

MTA DOKTORI PÁLYÁZAT

Doktori értekezés

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

Tisztelt Olvasó!

Engedje meg, hogy bemutathassam az értékezésben szereplő két szakterület színrelépésének pillanatát rögzítő eredeti angol nyelvű rövid leírást! Az igazságügyi genetika 1984-ben, a humán evolúciós vizsgálatok első szikrája (Mitochondrial DNA and human evolution) pedig 1987-ben pattant ki. De utóbbi igazából akkor vált ismertté, amikor az Éva hét lánya című könyv megszületett.

Forensic Genetics: The 'eureka' moment

"My life changed on Monday morning at 9.05 am, 10 September 1984. What emerged was the world's first genetic fingerprint. In science, it is unusual to have such a 'eureka' moment. We were getting extraordinarily variable patterns of DNA, including from our technician and her mother and father, as well as from non-human samples. My first reaction to the results was 'this is too complicated', and then the penny dropped and I realised we had genetic fingerprinting."

It opened up a new area of science. The research team immediately grasped its applications, including crime, paternity and identical twins, as well as work on conservation and diversity among non-human species."



The most common approach is the comparison of homologous sequences for genes using sequence alignment techniques to identify similarity. Another application of molecular phylogeny is in DNA barcoding, wherein the species of an individual organism is identified using small sections of mitochondrial DNA or Y-DNA. Another application of the techniques that make this possible can be seen in the very limited field of human genetics, such as the

ever-more-popular use of genetic testing to determine a child's paternity, as well as the emergence of a new branch of criminal forensics focused on evidence known as genetic fingerprinting.

Ancient History: The seven daughters of the Eve

The Science That Reveals Our Genetic Ancestry (W.W. Norton & Company Inc, 2001)

The Seven Daughters of Eve is a thrilling work of science that reveals how biological research can enrich our tangled lives. It is a book that chronicles many of the most exciting developments in genetics over the past decade by a man who is not only a brilliant scientist but also a gifted and thoroughly engaging writer. It ultimately demonstrates how much more we still have to discover about the absorbing story of human evolution.

One of the most dramatic stories of genetic discovery since James Watson's *The Double Helix*—a work whose scientific and cultural reverberations will be discussed for years to come. In 1994 Professor Bryan Sykes, a leading world authority on DNA and human evolution, was called in to examine the frozen remains of a man trapped in glacial ice in northern Italy. News of both the Ice Man's discovery and his age, which was put at over five thousand years, fascinated scientists and newspapers throughout the world. But what made Sykes's story particularly revelatory was his successful identification of a genetic descendant of the Ice Man, a woman living in Great Britain today. How was Sykes able to locate a living relative of a man who died thousands of years ago? In the *Seven Daughters of Eve*, he gives us a firsthand account of his research into a remarkable gene, which passes undiluted from generation to generation through the maternal line. After plotting thousands of DNA sequences from all over the world, Sykes found that they clustered around a handful of distinct groups. Among Europeans and North American Caucasians, there are, in fact, only seven. This conclusion was staggering: almost everyone of native European descent, wherever they may live throughout the world, can trace their ancestry back to one of seven women, the Seven Daughters of Eve. Naming them Ursula, Xenia, Helena, Velda, Tara, Katrine, and Jasmine, Sykes has created portraits of their disparate worlds by mapping the migratory patterns followed by millions of their ancestors. In reading the stories of these seven women, we learn exactly how our origins can be traced, how and where our ancient genetic ancestors lived, and how we are each living proof of the almost indestructible strands of DNA, which have survived over so many thousands of years. Indeed, *The Seven Daughters of Eve* is filled with dramatic stories: from Sykes's identification, using DNA samples from two living relatives, of the remains of Tsar Nicholas and Tsaress Alexandra, to the Caribbean woman

whose family had been sold into slavery centuries before and whose ancestry Sykes was able to trace back to the Eastern coast of central Africa. Ultimately, Sykes's investigation reveals that, as a race, what humans have in common is more deeply embedded than what separates us.

TARTALOMJEGYZÉK

1.	RÖVIDÍTÉSEK.....	4
2.	ELŐSZÓ ÉS KÖSZÖNETNYILVÁNÍTÁS	6
3.	BEVEZETÉS	9
3.1.	DNS-UJJLENYOMAT: IGAZSÁGÜGYI GENETIKA (FORENSIC GENETICS)....	9
3.1.1.	Az igazságügyi genetika alapkérdései.....	10
3.1.2.	DNS polimorfizmus	14
3.1.3.	Az autoszómális mikroszatellita (STR) lokuszok jellemzése	16
3.1.4.	Származás-megállapítási vizsgálatok – apasági tesztek.....	18
3.1.5.	Bűncselekmény helyszíni minták – egyedi személyazonosítás	19
3.1.6.	A törvényszéki DNS vizsgálatok statisztikai interpretációja	21
3.1.7.	Ivari kromoszómák.....	22
3.1.7.1.	Az X-kromoszómás STR markerek	22
3.1.7.2.	Az Y-kromoszómás STR markerek	26
3.1.7.3.	Az Y-kromoszómás SNP markerek	28
3.1.8.	Az anyai örökség.....	30
3.1.8.1.	A mitokondriális DNS szerkezete	30
3.1.8.2.	A mitokondriális DNS vizsgálat jelentősége az igazságügyi genetikában ...	31
3.1.8.3.	Mitokondriális haplocsoport meghatározás	33
3.2.	ÖSTÖRTÉNET ÉS AZ EMBER GENETIKAI ÚTJA	34
3.2.1.	A modern ember kirajzása Afrikából („Out of Africa” theory).....	34
3.2.2.	Ádám és Éva randevúja az Édenkertben	37
3.2.3.	Éva genetikai hagyatéka.....	38
3.2.4.	Szüleink genetikai öröksége: Nukleáris DNS és őstörténet.....	41
3.2.5.	Ádám genetikai üzenete	42
3.2.6.	Színre lép a DNS alapú filogenetika	43
3.2.6.1.	Y-kromoszóma haplocsoportok elterjedése	48
3.2.6.2.	Y-haplocsoportok Európa korai történelmében	58
3.2.6.3.	Európa a földművelés előtt.....	60
3.2.6.4.	A földművelés Európába érkezése	62
3.2.6.5.	A bronzkor – Európa mai genetikai képének kialakulása.....	64
3.2.7.	A mitokondriális DNS (mtDNS) haplocsoportok rövid áttekintése.....	70

4.	ANYAG ÉS MÓDSZER	77
4.1.	Minták és populációk	77
4.2.	A DNS izolálása és koncentrációjának meghatározása.....	78
4.3.	Genotipizálás fragmenshossz analízissel	78
4.4.	Az Y-kromoszómális SNP lokuszok PCR-amplifikációja és genotipizálása TaqMan próbával (haplocsoport meghatározás).....	78
4.5.	Populációstatisztikai analízisek az autoszómális és X-STR lokuszokon	79
4.6.	Struktúra analízis.....	82
4.7.	G teszt.....	82
4.8.	Populációstatisztikai és genetikai analízisek az Y-STR és Y-SNP lokuszokon	82
4.9.	F-statisztika és AMOVA.....	83
4.10.	Filogenetikai fa szerkesztése	84
4.10.1.	MDS (Multidimensional Scaling)	84
4.10.2.	Filogeográfiai analízis (Phylogeographic Analysis)	84
4.10.3.	Hálózat analízis és divergálási idő becslés.....	85
4.10.4.	Bayes Skyline Plots (BSP) analízis	85
5.	AZ EREDMÉNYEK ÉS AZOK MEGVITATÁSA	86
5.1.	AZ IGAZSÁGÜGYI GENETIKA.....	86
5.1.1.	Autoszómális STR markerek vizsgálata	86
5.1.2.	X-rekombináció vizsgálata, X-STR markerekkel	94
5.1.2.1.	Linkage analízis a férfiak X-kromoszómális haplotípusa alapján	98
5.1.3.	X-kromoszómális STR markerek vizsgálata hiányos apasági ügyekben.....	99
5.1.4.	Y-STR lokuszok vizsgálata az igazságügyi alkalmazásban.....	106
5.1.5.	A mtDNS vizsgálat az igazságügyben	110
5.2.	ÖSTÖRTÉNET: GÉNEKBE VÉSETT TÖRTÉNELEM.....	111
5.2.1.	A férfi genetikai útja: Y-kromoszóma STR és SNP vizsgálata magyar és roma populációkban	111
5.2.1.1.	Haplocsoport eloszlás a recens magyar populációban	111
5.2.1.2.	Filogeográfiai analízis roma populációkban	117
5.2.1.3.	Median Joining Network analízis.....	119
5.2.1.4.	Az R1a haplocsoport keletkezése, terjedése és alcsoportjai	122
5.2.1.5.	Apai leszármazási vonalak összehasonlítása a magyar és a manysi népességben.....	126
5.2.1.6.	A székelyek eredetével kapcsolatos új nézőpontok	133

5.2.1.7.	Az Y-kromoszómális haplocsoportok összesített eloszlása a mai magyar népességben.....	139
5.3.	Az mtDNS VIZSGÁLATI EREDMÉNYEK.....	143
5.3.1.	174 eurázsiai populáció mtDNS elemzése iteratív rangkorrelációs módszer segítségével	143
5.3.2.	A karosi honfoglaló temetőből származó minták vizsgálata.....	144
5.3.3.	Az anyai ági mtDNS haplocsoportok a mai magyar népességben.....	152
5.3.4.	A teljes mtDNS genom vizsgálata európai és közel keleti populációkban	154
5.4.	A zenébe vésett gének, avagy a zenegenetika (music-genetics)	162
5.4.1.	A populációk genetikai kapcsolatai.....	166
5.4.2.	Populációk zenei kapcsolatai	169
5.4.3.	Zenei és genetikai korreláció.....	173
6.	IRODALOMJEGYZÉK.....	179
7.	KÖZLEMÉNYEK LISTÁJA.....	201
7.1.	A doktori értekezésben tárgyalt közlemények	201
7.2.	Egyéb közlemények	203

1. RÖVIDÍTÉSEK

aDNS	Ancient DNA
BOI	Budapesti Orvosszakértői Intézet
BSP	Bayes Skyline Plots
CE	Capillary Electrophoresis
CHgC	Correlating Haplogroup Clusters
CODIS	Combined DNA Index System
CRS	Cambridge Reference Sequence
rCRS	Revised Cambridge Reference Sequence
DNS	Dezoxiribonukleinsav
DTV	Dallam Vonaltípus Vektor
EDNAP	European DNA Profiling Group
ENFSI	European Network of Forensic Science Institutes
FBI	Federal Bureau of Investigation
FTDNA	Family Tree DNA
Hd	Hypothesis of Defence
Hp	Hypothesis of Prosecutor
HV	Hipervariábilis régió
ISFG	International Society for Forensic Genetics
ISOGG	International Society of Genetic Genealogy
IRC	Iterative Rank Correlation
ISZKI	Igazságügyi Szakértői és Kutató Intézetek
KYA	Kilo Years Ago
LCN	Low Copy Number
LD	Linkage Disequilibrium
LG	Linkage Group
LR	Likelihood Ratio
MEC	Mean Exclusion Chance
MDS	Multidimensional Scaling
MJ	Median Joining
MSY	Male Specific Y Chromosome

MRCA	Most Recent Common Ancestor
mtDNS	Mitokondriális DNS
NGS	Next Generation Sequencing
NRY	Non-recombining Region of Y
PCR	Polymerase Chain Reaction
PD	Power of Discrimination
PE	Power of Exclusion
PI	Paternity Index
PIC	Polymorphism Information Content
pM	Probability of Match
SOC	Self Organising Cloud
SOM	Self Organising Map
SNP	Single Nucleotide Polymorphism
STR	Short Tandem Repeat
TMRCA	Time to Most Recent Common Ancestor
X-STR	X-kromoszomális STR
Y-STR	Y-kromoszomális STR
YHRD	Y Chromosome Haplotype Database

2. ELŐSZÓ ÉS KÖSZÖNETNYILVÁNÍTÁS

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, valamint, hogy az igazságügyi genetika területén bevezetésre kerüljenek a legújabb tudományos eredményeken alapuló DNS vizsgálati eljárások, ideértve mind a műszerezettséget, mind a kapcsolódó nemzetközi módszertani előírások és követelmények hazai meghonosítását.

Az MTA doktori fokozat elnyerésére irányuló jelen értekezés bemutatja az igazságügyi genetika 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 becsülhető a hazai törvényszéki eljárások során.

Az igazságügyi genetikai vizsgálatokat a PhD fokozat megszerzését követően kezdtem el, ennek során sikerült kiépíteni egy olyan genetikai laboratóriumot a volt Igazságügyi Szakértői Kutató Intézetben (ISZKI) belül - Dr. Susa Éva volt főigazgató támogatásával - amely a nemzetközi igazságügyi genetikai közösség (forensic genetics community) elismert tagjává vált. Ebben a közegben elsősorban MSc és PhD hallgatókkal dolgoztam együtt. Az értekezés ennek az időszaknak az igazságügyi genetikai vonatkozású munkáit öleli fel. DNS laboratóriumunk a volt szerológiai laboratóriumból nőtte ki magát. 2000-ben egy asszisztenssel kezdtem el a DNS laboratórium kiépítését egy darab ABI310 készülékkel, mára „parányi” kis laboratóriumunk egy teljesen automatizált, DNS-profil gyártó „üzemmé” alakult át, meghatározva évente 30-40 ezer terhelti minta DNS-profilját a bűnügyi DNS adatbázis számára.

Visszatekintve az elmúlt közel 20 évre, elmondhatom, hogy munkámat sok szempontból szerencsés korszakban kezdtem. Akkor már volt internet, 2001-ben fejeződött be a „Human Genome Project” és az interneten keresztül a szekvencia adatok hozzáférhetőek voltak. Emellett számos nemzetközi konferencián vehettem részt, így megadatott a találkozás egy olyan rügyező és önállóvá váló tudományág, mint az igazságügyi genetika, alapítójával Professor Sir Alec Jeffreysszel, továbbá John M. Butlerrel, akik szinte láncreakcióban publikálták az újabb és újabb módszereket és szakkönyveket. Ugyancsak lehetőségem nyílt találkozni az Y-kromoszómális markerek nemcsak az igazságszolgáltatás, hanem a humán

migráció és evolúció céljából való vizsgálatát indító szakmai korifeusokkal (vagy ahogy magamban elneveztem őket „Kings of Y Chromosome”), mint Chris Tyler-Smith (Cambridge, Sanger Institute), Lutz Roewer (a Y Chromosome Haplotype Reference Database alapítója, Charité, Berlin), valamint Mark Jobling (Leicester University), aki Alec Jeffreysszel dolgozott együtt (időközben Jeffreys visszavonult). E körben említendő továbbá, Davis Comas (Universitat Pompeu Fabra, Barcelona), aki a Genographic Project egyik vezetője. Szerencsésnek mondhatom magam, hogy ezekkel a kiemelkedő szaktekintélyekkel kapcsolatba kerülhettem, együtt publikálhattam velük számos tudományos cikket a „Peopling of Europe” és a „Genographic Project” témakörökön belül.

Ugyancsak ebben az időszakban a molekuláris genetika számos területén végbement ugrásszerű fejlődésnek köszönhetően a vizsgálati módszerek cunamiszerűen szaporodtak. Ezek közé tartozik a PCR technika, a microchip-alapú vizsgálatok, a szekvenálási eljárások, fragmens analízisek kapilláris elektroforézis segítségével és a TaqMan Assay alapú vizsgálatok. Szerencsésnek mondhatom magam, hogy aktív éveim egybeestek ennek a tudományágnak a robbanásszerű fejlődésével és a digitalizáció, illetve az internet térnyerésének korszakával.

A jelen értekezésben bemutatott eredmények számos más kolléga és munkatárs segítségével nélkül nem jöhettek volna létre.

Mindenek előtt családomnak szeretnék köszönetet mondani. Férjemnek, Szvetnik Antalnak és a lányomnak Szvetnik Enikő Aminának. Támaszuk és segítségük nélkül nem élvezhettem volna a kutatás örömét. Hálás vagyok megértésükért, toleranciájukért, mivel sokszor csak „aludni” jártam haza - férjem szerint a „vendégházba”.

Az eredmények elérésében köszönetemet fejezem ki a laboratórium volt MSc diákjainak, érkezési sorrendjében említve szakdolgozó hallgatóimat: Petermann Orsolya, Rak Szilvia, Zalán Andrea, Szegner Katalin, Molnár Aliz, Horváth Gergely, Kugler Renáta, Toldy Emese, Farkas Zsuzsanna, Borbély Andrea, Erős Ágnes, Varga Dániel, Bárány Gusztáv.

Végzett PhD hallgatóm Dr. Vágó-Zalán Andrea, aki nemcsak biológusként, hanem matematikusként is a szakértői tevékenység és a kutató munka területén matematika-statisztikai felkészültségével óriási segítséget nyújtott és hozzájárult „parányi laboratóriumunk” iskolateremtő műhely alapjainak lerakásához.

Köszönettel tartozom Dr. Buzás Zsuzsannának és Dr. Fodor Andrásnak, akik témavezetőim voltak a PhD fokozat megszerzése során; nekik köszönhetem a molekuláris genetikai módszerek elsajátítását, amelyek a későbbi szakértői és kutatói munkám alapját képezték. Dr. Buzás Zsuzsanna szakmai tudása, önzetlen segítsége és kivételes emberi

tulajdonságai mély hatást gyakoroltak rám, mivel a PhD megszerzésére irányuló feladatokat munka mellett, szabadidőmben (szabadságom alatt is), kisgyermekes anyaként végeztem. Önfeláldozó segítsége nélkül nem tudtam volna megszerezni a PhD fokozatot. Dr. Fodor Andrásnak különösen hálás vagyok azért, hogy PhD hallgatóként számos nemzetközi konferencián vehettem részt, továbbá, hogy támogatásával nyugodt szívvel fordulhattam a szakma nagyjaihoz.

Hálával tartozom Vida Gábor professzor úrnak, akadémikusnak, az ELTE Genetikai Tanszék korábbi vezetőjének, aki bevezetett a genetika szépségeibe, akitől egyetemistaként hallgathattam az akkori oktatási programban szereplő összes genetikai tárgyat. A professzor úr előadói felkészültsége, a genetika iránti elhivatottsága, és integráló tudományos szemlélete lenyűgözött. Valójában erre vezetem vissza, hogy „beleszerettem a genetikába” és nem tudtam volna elképzelni az életemet a genetika/DNS nélkül. Kevés ember részesül abban a szerencsében, mint én, hogy a foglalkozása egyben a hobbija is.

Dr. Orosz László professzor úrnak, akadémikusnak, az ELTE Genetikai Tanszék volt vezetőjének, a „két lábon járó genetikai enciklopédiának” mély köszönettel tartozom, mivel bármilyen szakmai probléma kapcsán bármikor fordulhattam hozzá és azonnali eligazítást kaphattam.

Dr. Susa Évának, a volt ISZKI főigazgatójának mérhetetlen hálával és köszönettel tartozom, nemcsak magam, hanem a szakma nevében is, hogy meglátva a benne rejlő lehetőségeket, támogatta a szakterület fejlesztését és művelését a szűk költségvetési keretből, mindig és mindenkor! Az ő támogatása nélkül nemcsak, hogy nem rügyezett volna az igazságügyi genetika kis hazánkban, hanem „szegényebb lett volna” a nemzetközi tudományos adatbázis közel 40 publikációval.

Végül, de nem utolsósorban a szüleimnek szeretnék köszönetet mondani, akik esélyt adtak nekem, hogy létezhesek a napfényes Föld nevű bolygón és kutathassam az élet rejtélyét a genetika eszköztárának a segítségével. A természet örök törvénye szerint eltávoztak, de szívemben mindig is ott vannak és segítségüket kértem, amikor szükségesnek éreztem.

3. BEVEZETÉS

Ez a fejezet 2 részből áll. Először megszólaltatom a „Génekbe vésett vallomások”, a DNS üzenetek, közül azokat, amelyek megmondják, mi a genetikai személyi számunk, azaz a DNS-ujjlenyomatunk. Bemutatom hogyan fejlődött az igazságügyi genetika a DNS-ujjlenyomat felfedezésének pillanatától önálló tudománnyá, ideértve ennek alapkérdéseit, vizsgálati diszciplínáit, valamint a statisztikai interpretációk elméleti háttérét. Részletesebben bemutatom 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-távoli múltjába. Hihetetlen, de igaz, hogy az első fejezetben bemutatott apai és anyai leszármazási vonalak hátrányai válnak előnyeivé a második fejezetben szereplő vizsgálati diszciplínának, azaz a sikeres anyák és a sikeres apák genetikai történetének. Felvázolom az őstörténeti kutatások néhány aktuális problémáját és ismertetem az értekezésben bemutatott eredmények szempontjából releváns paradigmákat.

3.1. DNS-UJJLENYOMAT: IGAZSÁGÜGYI GENETIKA (FORENSIC GENETICS)

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

Az utóbbi évtizedekben az igazságügyi genetika különálló tudományterületté fejlődött. Az 1980-as években Prof. Sir Alec Jeffreys – a leicesteri egyetem genetikusa - kutatási területe növények mikroszatellita vizsgálata volt. Feltűnt számára, hogy ezek a markerek különböző növények esetében különböző mintázatot adnak, ezért kíváncsi lett, vajon emberek esetében milyen mintázatok adódhatnak és megkérte asszisztensét, hozzon vérmintát családja tagjaitól, mármint magától, édesanyjától és édesapjától. 1984. szeptember 10-én, hétfőn reggel 9 órakor nézte meg az eredményeket és "**Heurékát**" kiáltva a következő megállapítást tette: „egy rendkívüli változatos mintázatot kaptunk a vizsgált 3 személy esetében, valamint non-humán mintáknál is. Az eredményeket látva az első reakcióm az volt, hogy "ez túl komplikált", de később leesett a tantusz, és rájöttem, ez nem más, mint a genetikai ujjlenyomat. "

Ez volt, azaz pillanat, amikor az **igazságügyi genetika** a világra jött, úgy 34 éve.

A módszert, amely az ún. RFLP (Restriction Fragment Length Polymorphism) technikán alapult, először Angliában alkalmazták törvényszéki eljárásban 1985-ben [1, 2], egy nemi erőszaktevő és sorozatgyilkos azonosítására (Colin Pitchfork). Napjainkban az igazságügyi genetika a világ szinte minden országában az igazságszolgáltatási eszköztár szerves részét képezi [3]. A világ szinte minden országa bevezette az ún. *nemzeti bűnügyi DNS nyilvántartást*, ahol a bűncselekmények helyszínéről származó, be nem azonosított személyekhez tartozó biológiai anyagmaradványok DNS-profilja mellett akár több százezer bűnelkövető – tettesek - DNS-mintázatát is őrzik [4]. Persze az előbbiekhöz szükség volt egy korszakalkotó felfedezésre - a PCR technika bevezetésére – az excentrikus zseninek, Kary Mullisnak köszönhetően (1983-ban).

Az igazságügyi genetika, mint alkalmazott tudományág segítséget nyújt bűncselekmények felderítésében, továbbá származás-megállapítási vizsgálatokban - apasági, anyasági, testvérségi tesztekben - valamint ismeretlen holttestek, emberi testmaradványok azonosításában. Ma már a személyazonosítás a DNS-profillal szinte egyedileg lehetséges, hacsak nincs az illetőnek egypetéjű ikertestvére. Az apasági DNS vizsgálatokban a biológiai apaság pozitív bizonyítása is megközelítőleg 100 %-os bizonyossággal tehető meg. Az apaság, anyaság kizárása (negatív bizonyítása) pedig kategorikussá vált.

A molekuláris biológiai technológia és az informatikai háttér robbanásszerű fejlődése napjainkra lehetővé tette az igazságügyi genetikai módszereknek az említetteknél sokkal szélesebb körű felhasználását. A bűncselekmény helyszínén hátrahagyott biológiai anyagmaradványokból kinyert DNS mintákból kiolvashatók az elkövető bizonyos külső jellegei, mint a szemszín, hajszín, bőrszín, testmagasság, földrajzi és etnikai hovatartozás, azaz a „genetikai szemtanú” vallomása [5, 6, 7]. Az említett eredményekről az ismeretterjesztés és a média jóvoltából, valamint számtalan CSI (Crime Scene Investigation) TV sorozatnak köszönhetően a társadalom széles rétege értesülhetett.

3.1.1. Az igazságügyi genetika alapkérdései

Az igazságügyi genetikai vizsgálatok alapvető célja annak megállapítása, 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 annak csoportosulásának származási, leszármazási vagy más rokonsági kapcsolatait. Tehát 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 szakterülettől.

A humán igazságügyi genetikai vizsgálatok fő feladata annak a megállapítása, hogy a törvényszéki eljárás során, bűncselekmények helyszínén hátrahagyott biológiai anyagmaradványokból, valamint polgári peres eljárás esetében a személyektől levett mintákból (apasági, anyasági és rokonsági vizsgálatok) meghatározott DNS-profilok hozzárendelhetők-e egy adott személyhez. Ebből világosan következik az, 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 csak az adott, lehetőleg „egyetlen” személyre vagy egyedre jellemző. Emellett 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 története 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 vizsgálat során olyan markereket (jellegeket) kell megszólaltatni, amelyek egyrészt sokféle változatban (polimorfizmus) figyelhetők meg az adott népességben, másrészt függetlenek az adott ember életkorától, pillanatnyi fiziológiai állapotától és egyéb más külső tényezőktől.

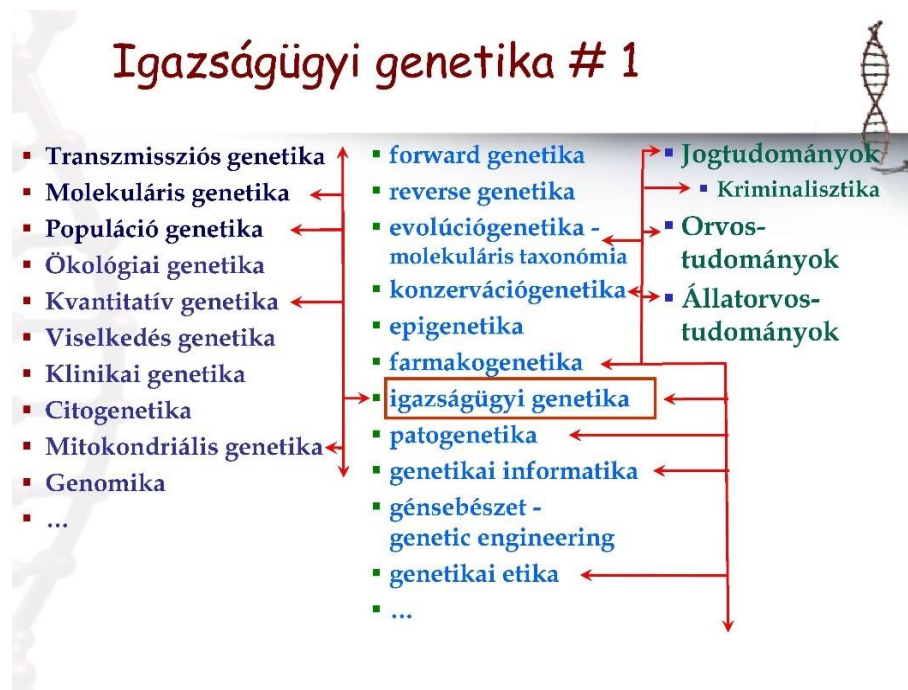
A humán populációk DNS vizsgálatával a következő kérdésekre lehet választ kapni:

1. az emberi populációk, mint egy-egy földrajzi régióban élő közösségek között a vizsgált markerek allélgyakorisági eloszlásában mutathatók-e ki 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/személyek 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. úgyszintén kiolvasható-e 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 (human population history)?

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. Ez a biostatistikai kiértékelés igen komplex is lehet, de egyúttal törvényszéki eljárás során a laikus személyek (bíró, ügyész, ügyvéd, ügyfelek) számára is érthetőnek és egyértelműnek kell lennie. A DNS vizsgálatok során generált adatok minősége és hibagyakorisága gyakorlati problémát okozhat, valamint a vizsgálat során keletkezett személyes genetikai adatok kezelése, védelme és törvénynek megfelelő felhasználása komoly etikai kérdéseket vethet fel a DNS adatbázisok kapcsán.

A DNS alapú azonosítással és származástannal foglalkozó igazságügyi genetikát kihívásai, problémafelvetései, módszertani és technikai fejlesztései, eredményei nemcsak passzív felhasználóként, hanem aktív közreműködőként is kapcsolják a genetikához, mint alaptudományhoz és a jogalkalmazáshoz egyaránt, ugyanakkor ez nem tekinthető a kettő egyszerű összegzésének [8]. Az igazságügyi genetika önálló tudományterületté nőtte ki magát, amely felhasználja a genetika módszereit, tudományos eredményeit. Önálló voltát támasztja alá azon körülmény is, hogy igazságügyi genetikai kutatások a genetika más területére vonatkozó információkat [pl. molekuláris genetika, evolúciógenetika, populációgenetika, apai és anyai leszármazási vonalak (filogenetika) feltérképezése] is szolgáltatnak (**1. ábra**), melyek önmagukban ugyanakkor nem jelentik az igazságügyi genetika célját.



1. ábra. Az igazságügyi genetika kapcsolatrendszere más tudományterületekkel¹

Törvényszéki szempontból nézve a genetika nem más, mint eszköz, melynek segítségével a jogi hipotézisek megerősíthetők vagy elvethetők. A tudományterület önállósodása – más, korábban határ - vagy multidiszciplináris tudományágak evolúciójával analóg módon – bizonyos tudományos koncepciók, besorolások és dogmák megváltozásával is jár. Habár a változások érvényre jutása eltérő mértékű lehet, 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 a törvényszéki perekben való bizonyítás lehetőségét [8].

A nemzetközi tudományos és szakmai közvélemény által elfogadott korszerű ismeretek, továbbá a jelenleg hozzáférhető laboratóriumi berendezések és számítástechnikai eszközök birtokában a vizsgálati módszerekkel szemben igen szigorú feltételek érvényesülnek, különösen nagy hangsúlyt kapnak a minőségbiztosítás követelményei és az akkreditációk. Az egységes Európai Gazdasági Térségen belül követelmény, hogy az igazságszolgáltatás azonos színvonalú és megbízhatóságú szakértői szolgáltatásra támaszkodva valósuljon meg.

¹ Dr. Pádár Zsolt PhD értekezése, 2006

3.1.2. DNS polimorfizmus

Az emberiség sokszínűsége főleg az emberek közötti különbözőségből ered. Populációgenetikai szempontból ezek a különbözőségek polimorfizmus mintázatban tárolódnak és adódnak át generációról generációra. Ezek a különbözőségek tesznek minket teljesen egyedivé, kivéve, ha van egypetjű ikertestvérünk. Az igazságügyi genetikában bűncselekmények tetteseinek vagy áldozatainak azonosítására alkalmazott DNS-ujjlenyomat vizsgálatok alapját a nem kódoló DNS szakaszok – intronok és intergénikus régiók - vizsgálata adja.

A DNS állomány döntő többsége (több mint 99,7%) az emberekben azonos, csak töredék része (kb. 0,3%, vagyis 10 millió nukleotid) egyénileg eltérő, és alkalmas az egyedek megkülönböztetésére genetikai módszerekkel. A DNS-ben tárolt információ személyazonosítási és rokonsági vizsgálati célú felhasználásának lehetőségét a variábilis régiók biztosítják [3, 4].

Az ember az egyetlen élőlény, aki megfejtette és leolvasta önmaga „gyártási receptjét”. A Humán Genom Projekt 1990-ben kezdődött, és 2001-ben tette közzé az emberi genom DNS szekvenciájának „nyers” változatát. A végleges referencia szekvenciát 2003 áprilisában hozták nyilvánosságra. Ez egy lélegzetállító pillanat volt a biológiai tudomány számára, azóta már sok más élőlény – állat- és növényfaj – teljes genetikai leírása ismertté vált. A projekt más nemzetközi kutatásokkal együtt felderítette a humán genom diverzitását, továbbá a diverzitás hátterében álló mutációs folyamatokat.

Az elmúlt két évtizedben keletkezett DNS kutatási ismeretek felelnek meg leginkább a törvényszéki eljárás által támasztott igen magas szintű követelményrendszernek. Így napjainkra a biológiai anyagmaradványok igazságügyi célú személyazonosításában a DNS-mintázat (DNS-profil) vizsgálata egyeduralkodóvá vált.

Az emberi egyedek sokféleségének egyik fő oka, hogy a természetben megtalálható szinte minden, az öröklődésért felelős DNS-lánc időnként mutációs változást „sz szenved el” és ennek eredményeképpen polimorfizmus alakul ki. A polimorfizmus adódhat egyetlen nukleotid eltéréséből (szerkezeti polimorfizmus) vagy egy bizonyos DNS szakasz méretbeli (hossz-polimorfizmus) eltéréséből. A szerkezeti polimorfizmus mérete egyetlen nukleotidtól (SNP: Single Nucleotide Polymorphism) több százezer nukleotidig terjedhet. Az emberi genomban általában 1000 bázispáronként egy nukleotid eltérés mutatható ki [3]. A hossz-polimorfizmust okozó markerek, mint például az STR (STR: Short Tandem Repeat)

lokuszok egyenlőtlenül szétszórva helyezkednek el a humán genomban [3]. Egy személy DNS mintájából különböző lokuszokon megállapított allélok együttesét DNS-mintázatnak, genotípusnak vagy DNS-profilnak nevezzük. A DNS-mintázat felfogható egy adott ember személyi számának genetikai megfelelőjeként. Azonban a személyi számmal ellentétben a DNS-profil egyedisége nem kizárólagos, azaz soha sem 100%-os, hanem statisztikailag megközelíti a 100%-ot. A Földön élő emberi sokaságon belül két személy nem rendelkezhet ugyanazzal a DNS-profillal, az elméleti lehetőség persze fennáll, de a valószínűsége elenyészően kicsi – ez természetesen a vizsgált markerek számától is függ –, kivéve, ha az illetők egypetűjű ikertestvérek.

A biológiai nyomból meghatározható DNS-profil egyediségének kérdését alapvetően több tényező is befolyásolja. Az egyik, hogy az emberi faj egyedei szaporodási közösségükben kisebb-nagyobb mértékben 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 [9]. Az egyediséget korlátozó további tényező, hogy a bűncselekmények elkövetése során keletkező biológiai nyom nem ritkán több személy mintájának keveredéséből jön létre.

Mindezen megfontolások alapján az igazságügyi genetikus szakértők – eltérve pl. az ujjlenyomat-szakértők által alkalmazott gyakorlattól – jelenleg nem kategorikus, hanem valószínűsítő véleményt adnak a személyazonosítás vagy származás-megállapítás során [10].

Két biológiai minta közös eredetének megállapítása a mintákból nyert DNS-profilok összehasonlításából áll. Ha a két DNS-profil minden egyes lokuszon megegyező (azonos eredet), vagy legalábbis átfedő mintázatot mutat (közvetlen ági rokon), akkor valószínűsíthető a közös eredetük. A biológiai rokonság vizsgálata során azt elemezzük, hogy a két feltételezett rokon személy DNS-profilja között van-e alléltípus egyezés, és ha igen, akkor ezek öröklődése megfelel-e Mendel törvényeinek.

3.1.3. Az autoszómális mikroszatellita (STR) lokuszok jellemzése

Az STR lokuszok olyan DNS-szakaszok, ahol 2-6 bázispárból álló motívumok ismétlődnek tandem módon egymás után. Az STR lokuszok ismétlődéseinek száma a személyek között nagy változatosságot mutat, és ez a polimorf tulajdonság teszi alkalmassá ezeket személyazonosítási célokra [3]. Egy STR lokusz különböző alléljai méretükben is különböznek egymástól – ahogy korábban említésre került – ezért az STR polimorfizmusokat a hossz-polimorfizmusok közé sorolják. Polimorf mikroszatelliták ezreit azonosították a humán genomban. Az STR repeat (ismétlődő) szekvenciákat ismétlődő egységük hosszának megfelelően nevezik el di-, tri-, tetra-, penta-, hexa-repeat egységeknek. Az igazságügyi genetikában leggyakrabban alkalmazott STR lokuszok általában tri- és tetranukleotidokból állnak.

