyarországi férfiból hat ugyanezzel a haplotípussal rendelkezett (4,6%), és mind a hat személy roma származású volt. Ha az igazságügyi genetikus szakértő csak az európai adatbázisban szereplő haplotípus-gyakoriságra hagyatkozott volna, az elkövető és a gyanúsított DNS-mintájának egyezésére vonatkozóan egy sokkal (két nagyságrenddel) magasabb valószínűségi hányados érték jött volna ki. Ez pedig azzal a súlyos következménnyel járt volna, hogy az igazságügyi szakértő a terhelt kárára téved, ami ellentétes az ártatlanság vélelmével.

4. ÚJ MEGÁLLAPÍTÁSOK, A KÖVETKEZTETÉSEK LEVONÁSA

A doktori értekezésemben tárgyalt populációs felméréssel Magyarországon elsőként végeztük el polimorf mikroszatellita (STR) lokuszok igazságügyi genetikai vizsgálatát. Úgy érzem, megteremtettünk egy olyan körülményrendszert, ami az elmúlt évtizedben hozzájárult a magyarországi igazságügyi genetikai szakértés fejlődéséhez mind a laboratóriumi technológia szintjén, mind elméleti, tudományos értelemben.

4.1.

A világ igazságügyi genetikai vérkeringéséhez azoknak a mikroszatellita polimorfizmusoknak az elemzésével csatlakoztunk, amelyeket az igazságügyi genetikus társadalom a genetikai ismeretek gyarapodásával befogadott és magáévá tett. Az általunk kiválasztott 17 autoszómás, egy X-kromoszómához kapcsolt és kilenc Y-kromoszómához kapcsolt lokusz döntő többsége, 23 lokusz a nemzeti bűnügyi DNS-adatbázisok ill. az Y-haplotípus referencia adatbázisok (YHRD) vizsgálati marker-készletét képezi.

A kiválasztott 27 STR lokusz elemzésével genetikai analízist (genotipizálást) végeztünk magyarországi populációkon. A referencia mintának választott kevert budapesti populációs mintán mindegyik lokuszt genotipizáltuk. Az igazságügyi aspektusú STR analízisekbe a világon elsőként vontuk be a roma populációt, amely Magyarország legnagyobb lélekszámú kisebbsége. A budapesti referencia és a baranyai roma populációs minta igazságügyi genetikai analízisét két lokusz kivételével valamennyi STR markerre nézve, referált nemzetközi tudományos folyóiratban publikáltuk. A budapesti referencia és a baranyai roma populációs adatbázis az európai igazságügyi STR munkacsoport által az elmúlt évtized végén felállított online Y-STR haplotípus referencia adatbázis (Eu-YHRD) részét képezi. Az adatbázis jelenleg (2003.08.15.) 84, zömében európai kaukázusi populációból származó 12.904 férfi 9-lokuszos haplotípusát tartalmazza. A baranyai roma mintánk mellett a 2000. évtől lehetőségünk nyílt egy másik, kelet-magyarországi roma populáció elemzésre is. Ugyancsak az ezredfordulótól populációs mintáink körét egy budapesti askenázi zsidó és egy fülöpszállási kiskun mintával is kiegészíthettük. Az öt populációs mintát alkotó, összesen 754 személyt legalább 22 STR lokuszon analizáltuk.

4.2.

Az igazságügyi genetikai vizsgáló-rendszerek eredményeinek megbízhatóságát a vizsgálatokat kivitelező laboratóriumoknak biztosítaniuk kell. E hitelesítés keretében, a genetikai analízisünk döntő többségében alkalmazott automatizált elektroforetikus rendszerek (ALF és ABI Prism 310 szekvenátorok) allélméret-meghatározási pontosságát és precizitását leteszteltük. Belső méretstandard alkalmazásával (ALF) ill. az eredetileg ajánlott belső méretstandardnak (GeneScan-500) egy másik, több fragmensből álló standarddal (CXR60-400) való helyettesítésével (ABI) az allélméret-meghatározás pontosságára irányuló vizsgálataink mindegyik analizált STR rendszer esetében kielégítő eredményre vezettek. Mindkét genotipizáló rendszerben a számított allélméret maximuma, az allél-ablak nem érte el az 1 nt-ot, egy-egy lokusz kivételével a teljes mérési tartományon belül 0,5 nt alatt maradt. A számított allél-méret szórása 0,15 nt-nál kisebbnek adódott. A kapott eredmények alapján az általunk kis mértékben továbbfejlesztett PCR-fragmens analizáló rendszerek alkalmasak az STR lokuszok 1 nukleotid pontossággal történő genotipizálására az igazságügyi genetikai szakértői gyakorlatunkban.

4.3.

Az igazságügyi genetikai vizsgálatok alapkövetelménye a genotípus pontos leírása. Az STR markerek esetében ez referencia-allélok (allél-létra) segítségével történik. Ha az ismeretlen allél nem felel meg egyik referencia-allélnak sem, a tipizálandó allélt – lehetőleg szekvenálással – jellemezni kell.

Az STR markereken végzett kriminalisztikai DNS-vizsgálataink során több lokuszon olyan allélokot is detektáltunk, amelyek a hazai és a külföldi populációs felmérések eredményei alapján igen ritkának számítottak vagy az észlelés időpontjában éppenséggel még nem voltak ismertek. Négy lokusz (D7S820, D13S317, HPRTB, DYS385) esetében ezen allélok közös vonása az volt, hogy a lokuszok allél-létráiból hiányoztak. Az ötödik esetben, a DYS19 markernél egy két-allélos genotípust figyeltünk meg. A ritka allélok döntő többségét a mikrovariáns (nem konszenzus), ún. interallélok tették ki. Négy lokuszon, összesen 12 allélvariáns típust detektáltunk, amelyek közül tizet szekvenálással is jellemeztünk. Mind az STR lokusz polimorf ismétlődő szakaszát, mind a repeateket határoló régiót érintő mutációkat is megfigyeltünk.

Nyolc DYS385 *-1 interallélt hordozó személy nagyban hasonló Y-haplotípusuk és legalább hét férfinál azonos Y-haplocsoportjuk alapján nem túlságosan régen élt közös

férfi őssel rendelkezhetett. Ez a nyolc személy négy populációs mintából származott. Ezek az eredmények a budapesti referencia, a baranyai roma, a budapesti askenázi és a kiskun populációk férfi tagjainak keveredésére (admixture) utalhatnak.

A mikrovariáns alléltípusok általunk detektált relatíve magas aránya felhívta a figyelmünket arra, hogy a PCR-fragmens analízisen alapuló genotipizáló rendszerek felbontásának növekedésével az alléltípusok száma is növekszik. A genotipizálási eredményeknek összehasonlíthatóaknak kell lenniük laboratóriumok között, és erre nézve az allélvariánsok növekvő száma egyre nagyobb kihívást jelent. Potenciálisan hasonló kompatibilitási problémát okozhat, ha a mutáció az STR lokusz PCR-primer kötő régiójában következik be, ahogy azt a D13S317 lokusz ritka 7-es alléljában láttuk. A primerkötő-régiót érintő megváltozás akár az allél elvesztését (drop-out) is okozhatja a PCR-amplifikálás során, amit ennél az allélnél meg is figyeltünk. Ugyanakkor az allélvariánsok megnövelhetik a DNS-vizsgálatok bizonyító erejét, ha a tettes és a gyanúsított DNS-mintája ugyanazt a ritka variánst hordozza.

4.4.

Az allél- és STR profil gyakorisági adatbázisokat tesztelni kell arra nézve, hogy felhasználhatók-e a DNS-vizsgálat bizonyító erejének korrekt megbecsléséhez. A valószínűségszámítás Bayes-elvén alapuló igazságügyi statisztikai interpretáció során alapvetően két DNS-profil, a tettes és a gyanúsított DNS-profiljának egyezési valószínűségét kell kiszámítani. Az STR profil egyezési valószínűség kalkulációjához számos paramétert és körülményt meg kell ismerni, amelyek közül a legfontosabbak: a vizsgált STR lokuszok polimorfizmusának foka, az allélok lokuszokon belüli és lokuszok közötti lehetséges asszociációja, ill. a teljes genetikai variancia megoszlása személyeken belül, a személyek között a populáción belül és a populációk között.

Az STR lokuszok polimorfizmus fokának megbecslésére irányuló vizsgálataink eredménye alapján megállapítottuk, hogy 13 autoszómás STR lokusz (az ún. CODIS USA bűnügyi adatbázis lokuszok) együttes analízise kellően hatékony a magyarországi igazságügyi személyazonosításra nézve. A hatékonysági értékek mindegyik magyarországi populációban rendkívül magasnak és egymással közel azonos nagyságrendűnek bizonyultak. A 13 lokuszra kombinált átlagos egyezési valószínűség értéke (pM) 10^{-13} – 10^{-15} között mozgott a populációs mintákban. A kombinált átlagos szülősségi kizáró erő (PE) ugyanerre a lokusz-készletre vonatkozóan 0,99998-0,99999 értéknek adódott. A 13 autoszómás CODIS STR marker analízisével potenciálisan

biztosítható a Magyarországon elkövetett bűncselekményekben érintett személyek individualizációja.

Az összesített populációs mintára (299 férfira) a teljes, kilenc Y-STR lokuszos haplotípus diverzitás 0,987-nek adódott. Ez azt jelenti, hogy két személyt véletlenszerűen kiválasztva a népességből 98,7% a várható valószínűsége annak, hogy a két személy eltérő haplotípust hordoz az Y kromoszómáján, tehát az igazságügyi vizsgálatok során egymástól megkülönböztethető, feltéve, ha a két személy nincs férfiágú rokonságban egymással. Az Y-STR haplotípusok igazságügyi megkülönböztető ereje az autoszómás STR profilokéhoz képest tehát több nagyságrenddel kisebbnek bizonyult. A kisebb személyazonosítási hatékonyságot azonban az Y-kromoszóma három, igazságügyi alkalmazásban rendkívül kedvező tulajdonsága ellensúlyozza. Férfi-specifikus, tehát a túlnyomó részben férfi elkövetők genetikai profilja a női genom jelenlétében is közvetlenül kimutatható. Erre esettanulmányunkban mutattunk be egy példát. Az Y-kromoszóma legnagyobb része nem rekombinálódik, ezért férfiágú rokonsági vonalak beazonosításához kiválóan alkalmas. Mindezek mellett és mindezek miatt populáció-specifikus jellege könnyebben detektálható, amelyet a roma populációra nézve mi is kimutattunk.

Az Y-haplotípusok és az egyedi haplotípusoknak az egyes populációkban megfigyelt aránya alapján a két roma populáció élesen elkülönült a többi populációtól. A detektált kisebb géndiverzitásért ezekben a populációkban elsősorban a haplotípusok közös leszármazása lehet a felelős, ami a populációk zártságára és beltenyésztésére utalhat. A budapesti referencia populáció haplotípus-diverzitás értéke hasonlóan bizonyult más kaukázusi populációkban detektált diverzitás értékekhez. Ez a megfigyelés a referenciául választott populációs mintánk igazságügyi célú felhasználásának megbízhatóságát bizonyította.

Az Y-haplotípusok általában populáció-specifikus megoszlást mutattak, mert ugyanazt a haplotípust nem vagy csak ritkán figyeltük meg a többi populációban. Ugyanakkor két haplotípus a roma populációkban extra-domináns volt. Ezeket a haplotípusokat egy bulgáriai roma populációban is megtaláltuk. E két haplotípus előfordulási gyakorisága nagyon kicsi volt az európai Y-STR adatbázis nem roma részében. Ezek a megállapítások alátámasztják azt a feltételezést, hogy a kérdéses haplotípusok roma specifikusak. Az Y-kromoszómás haplotípusok lehetséges populáció-specifikitása esélyt adhat a biológiai nyomot hátrahagyó személy etnikai hovatartozásának megbecsüléséhez az igazságügyi szakértői gyakorlatban.

4.5.

Az előző pontban a diploid, rekombinálandó genomban lévő STR lokuszok kombinált átlagos igazságügyi személyazonosító hatékonyságát bemutattuk. A hatékonysági értékek kiszámításakor feltételeztünk, hogy a DNS-profilban lévő allélok függetlenek egymástól. Ez a humán populációkban nincs feltétlenül így.

A lokuszon belüli potenciális allélikus asszociáció statisztikai tesztelése (Hardy-Weinberg) eredményeként az egyensúlytól való szignifikáns különbséget kaptunk a baranyai roma populáció TPOX, D3S1358, CSF1PO, D7S820, LPL lokuszai valamint a kelet-magyarországi roma minta D13S317 lokusza esetében. A Hardy-Weinberg egyensúlytól való jelentős, bár nem magasan szignifikáns eltérés legvalószínűbb magyarázata a mintavétel bizonytalansága és/vagy a roma populációk potenciálisan magasabb fokú beltenyésztése lehet. Ugyanakkor, a Bonferroni-procedúrával kapott eredményeink alapján nem vethettük el a HWE hipotézisét, ha az egyensúlyt a populációk teljes genomi szintjén vizsgáltuk. Ennek alapján helyesnek tűnt az az elképzelésünk, hogy a lokuszok genotípusának gyakoriságát az allélgyakorisági adatok alapján megbecsülhetjük.

Az allélok lokuszok közötti függetlenség-vizsgálatának (linkage disequilibrium tesztelés) eredményei alapján megalapozottnak tűnik az a feltételezés, hogy a baranyai roma populációs mintában genom szinten jelentős linkage disequilibrium van. A lokuszok zömének tényleges genetikai kapcsoltsága a lokuszok eltérő kromoszómás elhelyezkedése miatt kizárható. Így az általunk populációs szinten megfigyelt linkage disequilibriumot más tényezők, pl. a roma populáció potenciális beltenyésztése és/vagy genetikai strukturáltsága okozhatta. Korábban kimutatták, hogy kevert populációk esetében (a Wahlund-hatáshoz analóg módon) linkage disequilibrium lehet a teljes populációban annak ellenére, hogy az egyes alpopulációk egyensúlyban vannak.

A megállapításaink alapján, legalábbis négy populációs minta esetében helyesnek tűnt az az elképzelésünk, hogy a több lokuszos DNS-profilok gyakoriságát az allélgyakorisági adatok szabad kombinálásával megbecsülhetjük. A baranyai roma populációs minta genomi szintű kapcsoltsági egyensúlytalanságának hipotézisét az Expectation-Maximization algoritmuson alapuló valószínűségi hányados statisztika lokuszpárookra történő alkalmazásával állítottuk fel. Ez két problematikus pontot vethet fel. Először, az EM-algoritmus kimutatottan érzékeny a Hardy-Weinberg egyensúly meglétére vagy hiányára. A HWE-től való jelentősebb eltérés esetén a statisztika tévesen elvetheti a linkage equilibriumot. A baranyai roma populációban ez a jelenség nem zárható ki.

Másodszor, a lokuszpárokban detektált allélikus asszociáció nem feltétlenül jelentős mértékű, ha az asszociációt lokusz-trión, lokusz-kvartetten stb. belül vizsgáljuk. Ezt igazolják a nagy bűnügyi DNS-nyilvántartásokban elvégzett elemzések eredményei, amely szerint a vizsgált lokuszok számának emelésével, folyamatosan növekszik a személyazonosítás hatékonysága, nincs plató-effektus. E két megállapítás alapján feltételezhetjük, hogy a baranyai roma populáció detektált allélikus asszociációjának nincs lényeges hatása a DNS-profil gyakoriságának kiszámítására nézve. A probléma megnyugtató tisztázásához azonban további statisztikai genetikai analízisek szükségesek.

4.6.

Három populációstatisztikai teszt [G-statisztika, F-statisztika, molekuláris varianciaanalízis (AMOVA)] segítségével, páronként jellemeztük az öt vizsgált magyarországi populáció genetikai strukturáltságát annak tisztázása érdekében, hogy a budapesti referencia-adatbázisunk allél- és haplotípus-gyakorisági értékei vajon milyen mértékben nevezhető hitelesnek a magyar népeiségre nézve.

Az öt populációs minta genetikai összetételében, a lokuszokon mért allélgyakoriságok és allél-struktúrák és haplotípusok eloszlása tekintetében szingifikáns különbségeket detektáltunk.

Az autoszómás lokuszok esetében valamennyi statisztikai teszt a legtöbb szignifikáns eltérést a baranyai roma/budapesti referencia ill. a baranyai roma/budapesti askenázi populáció-párosban mutatta. A G- és F-statisztika alapján figyelemre méltó volt a roma populációk egymáshoz viszonyított nagyfokú heterogenitása. A roma populációknak az allélgyakoriságok szintjén tapasztalt heterogenitását az allélok molekuláris szerkezetén alapuló AMOVA csak csekély mértékben támasztotta alá. A populáció-párok közül a legkisebb genetikai tagozódást a budapesti referencia/kiskun párosban kaptunk.

A vizsgált 17 autoszómás STR között nem volt egy olyan marker sem, amely legalább egy tesztben ne adott volna szignifikáns eltérést. Azon 13 lokusz közül, amelyet mindegyik populációs mintán elemeztünk, a populáció-páronkénti összehasonlításban a legkisebb változatosságot a TH01 lokuszon tapasztaltunk. A legnagyobb genetikai variációt a 13 CODIS STR törzs-lokuszt közül a D21S11 lokuszon detektáltuk a páros összehasonlítások során.

Három populáció-párban, nevezetesen a budapesti referencia/baranyai roma, baranyai roma/askenázi és baranyai roma/kiskun mintapárban a 13 autoszómás lokuszra vonatkoztatott F-statisztikai értékek meghaladták a 0,02 szintet. Mivel nemzetközi

felmérések során nagyobb F_{ST} -értéket elsősorban izolált ill. eltérő genetikai eredetű népcsoportokban detektáltak, ezért megfigyeléseink a budapesti referencia, baranyai roma, askenázi, és a kiskun populáció eltérő genetikai múltjára ill. izoláltságára utalhat.

A filogenetikai vizsgálatunk eredménye is azt mutatta, hogy a nyolc autoszómás STR markerrel elemzett kaukázusi populáció-csoport (budapesti referencia, USA és lengyel populáció) valamint a baranyai roma populáció között viszonylag távolabbi a genetikai rokonság.

Ha a populációk páronként történő AMOVA-elemzése során a 13 autoszómás lokuszt együtt vizsgáltuk, nem kaptunk jelentős különbséget a budapesti referencia/kiskun párban, valamint a roma populációk között. A többi párosításban azonban a molekuláris varianciaanalízis szignifikáns különbséget jelzett.

Ha a populációkat nem párokban, hanem együtt vizsgáltuk, a megfigyelt F_{ST} -értékei alapján a 13 CODIS STR marker közül a D7S820 lokusz mutatta a legnagyobb variabilitást. Ugyancsak magas variációt kaptunk a TPOX lokusz esetében. A CSF1PO mikroszatellita mutatta a legkisebb változékonyságot az adatbázisok között. Ezen és a populáció-párokra kapott eredmények alapján az analizált 13 autoszómás STR marker közül a vizsgált magyar népességben a CSF1PO és TH01 lokusz bír a legkisebb, a D7S820 és D21S11 lokusz rendelkezik a legnagyobb információ-tartalommal a populációk elkülöníthetősége szempontjából. Az öt populáción mért F_{IS} -értékek nagyságából arra a következtetésre jutottunk, hogy a populációkon belüli beltenyésztés foka nem teljesen elhanyagolható. A magasabb F_{IS} értéket elsősorban az egyéneknek a populáción belüli nem véletlenszerű párválasztása okozhatta.

Egy szlovén, egy olasz és két spanyol populációs mintával való, lokuszonkénti összehasonlításban a budapesti referencia populáció csak az olasz mintától mutatott szignifikáns eltérést a G-statisztika eredménye alapján. A budapesti referencia populációs mintánk allélgyakoriság-eloszlása tehát nagyfokú hasonlóságot mutatott más európai népcsoportokkal. Ezek az eredmények egybevágtak azzal a tudományos nézettel, amely szerint a kevert magyar népesség az indoeurópai népekkel közeli genetikai rokonságban áll.

Az öt populáció páronkénti összehasonlításában – a 13 autoszómás STR egyszerre történő vizsgálatánál kapott eredménnyel összhangban – az Y-STR haplotípusokon elvégzett AMOVA nem mutatott jelentős különbséget a roma populációk között. Ezzel ellentétben mindkét roma minta – az askenázi populáció kivételével – mindegyik populációtól jelentősen eltért. Az AMOVA-analízis alapján nem volt jelentős különbség a

budapesti referencia és az askenázi populációs adatbázis között. Az askenázi populációra kapott eredményeink összhangban voltak azzal a közelmúltban publikált kutatási eredménnyel, amely szerint a zsidó populációk Y-kromoszómái nagyfokú hasonlóságot mutatnak a környezetükben lévő, nem-zsidó populációk Y-kromoszómájával, és nincs alapító hatásra utaló jel.

41 európai férfi populáció Y-STR haplotípusokra irányuló molekuláris varianciaanalízise alapján a balti, norvég és a budapesti referencia férfi populáció kis mértékben, a németországi törökök és a baranyai roma populáció jelentősebben elkülönült a vizsgált európai populációk többi tagjától. Az AMOVA alapján a balti populáció-csoport ill. török és a roma populáció elkülönülése statisztikailag szignifikánsnak bizonyult.

4.7.

Az öt magyarországi populációs mintánk megfigyelt genetikai strukturáltsága felhívta a figyelmünket arra, hogy a magyar népesség genetikai tagoltsága igazságügyi genetikai aspektusból jelentős lehet. A vizsgált populációk genetikai tagoltságának az igazságügyi statisztikai interpretációra gyakorolt potenciális hatását négy lépcsőben elemeztük. Mindegyik analízissel azt a bűnügyi helyzetet szimuláltuk, amikor a tettes és a gyanúsított egy személy, azaz a bűncselekmény során keletkezett biológiai nyom a gyanúsítottól származott. Az első három szakaszban, a profil-gyakoriság kiszámításakor nem vettük figyelembe a populációs minták általunk meghatározott genetikai variancia-értékeit.

