



**Mitokondriális DNS és mikroszatellita polimorfizmusok
igazságügyi genetikai aspektusú vizsgálata a magyar
népességben**

DOKTORI ÉRTEKEZÉS TÉZISEI

Készítette:

Egyed Balázs

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

2007

1. BEVEZETÉS

Egy személy DNS állománya egy másiktól nagyon nagy számú eltérést tartalmazhat, ezeknek a variációknak az egyszerű meghatározásával kézenfekvő eljárás nyerhető az individuumok DNS mintából való azonosítására. Önmagában egyetlen lókuszt sem rendelkezik csak egyetlen egyedre jellemző tulajdonsággal, azonban a megfelelő számú polimorf helyen megállapított allélok együttese – ún. genetikai profil – „genetikai személyi számként” értelmezhető. Az egyedek genetikájáról alkotott kérdések és megállapítások ugyanakkor nem fogalmazhatóak meg elszigetelten önmagukban, mivel a kérdésekre adandó válaszok – pl. hogy két DNS minta ugyanazon személytől származik-e – abban a populációban megfigyelt genetikai variációk szerkezetétől és szintjétől függ, melybe a kérdéses személy is tartozik. Ezért az egyedi variációkat mindig populációs összefüggésbe kell helyezni, és megfelelő statisztikai módszert kell alkalmazni a különböző hipotézisek tesztelésére.

A polimeráz lánreakció (PCR) felfedezésének évében, 1985-ben került sor a DNS-vizsgálatok első törvényszéki alkalmazására Angliában. A mikroszatelliták (Short Tandem Repeats) igazságügyben való alkalmazhatóságát is a PCR-technika gyors elterjedése tette lehetővé, a PCR termékek ciklus szekvenálásával pedig megnyílt az út a mitokondriális DNS szekvenciában rejlő polimorfizmusok irányába. A mikroszatellita lókusztok olyan nem átíródó DNS-régiók, ahol 2-6 bázispárból álló rövid szakaszok tandem módon ismétlődnek egymásután. A repeat egységek ismétlődéseinek száma az egyes egyének között nagy változatosságot mutat, ez a polimorf jelleg teszi őket alkalmassá genetikai személyazonosítási célokra. Az STR markerek jelenleg kulcsszerepet töltenek be igazságügyi célú alkalmazások területén, amit nagy számuk és viszonylag magas fokú polimorfizmusuk mellett egyszerű PCR amplifikálhatóságuknak köszönhetik. Az utóbbi években a mitokondriális DNS szekvencia-polimorfizmusainak igazságügyi szempontú analízise és alkalmazása rendkívüli mértékben megnőtt, jórészt annak köszönhetően, hogy a mitokondriális genom vizsgálati érzékenysége nagyságrendekkel nagyobb, mint a nukleáris DNS-ben lokalizálódó STR lókusztoké, mely tulajdonsága alkalmassá teszi degradálódott, akár több éve, évtizede bomlásnak indult biológiai anyagmaradványok (pl. csontok) azonosítására is.

2. CÉLKITŰZÉSEK

Allél- és profilgyakorisági adatbázisok nélkül az STR vizsgálatok, haplotípus adatbázisok nélkül a mitokondriális DNS analízis törvényszéki bizonyító erejének statisztikai interpretációja nem lehetséges. Munkám egyik alapvető célja volt ezért a mitokondriális DNS polimorfizmusok genetikai analízisének magyarországi igazságügyi gyakorlatba való bevezetése és hitelesítése, ami magába foglalta mitokondriális DNS haplotípusok automatizált generálását és ellenőrzött hitelesítését a magyar

populációban. Ugyanakkor a molekuláris genetikai technológia fejlődésével megjelenő multiplex PCR rendszerek és fluoreszcens detektálás lehetőségeit kihasználva multilokuszos (15-17 STR marker) analitikai módszerre épülő, relatíve gyors automatizált eljárással terveztük magyar populációs mintáink autoszóma STR alapú, igazságügyi célú statisztikai elemzését. A magyar populációkban generált (STR és mitokondriális) DNS-profilok genetikai analízise által olyan populációs referencia adatbázisok létrehozása volt a célom, amelyek segítségével a DNS azonosító vizsgálatok bizonyító ereje populációstatisztikailag megbecsülhetővé válhat a magyarországi törvényszéki eljárások során.