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

Nem minden STR allél mérete felel meg a repeat-egysége egész számú többszörösének. Vannak olyan, ún. mikrovariáns allélok is, amelyek az ismétlődéseket határoló régióban (flanking region) egy vagy több nukleotidot ért változást hordoznak, s emiatt a mikrovariáns allél mérete a konszenzusos allél méretével nem egyezik meg, pl. 23.2 allél az FGA lokuszon.

Az STR markerek igazságügyi célú alkalmazhatóságát – magas polimorfizmus fokuk mellett – PCR módszer segítségével sokszorozhatóságuk teszi lehetővé. A PCR módszer felfedezésével olyan eszköz került a genetikus szakértők birtokába, amelynek segítségével addig használhatatlannak tűnő biológiai anyagmaradványokat – pl. ujjlenyomat – is meg lehet „szólaltatni” az igazság kiderítése céljából. A PCR legnagyobb előnye, hogy a genomnak a szakértő által vizsgálni kívánt részéből laboratóriumi körülmények között egy-két órán belül több milliónyi másolat is készíthető, továbbá, hogy az ún. *multiplex PCR* segítségével egy PCR reakció során szimultán több STR lokusz is megsokszorozható. A megsokszorozandó DNS-szakasz lerövidítésével degradált DNS minták is „megszólaltathatóvá” váltak. A multiplex PCR ezen tulajdonságának köszönhetően akár néhány pikogramnyi vizsgálandó

DNS, azaz egyetlen sejt genetikai állománya (kb. 7 pg) is elegendő lehet a személyazonosítás számára. A PCR technikának ilyen magas fokú érzékenysége éppen az előbbiek miatt nemcsak előnyös, hanem hátrányos is lehet. Ha nem tartják be az igen szigorú vizsgálati előírásokat, a vizsgálandó mintában lévő parányi DNS-t felülfertőzheti (kontaminálhatja) a vizsgálatot végző személy saját DNS-e vagy valamely korábban vizsgált másik személy DNS-e. A szigorú szakmai és minőségirányítási követelmények betartása mellett viszont releváns vizsgálat végezhető akár 0,2-1 ng mennyiségű DNS minta felhasználásával. Az STR lokuszok DNS-fragmensanalízissel az esetek döntő többségében egyértelműen elkülöníthetők egymástól kapilláris elektroforézis készülékeken.

A polimorf STR lokuszoknak egy szűkebb köre – 16 marker – képezi a jelenlegi nemzeti bűnügyi DNS-profil nyilvántartások alapját. A Nemzetközi Rendőri Szervezet (Interpol), az Európai Igazságügyi Szakértői Intézetek Hálózata (ENFSI), valamint az USA Szövetségi Nyomozó Irodája (FBI) által vizsgálatra ajánlott markerek is ebből a körből kerülnek ki. Megindult a nemzeti adatbázisok nemzetközi szerveződése (Interpol DNA Unit, 2003), és együttműködése (Prüm Convention, 2005).

A Nemzetközi Igazságügyi Genetikai Társaság (ISFG, International Society of Forensic Genetics) több ajánlást is kidolgozott az igazságügyi genetikai célból elemzett STR lokuszok és alléljaik nevezéktanára vonatkozóan [11, 12].

A fentieknek megfelelően azt a DNS szálát kell figyelembe venni, amelyet a szakirodalomban vagy nyilvános adatbázisban először publikáltak, a szekvenciát 5'-3' irányban kell leolvasni. Az egységes lokusz nevezéktan érdekében követni kell a „D#S#” rendszert. A „D” DNS-t jelent, autoszómális lokuszok esetében az utána következő szám (1-től 22-ig) – ivari kromoszómás markerek esetében az utána következő betű (X vagy Y) – azt a kromoszómát jelöli, amelyen a lokusz található. Az „S” az egykópiás szekvencia („single copy sequence”) rövidítése, mely azt jelenti, hogy az adott szekvencia csak egyszer fordul elő a humán genomban, és az azt követő szám azt jelöli, hogy a lokuszt hányadikként írták le az adott kromoszómán. A szabályozás létrejöttékor már számos lokusz volt széles körben használatban, így ezek nem standard (pl. FGA, vWA, TH01 stb.) elnevezése továbbra is megmaradt.

Az allélokat a teljes ismétlődési egységek száma alapján kell elnevezni. Előfordulhatnak ún. mikrovariáns (köztes- v. inter-) allélok, melyek nem teljes „repeatet” is tartalmaznak, ilyen esetben a teljes ismétlődés száma után ponttal elválasztva kell a hiányos ismétlődés meglévő bázisainak számát feltüntetni. Léteznek nagyon bonyolult szerkezetű STR-ek, ezek többféle és eltérő méretű motívumból állnak, ezért igen nehéz az allélok megfelelő

elnevezése. Ezekben az esetekben az allélok hosszúsága között kell kapcsolatot keresni, alapul véve a már korábban elnevezetteket. Előfordulhat, hogy a határoló (un. „flanking region”) régióban inszerció vagy delécio történik, aminek következtében szintén köztes allélok jelennek meg. Ezek mikrovariáns alléloktól való megkülönböztetésére ajánlott a „+” (inszerció) ill. „-” (delécio) jelölés.

3.1.4. Származás-megállapítási vizsgálatok – apasági tesztek

A személyazonosítás mellett a DNS vizsgálat rutin eljárásnak számít azokban a 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 referencia mintát adó személy gyermekétől?”.

Az apasági tesztek alapja az, hogy a *de novo* mutációtól eltekintve, a gyermek minden vizsgált genetikai lokuszon egy allél anyával és egy allél apával egyező genotípust hordoz. Ezért a biológiai apaság megállapítása annak alapján történik, hogy megfelelő számú marker vizsgálata mellett a gyermek alléljai megtalálhatók-e a vélelmezett apa genotípusai között? Ebből kifolyólag az apasági vizsgálat kimenetele egyszerűen „kizárt” vagy „nem kizárt”. Ha a vizsgált férfi apaságát nem lehet kizárni a vizsgált lokuszok alapján, akkor biostatistikai számítással kell meghatározni az apaság valószínűségét (apasági valószínűségi százalék, apasági index).

A származás-megállapítást (*paternity testing*) végző laboratóriumok gyakran ugyanazokat az STR multiplexek kitekét használják, mint a személyazonosítást (*forensic casework*) végző laboratóriumok, velük ellentétben azonban nem teljes egyezést keresnek egy DNS-profillal, hanem lokuszonként a nem definiált, azaz obligát allél feltételezett személyi eredetét vizsgálják.

Hiányos szülőség (*deficiency cases*) esetén azonban nem vizsgálható a teljes trió (vélelmezett apa, anya, gyermek). Néha az anya, máskor a vélelmezett apa nem vizsgálható, de sokszor még ilyen esetekben is megállapítható a leszármazás a vizsgálandó lokuszok számának növelésével. Elvégezhetők mitokondriális, Y-kromoszómális és X-kromoszómális vizsgálatok is, melyek nemcsak a közvetlen leszármazók, hanem a távolabbi rokonok vizsgálatára is lehetőséget adnak. Mivel az apasági és rokonsági vizsgálatok a generációk közötti genetikai kapcsolatok feltárását jelentik, számításba kell venni a mutációk lehetőségét

is. A csírvonalban történő mutációk az anya petesejtjében vagy az apa spermiumában keletkezhetnek, és a zigóta keletkezésekor adódnak tovább, így minden apasági tesztnél, eltűnt személy, illetve tömegkatasztrófa áldozatának azonosításakor számolni kell a mutáció lehetőségével is. Éppen ezért fontos a mutációs ráták pontos meghatározása, amit csak nagyszámú meiózis (biztos szülő-gyermek párok) elemzéséből lehet megbecsülni. Érdeemes megjegyezni, hogy egy laboratórium minél több lokuszt vizsgál, annál nagyobb valószínűséggel találhat mutációt a nagy számok törvényének elve alapján. Mivel napjainkban gyakran több tucat STR lokuszt is vizsgálunk, ezért egyáltalán nem ritka, hogy a gyermek és a biológiai apa között akár két-három mutáció is tapasztalható [3, 4].

Ha eltűnt személy vagy tömegkatasztrófa áldozatát kell származási vizsgálattal azonosítani, akkor fordított is lehet a kérdés, „a gyermek genotípusának ismeretében kik lehetnek a szülők?”. Többnyire ritka, hogy mindkét szülőt vizsgálni lehessen, leggyakrabban csak az egyik szülő vagy a testvérek vizsgálhatók, így a *reverz szülősségi* vizsgálat általában nagyobb kihívást jelent.

3.1.5. Bűncselekmény helyszíni minták – egyedi személyazonosítás

Az igazságügyi genetikai vizsgálatokban a nagyobb allélszámnak és magasabb heterozigótia értékeknek köszönhetően leginkább mikroszatellita STR marker lehet potenciális jelölt, mint legalkalmasabb eszköz egymással nem rokon viszonyban álló személyek közötti genetikai különbségek feltárására. Egy személy DNS állománya egy véletlenszerűen kiválasztott másik személytől nagyon nagy számú eltérést is tartalmazhat. Ha ezek az eltérések, variációk viszonylag egyszerű módon meghatározhatók, kézenfekvő eljárást jelenthetnek személyek DNS mintából való azonosítására, és annak a konkluzív megállapítására, hogy két biológiai minta származhat-e egyazon személytől. Önmagában véve egyetlen lokusz – DNS-szakasz – sem rendelkezik kizárólag egyetlen egyedre jellemző tulajdonsággal, azonban a megfelelő számú polimorf helyen megállapított allélok együttese – a genetikai profil – „genetikai személyi számként” értelmezhető. A valós ügyek és a vizsgálatba bevont minták típusa, minősége adja meg azokat a praktikus okokat, melyek jelentősen befolyásolják a vizsgálatokhoz igénybe vett markerek számát és fajtáját [13].

A tudományos és a gyakorlati törvényszéki alkalmazások között különbséget kell tenni. A kutatásban megfelelőnek és megbízhatónak tartott eljárásokat az óvatos, tartózkodó jogi szemléletmód nem minden esetben fogadja el törvényszéki szempontból is megbízhatóan alkalmazhatónak. A törvényszéki DNS vizsgálatok eredménye mindig vitatható lesz, sokszor nem is az alkalmazott technológia miatt, hanem sokkal inkább a gyakorló személy azon

rátermettsége, képessége miatt, hogy a vizsgálatokat körültekintően és helyesen vitelezze ki [3]. A DNS-profil meghatározás ugyanis nem varázspálca, nem csalhatatlan törvényszéki eszköz, hanem egy összetett vizsgálati rendszer, mely ki van téve emberi hibáknak és magában foglalja a mintaszennyezés, mintacsere, valamint a téves szakértői interpretáció lehetőségét is [8].

Az egyedhez való kapcsolhatóság miatt a DNS vizsgálatok elsősorban a személyazonosítás eszközeként jelennek meg a gyakorlatban, a korábbi vércsoport és enzim fehérjék vizsgálatokhoz képest a DNS alapú vizsgálatok számos előnnyel rendelkeznek. A DNS-molekula viszonylagos védettsége és stabilitása a technikai előnyökkel párosulva sok esetben a természetes bomlási folyamatok ellenére is megoldást jelenthet. A nemi kromoszómák markerei, így az Y-kromoszóma, mint a humán genom kizárólag férfiakra jellemző szakasza az erőszakos bűncselekmények férfi elkövetőinek azonosítását tette még eredményesebbé. Ez leginkább azokban az eljárásokban alkalmazható, melyekben a kevert jellegű biológiai nyomok női eredetű sejtjeinek túlsúlya miatt az autoszómális markerek analízise nem informatív. A mitokondriális DNS hipervariábilis szakaszainak PCR alapú szekvencia-analízisével a DNS azonosítást a biológiai anyagmaradványok teljes körére kiterjesztették [14]. Az igazságügyi szempontból fontos megkülönböztetési – kizárási esélyerőnek részben a mitokondriális DNS maternális öröklődése miatti alacsonyabb voltát a sejtenkénti magasabb kópiaszám ellensúlyozza, ami szignifikánsan növeli a vizsgálatok sikerességi rátáját. A PCR technika sajátosságainak megfelelően rövid – kb. 100-150 bázispár – szakaszon belül is polimorf markerekkel vagy az amplikonok méretének redukálásával [3, 4] növelhető a vizsgálati érzékenység. Az extrémén kis mennyiségű DNS-t tartalmazó minták (LCN: low copy number DNA”) vizsgálhatósága a mikronyomok hagyományos maternális értelmezését kimutatási érzékenység és a származás vonatkozásában is kiterjesztette [13].

A DNS vizsgálatok eredménye kétféle módon használható fel büntető- vagy törvényszéki eljárások során. Egyrészt a bírósági szakaszban a bíró az igazságügyi genetikus szakértő véleményét, mint bizonyítékot értékeli, a vád és a védelem pedig igyekszik az eredmények minden aspektusát figyelembe véve részletekbe menően megvizsgálni. Másrészt, ha a DNS bizonyítékokat a büntetőeljárás során csak nyomozati információként használják fel, akkor azok nem feltétlenül jutnak el a tárgyaló terembe (pl. bűnügyi DNS nyilvántartás adatai), mivel az ügyben rendelkezésre állhatnak egyéb, nem DNS alapú közvetlen bizonyítékok is. Mindazonáltal minden esetben figyelembe kell venni azt a tényt, hogy egy helyszíni DNS mintával tökéletesen megegyező gyanúsított DNS-profil nem feltétlenül

jelenti a gyanúsított bűnösségét az adott ügyben, mivel az csak annyit demonstrál, hogy a kérdéses személy valamikor valamilyen kapcsolatban volt a helyszínnel [13].

3.1.6. A törvényszéki DNS vizsgálatok statisztikai interpretációja

Az igazságügyi genetikus szakértő tevékenysége során valószínűsítő szakértői véleményt ad, ha nincsenek kizáró kombinációk az összehasonlító minták között (pl. a bűncselekmény helyszínéről származó és a gyanúsítottól vett minta). A DNS bizonyíték statisztikai kiértékelésére a gyakoriságbecsléssel kombinált Bayes-elvű hipotézistesztelés a gyakorlatban legáltalánosabban elfogadott módszer [10]. A valószínűségszámítás Bayes-tételének esély formájú leírása szerint az alapkérdés (*hipotézis*) a következő:

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

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

- „mekkora annak a valószínűsége, 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* (H_p) és *ellenhipotézis* (H_d) teljesülésének feltételezésével számított valószínűségek arányát mint *valószínűségi hányadost* (LR, *Likelihood Ratio*) adjuk meg az igazságügyi DNS vizsgálatok statisztikai interpretációja során.

$$LR = \frac{H_p}{H_d}$$

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-profil egyezési 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.

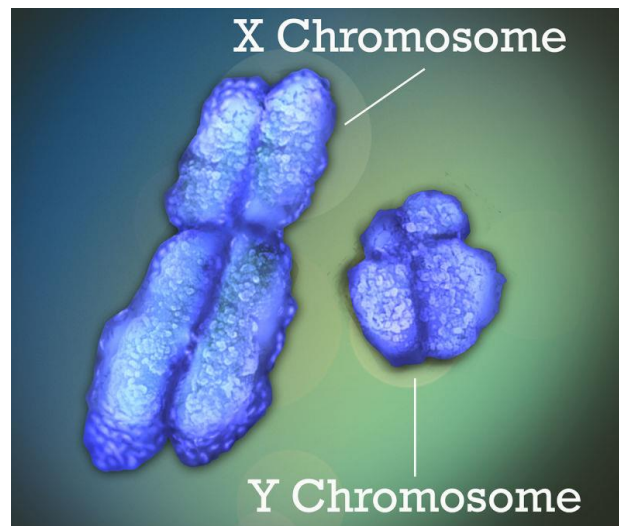
Amennyiben azonban számottevő az ún. *linkage disequilibrium* (LD), továbbá, ha populáció strukturálódás figyelhető meg a populációban a számított profilgyakoriságok tévesek lehetnek.

Felmerül a kérdés, lehetséges-e, hogy nem a gyanúsítottak, hanem egy közeli hozzátartozójának a DNS mintája került kimutatásra? Ez a kérdés azért merülhet fel, mert hozzátartozók nagyobb valószínűséggel rendelkeznek a gyanúsítottéval hasonló (ritkán megegyező, ha egypetűjű ikertestvére van) genotípussal, mint a *random* módon kiválasztott személyek. Ezzel a lehetőséggel is számolni kell az eredmények statisztikai interpretációja során. Fontos azt is hangsúlyozni, hogy az *ellenhipotézis* valószínűségi értékének számítása csak akkor korrekt, ha a statisztikai interpretáláshoz felhasznált populációs allél- vagy profilgyakorisági adatbázis típusát nem a terhelt személy etnikai hovatartozása, hanem a nyomozati információ (pl. a bűncselekmény elkövetésének helye, a sértett ill. a szemtanúk vallomása) határozza meg. A jogban az ártatlanság vélelme alkotmányos alapelv, miszerint senki sem tekinthető bűnösnek mindaddig, amíg a büntetőjogi felelősségét a bíróság jogerős határozata nem állapította meg. Ezért a szakértőnek a statisztikai interpretáció készítése során a gyanúsítottra nézve kevésbé terhelő LR értéket szükséges megadni. A napjainkban rutinszerűen alkalmazott multiplex STR kitek kb. $1:10^{10}$ - 10^{25} egyezési valószínűséget adnak.

3.1.7. Ivari kromoszómák

3.1.7.1. Az X-kromoszómás STR markerek

Az X-kromoszóma az emlősöket és az embert is beleértve, a kétféle ivart meghatározó, azaz szex kromoszóma egyike (a másik az Y-kromoszóma). Férfiak egy X kromoszómát, nők pedig két X kromoszómát hordoznak. Az X- és Y-kromoszómákat nemi kromoszómáknak szokás tekinteni (**2. ábra**), mivel majdnem tökéletesen meghatározzák az adott személy nemét. Az édesanyjától mindenki X kromoszómát örököl. Abban az esetben, ha az apjától is X-et kap, nőnemű, ha Y-t, hímnemű lesz az utód. Előfordulnak ritka kivételek, amelyek nem tárgyai a jelen értekezésnek.



2. **ábra.** A nemi kromoszómák²

Az ábrán a humán X- és Y-kromoszóma méretbeli különbsége látható.

Az X-kromoszóma 153 millió bázispárból, míg az Y-kromoszóma kb. 60 millió bázispárból áll. Így egyrészt a méretbeli különbségénél fogva, másrészt amiatt, hogy az X- és Y-kromoszóma legnagyobb részén – kivéve PAR1 és PAR2 régiót a kromoszómák két végén – egymással nem homológ DNS szekvenciák találhatók, az Y kromoszóma ún. *NRY* (Non-Recombining region Y) régiója rekombinációtól mentes marad. Ennek köszönhető az is, hogy az X-kromoszóma java része az igazságügyi genetikai gyakorlatban bizonyos típusú vizsgálatra – pl. hiányos apasági vizsgálatok – alkalmas.

Az X-kromoszómás markerek vizsgálata alkalmazható az egészségügyi gyakorlatban is. Ilyen esetekben az X-kromoszómához kapcsolódó betegség génjéhez közel lévő markerek vizsgálatával nyomon követhető a betegség családon belüli öröklődése, valamint megállapítható, hogy a születendő gyermek hordozza-e a betegséget okozó allélt, illetve az, hogy az egészséges családtagok közül ki hordozó és ki nem.

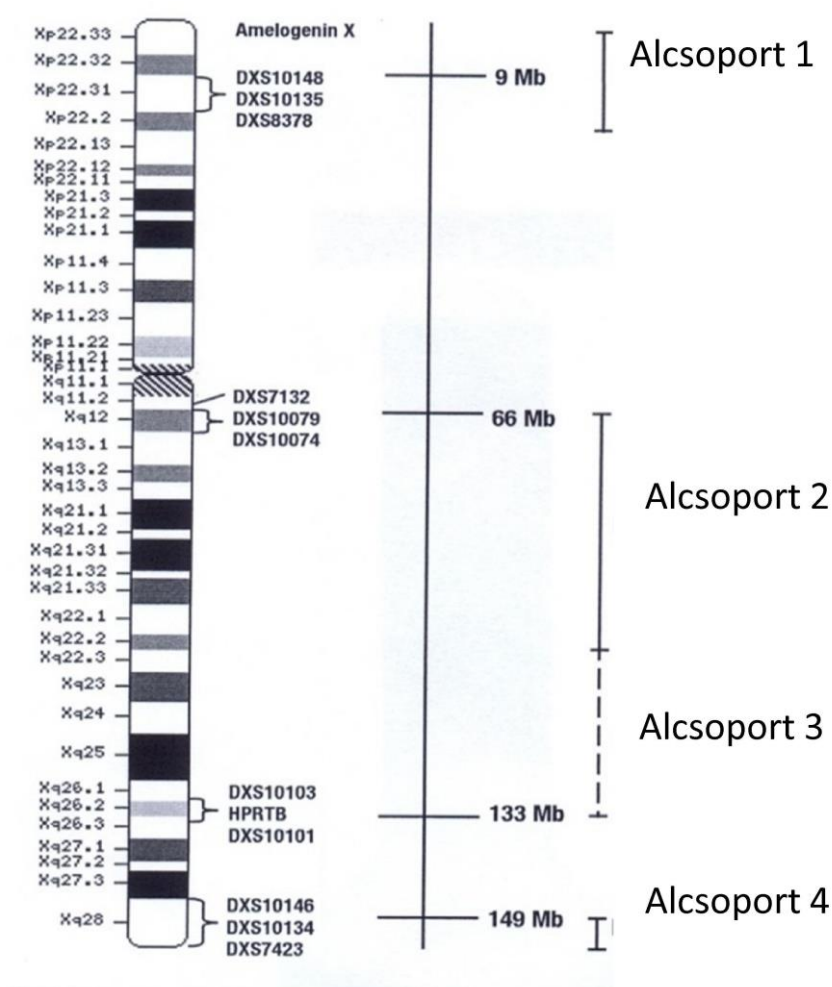
Standardizált vizsgálati vegyszercsomagok (kit) használatával lehetségessé vált már az X-STR lokuszok [15, 16, 17, 18, 19, 20] vizsgálata az igazságügyi genetikai gyakorlatban. A jelenleg igazságügyi genetikai célokra forgalmazott X-STR lokuszok vizsgálatára alkalmas kitben a lokuszok négy különböző alcsoportban (cluster) helyezkednek el, mindegyik alcsoporton belül 3-3 szorosan kapcsolt lokusz helyezkedik el, amelyek úgy viselkednek, mintha haplotípusok lennének. Nőkben a lokusz-triókon belül a rekombináció valószínűsége

² <http://gizmodo.com/the-y-chromosome-is-tiny-how-did-it-get-that-way-1702729542>

igen csekély, míg közöttük (kapcsoltsági alcsoportok között) lehetséges és gyakori a rekombináció, így az öröklődésük majdnem „függetlennek” tekinthető (**3. ábra**). A 3. ábrán látható a 4 kapcsoltsági alcsoport elhelyezkedése: alcsoport 1, alcsoport 2, alcsoport 3 és alcsoport 4.

Több tanulmány született X-kromoszómás rekombináció vizsgálat témában 3 generációs pedigre analízis segítségével [21, 22, 23, 24], amelyek igazolják a kapcsoltsági csoportok közötti gyakori rekombinációkat.

Az igazságügyi genetikában alkalmazott X-STR lokuszokra vonatkozó legfontosabb információk az **1. táblázatban** láthatók [24].



3. ábra. Az X-STR lokuszok kapcsoltsága³

Az ábra a vizsgálatban alkalmazott 4 alcsoporton belül szorosan kapcsolt lokusz trió nevét és helyét mutatja.

³ Investigator Agus X-12 QS handbook (módosítva)

1. táblázat. Az X-STR lokusz triók és lokalizációik

Marker	Kapcsoltsági alcsoport	Citogenetikai lokalizáció	Fizikai lokalizáció (Mb)	Genetikai lokalizáció (cM)
DXS10148	1	Xp22.31	9.198	19.84 ^a
DXS10135		Xp22.21	9.199	20.03 ^a
DXS8378		Xp22.31	9.330	20.20 ^b
DXS7132	2	Xcen	64.572	90.75 ^b
DXS10079		Xp12	66.632	90.82 ^a
DXS10074		Xp12	66.894	90.83 ^a
DXS10103	3	Xq26.2	133.246	149.37 ^a
HPRTB		Xq26.2	133.443	149.66 ^b
DXS10101		Xq26.2	133.482	149.75 ^a
DXS10146	4	Xq28	149.335	183.72 ^a
DXS10134		Xq28	149.401	183.96 ^a
DXS7423		Xq28	149.460	184.19 ^a

^a A megfelelő fizikai marker pozíciójából Rutgers Map Interpolator segítségével kiszámolva (<http://compgen.rutgers.edu/old/map-interpolator>)

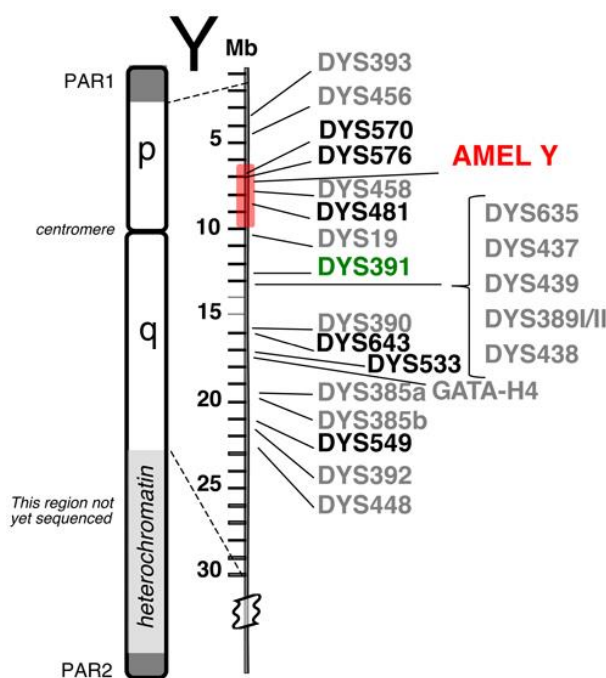
^b Az aktuális rekombinációs gyakoriság becsléséből Rutgers Map v.2 Kosambi mapping funkció segítségével származtatott értékek [25].

Rekombinációról csak akkor beszélünk, ha két X-kromoszóma van, vagyis csak nők esetében fordul elő. A rekombináció vizsgálatát célszerű olyan háromgenerációs családokban elvégezni, ahol nagyapák, leány-gyermekek és fiúunokák vannak. A nagyapában lévő X-kromoszóma egyben (haplotípus) adódik át leány gyermekébe és a leánygyermekben történik a rekombináció két X-kromoszóma között, amelyet a következő generációban, a fiúgyermekéknél lehet megfigyelni és megszámolni. Éppen ezért az X-STR markerek rokonsági vizsgálatokban való alkalmazásaihoz a nők két X-kromoszómája között történő rekombináció gyakoriságának pontos megismerése nélkülözhetetlen. A fent hivatkozott rekombináció vizsgálatok a következő rekombinációs gyakoriságokat mutatták ki: az 1. és 2. kapcsoltsági alcsoport között 0.387-0.5; a 2. és 3. kapcsoltsági alcsoport között 0.4-0.5 és 3. és 4. kapcsoltsági alcsoport között 0.25-0.367.

Az X-STR vizsgálatok rokonsági vizsgálatokban való alkalmazásához a kapcsoltsági alcsoporton belüli lokusz-triók előfordulási gyakoriságait (3 lokuszból álló haplotípus), valamint az 12 X-STR lokuszból álló egész X-kromoszómára vonatkozó haplotípusok gyakoriságait férfiakban fel kell mérni [18, 21, 22, 26].

3.1.7.2. Az Y-kromoszómás STR markerek

Az igazságügyi genetikában az Y-kromoszómán leginkább az STR lokuszok állnak az alkalmazás fókuszában [3], a rutinszerűen vizsgált Y-kromoszómás STR (Y-STR) lokuszok száma megközelíti az autoszómális STR-ekét. Ma már 23-27 igazságügyi alkalmazásra validált Y-STR-t tartalmazó kiteket forgalmaznak a gyártók személyazonosítás céljából (Promega, Life Technologies). Az Y-STR lokuszokon az egyes allélok nem véletlenszerű kombinációban vannak jelen. A kapcsolt allélok specifikus kombinációját haplotípusnak nevezzük [27]. Az Y-kromoszóma csak a férfiakban található, azért Y-kromoszómás STR-ekkel egy férfi haplotípusa olyan kevert mintákból is meghatározható, amelyekből a női sejtek túlnyomó többsége miatt az autoszómális STR lokuszok vizsgálata nem ad informatív eredményt [28]. A 4. ábra mutatja a szakértői rutinban leggyakrabban használt Powerplex Y23 kit összetételét (23 STR lokusz) és a lokuszok lokalizációját.



4. ábra. Az Y-STR lokuszok és lokalizációik⁴

Az ábra a rutin vizsgálatban alkalmazott 23 STR lokusz nevét és az Y-kromoszómán lévő elhelyezkedési sorrendjét mutatja.

A humán populációk és az evolúció történetének (human population and evolution history) felmérésénél, ha az Y-kromoszóma nyomon követése a cél, az autoszómás

⁴<https://worldwide.promega.com/resources/profiles-in-dna/2012/variability-of-new-str-loci-and-kits-in-us-population-groups/>

lokuszokkal ellentétben nem az allélgyakoriságot célszerű alkalmazni, hanem a haplotípus gyakoriságát [29, 30, 31]. A későbbiekben a filogenetikai vizsgálatokat tartalmazó fejezetben ez bővebb kifejtésre kerül.

Egy adott Y-STR haplotípust azonban a fentieknek megfelelően nem csak egyetlen személy hordozhat, hanem a kérdéses személy minden férfirokona apai ágon, így a gyanúsított vagy vélelmezett apa édesapja, nagyapja, fiútestvére, nagybátyja, fia, fiú unokája, fiú unokatestvére, stb. szintén ugyanazzal az Y-kromoszómás STR haplotípussal rendelkezik.

Az Y-STR lokuszok igazságügyi genetikában betöltött szerepét mutatja az is, hogy a 2000-ben létrejött nemzetközi Y-STR haplotípus referencia adatbázis mára (YHRD: Y Chromosomal Haplotype Reference Database, 2017/Oct/20, Release 55) közel 200.000 férfi haplotípusát tartalmazza különböző lokusz készlettel (www.yhrd.org).

A nemzetközi szakmai grémium kötelessége, hogy az YHRD részére minden vizsgáló intézmény átadja az adott populációban felmért Y-haplotípus adatokat, ami által a világ népességét minél jobban reprezentáló haplotípus adatbázis áll majd rendelkezésre, amely segíti a szakértői vélemények statisztikai interpretációját. Az igazságügyi genetikában, amennyiben valamely bűncselekményhez kapcsolható minta Y-STR haplotípusa és a gyanúsított haplotípusa megegyezik, úgy a haplotípus gyakorisága az említett nemzetközi Y-STR adatbázisban keresendő, mivel nemzeti adatbázisok általában nem léteznek. Ugyanúgy alkalmazható a haplotípus gyakoriság megkeresése a hiányos apasági ügyekben, ahol eltűnt vagy elhunyt a vélelmezett apa. Ilyen esetekben a vélelmezett apa valamelyik apai ági férfi rokona vizsgálható a perben szereplő fiúgyermekkel együtt. Az LR vagy PI számítása a keresett haplotípus gyakorisága alapján számolható ki. Ahogy korábban említésre került, a haplotípus vizsgálattal egy férfi személye kizárólagos jelleggel nem azonosítható, mivel családon belül apai ágon minden férfi azonos lesz, viszont a nyomozóhatóság munkája az adott családi csoportra szűkíthető.

3.1.7.3. Az Y-kromoszómás SNP markerek

Az Y-SNP lokuszok (haplocsoport) vizsgálata és a vizsgált populációk Y-kromoszómás 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 vizsgálata adhat információt az illető férfi földrajzi származásáról, amely a nyomozó hatóság munkáját segítheti. Az STR lokuszok mellett az Y-kromoszómás SNP lokuszok populációs és evolúciós tanulmányok informatív eszközei is, vagyis az igazságügyi genetikai vizsgálatokban mutatkozó hátrányaik előnynek bizonyulnak az evolúciós vizsgálatok szempontjából. Az Y-SNP lokuszok fontos szerepet játszanak a humán migrációval foglalkozó tanulmányokban, mivel lehetővé teszik az egyes populációk, illetve populáció-csoportok közötti lényeges különbségek hatékony feltárását [3].

Az Y-SNP allélok tipikusan „ősinek” („ancestral”) és „leszármazottnak” („derived”) nevezik, és egyszerű bináris formában tartják nyilván (0 = ősi, 1 = leszármazott). Az ősi allél megállapítása a csimpánz – az ember legközelebbi rokona – Y-kromoszóma DNS szekvenciájával való összehasonlítás alapján történt. Az SNP lokuszok előnyét stabilitásuk adja, mutációs rátájuk alacsony – kb. 10^{-8} mutáció/nukleotid/generáció [3] –, ellentétben az Y-STR lokuszok – $6,9 \times 10^{-4}$ mutáció/lokusz/generáció [32] – mutációs rátájával (evolúciós mutáció ráta), éppen ezért hosszabb idő távlatában is alkalmasak az apai leszármazási vonalak nyomon követésére.

A „haplocsoport” és „haplotípus” kifejezéseknek több, egymással átfedő definíciója létezik. A haplocsoportot bináris polimorfizmusok által meghatározott leszármazási vonalnak (NRY, Non-recombining Region of Y) is tekinthetjük. A „haplotípus” megnevezést azokra az egy haplocsoportba tartozó leszármazási vonalakra alkalmazzák, melyeket az NRY régióban elhelyezkedő STR lokuszok (Y-STR) variációi határoznak meg [33]. Az Y-kromoszómás SNP markerek alapján filogenetikai fa állítható fel, aminek a segítségével nyomon követhető az egyes haplocsoportok leszármazási sorrendje. A haplocsoportokat az „ABC” betűivel, az alcsoportokat a finomabb felbontás érdekében számok és betűk kombinációjával jelölik (pl. R1b1). A vizsgálati módszer fejlődését mutatja, hogy míg 2002-ben 243 SNP markerből és 153 haplocsoportból (Y Chromosome Consortium, 2002), 2008-ban már 599 markerből és 311 haplocsoportból [34] állt a filogenetikai fa és ez a fejlődés azóta is folyamatos. Mára az új

generációs szekvenálásnak (NGS) köszönhetően az alcsoportok bontása, új alcsoportok felfedezése szinte követhetlenné vált (1000 Genome Project, FTDNA, ISOGG) [35, 36, 37]. A haplocsoportok eredete emberi léptekben mérve nagyon régi idősakra, több ezer, esetenként több tíz- vagy százezer évvel ezelőtre nyúlik vissza. Ahogy az alcsoportok földrajzi specifikussága felszínre kerül, az alcsoportok kora a jelenkorhoz közeledik.

A haplocsoportok meghatározásának két módja is van. Egyik az, hogy a filogenetikai fa elejéről elindulva „végigsétálunk” a fa ágai mentén, amíg meg nem találjuk a mintához tartozó haplocsoportot. Ez idő- és költségigényes folyamat. A másik módszer az „út levágása” (shortcut), ez akkor alkalmazható, ha van előfeltevés arra vonatkozóan, melyik haplocsoportba tartozhat a vizsgált személy. Ebben az esetben csak a filogenetikai fa ágának végén lévő – azaz a haplocsoportot meghatározó – SNP markert vizsgálják.