Az átlagolt fenotípus-különbségi értékek (MVs) analízise során azt vizsgáltuk, hogy maximálisan négy STR lokusz analízise során a budapesti referencia adatbázis alkalmazása milyen hatással van a baranyai roma személyek STR profil frekvenciájára és vice versa. Eredményeink azt mutatták, hogy minden összehasonlítás során annak az allélgyakorisági adatbázisnak az alkalmazásakor kaptunk átlagosan nagyobb fenotípusgyakoriságot, amelyikből a vizsgált DNS-profil származott. Az eltérés mértéke a lokuszok fokozatos bevonásával kismértékben bár, de növekedett. Ezekből a megállapításainkból azt a következtetés vontuk le, hogy a magyar népesség genetikai szubstrukturáltsága az igazságügyi DNS-elemzés bizonyító erejének statisztikai interpretációját jelentősen befolyásolhatja, ha a vizsgált STR lokuszok száma 1-4 körül mozog. Ráadásul, ha nem a potenciális tettesek alpopulációjából választjuk ki az allélgyakorisági értékeket, akkor a gyanúsított kárára tévedhetünk, mivel az általunk

megfigyelt, átlagosan kisebb DNS-profil gyakoriság miatt az egyezési valószínűség értéke is átlagosan kisebb lesz.

A második lépcsőben, annak érdekében, hogy a bűnügyi szimulációba bevont DNS-profilok száma minél nagyobb legyen, a populációs mintákat alkotó személyek DNS-profiljainak egyesítésével egy nagy DNS-profil készletet (poolt) hoztunk létre. Már 358 budapesti referencia ill. baranyai roma személyt tartalmazó profil-készlet négy-négy lokuszon történő analízise is azt mutatta, hogy a különböző adatbázisokból számított profilgyakorisági értékek között akár majdnem két nagyságrend különbség is lehet. E vizsgálat eredményéből is azt a következtetést vonhattuk le, hogy a magyarországi bűncselekmények bizonyítása során elvégzett, kevés lokuszt érintő igazságügyi statisztikai interpretáció számos esetben méltánytalanul elfogult profilgyakoriságot eredményezhet akkor, ha az egyik populáció allélgyakorisági adatait a másik populációra alkalmazzuk.

Harmadik lépcsőben, a nagy igazságügyi személyazonosító erővel rendelkező multiplex STR genotipizáló-rendszerek alkalmazásával kíváncsiak voltunk arra, hogy a populációs adatbázis-választás fentiekben kifejtett hatása további három magyarországi populációs minta, és összesen 13 autoszómás STR lokusz (a CODIS törzs-lokuszek) bevonásával is érvényesül-e. Az elemzésbe bevont 754 DNS-profil egyezési valószínűségének reciprokát, a valószínűségi hányados értékeket (LR) ún. sörét- és Tippett-diagramokon értékeltük ki. A vizsgálat során valamennyi DNS-profil valószínűségi hányados értéke meghaladta az $LR = 10^{11}$ szintet, függetlenül attól, hogy melyik populációs adatbázis allélgyakorisági értékével számoltunk. Az $LR = 10^{11}$ szint (1 a százmilliárdhoz) a magyar népességben már rendkívül nagy igazságügyi bizonyító erővel rendelkezik. A valószínűségi értékek a baranyai roma populációval való összehasonlítások során szórtak a legjobban, amiért valószínűleg a baranyai populációnak és a többi populációnak – az allélgyakoriság-eloszlás szintjén tapasztalt – elkülönülése a felelős. A budapesti referencia populációval való összehasonlítást úgy is elvégeztük, hogy kiértékelés során csak azokat a profilokat vettük számításba, amelyeknek a budapesti referencia allélgyakorisági adatbázisból számolt LR-értékei meghaladták a többi adatbázisból kalkulált LR-értékeket. A négy lehetséges párosítás közül kettőben, a baranyai roma és a kiskun adatbázissal való összehasonlításban a profilok 11-12%-a mutatott min. két nagyságrend LR különbséget. Ez az alkalmazott modellünk alapján azt jelenti, hogyha a tettes a két populáció valamelyikéből származik, akkor az LR-érték megbecsülésekor várhatóan az esetek kb. 10%-ában tévednénk legalább két nagyságrendet

a vádlott kárára. Ez a tévedési arány és mérték igazságügyi kontextusban jelentősnek mondható.

Az öt magyarországi populáció igazságügyi statisztikai aspektusú összehasonlításának utolsó lépcsőjében azt vizsgáltuk, hogy az LR-értékek és különbségeinek eloszlása hogyan módosul, ha a profil egyezési valószínűség számítását a budapesti referencia adatbázis esetében az Ayres és Overall által javasolt módon korrigáljuk. A egyezési valószínűség számítás módosított formája figyelembe veszi a népesség egyes (igazságügyi értelemben a potenciális bűnelkövetők) alpopulációjában lévő személyek közötti genetikai korrelációt. A kiértékelés során itt már csak azokat a profilokat vettük számításba, amelyek LR értéke a budapesti referencia allélgyakorisági adatbázis használatával magasabbnak bizonyult. A budapesti referencia populációt tartalmazó négy lehetséges párosításban a profilok legfeljebb 2,5%-a mutatott min. két nagyságrend LR különbséget. Ez az alkalmazott alpopulációs modellünk alapján azt jelenti, hogyha az igazságügyi statisztikai interpretáció során a referencia budapesti allélgyakorisági adatbázissal történik a kalkuláció, akkor várhatóan az esetek legfeljebb 2-3%-ában tévednénk min. két nagyságrendet a terhelt kárára. Figyelembe véve azonban az LR értékek nagyságrendjét (10^{11} - 10^{25}) és az LR különbség-eloszlási görbék gyors lecsengését, levonhattuk azt a következtetést, hogy a detektált különbség igazságügyi szempontból elhanyagolható.

4.8.

A felmért öt magyarországi populáció genetikai összetételében különbségeket mutattunk ki. A budapesti referencia populációtól legnagyobb mértékben a baranyai roma populáció különbözött. Kíváncsiak voltunk, hogy ebben az extra esetben a két populációt alkotó személyek etnikai származását 1-17 autoszómás STR lokusz alléljaiból álló DNS-profiljuk alapján milyen sikerességgel lehet helyesen megbecsülni. E két populáció a magyar népeiséget és annak egy jelentős lélekszámú kisebbségét hivatott reprezentálni a szimulációban. Az STR markerek etnikai származás megállapító képességét e két magyarországi mintát tartalmazó populáció-csoport mellett másik hat külföldi populáció két csoportjában (osztrák kaukázusi/egyiptomi arab ill. négy USA populáció) is felmértük. Az STR profil gyakoriságára épülő etnikai eredet becslés (EEO) általunk választott megközelítése újnak nevezhető abban az értelemben, hogy a korábban leírt módszerekkel ellentétben a helyes és téves becslések várt aránya, így a procedúra megbízhatósága az eredményekből kiolvasható.

A magyarországi populáció-csoportban a az EEO helyes eredményt adott a személyek közel 90%-ára. Az EEO az osztrák/egyiptomi párra már csak 80%-ban volt megfelelő. A harmadik, USA-csoportban az ázsiai és az afro-amerikai populációra kapott 80% körüli sikeres azonosítás meglehetősen magas értéknek számít, figyelembe véve a lehetséges populációk számát. Az USA kaukázusi és a spanyol ajkú népességének külön-külön történő beazonosítása gyakorlatilag nem volt megbízható. Ha a kaukázusi vagy spanyol ajkú etnikai származást nem kívántuk tisztázni, akkor a két populációt alkotó személyek 93%-a sikeresen elkülöníthető volt az ázsiai és afro-amerikai populációtól. A helyes és téves etnikai eredet-becslések aránya a vizsgált lokuszok számának növelésével kb. 15 db lokuszig jelentősen emelkedett, 15-nél több lokusz bevonásával a hatékonyság már nem nőtt számottevően. Az eredményeink arra utalhatnak, hogy a globalizált társadalmakban teljesen izolált, érintetlen etnikai közösségek gyakorlatilag nem léteznek.

Az eredményeinkből levonhatjuk azt a következtetést, hogy az igazságügyi genetikai gyakorlatban vizsgált autoszómás STR markerek alkalmasak lehetnek a bűncselekmény elkövetésekor biológiai nyomot hátrahagyó tettes etnikai származásának megbecslésre Magyarországon is. Ugyanakkor az etnikai eredet-becslés általunk választott, viszonylag egyszerű módszerének van egy nagyon lényeges alapfeltétele. Nevezetesen, a EEO megbízhatóságára vonatkozó eredményeink csak akkor helyesek, ha azzal a feltételezéssel élünk, hogy az ismeretlen elkövető egyenlő *a priori* eséllyel tartozhat a bűncselekmény szempontjából releváns valamennyi populáció-csoportba. Ha pl. az egyik populáció lényegesen nagyobb a másikinál, a kiindulási valószínűség a nagyobb lélekszámú populáció esetén magasabb. Mindezek mellett a téves eredetbecslés viszonylag magas várt aránya miatt az eredmények interpretációjakor mind az igazságügyi szakértőknek, mind a nyomozó hatóságnak rendkívüli körültekintéssel kell eljárnia.

Eredményeink összefoglalása után megvalósultnak hiszem azt a célkitűzésemet, amely arra irányult, hogy polimorf mikroszatellita lokuszok genotipizálása révén felállítsunk egy olyan budapesti referencia populációs adatbázist, ami megfelelőnek bizonyul a magyarországon elkövetett bűncselekmények szakértői vizsgálatának korrekt statisztikai genetikai interpretációjára.

Megállapítottuk, hogy a magyar népesség genetikai szubstrukturáltsági foka lényeges hatással lehet a DNS-vizsgálat bizonyító erejének statisztikai kiértékelésére. Eredményeink azt mutatták, hogy az alpopulációs teória alapján a magyarországi népcsoportokban élő személyek genetikai korrelációját figyelembe kell venni a DNS-

profil egyezési valószínűség kalkulációja során annak érdekében, hogy a törvényszéki DNS-vizsgálat bizonyító erejének bemutatásakor ne történjen tévedés a terhelt kárára.

Az etnikai eredet megbecsülésének megbízhatóságára irányuló tesztvizsgálataink eredménye alapján megállapíthattuk, hogy az igazságügyi genetikai gyakorlatban vizsgált autoszómás STR markerek – bizonyos feltételek teljesülése esetén – alkalmasak lehetnek a bűncselekmény elkövetésekor biológiai nyomot hátrahagyó tettes etnikai származásának megbecsülésére Magyarországon is.

5. KÖSZÖNETNYILVÁNÍTÁS

A doktori értekezésemben bemutatott eredmények csapatmunka eredményei. Rengetegen dolgoztak és fáradoztak azon, hogy ezek az eredmények megszülethessenek. Ha esetleg bármelyikük is úgy érezné, hogy kihagytam őket az alábbi felsorolásból, az csak a véletlen műve és természetesen az én hibám.

Elsőként szeretném megköszönni mindazon személyeknek az együttműködését, akik biológiai mintáikkal nagylelkűen hozzájárultak a magyarországi igazságszolgáltatás e kicsiny, de véleményem szerint nagyon fontos szeletének (remélem) jobbá tételéért. Nélkülük ez a doktori értekezés nem jöhetett volna létre.

Köszönetet szeretnék mondani témavezetőmnek, dr. Lontainé dr. Santora Zsófiának, az ORFK Bűnügyi Szakértői és Kutatóintézet igazgatójának, hogy az igazságügyi genetikai szakértés magyarországi bevezetésében úttörő szerepet vállalva biztosította kutatásom feltételeit, felkarolta és támogatta mindazokat a kezdeményezéseimet, amelyek a szakértői és kutatómunkám megfelelőbb végrehajtására irányultak.

Köszönetet kell mondanom dr. Fodor Andrásnak, hogy az ELTE TTK Genetika Tanszéke felé köldökzsinórom maradt, miután az Alma Mater gondoskodó és fejlődést biztosító anyaméhét diplomám megszerzése után elhagytam.

Köszönetet mondok Pádár Zsolt, Woller János és Egyed Balázs kollégámnak a szakértői és kutatómunkámban adott rengeteg segítségért, a sok hasznos tanácsért és konzultációért, valamint azért, hogy mindig készséggel álltak a rendelkezésemre, ha a munkámmal kapcsolatban gondjaim adódtak.

Köszönettel tartozom az ORFK BSZKI Hemogenetikai Osztály összes dolgozójának, különös tekintettel az asszisztensekre, akik rengeteg segítséget adtak a vizsgálatok kivitelezésében.

Köszönöm Kontár Katalinnak és Hegedűs Zitának, az ELTE TTK két volt biológus hallgatójának, hogy segítséget nyújtottak a populációgenetikai felmérésekben.

Végül köszönet illeti mindazokat a személyeket és intézményeket, akik rendelkezésünkre bocsátották a populációs mintákat:

- Az Országos Vérellátó Szolgálat munkatársainak a budapesti referencia mintáért,
- dr. Béres Juditnak (Országos Epidemiológiai Központ) a kiskunsági mintákért,

- dr. Angyal Miklósnak (Pécsi Tudományegyetem) a baranyai roma mintákért,
- dr. Tordai Attilának és dr. Kalmár Lajosnak (Országos Gyógyintézeti Központ, Hematológiai és Immunológiai Intézet) az askenázi mintákért, és külön dr. Kalmár Lajosnak a vizsgálatok kivitelezésében nyújtott segítségéért,
- és végül, de nem utolsó sorban dr. Balogh Istvánnak (Debreceni Tudományegyetem, OEC) a kelet-magyarországi roma mintákért.

6. IRODALOMJEGYZÉK

- Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997)** Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25: 3389–3402
- Ayres KL, Overall ADJ (1999)** Allowing for within-subpopulation inbreeding in forensic match probabilities. *Forensic Sci Int* 103:207-216
- Balding DJ (1995)** Estimating products in forensic identification using DNA profiles. *J Am Stat Assoc* 90:839-844
- Balding DJ and Nichols RA (1994)** DNA profile match probability calculation: how to allow for population stratification, relatedness, database selection and single bands. *Forensic Sci Int* 64:125-140
- Bär W, Brinkmann B, Budowle B, Carracedo A, Gill P, Lincoln P, Mayr WR, Olaisen B (1997)** DNA recommendations. Further report of the DNA Commission of the ISFH regarding the use of short tandem repeat systems. *International Society for Forensic Haemogenetics. Int J Legal Med* 110:175-176
- Botstein D, White RL, Skolnick M, Davis RW (1980)** Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet.* 32:314-331
- Bouffard GG, Iyver LM, Idol JR, Braden VV, Cunningham AF, Weintraub LA, Mohr-Tidwell RM, Peluso DC, Fulton RS, Leckie MP, Green ED (1997)** A collection of 1814 human chromosome 7 specific STSs. *Genome Res* 7: 59–64
- Boutrand L, Egyed B, Füredi S, Moomers N, Mertens G, Vandenberghe A (2001)** Variations in primer sequences are the origin of allele drop-out at loci D13S317 and CD4. *Int J Legal Med* 114:295-297
- Brenner CH (1998)** Difficulties in the estimation of ethnic affiliation. *Am J Hum Genet* 62:1558-1560
- Brinkmann B, Carracedo A (eds) (2003)** *Progress in Forensic Genetics 9*. International Congress Series 1239. Elsevier, Amstredam New York Tokyo
- Brooks AJ (1999)** The essence of SNPs. *Gene* 234:177-186
- Budowle B, Chakraborty C (2001)** Population variation at the CODIS core short tandem repeat loci in Europeans. *Leg Med* 3:29-33

- Budowle B**, Chakraborty R, Giusti AM, Eisenberg AJ, Allen RC (1991) Analysis of the VNTR locus D1S80 by the PCR followed by high-resolution PAGE. *Am J Hum Genet.* 48:137-144.
- Budowle B**, Monson KL, Chakraborty R (1996) Estimating minimum allele frequencies for DNA profile frequency estimates for PCR-based loci. *Int J Legal Med* 108:173-176
- Budowle B**, Monson KL, Giusti AM, Brown BL (1994) The assessment of frequency estimates of Hae III-generated VNTR profiles in various reference databases. *J Forensic Sci* 39:319-352
- Budowle B**, Shea B, Niezgoda S, Chakraborty C (2001) CODIS STR loci data from 41 sample populations. *J Forensic Sci* 46:453-489
- Butler JM and Reader DJ** (2003) STRbase: Short Tandem Repeat DNA Internet Database. <http://www.cstl.nist.gov/biotech/strbase>
- Cavalli-Sforza LL**, Menozzi P, Piazza A (1994) *Geography of Human Genes*. Princeton University Press
- Chakraborty R** (1992) Sample size requirements for addressing the population genetic issues of forensic use of DNA typing. *Hum Biology* 64:141-159
- Chen H**, Lowther W, Avramopoulos D, Antonarakis SE (1994) Homologous loci DXYS156X and DXYS156Y contain a polymorphic pentanucleotide repeat (TAAAA)_n and map to human X and Y chromosomes. *Hum Mutat* 4:208-211
- Comey CT**, Koons BW, Presley KW, Smerick JB, Sobieralski CA, Stanley DM, Baechtel FS (1994) DNA extraction strategies for amplified fragment length polymorphism analysis. *J Forensic Sci* 39:1254-1269
- Corach D**, Filgueira Risso L, Marino M, Penacino G, Sala A (2001) Routine Y-STR typing in forensic casework. *Forensic Sci Int* 118: 131-135
- Csete K**, Shurenkamp M, Varga T (1996) The STR system HumVWA and HumACTB2 in a Hungarian population. *Int J Legal Med* 108:316-317
- Czeizel A**, Benkmann H-G, Goedde HW (1991) *Genetics of the Hungarian Population*. Springer-Verlag, Berlin, Heidelberg, New York
- Decorte R**, Cassiman J-J (1996) Evaluation of the ALF DNA sequencer for high-speed sizing of short tandem repeat alleles. *Electrophoresis* 17: 1542-1549
- Deka R**, Shriver MD, Yu LM, Ferrel RE, Chakraborty R (1995) Intra- and inter-population diversity at short tandem repeat loci in diverse populations of the world. *Electrophoresis* 16:1659-1664

- Edwards A**, Civitello A, Hammond HA, Caskey CT (1991) DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. *Am J Hum Genet* 49:746-756
- Egyed B**, Füredi S, Angyal M, Boutrand L, Vandenberghe A, Woller J, Pádár Z (2000) Analysis of eight STR loci in two Hungarian populations. *Int J Legal Med* 113:272-275
- Entrala C**, Lorente M, Lorente JA, Carlos Alvarez J, Moretti T, Budowle B, Villanueva E (1998) Fluorescent multiplex analysis of nine STR loci: Spanish population data. *Forensic Sci Int* 98:179–183
- Evett IW**, Lambert JA, Buckleton JS, Weir BS (1996) Statistical analysis of a large file of STR profiles of British Caucasians to support forensic casework. *Int J Legal Med* 109:173-177
- Evett IW**, Weir BS (1998) *Interpreting DNA evidence*. Sinauer Associates, Inc. Publishers, Sunderland, Massachusetts
- Excoffier L**, Slatkin M (1998) Incorporating genotypes of relatives into a test of linkage disequilibrium. *Am J Hum Genet* 62:171-180
- Excoffier L**, Smouse P, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131:479-491
- Felsenstein J** (1989) Phylip - phylogeny inference package (version 3.2). *Cladistics* 5:164-166
- Fowler JCS**, Burgoyne LA, Scott AC, Harding HWJ (1988) Repetitive deoxyribonucleic acid (DNA) and human genome variation – a concise review relevant to forensic biology. *J Forensic Sci* 33: 1111-1126
- Füredi S**, Angyal M, Kozma Z, Sétáló J, Woller J, Pádár Z (1997) Semi-automatic DNA profiling in a Hungarian Romany population using the STR loci HumVWFA31, HumTH01, HumTPOX, and HumCSF1PO. *Int J Legal Med* 110:184-187
- Füredi S**, Angyal M, Kozma Zs, Sétáló J, Woller J, Pádár Zs (1997b) Semi-automatic DNA profiling in a Hungarian Romany population using the STR loci HumVWFA31, HumTH01, HumTPOX and HumCSF1PO. *Int J Legal Med* 110:184-187
- Füredi S**, Budowle B, Woller J, Pádár Zs (1996) Hungarian population data on six STR loci - HumVWFA31, HumTH01, HumCSF1PO, HumFES/FPS, HumTPOX, and Hum HPRTB - derived using multiplex PCR amplification and manual typing. *Int J Legal Med* 109:100-101

- Füredi S**, Kozma Zs, Woller J, Pádár Zs, Angyal M, Bajnóczky I, Nishi K (1998) Population genetic data on four STR loci in a Hungarian Romany population. *Int J Leg Med* 112:72-74
- Füredi S**, Woller J, Pádár Zs (1995) Hungarian population data for the STR systems TH01 and VWA. *Int J Legal Med* 108:48-49
- Füredi S**, Woller J, Pádár Zs (1997a) A population study of the STR loci HumLPL, HumF13B and HumF13A01 in Hungary. *Int J Legal Med* 110:107-108
- Füredi S**, Woller J, Pádár Zs, Angyal M (1999) Y-STR haplotyping in two Hungarian populations. *Int J Legal Med* 113:38-42
- Garofano L**, Pizzamiglio M, Vecchio C, Lago G, Floris T, D'Errico G, Brembilla G, Romano A, Budowle B (1998) Italian population data on thirteen short tandem repeat loci: HUMTH01, D21S11, D18S51, HUMVWA31, HUMFIBRA, D8S1179, HUMTPOX, HUMCSF1PO, D16S539, D7S820, D13S317, D5S818, D3S1358. *Forensic Sci Int* 97:53–60
- Gehrig C**, Hochmeister M, Borer UV, Dirnhofer R (1995) Allele frequency of 3 chromosome Y-specific STR loci and their application to forensic casework. Proceedings of the 6th international symposium on human identification, Scottsdale, USA
- Gill P**, Brenner C, Brinkmann B, Budowle B, Carracedo A, Jobling MA, de Knijff P, Kayser M, Krawczak M, Mayr WR, Morling N, Olaisen B, Pascali V, Prinz M, Roewer L, Schneider PM, Sajantila A, Tyler-Smith C (2001) DNA Commission of the International Society of Forensic Genetics: recommendations on forensic analysis using Y-chromosome STRs. *Int J Legal Med* 114:305-309
- Guglielmino CR**, Béres J (1996) Genetic structure in relation to the history of Hungarian ethnic groups. *Human Biology* 68:335-355
- Guglielmino CR**, Silvestri A de, Béres J (2000) Probable ancestors of Hungarian ethnic groups: an admixture analysis. *Ann. Hum Genet.* 64:145-159
- Guo SW**, **Thompson EA** (1992) Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics* 48:361-372
- Hammond HA**, Jin L, Zhong Y, Caskey CT, Chakraborty R (1994) Evaluation of 13 short tandem repeat loci for use in personal identification applications. *Am J Hum Genet* 55:175-189