Korábbi populációgenetikai elemzések már felhívták a figyelmet arra, hogy a magyar népesség genetikailag erősen szeparált, ezért joggal vetődik fel az a kérdés, hogy egy referencia populációs adatbázis milyen feltételekkel és korlátokkal alkalmazható a magyar népességre. Ennek a tesztelésére több magyar, de nem feltétlenül magyarországi lokalizációjú populációs minta felmérését végeztem el az analízisbe bevont STR és mtDNS markerekkel, mindösszesen mintegy 1200 személyi mintát bevonva a populációs felmérésbe. Korábbi adatok nyomán a referencia populációs mintát a budapesti kevert populáció véletlenszerűen kiválasztott személyeiből állítottam össze. Megelőző vizsgálatok (vércsoportok, fehérjeanalízis) alapján a hazai roma és askenázi zsidó népesség mutatta a legnagyobb genetikai távolságot a többi magyar népcsoporttól, mely kérdést STR és mtDNS szinten is elemezni szándékoztam. A csíkszeredai székely és a gyimesi csángó populációk bevonása pedig lehetőséget teremtett arra, hogy olyan etnikumok is bekerüljenek a magyar népesség populációgenetikai felmérésébe, amelyek korábban nem vagy csak korlátozott mértékben kerültek feldolgozásra.

Munkám fő céljai a következők voltak:

1. Humán autoszóma STR markerek (15 tetra- és 2 pentanukleotid STR) multiplex vizsgálatával az összeállított populációs minták genotipizálása, allél- és profilgyakorisági táblázatok létrehozása;
2. A populációs felmérések során detektált mikrovariáns STR allélek jellemzése;
3. Az analizált STR polimorfizmusok géndiverzitásának, a polimorfizmus fokának, személyazonosító hatékonyságának megállapítása a magyar populációkban;
4. A vizsgált populációk Hardy-Weinberg egyensúlyi tesztelése az analizált STR markereken;
5. A humán mitokondriális DNS automatizált szekvencia-analízisének laboratóriumi bevezetése, a teljes kontroll régió szekvencia megállapítása az összeállított populációs mintákon, mtDNS haplotípus adatbázisok generálása;
6. A mtDNS kontroll régió genetikai diverzitásának, a polimorfizmus fokának, személyazonosító hatékonyságának megállapítása a magyar populációkban;
7. A felmérésbe bevont magyar populációk mtDNS haplocsoport szerkezetének felmérése;
8. A felmérésbe bevont magyar populációk genetikai strukturáltságának felmérése STR- és mtDNS-profilok alapján;

3. ANYAGOK ÉS MÓDSZEREK

A populációs felmérés hat magyar populációt, összesen 1194 populációs mintát ölelt fel. A biológiai mintákból (vér-, nyál-, szájnyalkahártya-törlet minták) a DNS kinyerése, a DNS-koncentráció meghatározása, az STR lókuszok multiplex-PCR sokszorozása, az STR genotipizálás DNS-fragmensanalízissel, a mikrovariáns STR allélok szekvencia-analízise standard protokollok szerint történt. A mtDNS teljes kontroll régiójának (16024-16569, 1-576) PCR alapú szekvenálását teljesen automatizált laboratóriumi eljárással végeztük, ami robotizált PCR amplifikációt és szekvenálást, valamint kizárólag elektronikus adattranszportot foglalt magába.

A mtDNS haplotípusok haplocsoportba sorolása a kontroll régió szekvencia-motívumai alapján haplocsoport-specifikus vagy haplocsoporthoz kapcsolt szekvencia polimorfizmusok alapján történt.