A minta haplocsoportjára vonatkozó előfeltevést kétféle módon lehet felállítani, mindkét esetben a minta előzetes Y-STR haplotípus meghatározására van szükség. Az első módszer feltételezi, hogy már vannak minták, amelynek haplotípusa és haplocsoportja ismert. A vizsgálandó minta haplotípusát össze kell vetni a már ismert haplocsoportú haplotípusokkal és teljes, vagy majdnem teljes (egy- vagy kétlépéses mutáció egy vagy két lokuszon) egyezést keresni. Ha nem található ilyen egyezés, akkor online elérhető, „haplocsoport jósló” programot érdemes használni. Jelenleg három ilyen programot lehet igénybe venni: Y-DNA Haplogroup Predictor - NEVGEN (<http://www.nevgen.org/>), World Haplogroup & Haplo-I Subclade Predictor (<http://members.bex.net/jtcullen515/haplotest.htm>) és With Athey's Haplogroup Predictor (Athey, 2006). A három program a haplotípusok többségéhez nagy valószínűséggel tud haplocsoportot rendelni. Néhány esetben ezek a programok sem adnak megbízható előrejelzést a haplocsoportot illetően - főleg ázsiai eredetű haplocsoportok esetén - ilyenkor a már elsőnek említett megoldást kell alkalmazni, azaz „végig kell sétálni” a fa ágai mentén.

Az igazságügyi genetika művelői bizonyos értelemben szerencsés helyzetben vannak, mivel mindkét diszciplína – igazságügyi genetika és filogenetika – vizsgálati eszközei azonosak, így módon a tudomány területekből nemcsak gyakorlati, hanem elméleti szinten is „csemegézhetnek”.

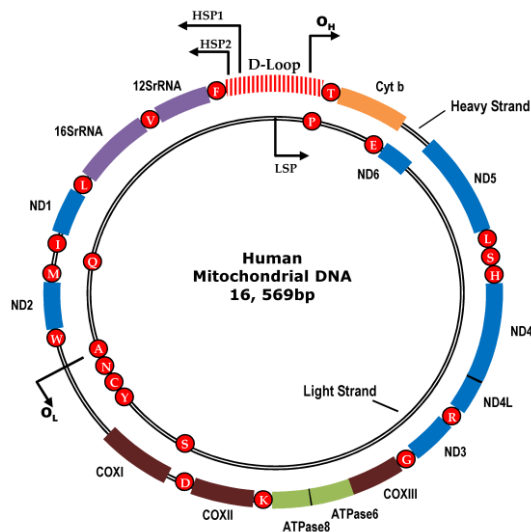
A téma bővebb interpretálása az Östörténet fejezetben található.

3.1.8. Az anyai örökség

3.1.8.1. A mitokondriális DNS szerkezete

A humán mitokondriális genom DNS szekvenciáját először 1981-ben írták le [38], és az első leírójának neve, illetve a laboratórium helyszíne után szokás Anderson vagy Cambridge referencia szekvenciaként (CRS) emlegetni. A humán mitokondriális genom egy kb. 16569 bp hosszú cirkuláris DNS-molekula, ami egy pirimidin bázisokban gazdag „könnyű” és egy túlnyomórészt purin bázisokból álló „nehéz” szálból áll (5. ábra). A molekula 15447 bp hosszú kódoló szakaszán 37 gén lokalizálódik, intronok nélkül összetömörítve. A 37 génből 22 gén transzfer RNS-t kódol (tRNS), 2 gén riboszómális RNS-t (12S és 16S rRNS), 13 gén pedig az oxidatív foszforilációban és ATP-szintézisben közreműködő enzimeket. A mitokondriumok normális működéséhez szükséges összes többi gén a sejtmagi DNS-ben lokalizálódik. Később, 1999-ben Andrews és munkatársai [39] újra elemezték a szekvenciát, amely azóta a rCRS (revised Cambridge Reference Sequence) néven lett ismert.

2012-ben Behar és munkatársai javasolták [40], hogy használják a RSRS (Reconstructed Sapiens Reference Sequence) jelölést a rCRS helyett, ahol a referencia szekvencia egy valaha élt afrikai nő mtDNS-e. A ma élő összes nő származtatható az említett afrikai ősanától, azaz a „mitokondriális Évától”. Ennek ellenére az RSRS használatának indokolt volta még vitatott.



5. ábra. A humán mtDNS vázlatos szerkezete és a gének egyszerű lokalizációja⁵
Az ábrán látható D-loop régió (hipervariábilis régió) szekvenciájának a rCRS szekvenciától való eltérése adja meg a haplotípust igazságügyi genetikai szempontból.

⁵<https://www.intechopen.com/books/molecular-considerations-and-evolving-surgical-management-issues-in-the-treatment-of-patients-with-a-brain-tumor/understanding-mitochondrial-dna-in-brain-tumorigenesis>

A molekulában természetes variációként megfigyelhető egy-két bázist érintő inszerciós/deléciós mutációk miatt a molekula összes nukleotidának száma egyénenként, sőt sejtenként, molekulánként is (hossz-heteroplazmia) változhat. A mitokondriális DNS-molekulát a kódoló régió kívül egy kb. 1122 bázispár hosszúságú nem kódoló szakasz az ún. **hipervariábilis régió** alkotja. A hipervariábilis régiót gyakran „D-loop” szakasznak is nevezik, utalva a DNS replikáció során elektronmikroszkóp alatt látható képére. Mivel a D-loop nem kódol génterméket, kisebb rajta a szelekciós nyomás, ezért az egyedek között megfigyelt nukleotid variabilitás és polimorfizmus lényegesen nagyobb fokú itt, mint a kódoló régióban. A standard referencia szekvencia sorszámozása önkényesen a hipervariábilis régió közepén kezdődik kettéválasztva azt egy 16024-16569 bázisok közötti és egy 1-576 bázisok közötti szakaszra. Az igazságügyben általánosan vizsgált szekvenciák az ún. **hipervariábilis régiók** is ezekre a szakaszokra esnek: a hipervariábilis 1. régió (HV1) a 16024-16365 szakaszra, a hipervariábilis 2. régió (HV2) a 73-340 bázistartományra terjed ki (**5. ábra**), bár ezek a határok nem szigorúan meghatározottak, függenek az alkalmazott vizsgálati módszertől és a vizsgáló laboratóriumtól. A mitokondriális DNS variánsok (haplotípusok) megállapítása és leírása a revised Cambridge referencia szekvencia (rCRS) könnyű szálához viszonyítva történik, így pl. a 16126 nukleotid pozícióban észlelt citozin bázist 16126C jelzéssel írjuk le (vagy T16126C). Magától értetődő, hogy a referencia szekvenciától eltérést nem mutató nukleotid bázisokat nem kell jelölni.

3.1.8.2. A mitokondriális DNS vizsgálat jelentősége az igazságügyi genetikában

Az elmúlt években a mitokondriális DNS (mtDNS) analízis igazságügyi célú alkalmazása és jelentősége rendkívüli mértékben megnőtt, mivel olyan előnyös tulajdonságokkal is rendelkezik, melyekkel a nukleáris (sejtmagi) DNS állomány nem bír. A mtDNS átlagosan több ezer példányban található meg egy-egy szomatikus sejtben [41], a mitokondriumok kettős membránnal körül vannak véve, ezért a benne lévő DNS degradáció hatásától védett, ami különösen fontos az igazságügyi genetikai gyakorlatban.

Magasabb kópiaszámának köszönhetően vizsgálati érzékenysége nagyságrendekkel nagyobb a nukleáris DNS-hez képest, és ennél fogva a nagymértékben degradálódott, bomlott állapotú biológiai mintákból (pl. csontmaradványok, szőrképletek) való azonosításra is alkalmas. Kizárólag anyai ágon öröklődik [42, 43, 44], nem rekombinálódik [42, 45], mely tulajdonságok jelentősen megnövelik a biológiai anyagmaradványok beazonosításához

referencia mintaadóként számításba vehető személyek számát, hiszen az összes élő anyai ágon rokon személy ugyanazzal a mitokondriális DNS haplotípussal rendelkezik az esetleges mutációs eseményeket nem számítva. Ez az oka annak is, hogy egy mitokondriális DNS haplotípus nem egyedi, ami – az az Y-STR haplotípusokhoz hasonlóan – az egyezési valószínűségek mértékét csökkenti. A sejtmagi DNS állományhoz képest a mitokondrium kevésbé hatékony DNS javító mechanizmussal rendelkezik, ami magasabb mutációs rátát eredményez. A mitokondriális DNS replikációját végző polimeráz enzim pontos leolvasásbeli (proof-reading) hiányossága miatt nő a lehetséges mutációk gyakorisága. A sejtmagi DNS-nél kb. tízszer nagyobb mutációs ráta eredményeképpen jöhetnek létre a különböző variációk az egyébként azonos anyai vonalakból [13]. Cirkuláris jellege miatt viszont a molekula kevésbé fogékony exonukleázok támadása iránt, ami segít megőrizni kompakt állapotát és ezzel együtt vizsgálhatóságát, azonosíthatóságát.

A mtDNS haplotípusok pontos meghatározása nemzetközileg egységesített nevezéktan segítségével történik. Az igazságügyi genetikai vizsgálatok során kapott DNS szekvenciák leírásának nevezéktanát, a vizsgálatok végzéséhez elengedhetetlen minőségirányítási követelményeket, a szekvencia eredmények interpretálásának szabályait az ISFG (International Society for Forensic Genetics) és az EDNAP (European DNA Profiling Group) ajánlásokban, útmutatókban összegezte [46].

Igazságügyi aspektusból a heteroplazmia jelensége az egyik leginkább problémás vonása a mitokondriális genomnak. Ugyanis a mitokondriális DNS állomány egyénen belül is diverzifikálódhat, ugyanis az egyes DNS-molekulák a sejtekben egymástól függetlenül replikálódnak (nincsenek kitéve a mitózisnak és a meiózisnak), egymással nem rekombinálódnak, és az osztódásuk során a mutációs ráta lényegesen nagyobb, mint a nukleáris genoménál. Kettő (vagy több) mitokondriális DNS típus szubpopulációjának jelenlétét egy egyénen belül heteroplazmiának nevezzük. Ezek a szekvencia variánsok egymástól függetlenül

replikálódnak és szegregálódnak. Bizonyos esetben egy-egy bizonyíték erősségét is növelheti, mint például az első, igazságügyben dokumentált pontmutációs heteroplazmia esetén [47] II. Miklós cár és családja földi maradványainak azonosítása során. A szubsztitúciós heteroplazmiákon kívül gyakran megfigyelhető a policitozin szálak (mind a HV1 mind a HV2 régióban) hossz-heteroplazmiája is [48].

Az igazságügyi genetikán kívül számos további tudományág tanulmányozza a mitokondriális DNS variációkat. Orvosi genetikával foglalkozó szakemberek számos betegség és kórkép okaként írták le a mtDNS valamilyen mutációját [49]. Evolúciobiológusok

és filogenetikusok fajok közötti és fajokon belüli variációk vizsgálatával tárják fel a rokonsági viszonyokat. A széles nyilvánosság számára is izgalmas eredmények születtek pl. a neandervölgyi ember maradványainak vizsgálata során [50], Ötzi mtDNS genomjának feltárása [51] vagy a „mitokondriális Éva” meghatározása [42, 52, 53] során.

Az igazságügyi célú mtDNS vizsgálat a célnak megfelelő populációs adatbázisokat igényel, amelyekkel megbecsülhető a kérdéses DNS-profilok (haplotípusok) gyakorisága az adott populációban. Ha egy eset kapcsán a bizonyíték és a gyanúsított mtDNS-haplotípusa megegyezik, a megfelelő referencia adatbázis alapján kalkulált haplotípus gyakoriság konzervatív becslését adja azon valószínűségnek, hogy a bizonyítékból kimutatott haplotípus a gyanúsítottól, vagy a populáció egy véletlen személyétől származik. Az elmúlt évek populációgenetikai kutatásai jól meghatározott különbségeket tártak fel a különféle eredetű populációk mtDNS haplotípus megoszlásában és ez az oka annak, hogy adott bizonyíték esetén a konkrét ügy egyéb körülményei is befolyásolják az alkalmazandó referencia adatbázist. Walther Parson jóvoltából, aki az innsbrucki egyetem professzora, létezik egy mtDNS adatbázis (<http://empop.online/>), amelynek segítségével kereshető a vizsgált haplotípus gyakorisága és annak alapján számolható valószínűségi hányados a vizsgált személy tekintetében [54, 55].

3.1.8.3. Mitokondriális haplocsoport meghatározás

Az mtDNS haplocsoport vizsgálata és a vizsgált populációk haplocsoport-összetétele igazságügyi genetikai szempontból – akárcsak az Y kromoszómális haplocsoportok esetében – nem bír annyira meghatározó jelentőséggel, mivel az mtDNS haplocsoportok száma az mtDNS haplotípusok számánál jóval kisebb. Egy mitokondriális haplocsoportba sok, különböző mitokondriális haplotípus tartozik, ezért a haplocsoportok vizsgálata leginkább kizárás megállapítására alkalmas. Egyezés esetén megfelelően magas valószínűségi hányados sem származási, sem bűnügyekben nem szolgáltatható. A mtDNS haplocsoportok mind a hipervariábilis régióban, mind a kódoló régióban egy-egy meghatározott pozícióban található nukleotidok jelenlétének összességével definiálhatók. A mtDNS haplocsoportok meghatározása informatív eszközként szolgál populációs és evolúciós tanulmányok során. Ezek fontos szerepet játszanak a humán migrációval foglalkozó tanulmányokban, lehetővé teszik az egyes populációk, illetve populáció-csoportok közötti lényeges különbségek hatékony feltárását [3], a populációk történetének, vándorlásának vizsgálatát.

3.2. ÖSTÖRTÉNET ÉS AZ EMBER GENETIKAI ÚTJA

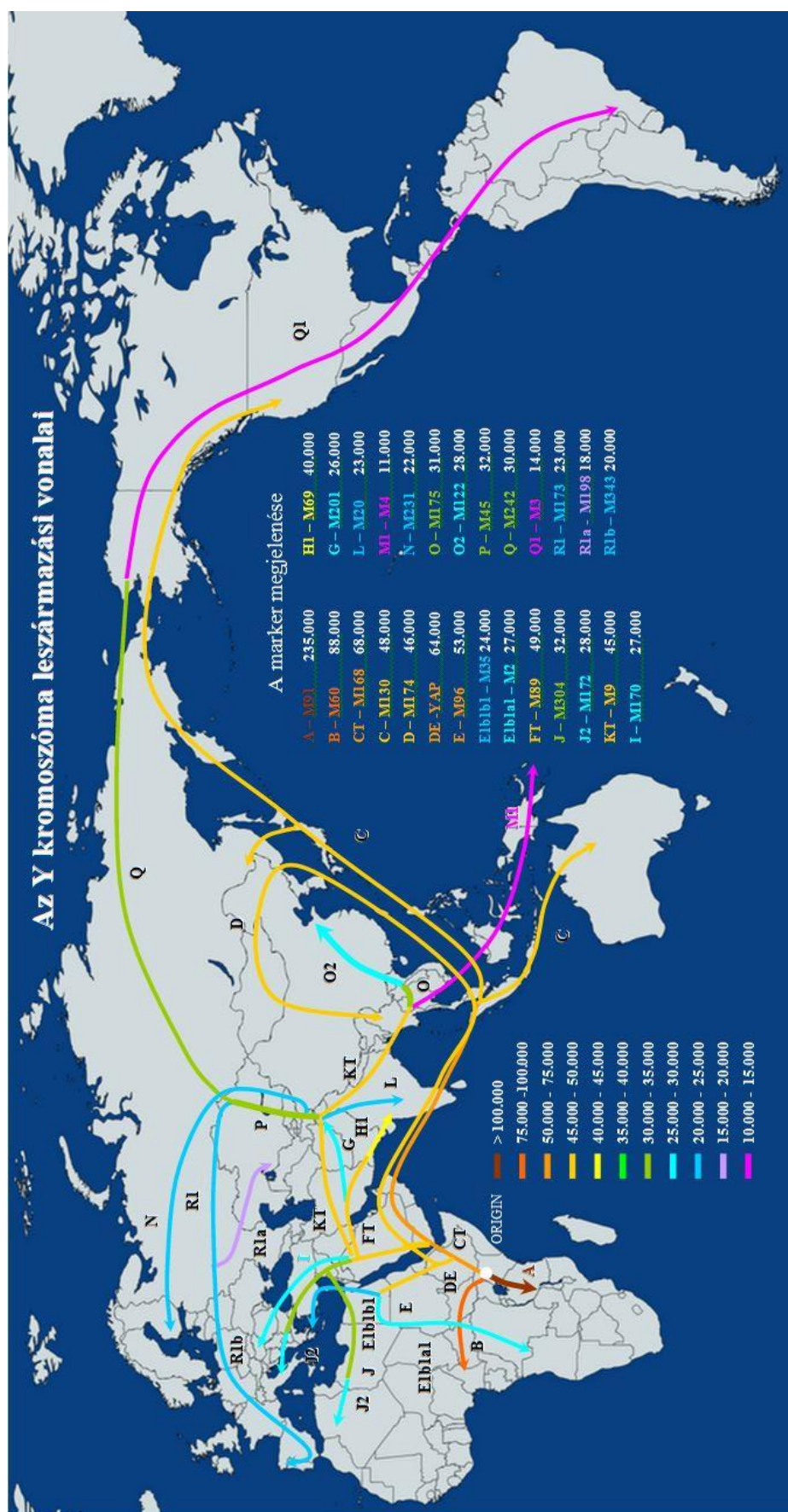
Richard Dawkins: Mindenki, akire csak pillantásuk esik, mikor kilép az időgépből, vagy egyetemes őse az emberiségnek, vagy senkinek sem őse.

3.2.1. A modern ember kirajzása Afrikából („Out of Africa” theory)

Manapság már vitathatatlan tény, hogy a modern ember (*Homo sapiens sapiens*) mintegy 60 ezer évvel ezelőtt hagyta el Afrikát – az emberiség bölcsőjét – és röpké 5-6 ezer év alatt megjárta a földkerekség minden zegét-zugát [56]. Egyes csoportok letelepedtek Arábiában, mások dél felé vándoroltak átszelve a mai Irak, Irán és Afganisztán területét és eljutottak Közép-Ázsiába, nagy valószínűséggel ugyan ebből a csoportból származó emberek eljutottak a Csendes-óceán szigetvilágába, valamint Ausztráliába. A harmadik csoport Európa felé vette az útját. A korai ázsiai lakók egy része Európát úgy 30-45 ezer éve kezdte benépesíteni.

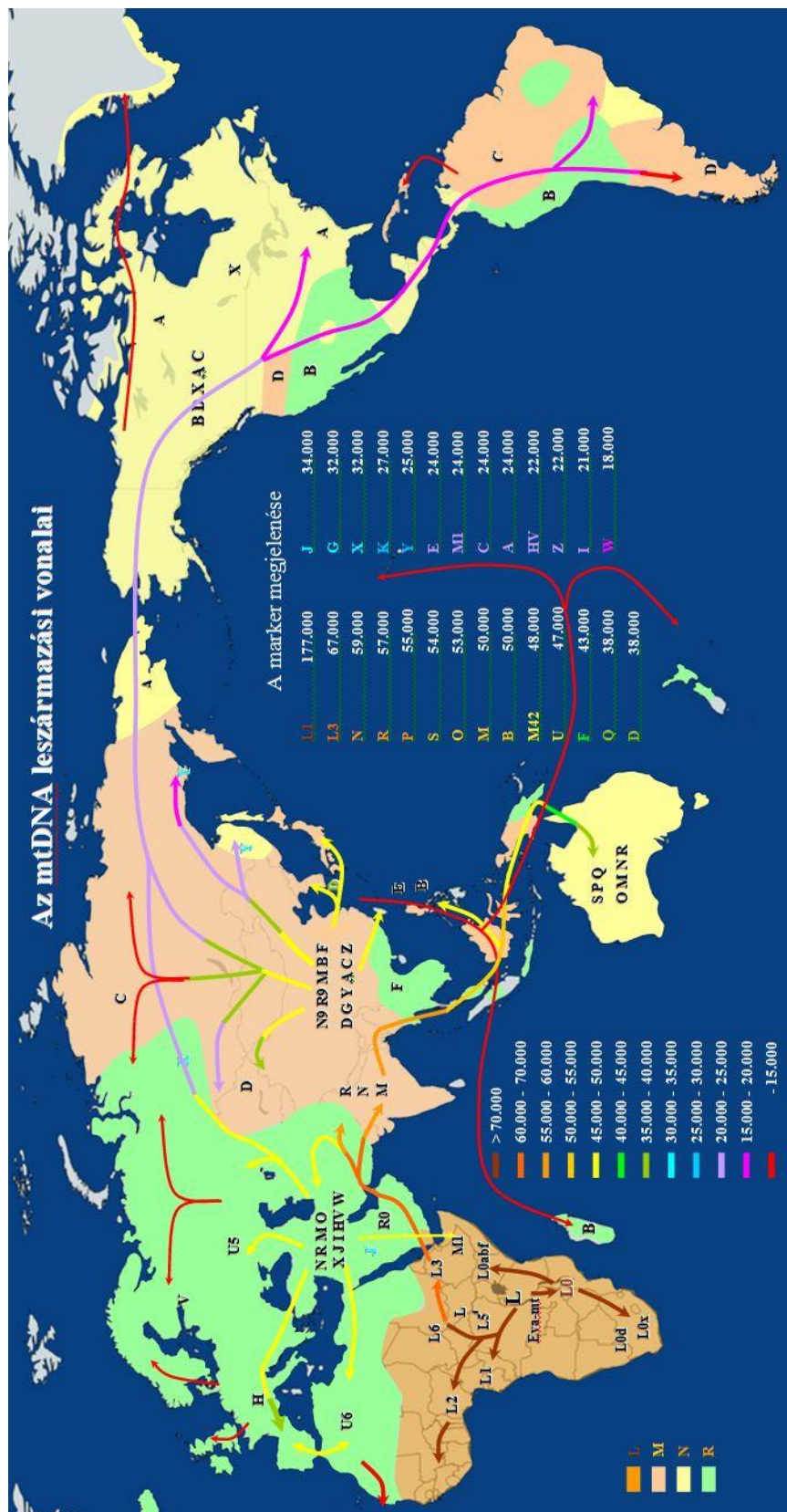
A 21. század küszöbén új vizsgáló technológia jött létre: a DNS alapú filogenetika (*phylogenetics*) vagy genealógia. A DNS, az örökítő anyag, egy gigantikus digitális könyv, amely rögzíti az emberi szervezet felépítésének, működésének instrukcióit és saját fajunk receptjét. Létezik ún. sejtmagon kívüli genom is - a *mitokondriális DNS* (mtDNS), mely értelemszerűen a mitokondriumokban helyezkedik el. A kizárólag apai ágon öröklődő Y-kromoszóma, valamint a kizárólag anyai ágon öröklődő mitokondriális DNS (mtDNS) nukleotid összetételének vizsgálatával lehetővé válik az összes eddig élt és ma élő embernek egy óriási „globális családfán” történő elhelyezése [34, 57]. 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 nagyvonalú 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 sejteikben hordozzák történelmünk egy fontos darabját. A ma élő emberekből következtethetünk őseink nagyszabású migrációs útvonalaira és a maiak egymáshoz képest, történelmi léptékben mérhető rokonsági viszonyaira (6. és 7. ábra).

Az *Y-kromoszóma* és a *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. A DNS alapú filogenetika által vizsgált tulajdonságok egy olyan stabil rendszert alkotnak, amely lehetővé teszi az időben és térben távoli népeségek közötti bonyolult kapcsolatok feltérképezését.



6. ábra. Az apai leszármazási vonalak elterjedése

Az ábrán látható színes vonalak az ősapák vándorlását mutatják (bal oldali színes skála), a marker megjelenési idejét (az ősapák születése) a jobb oldali színskála mutatja



7. ábra. Az anyai leszármazási vonalak terjedése

Az ábrán látható színes vonalak az őspanyák vándorlását mutatják a terjedési időszakban (bal oldali színskála), a marker megjelenési idejét (őspanyák születése) a jobb oldali színes skála mutatja

Remélem, az értekezés olvasója rávezethető arra, hogy távoli őseink, a modern emberek közös utat tettek meg az „édenkertből” – Fekete Afrikából – kiindulva; hol rövidebb, hol hosszabb ideig tartózkodva egy-egy földrajzi régióban, elhozták 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...

3.2.2. Ádám és Éva randevúja az Édenkertben

A modern ember Afrikában, az emberiség bölcsőjében született, aminek ténye manapság vitathatatlan. A Biblia metaforájával élve Ádám és Éva meghitt randevúja valahol a kelet-afrikai szavannán történhetett, közel 200 ezer évvel ezelőtt. A képzelet szárnyaira hagyatkozva mondhatjuk, hogy Ádám egy holdfényes éjszakán szerelmet vallott Évának az ún. *csettintő nyelven*. Egyes elméletek szerint a csettintő nyelvek a legősibb nyelvek közé tartoznak, és feltételezhető, hogy mai beszélőik, pl. a busmanok és koisanok, lehetnek a ma élő emberek őseinek legközelebbi leszármazottjai/rokonai.

Őseink „születésének” időbeli becslése a DNS-ben található mutációkon alapuló genetikai időszámítási módszerekkel történik [32, 58]. A modern ember mintegy 60-70 ezer évvel ezelőtt rajzott ki Afrikából (**6. és 7. ábra**). Ennek feltételeként, motivációs háttérként nyomós okunk van valószínűsíteni a népesség szaporodását, a csoportkohézió formáinak kialakulását, az éhínséget, az éghajlat változását, vagy új technológia kialakulását, pl. esetleg kezdetleges, tengeri helyváltoztatásra alkalmas eszköz készítését. Ha ez így van, akkor a modern ember Ázsia meghódítását Afrikából kiindulva az Indiai-óceán partja mentén haladva teljesítette, míg Ausztráliát Délkelet-Ázsia felől érte el. A lényeg az, hogy az ember csakhamar bejárta a Föld minden zegét-zugát. Ez a felfedező időszak meglepően rövid volt, egyes számítások szerint alig 5-6000 évet vett igénybe.

A „National Geographic Project” és az IBM összefogása révén egy genetikai kutatási program (A „Te genetikai utad” – Your Genetic Journey) vette kezdetét 2005-ben [59]. A projekt azt a célt tűzte maga elé, hogy a lehető legaprólékosabban feltárja azokat az útvonalakat, melyeket az egyes embercsoportok követtek, mikor Afrikából kiindulva benépesítették a földkerekséget. A molekuláris genetika ugrásszerű fejlődésének köszönhetően ma már nemcsak Ádámtól – azaz minden mai férfi közös ősétől – követhetjük nyomon a férfiak genetikai útvonalát az Y-kromoszóma vizsgálatával, hanem Éva vándorlását is megismerhetjük a mitokondriális DNS elemzésével. Ádám és Éva természetesen nem

azonosak bibliai névrokonaikkal, és nem is egy-egy embert jelölnek, hanem hipotetikus őseink csoportját. Együttes nyomon követésükkel, kétszeres biztonsággal határozhatjuk meg, merre haladtak őseink és mikor kezdték benépesíteni az egyes földrajzi régiókat.

A modern ember Európába érkezése megpecsételte az addig itt élő neandervölgyi emberek sorsát, akik ugyan még pár ezer évig valószínűleg együtt élhettek a modern emberrel. A legújabb kutatások szerint a mai emberek egy része kb. 2-4%-ban hordozza a neandervölgyi ember genetikai hagyatékát [60, 61].

Az emberiség őstörténete emberek és populációk vándorlásának történelme. Minden egyes fontosabb demográfiai esemény nyomot hagyott hátra a populációk genetikai diverzitásában. Ha egy populáció mérete csökken, genetikai diverzitása is csökken, ha a 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.

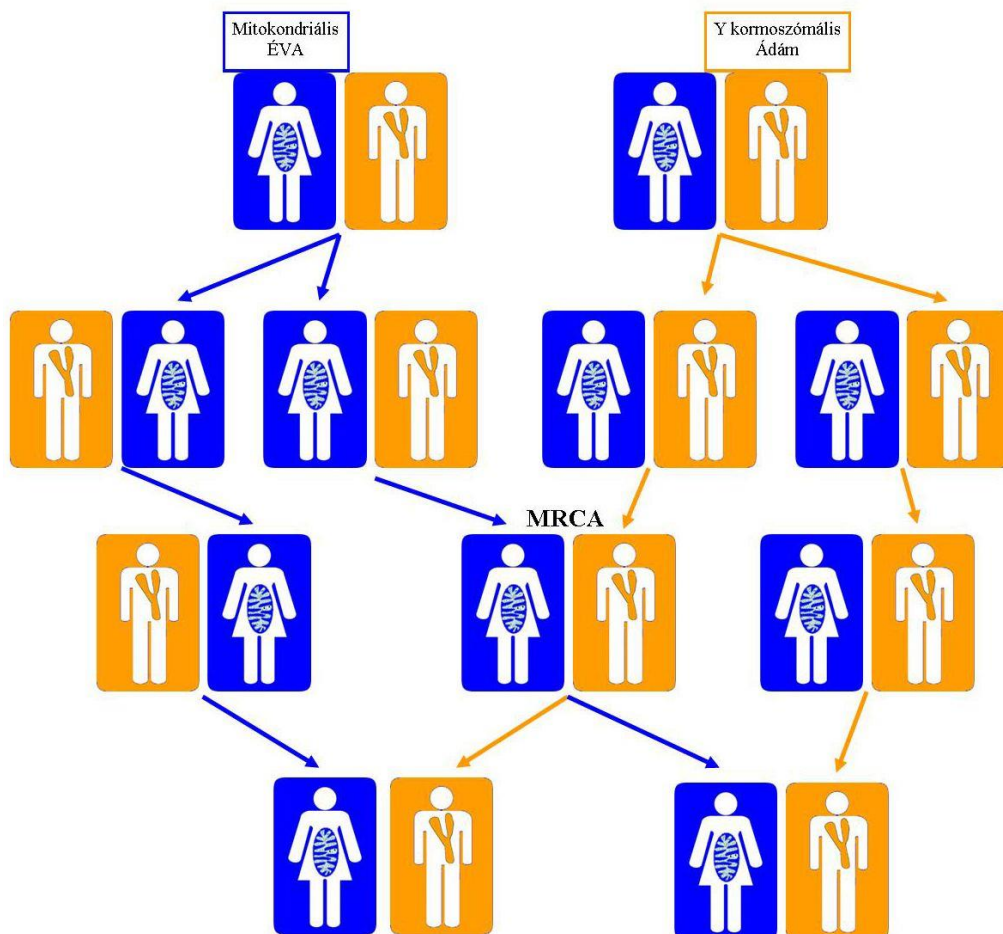
Most megkísérlem végigkalauzolni az olvasót ezen a több tízezer éves genetikai utazáson, az után fürkészve, vajon melyek azok a hírhozó DNS nyomok, amelyek feltárják nekünk eme történet részleteit.

3.2.3. Éva genetikai hagyatéka

Bryan Sykes, az Oxfordi Egyetem professzora „*Éva hét lánya*” című könyvéből idézve: *„Néhány nemzedékkel ezelőtt a legtöbb anyai ágú kapcsolat írásos nyomai teljesen eltűnnek, tehát nemigen tudnánk válaszolni a kérdésre. A DNS azonban nem felejt. A mitokondriális DNS, amely sajátos módon kizárólag női vonalon öröklődik, pontosan visszavezeti ezt a vonalat az időben. Mivel pedig a mitokondriális DNS-szekvencia véletlenszerű mutációkkal, bár egészen lassan, de változik, ezt valamiféle óraként használhatjuk”.*

A DNS az élő sejtekben két helyen fordul elő, döntően a sejtmagban és jóval kisebb mennyiségben a mitokondriumokban. Az egyszerűség és könnyebb logikai áttekintés kedvéért elsőként essék szó a mitokondriumról.

Tudjuk, hogy mind a petesejt, mind a spermium tartalmaz mitokondriumokat. Viszont megtermékenyítés után a spermiumból származó mitokondriumokat valamilyen, eddig részleteiben még ismeretlen, mechanizmus elpusztítja – vagy a petesejtbe be sem jutnak – így a zigótában egyedül az anya petesejtjéből származó mitokondriumok maradnak meg. A mitokondriális DNS-t tehát kizárólag az anyától kapja az utód, az ezek által meghatározott tulajdonságok csak anyai úton örökítődnek tovább a következő generációba (**8. ábra**).



8.ábra. A mtDNS és az Y-kromoszóma öröklődése
Az ábra bal oldalon az anyai, jobb oldalon az apai leszármazási
vonalak folytonosságát illusztrálja.

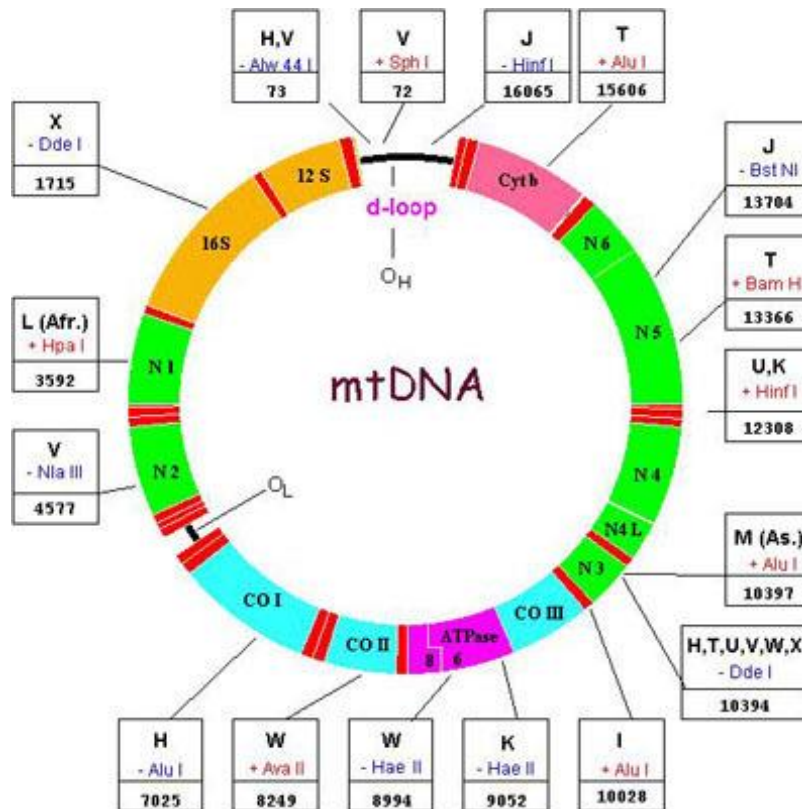
A **8. ábrán** látható, hogy ha egy őspanának és őspanának minden egyes generációban legalább egy leány vagy fiú gyermeke születik, akkor a ma élő személyek vizsgálatával az ősök genetikai üzenete megfejthető. Az ábra azt is mutatja, hogy ha két genetikailag eltérő embercsoport a vándorlás során összehatalkozik, és azután szétválk, akkor magukkal viszik

az egyik vagy a másik csoportra jellemző genetikai tulajdonságokat. A legközelebbi közös ős (MRCA, Most Recent Common Ancestor) óta eltelt idő matematikai módszerrel kiszámítható.

Genealógiai célból a mtDNS-nek egy gyakran mutálódó szakaszát, a *hipervariábilis szakaszt* (kontroll régió) vizsgálták sokáig, de ma már a teljes mitokondriális genom szekvenálása a cél. Az e szakaszban tárolt információk nagyon változatosak, mivel az ott keletkezett mutációk megőrződnek és nincsenek kitéve szelekciós nyomásnak. Filogenetikai szempontból pont ez a régió hordozza őseink genetikai szignálját, azaz Éva genetikai üzenetét leányainak.

A vizsgált személynek “rCRS”-től való eltéréseinek összessége határozza meg a haplotípust. A haplocsoportot hasonló haplotípussal rendelkező személyek csoportja alkotja, 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 (**9. ábra**).

A mtDNS haplocsoportokat – ugyanúgy, mint az Y-kromoszómális haplocsoportokat – alfabetikus szimbólummal jelölik, jóllehet a hasonló betűk nem ugyanazt jelentik.



