



MTA DOKTORI ÉRTÉKEZÉS TÉZISEI

Génekbe vésett vallomások: DNS-ujjlenyomat és őstörténet

Pamjav Horolma

Nemzeti Szakértői és Kutató Központ
Bűnügyi Igazságügyi Szakértői Igazgatóság
Genetikai Szakértői Intézet
Referencia-minta Vizsgáló Osztály

Budapest, 2018

I. ELŐSZÓ

Pályafutásom során arra törekedtem, hogy létrehozzak egy olyan genetikai laboratóriumot, amely nemcsak hazai, hanem nemzetközi szinten is hallat magáról magasszínvonalú szakmai és tudományos tevékenysége folytán. Továbbá – más szakértők törekvéseivel egybeeső – szándékom volt, hogy az igazságügyi genetikai területén bevezetésre kerüljenek a legújabb tudományos eredményeken alapuló DNS vizsgálati eljárások, ideértve mind a műszerezettséget, mind a kapcsolódó nemzetközi módszertani előírások és követelmények meghonosítását. Az MTA doktori fokozat elnyerésére irányuló jelen értekezés bemutatja az igazságügyi genetikai területén a szerző által végzett alkalmazott kutatások koherens részét képező, a humán DNS vizsgálatokon alapuló rutin szakértői tevékenységek nélkülözhetetlen alapját jelentő allél- és haplotípus gyakoriságok felmérését a magyar népességben, a referencia adatbázis létrehozását, valamint a hazai igazságügyi genetikai gyakorlatba történő bevezetését, hitelesítését. Ezen adatok segítségével a DNS vizsgálaton alapuló eredmények bizonyító ereje statisztikailag igazolhatóbb (becsülhető) a hazai törvényszéki eljárások során. Az igazságügyi genetikai gyakorlatban használt módszerek, valamint nemhez kötődően öröklődő DNS tulajdonságok *ún., haploid markerek* vizsgálati eredményei közvetlen alapot adnak a humán populációk prehistorikus és historikus időszakban történt migrációinak útvonalfeltérképezéséhez: külön-külön, mint anyai és apai leszármazási vonalak, és együtt is, mint egy-egy populáció genetikai története. A kutatási területre több szakmai elnevezés használatos: *populációgenetika, filogenetika, genealógia és őstörténet*.

II. BEVEZETÉS

Az értekezés 2 részből áll. Az első fejezetben először megszólaltatom a „Génekebe vésett vallomások”/DNS üzenetek közül azokat, amelyek megmondják, mi a genetikai „személyi számunk”, azaz a DNS-ujjlenyomatunk. Prezentálom a DNS-ujjlenyomat elemzésének genetikai logikáját, valamint az apai és anyai leszármazási vonalak szerepét a személyazonosításban.

A második részben „kifaggatom” azokat a „géneket”, amelyek beavatnak bennünket a humán migrációs, evolúciós folyamatok fontos mozzanataiba, valamint a filogenetikai történetek közeli és távoli múltjába. Hihetetlen, de igaz, hogy az első fejezetben bemutatott apai és anyai leszármazási vonalak hátrányai a második fejezetben szereplő vizsgálati diszciplínában előnyökké válnak, azaz a sikeres anyák és a sikeres apák genetikai történetét tárják elénk. Ugyanitt Eurázsia különböző népei zenei örökségének elemzésére is sor kerül, bemutatva a népzene-kultúrák közötti "genetikai" korrelációk hálózatát is. Válasz adható a kérdésre, hogy

közös zenei örökséggel rendelkező kultúrák genetikai értelemben is szorosan összefüggenek-e. Ezzel talán egy új interdiszciplináris kutatási területet is megalapoztunk, a "zenegenetikát".

II.1. DNS-UJJLENYOMAT: IGAZSÁGÜGYI GENETIKA (FORENSIC GENETICS)

Cavalli-Sforza: A fő genetikai különbség nem a népek vagy az emberfajták, hanem az egyedek között van.

Az igazságügyi genetikai vizsgálatok alapvető kérdése, hogy a törvényszéki eljárások során bizonyítékként szolgáló biológiai nyom milyen biztonsággal eredeztethető egy adott fajtól, annak egy egyedétől, populációjától. A kérdésfeltevés tárgya azonban nemcsak a biológiai anyagmaradvány, hanem egy egyed vagy azok egy csoportja (populáció, etnikai csoport) is lehet, vizsgálva az egyed vagy az egyedek csoportosulása származási, leszármazási vagy más rokonsági kapcsolatait. Azaz az igazságügyi genetika a törvényszéki eljárásokba vonható összes genetikai származtatással kapcsolatos kérdést/vizsgálatot magában foglalja. Egy apasági vizsgálat is genetikai származtatásnak tekinthető, hiszen ebben az esetben egy személyt származtatunk egy feltételezett szülőpártól. Etnikai csoportok anyai vagy apai ágú rokonsági viszonyainak tisztázása is genetikai származtatás, hiszen ez esetben az egyedek alkotta csoportok feltételezett vagy ismert közös őseinek csoportjától való származási lehetőségeit vizsgáljuk. Tehát, ha a vizsgálat tárgyköre törvényszéki eljárásba vonható, akkor az az igazságügyi genetikai diszciplínába tartozik, ez különbözteti meg az igazságügyi genetikát a többi genetikai szakterületről.

A humán igazságügyi genetikai vizsgálatok fő feladata annak megállapítása, hogy a törvényszéki eljárás során, a bűncselekmények helyszínén hátrahagyott biológiai anyagmaradványokból, valamint polgári peres eljárás esetében a személyektől levett mintákból (apasági, anyasági rokonsági vizsgálatok) meghatározott DNS-profilok hozzárendelhetők-e egy adott személyhez. Ebből világosan az következik, hogy a DNS-profil személy- vagy egyedazonosítási célokra történő felhasználása csak akkor lehet eredményes, ha a DNS-ből tévedést kizáró bizonyossággal kiolvasható olyan tulajdonság, amely egyedül csak az adott személyre jellemző. Ennek a tulajdonságnak időben és térben viszonylag stabilnak, változatlanak kell lennie, hogy a genetikai-szakértői vizsgálatok során megbízható eredmény születhessen, mivel bizonyos bűncselekmények esetében akár több tíz évvel a történés után is sor kerülhet a DNS minta szakértői vizsgálatára, pl. az áldozat csontmaradványainak előkerülése okán. Az így nyerhető objektív természettudományos bizonyíték a bírósági döntéshozatalt segíti. A DNS vizsgálathoz olyan markereket (jellegeket) kell megszólaltatni, amelyek egyrészt sokféle

változatban figyelhetők meg az adott népességben, másrészt függetlenek az adott ember életkorától, pillanatnyi fiziológiai állapotától és egyéb külső tényezőktől.

A humán populációk DNS vizsgálatával a következő kérdésekre lehet választ kapni:

1. az emberi populációk, mint egy-egy földrajzi régióban élő közösségek között a vizsgált markerek allélgyakorisági eloszlásában kimutathatók-e eltérések?
2. a különböző földrajzi régiókban élő populációk (etnikumok és rasszok) allélgyakorisági adatai között kimutathatók-e karakterisztikus különbségek (földrajzi régiókban vagy kontinenseken)?
3. milyen biztonsággal használhatók fel a vizsgált/megszólaltatott markerek genetikai variabilitásai egyedek személyazonosítására és/vagy személyek egymás közötti rokonságának megállapítására (apa-gyermek, anya-gyermek, testvérek stb.)?
4. az egyének vagy biológiai anyagmaradványok DNS állományának vizsgálata alapján kiolvashatók-e a minta tulajdonosának külső fenotípusos jellegei, mint például nem, szemszín, hajszín, bőrszín, testmagasság?
5. kiolvasható-e úgyszintén az egyén etnikai eredete vagy egy bizonyos földrajzi régióból való származása?
6. meghatározható-e, hogy egyes egyének vagy csoportok ugyanazt a leszármazási vonalat képviselik-e apai vagy anyai vonalon?

A fenti kérdések igen nagy társadalmi és jogi jelentőséggel bírnak, mivel ma már a DNS alapú vizsgálatok a bűncselekmények felderítésében és a rokonsági tesztekben az igazságszolgáltatás alapvető eszközévé váltak. Az egyének genetikai struktúrájára vonatkozó kérdések nem fogalmazhatók meg pusztán önmagukban, mivel a kérdésekre adható válaszok az adott populációra jellemző genetikai variabilitástól függenek, 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.

Az igazságügyi genetika mára önálló tudományterületté nőtte ki magát, amely felhasználja a genetika módszereit, tudományos eredményeit. Önálló voltát támasztja alá az a körülmény is, hogy igazságügyi genetikai kutatások a genetika más területére vonatkozó információkat (pl. molekuláris genetika, evolúciógenetika, populációgenetika, filogenetika, apai és anyai leszármazási vonalak feltérképezése) is szolgáltatnak, a kutatások önmagukban ugyanakkor nem jelentik az igazságügyi genetika célját. Törvényszéki szempontból nézve a genetika nem más, mint eszköz, melynek segítségével a jogi hipotézisek megerősíthetők vagy elvethetők. Az igazságügyi genetika, mint az alkalmazott tudományok autonóm részterülete napjainkra szinte

minden országban az igazságszolgáltatás részévé vált, és számos vonatkozásban kiszélesítette a bűncselekmények felderítésének és törvényszéki perek bizonyításának lehetőségét.

Az elmúlt 2 évtizedben keletkezett DNS kutatási ismeretek felelnek meg leginkább a törvényszéki eljárás által támasztott, igen magas szintű követelményrendszernek, így napjainkban a biológiai anyagmaradványok DNS-mintázat (DNS-profil) vizsgálata az igazságügyi célú személyazonosításban a egyeduralkodóvá vált. A biológiai nyomból meghatározható DNS-profil egyediségének kérdését alapvetően több tényező befolyásolja. Az egyik, hogy az emberi nem 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 vizsgálandó lokuszok polimorfizmusának mértéke populációról populációra és földrajzi területenként is változhat. Az egyediséget korlátozó további tényező, hogy a bűncselekmények elkövetése során keletkező biológiai nyom nem ritkán több személy mintájának keveredéséből jöhet létre. Mindezen megfontolások alapján az igazságügyi genetikus-szakértők – eltérően pl. az ujjlenyomat-szakértők által alkalmazott gyakorlattól – jelenleg nem kategorikus, hanem valószínűsítő véleményt adnak a személyazonosítás vagy származás-megállapítás során.

Az emberiség sokszínűsége főleg az emberek közötti különbözőségekből ered. Populációgenetikai szempontból ezek a különbözőségek polimorfizmus mintázatban tárolódnak és adódnak át generációról generációra. Ezek a különbözőségek tesznek minket teljesen egyedivé, kivéve, ha van egypetűjű ikertestvérünk. A DNS-ujjlenyomat vizsgálatok alapját a nem kódoló DNS szakaszok – intron és intergenikus régió - vizsgálata adja. Polimorf mikroszatelliták ezreit azonosították a humán genomban. Az STR (Short Tandem Repeat) szekvenciákat ismétlődő egységük hosszának megfelelően nevezik el: a di-, tri-, tetra-penta-, hexa-repeat egységeknek. Az igazságügyi genetikában leggyakrabban alkalmazott STR lokuszok általában tri- és tetranukleotidokból állnak, ezek a legalkalmasabbak az egymással nem rokoni viszonyban álló személyek közötti genetikai különbözőségek feltárására, a lényegesen nagyobb allélszámnak és heterozigóciás értékeknek köszönhetően. A személyazonosítás mellett a DNS vizsgálat rutin eljárásnak számít azon származási és rokonsági ügyekben is, ahol egymással potenciális rokonságban álló személyek profilját vetjük össze. Ezek közé tartozik a hagyományos szülősegi vizsgálat, ahol általában a biológiai apaság eldöntése a kérdés, vagyis a „ki a gyermek apja?” kérdés megválaszolása. Eltűnt személyek és tömegkatasztrófák áldozatainak azonosítása céljából fordított (reverz) szülősegi vizsgálat is végezhető, ilyen ügyekben pl. az a kérdés, „származhatnak-e a biológiai maradványok a referenciamintát adó személy gyermekétől?”.