A populációgenetikai alapértékeket mind az STR, mind a mtDNS adatok vonatkozásában standard formulák alapján számoltuk ki. A mtDNS szekvencia-adatok populáció-statisztikai kiértékelését az *Arlequin ver.2001* szoftver felhasználásával végeztük. A mtDNS haplotípuspárok közötti eltérések kalkulációjához a *LISA* és a *PAUP* v4.0* számítógépes programokat vettük igénybe.

A Hardy-Weinberg egyensúly (HWE) tesztelését a Fisher-féle exact tesztnek továbbfejlesztett változata segítségével az *Arlequin ver.2001* populációgenetikai szoftverrel végeztük el. A HWE tesztelése során Bonferroni-korrekción is alkalmaztunk.

A populáció-párok homogenitás-vizsgálatát komputerezált G-statisztikával az *R×C* szoftver felhasználásával végeztük, az egyes populációk között a genetikai profilok interpopulációs variabilitásának becsléséhez hagyományos Wright-féle F-statisztika analízist és molekuláris varianciaanalízist (AMOVA) használtunk. A számításokat a *Genetic Data Analysis (GDA v.1.1)* és az *Arlequin ver.2001* szoftver segítségével végeztük el.

A mitokondriális DNS haplotípusadatok median joining network analízisét a *Network 4.2* számítógépes program felhasználásával a 360 erdélyi (székely és gyimesi csángó) populációs mintán végeztük el. A csángó és a székely populáció közötti migrációs rátára a *MIGRATE* szoftver felhasználásával következtettünk a közös leszármazás teóriáját alkalmazva.

4. ÚJ MEGÁLLAPÍTÁSOK (TÉZISEK)

4.1. Az igazságügyi genetikai célú populációs felméréseink során hat magyar populációs minta összesen 1194 személyének autoszóma STR-alapú genotipizálását végeztük el. A genetikai analízishez használt STR lókuszok körét úgy alakítottuk ki, hogy minden egyes minta legalább 13 lókuszon, az ún. CODIS törzs-markereken (*CSF1PO, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, FGA, TH01, TPOX, VWA*) meg legyen tipizálva.

Referencia populációs mintánkat a budapesti kevert populációból véletlenszerűen kiválasztott 223 személy jelentette, akiket az összes kiválasztott marker (az előbbieken túl még a *D2S1338*, *D19S433*, *Penta E*, *Penta D* lókuszok) genotipizálásával jellemeztünk [1, 3, 4]. A budapesti referencia populáció analízise a felmérésbe bevont és a laboratóriumunkba bevezetett összes multiplex STR rendszerrel (*ProfilerPlus*, *Identifiler*, *PowerPlex16*) megtörtént, így ezt a felmérést kisebb léptékű primer konkordancia tesztelésnek is tekinthetjük a kevert magyar populációban. A különböző multiplex rendszerekkel az összes mintán átfedő markereken kapott megegyező genotípusok egybevágnak a nemzetközi adatbázisokon eddig elvégzett konkordancia felmérések eredményeivel.

A világon elsőként vontunk be nem magyarországi, de magyar populációkat (székely és gyimesi csángó) igazságügyi aspektusú autoszóma STR analízisbe. A populációkban az összes mikroszatellita marker genotipizálását elvégeztük [5]. Az erdélyi eredetű populációs minták mellett egy-egy dél-magyarországi (Baranya megye) és kelet-magyarországi (Debrecen környéki) roma populáció, valamint egy budapesti askenázi zsidó populáció elemzésére is sor került 15 autoszómás STR markerre nézve [4].