- Heyer E**, Puymirat J, Dieltjes P, Bakker E, de Knijff P (1997) Estimating Y chromosome specific microsatellite mutation frequencies using deep rooting pedigrees. *Hum Mol Genet* 6:799-803
- Higuchi R**, von Beroldingen CH, Sensabaugh GF, Erlich HA (1988) DNA typing from single hairs. *Nature* 322:543-546
- Interpol DNA Unit (2003)** Global DNA Database Inquiry 2002: Results and Analysis. I.C.P.O. Interpol, General Secretariat
- Jeffreys AJ**, Wilson V, Thein SL (1985a) Hypervariable minisatellite regions in human DNA. *Nature* 314:67-72
- Jeffreys AJ**, Wilson V, Thein SL (1985b) Individual-specific "fingerprints" of human DNA. *Nature* 316:76-79
- Jobling MA**, Pandya A, Tyler-Smith C (1997) The Y chromosome in forensic analysis and paternity testing. *Int J Legal Med* 110:118-124
- Jones DA (1972)** Blood samples: probability of discrimination. *J Forensic Sci Soc* 12:355-359
- Kanter E**, Baird M, Shaler R, Balazs I (1986) Analysis of restriction fragment length polymorphisms in deoxyribonucleic acid (DNA) recovered from dried bloodstains. *J Forensic Sci* 31 : 403-408
- Kayser M**, Caglià A, Corach D, Fretwell N, Gehrig C, Graziosi G, Heidorn F, Herrmann S, Herzog B, Hidding M, Honda K, Jobling M, Krawczak M, Leim K, Meuser S, Meyer E, Oesterreich W, Pandya A, Parson W, Penacino G, Perez-Lezaun A, Piccinini A, Prinz M, Schmitt C, Schneider PM, Szibor R, Teifel-Greding J, Weichhold G, Knijff P de, Roewer L (1997a) Evaluation of Y-chromosomal STRs: a multicenter study. *Int J Legal Med* 110:125-133
- Kayser M**, de Knijff P, Dieltjes P, Krawczak M, Nagy M, Zerjal T, Pandya A, Tyler-Smith C, Roewer L (1997b) Applications of microsatellite-based Y chromosome haplotyping. *Electrophoresis* 18: 1602-1607
- Kayser M**, Krawczak M, Excoffier L, Dieltjes P, Corach D, Pascali V, Gehrig C, Bernini LF, Jespersen J, Bakker E, Roewer L, de Knijff P (2001) An extensive analysis of Y-chromosomal microsatellite haplotypes in globally dispersed human populations. *Am J Hum Genet* 68:990-1018
- Kayser M**, Roewer L, Hedman M, Henke L, Henke J, Brauer S, Krüger C, Krawczak M, Nagy M, Dobosz T, Szibor R, de Knijff P, Stoneking M, Sajantila A (2000) Characteristics and frequency of germline mutations at microsatellite loci from the

human Y chromosome, as revealed by direct observation in father/son pairs. *Am J Hum Genet* 66:1580-1588

Kayser M, Sajantila A (2001) Mutations at Y-STR loci: implications for paternity testing and forensic analysis. *Forensic Science Int* 118: 116-121

Kimpton CP, Oldroyd NJ, Watson SK és mtsai(1996) Validation of highly discriminating multiplex short tandem repeat amplification systems for individual identification. *Electrophoresis* 17:1283-1293

Klintschar M, Füredi S., Egyed B, Reichenpfader B, Kleiber M (2003) Estimating the ethnic origin (EEO) of individuals using short tandem repeat loci of forensic relevance. In: Brinkmann B, Carracedo A (eds) *Progress in Forensic Genetics* 9. International Congress Series 1239. Elsevier, pp. 53-56

Klintschar M, Ricci U, Al Hammadi N, Reichenpfader B, Ebner A, Giovannucci Uzielli ML (1998) Genetic variation at the STR loci D12S391 and CSF1PO in four populations from Austria, Italy, Egypt and Yemen. *Forensic Sci Int* 97:37-45

Kozma Z, Nagai A, Woller J, Füredi S, Sétáló J (1998) Fluoresce based co-amplification and automated detection of the STR loci HUMFIBRA and HUMD21S11 in a Hungarian Caucasian population sample. *Int J Legal Med* 111:103-104

Lazaruk K, Walsh PS, Oaks F, Gilbert D, Rosenblum BB, Menchen S, Sceibler D, Wentz HM, Holt C, Wallin J (1998) Genotyping of forensic short tandem repeat (STR) systems based on sizing precision in a capillary electrophoresis instrument. *Electrophoresis* 19:86-93

Lászik A, Sótonyi P, Rand S, Hohoff C (2001) Frequency data for the STR locus ACTBP2 (SE33) in eight populations. *Int J Legal Med* 115:94-96

Lee JC-I, Chen C-H, Tsai L-C, Linacre A, Chang J-G (1997) The screening of 13 short tandem repeat loci in the Chinese population. *Forensic Sci Int* 87:137-144

Lewis PO, Zaykin D (2001) Genetic Data Analysis: Computer program for the analysis of allelic data. Version 1.0 (d16c). Free program distributed by the authors over the internet from <http://lewis.eeb.uconn.edu/lewishome/software.html>

Lygo JE, Johnson PE, Holdaway DJ és mtsai (1994) The validation of short tandem repeat loci for use in forensic casework. *Int J Legal Med* 107:77-89

Martin P, Alonso A, Budowle B, Albarrán C, García O, Sancho M (1995) Spanish population data on 7 tetrameric short tandem repeat loci. *Int J Legal Med* 108:145-149

- Mathias N**, Bayes M, Tyler-Smith C (1994) Highly informative compound haplotypes for the human Y chromosome. *Hum Mol Genet* 3:115-123
- Mertens G**, Gielis M, Mommers N, Mularoni A, Lamartine J, Heylen H, Muylle L, Vandenberghe A (1999) Mutation of the repeat number of the HPRTB locus and structure of rare intermediate alleles. *Int J Legal Med* 112:192-194
- Meyer E**, Wiegand P, Brinkmann B (1995) Phenotype differences of STRs in 7 human populations. *Int J Legal Med* 107:314-322
- Michalakis Y**, Excoffier L (1996) A generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. *Genetics* 142:1061-1064
- Micka KA**, Sprecher CJ, Lins AM és mtsai (1996) Validation of multiplex polymorphic STR amplification sets developed for personal identification applications. *J Forensic Sci* 41:582-590
- Miścicka-Śliwka D**, Czarny J, Grzybowski T, Woźniak M (1998) Population genetics of 14 STRs: vWA, TH01, TPOX, CSF1PO, D5S818, D13S317, D7S820, D16S539, F13A01, FESFPS, F13B, LPL, D3S1358 and FGA in the Pomerania-Kujawy region of Poland. In: Olaisen B, Brinkmann B, Lincoln P (eds) *Progress in Forensic Genetics 7. International Congress Series 1167*. Elsevier, pp. 261-263
- Möller A**, Wiegand P, Grüşchow C, Seuchter SA, Baur MP, Brinkmann B (1994) Population data and forensic efficiency values for the STR systems HumVWA, HumMBP and HumFABP. *Int J Leg Med* 106 : 183-189
- Monson KL**, Budowle B (1998) Effect of reference database on frequency estimates of polymerase chain reaction (PCR)-based DNA profiles. *J Forensic Sci* 43:483-488
- Moscetti A**, Boschi I, Dobosz M, Destro-Bisol G, Pescarmona M, D'Aloja E, Pascali VL (1995) Fluorescence-based classification of microsatellites using a single-wavelength semiautomatic sequencer: genotype assignment and identity tests by analysis of comigrating peak profiles. *Electrophoresis* 16:1875-1880
- National Commission of the Future of DNA Evidence (2001)** The future of Forensic DNA testing : Prediction of the Research and Development Working Group. Washington
- Nei M (1972)** Genetic distance between populations. *Am Nat* 106:283-291
- Nei M (1973)** Analysis of gene diversity in subdivided populations. *Proc Natl Acad Sci USA* 70:3321-3323
- Nei M**, Roychoudhury AK (1974) Sampling variances of heterozygosity and genetic distance. *Genetics* 76:379-390

- NRC** National Research Council (1996) The evaluation of forensic DNA evidence. National Academy Press Washington, D.C.
- Ohno Y**, Sebetan IM, Akaishi S (1982) A simple method for calculating the probability of excluding paternity with any number of codominant alleles. *Forensic Sci Int* 19:93-98
- Olaisen B**, Bar W, Brinkmann B, Budowle B, Carracedo A, Gill P, Lincoln P, Mayr WR, Rand S (1998) DNA recommendations 1997 of the International Society for Forensic Genetics. *Vox Sang* 74:61-63
- Ota T**, **Kimura M** (1973) A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in finite population. *Genet Res* 22:201-204
- Paracchini S**, Arredi B, Chalk R, Tyler-Smith C (2002) Hierarchical high-throughput SNP genotyping of the human Y chromosome using MALDI-TOF mass spectrometry. *Nucleic Acids Res* 30:e27
- Parsons TJ**, Weeden VW (1997) Preservation and Recovery of DNA in Postmortem Specimens and Trace Samples. In W.D. Haglund and M.H. Sorg (eds.): *Forensic Taphonomy. The Postmortem Fate of Human Remains*. CRC Press, New York London Tokyo
- Pestoni C**, Lareu MV, Rodriguez MS és mtsai (1995) The use of the STRs HumTH01, Hum VWA31/A, HumF13A1, Hum FES/FPS, HumLPL in forensic application: validation studies and population data for Galicia (NW Spain). *Int J Legal Med* 107:283-290
- Puers C**, Hammond HA, Caskey CT, Lins AM, Sprecher CJ, Brinkmann B, Schumm JW (1994) Allelic ladder characterization of the short tandem repeat polymorphism located in the 5' flanking region to the human coagulation factor XIII A subunit gene. *Genomics* 23:260.
- Quintana-Murci L**, Krausz Cs, McElreavey K (2001) The human Y chromosome: function, evolution and disease. *Forensic Science Int* 118: 169-181
- Reliagene** Technologies, Inc. Formerly GenTest Laboratories (2001) Y-PLEX™6 "Instruction Manual v2.0"
- Roewer L**, Arnemann J, Spurr NK, Grzeschik KH, Epplen JT (1992) Simple Repeat Sequences on the human Y chromosome are equally polymorphic as their autosomal counterparts. *Hum Genet* 89:389-394

- Roewer L, Epplen JT (1992)** Rapid and sensitive typing of forensic stains by PCR amplification of polymorphic simple repeat sequences in case work. *Forensic Sci Int* 53:163-171
- Roewer L, Krawczak M, Willuweit S, Nagy M, Alves C, Amorim A, Anslinger K, Augustin C, Betz A, Bosch E, Caglia A, Carracedo A, Corach D, Dekairelle A-F, Dobosz T, Dupuy BM, Füredi S, Gehrig C, Gusmao L, Henke J, Hidding M, Hohoff C, Hoste B, Jobling MA, Kargel HJ, de Knijff P, Lessig R, Liebeherr E, Lorente M, Martinez-Jarreta B, Nieves P, Nowak M, Parson W, Pascali VL, Penacino G, Ploski R, Rolf B, Sala A, Schmidt U, Schmitt C, Schneider PM, Szibor R, Teifel-Greding J, Kayser M (2001)** Online reference database of European Y-chromosomal short tandem repeat (STR) haplotypes. *Forensic Sci Int* 118:106-113
- Ruitberg CM, Reeder DJ, Butler JM (2001)** STRBase: a short tandem repeat DNA database for the human identity testing community. *Nucleic Acids Res* 29:320-322
- Santos FR, Gerelsaikhan T, Munkhtuja B, Oyunsuren T, Epplen J, Pena SDJ (1996)** Geographic differences in the allele frequencies of the human Y-linked tetranucleotide polymorphism DYS19. *Hum Genet* 97:309-313
- Santos FR, Tyler-Smith C (1996)** Reading the human Y chromosome: the emerging DNA markers and human genetic history. *Brasilian Journal of Genetics* 19, 4: 665-670
- Schneider PM, Meuser S, Waiyawuth W, Seo Y, Rittner C (1998)** Tandem repeat structure of the duplicated Y-chromosomal STR locus DYS385 and frequency studies in the German and three Asian populations. *Forensic Sci Int* 97:61-70
- Schneider S, Roesli D, Excoffier L (2000)** Arlequin ver. 2.000: A software for population genetics data analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland. URL: <http://anthro.unige.ch/arlequin>
- Schumm JW, Lins AM, Micka KA, Sprecher CJ, Rabbach DR, Bacher JW (1997)** Automated fluorescent detection of STR multiplexes - development of the GenePrint, PowerPlex and FFFL multiplexes for forensic and paternity applications. In: *Proceedings from the First European Symposium on Human Identification, 1996*. Promega Corporation, Madison, WI, pp. 90-104
- Shriver MD, Smith MW, Jin L, Marcini A, Akey JM, DeKa R, Ferrel RE (1997)** Ethnic-affiliation estimation by use of population-specific DNA markers. *Am J Hum Genet* 60:957-964
- Slatkin M, Excoffier L (1996)** Testing for linkage disequilibrium in genotypic data using the Expectation-Maximization algorithm. *Heredity* 76:377-383

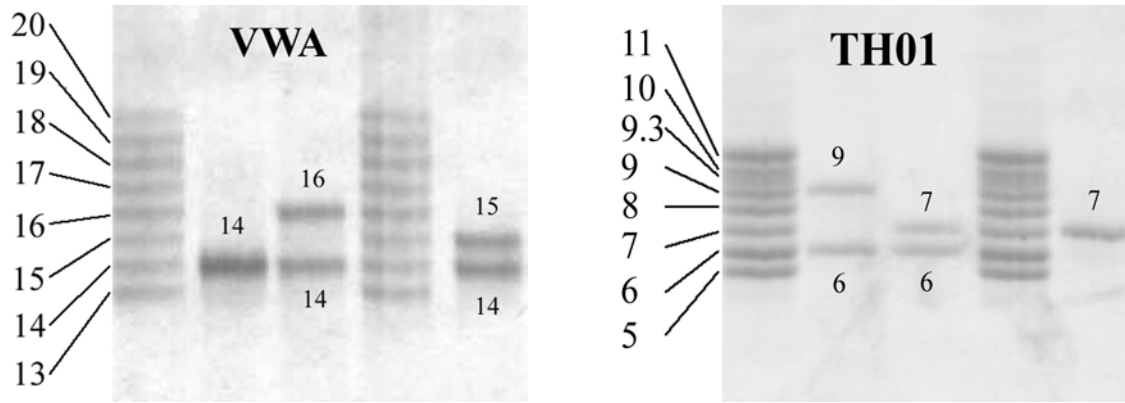
- Spurdle AB, Jenkins T (1992)** The Y-chromosome as a tool for studying human evolution. *Curr Opin Genet Dev* 2:487-491
- Strachan T, Read AP (1999)** *Human Molecular Genetics*, 2nd Edition, BIOS Scientific Publishers Ltd.
- Thomas MG, Eeale ME, Jones AL, Richards M, Smith A, Redhead N, Torroni A, Scoyyari R, Gratrix F, Tarekegn A, Wilson JF, Capelli C, Bradman N, Goldstein DB (2002)** Founding Mothers of Jewish Communities: Geographically Separated Jewish Groups Were Independently Founded by Very Few Female Ancestors. *Am J Hum Genet* 70:1411-1420
- Underhill PA, Shen P, Lin AA, Jin L, Passarino G, Yang WH, Kauffman E, Bonne-Tamir B, Bertranpetit J, Francalacci P és *mtsai* (2000)** Y chromosome sequence variation and the history of human populations. *Nature Genet* 26:358–361
- Urquhart A, Kimpton CP, Downes TJ, Gill P (1994)** Variation in short tandem repeat sequences – a survey of twelve microsatellite loci for the use as forensic identification markers. *Int J Legal Med* 107:13-20
- Valdes AM, Slatkin M, Freimer NB (1993)** Allele frequencies at microsatellite loci: the stepwise mutational model revisited. *Genetics* 133:737-749
- Vandenberghe A, Füredi S, Boutrand L, Mertens G (2000)** Structure of rare alleles of locus HPRTB in the Hungarian population. In: Sensabaugh GF, Lincoln PJ, Olajsen B (eds) *Progress in Forensic Genetics* 8. Elsevier 2000, pp. 43-45
- Walsh PS, Varlaro J, Reynolds R (1992)** A rapid chemiluminescent method for quantitation of human DNA. *Nucleic Acids Research* 20:5061-5065
- Weber JL, May PE (1989)** Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am J Hum Genet* 44:388-396
- Weir BS (1996)** *Genetic Data Analysis II*. Sinauer Associates, Inc. Publishers, Sunderland, Massachusetts
- Weir BS, Cockerham CC (1984)** Estimating *F*-statistics for the analysis of population structure. *Evolution* 38:1358-1370
- Wiegand P, Budowle B, Brinkmann B (1993)** Forensic validation of the STR systems SE 33 and TC11. *Int J Leg Med* 105 : 315-320
- Woller J (1995)** DNS a kriminalisztikában. ORFK Oktatási és Kiképző Központ kiadványa, Budapest
- Woller J, Furedi S, Padar Zs (1995)** AMPFLP analysis of the VNTR loci D1S80 and ApoB in Hungary. *Int J Legal Med* 107:273-274

- Woller J, Füredi S, Pádár Zs. (1997)** Polimeráz láncreakción alapuló DNS vizsgálatok a magyar igazságügyi gyakorlatban. Orvosi Hetilap 138:3223-3228
- Wright S (1951)** The genetical structure of populations. Ann Eugen 15:323-354
- Zaharova B, Andonova S, Gilissen A, Cassiman J-J, Decorte R, Kremensky I (2001)** Y-chromosomal STR haplotypes in three major population groups in Bulgaria. Forensic Sci Int 124:182-186
- Zupanic I, Balazic J, Komel R (1998)** Analysis of nine short tandem repeat (STR) loci in the Slovenian population. Int J Legal Med 111:248–250

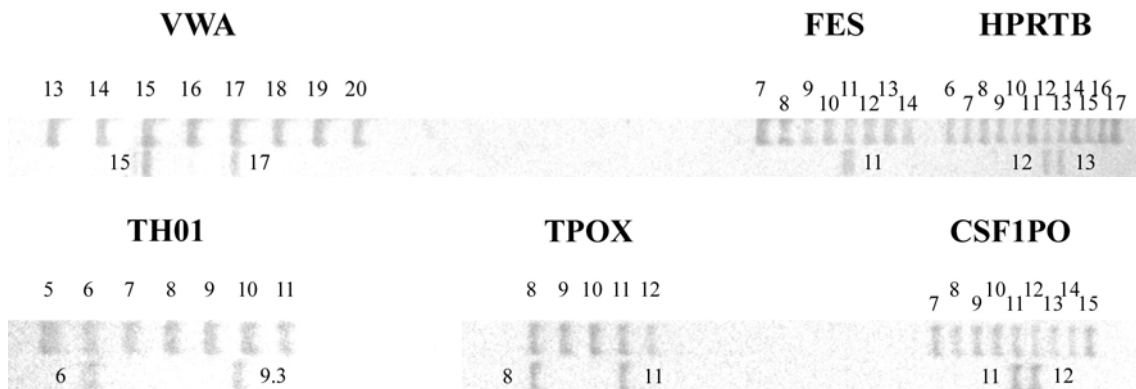
7. FÜGGELÉK

7.1. Ábrák

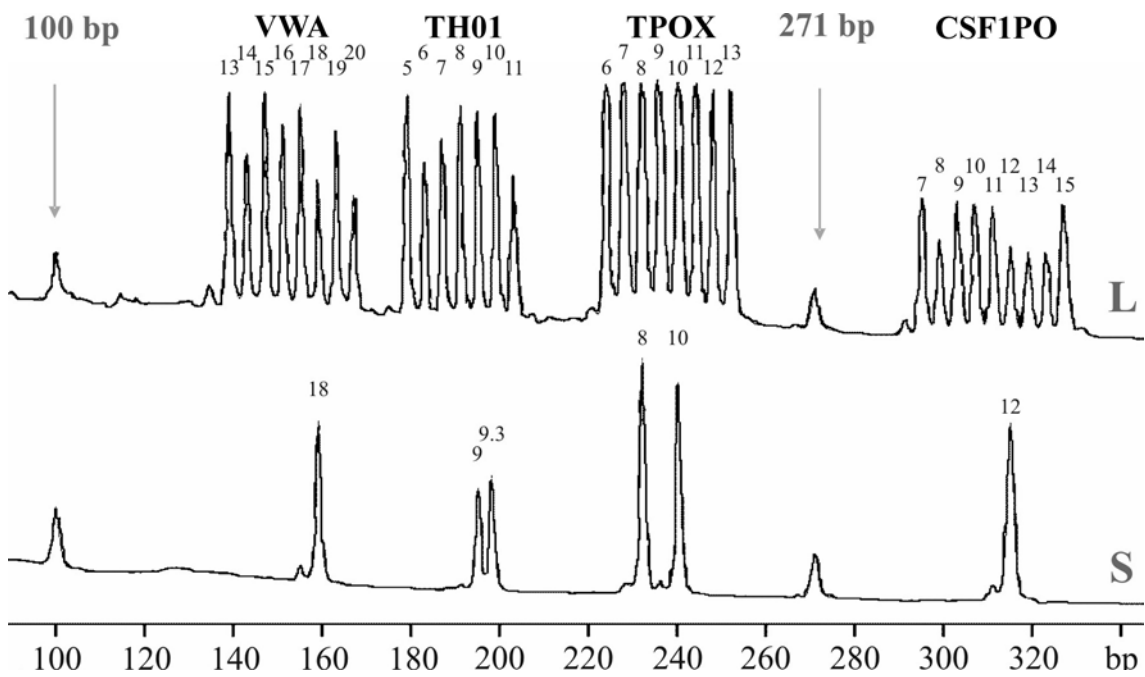
F.1. ábra VWA és TH01 (szimplex) (Füredi és mtsai, 1995)