9.ábra. A mtDNS haplocsoportokra jellemző nukleotidok helyei⁶.

A haplocsoport besorolásához szükséges pozíciók, amelyeken az ábrán látható restriktációs endonukleázok hasító helyei találhatóak (pl. az H haplocsoportra a mtDNS 7025. pozíciójában a COI génben egy AluI hasító hely jellemző).

3.2.4. Szüleink genetikai öröksége: Nukleáris DNS és őstörténet

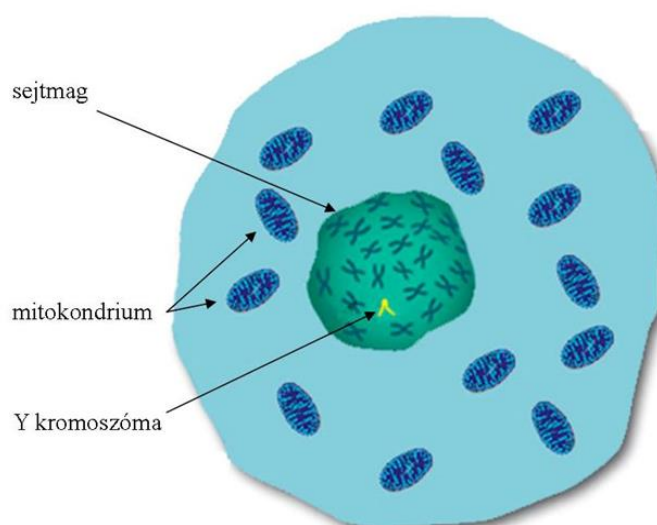
Az ember nukleáris genomját 23 kromoszómapár, azaz 46 kromoszóma alkotja. A 23 pár kromoszóma közül 22 pár autoszómális, a 23. pár a nemi kromoszóma. A Humán Genom Konzorcium szerint a genomban 25-30 ezer gén található [62]. A többi DNS -t a szakirodalom „nem kódoló DNS-nek” nevezi. Ahogy a 3.1.2. fejezet is említi, az igazságügyi genetikában bűncselekmények tettesének azonosítására alkalmazott ún. *DNS-ujjlenyomat* vizsgálatnak az alapja is ezeknek az –autoszómális kromoszómákon lokalizálódott – nem kódoló DNS-ekben tárolt információknak köszönhető.

A rekombinációnak köszönhetően a nukleáris DNS által közvetített genetikai üzenet általában 1-2 generációra visszamenőleg szolgáltat leírást múltunkról, vagyis mindkét szülő felmenőiről. Bár mostanság többszázezer szorosan kapcsolt autoszómális SNP vizsgálatával (microarray) vagy a teljes genom szekvenálás (NGS) segítségével lehetségessé vált sokkal

⁶<https://www.geni.com/projects/mtDNA-Haplogroups/3921>

több információhoz jutni – ahogy azt a nemzetközi kooperációban készült tanulmányunk is mutatja [63] – mint más autoszómális markerek, például STR-ek felhasználásával. Ezen eredmények segítségével és összehasonlításával a recens és ősi humán, valamint archaikus emberszabásúak (neandervölgyi, denisovai ősemberek) közötti rokonsági viszonyokat és az egymással való keveredés mértékét is lehet tisztázni (admixture analysis).

Egy kivétel viszont a nukleáris DNS tekintetében is van, mégpedig az egyik nemi kromoszóma: nevezetesen az Y-kromoszóma (**10. ábra**), amely régmúlt időkről mond el történeteket. A következő fejezeteket a szerző az Y-kromoszómának szenteli.



10. ábra. A sejt és a genomok vázlatos képe
Az Y-kromoszóma a sejtmagi DNS-hez tartozik, míg az mtDNS a sejtmagon kívül a citoplazmában található.

3.2.5. Ádám genetikai üzenete

A férfi nemi kromoszómái az X és az Y, mely alakban és méretben is eltérnek. Az Y-kromoszóma méretben sokkal kisebb és satnyább, mint az X-kromoszóma. Az Y-kromoszóma kb. 57 millió bázispárból, míg az X-kromoszóma kb. 156 millió bázispárból áll [64]. Egyrészt a méretbeli különbségből adódóan, másrészt annak köszönhetően, hogy az Y-kromoszóma legnagyobb részén az X-től eltérő, azzal nem homológ DNS található, az Y-kromoszóma legnagyobb része rekombinációtól mentes. Csak az Y-kromoszóma két végén lévő kis szegmenseken (ún. PAR1 és PAR2 régió) és az X-kromoszóma két végén lévő ugyanazon szakaszok között jöhet létre némi rekombináció. A Y-kromoszóma tehát az

mtDNS-sel analóg módon, de csak apai vonalon öröklődik, ezért alkalmas apai leszármazási vonalak egyértelmű követésére. Az Y-kromoszóma apáról fiúra örökön-örököké rekombináció nélkül továbbítódik egyik nemzedékről a másikra, ha minden nemzedékben legalább egy fiú születik (**8. ábra**). az anyai leszármazási vonalakkal – mtDNS-ben tárolt genetikai információ nyomon követésével – analóg módon az apai leszármazási vonalak is feltérképezhetők az Y-kromoszóma által közvetített genetikai üzenetek segítségével. Csak legyen eszköz és anyagi forrás ezeknek az izgalmas és felülírhatatlan üzeneteknek a megfejtéséhez, hogy olvashassunk ősapáinkról. Ez az ősapák üzenete a fiaiknak, azaz Ádám genetikai üzenete.

3.2.6. Színre lép a DNS alapú filogenetika

A következőkben a filogenetikai szempontból fontos fogalmakat tárgyaljuk. Maga a filogenetika témaköre magába foglalja a humán evolúciós és migrációs folyamatokat, vagyis az őstörténeti folyamatokat is. A DNS alapú őstörténeti vizsgálat nagyon fiatal tudományág. Ehhez a vizsgálathoz szükségszerű volt felfedezni a kutatóknak, hogy melyek azok a DNS részek, amelyek kizárólag anyai és apai leszármazási vonalakat alkotva évezredekken vagy évszázadokon keresztül generációról generációra úgy adódnak tovább, hogy alapvetően nem keverednek az anyai és apai DNS tulajdonságok. Az áttörést 1987-ben Rebecca Cann, Mark Stoneking és Allan Wilson *Nature* folyóiratban megjelent „*Mitokondriális DNS és humán evolúció*” című munkája hozta el [42]. A szerzők öt különböző földrajzi populációból 147 fő mitokondriális DNS-ét elemezték, és a következő megállapítást tették: eredményeik szerint mindezen mitokondriális DNS-ek ugyanattól, a feltehetően mintegy 200 000 évvel ezelőtt Afrikában élt nőtől erednek. Az afrikai populációt kivéve minden más vizsgált populációban többféle leszármazási vonal volt megtalálható, utalva arra, hogy az egyes területeken történt kolonizáció nem egyszeri eseményként zajlott le.

Így született meg az őszanya, a „mitokondriális Éva” fogalma és a mitokondriális DNS-ben rögzített történelme, az az egyedül álló üzenet a múltból, amely nemzedékről nemzedékre szinte változatlanul adódik át, csupán a molekuláris óra ketyegésével hébe-hóba történő mutációkkal gazdagodva. Ezekről a mutációkról szinte pontosan tudható az is, hogy közülük melyik milyen földrajzi területen keletkezhetett.

Az 1990-es években a kutatók nagy lázzal kerestek olyan genetikai markereket, amelyek segítségével – az akkor már ismert, mtDNS alapján vizsgált anyai leszármazási

vonalakhoz hasonlóan – feltérképezhetők lennének az apai leszármazási vonalak is. A választ Semino és munkatársai által 2000-ben a Science folyóiratban közzétett nemzetközi együttműködésben készült „*A paleolitikus Homo sapiens sapiens genetikai öröksége a jelenkori európaiakban az Y-kromoszóma szempontjából*” című cikke foglalta össze [65]. Ettől az időponttól kezdve lehetőség nyílt ősapáink üzenetének olvasására.

A mtDNS esetében a mutációk – nukleotid cserék – főleg a kontroll régióban találhatók, míg az Y-kromoszóma esetében a kromoszóma hosszában egyenletesen elosztva fordulnak elő. Az Y-kromoszómán olyan nukleotid pozíciókon található bázispárok azonosítására törekszünk, amelyek egy-egy sikeres ősapára utalva, a férfinépeséget releváns részekre bontják. Az ilyen lokuszt vagy bázist Y-SNP-nek nevezzük, ezekkel definiáljuk az apai leszármazási vonalakat (haplocsoport).

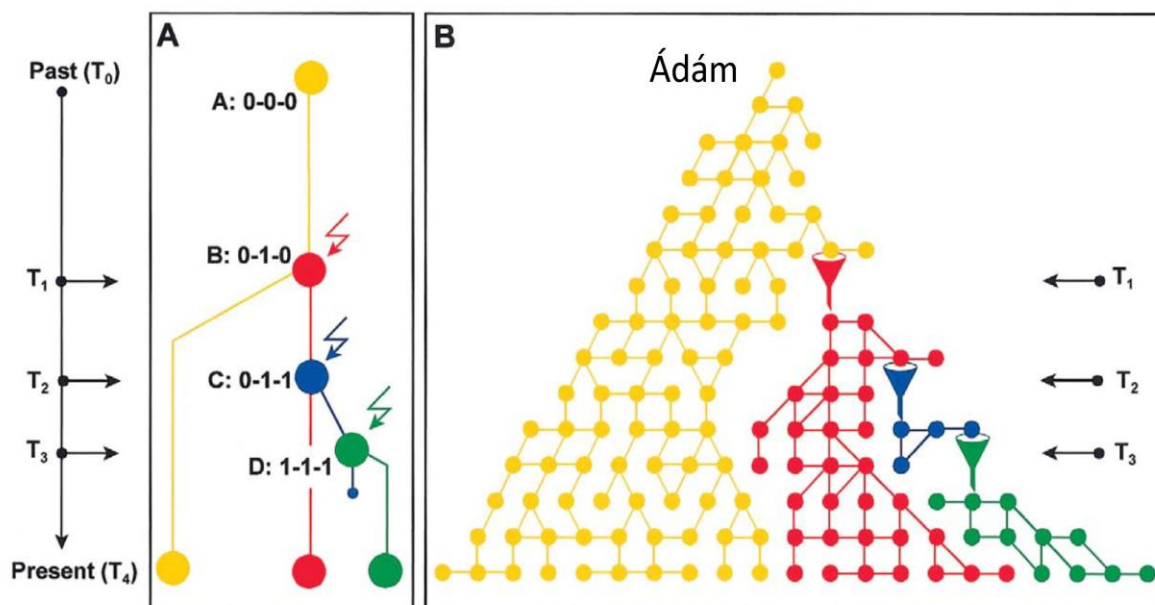
Az Y-SNP megléte/hiánya alapján minden férfiről egyértelműen eldönthető, hogy egy adott haplocsoportba tartozik-e vagy sem. A haplocsoport az egy apai leszármazási vonalhoz tartozó férfiak csoportját jelenti. Példaként bemutatjuk az „*eurázsiai Ádám*” haplocsoportját, aki a ma élő összes nem afrikai férfi őse. Őt az úgynevezett M168 SNP-vel definiáljuk. Az *eurázsiai Ádám* Y-kromoszómáján a 14 813 991 pozícióban keletkezett egy mutáció, C → T szubsztitúció, kb. 65-70 ezer évvel ezelőtt. Annak az esélye, hogy pont ebben a pozícióban, egy másik férfiban is keletkezzen ugyanilyen mutáció, 1:57 millióhoz, szinte a nullával egyenlő. Ezért is hívjuk ezt „*egyszeri mutációs eseménynek*” (*unique mutation event*) [33]. Ennek az ősapának a férfi leszármazottjai ma döntően az eurázsiai térség lakói. Ez azt jelenti, hogy sok olyan férfi van, akinél ugyanazon a helyen, ugyanaz az SNP mutáció megtalálható, vagyis mind egy ősapától származnak és a haplotípusaik is hasonlóak.

A haplocsoportokat alfabetikus szimbólumokkal jelöljük (A-T), hasonlóan az anyai leszármazási vonalakhoz, de mindig fontos megjegyezni, hogy éppen melyikre vonatkozó ábécét használjuk: a mtDNS-ét vagy az Y-kromoszóma SNP-jét, mert teljesen másra utalnak a hasonló betűk! Természetesen egy haplocsoporton belül létezhetnek alcsoportok (más pozícióban lévő SNP jelenléte), és azokon belül is további alcsoportok (még egy másik pozícióban lévő SNP jelenléte), ami azt eredményezi, hogy a haplocsoportok egy, a minden férfi őst reprezentáló „A” haplocsoportból induló fa-struktúrába (ld. később a 12. ábrát) rendeződnek [34].

A férfiak esetében a haplotípusok meghatározása – ahogy az igazságügyi genetikában is – az Y-STR vizsgálatokon alapszik. Minél több STR lokuszt vizsgálunk, genetikailag annál precízebben azonosítható a vizsgált férfi. Ezen STR markerek összessége adja meg a haplotípust.

Az Y-kromoszóma evolúciójában a molekuláris óra ketyegésével az egy-egy lokuszon lévő allélok repeat száma általában eggyel csökken vagy nő, ez az ún. *stepwise mutation model*. Ha két férfi közeli rokona egymásnak (pl. nagyapa és unoka), akkor szinte majdnem azonosak a haplotípusaik, viszont, ha a közös őstől történt leválás óta sok idő eltelt, akkor több lokuszon és több lépéssel eltérhetnek az allélok. Ugyanúgy, mint az mtDNS esetében az egy-egy haplocsoportba tartozó férfiak közel azonos haplotípussal rendelkeznek.

Az Y-STR haplotípusok azért különösen fontosak, mert segítségével megbecsülhetjük egy adott haplocsoport korát, illetve a haplocsoporton belüli férfiak egymás közötti apai vonalú rokonsági kapcsolatait az ún. *Median Joining Network* analízis segítségével [66]. Minél régebbi egy haplocsoport, annál nagyobb a diverzitása az ősapa megszületése óta eltelt idő függvényében [33]. A kutatók a világ különböző régióiból származó több mint 2000 férfi és főemlős (csimpánz) DNS szekvencia analíziséből állapították meg, hogy az apai leszármazási vonalak (haplocsoportok) hol és mikor keletkeztek, és melyek azok a genetikai markerek, amelyek jellemzik az adott ősapákat (**11. ábra**). Az Y-kromoszómális Ádám, aki „A” ősapának nevezhető (A haplocsoporthoz tartozik), a piramis tetején található és kb. 235 ezer éve született (T_0 időpontban). Az idő haladtával Ádám férfiági leszármazottjai felszaporodtak (az ábrán sárga körök) és valamelyik leszármazottja „B” ősapává (B haplocsoport) vált kb. 88 ezer éve (T_1 időpontban). A piros körök reprezentálják a „B” ősapa leszármazottjait. Tovább telt-múlt az idő és a „B” ősapa egyik leszármazottja „C” ősapává vált kb. 48 ezer éve (T_2 időpontban), amelyet a kék körök jeleznek. Az idő tovább ketyegett és kb. 46 ezer éve (T_3 időpontban) a „C” ősapa egyik leszármazottja „D” ősapaként jött világra. Látható az is ábrán, hogy minél „idősebb” egy ősapa, a leszármazottjaiban a haplotípusok annál inkább eltérnek az STR mutációnak köszönhetően, azaz annál sokfélebbek lesznek (lásd a sárga körök a piramis tetejéről indulva sokfelé ágaznak). Az is megfigyelhető, hogy pl. a „D” ősapa leszármazottjai nem annyira sokfélék, mint az „A” ősapáé, mivel a „D” egy viszonylag „fiatal” ősapa és nem volt elég idő arra, hogy az STR alapú haplotípusai diverzebbé váljanak.



Ősapák (SNP mutációk
keletkezése) születése

Ősapák (haplocsoportok) fiainak (haplotípusok)
születése, szaporulata (diverzifikáció) az idő függvényében

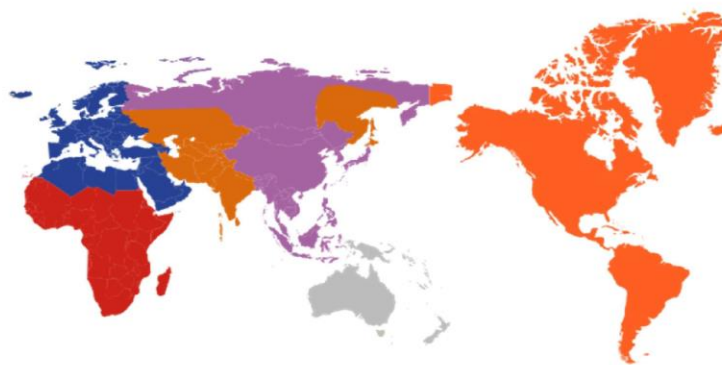
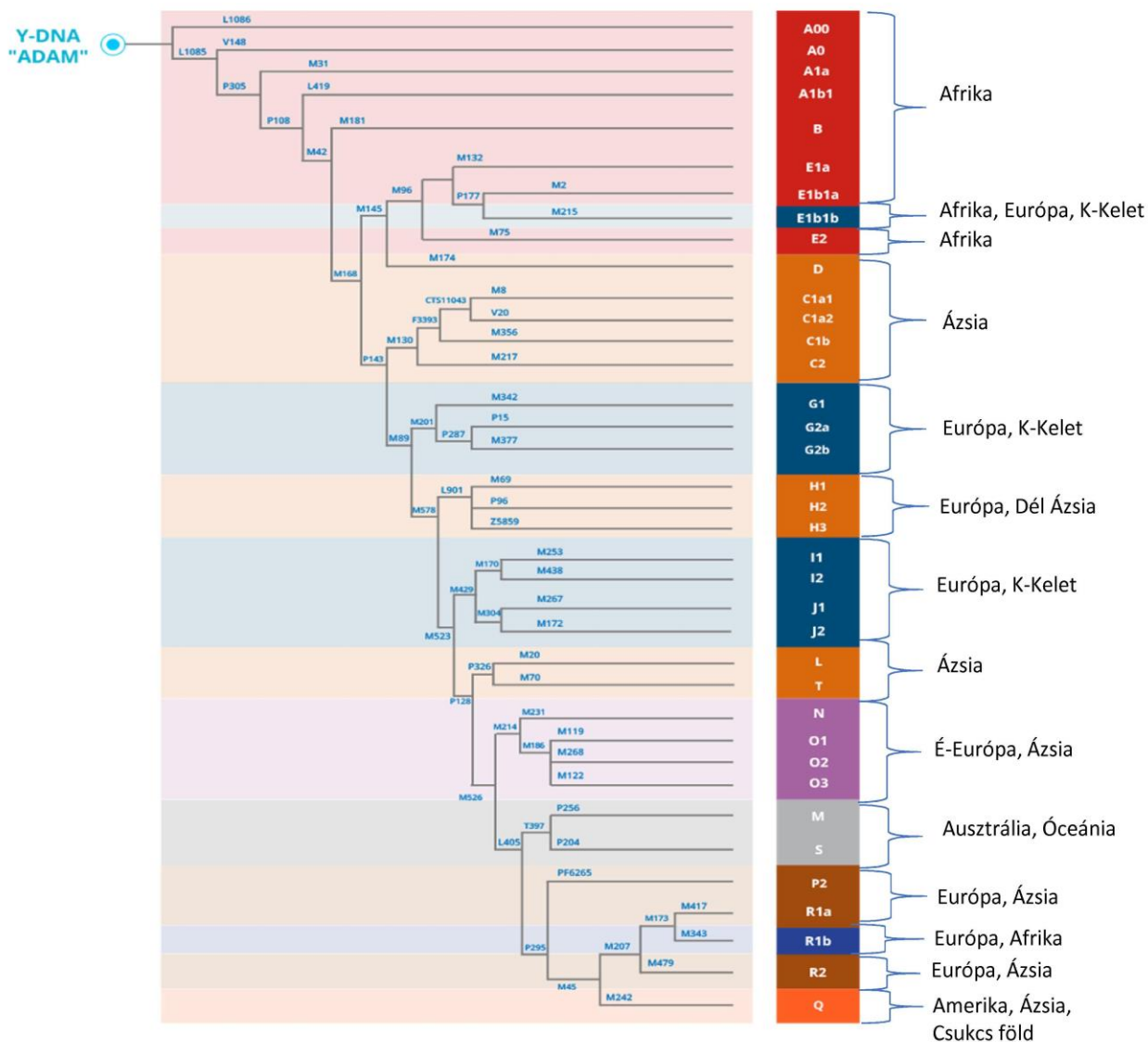
11. ábra. A haplocsoportok időbeli kialakulásának modellezése⁷

T_0 - T_4 : idő; A: A haplocsoport (sárga); B: B haplocsoport (piros);

C: C haplocsoport (kék); D: D haplocsoport (zöld)

A kizárólag apai ágon öröklődő Y-kromoszóma vizsgálatával így módon lehetővé vált az összes eddig élt és ma élő embernek egy óriási „globális családfán” történő elhelyezése. Az ősapák ma élő leszármazottjait a globális családfán születési sorrendben elhelyezve a **12. ábra** mutatja.

⁷Peter de Knijff, Am. J. Hum. Genet. 67:1055-1061, 2000 (módosítva)



12. ábra. Egyszerűsített filogenetikai fa⁸

Az ősapákat jellemző SNP mutációk (pl. a Q haplocsoportot M242 SNP mutáció jellemzi) és a haplocsoportok (A-Q) eloszlása a világtérképen (a haplocsoport színe azonos a földrajzi régióval)

⁸<https://genographic.nationalgeographic.com/tree-updates> (módosítva)

3.2.6.1. Y-kromoszóma haplocsoportok elterjedése

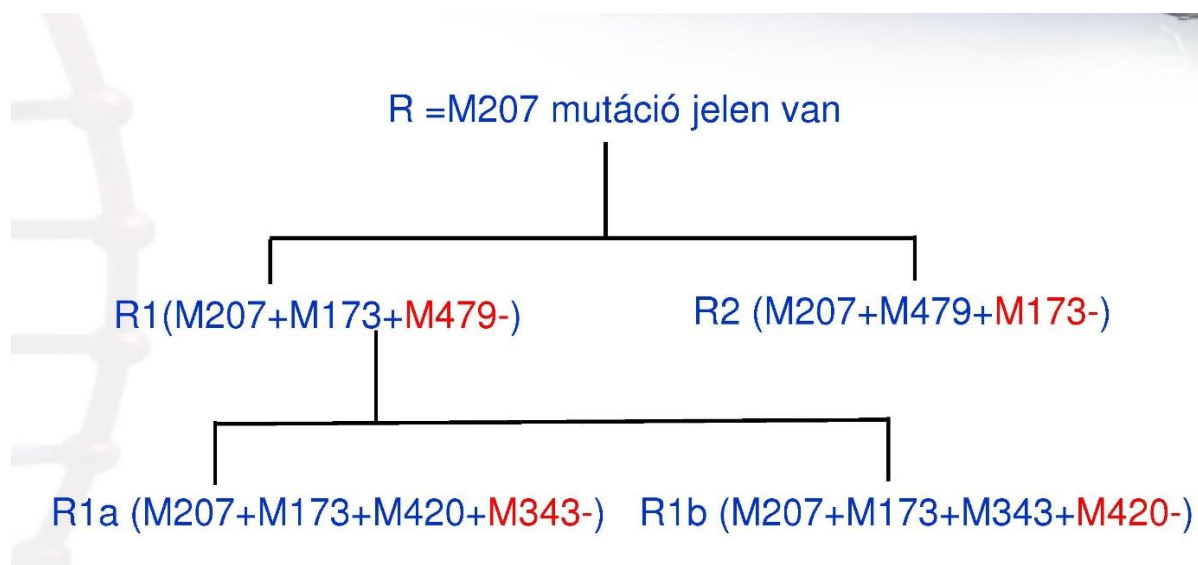
A jelen fejezet saját adatgyűjtés és ábrák alapján röviden ismerteti a tanulmány szempontjából fontosabb apai haplocsoportok – apai leszármazási vonalak – elterjedését és a haplocsoportot definiáló SNP „születésének” legvalószínűbb keletkezési helyét, idejét és a fontosabb expanziós térségeket. A haplocsoportok valaha élt hús-vér emberekre, olyan de facto sikeres ősapákra mutatnak, akiknek fiai, unokái és távolabbi utódai az apai leszármazási vonalak mentén széles körben elterjedtek. Ez a siker összefügghet egy új technikai vívmánnyal, vagy egy addig lakatlan földrajzi régió meghódításával, úgyszintén az éghajlat kedvezőbbé válásával. A korai emberi közösségek alacsony létszámát figyelembe véve nem zárható ki a szerencse, vagy a történészek elől örökre rejtve maradó egyéni teljesítmények lehetősége sem. Bármilyen is volt az oka a sikernek, az valamely apai vonal adott közösségen belüli meghatározóvá válásában nyilvánult meg.

Mivel az Y-kromoszóma csak a férfiakban található meg, ezért azt csak a biológiai apjától öröklí minden fiú. Ez azonban nem jelenti azt, hogy az Y-kromoszóma teljesen azonos volna apa és fiú között. A DNS másolási folyamatokba mindig csúszhatnak kisebb változások, és ezért az Y-kromoszómában, a DNS-lánc bizonyos pontjaiban eltérés – mutáció – alakulhat ki apa és fia között. Ezek az eltérések viszont tovább öröklődnek a fiú(k)ra a fiú unoká(k)ra, és így tovább. A haplocsoport az egyes sikeres ősapákban lezajlott SNP mutációval azonosítható. A filogenetika által azonosított, haplocsoportot meghatározó mutációk mai tudásunk szerint semmilyen fenotípusos tulajdonságot nem határoznak meg, ember voltunk szempontjából redundáns információt hordoznak. A mutációk jelölésére általában egy betűt és a betűt követő 2, 3 vagy több számjegyű kódot használnak (pl.: Z93, M207, L1034, P43).

Magukat a haplocsoportokat betűk és számok kombinációjaként jelöli a tudomány, amelynek első tagja önkényes, a további betűk és számok pedig az alcsoportok leszármazási sorrendjében következnek. Például az R haplocsoport az M207 mutáció (SNP) meglétével azonos, tehát minden férfi, akinél az M207-es mutáció megtalálható az Y-kromoszóma DNS-ének egy adott pozíciójában, az R haplocsoportba tartozik. Egyértelműség kedvéért nemritkán a haplocsoport mellé, odaírják kötőjellel a mutáció azonosítóját is. Így az R haplocsoportot jelölhetjük R-M207-nak is. Ez a jelölésmód különösen akkor hasznos, ha a haplocsoport elnevezése már meglehetősen hosszú, mert sokadik mélységű alcsoportot jelöl (pl. R1a1a1b2-Z93 helyett R1a-Z93 vagy R-Z93).

Folytassuk most tovább az R haplocsoport alcsoportokra való tagolását (**13. ábra**). Az M207-es mutációt hordozó férfiak egy része emellett az M173-as mutációt is hordozza, őket jelöli az R1-M173 alcsoport. Azok az M207-es mutációt hordozó férfiak, akik az M479 SNP-t is hordozzák, alkotják az R2-M479 alcsoportot. Az R1 és R2 haplocsoportokat megtestesítő ősapa az R haplocsoportot reprezentáló ősapa két sikeres, egymástól független, tisztán apai vonalú leszármazottja. Az R1 és R2 haplocsoportok között tehát „fivér” jellegű a kapcsolat, ami abban nyilvánul meg, hogy az R2-M479 alcsoportban nem található meg az M173-as és az R1-M173-as csoportban pedig az M479 mutáció (**13. ábra**). Elvileg természetesen elképzelhető, hogy két különböző személyben ugyanaz a mutáció lejátszódjon, de ennek a valószínűsége a gyakorlatban szinte nullával azonos.

Természetesen az R1 haplocsoport is tovább bomlik, nevezetesen az R1a és R1b alcsoportokra, és az R1a illetve R1b haplocsoportok is további alcsoportokra bomlanak (**13. ábra**). Végeredményben így egy olyan nagyvonalú leszármazási fát kapunk, amelynek gyökerében, minden ma élő férfi ősapja (R apa) található, és az egyes csomópontokat az egyes sikeres ősapákra utaló mutációk határozzák meg. Az alábbi ábra szemlélteti a fent leírtakat (**13. ábra**):



13. ábra. Az R haplocsoport leszármazási vonalai
 +: a jel előtt lévő SNP mutáció detektálható (kék színű SNP)
 -: a jel előtt lévő SNP mutáció nem detektálható (piros színű SNP)

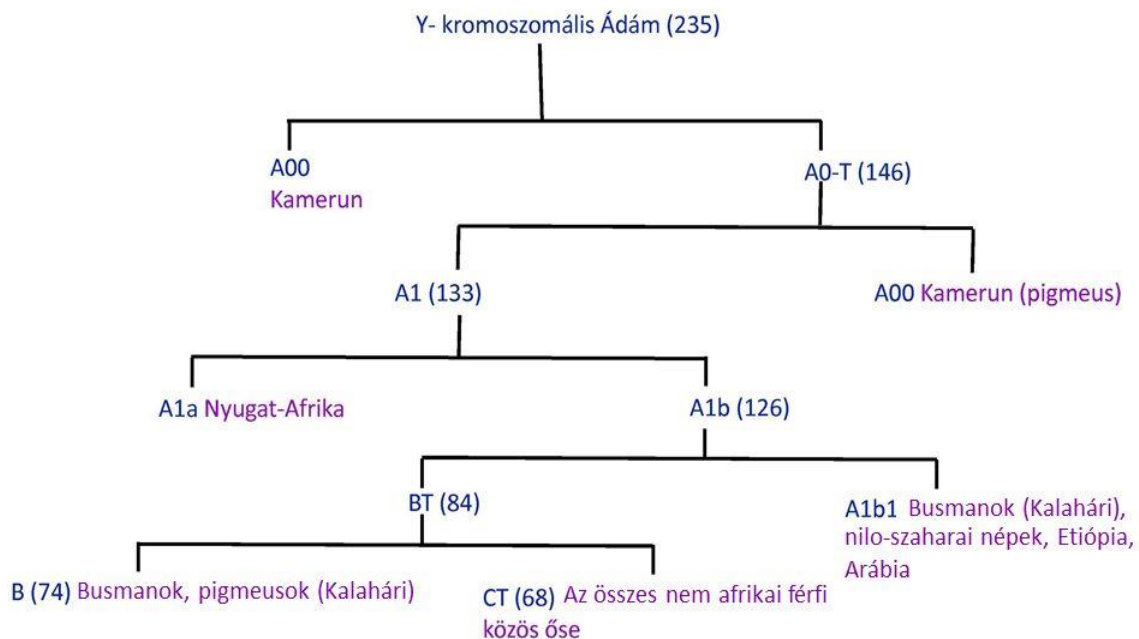
A korai tanulmányok még nem tudták megbízhatóan meghatározni a haplocsoportok földrajzi eredetét. Sokáig azt hitték, hogy egy haplocsoport ott keletkezett, ahol a leggyakoribb vagy éppen legdiverzebb a jelenléte. Elképzelhető azonban, hogy egy haplocsoport végül mégsem ott szaporodik el, ahol keletkezett, hanem születésétől távoli

területen. Ma már tudjuk, hogy egy haplocsoport nagy valószínűséggel ott keletkezett, ahol a legidősebb „ősapák” és „fivérek” nyomai megtalálhatók. Például az R-M207 haplocsoport nagy valószínűséggel Ázsiában keletkezett annak ellenére, hogy az R1a és R1b ma Európa két leggyakoribb haplocsoportja. Maga az R2 közép-ázsiai eredetű. Az R „fivére”, a Q haplocsoport szintén Közép- vagy Belső-Ázsiában keletkezett. Hasonló érveléssel támasztható alá, hogy az uráli nyelvcsalád szempontjából fontosnak tűnő N haplocsoport is kelet-ázsiai eredetű, a dél kínai Yunnan tartományhoz köthető. Az N „fivére” az O haplocsoport csak Kelet-Ázsiában található meg, hasonlóan az N és az O közös őséhez. Mindez a gyakorlatban nem kevesebbet jelent, mint azt, hogy Európát meghatározó impulzusok érték Ázsia felől az apai vonalak tekintetében. Az természetesen kérdés, hogy ezeknek a nagyon régi folyamatoknak mennyi közük van a ma élő népekhez, a jelen kor meghatározó nyelvcsaládjaihoz, amíg ez a kérdés nem kellő mértékben tisztázott, értelemszerűen nem is szabad ebből a tényből messzemenő következtetéseket levonni.

A korbecslés a filogenetika legvitatottabb területe. A különböző mutációs ráták, és így a haplocsoportok különböző mutációs rátákkal számolt kora között olyan nagy különbségek vannak, hogy alapvetően csak relatív kronológiáról beszélhetünk azok keletkezésére vonatkozóan. Nagy biztonsággal csak annyit mondhatunk, hogy két haplocsoport közül melyik az idősebb, és az idősebb jóval idősebb-e vagy sem, mint a másik. Mindenesetre a korszámítás komoly fejlődésen ment keresztül az utóbbi években. Tájékoztató jelleggel az Yfull leszármazási táblája alapján közölhető az Y haplocsoportok vélhető kora (<http://www.yfull.com/tree/>) [69], mivel vélelmezhető, hogy módszertana (nem STR alapú, hanem az Y-DNS szekvenálása alapján becsült kor), és a nagyszámú vizsgált minta révén ez a megközelítés tűnik a legmegbízhatóbbnak. Fel kell hívni a figyelmet arra, hogy amikor népnéveket vagy kormegjelölést (általában évezredben) írunk, az történeti népnévek esetében azt jelenti, hogy a régészet, történettudomány által e nép tagjának vélt személyek csontleleteiből származnak az adott minták; míg a jelenlegi népek esetében, hogy a mintaadók magukat az adott nép tagjának vallották. A kormegjelölés pedig arra utal, hogy mikor alakulhatott ki valamelyik apai leszármazóban az adott mutáció, majd lassan a leszármazás generációi alatt hogyan terjedt el. A haplocsoportok földrajzi eloszlásai, keletkezési helyei és becsült korai a következők [34,35, 36, 37, 59, 67, 68, 69]:

Az **A haplocsoport** (M91 SNP, Y-kromoszómális Ádám): a legősibb csoport, kis gyakorisággal Afrika-szerte előfordul, főként a busmanok között, Etiópiában, valamint a Nílus felső folyásánál élő nilóták körében. A legősibb, kameruni „A00” csoport leválása a

többi ma élő férfi őstől jelenlegi ismereteink szerint 235 ezer évvel ezelőttre tehető (**14. ábra**).