4.2. Populációs felméréseink során négy mikroszatellita lókuszon (*D7S820*, *D13S317*, *D18S51*, *FGA*) összesen hatféle olyan allélvariánst detektáltunk, melyek egyszerű fragmensanalízissel nem voltak definiálhatók. A mikrovariáns STR alléleket PCR termékük klónozásával és szekvenálásával jellemeztük [1, 2, 4]. A szekvencia-analízis során az STR lókusz polimorf ismétlődő szakaszát, valamint a repeatek határoló régióit érintő mutációkat is megfigyeltünk. A *D13S317* lókusz ritka 7-es alléljában egy 4 nukleotidos (TGTC) szakasz deléciója a reverz primer kötődési helyén az allél elvesztését (drop-out) okozta a módosított primerrel történő PCR-amplifikálás során [2]. A *D7S820* lókuszon detektált mikrovariáns allélek pontos DNS szekvenciáinak leírásával lehetőség nyílt a *D7S820* lókusz kromoszóma lokalizációjának pontosabb meghatározására a 7q11.23-q21 kromoszóma-szakaszra [1]. A kelet-magyarországi roma populáció több, egymástól látszólag független mintájában az *FGA* lókuszon leírt 24.1 mikrovariáns allél jelenléte, ennek az allélnek az alacsony előfordulási gyakorisága (<1%) a populációkban beltenyészetet vagy genetikai sodródást jelezhet a roma szub-populációban [4].

4.3. A markerek polimorfizmusának fokát jellemző paraméterek alapján a magyar népességben a DNS-alapú személyazonosításhoz az alkalmazott 17 autoszóma STR marker közül a *Penta E*, *FGA* és *D18S51* mikroszatelliták bírnak a legnagyobb igazságügyi megkülönböztető ill. individualizáló erővel, míg a *TPOX* lókusz mutatta a legalacsonyabb értéket.

Az elemzett populációkat egymással összevetve, általában alacsonyabb diverzitási értékeket kaptunk a roma szub-populációkban és a csángóknál, míg a budapesti referencia és askenázi zsidó populációban valamint a székely népesség esetében magasabb diverzitási értékek adódtak.

A hat populációs minta allélgyakoriság-eloszlása többé-kevésbé különbözött egymástól. Az allélgyakoriság-eloszlásokban megfigyelt jellegbeli eltérések heterogén alapító hatásokra és/vagy genetikai sodródási eseményekre utalhatnak a szub-populációkban. Az allélgyakoriság diagramok alapján nem állapítható meg etnikum specifikus allél vagy eloszlásgörbe egyetlen egymarkeren sem.

Felmértük az analizált STR lókuszek polimorfizmusának fokát minden egyes populációs mintában [1, 4, 5]. A 13 autoszómás (az ún. CODIS törzs-) lókuszra kombinált átlagos egyezési valószínűség (pM) 10^{-14} – 10^{-15} értékek között mozgott. A kombinált átlagos szülősegi kizáró erő (PE) ugyanerre a lókuszkészletre vonatkozóan 0,999983-0,999994 értéknek adódott. A 13 STR lókusból álló CODIS-rendszerrel az igazságügyi személyazonosítás hatékonysága mindegyik populációban rendkívül magasnak és egymással közel azonos nagyságrendűnek bizonyult. A 13 STR marker analízisével potenciálisan biztosítható a Magyarországon elkövetett bűncselekményekben érintett személyek individualizációja.

4.4. A lókuszekon belüli potenciális allélikus asszociáció statisztikai tesztelése (Hardy-Weinberg egyensúly) eredményeként az egyensúlytól való szignifikáns különbséget kaptunk a baranyai roma populáció D3S1358, D7S820, D8S1179, CSF1PO, Penta D, TPOX lókuszei [1, 4], a kelet-magyarországi roma minta D13S317 lókusza [4], valamint a gyimesi csángó népesség D5S818, D21S11 lókuszei [5] esetében. Ugyanakkor a Bonferroni-eljárás alkalmazásával az Exact-teszt analízise során kapott valószínűségi értékek meghaladták a teszt sorozatok egyedi szignifikancia szintjét. Mindezek alapján sem a roma mintákra, sem a csángó populációra nem vethetjük el a HWE meglétét, ha a azt a populációk teljes genomi szintjén vizsgáljuk.