Az Y-kromoszómának, mint a humán genom kizárólag férfiakra jellemző szakaszának vizsgálata, az erőszakos bűncselekmények férfi elkövetőinek azonosítását tette még eredményesebbé. Ez leginkább azokban az eljárásokban érhető tetten, amelyekben a kevert jellegű biológiai nyomok női eredetű sejtjeinek túlsúlya miatt az autoszómás markerek analízise nem informatív. Egy adott Y-STR haplotípust nem csak egyetlen személy hordozhat, hanem a kérdéses személy minden apai ági férfirokona, így a gyanúsított vagy a vélelmezett apa édesapja, nagyapja, fiútestvére, nagybátyja, fia, fiú unokája, fiú unokatestvére stb. szintén ugyanazzal az Y-kromoszómális STR haplotípussal rendelkeznek. Az Y-SNP lokuszok (haplocsoport) vizsgálata és a vizsgált populációk Y-kromoszómális haplocsoport-összetétele igazságügyi genetikai szempontból nem bír meghatározó jelentőséggel, mivel az Y-SNP haplocsoportok száma az Y-STR haplotípusok számánál jóval kisebb. Egy haplocsoportba sok, különböző Y-STR haplotípus tartozik, ezért a haplocsoportot definiáló SNP lokuszok vizsgálata leginkább kizárás megállapítására alkalmas, egyezés esetén megfelelően magas valószínűségi hányados sem származási, sem bűnügyekben nem szolgáltatható. Viszont az Y-SNP vizsgálat adhat információt az illető földrajzi származásáról, amely a nyomozóhatóság munkáját segítheti. Az Y-kromoszómális SNP-k az STR lokuszok mellett populációs és evolúciós tanulmányok informatív eszközei is. Fontos szerepet játszanak a humán migrációval foglalkozó tanulmányokban, mivel lehetővé teszik az egyes populációk, illetve populáció-csoportok közötti lényeges különbségek hatékony feltárását.

A mitokondriális DNS hipervariábilis szakaszainak PCR alapú szekvencia-analízisével a DNS alapú azonosítást a biológiai anyagmaradványok teljes körére kiterjesztették. Az igazságügyi szempontból fontos megkülönböztetési erő – kizárási esély – részben a mitokondriális DNS maternális öröklődése miatti alacsonyabb voltát a sejtenkénti magasabb kópiaszám ellensúlyozza, amely szignifikánsan növeli a vizsgálatok sikerességi rátáját.

Az értekezésben szereplő X-STR lokuszok a kromoszómán négy különböző kapcsoltsági alcsoportban (klaszter) helyezkednek el, mindegyik kapcsoltsági alcsoporton belül 3-3 szorosan kapcsolt lokusz található, amelyek úgy viselkednek, mintha haplotípusok lennének. Nőkben a lokusz-triókon belül a rekombináció valószínűsége igen csekély, míg közöttük (alcsoportok között) a rekombináció gyakori. Az X-STR vizsgálatok rokonsági tesztekben való alkalmazásához a kapcsoltsági alcsoporton belüli lokusz-triók előfordulási gyakoriságait (3 lokusból álló haplotípus), valamint férfiakban a mind a 12 X-STR lokuszra vonatkozó haplotípus gyakoriságokat kell felmérni.

Az igazságügyi genetikus-szakértő valószínűsítő szakértői véleményt ad, ha nincsenek kizáró kombinációk az összehasonlító minták között (pl. a bűncselekmény helyszínéről

származó és a gyanúsítottól vett minta). A gyakoriság becsléssel kombinált Bayes-elvű hipotézisvizsgálás a gyakorlatban legáltalánosabban elfogadott módszer a DNS-bizonyíték statisztikai kiértékelésére. A valószínűségszámítás Bayes-tételének esély formájú leírása szerint ugyanis az alapkérdés (*hipotézis*) a következő:

- „mekkora a valószínűsége annak, hogy a DNS-profil XY-től származik?”
- „mekkora a valószínűsége annak, hogy a vélelmezett apa a biológiai apja a gyermeknek?”

A feltett alapkérdésen – hipotézisen – kívül mindig legalább egy másik, ún. *ellenhipotézist* is fel kell állítani:

- „mekkora a valószínűsége annak, hogy a DNS-profil az adott népességből véletlenszerűen kiválasztott másik személytől származik?”
- „mekkora annak a valószínűsége, hogy a gyermek biológiai apja valaki más az adott népességből?”

Az első kérdés általában a vád (*prosecutor*) feltevésére, a második kérdés pedig általában a védelem (*defence*) hipotézisére vonatkozik. A *hipotézis* (*Hp*) és *ellenhipotézis* (*Hd*) teljesülésének feltételezésével számított valószínűségek arányát, mint *valószínűségi hányadost* (LR, *Likelihood Ratio*) adjuk meg az igazságügyi DNS-vizsgálatok statisztikai interpretációja során. A vérrokonság (pl. apaság) genetikai tesztelése esetében a valószínűségi hányadost apasági indexnek hívják (PI, *Paternity Index*). Ebben az esetben a vád hipotézise a *felperesi* és a védelem hipotézise az *alperesi* hipotézisnek felel meg. Amennyiben a biológiai nyom csak egy személy genetikai anyagát tartalmazza, abban az esetben a valószínűségi hányados – legegyszerűbb formájában – a DNS-profilegyezés valószínűségének (Match Probability) reciprokával egyezik meg. Alapvető gyakorlat függetlenül öröklődő lokuszok esetében a „product rule” alkalmazása, vagyis az egyes lokuszokon kimutatott allélok populáción belüli gyakorisági értékeinek összeszorozása. Amennyiben azonban számottevő az ún. *linkage disequilibrium* (LD), továbbá, ha populáció strukturálódás figyelhető meg a populációban, a számított profil gyakoriságok tévesek lehetnek.

II.2. ÖSTÖRTÉNET ÉS AZ EMBER GENETIKAI ÚTJA

Richard Dawkins: Mindenki, akire csak pillantásuk esik, mikor kilép az időgépből, vagy egyetlen őse az emberiségnek, vagy senkinek sem őse.

Manapság már vitathatatlan tény, hogy a modern ember (*Homo sapiens sapiens*) mintegy 60 ezer évvel ezelőtt hagyta el Afrikát, és röpké 5-6 ezer év alatt eljutott a földkerekség minden zegébe-zugába. Egyes csoportok letelepedtek a mai Arábiában, míg mások dél felé vándoroltak,

átszelve a mai Irak, Irán és Afganisztán területét így eljutva Közép-Ázsiába. Nagy valószínűséggel az utóbbi csoportból származó emberek eljutottak a csendes-óceáni szigetvilágba, valamint Ausztráliába. A harmadik csoport Európa felé vette útját.

A 21. század küszöbén új vizsgáló technológia jött létre: a DNS alapú filogenetika (*phylogenetics*) vagy genealógia. A kizárólag apai ágon öröklődő Y-kromoszóma, valamint a kizárólag anyai ágon öröklődő mitokondriális DNS (mtDNS) nukleotid összetételének vizsgálatával lehetővé vált az összes eddig élt és ma élő ember egyetlen óriási „globális családfán” történő elhelyezése. A globális családfa egyben felvázolja azt a hosszú utat is, amelyet már létező fajként tettünk meg az emberiség bölcsőjétől, Afrikától, a földkerekség legeldugottabb zugáig. A DNS alapú genetika tudománya nagyívű térképet nyújt őseink vándorlási útvonalairól, valamint hozzávetőlegesen dátumokat is tud rendelni az egyes meghatározó földrajzi állomásokhoz. Ez nem kevesebbet jelent, mint hogy a ma élő emberek sejtjeikben hordozzák történelmük egy fontos darabját. A ma élő emberekből következtethetünk őseik nagyívű migrációs útvonalaira és a maiak egymáshoz viszonyított, történelmi léptékben mérhető, rokonsági kapcsolataira. Az Y-kromoszóma és az mtDNS forradalmi jelentősége abban rejlik, hogy több évezredes időtávban is viszonylag sértetlenül adódik tovább tisztán anyai vagy apai leszármazási vonalon, így nincs kitéve a *rekombináció* következményeinek. Remélhető, az értekezés olvasója rávezethető arra, hogy távoli őseink, az első modern emberek, közös utat tettek meg az „Édenkertből” – fekete Afrikából – kiindulva; egy-egy földrajzi régióban hol rövidebb, hol hosszabb ideig tartózkodva, elhozva nekünk genetikai nyomaikat a távoli és közelebbi múltból oda, ahol ma élünk. Csak nézőpont kérdése, hogy térben és időben, mikortól és honnan kiindulva keressük genetikai őseinket, rokonainkat, hiszen a filogenetika egyik fő következtetése, hogy az emberiség egyetlen nagy családot alkot...

Az emberiség őstörténete emberek és populációk vándorlásának története. Minden egyes fontosabb demográfiai esemény nyomot hagyott hátra a populációk genetikai diverzitásában. Ha egy populáció mérete csökken, genetikai diverzitása is csökken, ha egy populáció mérete nő, genetikai diverzitása is nő. Ha a vándorló népek egymással keverednek, az a népek között nagyfokú hasonlóságot, míg az izoláció genetikai egyediséget eredményez. Ezek a demográfiai bélyegek generációról generációra adódnak át, ily módon a ma élő emberek genetikai állománya visszatükrözi demográfiai múltjukat. Tehát történelmünk mintegy meg van írva DNS-ünkben.

Genealógiai célból az mtDNS-nek egy gyakran mutálódó szakaszát, a *hipervariábilis szakaszt* (kontroll régió vagy D-loop) vizsgálták sokáig, de ma már a teljes genom szekvenálása a cél. Ez a szakasz nincsen kitéve szelekciós nyomásnak, az itt keletkezett mutációk

megőrződnek, így az itt tárolt információk nagyon változatosak. Filogenetikai szempontból pont ez a régió hordozza őseink genetikai szignálját, azaz Éva genetikai üzenetét leányainak.

A nukleáris DNS által közvetített genetikai üzenet a rekombinációnak köszönhetően általában 1-2 generációra visszamenőleg szolgáltat leírást múltunkról, vagyis mindkét szülő felmenőiről. Ez alól csak a nukleáris DNS egy kicsiny része kivétel, mégpedig az egyik nemi kromoszóma: nevezetesen az Y-kromoszóma, amely „mesél” a régmúlt időszakról. A Y-kromoszóma ugyanis az mtDNS-hez hasonlóan, de kizárólag apai vonalon öröklődik, és így alkalmas apai leszármazási vonalak egyértelmű követésére. Az Y-kromoszóma apáról fiúra örökön-örökké rekombináció nélkül továbbítódik egyik nemzedékről a másikra, ha minden nemzedékben legalább egy fiú születik. Analóg módon az anyai leszármazási vonalak mtDNS-ben tárolt genetikai információ segítségével történő nyomon követéséhez, az apai leszármazási vonalak is feltérképezhetők az Y-kromoszóma által közvetített genetikai üzenetek segítségével. Ez az ősapák üzenete fiaiknak, azaz Ádám genetikai üzenete.

A tanulmány szempontjából fontosabb apai haplocsoportok – apai leszármazási vonalak – elterjedését és a haplocsoportot definiáló SNP "születésének" legvalószínűbb keletkezési helyét és a fontosabb expanziós térségeket bővebben az értekezés tartalmazza. A haplocsoportok hús-vér emberekre, olyan de facto sikeres ősapákra mutatnak, akiknek fiai, unokái és távoli utódai az apai leszármazási vonalak mentén széles körben elterjedtek.