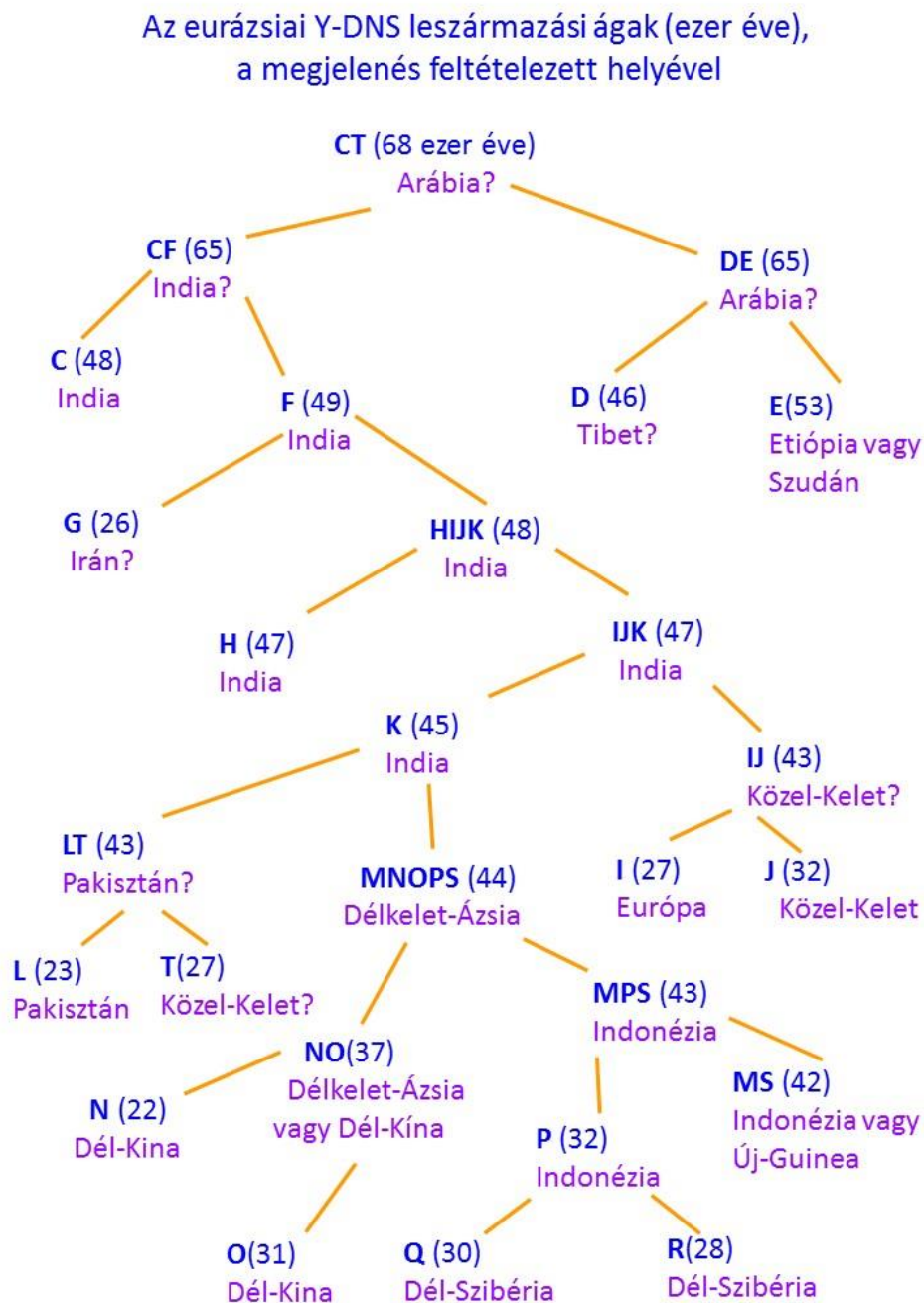
14. ábra. A korai (afrikai) Y-haplocsoport ágak és a csoportok kora ezer évben

A **B haplocsoport** (M60): az A-hoz hasonlóan csak Afrikában fordul elő, elsősorban a pigmeusok és busmanok között. Elválása a többi csoport közös ősenek számító CT csoporttól 84 ezer évvel ezelőttre tehető (**14. ábra**).

A **CT haplocsoport** (M168): a C-től T-ig terjedő haplocsoportok, azaz a ma élő összes nem afrikai férfi közös őse. Két leszármazottja van, a CF és a DE csoport. Ez a haplocsoport az első olyan férfinak a genetikai nyoma, aki elhagyta Afrikát kb. 68 ezer évvel ezelőtt (**15. ábra**).

A **DE haplocsoport** (YAP): a D és E csoportokat foglalja magába. Valószínűleg Kelet-Afrikában, Etiópiában alakult ki 65 ezer éve, de az arábiai eredet sem zárható ki. Érdekessége, hogy D alcsoportja csak Ázsiában található meg, míg az E alcsoport Afrika domináns haplocsoportja.

A **D haplocsoport** (M174): az Andamán-szigeteki őslakók és az ajnuk között található meg a legjelentősebb arányban (a népességhez viszonyított mintaszámban), ezenkívül a tibeti és a japán férfiak nagyjából egyharmada is ide tartozik. Elszórtan előfordul egész Délkelet-Ázsiában és Közép-Ázsiában. Nagyon ősi kőkorszaki vándorlás maradványa, talán Közép-Ázsiában jött létre 46 ezer éve, az eddigi szakirodalmi vizsgálatok alapján hordozói észak felől kerültek meg a Himaláját/Tibetet, ugyanis Indiában nem található meg.



15. ábra. Az „eurázsiai Ádám” leszármazási vonalai

Az **E haplocsoport**hoz (M96) tartozik a mai afrikai férfi lakosság túlnyomó többsége, függetlenül attól, hogy negroid, europid, etiopid vagy nilotid alrasszba tartozik-e. Valószínűleg Etiópiában jött létre kb. 53 ezer éve. Fontosabb alcsoportjai a következők:

E1a (M33): Maliban a szongáj férfiakra jellemző, de máshol is előfordul Nyugat-Afrikában. 52 ezer éve vált el az E1b-től, de a mai E1a férfiak közös őse csak 19 ezer éves.

E1b1a1 (M2): a fekete-afrikaiakra jellemző csoport, a niger-kongói (köztük a bantu) nyelveket beszélők között találták meg legnagyobb számban. Valószínűleg történelmi

keveredés miatt előfordul a pigmeusok, busmanok, nilóták és a csádi népek között is. A legtöbb afroamerikai is ebbe a csoportba tartozik. A mai M2 mutációt hordozó férfiak közös őse 27 ezer éve élt.

E1b1b1 (M35): az afroázsiai („sémi-hámi”) nyelveket beszélő népekre jellemző, de előfordul Európában is. Kb. 24 ezer éve jött létre, valószínűleg Etiópiában. A koszovói albánok közel fele az Európára jellemző E-M78-as alcsoportjába tartozik, de gyakori a görögök, szerbek, bolgárok, olaszok és spanyolok között is. Magyarországon a férfiak kb. 7%-a tartozik az E-M78 alcsoportba.

A **CF haplocsoport (P143):** a C és F csoport közös őse, jelenlegi ismereteink szerint Indiában alakulhatott ki kb. 65 ezer éve.

A **C haplocsoport (M216):** a tudomány mai állása szerint 48 ezer éve Dél-Indiából kiindulva a part mentén benépesítette Délkelet-Ázsiát, Ausztráliát és Északkelet-Ázsiát, később Észak-Amerikába is eljutott. A ritka C-M105 alcsoportba tartoznak az ausztrál őslakók, a polinézek, a dél-indiai őslakos veddák, valamint kis számban előfordul a japánok és ajnuk között is. A C-M217 alcsoport viszont sikeresen terjedt tova Északkelet-Ázsiából: a na-dené indiánok (pl. atapaszák, navahók) vitték Észak-Amerikába, a mongol és tunguz népek pedig Közép-Ázsiába és Szibériába. A mongol és kazah férfiak többsége ide tartozik, Dzsingisz kán is valószínűleg a C-M217-et hordozhatta. Lehetséges, hogy a hunok egy része is ebbe a csoportba tartozott, mivel ezt a haplocsoportot jelentős arányban találták meg a hun csontminták között. Magyarországon közel ezer férfi között eddig csupán egyszer fordult elő. A szakirodalom nagyjából egyetértő álláspontja szerint az ősi C-M105 csoport lehetett Európa első benépesítője is, a 7000 évvel ezelőtti kárpát-medencei leletek között 11%-ban van jelen, mára azonban teljesen eltűnt erről a területről.

Az **F haplocsoport (M89):** valószínűleg India középső vagy déli részén alakult ki kb. 48 ezer éve. Ősi formájában az indiai gondi törzseknél fordul elő a Dekkán-fennsíkon. F1 és F4 alcsoportjai Srí Lankán, F2 alcsoportja Kelet-Ázsiában fordul elő, az F3 pedig Hollandiában. Az F alcsoportja, továbbá a G, H és IJK csoport is (ld. a következőkben), valamint azok összes leszármazottja összességében az F csoportba tartozik, azaz összességében az emberiség kb. 90%-a. Ősi F* formájában viszont nagyon ritkán fordul elő (Portugáliától Malajziáig). Jelentős arányban volt jelen ez a haplocsoport Európa korai földművelői között, a Kárpát-medencében például e korszakban 25%-ra tehető az aránya, de mára nem maradt képviselője, eltűnt.

A **G haplocsoport (M201):** kb. 26 ezer éve jött létre a Kaukázus déli részén vagy Iránban. A G1-M342 csoport (19 ezer év) Kazahsztánban az argün törzs tagjai között a leggyakoribb, a

G2-P15 (18 ezer év) csoport pedig a kaukázusi oszétek (60%), grúzok (50%), északnyugat-kaukázusi nyelvű népek (abházok, cserkeszek 40-50%), valamint a szomszédos török nyelvű karacsajok és balkárok (30-32%) között a leggyakoribb. Európában Korzikán (21%) és Szardínián (13%) van a G2-P15 gyakorisági csúcsa. Magyarországon a G2-P15 előfordulási aránya ma 5-6%, míg a i.e. 6000-4000 közé datálható korai földművelők között a Kárpát-medencében még 34%-os aránya volt. A G1-M342 csoport a magukat magyarnak valló mintaadók között nem fordul elő.

A H haplocsoport (M2939): jelenlegi ismereteink szerint Indiában jött létre kb. 47 ezer éve. A helyi törzsek és az alsóbb kasztok tagjai között gyakori, az összes indiai egyharmada tartozik ide. Indián kívül kizárólag a roma és más indiai eredetű népességek körében található meg. A H-M82 alcsoportja a középkorban a vándorcigányokkal jutott el Európába. A magyarországi roma lakosságban is jelentős az aránya (az elszigeteltségtől függően 30–60%).

Az IJ haplocsoport: az I és J csoportok közös őse, kb. 43 ezer éve jöhetett létre Iránban, Anatóliában vagy a Termékeny Félhold területén.

Az I haplocsoport (M170): kb. 27 ezer éve keletkezett és még a jégtakaró legnagyobb kiterjedése (LGM) előtt elérte Európát. Az I apai vonal az európai őslakosokra jellemző. Egész Európában megtalálható a következő nagyobb alcsoportokkal:

I1-M253: ma a norvég, svéd, nyugat-finn, dán és észak-német férfiak 30-40%-át adja, de Hollandiában, Dél-Németországban, Angliában és Észtországban is gyakori (15% körül). A germán népmozgások által hagyományosan érintett többi területen is jelen van általában 10% alatt. Európán kívül gyakorlatilag nem vagy csak elenyésző arányban található meg, nagy valószínűséggel a bevándorlók vitték magukkal. Az I2-től már 27 ezer éve elvált, de a ma élő I1 férfiak egyetlen 4700 évvel ezelőtti férfitől származnak. Elterjedése teljesen a germánokhoz köthető. A magukat magyarnak valló mintaadó férfiak kb. 7-8%-a hordozza ezt a genetikai vonalat.

I2-P37: az M26 és M423 csoportok (ld. alább) közös őse, a mezolitikus Európa tipikus markere.

I2-M26: jelenlegi ismereteink szerint 8500 éves csoport, Nyugat-Európában (Franciaországban, Ibériában, Itáliában és a Brit-szigeteken 1-2%) kismértékben elterjedt, de szardíniai lakosság 41%-át és a baszkok 6-8%-át adja.

I2-M423: 14 ezer éves jelenlegi ismereteink szerint, Nyugat-Európában kismértékben fordul elő. Egy fiatalabb alcsoportja, a 2200 éves I2-CTS10228-as alcsoport, viszont jelentős szerepre tett szert Kelet-Európában. Gyakorisága a Dinári-Alpokban eléri a 40%-ot, de a Balkán többi népében (szlovének, horvátok, románok, szerbek, albánok, macedónok,

bolgárok) is 20% körüli a jelenléte. Jelentős, 15% körüli a gyakorisága a magyar, ukrán, belorusz és dél-orosz férfiak között, majd előfordulása 8-10%-ra csökken a nyugati szlávok, görögök és közép-oroszok körében.

I2-M223: Európában ma a szász és a holland népességre a legjellemzőbb ez a 12 ezer évesre becsült csoport. Kis számban (3-4%) Magyarországon is előfordul.

A **J haplocsoport** (M304): jelenlegi ismereteink szerint kb. 32 ezer éve jött létre a Közel-Keleten, ahonnan két nagy alcsoportja is elterjedt. Ősi J* formában csak a Jementől délre fekvő Sokotra szigetén található meg az Indiai-óceánon.

J1 haplocsoport (M267): Kurdisztán környékén jelenhetett meg 18-19 ezer éve, majd a szakirodalom következtetése szerint a mezőgazdasággal és állattartással terjedt el a térségben. Ma a leggyakoribb az északkelet-kaukázusi népek között (cssecsen, ingus, dagesztáni népek), illetve a jemeniek között (80%). Nagy arányban fordul elő a zsidók és arabok között, az izraelita királyi genetikai vonal (ún. Kohen haplotípus) is ide tartozik. Úgy tűnik, hogy Észak-Afrikában és a mediterrán térségben is az arab hódítással, illetve a szefárd diaszpórával terjedhetett el. Kelet-Európában jellemzően az askenáziak körében fordul elő.

J2 haplocsoport (M172): kb. 28 ezer éve jött létre Anatólia és Irán vidékén. A mediterrán térségben a görög gyarmatokhoz is köthető. Ibériában, Olaszországban, Albániában, Görögországban, Törökországban, Iránban, a Kaukázusban és a Termékeny Félhold területén kb. a férfiak egynegyede tartozik ide, elterjedt még Közép-Ázsiában és az indiai felsőbb kasztok között is. A magukat magyarnak valló mintaadók között 6-10% a gyakorisága.

A **K haplocsoport** (M9): az L, M, N, O, P, Q, R, S és T csoportok közös őse, kb. 45 ezer éve jött létre Indiában. Az európaiak, indiaiak és a mai kelet-ázsiaiak többsége, tehát az összes ma élő férfi kb. háromnegyede a K valamely alcsoportjába tartozik.

Az **L haplocsoport** (M11): jelenlegi ismereteink szerint kb. 23 ezer éve jött létre Pakisztán déli részén, Irántól Dél-Indiáig elterjedt az Indiai-óceán partján, ezen kívül Közép-Ázsiában is előfordul. A beludzso (Pakisztán) és drúzok (Libanon) közel egyharmada ebbe a csoportba tartozik. Valószínűleg – ha nem is kizárólagosan – az indus-völgyi civilizáció is ehhez a csoporthoz köthető. Kis számban (1% alatt) Magyarországon is előfordul.

A **T haplocsoport** (M70): jelenlegi ismereteink szerint kb. 27 ezer éve jött létre. Leghíresebb „tagja” Thomas Jefferson és Bryan Sykes. Nagyon érdekes az eloszlása, ugyanis az Indiai-óceán nyugati részén körívesen található meg (Tanzánia, Szomália, Egyiptom Omán, Irak, Dél-India), illetve a Földközi-tenger partvidékén is elterjedt, valószínűleg a föníciai gyarmatosítás idején. Eredete nehezen határolható be, de valószínűleg a hajózásnak jelentős szerepe volt elterjedésében. Magyarországon 1% körüli az előfordulása.

Az **M haplocsoport** (P256): a melanézok (pápuák, Salamon-szigetekiek stb.) között gyakori, valószínűleg a mezőgazdasággal együtt (táró- és jam-gyökér) terjedt el a szigetvilágban. Az M haplocsoport kb. 11 ezer éve alakult ki.

Az **S haplocsoport** (M230): Indonéziában, a pápuáknál és Óceániában fordul elő. Az M és S csoportok közös őse jelenlegi ismereteink szerint kb. 43 ezer éve vált el a P csoporttól, amely a Q és R csoportok őse. Ez nem kevesebbet jelenthet, mint hogy az Európában domináns R1a és R1b haplocsoportokhoz tartozó, és a Q haplocsoportokhoz tartozó amerikai őslakosok távoli őse valamikor Délkelet-Ázsiában élt.

Az **NO haplocsoport** (M214): az N és O csoportok közös őse, ősi formájában (NO*) alacsony gyakorisággal Északkelet-Ázsiában fordul elő (tunguzok, koreaiak, japánok). 37 ezer éve jelenhetett meg Közép- vagy Belső-Ázsiában.

Az **N haplocsoport** (M231): kb. 22 ezer éve jelent meg a mai Kína területén (Yunnan tartomány) a jégkorszak végén. Szibéria lakóira és az uráli népekre jellemző, de Kínában és Közép-Ázsiában is előfordul. Alcsoportjai:

N1a-M128: kb. 4500 éves, Mandzsúriában és Kína egyes kisebbségei között fordul elő.

N1b-P43: jelenleg a szamojéd népek között a leggyakoribb (a nyenyecék fele és szinte minden nganaszán ide tartozik), de Közép-Ázsiában is előfordul szórványosan, korát 5000 évre becsülik.

N1c-Tat (M46): valószínűleg a Bajkál-tó vidékén jött létre 14 ezer éve. A finn-permi és a balti népek között nagyon gyakori (a finneknél 60% fölött). A másik fő elterjedési területe Kelet-Szibéria, ahol a férfiak több mint fele ebbe a csoportba tartozik (jakutok, csukcsok, ázsiai eszkimók, burjátok). Magyarországon ritka, 2% alatti az előfordulása, de a székelyeknél az eddigi minták alapján eléri a 6%-ot.

Az **O haplocsoport** (M175): a kelet-ázsiai (mongolid) férfiak többsége az O csoportba tartozik, egyúttal szinte minden ebbe a csoportba tartozó férfi kelet-ázsiai, tehát a kapcsolat „függvényszerű”. Kb. 31 ezer éve jelent meg Közép- vagy Belső-Ázsiában. Nagy alcsoportjai a mezőgazdasággal terjedtek el szerte Kelet-Ázsiában.

O-M119: a tajvani őslakosok, illetve az ausztronéz nyelveket beszélő népek (filippínók, malájok, indonézok) jellemző haplocsoportja, de Kínában (15%) és Mongóliában is gyakori. Jelenlegi ismereteink szerint 13 ezer éve jelent meg, majd ezt követően terjedt el.

O-M95: az ausztróázsiai nyelveket beszélő népekre jellemző (khmerek, vietnámiak, mundák), de az őslakosok egész Délkelet-Ázsiában elterjesztették, 11 ezer éves ez a csoport.

O-SRY465: Koreában és Japánban gyakori (15-20%) csoport, amely 27 ezer éve válhatott el a többi O csoporttól.

O-M122: a 28 ezer évesre becsült csoport tagjait a szakirodalom a rizstermesztés elterjesztőivel köti össze Kelet-Ázsiában, ma már természetesen nemcsak rizstermesztő népességben található meg. Az O-M122 haplocsoportba tartoznak még japánok (20%), mongolok (18-23%), ujugrok (12%), kazahok (9%) és üzbégek (4%) is. Az O-M134 alcsoportja a sino-tibeti népek fő haplocsoportja, a kínaiak negyede ide tartozik (a kínai férfiak összesen több mint fele O-M122). Az O-M7 alcsoport a hmong-mien (miao-jao) népekre jellemző.

A **P haplocsoport** (M45): kb. 32 ezer éve jött létre, jelenlegi ismereteink szerint Dél-Szibériában, a Q és R csoportok közös őse.

A **Q haplocsoport** (M242): kb. 30 ezer éve jelent meg, jelenlegi ismereteink szerint Dél-Szibériában. A jég visszahúzódásával a Q férfiak Szibérián át Észak-Amerikába vándoroltak még a jégkorszak vége előtt. A legtöbb amerikai őslakos a Q-M3 alcsoportba tartozik (aztékok, maják, inkák stb.), kivéve a na-dené indiánok egy részét. A Q többi alcsoportja kismértékben egész Euráziában elterjedt. Az afganisztáni hazara népcsoport egyik meghatározó komponense. Európába feltehetően az avarokkal és a türk népekkel érkezhettek (a horvátországi Hvar szigeten 15%). A magyaroknál 1-3% körüli (az egyes mintavételek eltérő eredményekkel jártak).

Az **R haplocsoport** (M207): Kb. 28 ezer éve jelent meg, jelenlegi ismereteink szerint Dél-Szibériában. Európa, Közép-Ázsia és India leggyakoribb haplocsoportja, ide tartozik a magyarok több, mint fele. A gyarmatosítás óta az amerikaiak többsége is az R alcsoportjaiba tartozik. Fontosabb alcsoportjai:

Az **R2 haplocsoport** (M124): kb. 16 ezer éve jelent meg – a szakirodalom nagyjából egyetértő álláspontja szerint – Közép-Ázsiában és onnan terjedt el Indiában, ahol a férfiak 10-15%-ában fordul elő. Gyakori még egyes kurd csoportok körében és például a csecseneknél a Kaukázusban. A romák nagy része is az R2 csoportba tartozik.

Az **R1a-M198:** A szláv, finnugor és türk nyelvű népek (Lengyelországtól Pakisztánig félkörívben, átlagosan kb. 50%), valamint az indiaiak jellemző csoportja (a mintavételek alapján a felső kasztokban 30-40%). A mai magyar férfiak egynegyede tartozik ide. Eredete és elterjedése szempontjából ez az egyik legvitatottabb csoport. A legvalószínűbb kialakulási helyszínének Nyugat-Íránt tartják, de nem zárható ki teljesen a Baltikum keleti része sem (ld. lejjebb). A csoport korát 8500 évesre becsülik. Alcsoportjai közül az R1a-Z283 (5000 év) a balto-szláv, finnugor és norvég népességekre jellemző, az R1a-Z93 (4700 év) pedig az indoiráni és törökségi nyelvű népekre.

Az **R1b-P25 (M343)**: kb. 20 ezer éve jelent meg a Közel-Keleten vagy a Volga-vidéken. Alcsoportjaiba tartozik a nyugat-európai (és ebből következően az amerikai) férfiak kb. kétharmada, a magyarok kb. 20%-a. Meglepő módon R1b-V88 (12 ezer év) alcsoportú férfiak alkotják a csádi nyelvű afrikai népek többségét (hauszák, észak-kameruniak), ahová őseik a feltételezések szerint Egyiptomon keresztül juthattak el.

Az **R1b-M269**: Az R1b szinte egész Európára jellemző alcsoportja, kb. 6500 éve jelent meg a Volga-vidéken. Három nagy alcsoportja az R1b-Z2105 (L23), az R1b-U106 és az R1b-P312.

Az **R1b-Z2105 (L23)**: Az ún. „keleti” R1b becslések szerint 6100 éve jelent meg, és legnagyobb számban az örmények, albánok, görögök, más balkáni és délnyugat-ázsiai népek, valamint a baskírok között fordul elő. Magyarországon 6% körüli a gyakorisága.

Az **R1b-U106**: kb. 4900 éve jelent meg a germán-lengyel alföldön. A történelmi korokban, a szakirodalom zömének feltételezése szerint jellemzően a germán népekkel, terjedt el egész Európában. Ma Hollandiában a leggyakoribb, Magyarországon 6% körüli a gyakorisága.

Az **R1b-P312**: kb. 4400 éve jelent meg Dél-Németországban vagy a Kárpát-medencében. Többségében alcsoportjaiba tartoznak a baszkok és az egykori kelta, ma újlatin nyelveket beszélő népek Írországtól Portugálián át Olaszországig. Magyarországon 7% körüli a gyakorisága.

3.2.6.2. Y-haplocsoportok Európa korai történelmében

A következő alfejezetekben saját adatgyűjtésemet és az abból készített ábrákat mutatom be. Az már eddig ismertettekből is látható, hogy haplocsoportok aktuális elterjedése és vándorlási útvonala alapján önmagában még nem állapítható meg a mozgások időmélisége. Azaz nem tudjuk, hogy egy adott haplocsoporthoz vagy több haplocsoporthoz köthető migráció pontosan mikor történt, így azt sem tudhatjuk, milyen történelmi korhoz vagy régészeti kultúrához köthető a vizsgált népességmozgás, ami foglalkoztatja a legtöbb történelem iránt érdeklődő embert.

Közhely, hogy a csontok nem beszélnek, de a genetika mégis vándorlásukról, rokonságukról faggatja őket. Talán nem is sikertelenül. A 2010-es években szerencsére egyre gyakoribbá váltak a régészeti csontleletekből sikeresen kivitelezett DNS vizsgálatok, aminek forradalmi hatása lehet a történelem- és nyelvtudományra. A csontleletekből kivont DNS minták vizsgálata ugyanis segít annak eldöntésében, hogy két időben és/vagy térben különböző régészeti kultúra között milyen szegmensekben, milyen mértékű genetikai kapcsolat áll fenn.

A történelmi folyamatok jobb megértését szolgálja az is, ha azonosítani tudjuk, hogy egy népesség demográfiai növekedése, expanziója mikor és hol zajlott le, illetve egyes régészeti leletek alapján egységesnek tekinthető népességben milyen eltérő eredetű genetikai csoportokat és milyen arányban találunk. Mindkét kérdés vizsgálható a genetika segítségével. Meghatározható, hogy egy adott haplocsoport ma élő képviselőinek közös őse hozzávetőlegesen mikor élt, és mikor kezdődött a csoport demográfiai expanziója. Ez utóbbi értelemszerűen akár jelentősen is eltérhet a csoport teljes korától. Ugyanis, ha valamilyen speciális technológiai vagy környezeti változás hatására rohamosan javulnak egy csoport életfeltételei, akkor exponenciális növekedést tapasztalhatunk az utódok számában. A filogenetika tehát új szempontok vizsgálatára is képes.

A folyamat komplexitásának érzékeltetésére jó – bár nem európai populációra vonatkozó – példa a történelem előtti Japán, ahol i.e. 14500 és 300 között feltételezhetően a mai ajnu népességhez hasonló genetikai arculatú, a *dzsomon (Jomon) kultúrát* létrehozó vadászó-gyűjtögető kultúra élt kőkorszaki fejlettségi szinten. I.e. 300 körül kezdtek el a szigetekre bevándorolni a vaskori földműves kultúrával rendelkező *jajoi kultúrát* hordozó csoportok. Ezek jóval magasabb földművelési és fegyverkészítési technológiai szintjük miatt intenzívebb szaporodásra voltak képesek, ki tudták szorítani eddigi élőhelyeikről a dzsomon kultúra hordozóit, bár a csontleletek genetikai adatai alapján kis részben asszimiláltak is azokat. Ugyanakkor a kőkorszaki lakosság genetikai nyomai is bekerültek keveredés útján a jajoi népességbe, így az ezt a kultúrát hordozó emberek génállománya is részt tudott venni a demográfiai expanzióban. Haplocsoportokra lefordítva elmondható, hogy a mai japán férfiak 35%-a tartozik az ajnu és ryukyu-szigeteki őslakosokban domináns D csoportba, míg 52%-uk a Kelet-Ázsia teljes területén domináns O csoportba. Ezt a genetikai szakirodalom migrációs elmélete a fenti népességvándorlással magyarázza. Tehát a kőkorszakban kialakult apai leszármazási vonalak leegyszerűsítve 3:5 arányban fennmaradtak a hódítókkal történő keveredés következtében. A ma élő D haplocsoportba tartozó férfiak jelentős részének közös őse azonban feltételezhetően egy 2300 évvel ezelőtti időszakra tehető vissza, amikor a vaskori demográfiai expanzió a D haplocsoportú férfiak reprodukciós képességét is megnövelte. Ez nem áll ellentétben azzal a feltételezéssel, hogy a legkorábbi D haplocsoportú férfiak akár már 16 ezer éve a szigeten voltak, és nekik ma is lehetnek a demográfiai expanzióban részt nem vett leszármazottjaik.⁹ A genetikai megoszlások tehát nem feltétlenül magyarázhatók pusztán a migrációval [70, 71].

⁹A Japán FTDNA projektben lévő D1b1a-M125 alcsoport (a japánok 14,3%-a) közös őse 2100-2600 évesre becsülhető, míg a D1b1d alcsoporté legalább 3700-4600 év. Így logikusnak tűnik, hogy az eleve több tagot

3.2.6.3. Európa a földművelés előtt

Saját adatgyűjtésem alapján is elmondható, hogy rendkívül kevés genetikai adat áll rendelkezésünkre Európa korai történelméből, de ez a kis számú, független helyszínekről származó adat meglehetősen egyértelmű összhangban áll egymással. Így, ha óvatosan is, de lehetséges bizonyos trendek, észrevételek megfogalmazása.

Az európai őskőkorszakból különösen kevés lelet áll rendelkezésre. Értelemszerűen ez a csontleletek hatalmas időszakot átölelő megfelelő állagmegőrzésének nehézségéből fakad. Oroszország déli részén, a Don-folyónál (Kosztienki falu) egy 36-38 ezer éves lelet a C-F3393 csoportba tartozónak bizonyult [72], míg a mai Románia területén (Oase barlang) ebből az időszakból egy (alcsoportokra nem tesztelt) F* haplocsoportba tartozó mintát találtak [73]. A C-F3393 csoport ma leginkább az ausztráliai és óceániai őslakók között, valamint kismértékben Japánban van jelen. Európában ma ezrelékekben is alig mérhető a gyakorisága, de egy mezolit észak-ibériai leletben (LaBrana) is előfordult még, valamint a földművelés kezdetéről származó kárpát-medencei (Kompolt, Magyarország) mintában is. Ezért felmerült az a feltételezés, hogy Európa első kőkorszaki lakói 40 ezer évvel ezelőtt az ausztrál őslakókhoz hasonló genetikai összetételű emberek voltak, akik Ausztrália benépesítésével egy időben érkeztek Európába. A C-F3393 csoport nem tévesztendő össze a C-M217 csoporttal, amely Északkelet-Ázsiában a mongolok és mandzsu-tunguz népek domináns apai DNS tulajdonságát adja, és a mongol expanzióval a középkorban egész Euráziában elterjedt, de korábbi jelenlétét csontleletekből nem sikerült eddig kimutatni.

A középső kőkorszak (mezolitikum, kb. 11500 éve) időszakából több genetikai leletet is ismerünk, ezeket az I haplocsoport dominálja. Svédországban (Motala) öt 7500-8000 éves lelet közül mindegyik az I2 alcsoportba tartozott, azon belül négy az I2-P37 alcsoportba, egy pedig az I2-L597 alcsoportba [74, 75]. Luxemburgban (Loschbour, szikla menedék) egy 7000 éves minta ugyancsak az I2-P37-es csoportba tartozott [75, 76]. Az I2-P37 (az I2-M26 és I2-M423 közös őse) érdekessége, hogy Nyugat-Európában ma alacsony frekvenciával elterjedt (Franciaországban, Ibériában, Itáliában és a Brit-szigeteken 1-2%), de Szardínián a lakosság 41%-át adja, a baszkok között pedig 6-8% a gyakorisága [77], utalva arra, hogy e népek többet megőriztek a mezolit örökségből. Az I2-P37 legkorábbi előfordulása Európában még a mezolitikum előtti időre, 13600 évvel ezelőttre tehető, egy svájci barlangban (Bichon) találták meg [78].

számláló D1b1a csoport közös őse részt vett a vaskor eleji demográfiai robbanásban, míg a D1b1d alcsoporté nem. Összehasonlításképpen az O csoport tipikus, a japánok 22%-át adó alcsoportjának (O2b1a-47z) kora 2200-2800 évre becsülhető, tehát a D1b1a-hoz hasonlóan összhangban van a jajoi nép Japánba érkezésével.

Az I2-CTS10228 (köznevéen „dinári-kárpáti”) alcsoport legkorábbi közös őse 2200 évvel ezelőttre datálható, így ebben az esetben nem arról van szó, hogy a mezolitikus népesség Kelet-Európában ilyen mértékben fennmaradt volna, hanem arról, hogy egy, a mezolitikus csoportoktól származó szűk család az európai vaskorban (i.e. 1200 év) sikeresen integrálódhatott egy olyan társadalomba, amely hamarosan erőteljes demográfiai expanzióba kezdett. Ez is mutatja, hogy nem feltétlenül népek, hanem gyakran családok sikerével, nemzetségek elterjedésével is számolni kell, és ezt a jelenlegi etnikai identitással összefüggésbe hozni lehetetlen. A csoport elterjedése alapján valószínűsíthető, hogy részt vett a szláv népek migrációjában, ilyen módon vált az R1a után legdominánsabb csoporttá a mai Kelet-Európában. Nyugat-Európából viszont teljes mértékben hiányzik I2-CTS10228 alcsoport, kivéve a kora középkorban szláv nyelvű kelet-német területeket. A ős- és középső kőkorszakból származó (paleolitikus és mezolitikus) csontleletekből meghatározott Y-haplocsoportok a **16. ábrán** kerültek feltüntetésre.



16. ábra. Az európai paleolitikus/mezolitikus csontleletek Y-haplocsoportjai
A leletek származási helyei a jelenlegi országok területén vannak feltüntetve,
a színes alakzatok az Y haplocsoportokat jelölik.

Teljesen más kép rajzolódik ki Északkelet-Európa mezolitikumában, ahol Karéliában egy 7000-7500 éves [75], Szmolenszk környékén pedig egy 6000 éves minta tartozik az R1a csoportba [79]. A Volga-kanyarulatban pedig Szamara vidékén egy 7500 éves R1b mintát találtak [75, 80]. Mivel az R1a és R1b a bronzkorban kapott jelentős szerepet Európa benépesítésében, ezért további történetükről ott esik szó.

3.2.6.4. A földművelés Európába érkezése

A történészek között hosszú ideig vita tárgyát képezte, hogy Európába a földműveléssel együtt új népesség is költözött-e, vagy alapvetően a helyi lakosság tanulta meg az új technológiát kulturális diffúzió útján. Már a legelső csont-DNS vizsgálatok nyilvánvalóvá tették, hogy nagy népmozgások zajlottak a prehisztórikus Európában, ugyanakkor a kutatások bizonyítják a régebbi lakók bizonyos fokú továbbélését is a megváltozott kulturális környezetben. Genetikai adatok alapján tehát egyik elmélet sem talált teljes megerősítést a másikkal szemben, mindkét hipotézis létjogosultsága – bár eltérő mértékben – de igazolódni látszik. Ez jól példázza, hogy a nagy ívű történeti modellek és az ezekkel ellentétes más elképzelések nem feltétlenül kizárják, hanem gyakran inkább kiegészítik egymást. Érdeemes odafigyelni az arányokra is, ezért a következőkben röviden tekintsük át az értelmezés nehézségének szempontjából nagyon tanulságos folyamatot.

A Starčevo-Körös kultúra (i.e. 5800-4500) a Balkánról a Duna-medencébe érkező földművelőket jelzi. A Starčevóban vizsgált 13 Y-DNS mintából 8 fő a G2a haplocsoportba, 3 fő az F*, 1 fő az I-M170 és további 1 az I2-P37 (dinári) haplocsoportba tartozott [81]. A Starčevóból észak felé induló, a dunántúli *vonaldíszes kerámia* (*Linear Pottery*) kultúrából vizsgált 4 fő közül 2 fő G2-P15, 1 fő F* és 1 fő I1 haplocsoportú volt [81]. A neolit Kárpát-medencét vizsgáló másik tanulmány [82] a *Körös kultúrában* egyetlen I2-P37-es mintát tudott azonosítani, míg a középső neolitikumban 1 db I2-P37-et és 2 db C-F3393-at. A mai Németország területéről a *vonaldíszes kerámia* (*Linear Pottery*) korai korszakából, 7000 évvel ezelőttről 4 db G2-P15 és 1 db T lelet ismert, a középső neolitikumból 1 db I2a-P37 és 2 db olyan minta, amelyet nem sikerült elegendően sok SNP-re tesztelni [75]. Az i.e. 5000-3400 között virágzó *lengyel kultúrából* 1 db J2-M172, 1 db E1b-M78, 1 db C-M216 haplocsoportba tartozó minta került elő [83].

A földművelés másik elterjedési vonala a Földközi-tenger térségében volt, ahol a *benyomósos kerámia* (*Cardium Pottery*) kultúrája terjedt el i.e. 6400 körülől. A kultúrához tartozó 6 vizsgált ibériai férfi közül 5 fő a G2-P15, 1 fő pedig az E1b-V13 alcsoportba tartozott [84]. Egy másik tanulmány az ibériai neolitikumban 5900-7200 évvel ezelőttről 3 db I2-P37, 1 db F*, valamint 1 db R1b-V88 mintát azonosított [75]. 5000 éves dél-franciaországi sírokból származó 22 férfi leletéből 20 fő G2-P15, 2 fő pedig I2-P37 haplocsoportba tartozott [84].