4.5. Magyarországon elsőként végeztük el a humán mitokondriális DNS kontroll régió polimorfizmusainak populációs felmérését. Egy magyar mitokondriális DNS-profil referencia adatbázis létrehozása érdekében négy populációs mintán – budapesti referencia, baranya megyei roma, csíkszeredai székely, gyimesi csángó – a mtDNS teljes kontroll régió szekvenciáját határoztuk meg összesen 776 személy vonatkozásában egy teljesen automatizált laboratóriumi eljárással [6, 7, 8]. A szekvencia-adatok generálása során alkalmazott módszer nullára redukálta az ún. „fantom rekombinációk” és hibás szekvenciák adatbázisba való kerülésének esélyét. A kettős ellenőrzésen és filogenetikai analízisen átesett DNS szekvenciák haplotípusai bekerültek az on-line hozzáférésű EMPOP igazságügyi célú mtDNS haplotípus adatbázisba (<http://www.empop.org>), ahol azok szabadon hozzáférhetőek [8].

Mitokondriális DNS analízisbe a világon elsőként vontuk be a hazai roma népeiséget, valamint romániai területi elhelyezkedésű magyar populációkat (székely és gyimesi csángó).

4.6. Az analízisbe bevont négy magyar populációs minta drámai módon különbözött egymástól mitokondriális DNS szinten. A budapesti kevert (referencia) népeiségben megfigyelt haplotípus diverzitás kisebb vagy jelentős mértékben csökkent a többi populációban [8]. A körülbelül azonos mintaméret ellenére a budapesti kevert populációs minta több mint háromszor annyi különböző megfigyelt haplotípussal rendelkezik, mint a Baranya Roma populáció és ötször több egyedi haplotípussal [6]. A csángó populációban szignifikánsan kevesebb különböző és egyedi haplotípus volt megfigyelhető a székely populációhoz képest, amellyel pedig szoros geográfiai szomszédságban és feltételezhetően genetikai kapcsolatban is áll [7]. A gyimesi csángó és a baranya megyei roma populációkban a két másik populációhoz viszonyítva tapasztalt szignifikánsan kevesebb megfigyelt haplotípus genetikai izolációs hatásokra utalhat ezekben az etnikai csoportokban. A budapesti referencia és a székely populációkban kalkulált genetikai diverzitási paraméterek megfelelnek az átlagos európai populációkban kalkulált adatoknak, a csángó és a roma populációkban viszont extrém alacsony értékek adódtak, melyek szembeütő módon eltérnek a publikált nemzetközi populációs adatoktól [8].

A leggyakrabban megfigyelt haplotípus mind a négy populációban más-más volt. A budapesti referencia populációban megfigyelt leggyakoribb haplotípus a legelterjedtebb nyugat-eurázsiai haplotípussal egyezik meg, és kisebb-nagyobb gyakorisággal mind a négy populációban előfordult. Ezen a haplotípuson kívül egyetlen egy fordult még elő mind a négy populációs mintában, viszont lényegesen kisebb gyakorisággal. A gyimesi csángó és a baranya megyei roma népeiség haplotípus kompozíciója feltűnő mértékben eltért egy átlagos nyugat-európai populáció képétől. A roma és a csángó populációkban megfigyelt mitokondriális DNS szintű homogenitás arra is utalhat, hogy ezekben a szub-populációkban a populációk közötti gén áramlás (gene flow) korlátozott.

4.7. A budapesti referencia, a baranya megyei roma, a csíkszeredai székely és a gyimesi csángó populációk haplocsoport szerkezete rendkívüli mértékben különbözött egymástól [8]. A budapesti referencia populáció haplocsoport eloszlása nagyon hasonló az átlagos nyugat-európai népeiségeket jellemző eloszlásra, a populációs mintában megfigyelt összes szekvencia több mint 90%-a besorolható valamelyik nyugat-eurázsiai haplocsoportba [6].

A haplocsoport eloszlás a székely populációk esetében is általában jól követi az átlag európai népeiségekre jellemzőt, viszont számottevően eltér attól a gyimesi csángóknál. Ennek oka az itt élő emberek eredetében és történelmében keresendő, amely jelentős hatást gyakorolhatott genetikai állományukra is. A csángó populáció torzult haplocsoport eloszlása szigetszerű genetikai izolációra

utalhat. Az ázsiai eredetű mitokondriális DNS haplocsoportok relatíve magas aránya a székely populációban (~ 8%) a székelységet a középkorban ért ázsiai eredetű nomád behatásra utal [7].