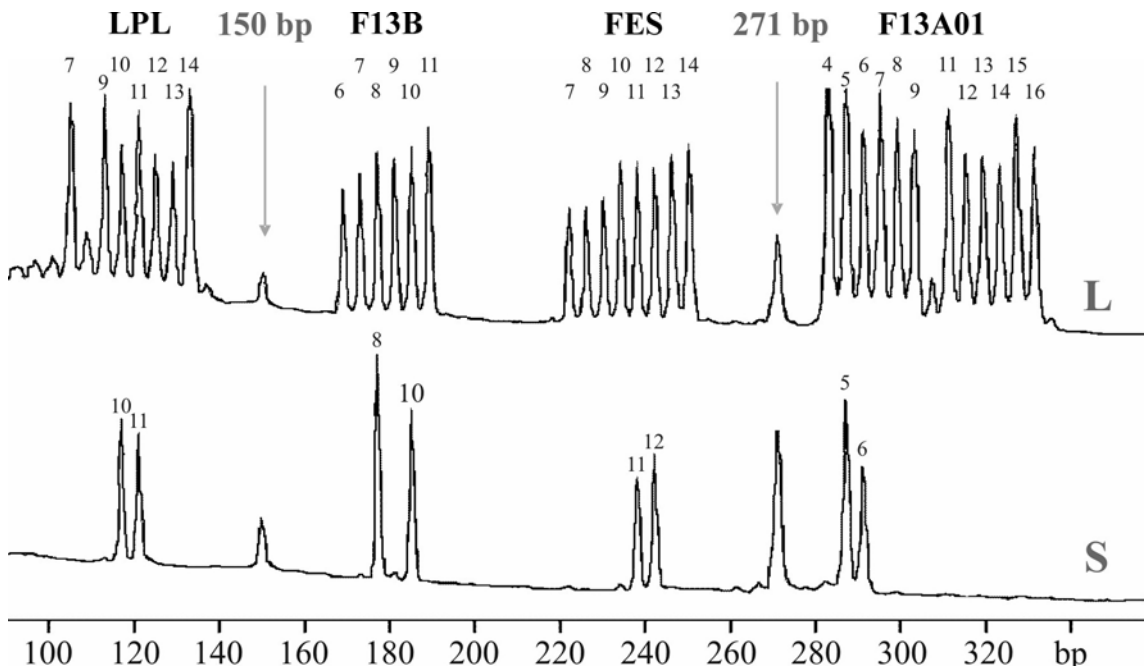
F.2. ábra VWA-FES-HPRTB és a TH01-TPOX-CSF1PO (triplex) (Füredi és mtsai, 1996)



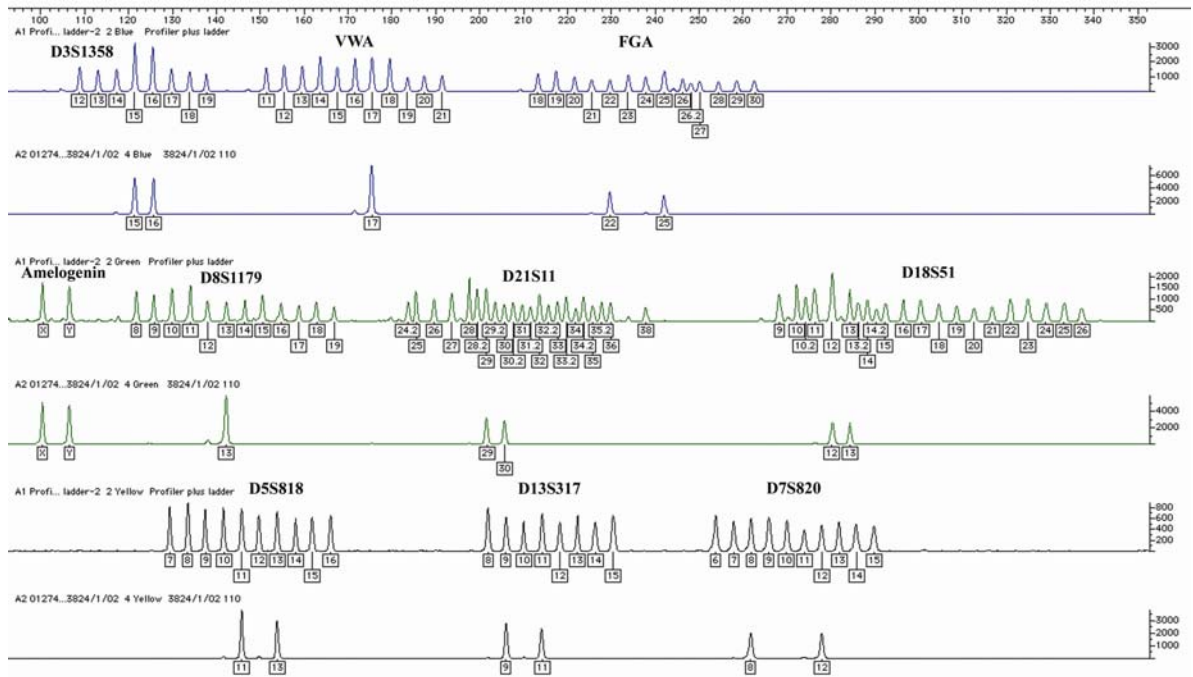
F.3. ábra VWA-TH01-TPOX-CSF1PO (kvadruplex) (Füredi és mtsai, 1997b)



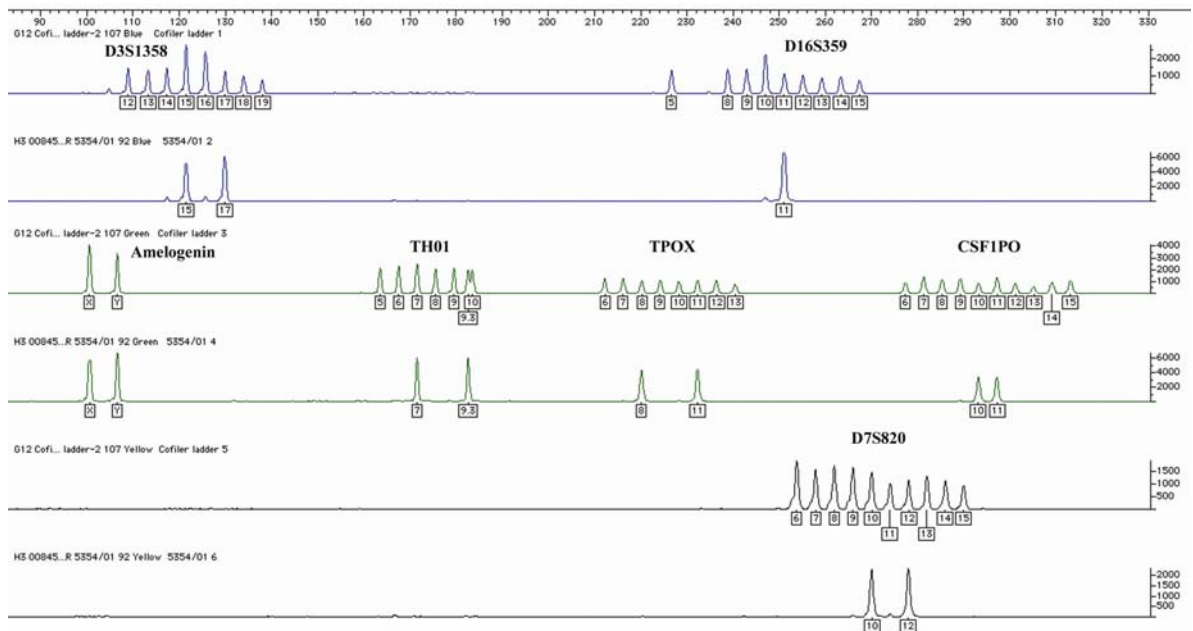
F.4. ábra LPL-F13B-FES-F13A01 (kvadruplex) (Füredi és mtsai, 1997a)



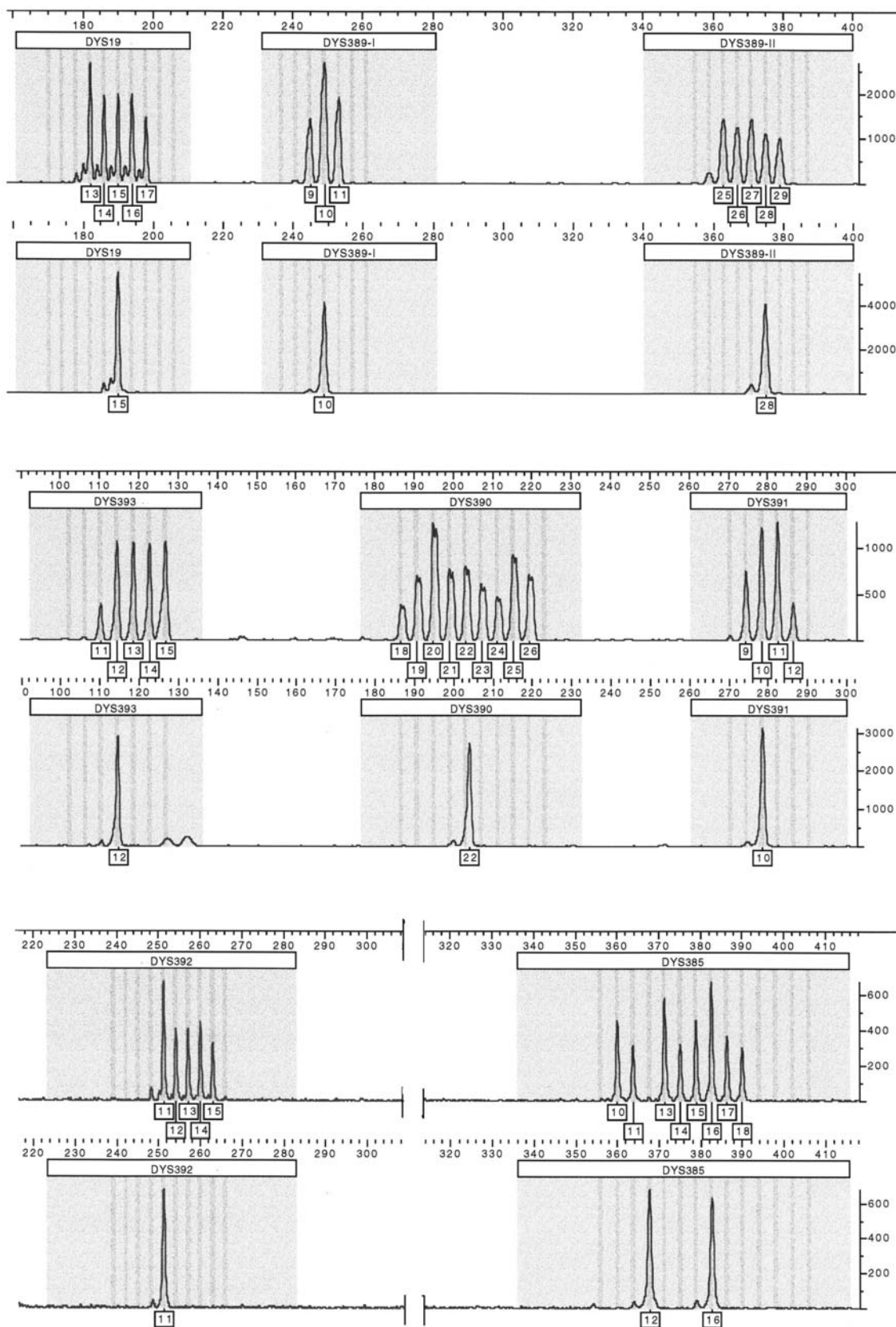
F.5. ábra D3S1358-VWA-FGA-Amelogenin-D8S1179-D21S11-D18S51-D5S818-D13S317-D7S820 (Profiler Plus multiplex) (Egyed és mtsai, 2000)



F.6. ábra D3S1358-D16S359-Amelogenin-TH01-TPOX-CSF1PO-D7S820 (Cofiler multiplex)

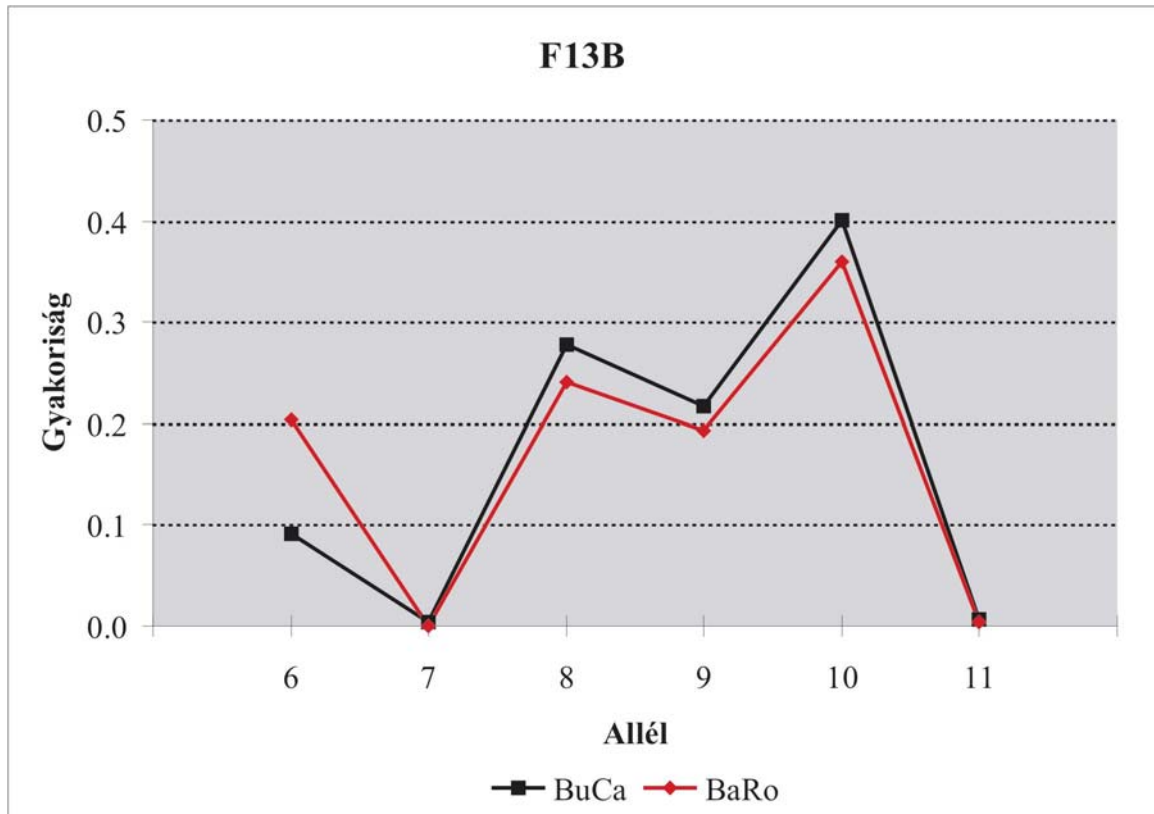


F.7. ábra DYS19-DYS389-I/II (Triplex I), DYS393-DYS390-DYS391 (Triplex II) és DYS392-DYS385 (Duplex) (Füredi és mtsai, 1999)

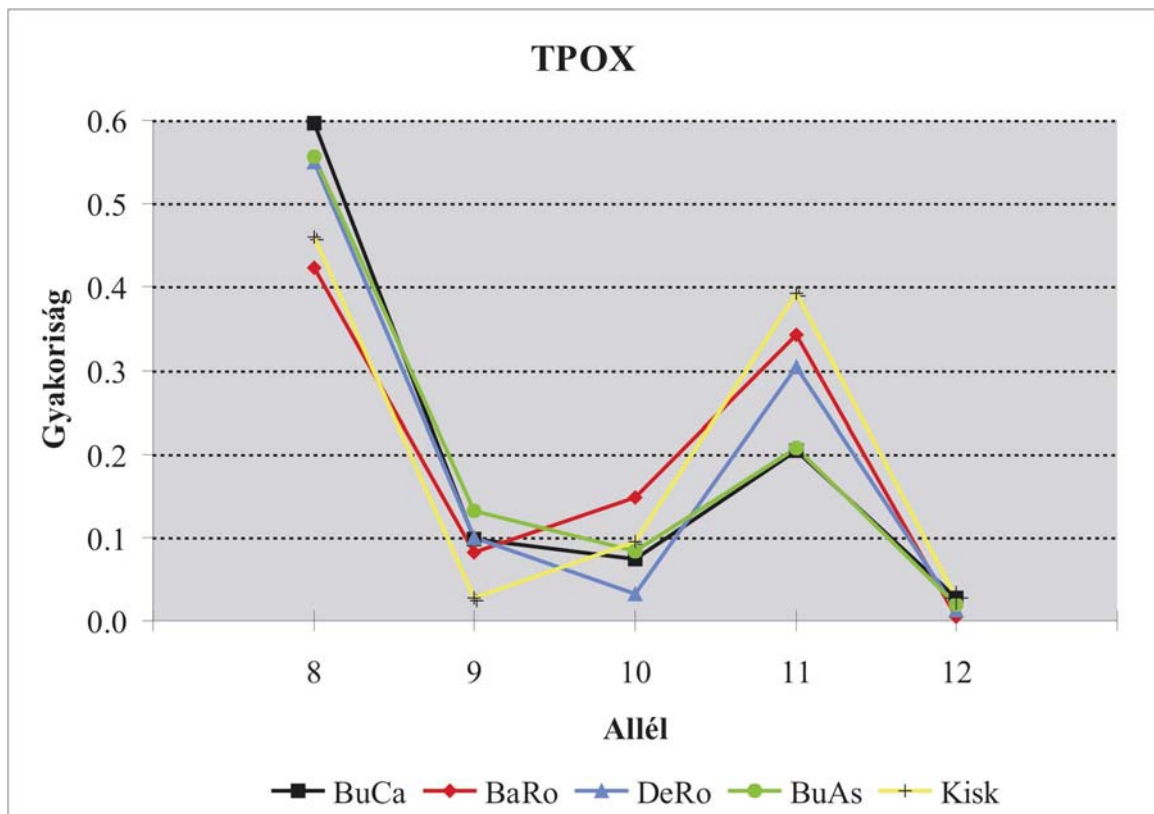


A DYS389-II lokusz esetében az eredeti alléljelölés van feltüntetve, a jelenleg érvényes nevezéktan szerint az allélrepeatek számához hármat hozzá kell adni (Roewer és mtsai, 2001)

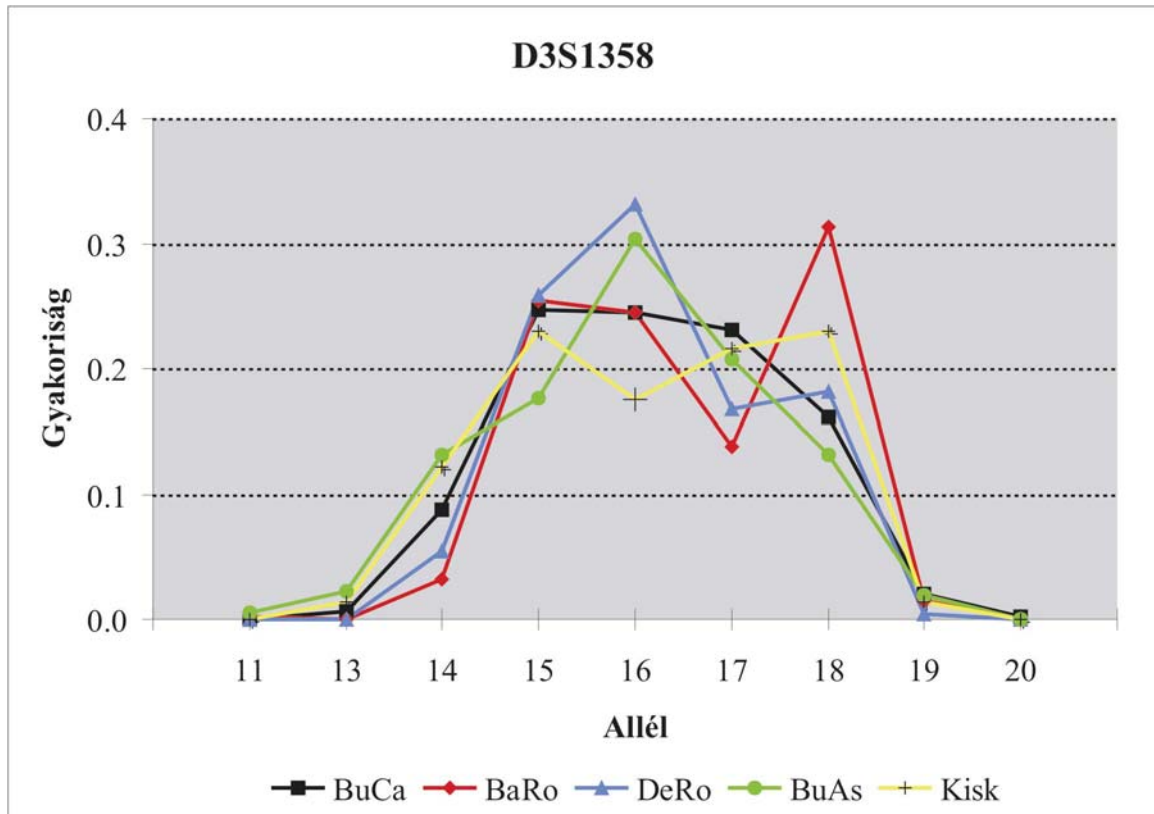
F.8. ábra Az F13B lokusz allélgyakoriság-eloszlása