A földművelés duna-völgyi és földközi-tengeri elterjedése azt mutatja, hogy a korai földművelőkre jellemző G2a, I2 és F* haplocsoport relatív többsége újonnan érkezett

jövevény volt Európában, akik eredetileg a Közel-Kelet és a Kaukázus vidékéről származtak. Ezt támasztja alá a mai Törökország területén feltárt korai földművelők csontleleteinek mintáiból végzett vizsgálat, melyek közül 9 db G2a, 2 db H2 (korábban ez az F*-hoz tartozott), valamint 1-1 db J2a, C1, I2b és I* haplocsoportú minta került elő [80]. A G2-P15 alcsoport ma egyértelműen a Kaukázusban a legmeghatározóbb, nagy valószínűséggel onnan is származik. Az F* (azaz M89 mutációt igen, de a G, H, I, J és K csoportok mutációit nem hordozó) csoport ma leggyakrabban egyes közép-indiai nomád törzseknél fordul elő, előfordulása Európában egy ezrelék alatti. A neolit csontleletből meghatározott Y-haplocsoportok a **17. ábrán** kerültek feltüntetésre.



17. ábra. A korai neolit csontleletekből meghatározott Y haplocsoportok¹⁰

A feltárt leletekhez tartozó régészeti kultúrák fel vannak tüntetve a térképen, az alakzatok formája az Y haplocsoportokat, a kitöltő szín az alcsoportokat jelöli.

¹⁰https://www.eupedia.com/genetics/starcevo_culture.shtml (módosítva)

A mai európai népességben a korai földművelők közvetlen férfi leszármazottjai csekély számú kisebbséget alkotnak, Szicílián és Korzikán kívül arányuk sehol nem éri el a 10%-ot. 5 és 10% között vannak jelen Olaszországban, Görögországban, Svájcban és Ibéria nyugati részén. 3-5% közé esik a gyakoriságuk Ibéria többi részén, Németországban, Magyarországon, Bulgáriában és Romániában, valamint a Volga-Káma vidékén a tatárok és baskírok között. Európa többi részén 1-3% közötti az előfordulásuk [83].

A neolitikum kapcsán érdemes kitérni az I1-M253 csoport történetére is. Az I1-M253 eddigi ismereteink szerint a dunántúli *vonaldíszes kerámia (Linear Pottery) kultúrájában* fordult elő először 7000 évvel ezelőtt, tehát elképzelhető, hogy a csoport később vándorolt a Kárpát-medencéből észak felé. Feltételezhető, de szintén nem bizonyítható, hogy az I1 esetében is arról van szó, hogy egy kis mezolitik csoport átvette a földművelés technológiáját és így fenn tudott maradni a genetikai öröksége. Mivel a csoport ma élő férfitagjainak közös őse STR és SNP alapú becslések szerint is viszonylag fiatal kb. 3500-4000 év [85, 86], mai eloszlása már egy bronzkori dél-skandináviai demográfiai robbanás és vaskori migráció eredménye lehet.

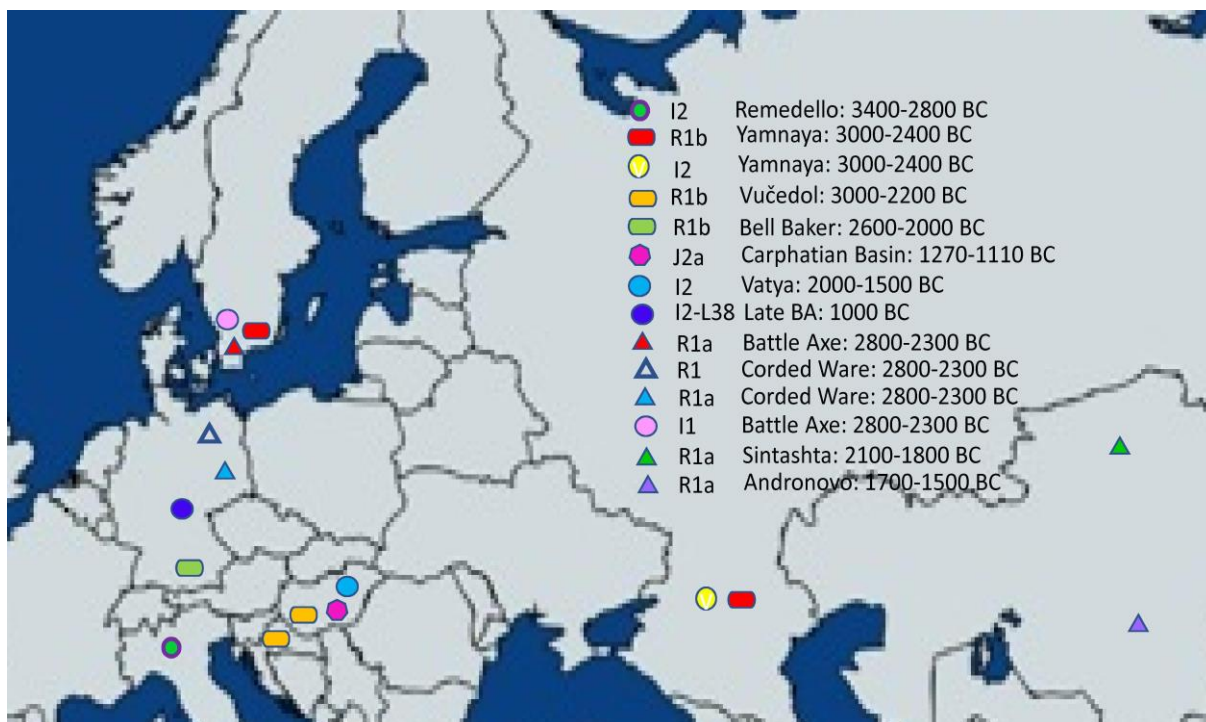
Északkelet-Európában a kezdődő neolitikum is eltérő haplocsoport eloszlási képet mutat, a nyugat-oroszországi Pszkov és Szmolenszk megyékben feltárt 3 db 4500 éves csontlelet közül 2 fő az R1a-M198, 1 fő pedig az N-Tat csoportba tartozott, amely párosítás ma az oroszországi finn-permi népekre, valamint az őket asszimiláló észak-oroszokra jellemző [79].

3.2.6.5. A bronzkor – Európa mai genetikai képének kialakulása

A bronzkorban (i.e. 3500) kapott jelentős szerepet az eddigi részekben ritkán említett Északkelet-Európa, azaz a mai Oroszország európai területe. Áttörést jelentett Európa benépesülésének kutatásában, amikor Szamara környéki 5000-5200 éves gödörsírokban – az ún. *gödörsíros kultúra (Kurgan vagy Yamnaya culture)* sírleleteiben – hét mintából mind a hét a ma Nyugat-Európában gyakori R1b haplocsoportba tartozónak bizonyult [75]. Egy másik kutatás pedig a mai Volgográdban és attól délre hat mintából ötöt R1b és egyet I2-M223 haplocsoportba tartozónak talált a *jamnája* (gödörsíros) kultúrában¹¹ [74]. A kutatások nemcsak a jamnája kultúra képviselőinek döntően R1b mivoltát igazolták, hanem azt is, hogy

¹¹ A gödörsíros vagy más néven jamnája kultúra a korai bronzkorban (i.e. 3000-2400) foglalta el a Fekete-tengertől északra és keletre húzódó sztyeppeövezetet a Dnyeszter folyótól a Dél-Urál-hegység nyugati előteréig, azaz a mai Moldovától Dél-Baskíriáig. A kultúra hordozói nagyállattartó népek voltak, akik kismértékben földet is műveltek a folyóvölgyekben, fejlett fémművességük és fazekasságuk volt. Halomsíraikban az elhunytat okkerrel fedték be.

az európai bronzkori kultúrák népességére a jamnája kultúra jelentős genetikai hatást gyakorolt. Hozzá kell tenni, hogy a Szamara környéki hét R1b közül öt (ld. 18. ábrát) az R1b-Z2105 (R1b-L23), ún. „keleti” R1b alcsoportba tartozott; a két másik minta pedig a még keletre és nyugatira szét nem vált, ősi R1b jelleget mutatta. A „keleti” alcsoport ma a baskírok és az örmények között a leggyakoribb, de előfordul a Kaukázusban, Anatóliában, valamint kisebb gyakorisággal a Balkánon, a szláv és a finnugor népek között is [87]. Ezzel szemben a nyugat- és közép-európai, valamint albán R1b haplocsoportba tartozó férfiak többsége az R1b-L51/M412 alcsoportba tartozik, amely először bizonyítottan a közép-európai bronzkorban fordult elő [88]. A bronzkori csontleletekből meghatározott Y-haplocsoportok a **18. ábrán** kerültek feltüntetésre.



18. ábra. A bronzkori csontleletekből meghatározott Y haplocsoportok

A leletek származási helyei a jelenlegi országok területén vannak feltüntetve, az alakzatok formája az Y haplocsoportokat, a kitöltő szín az alcsoportokat jelöli.

A jamnája népesség útja Közép-Európa felé genetikailag dokumentált, mert a legfrissebb kutatások szerint az R1b haplocsoport már a mai Horvátország és Magyarország területéről feltárt *Vučedol kultúrában* (i.e. 3000-2200) is előfordult [83], továbbá a *harangedényes* (*Bell Beaker*¹²) *kultúrájának* németországi leleteiben kizárólagosan R1b

¹² A korai bronzkorhoz tartozó harangalakú edények kultúrája i.e. 2600-2000 között terjedt el egész Közép- és Nyugat-Európában (Kárpát-medence, Csehország, Németország és Alpok, Itália, Ibéria, Gallia, Brit-szigetek)

fordult elő [83]. A *harangedényes kultúra* (Bell Beaker) i.e. 2600-2000 között nagyon gyorsan elterjedt egész Nyugat-Európában, ahol az R1b haplocsoport ma is a férfiak több mint 50%-át adja (az Atlanti-óceán partján pedig több helyen a 80%-ot is eléri) [89]. Az eredmények azt is jól mutatják, hogy az R1b a korábbi mezolit és neolit népességet szorította ki, mert a *harang alakú edények* (Bell Beaker) előtti észak-itáliai *Remedello-kultúra* (i.e. 3400-2800) 3 mintája közül mindhárom a ma Szardínián leggyakoribb I2-M26 alcsoportba (I2-P37-en belül) tartozott. A bronzkor előtti Kárpát-medencében Allentoft és mtsai, 2015 [74] adatai szerint a vizsgált minták között 1 db alcsoportra nem tesztelt I2, 1 db I2-P37, 1 db I2-M223 és 1 db G2-P15 volt az arány. Az R1b után később még újabb haplocsoportok is megjelentek, a Kárpát-medenceében egy darab késő bronzkori (i.e. 1270-1110) J2a mintát találtak [82].

A genetikai adatok alapján is valószínűsíthető, hogy a *jamnája kultúrát* hordozó népesség expanzív volt, már korán beolvastott magába más, a mezolitikum óta Közép- és Kelet-Európában élt népességeket is, erre utal az 1 db I2-M223 jelenléte a Don és Volga közötti területen [74]. Ugyanez a csoport bukkant fel a bronzkori *Vatya-kultúrában* (i.e. 2000-1500) Magyarországon, valamint a bronzkori horvátországi és magyarországi *Vučedol kultúrában* is az R1b-vel együtt [83]. A ma nagyon ritka I2-L38 csoport legnagyobb jelentőségű csontleletét Közép-Németországban találták meg, a lelet i.e. 1000 körülre (*Late Bronze Age: Late BA*) datálható [90]. Az I2-M223 alcsoportok mozgása egyelőre nem vázolható fel pontosan, de úgy tűnik, hogy a mezolitikumban széles körben elterjedt I2-P37 „fivérével” szemben az I2-M223 döntően inkább az R1b férfiakkal együtt terjedt Európában [90].

Ugyancsak jamnája genetikai hatást mutat a közép-európai *harangedényes* (Bell Beaker: i.e. 2600-2000) *kultúra* mellett a *zsinórdíszes kerámia* (Cord Ware, i.e. 2800-2300) *kultúrája* is, amely a bronzkor elején a mai Dánia és Oroszország középső területei között terjedt el a Baltikum déli részén [74]. Allentoft és mtsai, 2015 által publikált adatok szerint az innen származó minták között 4 db R1a-M417, 1 db R1b-P25 található és 1 db R1 a kultúra svédországi kiterjesztését jelentő *csatabárdos* (Battle Axe, i.e. 2800-2300) *kultúrából* pedig már a ma széles körben (Skandináviától Indiáig) elterjedt R1a-Z645 még differenciálta szakaszát testesíti meg. Korábban Németországban 3 db egymással rokon R1a-M198 mintát találtak a *zsinórdíszes* (Cord Ware) korból [91]. A *zsinórdíszes kerámia kultúrájában* fellelhető R1a-dominancia összhangban van az R1a csoport mai elterjedésével, hiszen jelenleg Skandináviában és Észak-Németországban 12-24%, a nyugati és keleti szlávoknál pedig 40-55% a gyakorisága [92].

A Fekete-tenger északi előterében az i.e. 2800-2300 között virágzó ún. *csatabárdos kultúrában* (*Battle Axe*) a már említett R1a haplocsoportba tartozón kívül még 2 db R1b, 1 db R1b-U106 és 2 db I1-M253 mintát is találtak [74].

Dél-Urálban (Közép-Ázsia) a bronzkori *Szintasta-kultúrából* (*Sintashta*, i.e. 2100-1800) és az azt követő *Andronovo-kultúrából* (*Andronovo*, i.e. 1700-1500) feltárt csontminták mindegyike R1a-M198 csoportba tartozó férfitől származik, az R1b nem volt megtalálható [74].

A ma rendelkezésre álló csont-DNS eredmények azt valószínűsítik, hogy Európa férfi génállománya a bronzkor végére már sok szempontból a maihoz hasonlóvá vált, jelentős mértékű külső hatás már nem érte, a vaskorban (népvándorlás) inkább belső átrendeződés történt. Egy nemzetközi együttműködésben végzett saját kutatás alapján, mely során az Y-kromoszóma új generációs szekvenálása (NGS) történt, elmondható, hogy a mai görög, török és lapp (számi) népek kivételével minden vizsgált európai népesség¹³ 2100-4200 éve élte meg népességminimumát és ezután kezdett folyamatos demográfiai expanzióba [85]. Ez középként az i.e. 1200 körüli időszakot jelenti, amikor az *urnamezős kultúra* (*Urnfild culture*) megjelent Közép-Európában, valamint komoly migrációs hullámok is végigsöpörtek Dél-Európán és a Közel-Keleten (ún. *Bronze Age collapse*).

Összegzésként tehát elmondható, hogy Európa benépesülése négy meghatározó migrációs hullámmal írható le. Ez természetesen nem jelenti azt, hogy nem voltak más, demográfiai szempontból kisebb jelentőségű Európát érő migrációk, amelyek ne árnyalnák az alábbiakban vázolt átfogó képet.

A genetikai adatok azt látszanak megerősíteni, hogy az első hullámmal, olyan őskőkorszaki C-F3393 népesség érkezett Európába, amely az Ausztráliát is meghódító ősi embercsoportokkal mutat rokonságot. A mai európai népességben való sporadikus jelenlét és az időbeli távolság nagy volta miatt értelmetlennek tűnik e csoport lehetséges nyelvének vizsgálata.

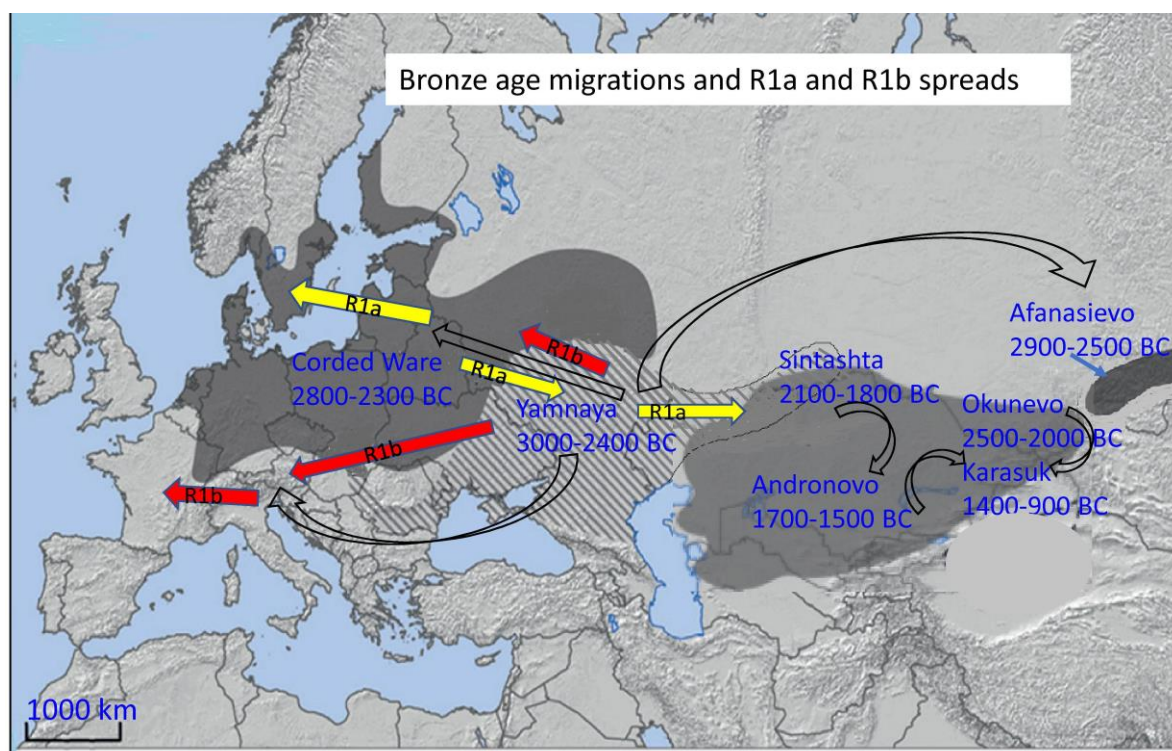
Érdekesebb kérdésnek tűnik a mezolitikumban meghatározó szerepet játszó I haplocsoportú népesség nyelve. Mivel a latin előtti szárd nyelvet egyesek a baszkkal hozzák párhuzamba, az I2-P37 archaikus példái pedig ma Szardínián és a baszkok között találhatók meg legnagyobb mértékben, elképzelhetőnek tartják [93], hogy az I2-P37 haplocsoportot hordozó mezolit népesség egyetlen mai leszármazottjai a baszk nyelvet beszélők lennének Európában (máshol ugyanis a haplocsoportot hordozó családok apai ágon kihaltak). Mindez

¹³ Ezek: norvég, dán, fríz, angol, ír, Orkney-szigeteki, baszk, spanyol, francia, toszkán, bajor, magyar, szerb.

azt is jelentené, hogy Szardínián nagyobb arányban kellene találnunk baszk eredetű helynév anyagot, de ennek legjobb tudomásunk szerint nincs nyoma.

A korai földművelők csontanyagát a Kaukázus vidékéről származó G2a haplocsoport jellemzi, és ezért hipotézisként a populációgenetikai szakirodalomban megfogalmazták, hogy a Mezopotámiából származó korai földművelők esetleg valamilyen kaukázusi nyelven beszéltek [80, 94]. 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 ez az arány, Szicília és Korzika szigetén pedig 10 és 15% között mutatható ki a jelenlétük [94].

Végezetül – a leginkább elfogadott populáció-genetikai hipotézis szerint – a mai Európa népességének összetétele döntő mértékben a kora bronzkori, feltehetően indoeurópai nyelveket beszélő népek migrációs hullamai következtében alakult ki, párhuzamosan a korábbi mezolitikumi népesség jelentős eltűnésével. A bronzkori európai régészeti kultúrák hordozóit jellemző R1a és R1b haplocsoportok együtt alkotják a mai európai férfi lakosság abszolút többségét (19. ábra).



19. ábra. A bronzkori humán migrációk és az R1a és R1b haplocsoportok elterjedése¹⁴
Az üres fekete nyilak az adott régészeti kultúrák terjedését/kapcsolatát jelölik.

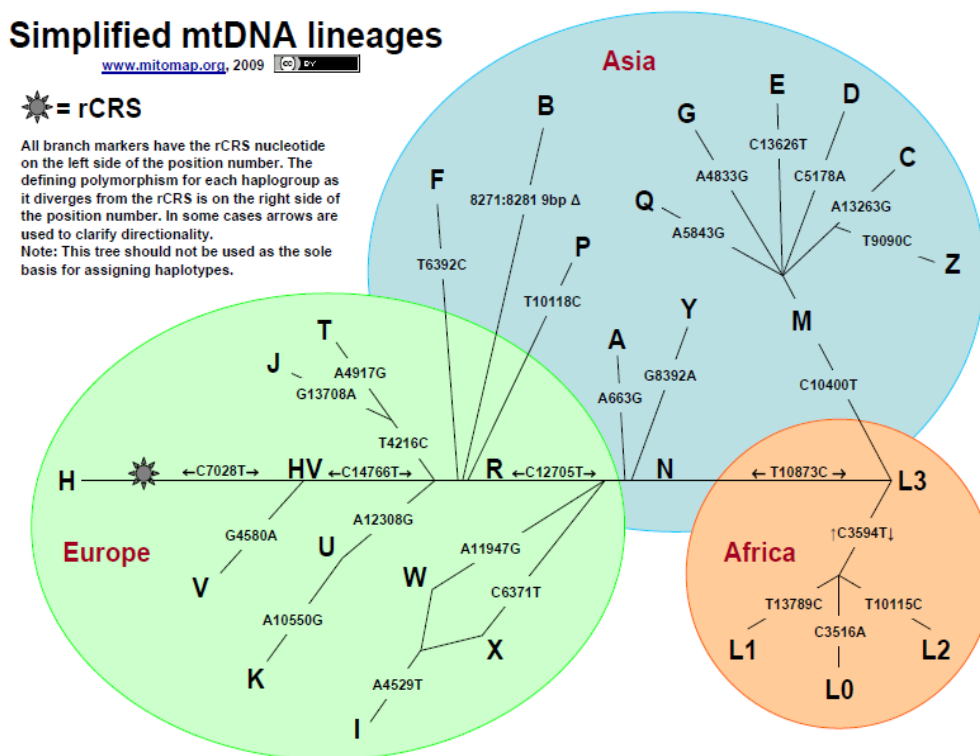
¹⁴ Allentoft et al. 2015 (módosítva)

A lovakat korán házasító kelet-európai sztyeppei ló- és nagyállattartó népek nyugat-délnyugati irányban meghódították egész Európát, a bronzkor végére (i.e. 1200) háttérbe szorítva az ottani korai földművelő és halász-vadász népcsoportokat. Az indoeurópaiktól eltérő nyelvek Európában a középkorra jórészt csak a *gödörsíros kultúrától* északra fekvő erdőövezetben (uráli nyelvek), valamint Baszkföldön és a Kaukázusban maradtak meg [93, 95]. Ez a nyelvészeti adat azonban nem feltétlenül hozható összefüggésbe a genetikai adatokkal, hiszen tudjuk, hogy a görög-római történelem korszakában és azt követően megannyi más nyelvű nép is megfordult Európában, tehát a forrásokból is ismert történelmi folyamatok nem támasztják alá ezt a leegyszerűsítést. A jelenlegi nyelvi megoszlásból visszakövetkeztetni nem lehet, hiszen pl. magyarok is a középkor folyamán érkeztek a Kárpát-medencébe, és nyelve sem indoeurópai.

A jelenleg rendelkezésre álló genetikai adatok alapján Európa népességének gyakori „kicserélődésére” kell következtetnünk, ami szintén további vizsgálatra szoruló feltételezés.

3.2.7. A mitokondriális DNS (mtDNS) haplocsoportok rövid áttekintése

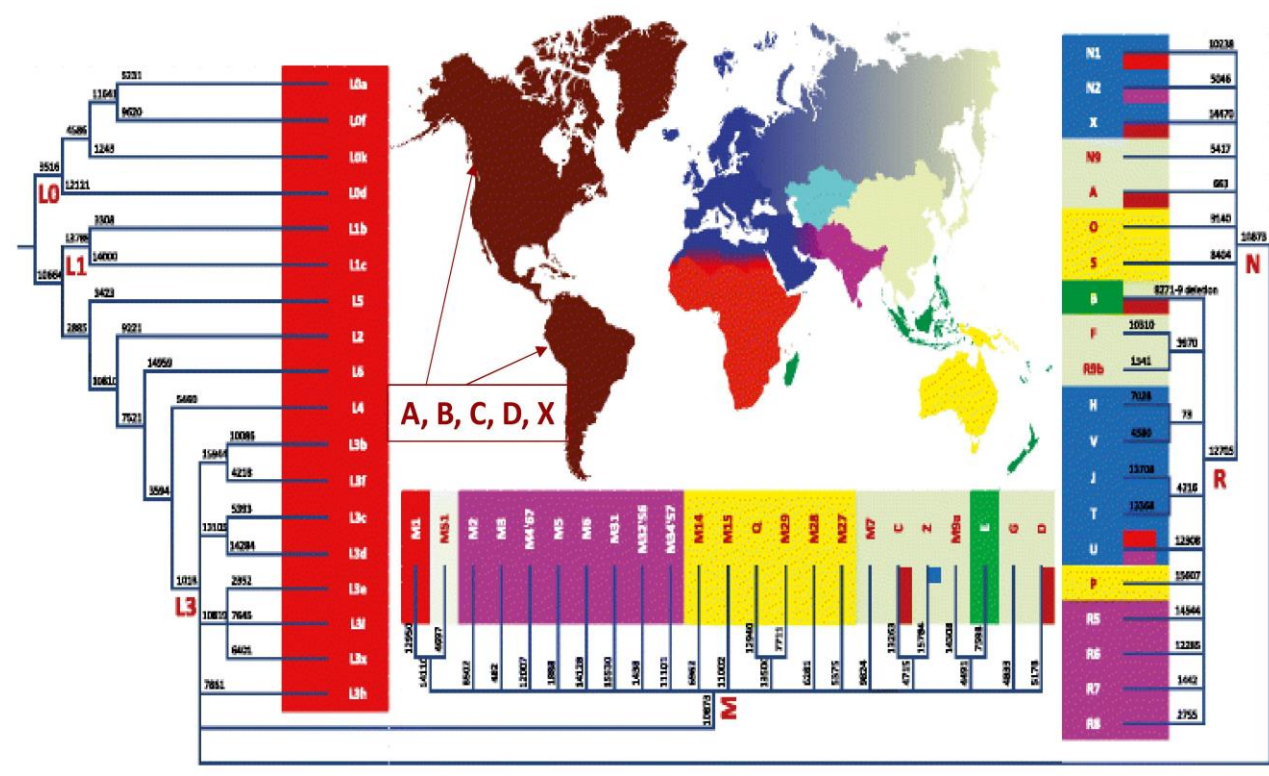
Az Y-kromoszómális haplocsoportokkal szemben az anyai ági leszármazást jelző mtDNS haplocsoportok esetében az elnevezések sajnos nem követnek leszármazási logikát, azaz nem az „A” csoport az összes többitől legkorábban levált csoport, hanem a csoportok felfedezési sorrendjében nevezték el azokat. Emiatt a következő, fastruktúrának megfelelő történeti áttekintés nem ABC-sorrendben történik, hanem a csoportok vélt „ősiségének” és elterjedésének sorrendjében (**20. ábra**). A struktúrát és a korbecsléseket is közöljük [40, 57, 96]. Figyelembe kell venni, hogy az mtDNS vonalak elterjedése ritkábban köthető konkrét népmozgásokhoz, mint az Y-kromoszómális vonalaké, így a mitokondriális haplocsoportok népek és nyelvcsaládok helyett gyakran inkább földrajzi régiókhoz köthetők. Az alábbiakban ismertetem a saját adatgyűjtés alapján készült a mtDNS haplocsoportok eloszlására és elterjedésére vonatkozó összefoglalót.



20. ábra. A mtDNS haplocsoportok kialakulása, előfordulása és az adott haplocsoportra jellemző mutációk¹⁵

¹⁵ (https://www.google.hu/search?q=discovery+of++mtDNA+lineages&client=firefox-b&biw=1536&bih=708&source=lnms&tbm=isch&sa=X&ved=0ahUKEwiw4LHDIYPSAhUJKMAKHfE8DbEQ_AUIBigB#imgsrc=n-d1CTi7FbDG5M:)

Az **L haplocsoport** (L0, L1, L2, L3, L4, L5, L6): a legősibb női ág, amely Afrikában és Afrikából elszármazott népesekben van jelen. A ma élő összes női vonal közös ősének, a kb. 177 ezer éve élt „mitokondriális Évának”, haplocsoportja is ez. Nagyon ritkán előfordul Afrikán kívül is, elszigetelt helyi népesekben. Az Afrikán kívül elterjedt M és N csoportok (21. ábra) az L3-as alcsoportból váltak ki 67 ezer éve.



21. ábra. A mtDNS haplocsoportok filogenetikai fája¹⁶

Az egy adott ágat definiáló markerek a kódoló régióból (számok a pozíciói) színes háttérrel vannak kiemelve. A színek a térképen a haplocsoportokat jelölik

Az **M haplocsoport**: kb. 50 ezer éve jelent meg Indiában. Számos alcsoportja van, amely nem kapott külön betűs jelölést, csak számokat. Az önálló betűkkel jelzett alcsoportjait (**Q, G, E, D, C, Z**) külön kiemeljük (20. ábra). Az M számozott alcsoportjai Európában alig fordulnak elő, viszont az európai romák között a 30%-ot is elérhetik.

M1 alcsoport: 24 ezer éves, Észak- és Kelet-Afrikában van jelen, valószínűleg az emberek Ázsia felől Afrikába történt visszaáramlására utal. Közeleli rokona a Kínában előforduló M20 és a Kambodzsában előforduló M51.

¹⁶ Kivisild T (2015). "Maternal ancestry and population history from whole mitochondrial genomes". Investig Genet (módosítva)

Az **M2, M3, M4, M5, M6, M30, M33, M34, M35, M39, M40, M41**: Dél-Ázsiára (India) jellemző ágak. Pakisztán és India női ágainak 47-65%-a esik valamely M alcsoportba.

Az **M31 és M32**: az Andamán-szigeteken jellemző csoportok.

Az **M7** 45 ezer éves, főként Kelet-Ázsiában terjedt el (Japán, Kína, Vietnam, Laosz)

Az **M9, M11, M14, M15**: főként Kínában és Tibetben fordul elő. A 39 ezer éves M9 alcsoportja az **E**, melynek kora 24 ezer év és az ausztronéz migrációval terjedt el (a Fülöp-szigeteken 20%, Tajvanon, Malajziában és Indonéziában 10-12%-os az előfordulása).

Az **M17, M19, M21, M24, M80**: Délkelet-Ázsiában előforduló csoportok.

Az **M27, M28, M29 és Q**: Melanéziában és Pápua Új-Guineában előforduló alcsoportok. A Q kora 38 ezer év, a pápuák 29%-a tartozik ide, az ausztrál őslakóknak pedig 3%-a.

Az **M42**: 48 ezer éves, az ausztrál őslakók között előforduló csoport.

Az **M23**: Madagaszkáron előforduló 6000 éves csoport (Indonéziából érkezhettek).

Az **M8**: Kelet-Ázsiában és Dél-Szibériában kis gyakorisággal, Kamcsatkán gyakrabban előforduló 37 ezer éves alcsoport. Leszármazottja a C és Z közös ősenek számító 36 ezer éves CZ csoport.

A **C haplocsoport**: 24 ezer éves, az eddigi szakirodalmi adatok szerint a Bajkál-tó környékén jöhetett létre. Északkelet-Szibériában (jukagirok, jakutok, evenkik) 62-67%-os gyakoriságú női vonal. A tuvaiaknál és nganaszánoknál 50% az előfordulása. 30%-nál kisebb arányban egész Belső- és Közép-Ázsiában, továbbá az amerikai őslakók között, valamint Kelet-Európában is előfordul. Gyakoriságát a székelyek között 5%-nak, a csángóknál 0,5%-nak mérték az eddigi mintavételek alapján.

A **Z haplocsoport**: 22 ezer éves lehet, egész Euráziában elterjedt, de nagyon ritka. Gyakorisági maximuma a kazahok között (11%) van, jellemző még a csukcs-kamcsatkai népekre is 6-7%-ban. Számos uráli és altáji népben előfordul 5% alatti gyakorisággal, de a magyarok között az eddigi mintavételek alapján nincs jelen.

Az **M12**: kb. 31 ezer évvel ezelőtt keletezett, a G csoport „nővére”, Kelet-Ázsiában fordul elő.

A **G haplocsoport**: 32 ezer éves, a csukcs-kamcsatkai népek jellemző csoportja (27-68%). Más szibériai és belső-ázsiai népek között jellemzően 5-12%-os gyakoriságú. Ritkán az obi-ugorok és székelyek között is előfordul (aránya az 1%-ot nem éri el az eddigi mérések alapján).

A **D haplocsoport**: a becslések szerint 38 ezer éves, az M-nek önálló, kelet-ázsiai alcsoportja. A koreaiak, japánok és mongolok között, valamint a hsziungnu (ázsiai hun) csontleletekben 30-40%-os gyakoriságú. A többi türkségi és mandzsu-tunguz népben, valamint a kínaiak és

tibetiek között 10-25%-ban fordul elő. Jelen van az amerikai őslakosok között is. Közép-Ázsiában gyakorisága 7% alá, Kelet-Európában 2% alá csökken, a magyarok között eddigi ismereteink szerint nincs jelen.

Az **N haplocsoport**: becsült kora 59 ezer év, ekkor vált el az L3 csoport többi ágától, Afrikán kívül az egész világon elterjedt alcsoportjaival együtt. Az M-hez hasonlóan egyes alcsoportjait csak számok jelölik, míg mások saját betűjelet kaptak. Őshazája az M és R „nagy csoportokhoz” hasonlóan talán India lehet.

N1 alcsoport: a Közel-Keletre jellemző 52 ezer éve keletkezett haplocsoport. Egyes alcsoportjai Indiába (N1d) és Kelet-Afrikába (N1a) is eljutottak. Arról nevezetes, hogy a korai európai földművelők között a neolitikumban gyakran előfordult (a Kárpát-medencében még a honfoglalók között is jelen volt), a mai lakosságban azonban ritka. A csángóknál közel 4%-os előfordulása.

Az **I haplocsoport**: Az N1 önálló betűt kapott, kb. 21 ezer éve keletkezett alcsoportja. Valószínűleg a Közel-Keletről terjedt el széles körben. A kenyai kusita pásztornépek között 8-23%-os előfordulása, de Afrikában máshol nincs jelen. 2-5% közötti előfordulása Iránban, Afganisztánban, Azerbajdzsánban és Törökországban, 1% alatti Dél-Ázsiában. Gyakorisága Finnországban egy tanulmány szerint 10%-os, ezen kívül Skóciában és a környező szigeteken, Izlandon és Norvégiában, valamint Franciaország egyes részein haladja meg a 4%-ot. Mivel viking csontleletek között 10% fölötti gyakoriságú, Észak-Európában vélhetően velük terjedt el. Kelet-Európában jellemzően 0-3% közötti arányban található meg, két elszigetelt népességben (Krk szigeti horvátok, kárpátaljai lemkók) gyakorisága eléri a 11%-ot.

N2 alcsoport: 45 ezer éves, ritka alcsoport, nagyobb mértékben az önálló betűt kapott W alcsoportja terjedt el.

A **W haplocsoport**: 18 ezer évesre becsült (tehát az utolsó jégkorszakban keletkezett), magyar szempontból különleges, ugyanis a magyarországi magyarok (4%), székelyek (5%) és csángók (8%) között is viszonylag gyakori. Hasonló frekvencia értékek csak a finneknél (10%), ujjuroknál és tádzsikoknál (6%), pamíri népeknél (4-5%), szlovéneknek, délolaszoknál és grúzoknál (5%) fordulnak elő. Kisebb mértékben, mintegy 1-4% közötti gyakorisággal egész Európában és a Közel-Keleten is megtalálható, ahogy a korai európai földművelő csontleletek között is.