A baranya megyei roma populáció haplocsoport eloszlása sokkal inkább az európai cigány népségekből megfigyeltekre hasonlít az M haplocsoport nagyon nagy arányú (>30%) reprezentációjával, a maradék szekvencia nagy része pedig általában besorolható valamelyik nyugat-eurázsiai haplocsoportba. A baranyai roma populációban megfigyelt leggyakoribb haplotípus rendelkezik a legnagyobb prevalenciával az összesített európai roma populációs mintában is [6].

4.8. A populációk STR-profiljaiban és mtDNS haplotípus szerkezetében rejlő genetikai strukturáltság felmérése annak tisztázása érdekében történt, hogy a budapesti kevert népségből álló referencia-adatbázisunk allél- és haplotípus-gyakorisági értékei vajon milyen mértékben fogadhatók el hitelesnek a teljes magyar népségre nézve.

4.8.1. A hat populációs minta autoszóma STR szintű genetikai összetételében, a mikroszatellita lókuszokon mért allélgyakoriságok és allél-struktúrák eloszlása tekintetében szingifikáns különbségeket detektáltunk. 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 is, amit viszont az allélok molekuláris szerkezetén alapuló AMOVA csak csekély mértékben támasztott alá. A populáció-párok közül a legkisebb genetikai tagozódást a budapesti referencia/askenázi párosban kaptuk. A populáció-párok között kalkulált egylókuszos F_{ST} Φ -statisztika indexek lényegesen magasabb értékűnek adódtak a négy magyarországi populációs minta vonatkozásában, mint amit előzőleg egyéb európai populációkban megfigyeltek. Mivel nemzetközi felmérések során nagyobb F_{ST} és Φ_{ST} értéket elsősorban izolált ill. eltérő genetikai eredetű népcsoportokban detektáltak, ezért megfigyeléseink a magyarországi populációk eltérő genetikai múltjára ill. izoláltságára utalhat [1, 4].

Az autoszóma STR markerek többségének allélgyakoriság eloszlása a budapesti referencia populációnak a székely és csángó populációk viszonylatában szignifikáns eltérést mutatott a G-statisztika teszt alapján, ami az erdélyi és a budapesti populáció részben eltérő genetikai szerkezetére utalhat. Ez az eltérés a székely/csángó populáció-párban is megfigyelhető. A populáció-párok között kalkulált egylókuszos F_{ST} és Φ_{ST} értékek a lókuszok túlnyomó többségén nem erősítették meg az allélgyakoriság eloszlásban megfigyelt fenotípusos eltéréseket, ami alapján lényegesen kisebb mértékű szub-struktúráldás tételezhető fel a budapesti referencia, a csíkszeredai székely és a gyimesi csángó populációk között, mint a négy magyarországi populációban tapasztaltak [5].

A vizsgált 17 autoszómás STR között nem volt egy olyan marker sem, amely legalább egy tesztben ne adott volna szignifikáns eltérést. Mindegyik populációs mintán elemzett 15 marker közül, a populáció-páronkénti összehasonlításban a legkisebb változatosságot a D5S818, CSF1PO és TH01 lókuszon tapasztaltuk. A legnagyobb genetikai variációt a D3S1358, D18S51, Penta E és TPOX lókuszon detektáltuk. Ha populációkat nem párokban, hanem együtt vizsgáltuk, a megfigyelt F_{ST} értékek alapján a D7S820 lókuszt mutatta a legnagyobb variabilitást. Ugyancsak magas variációt kaptunk a TPOX és Penta D lókuszek esetében. A adatbázisok között a legkisebb változatosságot a CSF1PO mikroszatellita mutatta. Mindezek alapján az analizált 17 autoszómás STR marker közül a magyar népességben a CSF1PO és TH01 lókuszt bír a legkisebb, a D7S820 és TPOX lókuszt rendelkezik a legnagyobb információ-tartalommal a populációk elkülöníthetőségére vonatkozólag. Az öt populáción mért F_{IS} értékek nagyságából arra is lehet következtetni, hogy a populációkon belüli beltenyésztés foka nem teljesen elhanyagolható.