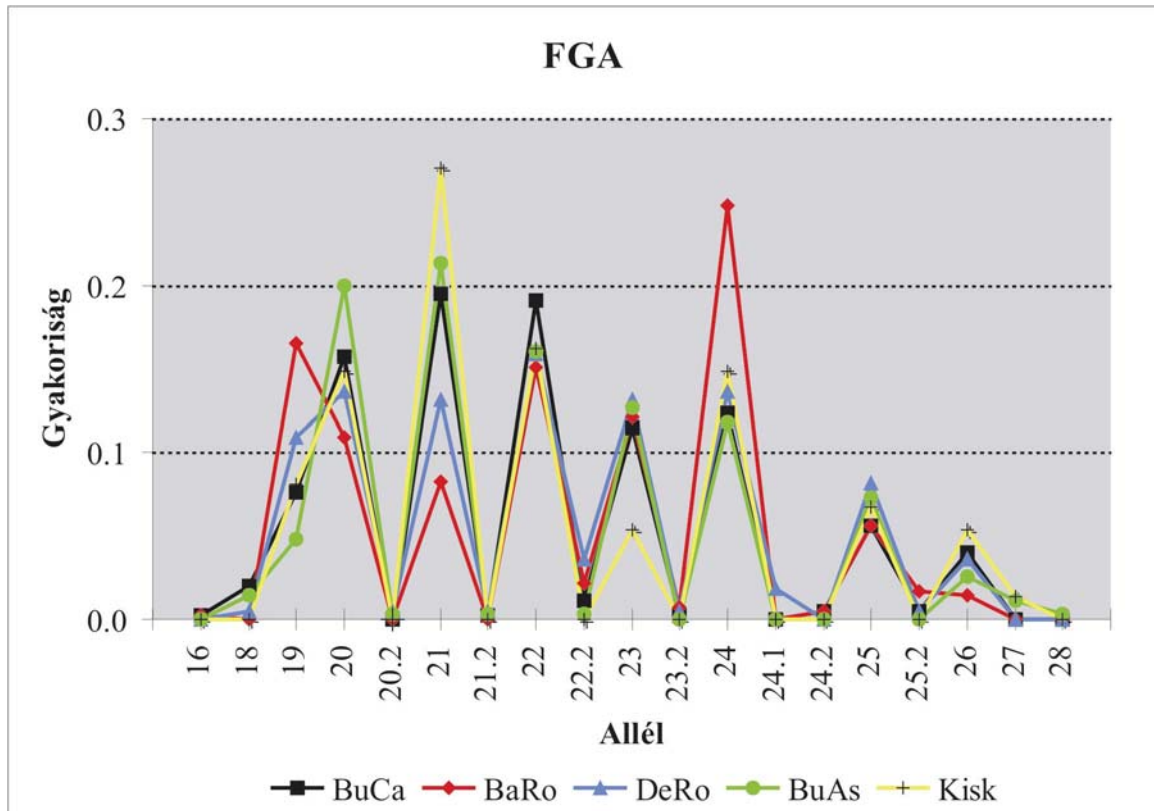
F.9. ábra A TPOX lokusz allélgyakoriság-eloszlása



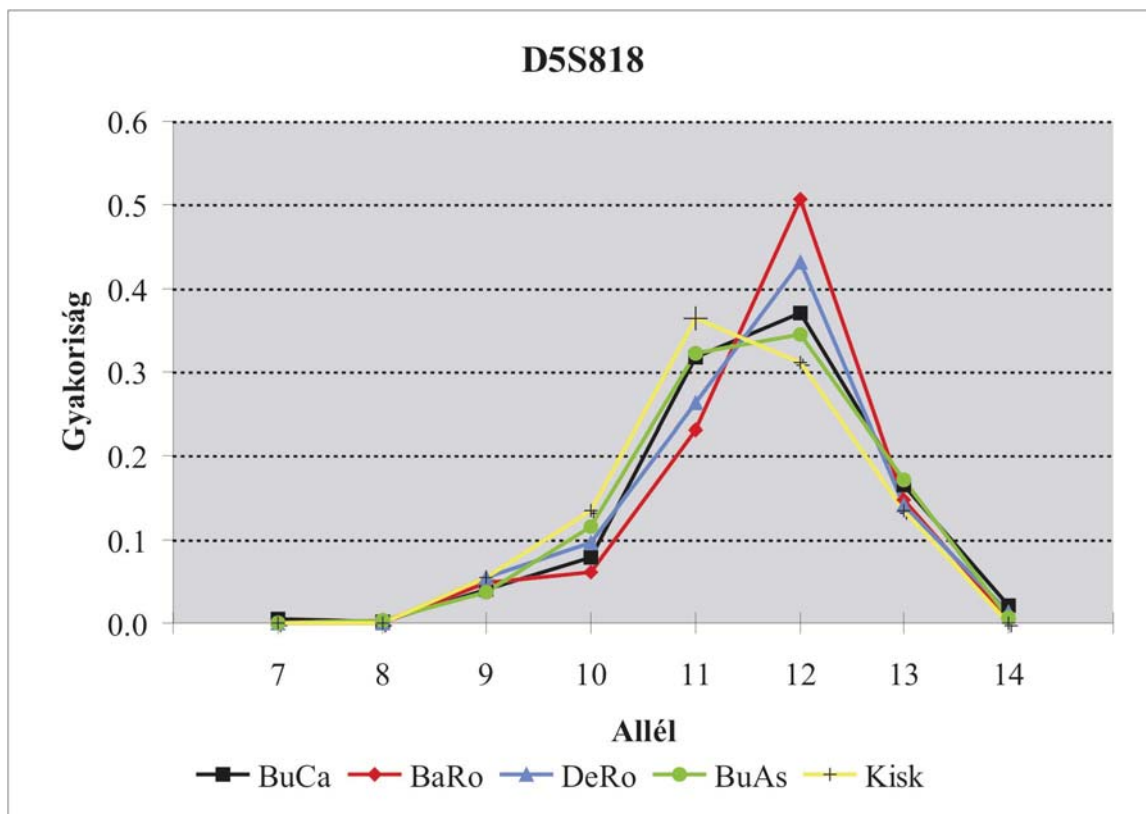
F.10. ábra A D3S1358 lokusz allélgyakoriság-eloszlása



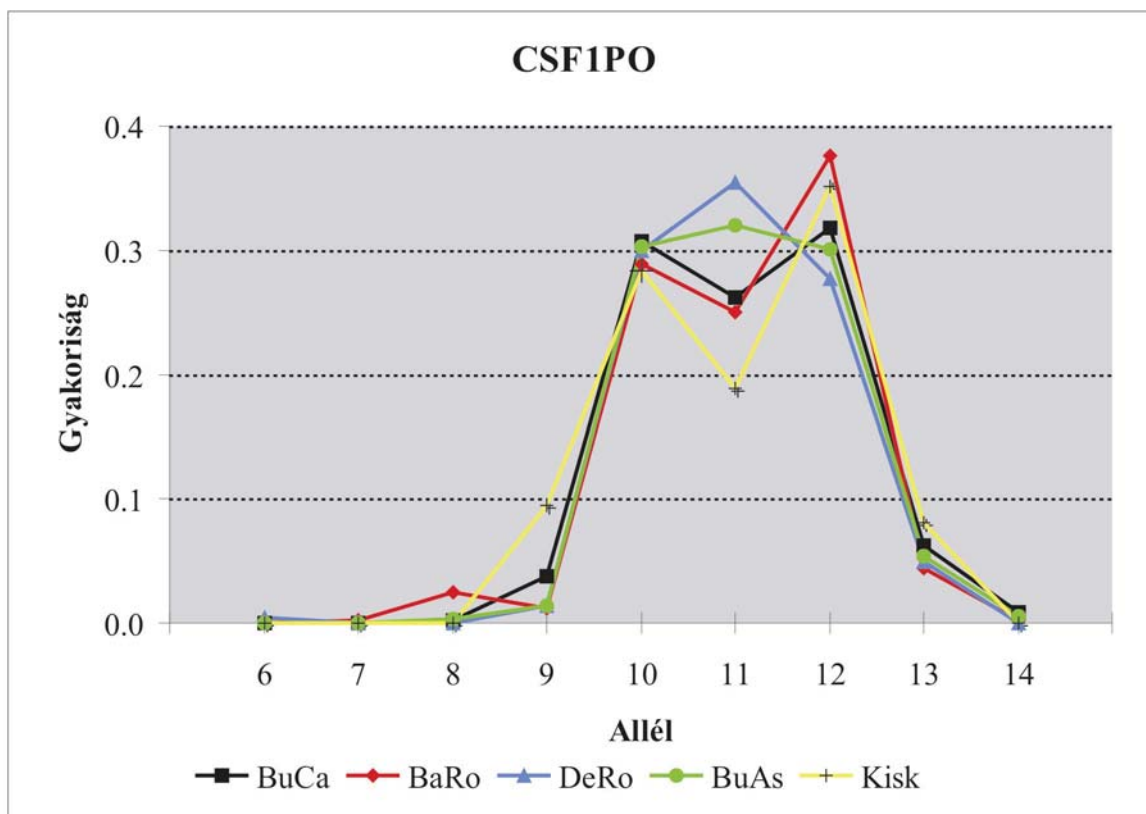
F.11. ábra Az FGA lokusz allélgyakoriság-eloszlása



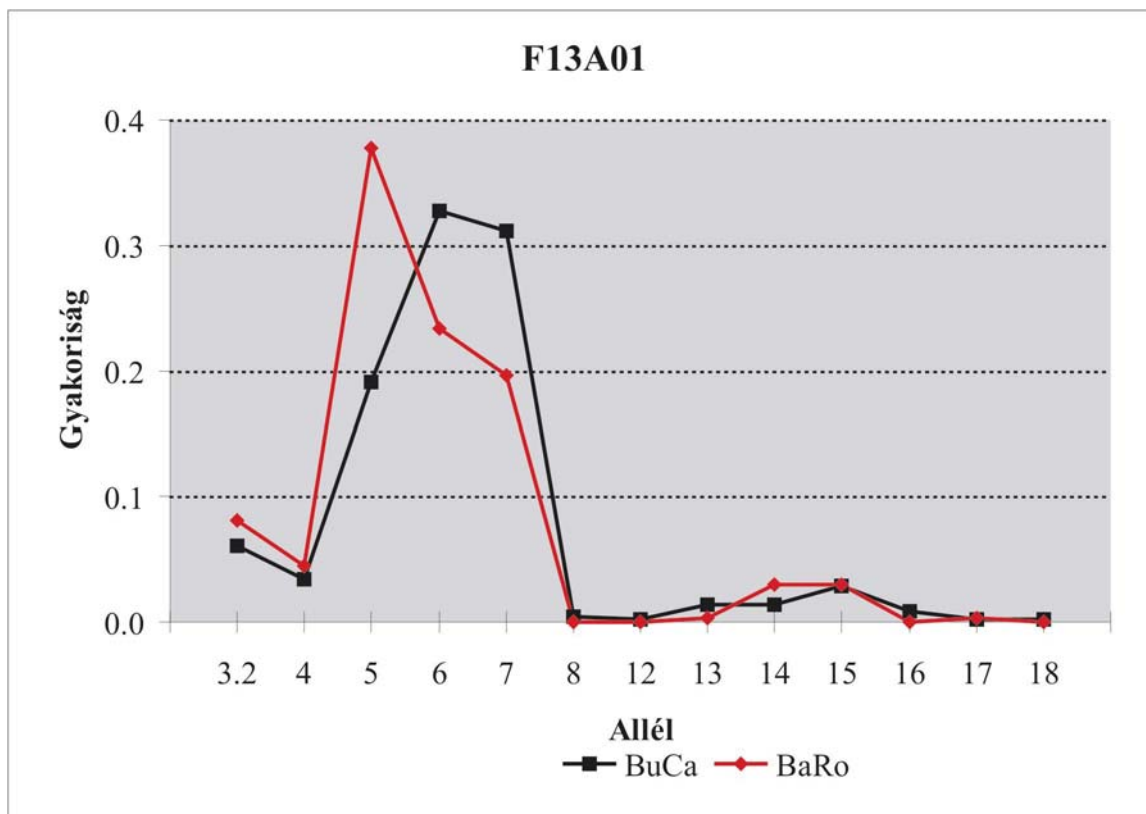
F.12. ábra A D5S818 lokusz allélgyakoriság-eloszlása



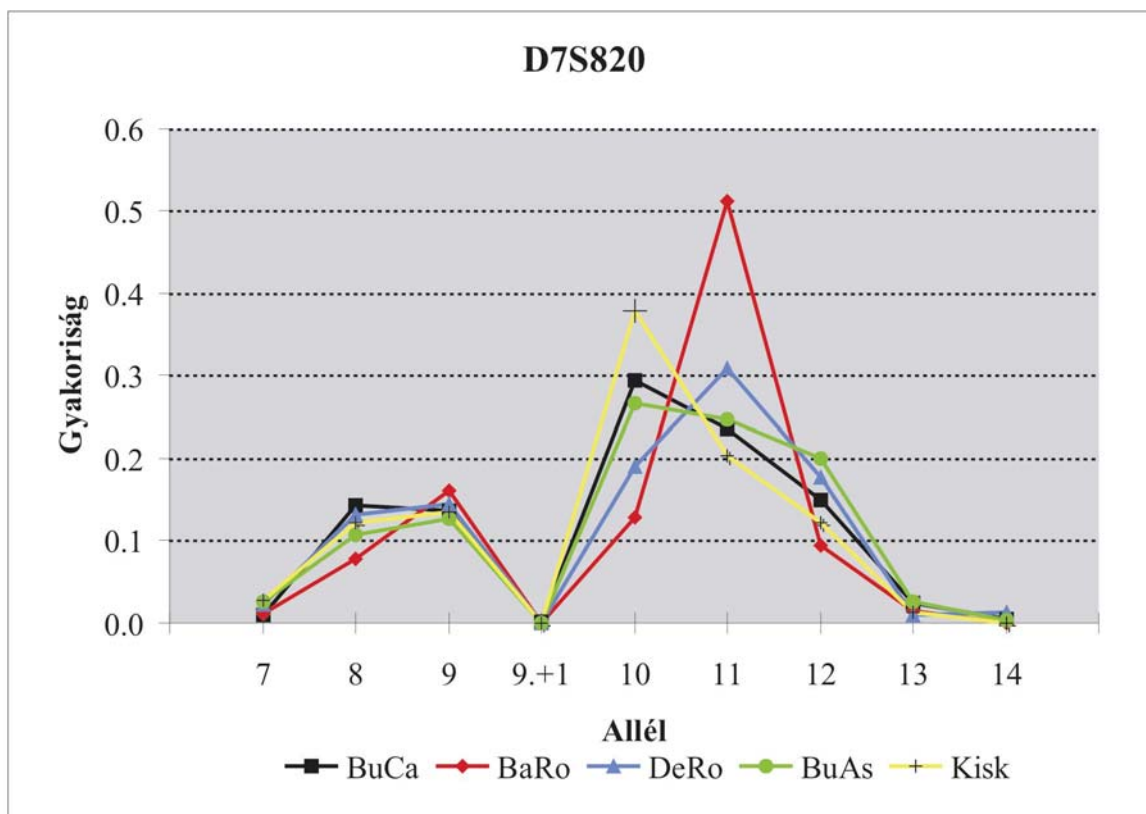
F.13. ábra A CSF1PO lokusz allélgyakoriság-eloszlása



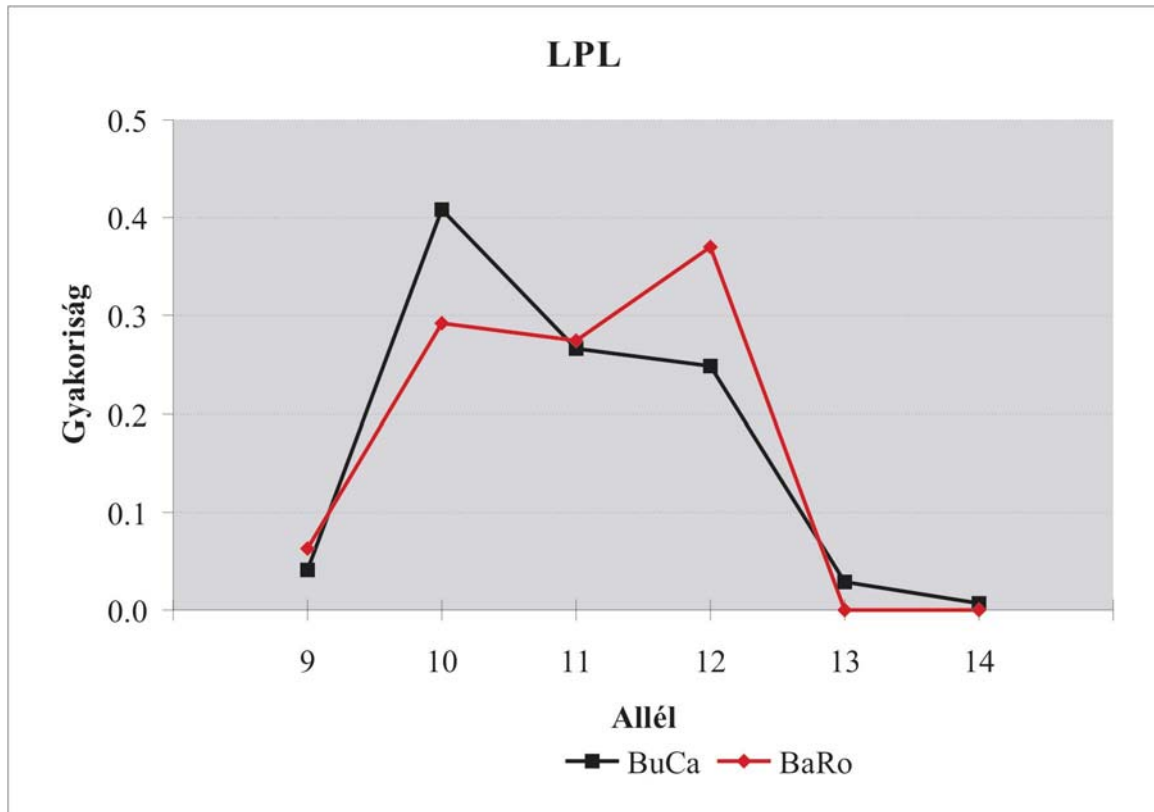
F.14. ábra Az F13A01 lokusz allélgyakoriság-eloszlása



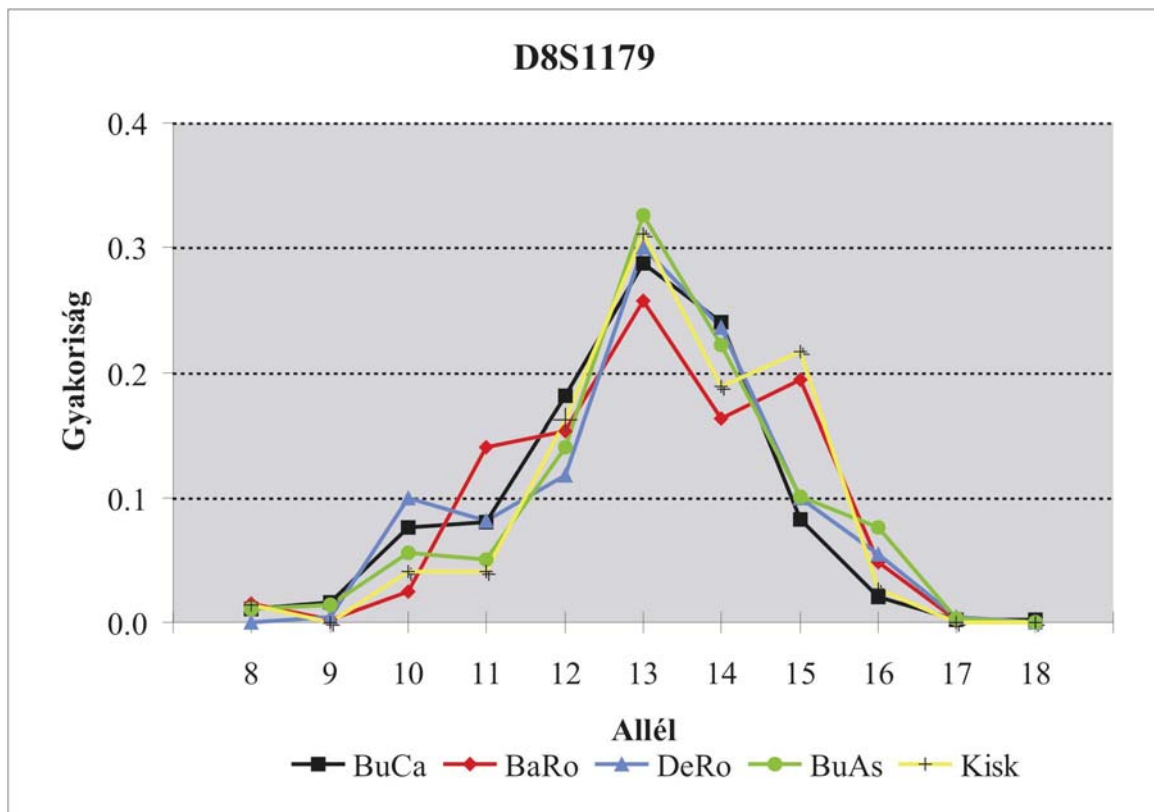
F.15. ábra A D7S820 lokusz allélgyakoriság-eloszlása