A haplocsoport az egy apai leszármazási vonalhoz tartozó férfiak csoportját jelenti. Az adott Y-SNP megléte/hiánya alapján minden férfiről egyértelműen eldönthető, hogy egy adott haplocsoportba tartozik-e vagy sem. Példaként bemutatjuk az „eurázsiai Ádám” haplocsoportját, aki a ma élő összes nem-afrikai férfi őse. Őt az úgynevezett M168 SNP-vel definiáljuk. Az *eurázsiai Ádám* Y-kromoszómáján a 14.813.991 pozícióban keletkezett egy mutáció, azaz C → T szubsztitúció történt, kb. 65-70 ezer évvel ezelőtt. Annak az esélye, hogy pont ebben a pozícióban, egy másik férfiban is keletkezzen ugyanilyen mutáció, 1:57 millióhoz, vagyis rendkívül csekély. Ezért is hívjuk ezt „egyszeri mutációs eseménynek” (*unique mutation event*). Ennek az ősapának a férfi leszármazottjai ma döntően az eurázsiai térség lakói. Magukat a haplocsoportokat betűk és számok kombinációjaként jelöli a tudomány, amelyek első tagja önkényes, a további betűk és számok pedig az alcsoportok leszármazási sorrendjében következnek (pl. CT-M168, N1c-M46, R1a-Z93 stb.)

A korbecslés a filogenetika legvitatottabb területe. A különböző mutációs ráták, és így a haplocsoportok különböző mutációs rátákkal számított korbecslései között nagy különbségek vannak, így alapvetően csak relatív kronológiáról beszélhetünk azok keletkezésére vonatkozóan. Nagy biztonsággal csak annyi állítható, hogy két haplocsoport közül melyik az idősebb, és az

idősebb jóval idősebb-e vagy sem, mint a másik. A korszámítást (TMRA: Time to Most Recent Common Ancestor) STR mutációs rátával (evolúciós vagy pedigré) vagy SNP alapú NGS (Next Generation Sequencing) módszerrel szokták számolni.

Megállapítást nyert, hogy Európa benépesülése négy (paleolitikum, mezolitikum, neolitikum és bronzkor) meghatározó migrációs hullámmal írható le. A genetikai adatok azt látszanak megerősíteni, hogy az első hullámmal (30-40 ezer éve) olyan őskőkorszaki népesség (C-F3393, F*-M89 ősapja) érkezett Európába, amely az Ausztráliát is meghódító ősi embercsoportokkal mutat rokonságot.

Úgy 27-30 ezer éve született az I*-M170 ősapja Európában. Az I apai vonal az európai őslakókra jellemző. Ma is megtalálható egész Európában az alcsoportjaival együtt.

A földművelés a Termékeny Félholdról 2 útvonalon terjedt Európába úgy i.e. 6-5 ezer éve: az egyik útvonal a Balkán felől a Duna-medencébe érkező földművelőket jelentette – régészetiileg a *Starčevo-Körös kultúrával* (i. e. 5800-4500) jellemezhető. A földművelés másik elterjedési vonala a Földközi-tenger térségében volt, ahol a *benyomásos kerámia (Cardium Pottery) kultúrája* terjedt el i.e. 6400 körülől. A korai földművelők csontanyagának többségét a Kaukázus vidékéről származó G2a haplocsoport jellemzi. A mai Európában általában 1-3% között van jelen a G2a haplocsoport, de egyes területeken, így Olaszországban, Görögországban, Svájcban és Ibéria nyugati részén 5-10%-os arányban, Szicília és Korzika szigetén pedig 10 és 15% között mutatható ki a jelenlétük.

Végezetül Európa mai népességének Y haplocsoport összetétele a leginkább elfogadott populációgenetikai hipotézis szerint a bronzkorban (i.e. 3000-1500 év) indoeurópai nyelveket beszélő népek migrációs hullámai következtében alakult ki, párhuzamosan a korábbi mezolitikumi népesség jelentős eltűnésével. A bronzkori európai régészeti kultúrák (*Yamnaya, Vučedol, Bell Baker, Battle Axe, Corded Ware, Sintashta, Andronovo*) hordozóit jellemző R1a és R1b haplocsoportok együtt alkotják a mai európai férfi lakosság abszolút többségét. A rendelkezésre álló genetikai adatok alapján Európa népességének gyakori „kicserélődésére” kell következtetnünk, amely további vizsgálatra szoruló feltételezés.

Genealógiai célból a mtDNS-nek egy gyakran mutálódó szakaszát, a *hipervariábilis szakaszt* vizsgáljuk. A vizsgált személynek egy referencia szekvenciától – az „rCRS”-től – való eltéréseinek összessége határozza meg a haplotípust. A haplocsoportot hasonló haplotípussal rendelkező személyek csoportja adja meg, akik egy közös ősanától származtathatók. A mtDNS-ben egy vagy több pozícióban lévő bizonyos nukleotidok eltéréseinek (inszerciók, deléciók, szubsztitúciók) összessége (haplotípus) határozza meg a rCRS-től való eltérést. A kódoló régióban viszont egy-egy adott pozícióban lévő bizonyos nukleotidok megléte szükséges az adott

személy haplocsoportjának meghatározásához. Az Y-kromoszómális haplocsoportokkal szemben az anyai ági leszármazást jelző mtDNS haplocsoportok esetében az elnevezések sajnos nem követnek leszármazási logikát, azaz nem az „A” csoport az összes többtől legkorábban levált csoport, hanem a csoportok felfedezési sorrendjében nevezték el azokat.

III. ANYAG ÉS MÓDSZER

Az elmúlt évek során az autoszómális, X és Y-kromoszómális STR (mikroszatellita), illetve Y-kromoszómális SNP lokuszokon végeztünk felméréseket különböző populációs mintákon.

A kereskedelembe kapható különböző STR kitek által tartalmazott lokuszok számának növekedésével és/vagy új lokuszok bevonásával összhangban vizsgáltunk különböző egyedszámú magyar populációs mintákat. Az autoszómális STR vizsgálatok tekintetében 4213, 254, 455 és 21473 (összesen: 26395) magyar minta, valamint 738 mon-khmer nyelvi csoportba tartozó thaiföldi személy, azaz összesen **27133** minta vizsgálata történt meg.

Az X-kromoszómális STR-lokuszok vizsgálata esetében a magyar populációban 384 nem rokon, random szelektált személytől (219 férfi és 185 nő) történt mintavétel. A rekombinációs vizsgálatokhoz 45 háromgenerációs családból 155 személyt vizsgáltunk

Az Y-kromoszómális STR és SNP lokuszokra vonatkozó felmérések során különböző populációkban összesen 2045 mintán történt vizsgálat (11-23 Y-STR lokusz és több mint 50 SNP). Ezenkívül nemzetközi együttműködés keretében részt vettünk a gyorsan mutálódó STR-ek (Rapidly Mutating Y-Chromosomal Short Tandem Repeats) vizsgálatában 103 személy genotipizálásával 13 RM lokuszon, valamint 23 Y-STR vizsgálatában 143 magyar és 101 roma minta tesztelésével.

A magyar populáció genetikai struktúrájának felméréséhez különböző számú (8-16) STR-lokuszra vonatkozó allélgyakorisági, PIC, PD, PM, HET, $MEC^{Krüger}$ és $MEC^{Kishida}$ értékeket külön-külön határoztuk meg. Az Y-STR és Y-SNP lokuszokon pedig haplotípus és haplocsoport gyakoriságot, illetve diverzitást számoltunk, továbbá F-statisztika és AMOVA analízist végeztünk.

A vizsgált populációk és genetikai struktúrák felmérésére és a populációk közötti rokonsági kapcsolatok feltárására a következő módszereket használtuk: Struktúra analízis, G-teszt, MDS (Multidimensional Scaling), Network analízis, filogeográfiai analízis, Bayesian Skyline Plots (BSP) analízis.

IV. AZ EREDMÉNYEK ÉS AZOK MEGVITATÁSA

IV.1. IGAZSÁGÜGYI GENETIKA

IV.1.1. Autoszomális STR markerek vizsgálata:

A törvényszéki eljárásokban a DNS bizonyíték biostatistikai szempontú interpretációja szolgáltatja a bizonyító erőt a jogalkalmazók számára. A bizonyítékok együttes mérlegelése során kritikus tényező lehet egy biológiai anyagmaradvány megbízható, nagy „statisztikai erővel” bíró származtatása. A biostatistikai számítások elvégzését ma már nemcsak szakmai, hanem jogszabályi előírások is szabályozzák, kötelező elemei a szakértői véleménynek. A biostatistikai számítások Bayes elvű megközelítése szükségessé tette a hazai populációt minél jobban reprezentáló referencia adatbázis létrehozását.

Egy olyan, 21472 személy DNS-profiljából álló adatbázist hoztunk létre, amelyen legjobb tudomásunk szerint még nemzetközi szinten sem született. A büntetőügyekben ebből az adatbázisból származó allélgyakoriságokat használva az ún. DNS-profilegyezési valószínűség nagyon kicsi, a DNS-profil szinte az illető genetikai személy számának tekinthető, azaz a DNS-ujjlenyomata. Ezen allélgyakorisági adatok elfogulatlanabbak, pontosabbak, így módon a nagyobb adatbázis potenciálisan lehetővé teszi alacsonyabb allélgyakoriságok használatát.

Az allélgyakoriságon kívül a HWE p érték, minimum allélfrekvencia, valamint igazságügyi hatékonysági és statisztikai paraméterek (PIC, H_E , H_O , PD, p_M , PE) szintén a számítások részét képezték.

IV.1.2. X-rekombináció és X-STR markerek vizsgálata

Magyarországon elsőként vezettük be az X-STR vizsgálatokat és máig nincs más intézet, amely alkalmazná ezt a vizsgálatot. Az X-STR vizsgálatok, ahogy a vizsgálható lokuszok számai lehetővé tették, először 4 STR, utána 8 STR, majd 12 STR lokuszra történtek ugyanazon mintakészleten (219 férfi és 185 nő). Ennek eredményeként a férfiak vizsgálatával a magyar populáció X-STR haplotípus gyakoriságait az egész X-kromoszómára nézve, valamint a kapcsoltsági alcsoporton belül a lokusz-trióra vonatkozóan is felmértük. Ezen adatok segítségével igazságügyi genetikai szakértői véleményt tudunk adni az olyan *ún. hiányos apasági ügyekben*, amikor a vélelmezett apa elhunyt vagy eltűnt, és a perben szereplő gyermek leány. Ilyen esetekben a vélelmezett apa X-kromoszómáját vizsgáljuk indirekt módon az

elhunyt/eltűnt édesanyjának vagy törvényes lánygyermekének a vizsgálatával, a perben szereplő lánygyermekkel együtt.

Három generációs családokban végzett rekombinációs vizsgálatunk szerint az X-kromoszómán található 4 kapcsoltsági alcsoport között gyakran történik rekombináció: az 1. és a 2. kapcsoltsági alcsoport között 38,7%-ban, a 2. a 3. között 41,0%-ban, a 3. és a 4. között 36.1 %-ban. Viszont nagyon ritka a kapcsoltsági alcsoportokon (lokusz trió) belüli rekombináció.

IV.1.3. Y-STR vizsgálata az igazságügyi alkalmazásban

Az elmúlt 10 évben 2045 személy – nemcsak magyar és a hazai roma, hanem más eurázsiai populációk - Y-STR haplotipizálását végeztük el. A haplotípus/haplocsoport adatok egyetlen szakmai, nemzetközi Y-haplotípus referencia adatbázisba kerültek (www.yhrd.org), ezáltal nemcsak a hazai, hanem nemzetközi szakmai színvonal növeléséhez is hozzájárultunk. Az igazságügyi genetikusok szakértői vélemény készítése során az YHRD adatbázisban keresik meg a kérdéses személy haplotípus gyakoriságát az LR számításhoz. A legjobb tudomásunk szerint ennyi haplotípus adat más hazai intézmény által nem került be az említett adatbázisba.

Az igazságügyi genetikai gyakorlatban Y-STR vizsgálatot akkor célszerű végezni, amikor 2 személy DNS-e keveredett és a női DNS túlsúlyban van. Ilyen eset lehet például nemi erőszak után biztosított női hüvelyváladék, ahol értelemszerűen az áldozattól származó laphámsejtek jóval nagyobb arányban találhatóak, mint a férfi hímvarsejtjei. Azonkívül erőszakos bűncselekmény során dulakodásból adódóan a női sértett körme alá kerülhetnek férfitől származó bőrsejtek, melyeket az ún. „körömkaparék” tartalmaz. Az ilyen bűncselekményekből származó minták vizsgálatánál célszerű csak a férfitől származó Y-STR lokuszok tesztelésére összpontosítani.