Ha a populációk páronként történő AMOVA-elemzése során a 15 ill. 17 autoszómás lókuszból álló profilok molekuláris variancia tartalmát együtt vizsgáltuk, nem kaptunk jelentős különbséget a budapesti referencia/debreceni roma párban, valamint a budapesti referencia populáció székely és csángó populációkkal alkotott populáció-párosai, valamint a két erdélyi mintacsoport között. A többi párosításban azonban a molekuláris varianciaanalízis szignifikáns különbséget jelzett.

A budapesti referencia populáció több európai populációs mintával való, lókuszonkénti összehasonlításában a G-statisztika teszt alapján csak kevés eltérést tapasztaltunk [1, 4]. Budapesti referencia populációnk allélgyakoriság-eloszlása tehát nagyfokú hasonlóságot mutat más európai népcsoportokkal. Ezek az eredmények egybevágóan azzal a tudományos nézettel, amely szerint a kevert magyar népesség az indoeurópai népekkel közeli genetikai rokonságban áll.

4.8.2. Négy magyar populáció együttes mtDNS szekvencia-adatainak AMOVA analízise alapján a szekvenciák közötti összes variabilitás 94,8%-a köszönhető a populációkon belüli különbözőségeknek és a teljes variancia 5,2%-a vezethető vissza az interpopulációs eltérésekre. Ez az eredmény rendkívüli módon eltér azoknak a hasonló kalkulációknak az adataitól, melyeket nyugat-európai populációkban végeztek. A magyar populációk között megfigyelt nagyarányú variancia szignifikáns különbséget jelez a vizsgált népségek genetikai struktúrájában [8].

A populáció-párok között kalkulált F_{ST} értékek alapján a baranya megyei roma populáció a legnagyobb genetikai távolságot mutatta az összes többi magyar népcsoporttal szemben [6]. A másik mintanépség a gyimesi csángó volt, mely hasonló mértékű szignifikáns különbséget mutatott a többi magyar populációhoz viszonyítva. A vizsgálatok alapján a székely népesség genetikai struktúrája nagyon hasonlít a budapesti referencia populációhoz, ami viszont az átlagos nyugat-eurázsiai népséggel teljesen megegyező jelleget mutat [7].

Annak ellenére, hogy a baranya megyei roma mtDNS adatok haplocsoport szerkezete lényegileg megegyezik az európai roma népeiségekben megfigyelttel, alapvető haplotípus gyakoriságbeli különbségek voltak kimutathatók a magyar és európai roma populációk között. Ezt a genetikai különbséget megerősítette a baranya megyei és a többi európai roma populáció között páronként végzett F-statisztika nemzetközi adatok alapján igen magasnak minősülő értékei. Az összevont roma populációs mintában megfigyelt összes variancia 5,15%-a volt tulajdonítható a roma populációk közötti különbségeknek. A roma populációknak haplocsoport reprezentáció szinten megfigyelhető hasonlósága ellenére megállapítható, hogy léteznek olyan faktorok, melyek korlátozzák az anyai génvonalak egyes roma közösségek közötti áramlását [6].

A székely és csángó populáció mtDNS szekvencia adatait egy-egy görög és macedón populációs adatbázissal is összevetettük. A populáció-párok között kalkulált standardizált gényakoriságok variancia tartalmának (F_{ST}) összehasonlítása a csángó populációs minta mtDNS struktúrájának kifejezett különbségét mutatta az összes többi populációval szemben. Ezzel szemben a székely populáció mitokondriális DNS struktúrája jól illeszkedik a nyugat-eurázsiai típusokba [7].

A mtDNS haplotípusok hálózat-analízise alapján a csángó emberek nagyobb arányú beáramlására lehet következtetni a székely populációba, mint ellenkezőleg, ami aszimmetrikus migrációs rátát tételez fel a két népesség között. Ezt a hipotézist megerősítette a migrációs ráta becslése, mely szerint a csángó bevándorlás a székely populációba nagyobb arányú, mint fordítva [7].