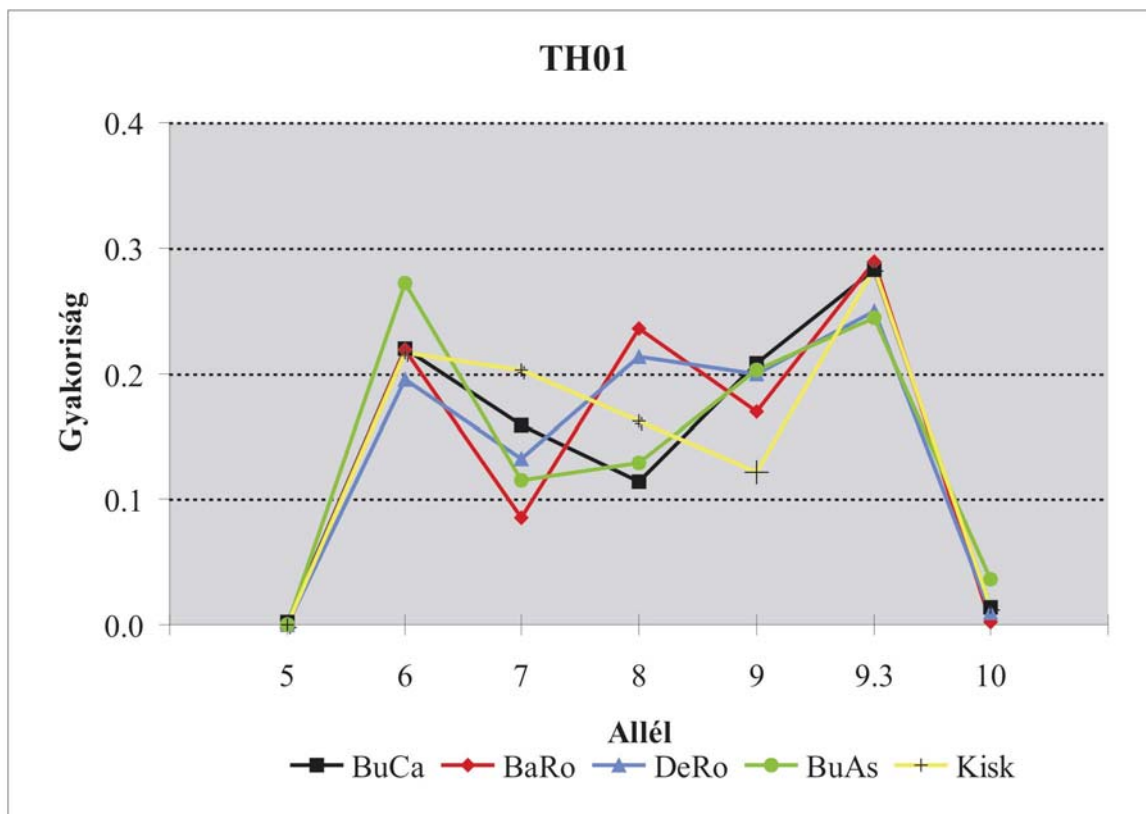
F.16. ábra Az LPL lokusz allélgyakoriság-eloszlása



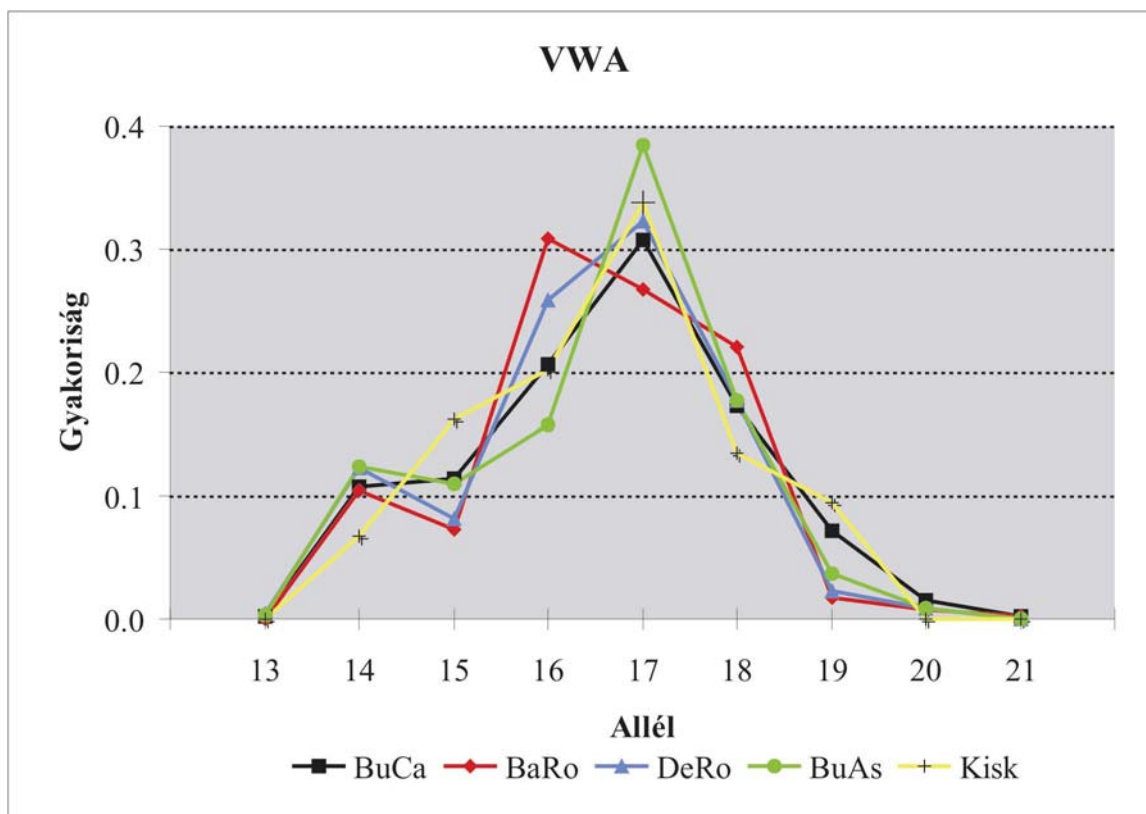
F.17. ábra A D8S1179 lokusz allélgyakoriság-eloszlása



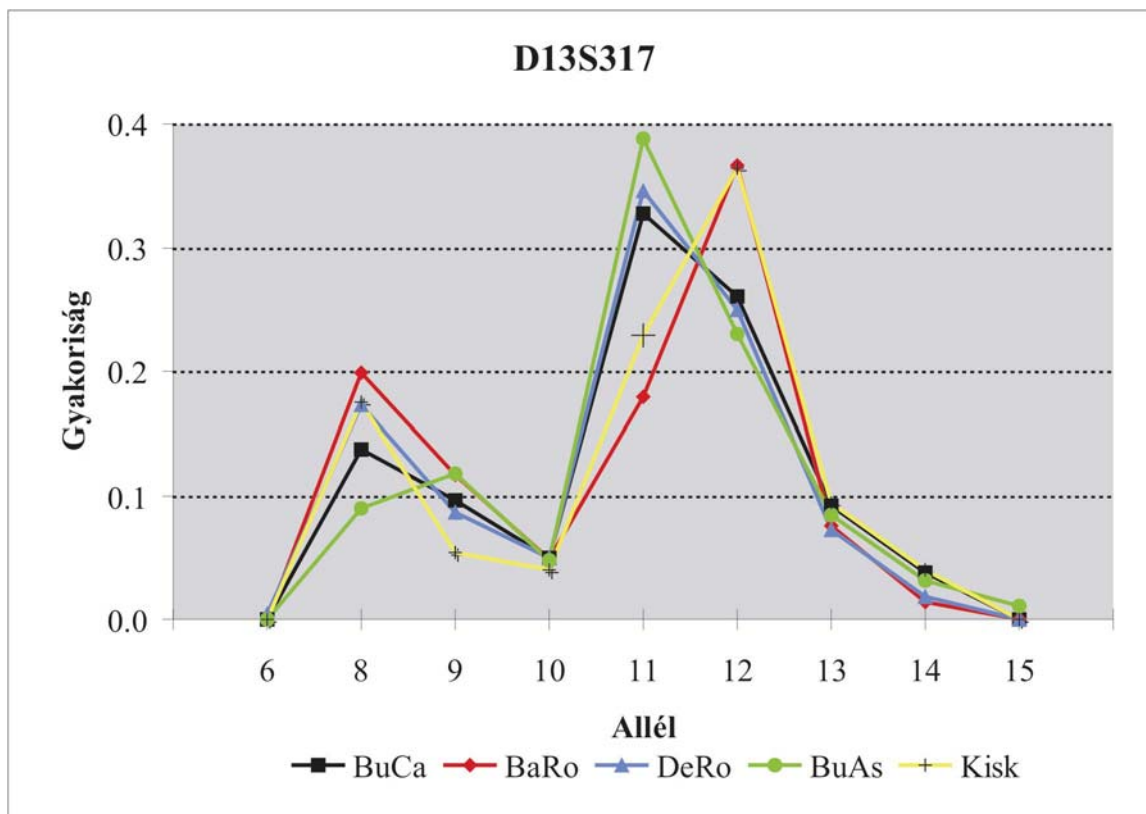
F.18. ábra A TH01 lokusz allélgyakoriság-eloszlása



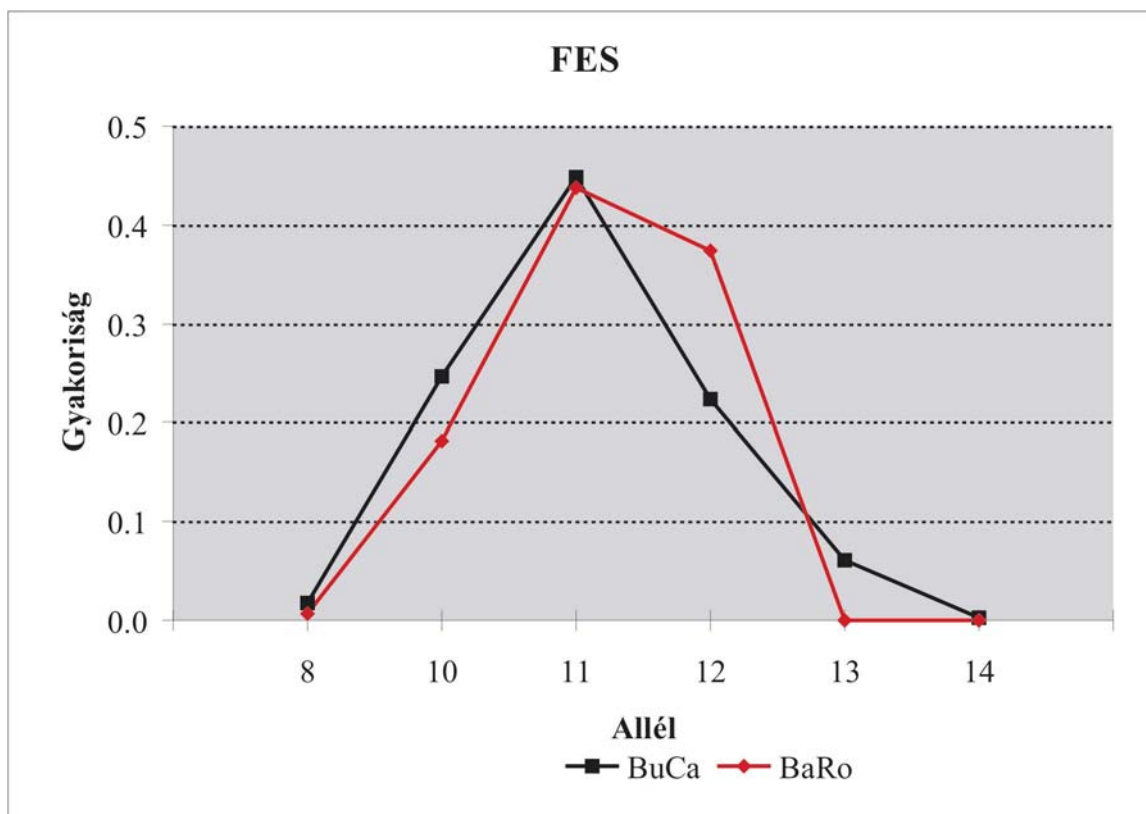
F.19. ábra A VWA lokusz allélgyakoriság-eloszlása



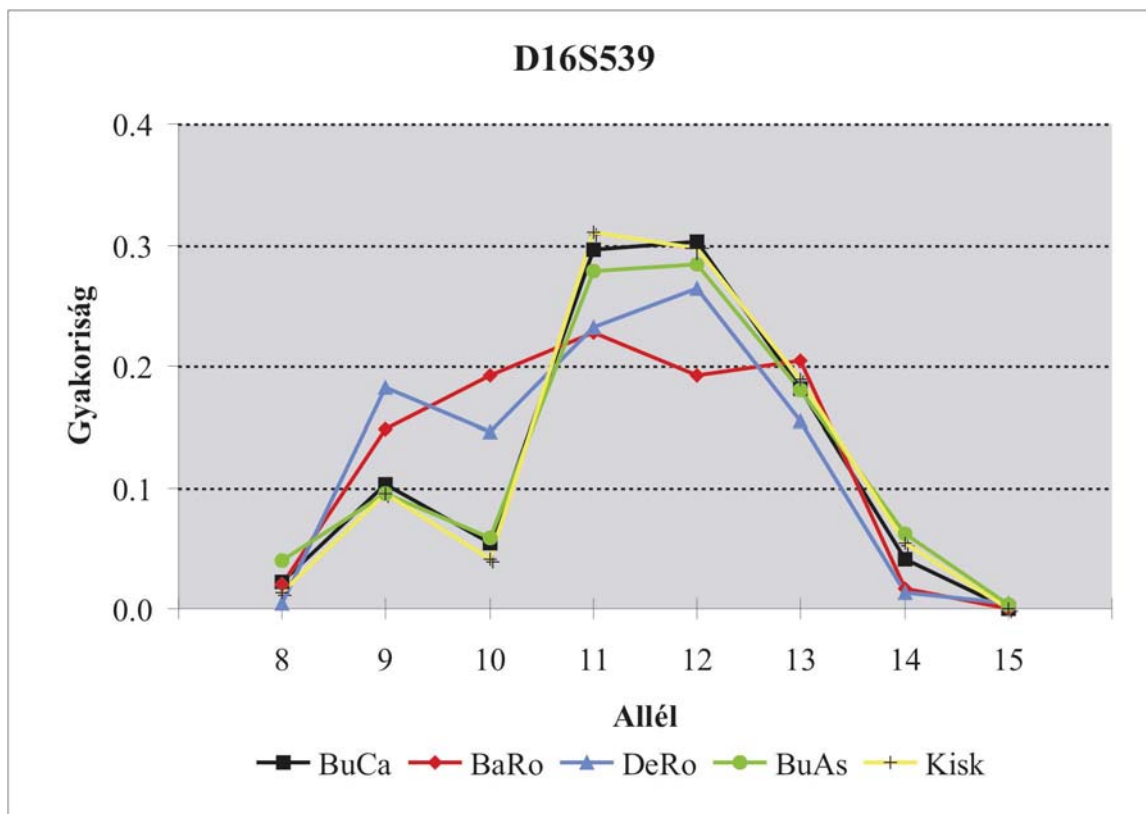
F.20. ábra A D13S317 lokusz allélgyakoriság-eloszlása



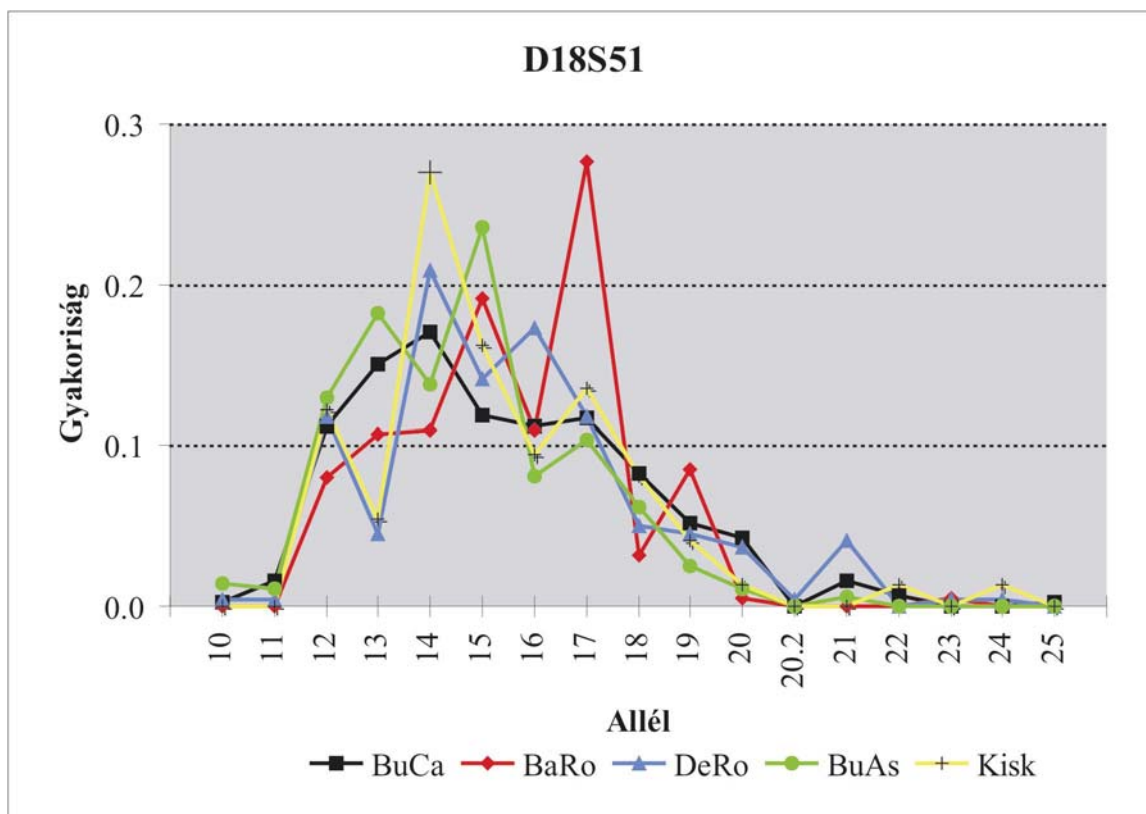
F.21. ábra A FES lokusz allélgyakoriság-eloszlása



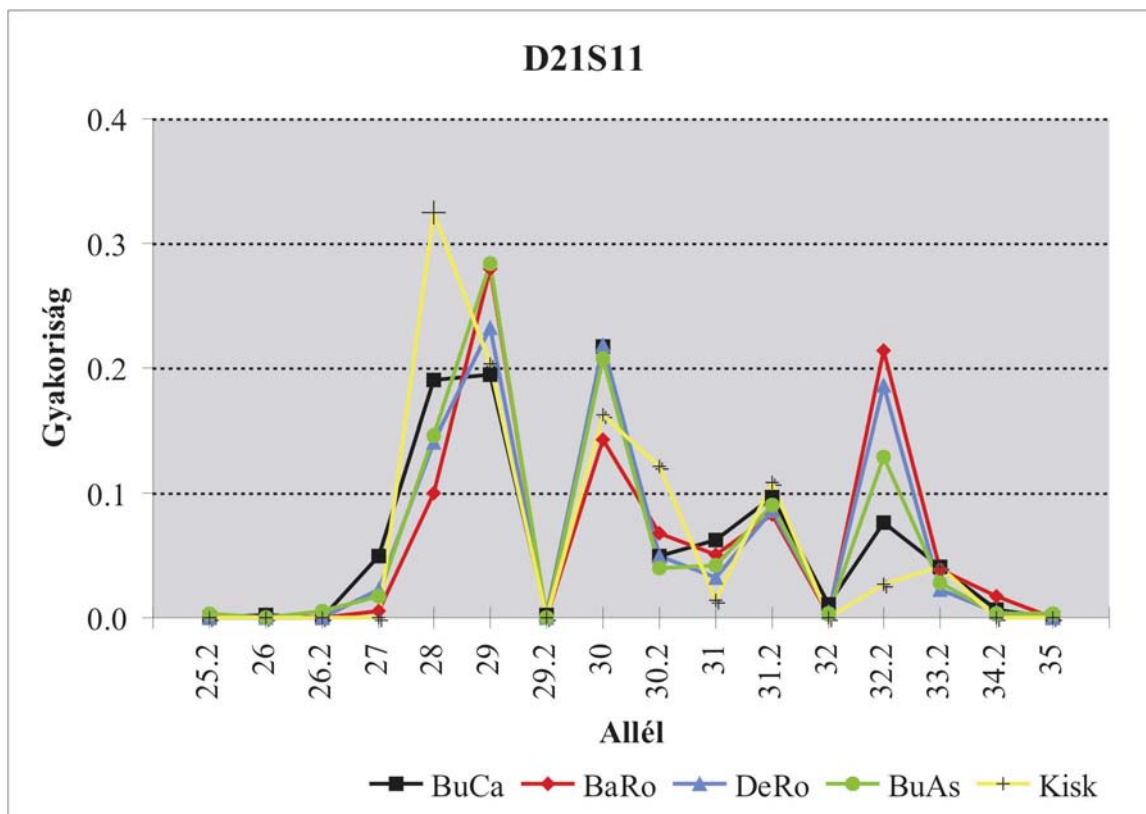
F.22. ábra A D16S539 lokusz allélgyakoriság-eloszlása