N8, N10, N11: Kínában és Délkelet-Ázsiában előforduló, az utolsó jégkorszak után keletkezett haplocsoportok.

N12 (=O), N13, N14: az ausztrál őslakók között gyakori csoportok (ott 22%-ban vannak jelen).

N21: a malájokra jellemző csoport, előfordulása máshol elhanyagolható.

N22: Délkelet-Ázsiában és a japánoknál fordul elő.

N9 alcsoport: 46 ezer évesre becsült távol-keleti csoport, amelynek három jelentősebb alcsoportja van. A kb. 18 ezer éve keletkezett **N9a** Kelet-Ázsiában (kínaiak) és Közép-Ázsiában fordul elő; a mintegy 15 ezer éves **N9b** pedig az Amur torkolatvidékén és Japánban. Harmadik alcsoportja az **Y** nevet kapta.

Az **Y haplocsoport:** becslések szerint 25 ezer éves, két alcsoportja közül a 15 ezer éves **Y1** alcsoport az Ohotszki-tenger környékén, az Amur torkolatánál gyakori (nivh 66%, ulcsi 38%, negidál 21%, korják, evenki, udege 8-10%). Koreában, Kínában és a mongol népek között 1-3% körüli előfordulású. Az **Y2** alcsoport 7000 éves és a délkelet-ázsiai ausztronéz népek között 1-4%-ban fordul elő. Előfordul a székelyeknél is, nagyon kis számban.

Az **S haplocsoport:** kb. 54 ezer éve keletkezett, az ausztrál őslakóknál 34%-ban, az Amur-vidéken élő udegeknél 15%-ban, az ulcsiknál 5%-ban fordul elő, más kelet-ázsiai népességből nem ismert.

Az **A haplocsoport:** 24 ezer évre becsült, északkelet-ázsiai csoport. 15 ezer éves **A2** alcsoportja csak az amerikai őslakók és a Csukcs-félszigeti népek (25-47%) között fordul elő. Az eszkimók között 80-96% közötti gyakoriságú, az észak-amerikai indián népek között magasabb, a dél-amerikai indiánok között alacsonyabb, a szóródást tekintve 17-70% között változó arányban előforduló haplocsoport. A nem **A2**-höz tartozó ázsiai alcsoportok összesen 13-15%-ot tesznek ki a tibetiek, a mongolok és a hsziungnu (ázsiai hun) csontleletek között. 6-10%-os gyakoriságú a japánok, koreaiak, kínaiak és jenyiszeji ketek között. Az altáji nyelvű népek körében aránya jellemzően 3-6% közötti. A Közel-Keleten és Kelet-Európában ritkán, 3% alatt fordul elő, a székelyeknél és csángóknál is kis számban található meg.

Az **X haplocsoport:** 32 ezer évesre becsült, ritka, de sokfelé elterjedt, vitatott eredetű csoport. Legmagasabb arányban **X1** és **X2** alcsoportja a közel-keleti drúzok és grúzok között található meg, az **X1** jellemzően a közel-keleti és észak-afrikai régióban van jelen. Az **X2** egy 19 ezer évesre becsült alcsoportjának különlegessége, hogy Európában és Észak-Amerika keleti részén van jelen, de Szibériában alig fordul elő. Egyesek szerint elképzelhető, hogy az **X2** női vonal hordozói a jégkorszakban valahogy eljutottak Európából Észak-Amerikába (az elmélet egyelőre nem tekinthető bizonyítottnak). Európában jellemzően minden népben 1-6%-ban előfordul, beleértve a magyarokat, székelyeket, csángókat is.

Az **R haplocsoport:** az N csoport nagy, tipikusan nyugat-eurázsiai, mintegy 57 ezer éve keletkezett alcsoportja, a szakirodalom szerint valószínűleg Indiában jött létre. Nagyon sok, csupán számokkal jelölt alcsoportja van, míg betűkkel jelölt alcsoportjai között a legtöbb a

tipikusan nyugat-eurázsiai (kaukázusi vagy „fehér”) embertípust fed le. A magyarok több, mint 90%-a az R csoport valamely alcsoportjába tartozik.

R5, R6, R7, R8, R30, R31: tipikusan dél-ázsiai (indiai) alcsoportok. Ezek az indiai női vonalak 16-20%-át fedik le.

R3: Örményországban megtalálható alcsoport.

R12, R14, R21, R22, R23: a délkelet-ázsiai szigetvilágban előforduló alcsoportok.

R9 csoport: Délkelet-Ázsia szigetvilágában fordul elő, külön betűs alcsoportja az **F**.

F haplocsoport: 43 ezer éves csoport, Kelet-Ázsiára jellemző, szinte minden népességben 10-26% közötti előfordulással. Közép-Ázsiában az altáji migrációt jelzi 1-8% közötti előfordulással (egyes kisebb altáji türk népeknél 25-40% is lehet). A magyarok között nincs jelen, de Hvar szigetén Horvátországban megtalálták (amit a szakirodalomban sokan ún. avar kapcsolatnak tartanak).

R11 csoport: Délnyugat-Kínában, Indiában és Japánban fordul elő. Közeli kapcsolat fűzi a **B** csoporthoz.

A **B haplocsoport:** 50 ezer éve alakult ki, elsősorban az ausztronéz népekre (14-44%), az amerikai őslakókra (22-28%), a vietnamiakra (28-31%) és a laosziakra (25%) jellemző. A kínaiak, koreaiak és japánok között is 10-20% közötti. Közép-Ázsiában és Dél-Szibériában 1-15% közötti előfordulású, a mai magyarok és a honfoglaló csontok között is jelen van (1%).

A **P haplocsoport:** 55 ezer évesre becsült, domináns Pápua Új-Guineában (56%) és a fülöp-szigeteki negritó (aeta) lakosságban (40%), de előfordul az ausztrál őslakóknál (15-16%) is.

Az **U haplocsoport:** a mintegy 47 ezer éve alakult ki, a szakirodalom szerint valószínűleg Indiából terjedt el széles körben. A magyarok, székelyek és csángók között az U (K nélkül) összesen 15-18%-os előfordulású. Alcsoportjai közül a 30 ezer éves, **U5** Európa kőkorszaki halász-vadász lakosságára jellemző (a lappoknál 46%), az **U4** pedig Skandináviától Közép-Szibériáig terjedt el a kőkorszakban. Az **U4** és **U5** Északkelet-Európa uráli népei, valamint a csuvasok és baskírok között jellemzően 24-36%-os előfordulású, de Európa többi részén és Közép-Ázsiában is meghaladja a 10%-ot.

Az **U1** csoport Grúziától Indiáig terjedt el. Az **U2** és **U9** csoportok Dél-Ázsiában fordulnak elő. Az **U6**-os alcsoport Nyugat-Afrikában a berberekre (30%) és a Kanári-szigetek őslakóira (guancsok) jellemző. Az **U7** Indiától Nyugat-Európáig és a Közel-Kelettől Szibériáig alacsony gyakorisággal található meg. Az **U3**-as és **U8**-as csoport a Közel-Keleten és Európában fordul elő.

A **K haplocsoport:** az **U8** alcsoport külön betűt kapott, mintegy 27 ezer éve kialakult alcsoportja. A zsidó diaszpórában és a korai európai földművelő népesség között 15-38%

közötti előfordulását, Cipruson és a csángók között pedig 23%. A közel-keleti Termékeny Félhold mai lakóinál aránya 9-12%, de 10% alatt egész Európában és Közép-Ázsiában is megtalálható. A magyarok és székelyek között ennél kisebb, 2-5% az előfordulása.

Az **R2 haplocsoport**: Törökországtól Indiáig fordul elő, legközelebbi csoportja a 47 ezer évesre becsült **JT** csoport, amely a **J** és **T** csoportok közös őse.

A **J haplocsoport**: A 34 ezer éve kialakult **J** csoport maximumát Irakban és Szaúd-Arábiában mérték (21%), Közép-Ázsiában, Európában és Észak-Afrikában átlag 5-15% között fordul elő. A magyarok, székelyek és csángók között 8-13% közötti gyakoriságú.

A **T haplocsoport**: a 25 ezer éve kialakult **T** csoport legnagyobb gyakorisága a korai földművelő és bronzkori kelet-európai csontleletek között mérhető (14-28%), amelyhez hasonló mértékben ma csak az udmurtok (24%), ukránok és macedónok (16%) között, valamint a Dél-Kaukázusban (14-15%) fordul elő. Előfordulása 4-13% között mérhető minden más helyen, ahol az europid (kaukázusi) típus elterjedt. Gyakorisága a magyarok és csángók között kisebb, 2-4%-os, míg a székelyeknél 9-10%-ban mérhető.

A **R0 (pre-HV)** csoport: főként Délnyugat-Ázsiában, az Arab-félszigeten fordul elő.

A **HV haplocsoport**: mintegy 22 ezer évesre becsült csoport, dominanciája a kaukázusi típusúnak tekintett népesség által lakott területre jellemző. **HV1** alcsoportja a Közel-Keleten, a **HV2** Dél-Ázsiában, a **HV3** Kelet-Európában van jelen. A **HV*** (nem **H** és **V**) csoport gyakorisága a magyar népességben 4-10% közötti. Két külön betűt kapott alcsoportja a **H** és a **V** (eredetileg **HV0**).

A **H haplocsoport**: 13-15 ezer éve Dél Franciaországban keletkezett csoport, amely domináns Európában. Gyakorisága a lappokat kivéve (elenyésző) minden európai népnél 32-63% közötti: a magyaroknál 47%, a székelyeknél 33%, a csángóknál 23%. Közép-Ázsiában, Észak-Afrikában és a Közel-Keleten 10-30% közötti, de még Dél-Szibériában és Belső-Ázsiában is 2-8% közötti.

A **V haplocsoport**: kb. 10 ezer éve alakult ki, frekvenciája a lappoknál a 40%-ot is eléri. A marik és finnek között, valamint a baszkoknál és tuaregeknél is 10-12% közötti gyakoriságú, Európa többi népénél jellemzően 1-7%-ban található meg, a magyarok között ritkán, 1% körül.

4. ANYAG ÉS MÓDSZER

4.1. Minták és populációk

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

Kereskedelembe kapható különböző STR kitek által tartalmazott lokusz számainak növekedésével és/vagy új lokuszok bevonásával összhangban vizsgáltunk különböző számú személyből álló magyar populációs mintákat.

Az autoszómális STR vizsgálatok tekintetében összesen: 26395 magyar minta (4213, 254, 455 és 21473), 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 három generációs családból **155** személy vizsgálata történt meg.

Az Y-kromoszómális STR és SNP lokuszokra vonatkozó felmérést különböző populációkban végeztük el: Magyarország különböző régióitól származó 341 magyar (YHRD accession number: YA003185, YA004098), 95 székely (csak SNP, YHRD accession number: YA002983), 95 csángó (csak SNP, YHRD accession number: YA002984), 145 bodrogközi magyar (YHRD accession number: YC00032), 137 magyarországi roma (YHRD accession number: YA003188), 29 tiszavasvári oláh roma (YHRD accession number: YA003658), 39 tokaji oláh roma (YHRD accession number: YA003660), 19 taktaközi romungro roma (YHRD accession number: YA003659), 62 szlovák roma (YHRD accession number: YA003186), 53 kazahsztáni madjar (YHRD accession number: YA003998), 96 manyisi (YHRD accession number: YA004169, YA004170), 115 üzbég (YHRD accession number: YA003994), 61 üzbegisztáni madjar (YHRD accession number: YA003995), 127 mongol (YHRD accession number: YA003997), 88 burját (YHRD accession number: YA003996), 301 malajziai indiai (csak SNP, YHRD accession number: YA003277), 241 thaiföldi YHRD (accession number: YC000306). Összeségében **2045** mintán vizsgáltunk. Ezen felül nemzetközi együttműködés keretében részt vettünk gyorsan mutálódó Y-STR lokuszok (Rapidly Mutating Y-Chromosomal Short Tandem Repeats) vizsgálatában 103 személy 13 RM lokuszon, valamint 143 magyar és 101 roma minta 23 Y-STR lokuszon történő tesztelésével (összesen **347** személy).

4.2. A DNS izolálása és koncentrációjának meghatározása

A friss, EDTA-val alvadásgátolt vérmintákból a DNS izolálása a gyártó cég utasításainak megfelelően QIAamp DNA Blood Mini Kittel történt.

A nyálmintavevő pálcával (Whatman OmniSwab) levett szájnyálkahártya törletekből Chelex módszerrel [3], valamint a DNAIQ (Promega) módszerrel nyertük ki a DNS-t.

A szűrőpapírra, illetve FTA papírra csepegtetett és beszárított vérmintákból szerves extrakciós módszerrel [97] és ultraszűrés-koncentrációval (Microcon-100, Amicon, Millipore) történt a DNS izolálás. A mintában jelen lévő humán DNS (minőséget is viszonylag pontosan tükröző) mennyiségi meghatározása real-time PCR technikával történt Quantifiler Human kit, ABI 7500 Real-Time PCR System készülék és Sequence Detection Software 1.2.3. (Life Technologies) alkalmazásával a gyártó cég utasításainak megfelelően történt [98].

A szájnyálkahártya törlét vételhez a fent említett OmniSwabon kívül 2 másik mintavevőt (Bode DNA Collector és Whatman EasiCollect) is használtunk, amelyek autoszómális és Y-STR vizsgálat esetén direkt PCR-re alkalmasak (DNS izolálás nélküli PCR technológia), az X-STR és Y-SNP esetében DNS izolálás a fent leírt módon történt.

4.3. Genotipizálás fragmenshossz analízissel

A felsokszorozott DNS-fragmentumok méret szerinti elválasztása denaturáló kapilláris gélelektroforézissel ABI Prism 310 és ABI 3130 Genetic Analyzer (Life Technologies) készülékeken a felhasznált kitek használati utasításának megfelelően történt (Biotype AG, Germany; Promega, USA). Az adatok kiértékeléséhez, és a minták genotipizálásához az ABI Prism 310 Genetic Analyzer készüléken GeneScan Analysis v3.1. és Genotyper v2.6.2 szoftvereket (Life Technologies), az ABI 3130 készüléken Genemapper ID v3.2 (Life Technologies) szoftvert használtunk.

4.4. Az Y-kromoszómális SNP lokuszok PCR-amplifikációja és genotipizálása TaqMan próbával (haplocsoport meghatározás)

A vizsgált Y-SNP markerek és a hozzájuk tartozó haplocsoportok – CT-M168; C-M216; C, FT (xD,E)-P143; C1-M105; C2-M38; C3-M217; DE-YAP; D-M174; E-SRY4064; E1b1b1-M35; E1b1b1a-M78; E1b1b1b-M81; E1b1b1c-M123; E1b1a-M2; FT-M89; G-M201; G1-M285; G2-P15 (L156); H1-M52; H1a-M82; I-M170; I1-M253; I2a1-P37.2; I2a2-M223 (I2b); J-M304; J1-M267; J2-M172; J2a2-M67; J2b-M12 (M102); KT-M9; L-M11; M1-M106; N-M231; N1c-Tat (M46); N1c-L708, N1c-L1034, N1c-VL29, N1c-Z1936, N1c-

F4205; N1b-P43, O-M175; O3-M122; Q-M242; PT-M45; R-M207; R1-M173; R1a-M198; R1a-M458; R1a-Z280; R1a-Z93; R1b-P25 (M343); R1b-U106; R1b-M412; R1b-P312; R2-M124 és T-M170 [38, 99, 100, 101, 102, 103, 104, 105, 106, 107]. A nevezéktant az ajánlásoknak megfelelően alkalmaztuk [31, 38].

Az Y-haplocsoportokat 2005-ben kezdtük vizsgálni - abban az időben az Y-SNP vizsgálatok nemzetközi szinten is kezdeti stádiumban voltak - ezért a 2002-ben publikált Y-SNP filogenetikai fa minden fő haplocsoportjához választottunk markert, kivéve az **A** és **B** haplocsoportokat (ezek csak Afrikában fordulnak elő).

A fő haplocsoportok magyar populációban történt meghatározása után a marker választás olyan szempontok szerint történt, hogy a nagy gyakorisággal előforduló fő haplocsoportokon belül vizsgálható legyen azoknak az alhaplocsoportoknak a jelenléte, melyeknek feltételezhető az előfordulása a magyar populációban: E1b1b1-M35; E1b1b1a-M78; E1b1b1b-M81; E1b1b1c-M123, G2-P15, I1-M253; I2a-P37.2; I2b-M223, J2-M172; J2a2-M67; J2b-M12 (M102), R1-M173; R1a1-M198; R1b1-P25. A H1a-M82 alhaplocsoport előfordulása nem jellemző Európában, ám a roma populációkban igen, ezért ezt a markert is vizsgáltuk. A jellegzetesen Ázsiában előforduló fő haplocsoportok alhaplocsoportjait is vizsgáltuk, mivel előfordulásuk lehetősége fennállt a malajziai indiai populációban: C1-M105; C2-M38; C3-M217; C4a-M210, M2a-SRY9138, O1a-M119; O3-M122, R2-M124.

A genotipizáláshoz TaqMan próbákat (Applied Biosystems) használtuk, 3-5 ng DNS-templát, ABI 7500 Real-Time PCR System és GeneAmp 9700 PCR System készülékek (Applied Biosystems) alkalmazásával. A PCR termékek relatív fluoreszcencia detektálását ABI 7500 Real-Time PCR System készülékkel végeztük, a genotipizáláshoz a Sequence Detection software v1.2.3-at alkalmaztuk a gyártó utasításainak megfelelően [108].

4.5. Populációstatisztikai analízisek az autoszómális és X-STR lokuszokon

A magyar populáció genetikai struktúrájának felméréséhez különböző számú (8-16) STR-lokuszra vonatkozó allélgyakoriság, PIC, PD, PM, HET, $MEC^{Krüger}$ és $MEC^{Kishida}$ értékeket külön-külön határoztuk meg.

Allélgyakoriság

a lokuszokhoz tartozó allélok megfigyelt gyakorisága egy populációban.

A PIC (Polymorphism Information Content)

érték azt fejezi ki, hogy egy adott lokusz az allélok mérettartományának és gyakoriságának ismeretében mennyire informatív.

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

ahol p_i és p_j az i -edik és j -edik allél megfigyelt gyakorisága [109].

A HET^{expected} (Heterozygosis)

érték kifejezi, hogy a vizsgált egyéneknek várhatóan mekkora hányada lesz heterozigóta a vizsgált lokuszra nézve.

$$HET^{\text{expected}} = 1 - \sum_{i=1}^n p_i^2$$

ahol p_i az i -edik allél megfigyelt gyakorisága [10].

A PD (Power of Discrimination)

érték kifejezi, hogy a populációból véletlenszerűen kiválasztott két személyt vizsgálva, mennyi annak az átlagos valószínűsége, hogy a két személy a vizsgált lokuszon eltérő genotípust hordoz.

$$PD = 1 - \sum_{i=1}^n x_i^2$$

ahol x_i az i -edik genotípus vagy haplotípus megfigyelt gyakorisága [110].

Több lokuszra kombinálva

$$PD_{\text{komb}} = 1 - \prod_{i=1}^n (1 - PD_i)$$

A számítást az X-kromoszómális vizsgálatok esetében a női és férfi mintákon külön kell végezni, mivel a lokuszok öröklődése a két esetben különböző (PD^{male} , PD^{female}).

A PM (Probability of Match)

érték kifejezi, hogy a populációból véletlenszerűen kiválasztott két személyt vizsgálva, mennyi annak az átlagos valószínűsége, hogy a két személy a vizsgált lokuszon megegyező genotípust hordoz. A számítást X-kromoszómális lokuszok esetében a női és férfi mintákon külön kell végezni, mivel a lokuszok öröklődése a két nem esetében különböző (PM^{male} , PM^{female}).

$$PM = 1 - PD$$

Több lokuszra kombinálva

$$PM_{\text{komb}} = \prod_{i=1}^n PM_i.$$

A MEC^{Krüger} (Mean Exclusion Chance)

érték hiányos apasági vizsgálatok esetén kifejezi, ha az anyát, annak lány gyermekét és a vélelmezett apa anyját vizsgáljuk X-kromoszómális lokuszokra nézve, mennyi annak a valószínűsége, hogy a populációból véletlenszerűen kiválasztott férfi, mint vélelmezett apa kizárt. Ugyan ezt képletet használjuk az autoszómális STR esetében is anyára, gyermekekre és vélelmezett apára vonatkozóan.

$$MEC^{Krüger} = \sum_{i=1}^n p_i \cdot (1 - p_i)^2 + \sum_{i=1}^{n-1} \sum_{j=i+1}^n p_i \cdot p_j \cdot (3p_i + 3p_j - 4)$$

ahol p_i és p_j az i -edik és j -edik allél megfigyelt gyakorisága [111].

A MEC^{Kishida} (Mean Exclusion Chance)

érték nem hiányos apasági vizsgálatok (X-kromoszóma) esetén kifejezi, ha az anyát, lány gyermekét és a vélelmezett apát vizsgáljuk, mennyi annak a valószínűsége, hogy a populációból véletlenszerűen kiválasztott férfi, mint vélelmezett apa kizárt.

$$MEC^{Kishida} = \sum_{i=1}^n p_i^3 \cdot (1 - p_i) + \sum_{i=1}^n p_i \cdot (1 - p_i)^2 + \sum_{i=1}^{n-1} \sum_{j=i+1}^n p_i \cdot p_j \cdot (p_i + p_j)$$

ahol p_i és p_j az i -edik és j -edik allél megfigyelt gyakorisága [112].

Az egyes lokuszokra kiszámított MEC értékeket az alábbi képlet szerint **kombinálhatjuk** az összes vizsgált lokuszra:

$$MEC_{\text{komb}} = 1 - \prod_{i=1}^n (1 - MEC_i)$$

A HWE P érték (Hardy-Weinberg egyensúly)

megmutatja, hogy a vizsgált lokusz Hardy-Weinberg egyensúlyban van-e, autoszómális és X-kromoszómális lokuszok esetén csak női mintákra értelmezhető.

A Hardy Weinberg egyensúly tesztelésére az Arlequin 2.001 szoftvert használjuk [113] a következő beállításokkal: No. of steps in Markov chain = 100000, No. of Dememorisation Steps = 10000.

4.6. Struktúra analízis

Mivel a vizsgált minták a bűnügyi DNS adatbázis mintáiból álltak (21473 minta), ezért kíváncsiak voltunk a szubpopulációs struktúrára, amelyet STRUCTURE 2.3.4 szoftver [114] segítségével elemeztünk. A szoftverrel a legnagyobb számú lehetséges genetikai csoportot becsültük (clusters, K) az összes vizsgált személy tekintetében. K értékeket generáltunk 1-től 100-ig, 10000 (burn-in) szakaszt 10000 MCMC iteráció követett és 10 futtatást végeztünk $K \leq 35$ tekintetében. $K=35$ után azonos számú beégést és iterációt használtunk, de 3 futást alkalmaztunk $K=100$ -ig az eredmény fájlok nagy mérete miatt. Admixture modellt használtunk korrelált allél frekvenciával. A Bayes statisztika eredményét STRUCTURE HARVESTER [115] programmal elemeztük, amelybe az Evanno módszer (ΔK) [116] van beépítve.

4.7. G teszt

Annak vizsgálatára, hogy a külföldi és más hazai populációs adatok és a vizsgált magyar allélgyakorisági adatok között van-e szignifikáns eltérés, G tesztet alkalmaztunk.

4.8. Populációstatisztikai és genetikai analízisek az Y-STR és Y-SNP lokuszokon

A genetikai struktúra felméréséhez kiszámítottuk az Y kromoszóma nem rekombinálandó régiójában elhelyezkedő kezdetben 12, majd később 23 Y-STR és több mint 50 Y-SNP lokuszra vonatkozó haplotípus- és haplocsoport gyakoriságokat, illetve haplotípus- és haplocsoport diverzitást (Nei-féle géndiverzitás). A vizsgált populációk és a referált irodalmi populációs adatok összehasonlításához a populáció párokra vonatkozóan a haplotípus, illetve haplocsoport gyakoriságok alapján F-statisztikát és AMOVA-t is végeztünk. Az eredmények összefüggéseit MDS (Multidimensional Scaling) módszerrel elemeztük [117]. A földrajzi és genetikai távolságok összefüggésének ábrázolására a vizsgált populációk vonatkozásában filogeográfiai analízis történt [118]. Az egyes haplocsoportokhoz tartozó haplotípusok leszármazási viszonyainak ábrázolását és az esetlegesen egy közös ősről visszavezethető haplotípusok létrejöttéhez szükséges divergálódási idő becslését Network programmal [66] kiviteleztük.

Haplotípus gyakoriság

a 12-23 Y-STR lokuszhoz tartozó haplotípusok megfigyelt gyakorisága egy populációban.

Haplocsoport gyakoriság

az Y-SNP lokuszhoz tartozó haplocsoportok megfigyelt gyakorisága egy populációban.

A Haplotípus diverzitás (Haplotype Diversity) és haplocsoport diverzitás (Haplogroup Diversity)

kifejezi a vizsgált populációk géndiverzitását

$$h = \frac{n}{n-1} \cdot (1 - \sum x_i^2)$$

ahol n a kromoszómák száma és x_i az i -edik haplotípus vagy haplocsoport megfigyelt gyakorisága [119].

4.9. F-statisztika és AMOVA

A genetikai profilok interpopulációs variabilitásának becsléséhez hagyományos Wright-féle F-statisztikát és molekuláris varianciaanalízist (AMOVA) használtunk [120, 121, 122, 123]. A populáció párokra vonatkozó Fst értékeket és azok szignifikanciáját Arlequin 2.0 szoftverrel határoztuk meg.

Az AMOVA a genotípusok (haplotípusok) közötti molekuláris távolságon alapszik, amely egyenlő az egyes genotípusokat egymástól elválasztó mutációk minimális számával. Az elemzés eredményeképpen kapott Φ_{ST} érték a Wright-féle Fst értékekhez hasonlóan azt fejezi ki, hogy a populációban meglévő teljes molekuláris variancia mekkora része származik a populációk közötti varianciából. A távolságszámításnál az STR lokuszok lépésenkénti (azaz stepwise) mutációs modelljét állítottuk be [124]. A távolságmátrix létrehozásakor az STR allélok repeat-számbeli különbségeinek négyzetét alkalmaztuk [123]. A Φ_{ST} érték nagysága a populációk közötti molekuláris távolsággal arányos, statisztikai szignifikanciáját szimulációval teszteltük. A szimulációt végző szoftver az adott populáció párban előforduló genotípusokat (haplotípusokat) véletlenszerűen újra szétosztja a két populáció között. A szoftver beállításánál 100000 replikációt alkalmaztunk. Az aktuális Φ_{ST} értéket meghaladó szimulációs Φ_{ST} értékek gyakorisága a szimuláció P értéke. Ebben az esetben a $\{\Phi_{ST}=0\}$ hipotézist teszteltük a $\{\Phi_{ST}>0\}$ hipotézis ellenében. Ha a kapott $P<0,05$, akkor 95%-os szignifikancia-szint mellett a nullhipotézis elvethető, azaz a két vizsgált populáció között szignifikáns különbség van.

Mivel az AMOVA nem képes duplikált lokuszok kezelésére, ezért a DYS385 lokuszt nem elemeztük. A DYS39II lokusz tartalmazza a DYS389I lokusz ismétlődéseit is, ezért az

AMOVA-hoz a DYS389II lokuszon detektált repeat egységek számából kivontuk a DYS389I lokusz ismétlődéseinek számát.

4.10. Filogenetikai fa szerkesztése

A filogenetikai fa szerkesztését a genetikai távolságokból Median Joining módszerrel végeztük, a Phylip 3.6 programcsomag használatával [125].

4.10.1. MDS (Multidimensional Scaling)

A többdimenziós skálázás olyan módszer, melynek kiindulási adatait az objektumok „közelsége” adja. Az elemzés célja a megfigyelt egyedek közti hasonlóságot/különbözőséget visszaadó, összevont dimenziók (több megfigyelési szempontot egyszerre visszatükröző, egyszerűsített változók) megtalálása, hasonlósággént/távolsággént bármilyen jellegű hasonlóság/távolságmátrix használható. Az MDS az objektumok hatékony „újra elrendezésének” módja a megadott dimenziók számának megfelelő térben, végeredménye az objektumok olyan elrendezése, ami leginkább megfelel a köztük megfigyelt hasonlóságoknak/távolságoknak. Az MDS eredménye általában egy térbeli vagy síkbeli reprezentáció, ami pontok geometriai konfigurációját térkép-szerűen ábrázolja. A konfiguráció minden egyes pontja egy objektumnak felel meg, így a konfiguráció az adatok „rejtett struktúráját” tükrözi, és gyakran megkönnyíti az adatok megértését. Az adat-struktúra leképezésekor minél nagyobb a különbség (vagy kisebb a hasonlóság) két objektum között, a térbeli reprezentáción annál távolabb helyezkednek el [117]. A genetikai távolságok MDS formában történő ábrázolását ViSta 7.9.2.4 szoftverrel [126] készítettük.

4.10.2. Filogeográfiai analízis (Phylogeographic Analysis)

A populációk közötti hasonlóság származhat a közös eredetből, vagy kis földrajzi távolság esetén az utóbbi időkben történt keveredésükből. Általánosságban elmondható, hogy a földrajzi távolság miatti izoláció következtében az egyes népcsoportok közötti genetikai távolság arányos lehet a köztük lévő földrajzi távolsággal, ami látható, ha a populációk közötti földrajzi távolságok függvényében történik a genetikai távolságok ábrázolása [118].

4.10.3. Hálózat analízis és divergálási idő becslés

A Median Joining Network analízist [66] a Network 4.2. v. számítógépes program felhasználásával végeztük el. A Median Joining hálózatokat mindig egy adott haplocsoportba tartozó haplotípusokra vonatkozóan készítettük el. Az analízis előtt a DYS389II ismétlődési számából levontuk a DYS389I lokusz ismétlődési számát. Mivel a program nem tud duplikálódott lokuszt kezelni, ezért a DYS835 lokuszt kihagytuk az elemzésből. A szoftver a betáplált Y-STR haplotípusokból elkészíti az összes lehetséges legrövidebb és legegyszerűbb filogenetikai fát (az összes maximum parszimónia fát), azaz a Median Joining hálózatot. A fa csomópontjai az egyes haplotípusok, a csomópontok mérete arányos az azoknak megfelelő haplotípusok előfordulási gyakoriságával, míg az egyes csomópontokat összekötő vonalak (ágak) hossza a haplotípusok között lévő mutációs lépések számával arányos. A mutációs ráta ismeretében szoftverrel lehetséges megbecsülni a hálózaton belül ősinek kiválasztott csomópontból a belőle leszármazott csomópontok létrejöttéhez szükséges időt. A divergálódási idő becsléséhez az evolúciós mutációs rátát ($6,9 \times 10^{-4}$ mutáció/lokusz/25 év) [32], valamint a pedigre mutációs rátát ($2,5 \times 10^{-3}$ /lokusz/30 év) [58] is figyelembe vettük.

4.10.4. Bayes Skyline Plots (BSP) analízis

A BSP-t a BEAST v1.8.0 program alkalmazásával generáltuk [127]. Az MCMC minták 30 000 000 generáción alapultak, 1000 lépésenként naplózva, az első 3 000 000 generációt, elvettük, mint burn-int. Az elemzést a következő beállításokkal végeztük: tíz csoport, HKY szubsztitúciós modell, a mutáció sebességéhez egy korrelálatlan relaxed óra lognormális eloszlással ($2,21 \pm 0,17 \times 10^{-8}$ mutáció / nukleotid / év) [128] logaritmikus eloszlással és 30 éves generációs idővel [129, 130]. Ehhez az elemzéshez figyelembe vettük az egész mitokondriális genomot: az időhatárt figyelembe véve (<40 KYA) a sebesség és utóbbi standard hibájának számítását a Soares és mtsai., 2009 cikkben [128] feltüntetésre kerültek szerint végeztük, hogy meghatározzuk az arányt az egyes érintett csomópontokon. Ezen értékek átlagát használtuk fel a mtDNS és MSY haplocsoportok (NGS eredmények) idejének (TMRCA) becsléséhez [85, 86].

5. AZ EREDMÉNYEK ÉS AZOK MEGVITATÁSA

5.1. AZ IGAZSÁGÜGYI GENETIKA

Az elmúlt 20 évben az igazságügyi genetikai területen végzett vizsgálataink közül csak a legjelentősebbeket szeretném bemutatni, inkább az eredményeket igazságügyi szakértői gyakorlatban való alkalmazhatóságot prezentáló példákkal illusztrálva.

5.1.1. Autoszómális STR markerek vizsgálata

Az igazságügyi genetikai szakértéshez az adott populáció, jelen esetben a magyar népesség genetikai összetételének vizsgálata feltétlenül szükséges. A tudomány és a technológia fejlődésével a multiplex PCR módszer segítségével egyszerre, szimultán vizsgálható STR lokuszok száma évről évre növekedett, ami maga után vonta azt a kényszert, hogy újabb és újabb populációgenetikai vizsgálatok nélkül azok alkalmazása a rutin szakértői gyakorlatban lehetetlen. Természetesen ezen új lokuszok vizsgálatba bevonásával az ún. *DNS-profil egyezési valószínűségének* mértéke csökken (egyezési valószínűség kicsi), azaz egy-egy személy DNS-profiljának egyedisége inkább biztosított. Másrészt az eredmények összehasonlíthatók lesznek más munkacsoportok/laboratóriumok eredményeivel, ami nemcsak tudományos, hanem szakértői szempontból is rendkívül fontos, mivel a laboratóriumok közötti kompatibilitás a szakértői munka hitelének egyik zálogát jelenti.

Az elmúlt 4-5 évben az EU-n belül jelentős beruházások történtek annak érdekében, hogy a határokon túli bűnözést és a terrorizmust megakadályozandó, erősítsék a nemzetközi együttműködést és a genetikai információk cseréjét. Erőfeszítések történtek a prűmi szerződésben foglaltaknak az európai törvényekbe való átültetése érdekében. Amíg nem történik meg a szükséges technológiai harmonizáció [131], különbségek vannak az EU-tagországok bűnügyi DNS adatbázisainak befogadási és tárolási követelményeinek szabványosításában, addig ártatlan emberek és gyermekek DNS-profiljai nem lehetnek megfelelő és egyenlő mértékű biztonságban, továbbá nem garantálható a prűmi szerződés azonos szigorúságú végrehajtása [132, 133].

Már a Bevezetőben említésre került, hogy szinte minden országban létezik nemzeti bűnügyi DNS adatbázis, nemzetközi szinten pedig az amerikai CODIS (FBI) és az európai (prűmi) DNS adatcsere hálózat, több százezer/millió DNS-profil tárol. Az adatcsere támogatása érdekében az ún. *DNA Gateway* kifejlesztésre került sor, ami nemzetközileg

elismert standardként szolgálja az elektronikus DNS adatátvitelt az INTERPOL és annak tagországai között. A tagországi rendőrségek beküldhetik az elkövetők, helyszíni minták, eltűnt személyek és azonosítatlan holttestek DNS-profiljait az INTERPOL automatizált DNS adatbázisába. A résztvevő országok aktívan használják a DNA Gatewayt, mint eszközt, a bűnügyi nyomozásban, és rendszeresen figyelik a potenciális kapcsolatokat a DNS-profilokat beküldő országok között. A Nemzetközi DNS adatbázisban állandóan gyarapszik a tárolt DNS- profilok mennyisége, valamint bővül a használandó STR lokuszok listája, ami növeli a nemzeti igazságügyi DNS laboratóriumok vizsgálati palettáját is. Ehhez kapcsolódva vizsgáltuk az STR lokuszok allélgyakorisági adatait **26395** magyar mintán igazságügyi aspektusból. A minták közül 4922 mintát véletlenszerűen válogattunk, 21473 minta pedig a fogvatartottaktól és gyanúsítottaktól származott. Utóbbiak eredményei a nemzeti bűnügyi DNS adatbázisba kerültek, és a potenciális bűnelkövetők csoportját képezik. A magyar népességre vonatkozóan korábban nem állt rendelkezésre a potenciális bűnelkövetők STR allél- és profilgyakorisági adatbázisa. Eredményeink azzal is kecsegtettek, hogy összehasonlíthatók lesznek más munkacsoportok eredményeivel. Ez nemcsak tudományos, hanem szakértői szempontból is rendkívül fontos, mivel az igazságügyi genetikai laboratóriumok közötti kompatibilitás a szakértői munka hitelének egyik zálogát jelenti.