A budapesti referencia populáció mtDNS szintű genetikai struktúrája lényegében megegyezik az európai őslakos népességével, ami azt jelenti, hogy a budapesti kevert mintacsoport szekvencia adatai beilleszthetőek igazságügyi célú mitokondriális DNS adatbázisokba és felhasználhatók az igazságügyi célú összehasonlítások során [8].

5. KÖVETKEZTETÉSEK, ÖSSZEGZÉS

Populációs felmérésünk eredményeképpen megvalósult az a célkitűzés, hogy polimorf autoszómás mikroszatellita lokuszok és a mitokondriális DNS kontroll régiójának tipizálása révén felállítsunk egy olyan populációs referencia adatbázist, ami megfelelően bizonyulhat a magyarországi igazságügyi genetikai szakértői vizsgálatok korrekt statisztikai genetikai interpretációjára.

Megállapítható, hogy a 13 autoszómás CODIS STR marker analízisével potenciálisan biztosítható a Magyarországon elkövetett bűncselekményekben érintett személyek individualizációja.

Megállapításra került, hogy a magyar népesség genetikai szubstrukturáltsági foka lényeges hatással lehet a DNS-vizsgálat bizonyító erejének statisztikai kiértékelésére. Mivel hasonló adatokat mindeddig csak eltérő történelmi-genetikai eredetű etnikai csoportok vagy geográfiaileg-genetikailag

izolált populációk között mérték, a magyar populációban tapasztalt nagyfokú heterogenitás a vizsgált magyar szub-populációk diverz genetikai történelmével és/vagy izolációjukkal magyarázható.

A referencia populációként definiált budapesti kevert népességből származó mintacsoport genetikai struktúráját tekintve lényegében megegyezik az őshonos európai népességgel. Ennek köszönhetően a budapesti referencia populáció autoszóma STR- és mtDNS-profil adatai beilleszthetőek az igazságügyi célú, nemzetközi populációs DNS adatbázisokba és statisztikai adatai felhasználhatók az igazságügyi célú összehasonlítások során. Mindazonáltal eredményeink azt mutatják, 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-vizsgálatok bizonyító erejének interpretálása tévedéstől mentes legyen.

6. KAPCSOLÓDÓ KÖZLEMÉNYEK

- [1] **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
- [2] 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
- [3] **Egyed B**, Csikai M, Füredi S, Pádár Z (2005) *Population genetic data on the STR loci D2S1338, D19S433 and SE33 in Hungary*. J Forensic Sci 50(3):720-721
- [4] **Egyed B**, Füredi S, Angyal M, Balogh I, Kalmar L, Pádár Z (2006) *Analysis of the population heterogeneity in Hungary using fifteen forensically informative STR markers*. Forensic Sci Int 158:244-249
- [5] **Egyed B**, Füredi S, Pádár Z (2006) *Population genetic study in two Transylvanian populations using forensically informative autosomal and Y-chromosomal STR markers*. Forensic Sci Int 158:244-249
- [6] Irwin J, **Egyed B**, Saunier J, Szamosi G, O'callaghan J, Pádár Z, Parsons TJ (2006) *Hungarian mtDNA population databases from Budapest and the Baranya county Roma*. Int J Legal Med DOI 10.1007/s00414-006-0128-4
- [7] Brandstätter A, **Egyed B**, Zimmermann B, Duftner N, Pádár Z, Parson W (2007) *Migration rates and genetic structure of two Hungarian ethnic groups in Transylvania, Romania*. Ann Hum Genet doi: 10.1111/j.1469-1809.2007.00371.x
- [8] **Egyed B**, Brandstätter A, Irwin JA, Pádár Z, Parsons TJ, Parson W (2007) *Mitochondrial control region sequence variations in the Hungarian population: Analysis of population samples from Hungary and from Transylvania (Romania)*. Forensic Sci Int Genet 1:158-162