Apasági DNS vizsgálatokban csak akkor alkalmazható Y-kromoszómális lokusz, ha a perben szereplő gyermek fiú és a vélelmezett apa elhunyt vagy eltűnt. Ilyen esetekben a vélelmezett apa valamelyik apai ági férfi rokona (nagyapja, testvére, nagybátyja, unokatestvére stb.) vizsgálata célravezető. Ha a vizsgált fiúgyermek és a férfi azonos haplotípussal rendelkezik, akkor feltételezhetően azonos apai felmenőktől származnak (LR számítás).

V. ÓSTÖRTÉNET: GÉNEKBE VÉSETT TÖRTÉNELEM

Az evolúciós történelmi tanulmányok során nagy szerep jut az mtDNS és az Y-kromoszóma MSY régiója elemzésének, mivel az emberi genom ezen régiói egyik generációról a másikra változás nélkül adódnak át, megtartva egy-egy populáció anyai és apai leszármazási vonalát. A populációk oszthatnak közös mtDNS vagy Y-kromoszóma leszármazási vonalon, ami közös eredetre, vagy génáramlási folyamatokra utal. A populációk közötti genetikai hasonlóság származtatható a közös eredetből vagy a közelmúltbeli földrajzi közelségből eredő keveredésből. A populációk közötti genetikai távolságok általában korrelálnak a földrajzi távolságokkal, az ún. távolság-függő elszigeteltségi modell szerint, de ez nem minden esetben releváns.

V.1. A férfi genetikai útja: Y-kromoszómás STR és SNP lokuszok vizsgálata magyar és roma populációkban

Megállapítottuk, hogy a kárpát-medencei magyar, székely és csángó populációk Y-kromoszómális haplocsoport eloszlásai közel azonosak egymással, valamint a közép- és kelet-európai országok populációival. Elszigeteltebb populációkban (székely és csángó) viszont jobban megőrződtek a dél-szibériai eredetű apai tulajdonságok (N1c és Q haplocsoport). Összehasonlítva a roma populációkat egymással, a magyar és a malajziai indiai populációkkal az állapítható meg, hogy a roma férfiak az eredeti indiai génállományuk egy részét megőrizték: a jelenleg csak indiai eredetű populációkban előforduló H1a, R2, L haplocsoportokat. Génállományuk másik részét a Balkánon (kb. 1000 éve jelentek meg a Balkánon) vagy a befogadó országok populációjától szerezték. Ezeket az átfogó vizsgálatokat elsőként végeztük el a hazai szinten.

A magyar férfi népesség kb. 20% -a (E, G, J haplocsoport) a neolitikumban 8-10 ezer éve a Termékeny Félhold területéről érkezett földművelők, 27%-a európai őslakók (I haplocsoport), 46%-a bronzkori keleti bevándorlók, azaz indoeurópaiak (R1a, R1b), 6%-a roma kisebbség (H, R2), 1%-a finnugor eredetű nép (N1c) leszármazottja. A csíkszeredai székely és gyimesi csángó minták is hasonló haplocsoport eloszlást mutatnak azzal a kivétellel, hogy a hazai populációhoz képest magasabb arányban jelennek meg dél-szibériai eredetű haplocsoportok: a székelyeknél 8%-ban (N1c), a csángóknál pedig 4%-ban (Q). A J2b haplocsoport 5-6%-ban jelen van mindkét magyar etnikumban, de hiányzik a magyarországi magyarokból. A J2b Görögországban gyakoribb, Nyugat-Törökországban ritkább, és ezért a dél-európai neolitikus kolonizáció során

tengeri útvonalon érkezett markerként értelmezhető. Ugyanez a marker a bodroközi magyarokban (regionális) is jelen van 3,4%-ban, valamint a N1c is 6,2%-ban.

A hazai roma kisebbségekben (kevert roma, tiszavasvári, taktaközi stb.) attól függően, hogy mennyire izolált a populáció – természetesen függ még a mintaszámtól is – egyes tulajdonságok genetikai sodródással eltűntek (R1a hiánya a tiszavasvári romákban) vagy feldúsultak (H1a haplocsoport), illetve különböző arányban megjelennek a befogadó népességre jellemző haplocsoportok. A vizsgált roma populációk haplocsoport eloszlását összehasonlítva a malajziai indiai haplocsoport eloszlásával, megállapítható, hogy a H1a haplocsoport az összes vizsgált roma mintacsoportban és a malajziai indiaiakban is megfigyelhető. A H1a haplocsoportba tartozó Y-kromoszómák Median Joining hálózata alapján pedig az összes vizsgált roma mintacsoport osztozik egy közös H1a Y-STR haplotípuson. Ez a klaszter, azaz a közös haplotípuson való osztozás, egymással közeli rokonsági viszonyban lévő Y-kromoszómák egy csoportját reprezentálja a roma populációkban, ráadásul kirajzolódik az is, hogy a hálózatban minden egyes haplotípus egyazon közös haplotípus leszármazottja, függetlenül a mostani földrajzi helyzetétől. Ez a H1a haplocsoport indiai eredetével együtt alátámasztja a közös indiai őstől való leszármazást.

V.2. Az R1a haplocsoport keletkezése, terjedése és alcsoportjai

Nemzetközi szinten elsőként vontuk le azt a következtetést, hogy az R1a-M458 és az R1a-Z280 alcsoportok (ősapák) a magyar nyelvű populációkra jellemzők, míg az R1a-Z93 alcsoport a malajziai indiai és a magyar roma populációkra. Publikációnk az egyik legjelentősebb tanulmány, amelynek sikerült különbséget tennie az ázsiai és európai eredetű R1a kromoszómák között. Az R1a alcsoportokra való korai differenciálódásának területét valahol az eurázsiai sztyeppéken, a Közel-Keleten vagy a Kaukázus térségében feltételeztük, mivel ezek a területek Dél-Ázsia és Kelet-Európa között fekszenek.

Az R1a-M198 az Y kromoszómális haplocsoportok közül az egyik legjelentősebb. 2012-ig számos erőfeszítés történt a nagyszámú alcsoportjának SNP-alapú azonosítására és azok migrációs útjának feltérképezésére. Az R1a kromoszómák őskori és későbbi történelmi eseményeinek mélyebb értelmezése az eurázsiai populációkban korábban nem volt lehetséges az R1a haplocsoporton belüli új downstream SNP-k hiánya miatt. A kutatók azon vitatkoztak, hogy vajon hol születhetett az R1a ősapája, indiai vagy közép-ázsiai eredetű-e. Egyesek a jelenlegi Ukrajna területén lokalizálták az R1a-M198 ősapája születését és feltételezték, hogy onnan a *kurgán kultúrával* és az indoeurópai nyelvek terjedésével együtt terjedt el mind Európában, mind kelet felé. 2010-ben jelent meg egy tanulmány (Underhill és mtsai, 2010), amelynek az R1a-

M198 haplocsoporton belül sikerült két alcsoportot (R1a-M458-at Kelet-Európában, R1a-M434-et kisebb gyakorisággal Pakisztánban) azonosítani. Kutatócsoportunknak 2012-ben 217 Európából és Ázsiából származó R1a mintát vizsgálva további 2 alcsoportot sikerült beazonosítani, nevezetesen az R1a-Z280 és R1a-Z93 alcsoportokat.

A vizsgálat legfontosabb megfigyelése az, hogy az új bináris markerek hatékony eszköznek mutatkoznak az R1a-M198 haplocsoporton belüli alcsoportba sorolásra, mivel az elemzett R1a minták több mint 98%-a a három vizsgált alcsoport egyikébe tartozott. Az R1a-M458 haplocsoport legmagasabb frekvenciája Kelet-Európára (Magyarországra és a romániai magyar etnikumokra) korlátozódik, de gyakorlatilag hiányzik Ázsiából, ami egybeesik Underhill és mtsai, 2010 eredményével. Belső- és Közép-Ázsia átfedési zóna lehetett az R1a-Z280 és az R1a-Z93 leszármazási vonalak tekintetében.

Később, 2014-ben Peter Underhill (Stanford Egyetem) irányításával nemzetközi kooperációban átfogó tanulmányt végeztek 16244 férfi vizsgálatával 126 eurázsiai populációból, ami megerősítette eredményeinket, valamint további alcsoportokat írt le. A kutatók az R1a ősapa születési helyét a Közel-Keleten, a mai Irán közelében lokalizálták, ahol a parahaplocsoportok (más szóval a „vad” típusok) változatlanul megmaradtak ősi formájukban.

V.3. Apai leszármazási vonalak a magyar és a manysi népesség összehasonlítása alapján

Nemzetközi szinten elsőként publikáltuk, hogy az N-Tat haplocsoport egyik alcsoportja – az N-L1034 – összeköti a magyar és manysi népességet, továbbá potenciálisan elválasztja őket más népektől. A mai magyar populációban nagyon alacsony az N-Tat haplocsoport aránya a többi finnugor nyelvet beszélő népességhez képest. Az R1a-Z280 haplocsoporton belül számos magyar-manysi haplotípus egyezés detektálható, akárcsak az N1c-L1034 alcsoportban. Ezek alátámasztják, de archeogenetikai adatok hiányában nem bizonyítják az R1a-Z280 ősapa jelenlétét a korai finnugor nyelvű népesség génállományában.

A nyelvészek mind a hagyományos, mind a modern osztályozási próbálkozások alapján a magyar nyelv legközelebbi rokonainak az obi-ugor nyelveket (manysi, hanti) tekintik. 2012-ben Vándor Anna három hónapot töltött Nyugat-Szibériában a manysikat kutatva és 60 északi manysi férfitől vett mintát apai ágon öröklődő genetikai markerek vizsgálatához. A mintákat megvizsgáltuk és az eredményeket finn, székely, magyar, csángó, mongol, burját és üzbecg populációkra vonatkozó adatokkal hasonlítottuk össze. A 60 manysi mintában 86,7% volt az Észak-Euráziára jellemző N haplocsoport aránya, ezen belül a többség az N-P43 (63,3%), a

kisebbség az N-Tat (23,3%) alcsoportba tartozott. Más apai vonalú haplocsoport alacsony gyakorisággal fordult csak elő, így 6,7% volt a nyugat-európai és dél-kaukázusi jellegű R1b-M269 (a kaukázusi kapcsolat valószínűsíthető), valamint 5% a fent jellemzett R1a-M198 aránya. Egyetlen minta dél-szibériai eredetű, az amerikai őslakosságra tipikusnak mondható Q-M242 csoportba tartozott. Vizsgálataink eredménye szerint az N-Tat haplocsoporton belül elkülöníthető egy alcsoport, az N-L1034, amely a manysi és magyar/székely N-Tat minták között jelen van, míg az ugyancsak vizsgált burját, mongol és finn populációkból hiányzik. A FamilyTree DNA (amerikai DNS laboratórium) szakemberei az N-L1034 alcsoportot a Kárpát-medencén belül a Csíki-medencében, a Felső-Tisza vidékén, Sárréten és az Őrség burgenlandi részén azonosították, olyan területeken, ahol történelmi okok miatt nagyobb arányban maradhattak meg a honfoglalók leszármazottjai. A fenti észrevételek alapján feltételeztük, hogy az N-L1034 alcsoport ugorkori örökség lehet a magyar és manysi génállományban. A másik érdekesség az, hogy az N-L1034 alcsoport aránya magasabb a mai csíkszeredai székely populációban, mint a magyarországi magyar populációban. Ez a tény a székelyek eredetéről szóló tudományos vitákat abba az irányba billentheti el, hogy a székelyek mindig is magyarok voltak. Számos kutató gondolta úgy idáig, hogy a székely egy, a magyar törzsszövetséghez csatlakozott, eredetileg vélhetően török etnikumú népesség volt. Ezeket az elképzeléseket azonban nagy valószínűséggel újra kell értékelni a genetikai kutatások tükrében. Bár a kondai (déli) manysi minták publikálása még folyamatban van, az megemlíthető, hogy a földrajzi okok miatt igen nehezen megközelíthető kondai manysik között az északi manysikhoz képest szignifikánsan magasabb a modellünk szerint az ugor és finnugor etnogenezisben meghatározó szerepet játszó N-L1034 és az R1a-Z280 haplocsoport aránya. Ez összhangban van elvárásunkkal, miszerint a nehezebben megközelíthető, *de facto* endogám területeken élők jobban megőrzik eredeti génállományukat. Magyar szemszögből nézve az R1a-Z280 haplocsoportban számos magyar-manysi Y-STR haplotípus egyezés található, akárcsak az N1c-L1034 alcsoportban.