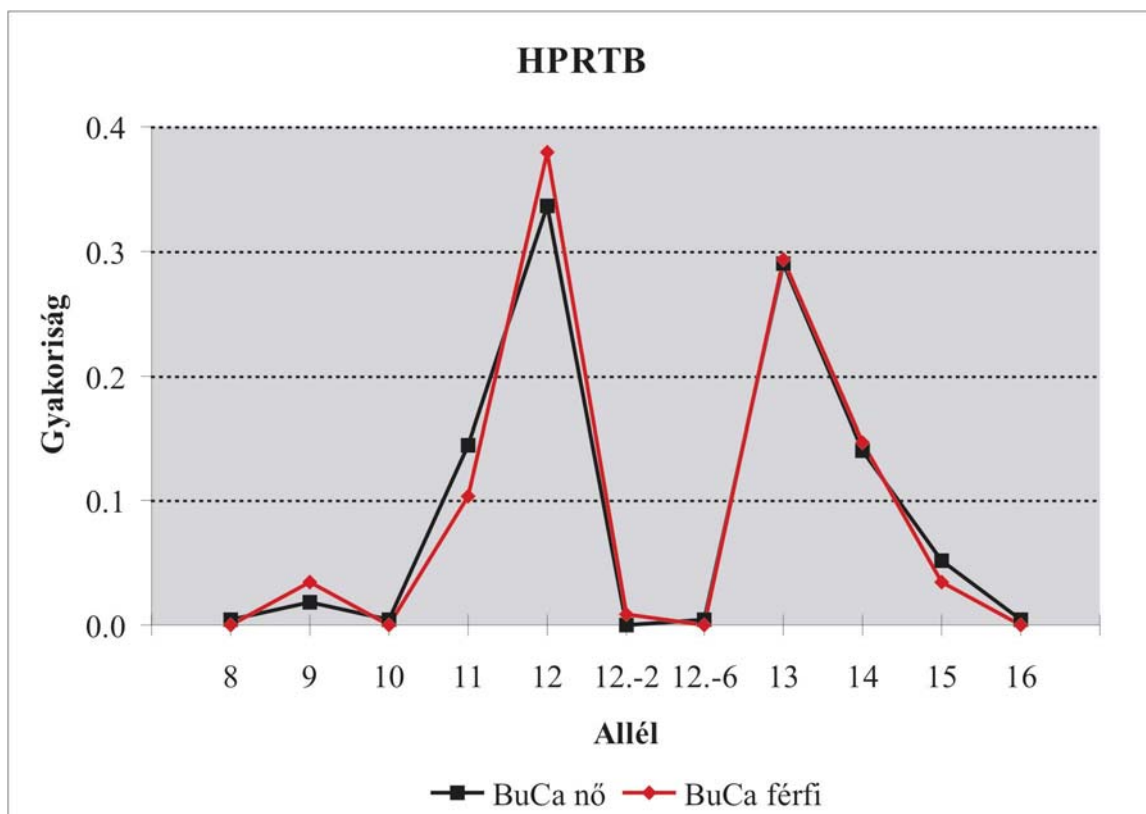
F.23. ábra A D18S51 lokusz allélgyakoriság-eloszlása



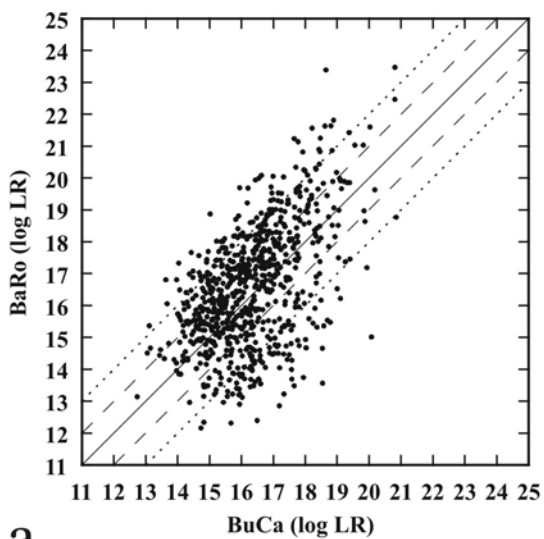
F.24. ábra A D21S11 lokusz allélgyakoriság-eloszlása



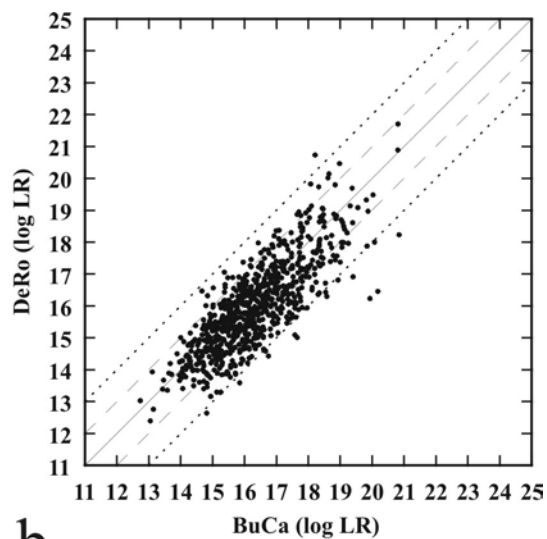
F.25. ábra A HPRTB lokusz allélgyakoriság-eloszlása



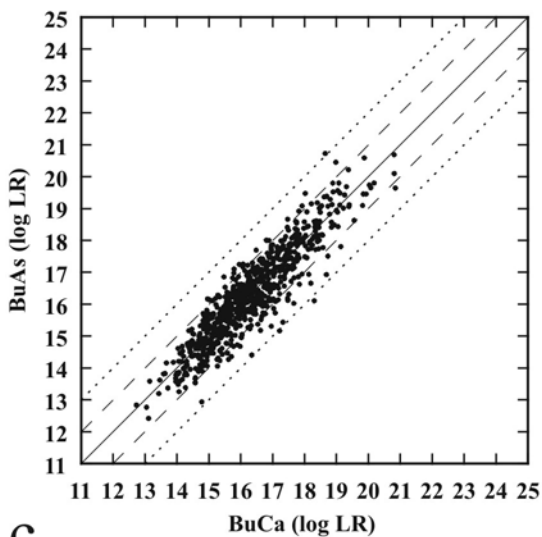
F.26a-f. ábra A vizsgált magyarországi populáció-párok 13 STR lokuszos sörét diagramjai I.



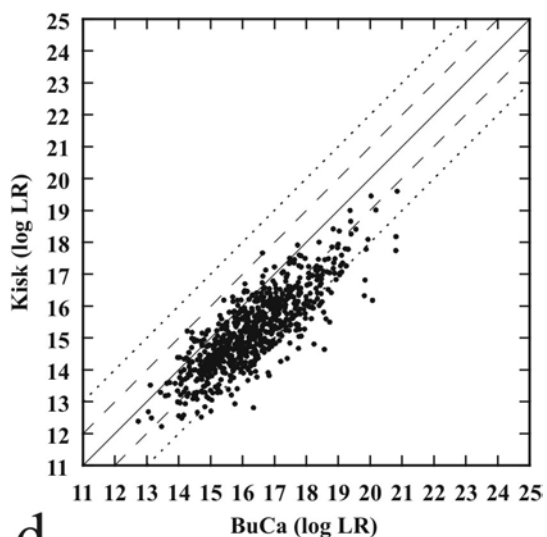
a,



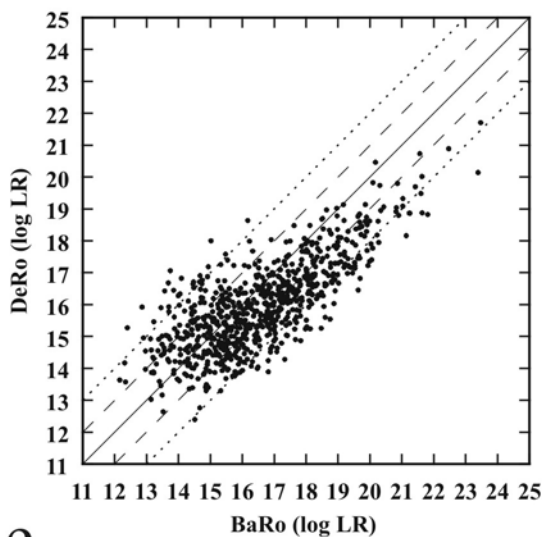
b,



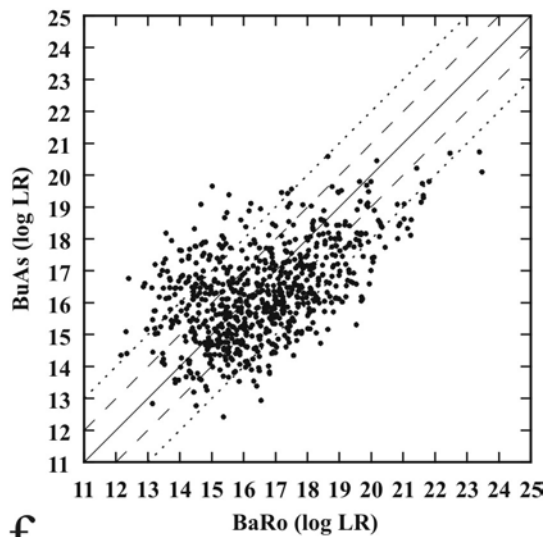
c,



d,

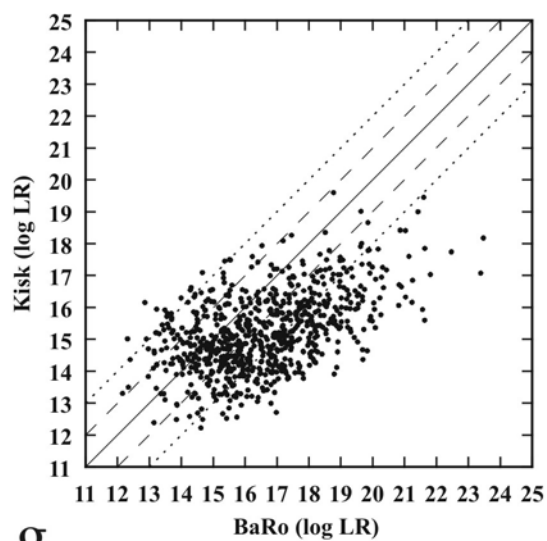


e,

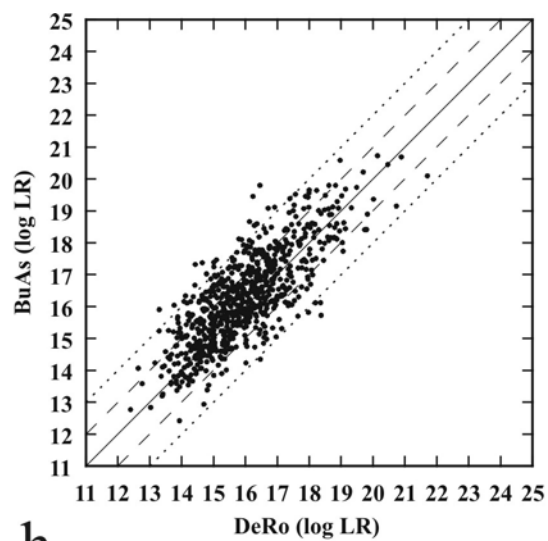


f,

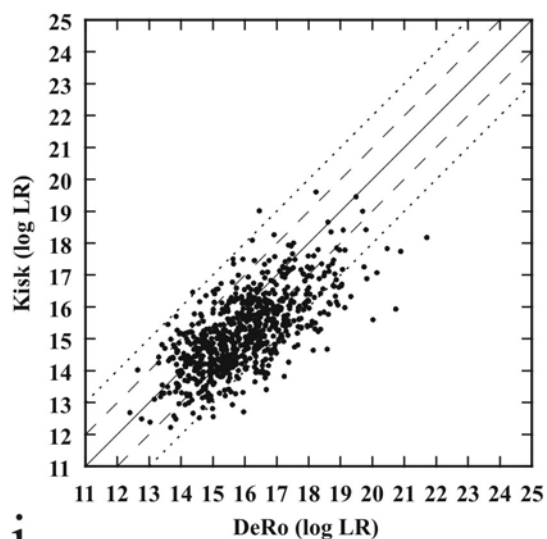
F.26g-j. ábra A vizsgált magyarországi populáció-párok 13 STR lokuszos sörét diagramjai II.



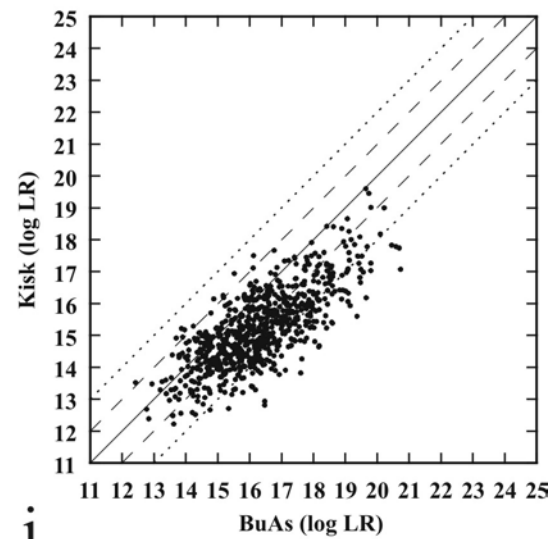
g,



h,



i,



j,

7.2. Táblázatok

	F13B		TPOX					D3S1358					FGA					D5S818				
	BuCa	BaRo	BuCa	BaRo	DeRo	BuAs	Kisk	BuCa	BaRo	DeRo	BuAs	Kisk	BuCa	BaRo	DeRo	BuAs	Kisk	BuCa	BaRo	DeRo	BuAs	Kisk
6	0,0919	0,2037																				
7	0,0045																	0,0045				
8	0,2780	0,2407	0,5964	0,4223	0,5500	0,5562	0,4595											0,0022			0,0028	
9	0,2175	0,1926	0,0987	0,0825	0,1000	0,1320	0,0270											0,0404	0,0485	0,0545	0,0365	0,0541
10	0,4013	0,3593	0,0740	0,1481	0,0318	0,0843	0,0946											0,0785	0,0607	0,0955	0,1152	0,1351
11	0,0067	0,0037	0,2040	0,3422	0,3045	0,2079	0,3919		0,0024		0,0056							0,3184	0,2306	0,2636	0,3230	0,3649
12			0,0269	0,0049	0,0136	0,0197	0,0270											0,3700	0,5073	0,4318	0,3455	0,3108
13								0,0067			0,0225	0,0135						0,1659	0,1481	0,1409	0,1713	0,1351
14								0,0874	0,0316	0,0545	0,1320	0,1216						0,0202	0,0049	0,0136	0,0056	
15								0,2466	0,2549	0,2591	0,1770	0,2297										
16								0,2444	0,2451	0,3318	0,3034	0,1757	0,0022	0,0024								
17								0,2309	0,1383	0,1682	0,2079	0,2162										
18								0,1614	0,3131	0,1818	0,1320	0,2297	0,0202		0,0045	0,0140						
19								0,0202	0,0146	0,0045	0,0197	0,0135	0,0762	0,1650	0,1091	0,0478	0,0811					
20								0,0022					0,1570	0,1092	0,1364	0,1994	0,1486					
20.2															0,0045	0,0028						
21													0,1951	0,0825	0,1318	0,2135	0,2703					
21.2													0,0022		0,0045	0,0028						
22													0,1906	0,1505	0,1591	0,1601	0,1622					
22.2													0,0112	0,0218	0,0364	0,0028						
23													0,1143	0,1214	0,1318	0,1264	0,0541					
23.2													0,0022	0,0073	0,0045							
24													0,1233	0,2476	0,1364	0,1180	0,1486					
24.1															0,0182							
24.2													0,0045	0,0049								
25													0,0561	0,0558	0,0818	0,0730	0,0676					
25.2													0,0045	0,0170	0,0045							
26													0,0404	0,0146	0,0364	0,0253	0,0541					
27																0,0112	0,0135					
n	446	270	446	412	220	356	74	446	412	220	356	74	446	412	220	356	74	446	412	220	356	74
Pmin	0,0114	0,0195	0,0107	0,0122	0,0215	0,0138	0,0660	0,0118	0,0123	0,0233	0,0157	0,0752	0,0125	0,0137	0,0250	0,0158	0,0694	0,0114	0,0118	0,0230	0,0145	0,0682
PD	0,8582	0,8713	0,7860	0,8216	0,7590	0,8099	0,7816	0,9259	0,8984	0,8979	0,9179	0,8985	0,9637	0,9588	0,9684	0,9561	0,9364	0,8794	0,8408	0,8772	0,8811	0,8576
PE	0,4523	0,4881	0,3585	0,4139	0,3325	0,3872	0,3516	0,5874	0,5270	0,5325	0,6033	0,6030	0,7219	0,7043	0,7584	0,7007	0,6830	0,4909	0,4251	0,4803	0,4948	0,4951
PIC	0,6535	0,6869	0,5443	0,6167	0,5286	0,5780	0,5524	0,7595	0,7151	0,7186	0,7686	0,7713	0,8462	0,8345	0,8691	0,8330	0,8206	0,6808	0,6165	0,6690	0,6868	0,6861
H(obs)	0,7220	0,7704	0,6054	0,7087	0,6182	0,6573	0,6757	0,7713	0,7233	0,7545	0,8539	0,8649	0,8565	0,8786	0,8545	0,8652	0,7568	0,7265	0,6553	0,7364	0,7472	0,7297
H(exp)	0,7074	0,7371	0,5880	0,6774	0,5963	0,6243	0,6335	0,7937	0,7584	0,7619	0,7999	0,8127	0,8635	0,8532	0,8851	0,8529	0,8508	0,7276	0,6631	0,7152	0,7343	0,7408
SE	0,0305	0,0379	0,0330	0,0326	0,0468	0,0363	0,0792	0,0271	0,0298	0,0406	0,0300	0,0641	0,0230	0,0247	0,0304	0,0265	0,0586	0,0298	0,0329	0,0430	0,0331	0,0720

F.1. táblázat Öt populációs minta STR allélgyakorisági értékei és populációgenetikai alapjellemezői I.

	CSFIPO					F13A01		D7S820					LPL		D8S1179					TH01				
	BuCa	BaRo	DeRo	BuAs	Kisk	BuCa	BaRo	BuCa	BaRo	DeRo	BuAs	Kisk	BuCa	BaRo	BuCa	BaRo	DeRo	BuAs	Kisk	BuCa	BaRo	DeRo	BuAs	Kisk
3.2						0,0605	0,0815																	
4						0,0336	0,0444																	
5						0,1906	0,3778													0,0022				
6			0,0045			0,3274	0,2333													0,2197	0,2184	0,1955	0,2725	0,2162
7		0,0024				0,3117	0,1963	0,0090	0,0121	0,0227	0,0253	0,0270								0,1592	0,0850	0,1318	0,1152	0,2027
8	0,0022	0,0243		0,0028		0,0045		0,1435	0,0777	0,1318	0,1067	0,1216			0,0112	0,0146		0,0112	0,0135	0,1143	0,2354	0,2136	0,1292	0,1622
9	0,0381	0,0121	0,0136	0,0140	0,0946			0,1368	0,1602	0,1455	0,1264	0,1351	0,0404	0,0630	0,0157	0,0024	0,0045	0,0140		0,2085	0,1699	0,2000	0,2022	0,1216
9.+1								0,0022																
9.3																								
10	0,3072	0,2888	0,3000	0,3034	0,2838			0,2937	0,1286	0,1909	0,2669	0,3784	0,4081	0,2926	0,0762	0,0243	0,1000	0,0562	0,0405	0,0135	0,0024	0,0091	0,0365	0,0135
11	0,2623	0,2500	0,3545	0,3202	0,1892			0,2354	0,5121	0,3091	0,2472	0,2027	0,2668	0,2741	0,0807	0,1408	0,0818	0,0506	0,0405					
12	0,3184	0,3762	0,2773	0,3006	0,3514	0,0022		0,1502	0,0947	0,1773	0,1994	0,1216	0,2489	0,3704	0,1816	0,1529	0,1182	0,1404	0,1622					
13	0,0628	0,0437	0,0500	0,0534	0,0811	0,0135	0,0037	0,0247	0,0146	0,0091	0,0253	0,0135	0,0291		0,2870	0,2573	0,3000	0,3258	0,3108					
14	0,0090	0,0024		0,0056		0,0135	0,0296	0,0045		0,0136	0,0028		0,0067		0,2399	0,1626	0,2364	0,2219	0,1892					
15						0,0291	0,0296								0,0830	0,1942	0,1000	0,1011	0,2162					
16						0,0090									0,0202	0,0485	0,0545	0,0758	0,0270					
17						0,0022	0,0037								0,0022	0,0024	0,0045	0,0028						
18						0,0022									0,0022									
n	446	412	220	356	74	446	270	446	412	220	356	74	446	270	446	412	220	356	74	446	412	220	356	74
Pmin	0,0114	0,0122	0,0227	0,0145	0,0694	0,0116	0,0193	0,0117	0,0116	0,0229	0,0147	0,0707	0,0110	0,0181	0,0124	0,0126	0,0239	0,0146	0,0735	0,0122	0,0128	0,0234	0,0146	0,0682
PD	0,8770	0,8554	0,8542	0,8531	0,8649	0,8991	0,8960	0,9292	0,8495	0,9283	0,9289	0,8985	0,8573	0,8467	0,9319	0,9435	0,9332	0,9297	0,9058	0,9142	0,9134	0,9217	0,9246	0,9102
PE	0,4844	0,4557	0,4411	0,4514	0,5143	0,5347	0,5405	0,5986	0,4673	0,6030	0,6041	0,5632	0,4408	0,4299	0,6246	0,6478	0,6327	0,6233	0,5922	0,5835	0,5603	0,5913	0,5926	0,5841
PIC	0,6810	0,6566	0,6473	0,6562	0,7027	0,7146	0,7182	0,7658	0,6477	0,7684	0,7697	0,7356	0,6431	0,6388	0,7812	0,7995	0,7862	0,7782	0,7599	0,7571	0,7410	0,7640	0,7625	0,7577
H(obs)	0,7175	0,7136	0,7091	0,7416	0,7568	0,7489	0,7556	0,7578	0,6214	0,7273	0,7640	0,7838	0,6547	0,6519	0,8386	0,7670	0,7909	0,7584	0,8378	0,8251	0,7864	0,7636	0,7584	0,7297
H(exp)	0,7316	0,7116	0,7079	0,7140	0,7549	0,7550	0,7567	0,7975	0,6818	0,8009	0,8015	0,7775	0,6994	0,7007	0,8089	0,8253	0,8142	0,8051	0,8012	0,7916	0,7792	0,7998	0,7961	0,8012
SE	0,0297	0,0316	0,0434	0,0339	0,0707	0,0288	0,0369	0,0269	0,0325	0,0381	0,0299	0,0684	0,0307	0,0394	0,0263	0,0265	0,0371	0,0297	0,0656	0,0272	0,0289	0,0382	0,0302	0,0656

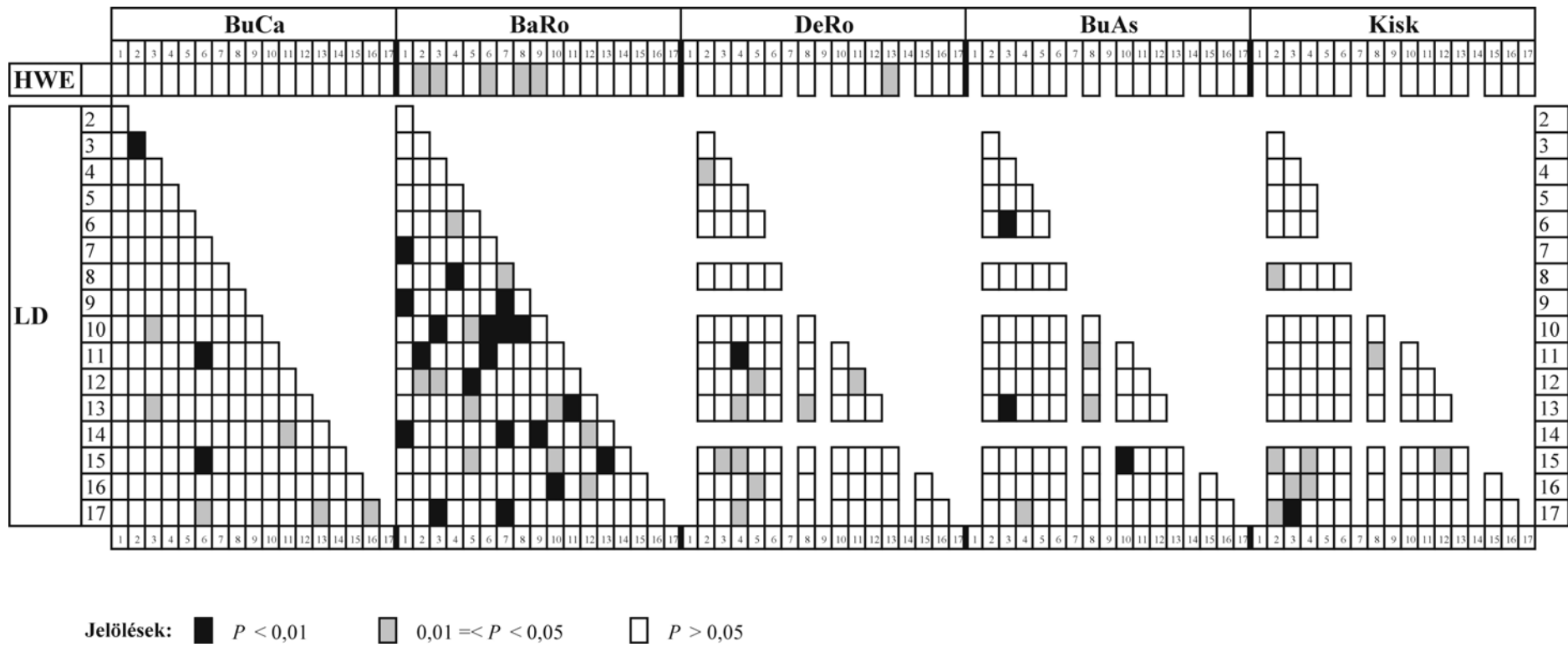
F.2. táblázat Öt populációs minta STR allélgyakorisági értékei és populációgenetikai alapjellemezői II.