V.4. Az mtDNS VIZSGÁLATI EREDMÉNYEK

Három tanulmány jelent meg a témában: a mitokondriális haplocsoportok gyakoriságán alapuló klaszterezési módszer kidolgozása, amelyet 174 eurázsiai populációs adaton teszteltünk, a szegedi egyetemen dolgozó kollégákkal közösen a honfoglaló temetőkből származó csontminták mitokondriális hipervariábilis régió szekvenciáján alapuló genetikai adatok prezentálása, valamint nemzetközi kooperációban hozzájárultunk 17 recens európai és közeli keleti populációból származó teljes mtDNS genom szekvencia adatok közzétételéhez.

V.4.1. 174 eurázsiai populáció mtDNS elemzése iteratív rangkorrelációs módszer segítségével

Egy teljesen új, iteratív rangkorrelációs algoritmus (IRC) segítségével a recens és az ősi mtDNS haplocsoportok eloszlásait, mint bizonyos ősi magpopulációk történelmi vándorlásának következményeit mutattuk ki. Az algoritmust alkalmazva 3 elkülönülő klasztert azonosítottunk jól interpretálható földrajzi régiókban (nyugati, keleti és szibériai), és megállapítottuk, hogy a vizsgált mtDNS haplocsoportok eloszlása jól értelmezhető e három klaszter különböző összetételeként.

Szimulációs eredményeink azt igazolták, hogy az ilyen migrációs folyamatok hatását az új iteratív rangkorrelációs algoritmus segítségével ki lehet mutatni, mivel az ilyen magpopulációk migrációja és a populációk jelenlegi haplocsoport összetétele a kezdeti populációkra jellemző haplocsoportok magas rangkorrelációját eredményezi. Az iteratív rangkorrelációs algoritmust alkalmazva három elkülönülő CHgC-t (Correlating Haplogroup Clusters: korreláló haplocsoport klaszter) azonosítottunk jól interpretálható földrajzi régiókban (nyugati, keleti és szibériai), és megállapítottuk, hogy a vizsgált mtDNS haplocsoportok eloszlása jól értelmezhető e három CHgC különböző összetételeként. Viszonylag nagyszámú ősi mtDNS minta bevonása az elemzésbe felbecsülhetetlen értékű eszközt jelentett az ókori népmozgások és a recens populációk élőhelyének összehasonlításában. Az *Andronovo*- és a *kurgán kultúrákhoz* köthető mtDNS adatok recens populációs eloszlásokkal való összehasonlítás alapján az a következtetés vonható le, hogy a recens kelet- és közép-európai populációk olyan összetett populációknak tekinthetők, amelyek a közel-keleti neolitikus forrásból származó nyugati szubsztrátum, továbbá egy *Andronovo* területéről eredő szibériai és a nyugati komponens keveredéséből jöttek létre. A recens és ősi mtDNS haplocsoport eloszlások szoros kapcsolata ezeken a területeken azt is valószínűsíti, hogy az uráli és altáji nyelveket beszélő nők jelentős szerepet játszhattak az *Andronovo*- és *kurgán-kultúrákban*. A módszert az ún. *stepping stone modell* segítségével validáltuk.

A módszert a honfoglalók genetikai összetételének elemzésére is alkalmaztuk, amely röviden a következő fejezetből ismerhető meg.

V.4.2. A karosi honfoglaló temetőből származó minták vizsgálata

Az új iteratív rangkorrelációs algoritmus segítségével, amellyel összehasonlítottuk a birtokunkban lévő mitokondriális DNS haplocsoport adatokat más ókori és recens

eurázsiai adatokkal, megállapítottuk, hogy a honfoglalók génállományának jelentős része valószínűleg egy régen konszolidált közép-ázsiai és dél-szibériai génállományból származhatott, és amely még mindig detektálható a recens magyar népességben. A honfoglaló magyarok másik genetikai rétegüket nyugatra vándorlásuk során különböző európai eredetű populációkkal való keveredésből szerezhették, ezek közül az egyik fontos elem a kaukázusi régióra mutat. Adataink azt mutatják, hogy a magyarok nemcsak a kontinensen már jelenlévő anyai haplocsoportokat birtokolták, hanem új ázsiai génekkel is frissítették az európai génállományt. Azok a recens populációk, amelyek genetikailag legközelebb állnak a honfoglaló magyarokhoz, ma indoeurópai nyelveket beszélnek.

A tanulmány a Karos III. temetőből származó 19 csontminta közül 17-ből sikeresen meghatározott anyai ági haplotípusokat mutatja be. A mitokondriális DNS vizsgálat eredménye a szegedi egyetemen dolgozó kollégáknak köszönhető.

A leggyakoribb haplocsoport a B volt, amely az A haplocsoporttal együtt a karosi mintapopuláció mintegy 30% -át tette ki. Ez a két csoport jelentett genetikai kapcsolatot Közép- és Kelet-Ázsiával. A fő haplocsoportok (H, U, T, J, X) nagy része eurázsiai eredetű; figyelemre méltó azonban, hogy két személy a H6 alcsoportba tartozott, amely szintén utalhat ázsiai kapcsolatra. A H6 a Közel-Keleten vagy a Kaukázusban keletkezett és ma a leggyakoribb haplocsoport Közép- és Belső-Ázsiában (21%), különösen az altájiakban (35%), Európában csak a bronzkorban jelent meg. A karosi honfoglalók anyai ági genetikai összetétele több forrásból származó kevert eredetet mutat. A mintacsoportban kimutatott haplotípusok mintegy 30-40% -a egyértelműen Kelet- és Közép-Ázsiából származik, legalábbis szélesebb időskálán belül. A B haplocsoport Median Joining hálózatainak adatai azt jelzik, hogy ez a komponens származhatott a mai altáji (török) nyelvet beszélő etnikai csoportok egy közös génállományából. Annak ellenére, hogy az ázsiai haplocsoportok jelentős hányada kimutatható, a honfoglalók elsősorban európai antropológiai jellegekkel rendelkeztek, ez a Karos III temetőre is igaz. Ez arra utalhat, hogy a honfoglalók generációkkal a honfoglalás előtt keveredhettek ázsiai népekkel a kelet-európai Pontuszi-sztyeppén. Az IRC és a SOC (Self Organising Cloud) módszerekkel történt, honfoglalókra vonatkozó elemzések azt mutatták, hogy a *Szintasta-Andronovo-* (SIA; i.e. 2100-1500) és a *késő-Baraba-* (BB3; i.e. 1000-800) kultúrák mintái a karosi mintákhoz álltak legközelebb, amelyek a nyugati és a szibériai génáramlás metszéspontjába estek. A SIA és a BB3 populációk földrajzi helyzete a történelmi várakozásoknak megfelel, ők már rendelkeztek a nyugati és szibériai CHG-k kiegyensúlyozott összetételével. Ez arra enged következtetni, hogy a honfoglalók génállományának alapkomponeense, amely a közép-ázsiai és dél-szibériai komponensek keveredéséből származik, több mint 2000 évvel a honfoglalás előtt alakulhatott ki

a SIA (*Sintashta-Andronovo*) és a BB3 (*Baraba*) horizont körül. A SOC elemzés, amelybe 111 recens és 35 ősi populáció mtDNS haplocsoport adatait vontunk be, 7 klasztert eredményezett. A 2. klaszter jelentős arányú nyugati és szignifikáns arányú szibériai CHgC-t tartalmazott. A karosi minták a 2. klaszterbe kerültek, melyet a nyugati CHgC-k dominanciája jellemez, de emellett jelentős arányú szibériai CHgC-t is tartalmaz. A 2. klaszteren belül azonban a karosi minták a szibériai komponensek magas aránya miatt viszonylag izoláltak, és genetikai távolságok alapján megközelítik a 3. klasztert, ahol a nyugati és szibériai komponensek aránya a leginkább kiegyensúlyozott. A 2. klaszter nyugati elemei sokkal hasonlóbbak egymáshoz, mint az 1. klaszter komponenseihez (a nyugati populációk), ami a klaszterek tagjai közötti genetikai kapcsolatokra utal. A karosi minták legközelebbi kapcsolata, mint azt a SOC elemzések kimutatták, a 2. és 3. klaszterek populációi között található.

Említésre méltó megfigyelés, hogy a 2. klaszter egy kivétellel az összes ősi és recens magyar populációt tartalmazza: AH1 (Ancient Hungarian 1), AH2 (Ancient Hungarian 2, Kr. u. 900), SEK (székely) és CSG (csángó), kivéve a HUB-t (Hungarian Budapest). Ez indokolt, mivel az ázsiai eredetű mitokondriális haplocsoportok (A, B, C és D) előfordulási gyakorisága 1% alatti a recens magyar mintákban és 1% fölötti a székely és csángó mintákban, ezenkívül az utóbbi két mintacsoport viszonylag magas arányban – 5-7,4%-ban – közép-ázsiai és belső-ázsiai apai haplocsoportokat is tartalmaz, ami megkülönbözteti őket a nyugat-európai populációktól. Az is figyelemre méltó, hogy a 35 ősi populáció közül 14 a 2. klaszterbe csoportosult és ez a szám 24-re nőtt az egyesített 2. és 3. klaszter populációiban, amely a Kárpát-medence majdnem összes ősi populációját tartalmazta. Ennek a váratlan eredménynek az esetleges magyarázata abban a közel-keleti neolitikus populációban keresendő, ami az MDS ábrán nem messze helyezkedik el a karosi temető mintáitól (ez azt is jelentheti, hogy mindezen populációk jelentős genetikai behatást kaphattak a korai neolitikumban a Termékeny Félhold területéről).

V.4.3. Az teljes mtDNS genom vizsgálata európai és közel-keleti populációkban

A demográfiai történéseket az mtDNS-szekvenciákon alapuló Bayesian Skyline Plots (BSP) alkalmazásával először populációkra, nem pedig leszármazási vonalakra, vonatkozóan rekonstruáltuk. Összehasonlítható adatként már rendelkezésre álltak ugyanezen populációk Y-kromoszómális (MSY régió) szekvenciái, melyet szintén a kutatócsoportunk tett közzé. A populációkra jellemző teljes mtDNS genom vizsgálatok hiányoztak Európában, így a férfiak és a nők demográfiai történeteinek szisztematikus összehasonlítására még nem adódott lehetőség. Az mtDNS-ben megfigyelhető mintázat

meglepően különbözik az MSY-ban megfigyeltékhez képest, ami összeegyeztethető az utolsó jégkorszak utáni terjeszkedéssel, és megerősíti a humán populációk bronzkori terjeszkedésének férfi-specifikus természetét Európában. Az mtDNS vonalak expanziói ugyanis a legtöbb európai populációban sokkal ősbbeek voltak, úgy 13 és 20 ezer évvel ezelőtti paleolitikus expanzió nyomait viselik.

A harmadik mtDNS tanulmányban nemzetközi kooperáció keretében egy populáció alapú vizsgálatban működünk közre, amely 17 populációból származó – populációként 20-20 – összesen 340 európai és közel-keleti egyén teljes mtDNS genomjának újra szekvenálására irányult.

Összehasonlítható adatként már rendelkezésre álltak ugyanezen populációk Y-kromoszómális (MSY régió) szekvenciái, melyet szintén a kutatócsoportunk tett közzé. Az mtDNS-ben megfigyelhető mintázat meglepően különbözik az MSY-ban megfigyeltékhez képest, ami összeegyeztethető az utolsó jégkorszak utáni terjeszkedéssel, és megerősíti a humán populációk bronzkori terjeszkedésének férfi-specifikus természetét Európában. A teljes adatkészletben a fő haplocsoportok a következő eloszlásban fordulnak elő: H (34,1%), U (17,9%), T (13,5%), J (9,1%), K (7,3%) és V (5,3%). A fennmaradó 12,6%-on sok kisebb haplocsoport osztozik, ezek közé tartoznak pl. az L2 és L3 vonalak (a palesztin és a spanyol mintában), amelyek jellemzően a Szaharától délre eső afrikai területen fordulnak elő.

A mitokondriális genomok (kódoló régió) filogenetikai analízisét összehasonlítottuk az MSY analízisével. A vizsgálatokhoz maximális parszimónia (MP) módszert alkalmaztunk. A haplocsoportok földrajzi eloszlása az MSY esetében sokkal inkább korlátozott, mint az mtDNS esetében: például az apai vonalak közül az R1b különösen nagy gyakorisággal van jelen északnyugaton, az R1a Észak-és Közép Európában, a J2 pedig Dél-Európában. A két maximum parszimónia (MP) fának nagyon eltérő a skálája, mivel az mtDNS-ben sokkal kisebb számú nukleotidot elemeztünk, mint az MSY-ban (1: 250 arány), mindkettő deep-rooting (mélyen gyökerező) kládokat (mtDNS: U, K, T2 haplocsoportok; MSY: E1b-M35, G2a-L31, I2-P215, J2-M172, L, T haplocsoportok), valamint csillagszerű mintázatokat mutatott, amely a populációk expanzióját jelenti. Az mtDNS esetében ezek a haplocsoportok a mintánk 51,7% -át reprezentáló H, J1, T1, V, míg az MSY esetében az I1-M253, N1c-M178, R1a-M198 és az R1b-M269 (együtt 64%). A fő különbség a belső ágak relatív hosszúsága, ami azt mutatja, hogy az mtDNS vonalak (anyai) expanziói sokkal ősbbeek, mint az MSY vonalaké (apai). Ezt támasztja alá a csillagszerű struktúrát mutató haplocsoportok legközelebbi közös őstől (TMRCA) való leválási idejének becslése is, amelyek mindegyike kevesebb 6 ezer évnél az MSY esetében (posztneolitikus), míg az mtDNS-esetében legalább 13 ezer év (paleolitikus). A török és a palesztin minták eltérnek a

többségtől a populáció sokkal régebbre, úgy 40 ezer évvel ezelőttre, tehető expanziója miatt. Az eredmények összhangban vannak a korábban megfigyelt kontinensszintű különbségekkel, amely a populációk nagy részére igaz, ugyanakkor rámutatnak a földrajzi mintázat hiányára is Európán belül. A legősibb közel-keleti mtDNS expanziók a korai eurázsiai benépesítéshez közelebbi időszakra datálhatók (40-50 ezer év). Ezek a mintázatok ugyanezen populációcsoporton belül éles ellentétben állnak az MSY BSP mintázataival, amely a legtöbb populáció esetében (13 a 17-ből) azok demográfiai múltját prezentálja. A populáció legkisebb mérete kb. 3 ezer évvel ezelőttre tehető (késői bronzkor), ezt követte a jelenlegi gyors expanzió. Minden egyes populációra diverzitási indexet is számoltuk. A számi mutatja a legalacsonyabb értéket, míg a legmagasabb értékeket a palesztin és a török minták adták, ami ismét összhangban van a BSP-mintázatokban megfigyelt ősi népességnövekedéssel. A 17 populációs adatunkban szereplő számi/lapp kilógó státusza egyértelmű, mivel nemcsak az egymással szorosan összefüggő mtDNS-szekvenciák száma magas köztük, hanem a máshol ritkán előforduló mtDNS és MSY haplocsoportoké is, illetve több egyénben is detektálhatók azonos MSY szekvenciák és Y-STR haplotípusok. Ezek a jellemzők összhangban vannak a populációméret tényleges növekedésének BSP-ben megfigyelt hiányával. A populációk genomszintű SNP analízise azt is megmutatja, hogy a számi populáció genetikailag differenciálódott az európaiakhoz képest, valamint kelet-ázsiai eredetű genetikai komponenseket hordoz. Adataink összhangban vannak az aDNS-adatokkal is, megerősítve az uniparentális genetikai folyamatok szerepét a mai európai demográfiai összetételében, miszerint a recens mtDNS diverzitás mintázatai nem utalnak vissza a bronzkori expanzióra, miközben a recens európai MSY diverzitásának nagy részét a bronzkori expanzió alakította át.

V.4.4. Zenébe vésett gének, avagy a zenegenetika (music-genetics)

*Kodály Zoltán: Pár hangnyi dallamok, mintha
kőbe vésvé állták volna századok viharát.*

Nemzetközi szinten elsőként a genetika, az adatbányászat, a mesterséges intelligencia kutatás és a muzikológia jelenlegi módszereit felhasználva 31 kultúra zenei kapcsolódási rendszerét vizsgáltuk Euráziában, és igazoltuk, hogy ezek a kapcsolatok a vizsgált kultúrák 82%-ában genetikai alapokra vezethetők vissza, míg a populációk szoros genetikai kapcsolatai nem feltétlenül megfelelő indikátorai a hasonló zenei kultúráknak (csupán 28%-ban). Továbbá feltártuk, hogy a közös zenei örökséggel rendelkező kultúrák többsége genetikai értelemben is szorosan összefügg, és ezzel megalapoztunk egy új interdiszciplináris kutatási területet, a "zenegenetikát".

Juhász Zoltánnal (KFKI, mesterséges intelligencia-kutató) közösen sejtettük meg a népzeneben rejlő „öröklődő géneket”. A legmegdöbbentőbb felismerés az volt, hogy az egymástól távoli két szakmában széles körben használt adat prezentálási módszer ugyanazon az elven működött, nevezetesen a Juhász Zoltán által zenei elemzésekre kifejlesztett módszer az *ún. nem-metrikus MDS* elvére támaszkodott, ami a genetikában is használatos. A kételkedések feloldására – ami igazságügyi szakértő mivoltamból fakad – olyan genetikai adatokkal (Fst genetikai távolságok) teszteltük a zenei MDS-t (SOM=Self Organizing Map), amelyekre már megvoltak a genetikában általánosan használt MDS módszerrel készült ábrák. Nagy örömmre zenei MDS módszerrel ugyanolyan populációs klasztereket kaptunk, mint a nem-metrikus MDS módszerrel, amely szaknyelven a módszer validálása. A különböző kultúrákból származó hagyományos népzenei dallamvilág számítógéppel való elemzése és összehasonlítása az egyik legújabb kutatási terület. Míg korábban jelentős számú genetikai vizsgálat történt a humán migráció történetével kapcsolatban, addig a népzene/népi dallamkincs és a genetika közötti kapcsolatokat soha nem vizsgálták, ezért fordítottuk figyelmünket a genetikai adatok és a népzenei dallamok közötti korrelációk vizsgálatára. A kifejlesztett szoftverrendszer a különböző népdalok közötti hasonlóságot jellemzi skálázható mértékegységgel. Ennek a rendszernek az alapötlete az volt, hogy a dallamkincs gyarapodik és fejlődik a melódiák végtelen variációja által. A variáció a legtöbb orális kultúra alapvető jellemzője, amely analóg a genetikai diverzitással. Eurázsia különböző nemzeteinek összehasonlító zenei elemzése során kirajzolódik egy jól interpretálható hálózat, a zenekultúrák "genetikai" kapcsolatainak hálózata. Ez arra a kérdésre vezetett, hogy ezek a kapcsolatok pusztán a véletlen művei-e, vagy esetleg a régmúlt idők népességeinek együtt élésével, illetőleg migrációjával hozhatók-e kapcsolatba? A zenei kapcsolatok bonyolult szerkezete mutat-e világos képet és tudjuk-e interpretálni ezt a kapcsolati rendszert genetikai adatokkal összevetve? A tanulmány az említett szoftver segítségével készült, 31 eurázsiai és észak-amerikai nép népdal kollekciójának kereszt-kulturális elemzését ismerteti, valamint ezzel párhuzamosan a népzenei csoportoknak megfelelő populációk közötti genetikai kapcsolatok feltárására is irányult. A zenei adatbázis azt a 31 kultúrát öleli fel, amelyek mindegyikéből 1000-2500 dallam digitalizált kottája állt rendelkezésre. A vizsgált kultúrák a következők: kínai, mongol, kirgiz, volgai régió (mari-csuvas-tatár-votják), szicíliai, bolgár, azeri, anatóliai, karacsáj, magyar, szlovák, morva, román, finn, norvég, német, luxemburgi, francia, holland, ír-skót-angol (egy csoport), spanyol, dakota, komi, hanti, horvát-szerb (Balkán csoport), kurd, orosz (a pszkovi területről), navahó és lengyel (két kisebb régió: Warmia és Kasubia). A zenei adatbázis vizsgálatával párhuzamosan, elkészítettük a megfelelő genetikai távolságmátrixokat is a széles körben használatos szoftvereket alkalmazva. A népzene

elemzésére kidolgoztunk egy rendszert, amely a mesterséges intelligencia egy általánosan alkalmazott típusán alapul a dallamkontúrok automatikus osztályozására - az úgynevezett önszerveződő térképet (SOM=Self Organization Map). Az első lépés a D-dimenziós dallamvonal-vektorok (D-elemű hangmagasság-idősor) szerkesztése a kották digitális kódjaiból. Mivel D mindegyik dallamra ugyanaz volt, a kontúrok összehasonlíthatók egymással egy közös euklideszi távolságfüggvény segítségével ugyanabban a D-dimenziós dallamtérben, függetlenül azok tényleges egyéni hosszától. A dallamvonal-vektorokban, a dallamok fő ritmikai jellemzői szintén leképződnek. Ugyanakkor a tempó is teljesen normalizálható ezzel a technikával, ez pedig a lényeg kiemelésének bevett, hasznos és általános eszköze a népzene tudományban. A különböző lejegyzési elvekből eredő problémák elkerülése érdekében a gyűjtemény összes dallama G alaphangra transzponálható. A legtöbb egyszólamú népdalkultúrában a dallamoknak van egy jól definiált alaphangja (tonika), mely a záróhanggal azonos; ezért ez a technika zeneileg releváns a leggyakrabban vizsgált esetekben. Ugyanakkor ez általánosságban nem igaz, főleg a nyugati népdalok esetében, ezért ezeknek az adatbázisoknak automatikus átvétele részben szakértők által, részben további algoritmusok által korrigálható. A betanítás után a SOM egy adott kultúra dallamainak kontúr vektoraival, a térkép különböző rácshelyeihez tartozó megtanult dallamvonal-típus-vektorokkal (DVT-k) képezi le az adott kultúrát jellemző legfontosabb kontúrokat. Más szóval, egy jól betanított SOM DVT-készlete egy olyan zenei nyelvet képvisel, amely optimális a saját kultúrájának "megértéséhez". Az idegi hálózatok terminológiájával élve, a DVT vektorok „receptoroknak” tekinthetők, amelyek „tüzelnek” vagy „aktiválódnak”, ha egy dallam kontúrja elég hasonlóknak tűnik hozzájuk. Két különböző „zenei nyelv” összehasonlításakor az egyik nyelv DVT-it próbáljuk aktiválni a másik nyelv DVT-i által. A közös zenei jellemzők az „idegen nyelv” DVT vektorjai által aktiválódnak, ugyanakkor az egyes kultúrák sajátosságai szintén leírhatók, azon DVT-k halmazaként, amelyeknek nincs kapcsolatuk a másik kultúrával.