	VWA					D13S317					FES		D16S539				
	BuCa	BaRo	DeRo	BuAs	Kisk	BuCa	BaRo	DeRo	BuAs	Kisk	BuCa	BaRo	BuCa	BaRo	DeRo	BuAs	Kisk
6								0,0045									
8						0,1368	0,1990	0,1727	0,0899	0,1757	0,0179	0,0074	0,0224	0,0194	0,0045	0,0393	0,0135
9						0,0964	0,1165	0,0864	0,1180	0,0541			0,1031	0,1481	0,1818	0,0955	0,0946
10						0,0493	0,0485	0,0500	0,0478	0,0405	0,2466	0,1815	0,0538	0,1917	0,1455	0,0590	0,0405
11						0,3274	0,1796	0,3455	0,3876	0,2297	0,4484	0,4370	0,2960	0,2282	0,2318	0,2781	0,3108
12						0,2601	0,3665	0,2500	0,2303	0,3649	0,2242	0,3741	0,3027	0,1917	0,2636	0,2837	0,2973
13	0,0022		0,0045	0,0028		0,0919	0,0752	0,0727	0,0843	0,0946	0,0605		0,1816	0,2039	0,1545	0,1798	0,1892
14	0,1076	0,1044	0,1227	0,1236	0,0676	0,0381	0,0146	0,0182	0,0309	0,0405	0,0022		0,0404	0,0170	0,0136	0,0618	0,0541
15	0,1143	0,0728	0,0818	0,1096	0,1622				0,0112						0,0045	0,0028	
16	0,2063	0,3083	0,2591	0,1573	0,2027												
17	0,3072	0,2670	0,3227	0,3848	0,3378												
18	0,1726	0,2209	0,1773	0,1770	0,1351												
19	0,0717	0,0170	0,0227	0,0365	0,0946												
20	0,0157	0,0073	0,0091	0,0084													
21	0,0022	0,0024															
n	446	412	220	356	74	446	412	220	356	74	446	270	446	412	220	356	74
Pmin	0,0123	0,0123	0,0250	0,0151	0,0671	0,0119	0,0122	0,0222	0,0148	0,0707	0,0113	0,0179	0,0120	0,0130	0,0239	0,0146	0,0650
PD	0,9263	0,9089	0,9020	0,9111	0,9102	0,9244	0,9170	0,9164	0,9138	0,8736	0,8327	0,7773	0,9051	0,9323	0,9228	0,9240	0,9014
PE	0,6162	0,5514	0,5659	0,5658	0,5892	0,5892	0,5699	0,5685	0,5660	0,5639	0,4296	0,3478	0,5626	0,6182	0,5983	0,5970	0,5510
PIC	0,7764	0,7314	0,7409	0,7368	0,7571	0,7551	0,7409	0,7402	0,7343	0,7358	0,6305	0,5608	0,7375	0,7819	0,7676	0,7622	0,7291
H(obs)	0,8296	0,7282	0,8545	0,8034	0,7027	0,7803	0,7136	0,6727	0,7753	0,7838	0,7085	0,6296	0,7982	0,8107	0,7909	0,7584	0,6486
H(exp)	0,8050	0,7702	0,7785	0,7693	0,7975	0,7866	0,7739	0,7763	0,7664	0,7786	0,6854	0,6384	0,7739	0,8122	0,8021	0,7941	0,7760
SE	0,0265	0,0293	0,0396	0,0316	0,0661	0,0274	0,0291	0,0397	0,0317	0,0683	0,0311	0,0414	0,0280	0,0272	0,0380	0,0303	0,0685

F.3. táblázat Öt populációs minta STR allélgyakorisági értékei és populációgenetikai alapjellemezői III.

	D18S51					D21S11					HPRTB		
	BuCa	BaRo	DeRo	BuAs	Kisk	BuCa	BaRo	DeRo	BuAs	Kisk	BuCa teljes	BuCa nő	BuCa férfi
8											0,0030	0,0047	
9											0,0242	0,0187	0,0345
10	0,0022		0,0045	0,0140							0,0030	0,0047	
11	0,0157		0,0045	0,0112							0,1303	0,1449	0,1034
12	0,1121	0,0801	0,1182	0,1292	0,1216						0,3515	0,3364	0,3793
12.-2											0,0030		0,0086
12.-6											0,0030	0,0047	
13	0,1502	0,1068	0,0455	0,1826	0,0541						0,2909	0,2897	0,2931
14	0,1704	0,1092	0,2091	0,1376	0,2703						0,1424	0,1402	0,1466
15	0,1188	0,1917	0,1409	0,2360	0,1622						0,0455	0,0514	0,0345
16	0,1121	0,1092	0,1727	0,0815	0,0946						0,0030	0,0047	
17	0,1166	0,2767	0,1182	0,1039	0,1351								
18	0,0830	0,0316	0,0500	0,0618	0,0811								
19	0,0516	0,0850	0,0455	0,0253	0,0405								
20	0,0426	0,0049	0,0364	0,0112	0,0135								
20.2			0,0045										
21	0,0157		0,0409	0,0056									
22	0,0067				0,0135								
23		0,0049	0,0045										
24			0,0045		0,0135								
25	0,0022												
25.2									0,0028				
26						0,0022							
26.2									0,0056				
27						0,0493	0,0049	0,0227	0,0169				
28						0,1906	0,0995	0,1409	0,1461	0,3243			
29						0,1951	0,2791	0,2318	0,2837	0,2027			
29.2						0,0022							
30						0,2175	0,1432	0,2182	0,2079	0,1622			
30.2						0,0493	0,0680	0,0500	0,0393	0,1216			
31						0,0628	0,0510	0,0318	0,0421	0,0135			
31.2						0,0964	0,0825	0,0864	0,0899	0,1081			
32						0,0112	0,0024	0,0045	0,0028				
32.2						0,0762	0,2136	0,1864	0,1292	0,0270			
33.2						0,0404	0,0388	0,0227	0,0281	0,0405			
34.2						0,0067	0,0170	0,0045	0,0028				
35									0,0028				
n	446	412	220	356	74	446	412	220	356	74	330	214	116
Pmin	0,0123	0,0133	0,0260	0,0154	0,0752	0,0126	0,0130	0,0254	0,0151	0,0707	-	0,0225	-
PD	0,9740	0,9520	0,9593	0,9573	0,9423	0,9578	0,9476	0,9365	0,9463	0,9277	-	0,9068	0,7356
PE	0,7640	0,6793	0,7377	0,7070	0,6992	0,7072	0,6687	0,6664	0,6599	0,6109	-	0,5420	0,5092
PIC	0,8722	0,8178	0,8554	0,8364	0,8303	0,8360	0,8101	0,8106	0,8045	0,7716	-	0,7218	0,6943
H(obs)	0,8341	0,8350	0,9000	0,8315	0,8649	0,8610	0,8058	0,8727	0,8034	0,7838	-	0,6542	-
H(exp)	0,8856	0,8388	0,8729	0,8554	0,8586	0,8545	0,8322	0,8358	0,8281	0,8093	-	0,7627	-
SE	0,0213	0,0256	0,0318	0,0264	0,0573	0,0236	0,0260	0,0353	0,0283	0,0646	-	0,0411	-

F.4. táblázat Öt populációs minta STR allélgyakorisági értékei és populációgenetikai alapjellemzői IV.



F.5. táblázat Öt populációs mintán felmért 13-17 autoszómás STR lokusz HWE és LD tesztelési eredménye. A lokuszok jelölése: 1: F13B, 2: TPOX, 3: D3S1358, 4: FGA, 5: D5S818, 6: CSF1PO, 7: F13A01, 8: D7S820, 9: LPL, 10: D8S1179, 11: TH01, 12: VWA, 13: D13S317, 14: FES, 15: D16S539, 16: D18S51, 17: D21S11

a,

P	BuCa																	BaRo																	DeRo																	BuAs																
	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	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
BaRo	[Black]																	[Black]																	[Black]																	[Black]																
DeRo	[Pattern]																	[Pattern]																	[Pattern]																	[Pattern]																
BuAs	[Pattern]																	[Pattern]																	[Pattern]																	[Pattern]																
Kisk	[Pattern]																	[Pattern]																	[Pattern]																	[Pattern]																

b,

F_{ST}	BuCa																	BaRo																	DeRo																	BuAs																	Kisk																
	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	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]																	[Pattern]																	[Pattern]																
BaRo	[Pattern]																	[Pattern]																	[Pattern]																	[Pattern]																	[Pattern]																
DeRo	[Pattern]																	[Pattern]																	[Pattern]																	[Pattern]																	[Pattern]																
BuAs	[Pattern]																	[Pattern]																	[Pattern]																	[Pattern]																	[Pattern]																
Kisk	[Pattern]																	[Pattern]																	[Pattern]																	[Pattern]																	[Pattern]																

c,

Φ_{ST}	BuCa																	BaRo																	DeRo																	BuAs																	Kisk																
	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	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]																	[Pattern]																	[Pattern]																
BaRo	[Pattern]																	[Pattern]																	[Pattern]																	[Pattern]																	[Pattern]																
DeRo	[Pattern]																	[Pattern]																	[Pattern]																	[Pattern]																	[Pattern]																
BuAs	[Pattern]																	[Pattern]																	[Pattern]																	[Pattern]																	[Pattern]																
Kisk	[Pattern]																	[Pattern]																	[Pattern]																	[Pattern]																	[Pattern]																

Jelölések: ■ $(F_{ST}, \Phi_{ST}) > 0,03; P < 0,01$ ■ $0,03 \geq (F_{ST}, \Phi_{ST}) > 0,02$ ■ $0,02 \geq (F_{ST}, \Phi_{ST}) > 0,01; 0,01 \leq P < 0,05$ □ $(F_{ST}, \Phi_{ST}) < 0,01; P > 0,05$

F.6. táblázat Az öt populációs minta páronkénti összehasonlítása 13-17 autoszómás STR lokuszra. **a**, G-statisztika; **b**, F-statisztika: F_{ST} és valószínűségi értékek; **c**, AMOVA: Φ_{ST} és valószínűségi értékek. A lokuszok jelölését lásd az F.5 táblázatban

7.3. Rövidítések jegyzéke

AMOVA	molekuláris varianciaanalízis (Analysis of Molecular Variance)
BaRo	a Baranya megyei roma populációs mintánk
BuAs	a budapesti askenázi populációs mintánk
BuCa	a budapesti kevert jellegű populációs referencia mintánk
CODIS	Combined DNA Index System (az USA bűnügyi DNS-adatbázis rendszere)
Cof	multiplex PCR-alapú, D3S1358-D16S539-Amel-TH01-TPOX-CSF1PO-D7S820 genotipizáló vizsgálati rendszer (Applied Biosystems AmpF/STR Cofiler kit)
CTT	triplex PCR-alapú, CSF1PO-TPOX-TH01 STR genotipizáló vizsgálati rendszer (Promega kit)
CTTV	kvadruplex PCR-alapú, CSF1PO-TPOX-TH01-VWA STR genotipizáló vizsgálati rendszer (Promega kit)
DeRo	a kelet-magyarországi roma populációs mintánk
EEO	etnikai származás becslése
EM	Expectation-Maximization algoritmus
Eu-YHRD	európai Y-STR haplotípus referencia adatbázis
F.	Az értekezés Függelék fejezete
FFFL	kvadruplex PCR-alapú, F13A01-FES-F13B-LPL STR genotipizáló vizsgálati rendszer (Promega kit)
h	Nei-féle géndiverzitás
HFV	triplex PCR-alapú, HPRTB-FES-VWA STR genotipizáló vizsgálati rendszer (Promega kit)
H_{exp}	becsült heterozigótaság
H_{obs}	megfigyelt heterozigótaság
HWE	Hardy-Weinberg egyensúly
Kisk	a fülöpszállási kiskun mintánk
LE	kapcsoltsági egyensúly
LD	kapcsoltsági egyensúlytalanság (linkage disequilibrium)
LR	valószínűségi hányados
MV	átlagolt fenotípus-különbségi érték

n	kromoszómák (allélok) száma
N	személyek száma
PD	igazságügyi megkülönböztető erő
PE	igazságügyi szülősegi kizáró erő
PIC	polimorfizmus információ tartalom
pM	igazságügyi egyezési valószínűség
p_{min}	minimális allélgyakoriság
P+	multiplex PCR-alapú, D3S1358-VWA-FGA-Amel-D8S1179-D21S11-D18S51-D5S818-D13S317-D7S820 genotipizáló vizsgálati rendszer (Applied Biosystems AmpF/STR Profiler Plus kit)
SD	szórás
SE	standard hiba
SNP	egynukleotid polimorfizmus (Single Nucleotide Polymorphism)
STR	Short Tandem Repeat

8. ÖSSZEFOGLALÁS

Célom a humán mikroszatellita (short tandem repeat, STR) polimorfizmusok genetikai analízisének a magyarországi kriminalisztikai gyakorlatba történő bevezetése, hitelesítése, valamint egy olyan magyar populációs referencia adatbázis létrehozása, amelynek segítségével az STR lokuszok vizsgálatán alapuló DNS-eredmények bizonyító ereje populációstatisztikailag megbecsülhető a magyarországi törvényszéki eljárások során.

Magyarországon elsőként végeztük el összesen 17 autoszómás, egy X-kromoszómához és kilenc Y-kromoszómához kapcsolt STR lokusz igazságügyi genetikai analízisét.

Öt magyarországi populációs mintán, összesen 754 személyen végeztünk genotipizálást. Referencia populációs mintánkat a budapesti kevert populációból véletlenszerűen kiválasztott 223 személy jelentette, akiket az összes kiválasztott marker (*F13B*, *TPOX*, *D3S1358*, *FGA*, *D5S818*, *CSF1PO*, *F13A01*, *D7S820*, *LPL*, *D8S1179*, *TH01*, *VWA*, *D13S317*, *FES/FPS*, *D16S539*, *D18S51*, *D21S11*, *HPRTB*, *DYS19*, *DYS389-I/II*, *DYS90*, *DYS391*, *DYS392*, *DYS393*, *DYS385*) genotipizálásával jellemeztünk. Az igazságügyi aspektusú STR analízisekbe a világon elsőként vontuk be a roma populációt. A baranyi roma minta esetében a *HPRTB* lokusz kivételével mindegyik lokuszt genotipizáltuk. A baranyai roma mintánk mellett egy kelet-magyarországi roma, egy budapesti askenázi zsidó valamint egy fülöpszállási kiskun populáció elemzésére is sor került 13 autoszómás (az ún. USA CODIS törzs-lokuszok) és kilenc Y-STR markerre nézve.

13 autoszómás STR marker (a CODIS törzs-lokuszok) analízisével biztosítottak tűnt a Magyarországon elkövetett bűncselekményekben érintett személyek individualizációja.

Az öt populációs mintában a lokuszokon mért allélgyakoriságok és az Y-STR haplotípusok eloszlása tekintetében számos esetben szignifikáns különbségeket detektáltunk. A budapesti referencia populációs mintánk autoszómás STR allélgyakoriság-eloszlása nagyfokú hasonlóságot mutatott más kaukázusi népcsoportokkal. A budapesti referencia és a baranyai roma populációt magában foglaló 41 európai férfi populáció Y-STR molekuláris varianciaanalízise alapján a baranyai roma populáció nagy mértékben elkülönült a vizsgált európai populációk többi tagjától.

A polimorf mikroszatellita markerek genotipizálása révén felállítottunk egy olyan referencia populációs adatbázist, ami megfelelőnek bizonyult a Magyarországon elkövetett bűncselekmények szakértői vizsgálatának korrekt statisztikai genetikai interpretációjához. Megállapítottuk, hogy a magyar népesség genetikai szubstrukturáltsági foka lényeges hatással lehet a DNS-vizsgálat bizonyító erejének statisztikai kiértékelésére. Eredményeink azt mutatták, hogy az alpopulációs teória alapján a magyarországi népcsoportokban élő személyek genetikai korrelációját figyelembe kell venni a DNS-profil egyezési valószínűség kalkulációja során annak érdekében, hogy a törvényszéki DNS-vizsgálat bizonyító erejének bemutatásakor ne történjen tévedés a terhelt kárára.

Az igazságügyi genetikai gyakorlatban vizsgált autoszómás STR markerek alkalmasak lehetnek a bűncselekmény elkövetésekor biológiai nyomot hátrahagyó tettes etnikai származásának megbecsülésére Magyarországon is.

9. SUMMARY

The aim of this study was to introduce and validate genetic analysis of human microsatellite (short tandem repeat, STR) polymorphisms to Hungarian forensic practice. In addition, our purpose was also to establish a Hungarian reference population database which could be used for interpreting the weight of DNA evidence presented at Hungarian court.

Forensic genetic analysis at 17 autosomal, one X-linked and nine Y-linked STR loci was performed for the first time in Hungary.

Altogether 754 individuals from five Hungarian populations were genotyped. The sample collected from the mixed Budapest population was chosen as our reference population sample. The reference Budapest population sample was characterized for all the loci (*F13B*, *TPOX*, *D3S1358*, *FGA*, *D5S818*, *CSF1PO*, *F13A01*, *D7S820*, *LPL*, *D8S1179*, *TH01*, *VWA*, *D13S317*, *FES/FPS*, *D16S539*, *D18S51*, *D21S11*, *HPRTB*, *DYS19*, *DYS389-I/II*, *DYS90*, *DYS391*, *DYS392*, *DYS393*, *DYS385*). Of the Romany ('Roma') populations living worldwide, the Romany individuals residing in Baranya county were the first Roma people to be genotyped at forensically relevant STR loci. The Baranya Romany population sample was investigated at all the loci, except for *HPRTB* marker. Forensic genetic analysis for the 13 autosomal loci (i.e. the CODIS core markers) was also performed in three other Hungarian population samples collected from Romany individuals living in eastern Hungary, from the Ashkenazi Jews residing in Budapest, and from Kiskun ('Small Cumanian') people living in a small village Fülöpszállás.

Based on our findings, the analysed 13 autosomal CODIS core loci seemed to be suitable for forensic individualization in Hungary.

The allele frequency and Y-STR haplotype distributions, in many cases, were significantly different between the analysed population samples. The population data of the reference Budapest sample were very similar to those observed in the other Caucasian populations. By performing analysis of molecular variance for Y-STR haplotypes of males from 41 European populations including our Budapest reference and Baranya roma samples, the Baranya roma population was found to be relatively distantly related to the rest of the populations of European descent.

The STR population database generated from the mixed Budapest population sample proved to be appropriate for the statistical interpretation of DNA evidence in Hungarian forensic cases. Our results suggest that the possibility of population differentiation should be taken into account on the shaping of expert opinions. In Hungarian forensic cases an overstatement of the strength of the DNA evidence against a defendant could be avoided by incorporating measures for genetic correlation detected in the forensically relevant subpopulations into the calculation of DNA-profile match probabilities.

Estimating of the ethnic origin of individuals appeared to be a promising approach by DNA-profiling of polymorphic STR loci in biological samples collected from Hungarian crime scenes.