A férfiak által közvetített genetikai örökség feltárására az Y-kromoszómális haplocsoportok gyakoriságai alapján számolt genetikai távolságokat (F_{st}) hasonlítottuk össze 42 populáció esetében és azokat nem-metrikus MDS térképen jelenítettük meg. Hogy a fenti megállapításokat a nők által közvetített genetikai örökség kontextusába helyezzük, hasonlóan a férfiakéhoz, kiszámítottuk az mtDNS haplocsoportok genetikai távolságait 56 populáció publikált forrásokból származó adatait használva és az eredményeket ismét MDS diagramként mutattuk be. A populációk klaszterbe sorolása azonos elveken alapult, tehát a férfiak és nők esetében is a közeli rokonnak tekinthető populációk genetikai távolságai 0,05-nél kisebbeknek kellett lenniük. A fentiekben ismertetett SOM-alapú analízis által feltárt zenei kapcsolatokat

szintén egy zenei MDS ábrán mutattuk be, ahol a 31 kultúrát reprezentáló csomópontok elhelyezkedését az MDS elve alapján határozzuk meg. Igyekeztünk arra törekedni, hogy azonosítsuk, a zenei kultúrák mely csoportjai hordoznak jelentős mennyiségű közös dallamot. Zenei szempontból a komplementer megközelítés az, hogy találjuk meg a 31 kultúra egy közös részalmazában egyidejűleg előforduló melódiák nagy csoportjait. E feladat véghezviteléhez meghatároztuk a melódia adatbázisunkban megjelenő fontos DVT-k összességét úgy, hogy egy nagyméretű SOM-t képeztünk, amelyet a korábban megtanult 31 nemzeti/területi DVT-k egyesített halmazával tanítottunk. Ennek a nagy "egyesített" DVT-készletnek a birtokában a fenti kérdés már algoritmikus keresésként fogalmazható meg a következőképpen: a 31 melódia-gyűjtemény (corpora) legnagyobb alcsoportjait keressük, amelyeknek a legtöbb közös eleme van az egyesített DVT gyűjteményben. Az eredményeket egy zenei térképen mutattuk be, ahol az egyesített DVT vektorok a sík (plane) pontjaihoz vannak hozzárendelve és az eredményként kapott pontrendszer egy MDS algoritmussal rendeztük el. Így hasonló DVT-k egymáshoz közelebbi helyekre kerülnek a síkon és a pontrendszer teljes elrendezése átfogóan tükrözi az egyesített DVT-halmaz hasonlósági viszonyait. Amikor a 31 kultúra azon 4-elemű részalmazait kerestük, amelyek a legnagyobb közös DVT-készlettel rendelkeznek, a volga-török-karacsáj-magyar, ill. a kínai-török-karacsáj-magyar csoportokra bukkantunk (46, illetve 45 teljesen közös fajta). Az ilyen közös típusok a dallamtérkép jól körülhatárolt területén helyezkednek el, következésképpen egy jól definiált zenei stílushoz tartozó folyamatos variációképződés eredményei. Összehasonlításként négy nyugati kultúra közös dallamtípusait is bemutattuk, ahol ezek a típusok az egyesített felhő teljesen más területein helyezkednek el, így teljesen más zenei formákat jelentenek. A nyugati típusok maximális közös eleme azt mutatja, hogy a nyugat-európai zenei mag sokkal kisebb, mint a fent említett csoportokban. A legtöbb közös DVT-t tartalmazó kultúrák héttagú csoportjainak keresése során ugyancsak a 4-elemű keresés alapkulturái találhatóak meg, kiegészülve a szicíliai, finn és dakota, ill. a mongol, volga-vidéki és szicíliai kultúrákkal. A vizsgált populációk közötti genetikai kapcsolatok feltárásához Fst genetikai távolságokat alkalmaztunk és a távolságmátrix általános szerkezetét MDS diagramokon ábrázoltuk. Mind az Y-kromoszómális, mind az mtDNS haplocsoportok eloszlására vonatkozó MDS ábrákat a genetikai adatok függvényében közel azonos számú populációból állítottuk elő. A vizsgált populációk mindkét MDS diagramon hat és hét különálló klasztert képeztek. Az egy klaszteren belül elhelyezkedő populációk többségének genetikai közelsége magyarázható azok földrajzi közelségével és/vagy közös történelmi múltjával. Általánosságban azt mondhatjuk, hogy két populáció zenei és genetikai jellemzői közötti kapcsolat egy korábbi fizikai és biológiai kapcsolatot jelez, ezért a zenei kapcsolat

visszavezethető az adott népcsoportok őseinek feltételezett kulturális kölcsönhatására. Ezért az egy zenei kultúrához tartozó populációpár közötti szoros genetikai összefüggését úgy határozzuk meg, hogy a megfelelő populációk között legalább az egyik genetikai távolságnak (mtDNS vagy Y-SNP alapú) kevesebbnek kell lennie 0,05-nél, ami meglehetősen szigorú követelmény, tekintve, hogy a genetikai távolság definíció szerint 0- és 1 közötti értéket vehet fel. Ez az Y-kromoszómális és a mitokondriális adatoknak csak 37% -ára teljesül. Összeszámoltuk azon esetek számát, ahol szoros kapcsolatot észleltünk: 1. genetikai (G); 2. zenei (M); 3. mind a zenei, mind a genetikai (X) távolságok tekintetében. Ennek alapján azon események feltételes valószínűségei, hogy

1. a szoros genetikai kapcsolat szoros zenei kapcsolatot is jelent $p(m|g)=X/G=0,28$,
2. valamint a szoros zenei kapcsolat szoros genetikai kapcsolatot is jelent $p(g|m)=X/M=0,82$.

Az első eset alacsony feltételes valószínűsége szerint a populációk szoros genetikai kapcsolata nem feltétlenül megfelelő mutató a zenei kultúrák hasonlósága szempontjából. Azonban a második esetben számolt feltételes valószínűség magas értéke azt igazolja, hogy a népzenei kultúrák jelentős hasonlósága szignifikáns genetikai kapcsolatokra utal a vizsgált esetek 82%-ában. Másképpen kifejezve, a zenei kapcsolatok fennállásából előre jósolhatók a genetikai kapcsolatok, kb. 82% -os valószínűséggel. A zenei és a genetikai kapcsolatok korrelációját a genetikai távolság küszöbértékének függvényében vizsgáltuk és bebizonyítottuk, hogy szoros zenei-genetikai kapcsolatként definiált genetikai távolság küszöbérték, melyet mi történetesen 0,05-ként határoztunk meg, valójában független a küszöbérték választásától. Továbbá megállapítottuk, hogy szoros genetikai kapcsolatok alapján nem lehet megjósolni szoros zenei kapcsolatokat, és ez szintén független a genetikai távolság küszöbértékétől. A zenei elemzés megmutathatja azt, hogy kölcsönösen intenzív érintkezés van kultúrákon belül, de a nagy földrajzi távolságok kérdéssé tehetik ezeket az eredményeket a genetikai adatokkal való összehasonlítás előtt. Az mtDNS adatok alapján számolt genetikai távolságok azonban összefüggést mutatnak a zenei eredményekkel: 0,0017, 0,024 és 0,024 a norvég-appalache (skót), norvég-magyar és magyar-appalache kapcsolatokban. Ezen adatok tükrében az előbbi népek közötti hasonló dallamok többé már nem tekinthetők „véletlenül hasonlóképpen komponáltak”, azaz három egymástól függetlenül fejlődő kultúra "véletlen egybeesésének". Érdeemes megemlíteni, hogy maga a zenei elemzés kizárja ezt a feltételezést, mert a három kultúrában rendkívül sok egymást átfedő dallamtípusra utal. Ezen túlmenően a példák azt is mutatják, hogy ezek az átfedések egyaránt jelentős mennyiségű, mindhárom kultúrában egyszerre jelentkező, közös zenei típust fednek le. A fenti valószínűségi analízis elvégzését a

mitokondriális és az Y-kromoszómális haplocsoportok alapján számolt genetikai távolságok tekintetében külön-külön is elvégeztük: $p(m|g)=0,27$ és $p(g|m)=0,77$ a mitokondriális, valamint $p(m|g)=0,36$ és $p(g|m)=0,35$ az Y-kromoszómális adatok esetében. Ezek az eredmények azt mutatják, hogy a szoros zenei kapcsolatok dominánsan női genetikai kapcsolatokat jeleznek. Lehet-e ezt a megfigyelést azzal magyarázni, hogy a menyasszonyok mitokondriális genetikai információja a vőlegényeik népeiben is megjelenik, vagy azzal, hogy az anyák átadják zenei anyanyelvüket gyermekeiknek, akár idegen környezetben is? Mindent összevetve, a nők sokkal fontosabb szereplők a zenei és genetikai információk átadásában, mint a férfiak, és evolúciós történetük során nagyobb mobilitással rendelkeztek. Az egyesített DVT-k egyidejű létezése az appalache-norvég-magyar változatokban további kérdéseket vet fel: Vannak-e olyan közös zenei stílusok, amelyek különböző kultúrákban egyidejűleg léteznek és ezek az adatok összefügghetnek-e párhuzamos genetikai kapcsolatokkal? Az ezekre a kérdésekre adott pozitív válasz a zenei kultúrák közös kristályosodási pontjait tárná fel és igazolná, hogy ezek a kapcsolatok történelmi vagy történelem előtti okokra visszavezethetők. Ezekre az izgalmas kérdésekre pozitív válasz adható, és kimutatható, hogy a teljesen közös melódiák legnagyobb részhalmaza a volgai-török-karacsáj-magyar és a kínai-török-karacsáj-magyar csoportok kvartettjeiben található, 46 és 45 teljesen közös típusként. Ezek a típusok a közös zenei térkép jól körülhatárolt részében helyezkednek el, ezért azt állíthatjuk, hogy ezek egy összetartozó zenei stílust képviselnek, például a legtöbbjüknek ereszkedő kontúrja van egy oktáv hangterjedelemmel. A 31 kultúra nagyobb részhalmaza közötti szisztematikus kereséssel az mutatható ki, hogy a legnagyobb átfedésekkel rendelkező csoportok a fenti kvartettek megfelelő kiterjesztései: a volgai-szicíliai-török-karacsáj-magyar-finn-dakota és a kínai-mongol-volgai-szicíliai-török-karacsáj-magyar kultúrák 7 elemű csoportjaiban a legelterjedtebb dallamtípusok száma 12 és 10 között van. A dakota és karacsáj genetikai adatok hiánya mellett is az előbbi csoport szoros zenei kapcsolatai teljesen korrelálnak az ismert genetikai adatokkal. A fent említett zenei kultúrák közös területe a zenei térképen teljesen elkülönül a német-(luxemburgi-lotaringiai)-francia-holland kvartettétől, ahol a genetikai távolságok - amelyek bármely pár esetében 0,03-nál kisebbek - szintén nagyon intenzív genetikai kapcsolatokat jeleznek. Azt a tényt, hogy egy teljesen elkülönülő terület detektálható a közös "nyugati" és "keleti" dallamformákhoz, két alapvetően különböző zenei nyelvcsoporthoz létezésével lehet értelmezni. A nyugati kultúrákban a teljesen közös típusok viszonylag alacsony száma (17) azt mutatja, hogy azok párosított nagy átfedései főként különböző típusokat tartalmaznak; ezért azok sokkal kisebb kiterjedésű magstílust alkotnak, mint a keleti csoport. Ezek a melódiák, amelyek két, az oktávból fokozatosan leereszkedő, nagy ambitusú DVT változatai, egy jól definiált közös stílust mutatnak

a két "keleti" csoport egyik tagjaként besorolt kultúrákban. Alig lehet elképzelni másfajta magyarázatot az ilyen közös zenei stílusra, mint egy közös "zenei ősnyelv" létezését, azaz többé-kevésbé független zenei változások kezdeti kristályosodási pontját. Ez egy hipotézis, amelynek bizonyításához további történelmi, régészeti stb. kutatásra van szükség, de azt állíthatjuk, hogy az itt tárgyalt genetikai korreláció ezt a hipotézist alátámasztja. A kapott eredmény könnyen elfogadható a közelmúltban ugyanazon a területen élők esetében, mint a nyugat-európai országok, szlovákok és magyarok, finnek és norvégok stb. Eredményeink jelentős része ugyanakkor szimultán létező zenei és genetikai hasonlóságot mutat olyan vizsgált populációk között, amelyek nagyon régóta nagyon távol élnek egymástól. Például a magyarok és a norvégok, a tatárok és a szicíliaiak, a szicíliaiak és a törökök stb. egyidejű zenei és genetikai kapcsolatai kihangsúlyozhatják, hogy a zenei kapcsolatok a nagyon régi időkből származhatnak, így a szóbeli zenei hagyományok őskori zenei formákat és stílusokat őrizhetnek. A genetika és a számítástechnika ill. muzikológia együttműködése létfontosságú volt ennek a "rejtett" kapcsolatnak a feltárásához. Megállapítottuk a legszélesebbközös ősi zenei stílust, amely számos különböző kultúrában egyszerre jelenik meg, hasonló dallamformák sokaságát megalapozva. Továbbá feltártuk, hogy a közös zenei örökséggel rendelkező kultúrák többsége genetikai értelemben is szorosan összefügg és ezzel megalapoztunk egy új interdiszciplináris kutatási területet, a "zenegenetikát". Az eredmények arra a következtetésre vezettek, hogy a történelmi zenei stílusok egy történelem előtti eredetű közös zenei „anyanyelvből” származnak. A tanulmányunk megjelenése óta (2012) több hasonló cikk és PhD disszertáció is született, amelyek megerősítették az általunk megfigyelt zenei-genetikai kapcsolatot, hogy a zenei hasonlóságok vertikálisan öröklődnek, nem pedig horizontálisan. Emellett a humán vizagálatokkal analóg módon a kutatók az állatvilágban (az állatok és madarak hangja/dala), valamint az emberek és állatok kreativitásával kapcsolatban is hasonló kutatásba fogtak. Reméljük, hogy egy újonnan keletkező interdiszciplináris tudomány területnek a magjait ültettük el, amik idővel ki is hajtanak...

VI. KÖZLEMÉNYEK JEGYZÉKE

VI.1. A doktori értekezésben tárgyalt közlemények (IF: 87.194)

1. Huang YZ, **Pamjav H**, Flegontov P, Stenzl V, Wen SQ, Tong XZ, Wang CC, Wang LX, Wei LH, Gao JY, Jin L, Li H: Dispersals of the Siberian Y-chromosome haplogroup Q in Eurasia. *Mol Genet Genomics*. 293:(1) pp.107-117. (2018). (IF: 2.979).
2. Batini C, Hallast P, Vågene AJ, Zadik D, Erikse HA, **Pamjav H**, Sajantila A, Wetton JH & Jobling MA. Population resequencing of European mitochondrial genomes highlights sex-bias in Bronze Age demographic expansions. *Sci Rep*. 7:(1) Paper12086. 8 p. (2017). (IF: 4.259).
3. **Pamjav H**, Fóthi Á, Fehér T, Fóthi E. A study of the Bodroglakó population in north-eastern Hungary by Y chromosomal haplotypes and haplogroups. *Mol Genet Genomics*. 292:(4) pp. 883-894. (2017) (IF: 2.979).
4. Neparáczki E, Juhász Z, **Pamjav H**, Fehér T, Csányi B, Zink A, Maixner F, Pálfi G, Molnár E, Pap I, Kustár Á, Révész L, Raskó I, Török T. Genetic structure of the early Hungarian conquerors inferred from mtDNA haplotypes and Y-chromosome haplogroups in a small cemetery. *Mol Genet Genomics*. 292:(1) pp. 201-214. (2017). (IF: 2.979).
5. Juhász Z, Fehér T, Németh E and **Pamjav H**. mtDNA analysis of 174 Eurasian populations using a new iterative rank-correlation method. *Mol Genet Genomics*. 291:(1) pp. 493-509. (2016). (IF: 2.979).
6. Batini C, Hallast P, Zadik D, **Pamjav H**, Sajantila A et al. Large-scale recent expansion of European patrilineages shown by population resequencing: *Nature Communication* 6: Paper 7152. 8 p. (2015) (IF: 11.329)
7. Hallast P, Batini C, Zadik D, Maisano Delser P, Wetton JH, Arroyo-Pardo E, Cavalleri GL, de Knijff P, Destro Bisol G, Dupuy BM, Eriksen HA, Jorde LB, King TE, Larmuseau MH, López de Munain A, López-Parra AM, Loutradis A, Milasin J, Novelletto A, **Pamjav H**, Sajantila A, Schempp W, Sears M, Tolun A, Tyler-Smith C, Van Geystelen A, Watkins S, Winney B, Jobling MA. The Y-Chromosome Tree Bursts into Leaf: 13,000 High-Confidence SNPs Covering the Majority of Known Clades. *Mol Biol Evol*. 32:(3) pp. 661-73. (2015). (IF: 13.649)
8. Szabolcsi Z, Farkas Z, Borbély A, Bárány G, Varga D, Heinrich A, Völgyi A, **Pamjav H**. Statistical and population genetics issues of two Hungarian datasets from the aspect of DNA evidence interpretation. *Forensic Sci Int Genet*. 19: pp. 18-21. (2015). (IF: 4.988).
9. Fehér T, Németh E, Vándor A, Kornienko IV, Csáji LK and **Pamjav H**. Y-SNP L1034: limited genetic link between Mansi and Hungarian-speaking populations. *Mol Genet Genomics*. 290:(1) pp. 377-386. (2015). (IF: 2.622)
10. Bíró A, Fehér T, Bárány G and **Pamjav H**. Testing Central and Inner Asian admixture among contemporary Hungarians. *Forensic Sci Int Genet*. 15: pp. 121-126. (2015). (IF: 4.988).
11. **Pamjav H**, Fehér T, Németh E, Pádár Z. Brief communication: New Y-chromosome binary markers improve phylogenetic resolution within haplogroup R1a1. *Am J Phys Anthropol*. 149:(4) pp. 611-615. (2012). (IF: 2.481)
12. **Pamjav H**, Juhász Z, Zalán A, Németh E, Damdin B. A comparative phylogenetic study of genetics and folk music: *Mol Genet Genomics*. 287: (4) pp. 337-349. (2012). (IF: 2.881)
13. **Pamjav H**, Kugler R, Zalán A, Völgyi A, Straky Z, Endrédi P, Kozma Z. X chromosomal recombination study in three-generation families in Hungary. *Forensic Science International: Genetics*, 6: (3) pp. e95-6. (2012).
14. Horváth G, Zalán A, Kis Z, **Pamjav H**. A genetic study of 12 X-STR loci in the Hungarian population. *Forensic Science International: Genetics*. 6: (1) pp. e46-47. (2012). (IF: 3.861).
15. **Pamjav H**, Zalán A, Béres J, Nagy M, Chang YM. Genetic structure of the paternal lineage of the Roma People. *Am J Phys Anthropol*. 145:(1) pp. 21-29. (2011). (IF: 2.824)
16. Zalán A, Béres J, **Pamjav H**. Paternal Genetic history of the Vlax Roma. *Forensic Sci Int Genet*. 5:(2) pp.109-13. (2011). (IF: 3.082)
17. Molnár A, Zalán A, Horváth G, **Pamjav H**. Allele distribution of the new European Standard Set (ESS) loci in the Hungarian population. *Forensic Sci Int Genet*. 5:(5) pp.555-556. (2011). (IF: 3.082).
18. Rak SÁ, Zalán A, Szabados G, **Pamjav H**. Population genetic data on 15 STR loci in the Hungarian population. *Forensic Sci Int Genet*. 5:(5) pp. 543-544. (2011). (IF: 3.082).
19. Bíró AZ, Zalán A, Völgyi A and **Pamjav H**. A Y-chromosomal comparison of the Majars (Kazakhstan) and Magyars (Hungary). *Am J Phys Anthropol*. 139:(3) pp. 305-310. (2009). (IF: 2.756).
20. Völgyi A, Zalán A, Szvetnik E, **Pamjav H**. Hungarian population data based on 11 Y-STR and 49 Y-SNP markers. *Forensic Sci Int Genet*. 3:(2) pp. e27-28. (2009). (IF:2.421)
21. Zalán A, Völgyi A, Brabetz W, Schleinitz D, **Pamjav H**. Hungarian population data of eight X-linked markers in four linkage groups. *Forensic Sci Int*. (2008). 175:(1) pp. 73-78. (IF:1.928)

22. Zalán A, Völgyi A, Jung M, Peterman O, **Pamjav H**. Hungarian population data of four X-linked markers: DXS8378, DXS7132, HPRTB, and DXS7423. *Int J Legal Med* 121:(1) pp. 74-77 (2007). (**IF: 3.030**)
23. Nagy M, Henke L, Henke J, Chatthopadhyay PK, Völgyi A, Zalán A, Peterman O, Bernasovská J, **Pamjav H**. Searching for the origin of Romanies: Slovakian Romani, Jats of Haryana and Jat Sikhs Y-STR data in comparison with different Romani populations. *Forensic Sci Int.* 169:(1) pp. 19-26. (2007). (**IF: 2.015**).

VI.2. Egyéb közlemények (IF: 46,197)

1. Kampuansai J, Völgyi A, Kutanan W, Kangwanpong D, Pamjav H. Autosomal STR variations reveal genetic heterogeneity in the Mon-Khmer speaking group of Northern Thailand. *Forensic Sci Int Genet.* 27: pp. 92-99. (2016). (**IF: 3.911**).
2. Mayer B, Silló P, Mazán M, Pintér D, Medvecz M, Has C, Castiglia D, Petit F, Charlesworth A, Hatvani Z, **Pamjav H**, Kárpáti S.: A unique LAMB3 splice-site mutation with founder effect from the Balkans causes lethal epidermolysis bullosa in several European countries. *Br J Dermatol.* 175(4):721-7. (2016). (**IF: 4.706**).
3. Martínez-Cruz B, Mendizabal I, Harmant C, de Pablo R, Ioana M, Angelicheva D, Kouvatsi A, Makukh H, Netea MG, Pamjav H, Zalán A, Tournev I, Marushiakova E, Popov V, Bertranpetit J, Kalaydjieva L, Quintana-Murci, Comas D. Origins, admixture and founder lineages in European Roma. *Eur J Hum Genet.* 24:(6) pp. 937-943. (2016). (**IF: 4.287**).
4. Haltrich I, Pikó H, **Pamjav H**, Somogyi A, Völgyi A, David D, Beke A, Garamvölgyi Z, Kiss E, Karcagi V, Fekete G. Complex X chromosome rearrangement associated with multiorgan autoimmunity. *Mol Cytogenet.* 8: Paper 51. 11 p. (2015). (**IF: 1.506**)
5. Juhász Z, Fehér T, Bárány G, Zalán A, Németh E, Pádár Z and **Pamjav H**. New clustering methods for population comparison on paternal lineages. *Mol Genet Genomics.* 290: (2) pp. 767-84. (2015). (**IF: 2.622**).
6. Ballantyne et al. Towards complete male individualization with rapidly mutating Y-chromosomal STRs. *Hum Mutat. Hum Mutat.* 35:(8) pp. 1021-32. (2014). (**IF: 5.340**).
7. Purps et al. A global analysis of Y-chromosomal haplotype diversity for 23 STR loci. *Forensic Sci Int Genet.* 12: pp. 12-23. (2014). (**IF: 4.604**).
8. Kozma Zs, Sandor G, Pamjav H, Huszar A. The human iris polymorphisms: computer-based and genetic assessments of human iris and possible applications in human identification. *AARMS* 12: (2) pp. 229-246. (2013).
9. Mendizabal et al. Reconstructing the Population History of European Romani from Genome-wide Data. *Curr Biol. Curr Biol.* 22:(24) pp. 2342-2349. (2012). (**IF: 9.494**)
10. Martinez-Cruz et al. Y-chromosome analysis in individuals bearing the Basarab name of the first dynasty of Wallachian kings. *PLoS One.* (7): (7) p. e41803. (2012). (**IF: 3.73**).
11. Kis Z, Zalán A, Völgyi A, Kozma Z, Domján L, Pamjav H. Genome deletion and insertion polymorphisms (DIPs) in the Hungarian population. *Forensic Science International: Genetic.* 6:(5) pp. e125-126. (2012). (**IF: 3.861**).
12. Zalan A, Pamjav H, Kokeny S, Volgyi A. A study of thirty autosomal Single Nucleotide Polymorphism markers in Hungarian population. *Leg Med (Tokyo).* 9:(3) pp. 171-174. (2007).
13. Keyser-Tracqui C, Crubezy E, Pamzsav H, Varga T, Ludes B. Population origins in Mongolia: genetic structure analysis of ancient and modern DNA. *Am J Phys Anthropol.* 131:(2) pp. 272-281. (2006).(**IF: 2.136**)
14. Varga T, Keyser C, Beer Z, Péntzes Z, Pamzsav H, Csete K, Ludes B. Short tandem repeat data analysis in a Mongolian population. *Leg Med (Tokyo).* 5 Suppl 1. S156-199. (2003).