Az autoszómális STR alapú vizsgálatok alapvető célja a humán STR polimorfizmusok genetikai analízisének bevezetése a magyarországi igazságügyi genetikai gyakorlatba, hitelesítése, valamint olyan magyar populációs referencia adatbázis létrehozása, amelynek segítségével az STR lokuszok vizsgálatán alapuló DNS-eredmények bizonyító ereje populációstatisztikai szinten becsülhető a magyarországi törvényszéki eljárások során.

A magyarországi publikált adatok közül a kb. 10 milliós lakosság 0,26%-át képező 21473 mintára kiterjedő adatbázis jelentőségét és vizsgálati eredményeit szeretném bemutatni [134].

A törvényszéki eljárásokban a DNS bizonyíték biostatisztikai szempontú interpretációja szolgálat bizonyító erőt a jogalkalmazók számára. A bizonyítékok együttes mérlegelése során kritikus tényező lehet egy biológiai anyagmaradvány megbízható, nagy „statisztikai erővel” bíró származtatása. A biostatisztikai számítások elvégzését ma már nemcsak szakmai, hanem jogszabályi előírások is szabályozzák, kötelező elemei a szakértői véleménynek. A biostatisztikai számítások Bayes elvű megközelítése szükségessé tette a hazai populációt minél jobban reprezentáló referencia adatbázis létrehozását. A hazai és a nemzetközi gyakorlatban rutinszerűen használt 15 STR lokusz (és szex meghatározás) magyarországi alléleloszlását korábban – a D19S433, D1S1656 és a D12S391 lokuszok kivételével – 4213

nem rokon személy vizsgálata alapján közöltük. A 21473 új és 4213 korábbi mintából álló adatbázisok populációstatisztikai paramétereit a **2. táblázat** tartalmazza.

A D19S433, D1S1656 és a D12S391 lokuszok kivételével az összes PIC érték magasabbnak mutatkozott a 21473 minta tekintetében, mint a korábban publikált 4213 minta esetében. A kivételt képező három lokusz esetében a csökkenés a kisebb vizsgálati populációnak volt tulajdonítható.

A populáción belüli heterozigócia csökkenés mértékével (F_{IS}) korrigálva a profilegyezési valószínűségek (pM) a korábban publikált adatokhoz képest 5 lokusz esetében magasabbak lettek. Az összesített pM érték majdnem azonos volt ($pM=1.06 \times 10^{-20}$) a 21473 minta esetében, a korábban publikált 4213 minta alapján számolt értékkel ($pM=1.06 \times 10^{-19}$) [135, 136]. A 4213 személyből álló populációs minta lokuszaira eső minimum allél frekvencia (p_{min}) értékeit újra kalkuláltuk, ezekre korábban 0,011 és 0,00064 közé eső értékeket kaptunk. A minimum allél frekvenciák a 21473 minta minden egyes lokusz tekintetében kisebbek voltak 0,0001339-nél, vagyis 1000-szer kisebbek, mint a 4213 mintát tartalmazó adatbázis esetében. Ez az eredmény azt jelenti, hogy a DNS-profilok összesített elméleti minimum frekvenciája 15 STR lokusz esetén $4,1 \times 10^{-92}$ a 21473 minta alapján, míg a korábban publikált 4213 minta alapján számolt érték $2,6 \times 10^{-77}$ volt.

A ritka allélt/allélokat is tartalmazó DNS-profilok populációs gyakoriságának becslésénél tehát az új adatbázis kevésbé torzít, megbízhatóbb értéket szolgáltat a tényleges előfordulásról. Nyolcvanöt új allélt találtunk az új adatsor (21473) elemzése során. A D12S391 lokuszon 62 új allélt detektáltunk, amelyek közül a 21.3 allél 34-szer fordult elő [134]. Érdekes, hogy ezt a mikrovariáns allélt a Promega cég által forgalmazott allélétra nem tartalmazta, ezért ennek bevezetését javasoltuk.

2. táblázat. 15 STR lokusz igazságügyi hatékonysági és statisztikai paraméterei 21473 minta alapján

	HW <i>p</i> -érték ^I	HW <i>p</i> -érték ^{II}	<i>F</i> _{IS}	<i>F</i> _{IS} <i>p</i> -érték	<i>F</i> _{st}	<i>p</i> _{min} *	<i>p</i> _{min} **	PIC	<i>H</i> _E	<i>H</i> _O	PD	<i>pM</i> _{ave}	PE
D3S1358	0.0351	0.20511	0.00165	0.32454	0.0004	0.0001244	0.0006358	0.75298	0.78745	0.78615	0.921626	0.078374	0.5736
D19S433	0.0010	0.06340	0.01695	0.00000	-0.0007	0.0001235	0.0050635	0.76067	0.78932	0.77594	0.927266	0.072734	0.5552
D2S1338	0.1243	0.10552	0.01006	0.00000	0.0003	0.0001322	0.0054915	0.86325	0.87590	0.86709	0.972144	0.027856	0.7288
D22S1045	0.0002	0.00336	0.01985	0.00000	-0.0005	0.0001190	0.0100760	0.69047	0.73217	0.71763	0.88696	0.11304	0.4561
D16S539	0.0000	0.00127	0.01642	0.00000	-0.0007	0.0001230	0.0039769	0.74916	0.78230	0.76946	0.920018	0.079982	0.5436
D18S51	0.8897	0.93533	0.00351	0.08602	0.0001	0.0001326	0.0007006	0.85996	0.87334	0.87027	0.970792	0.029208	0.7352
D1S1656	0.0000	0.00473	0.01371	0.00000	0.0012	0.0001339	0.0111216	0.88435	0.89399	0.88174	0.979296	0.020704	0.7583
D10S1248	0.0544	0.25833	0.00830	0.00587	-0.0005	0.0001232	0.0105303	0.74421	0.77929	0.77282	0.916694	0.083306	0.5496
D2S441	0.0000	0.00000	0.00857	0.00978	-0.0002	0.0001229	0.0099070	0.74060	0.77482	0.76817	0.915512	0.084488	0.5414
TH01	0.0248	0.05943	0.01302	0.00000	0.0005	0.0001236	0.0040831	0.75456	0.78789	0.77763	0.922269	0.077731	0.5582
vWA	0.0015	0.00011	0.01018	0.00293	0.0008	0.0001253	0.0006442	0.77829	0.80565	0.79745	0.935295	0.064705	0.5943
D21S11	0.1360	0.51876	0.01066	0.00000	0.0008	0.0001294	0.0006661	0.83198	0.84937	0.84031	0.960157	0.039843	0.6759
D12S391	0.0089	0.11330	0.01055	0.00000	0.0009	0.0001326	0.0112776	0.86741	0.87925	0.86998	0.973763	0.026237	0.7346
D8S1179	0.3060	0.43626	0.00709	0.01760	0.0003	0.0001261	0.0006581	0.78688	0.81128	0.80552	0.940225	0.059775	0.6093
FGA	0.0003	0.04544	0.01064	0.00000	0.0003	0.0001315	0.0006835	0.85577	0.86988	0.86063	0.96924	0.03076	0.7159
Összesített												5.9073x10 ⁻²⁰	

Hardy-Weinberg egyensúly (HW) egzakt teszt *p*-értékek, minimum allélfrekvencia (*p*_{min}): 21473 minta alapján **: korábban publikált adatok alapján, polimorfizmus információ tartalom (Polymorphism Information Content: PIC), várt heterozigócia (*H*_E), észlelt heterozigócia (*H*_O), *F*_{IS}, szignifikáns *F*_{st} értékek (*p*<0.05) két adat között, átlagos egyezési valószínűség (Match Probability: *pM*_{ave}) és szignifikáns *F*_{IS} értékekkel korrigálva. A megkülönböztető erő (Power of Discrimination: PD), kizárási esély (Exclusion Power: PE), és az összesített egyezési valószínűség (Average Match Probability: *pM*_{ave}).

A két adatbázis közötti lehetséges genetikai különbség vizsgálatára számolt statisztikai paraméterek szignifikánsak lettek, de nagyon alacsony értékűek. A 21473 és a 4213 személyből álló mintapopulációk között tehát kismértékű genotípus eloszlásbeli különbség volt kimutatható. A szignifikáns, páronkénti *F*_{st} értékek -0,0007 és 0,0012 közé sorolódtak. A két adatbázis közötti különbség tesztelésére egy szimulációs modellt készítettünk. A modellhez a rutin esetekből 101 teljes és parciális helyszíni profilt véletlenszerűen választottunk ki. A profilokkal azonos típusú bűnügyi esetet szimuláltunk. Az esetet úgy tekintettük, mintha 101 bűncselekmény helyszínén rögzített biológiai anyagmaradványból

meghatározott, olyan DNS-profil vizsgálnánk, amely teljes egyezést adott volna a referencia minta (gyanúsított) DNS-profiljával. A 101 DNS-profil egyezési valószínűségét (pM) a mindkét magyar populációs adatbázis allélgyakoriságaival kiszámoltuk és az eredményeket a **22A. ábrán** prezentáltuk. A pM értékeket korrigáltuk a p_{min} és HWE-től való eltérés mértékével (F_{IS}). A két adatbázis allélgyakoriságaival számolt pM értékek különbségét ΔpM -ként adtuk meg és a **22B. ábrán** mutattuk be.

A 101 profil közül 80 DNS-profil pM értékei magasabbak voltak, ha a 21473 személyt tartalmazó referencia adatbázist használtuk. A ΔpM értékek azonban nem voltak magasabbak a két adatbázis allélgyakorisági adataival számolt pM arány 27-szeresénél [$pM_{21473}/pM_{4213}=27$], ld. például az R083 profilt a **22B. ábrán**. A profil egyezési valószínűségek (pM) 95% -os konfidencia intervalluma a 21473 mintát tartalmazó nagyobb adatbázis esetében szűkebb, ez várható is volt, ugyanis a nagyobb mintaszám pontosabb genotípus gyakoriság becslést tesz lehetővé (**22A. ábra**, piros rombusz: új adattal számolt pM értékek; kék rombusz: a régi adattal számolt pM értékek).

A minimum allélfrekvencia (p_{min}) korrekciójának és a HWE-től való eltérésnek (F_{IS}) a hatását is teszteltük a 101 profilra kiszámolt egyezési valószínűségi értékeken. Ha a pM valószínűségeket F_{IS} értékkel korrigáltuk, akkor 27, ha nem, akkor 26 DNS-profil pM értéke mutatott szignifikáns különbséget a két különböző adatbázis allélgyakorisági adataival számolva. A két adatbázis genotípusos eloszlásának különbsége tehát a HWE-től való eltérés mértékével korrigálva lényegében nem változik. Mindössze 5 profil pM értékére volt szignifikáns hatással az F_{IS} korrekció.

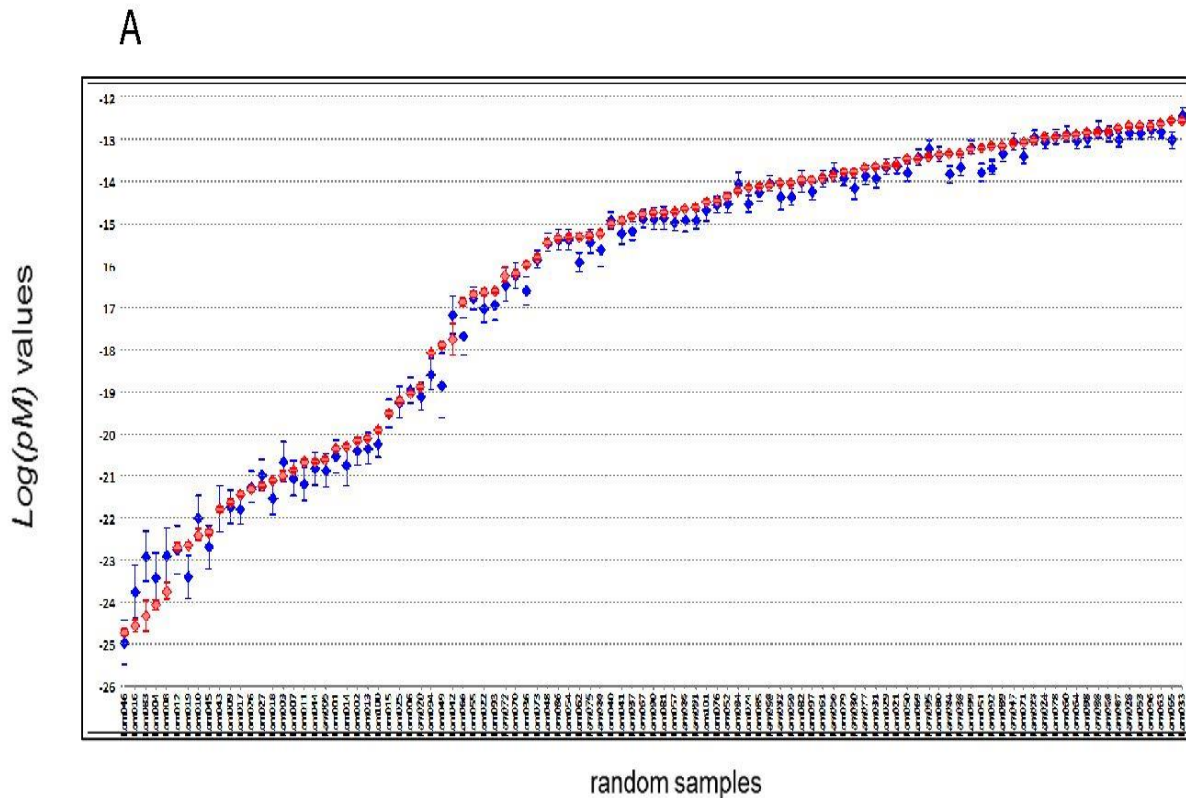
A p_{min} egyetlen minta esetében módosította a mindkét adatbázis alapján kalkulált pM értéket, mivel a mintában az FGA lokuszon detektált 24.1-es allél hiányzott mindkét adatbázisból. 6 mintánál a 4213-as adatbázis allélgyakorisági értékeivel számolva magasabb LR eredményt kaptunk, ha a p_{min} -t nem alkalmaztuk.

A **23. ábrán** egy személy DNS-profilját – 16 lokusz vizsgálatával - mutatjuk be.

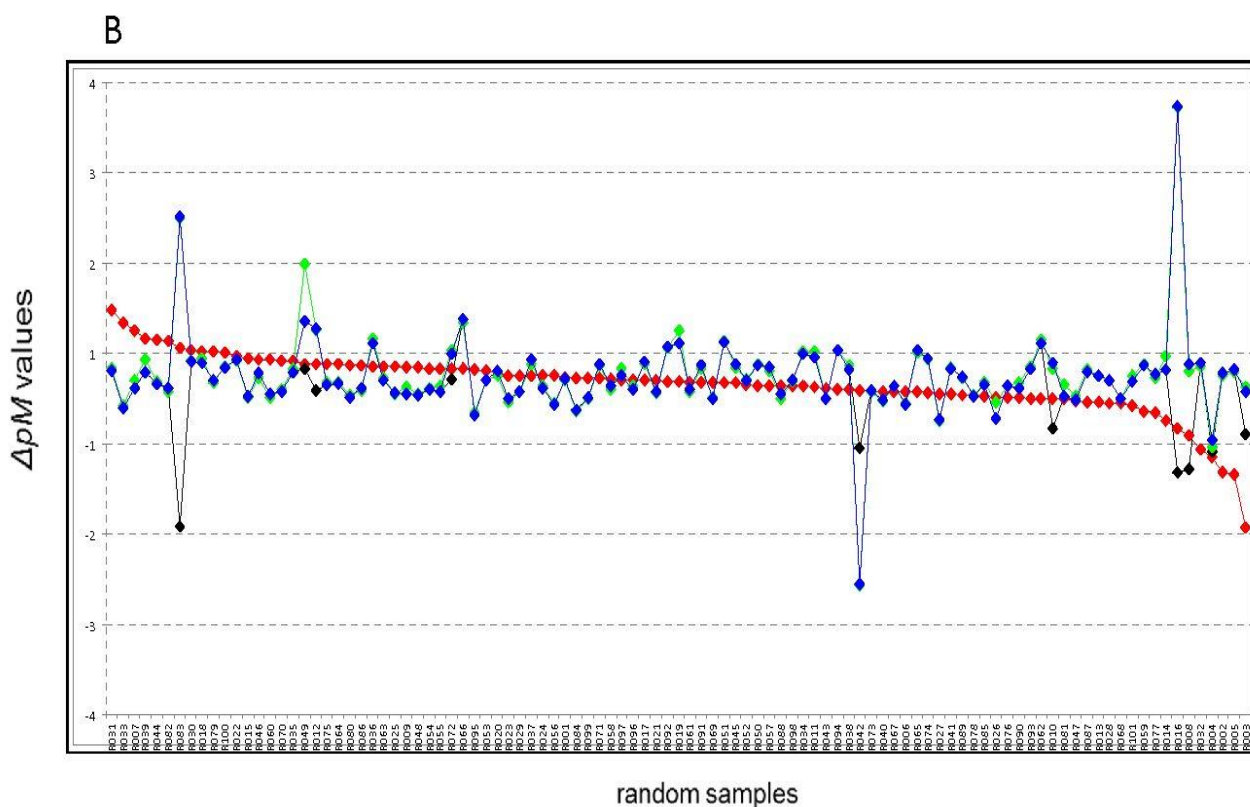
Összesítve tehát tapasztalatainkat, az F_{IS} értékkel való korrekció kevésbé befolyásolta a kapott pM értékeket, mint a p_{min} alkalmazása. A valószínűségi hányadosok (LR) számításnál javasoljuk a p_{min} alkalmazását a korábban publikált két populációs adat használata esetén [135, 136].

Egyértelműen megállapítható, hogy a nagyobb mintaszámú adatbázis használata sokkal pontosabb LR számítást tesz lehetővé, mint a korábban publikált referencia adatbázis. Az adatbank mintákban lehetségesen előforduló rokoni kapcsolatoknak strukturális hatása nem volt kimutatható, de bizonyos lokuszok esetén befolyással

lehettek a genotípusos egyensúly eltolódására. Ennek mértéke azonban nincs szignifikáns hatással az egyezési valószínűségek (pM) értékének számítására. A profilgyakorisági kalkulációk esetében nagyobb jelentőséggel bír a minimum allélfrekvenciával való korrekció betartása.



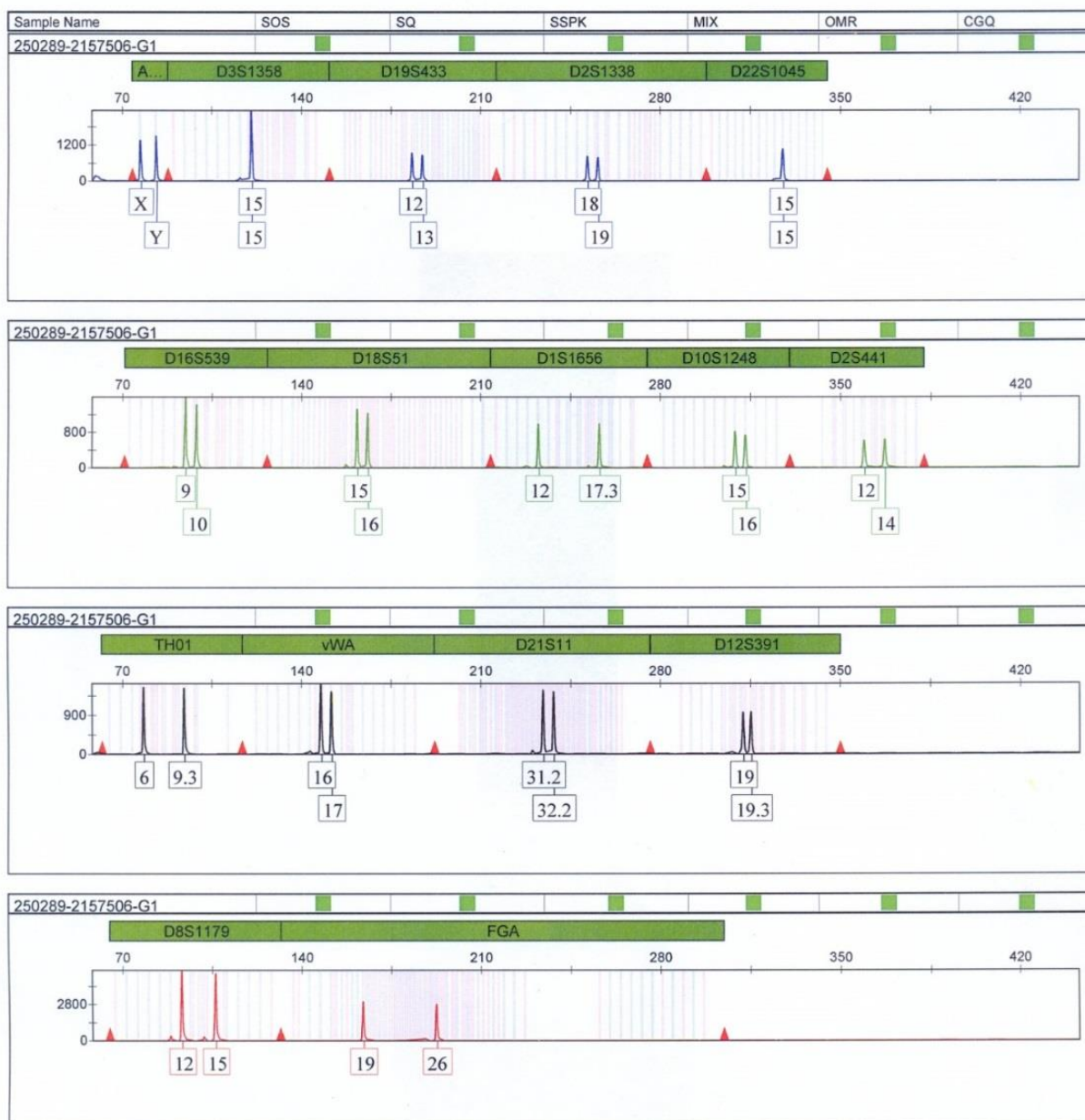
22A. ábra. 101véletlenszerűen kiválasztott minta becsült profilgyakorisága (pM). A becsült 10-logpM értékek (rombusz jelek) és azok 95%-os konfidencia intervalluma (vonalak határolta tartományban) az új adatokkal (piros szín) és a korábbi adatokkal (kék szín) számolva.



22B. ábra. A ΔpM számítása F_{IS} és p_{min} bevonásával (piros vonal), F_{IS} bevonásával és p_{min} nélkül (zöld vonal), p_{min} bevonásával és F_{IS} nélkül (fekete vonal), F_{IS} és p_{min} nélkül (kék vonal) az F_{IS} és p_{min} értékekkel korrigált ΔpM adatok sorrendjébe rendezve.

Az ábrán látható, hogy az F_{IS} és p_{min} értékek bevonásával sokkal megbízhatóbb LR érték számítható, mint azok bevonása nélkül.

Összefoglalva elmondható, hogy a magas (21473) és alacsonyabb (200-500) mintaszámú vizsgálati eredmények között nincs szignifikáns különbség. A büntető ügyekben ajánlatos a 21472 minta alapján számolt allélgyakoriságokat használni, mivel ezen adatok elfogulatlanabbak, pontosabbak, így módon a nagyobb adatbázis potenciálisan lehetővé teszi alacsonyabb allélgyakoriságok használatát.

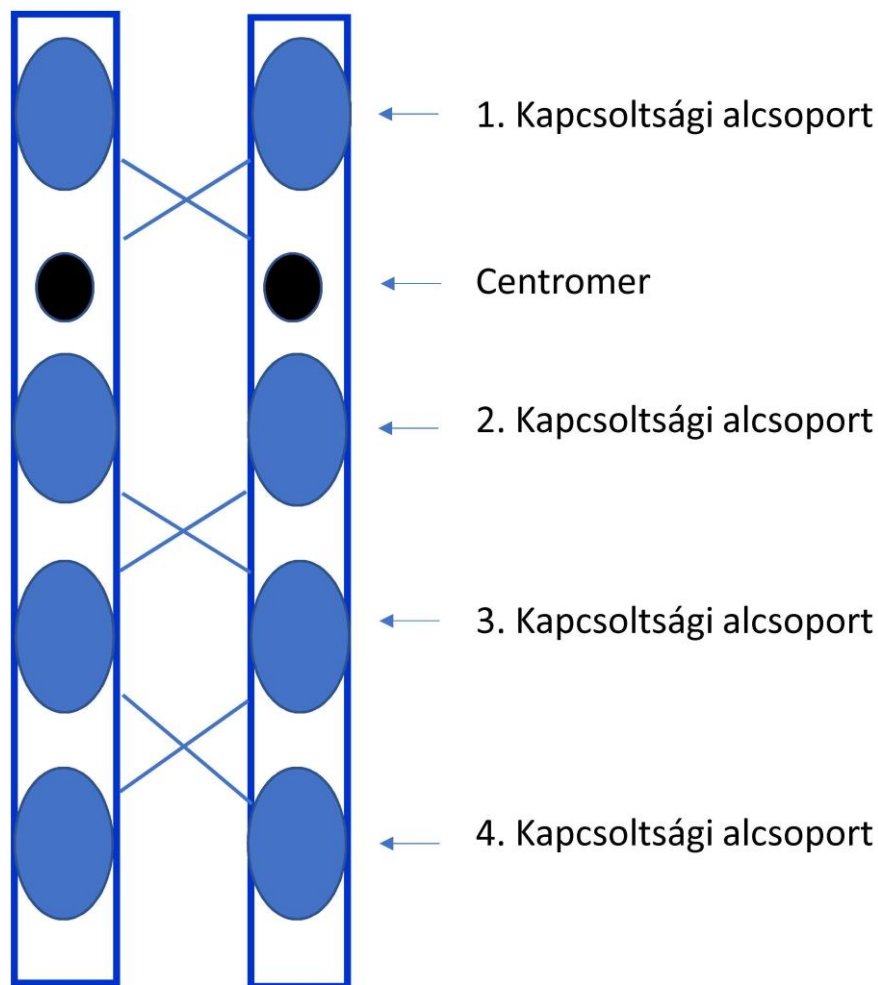


23. ábra. Egy személy DNS-profilja 16 STR lokusz vizsgálatával: a lokuszok nevei zöld háttéren, a lokuszok méretei bázisban a felső skálán láthatók.

5.1.2. X-rekombináció vizsgálata, X-STR markerekkel

Az X-kromoszómális STR vizsgálatok különös szerepet játszanak a kriminalisztikai és populációgenetikában. Ezek a markerek egyetlen példányban vannak jelen a férfiakban és két példányban a nőknél. Az X-kromoszómális lokuszok elemzése előnyös lehet hiányos apasági ügyekben, amikor az utód leány és a vélelmezett apa hiányzik, vagy az anyasági ügyekben. Az X-kromoszómális rekombináció vizsgálata nagy jelentőséggel bír, különösen akkor, ha az X-STR markereket törvényszéki rokonsági elemzésben vizsgáljuk.

Az X-STR-ek tulajdonságainak tanulmányozása nemcsak az allél és haplotípus frekvenciák pontos ismerete miatt érdekes genetikailag és statisztikailag, hanem a genetikai linkage equilibrium (LE) és a linkage disequilibrium (LD) szempontjából is. Ezek hatását a markerek öröklődésére szintén figyelembe kell venni az interpretáció során és a valószínűségi százalék számításánál, lévén ezek a markerek ugyanazon a kromoszómán helyezkednek el. A kereskedelembe kapható Investigator Argus X-12 marker készlet (Qiagen Németország) lehetővé tette a DXS10148-DXS10135-DXS8378, DXS7132 DXS10079-DXS10074, DXS10103-HPRTB-DXS10101 és DXS10146-DXS10134-DXS7423 mikroszatellita markerek tesztelését az X-kromoszóma négy kapcsoltsági alcsoportjára (**Cluster: CL**) egy multiplex reakcióban (**24. ábra**).



24. ábra. Az X-kromoszóma kapcsoltsági alcsoportjai

A tanulmány célja az X-kromoszóma rekombinációs eseményeire vonatkozó információk feltárása három generációs családokban (nagyapák lányaikkal és fiú unokáikkal együtt), valamint annak megismerése, hogy az X-STR-ek hogyan viselkednek a család és a populáció szintjén. Perifériás vért gyűjtöttünk 45 három generációs családból ($N = 155$ fő), amelyek Magyarország különböző régióiból származtak.

Korábban 219 magyar férfi vizsgálatával négy kapcsoltsági alcsoporton belül lokalizált 4, 8 és 12 X-STR lokusz haplotípus frekvenciáit publikáltuk [15, 16, 17]. A kapcsoltsági alcsoportok közötti megfigyelt rekombinációs rátát hasonlítottuk össze a genetikai távolságokból konvertált várható rekombinációs rátákkal (r) Haldane térképfüggvénye segítségével [21, 137]:

$$r_{1-2} = \frac{1}{2}(1 - e^{-2d_{1-2}})$$

ahol d_{1-2} átlagos crossing over gyakoriság a lokuszok (1. és 2. pont) között, r_{1-2} a rekombináció megfigyelt gyakorisága a lokuszok között. A fizikai távolságokat átszámoltuk genetikai távolságokra, feltételezve, hogy 1 Mb felel meg 1 cM-nek [138].

A rekombinációs eseményeket összegezve látható, hogy 11 esetben kaptuk azt az eredményt, hogy a meiózis során nem történt crossing over. Ilyenkor a teljes nagyanyai vagy nagyapai X kromoszóma adódik tovább két generáción keresztül. Az X-kromoszóma rövid és hosszú karján megfigyelhető rekombinációs események számát összesítve a **3. táblázat** tartalmazza.

3.táblázat. Rekombinációs események száma az X-kromoszóma rövid és a hosszú karján

rekombinációs események száma a rövid karon	rekombinációs események száma a hosszú karon	esetek száma
Xp, 0	Xq, 0	11
	Xq, 1	19
	Xq, 2	6
Xp, 1	Xq, 0	11
	Xq, 1	9
	Xq, 2	3

Mutációt viszonylag alacsony számban találtunk: három családnál a nagyapa és az anya között, két családnál pedig az anya és a gyermek között történt mutáció az ivarsejtképződés során. A mutációk minden esetben egy repeat eltérést jelentettek (példaként ld. a **4. táblázatot**).

4. táblázat. Példa öröklődő mutációra (mutáns allél lila háttérrel jelölve)

Család 6	CL 1			CL 2			CL 3			CL 4		
	DXS 10148	DXS 10135	DXS 8378	DXS 7132	DXS 10079	DXS 10074	DXS 10103	HPRTB	DXS 10101	DXS 10146	DXS 10134	DXS 7423
Nagyapa	25.1	30	11	14	23	18	19	12	27	41.2	36	14
Anya	25.1	29	11	14	23	18	19	12	27	41.2	36	14
	18	18	12	12	23	15	19	12	27	30	34	16
Fiúunoka1	25.1	29	11	12	23	15	19	12	27	30	34	16
Fiúunoka2	25.1	29	11	14	23	18	19	12	27	30	34	16

Kapcsoltsági alcsoporton belüli rekombináció is megfigyelhető volt, - bár az is lehetséges, hogy 2-lépéses mutáció történt - ez igen kis számban, összesen egy esetben fordult

elő, mégpedig a 2. kapcsoltsági alcsoportban (**5. táblázat**). Ezt természetesen számításba vettük a kromoszóma karokon megfigyelhető rekombinációs események összegzésekor.

5. táblázat. Példa kapcsoltsági csoporton belüli rekombinációra (zöld háttérrel jelölve)

Család 35	CL 1				CL 2				CL 3				CL 4		
	DXS 10148	DXS 10135	DXS 8378		DXS 7132	DXS 10079	DXS 10074		DXS 10103	HPRTB	DXS 10101		DXS 10146	DXS 10134	DXS 7423
Nagyapa	27.1	18	10		14	17	8		18	13	32		27	37	15
Anya	27.1	18	10		14	17	8		18	13	32		27	37	15
	25.1	19	11		16	19	14		18	11	30.2		28	39	14
Fiúunoka	27.1	18	10		14	19	14		18	11	30.2		27	37	15

Bizonyos esetekben nem lehet eldönteni, hogy rekombináció vagy mutáció történt. Ez abban az esetben fordul elő, amikor a két nagyszülői allél egy repeat-ben tér csak el egymástól (**6. táblázat**). A 2. kapcsoltsági alcsoportban a fiúunoka a DXS7132 lokuszon látszólag nagyanyai, míg a DXS10079 és DXS10074 lokuszokon nagyapai allélokat örökölt. Ez a kombináció kétféleképpen jöhetett létre. Az első esetben az anyában a meiózis során a nagyapai 16-os allél mutálódott 15-re, ami megegyezik a nagyanyai alléllal. Ebben az esetben a nagyapa haplotípusa öröklődött egy mutációval. A második esetben a szorosan kapcsolt DXS7132 és DXS10079 lokuszok között rekombináció történt az anyában a meiózis során, így a fiúunoka a 2. kapcsoltsági alcsoportban vegyesen örökölt nagyapai és nagyanyai allélokat. A szakirodalom szerint a második eset igen ritka [139], ezért nagy valószínűséggel az első eset következhetett be (**6. táblázat**). Ez összesen kétszer fordult elő a 45 család vizsgálata során.

6. táblázat Példa nem eldönthető esetre (a mutáció/rekombináció narancs színnel jelölve)

Család 43	CL 1				CL 2				CL 3				CL 4		
	DXS 10148	DXS 10135	DXS 8378		DXS 7132	DXS 10079	DXS 10074		DXS 10103	HPRTB	DXS 10101		DXS 10146	DXS 10134	DXS 7423
Nagyapa	19	25	11		16	21	16		18	12	32		40.2	40.3	15
Anya	19	25	11		16	21	16		18	12	32		40.2	40.3	15
	25.1	22	10		15	20	7		16	13	32		31	36	16
Fiúunoka	25.1	22	10		15	21	16		16	13	32		31	36	16

5.1.2.1. Linkage analízis a férfiak X-kromoszómális haplotípusa alapján

A linkage disequilibrium tesztelésére végzett Fisher-féle egzakt teszt eredményeit a **7. táblázat** tartalmazza. Látható, hogy szignifikáns p-értéket három esetben kaptunk: a HPRTB - DXS10103, DXS10101 - DXS10103, DXS10101 - HPRTB lokuszok között.

7. táblázat. A linkage disequilibrium tesztelésének p-értékei; Fisher-féle egzakt-teszt, a szignifikáns p-értékek ($p < 0,00091$ Bonferroni korrekcióval) vastag betűvel szedve

Lokusz	DXS 10148	DXS 10135	DXS 8378	DXS 7132	DXS 10079	DXS 10074	DXS 10103	HPRTB	DXS 10101	DXS 10146	DXS 10134	DXS 7423
DXS10148	-	0.81262	0.95441	0.64928	0.10319	0.84047	0.87412	0.95311	0.50424	0.29922	0.27861	0.94703
DXS10135		-	0.18360	0.79271	0.69192	0.57079	0.12315	0.33172	0.63789	0.18529	0.77662	0.78902
DXS8378			-	0.43241	0.89481	0.78866	0.24968	0.64002	0.12660	0.38015	0.14950	0.71909
DXS7132				-	0.09996	0.07579	0.04204	0.99908	0.67716	0.16042	0.83086	0.17072
DXS10079					-	0.03870	0.46422	0.21034	0.78212	0.23533	0.34331	0.43091
DXS10074						-	0.14690	0.78131	0.78300	0.25453	0.14268	0.83939
DXS10103							-	0.00915	0.00000	0.28295	0.15946	0.63301
HPRTB								-	0.00004	0.92155	0.48734	0.18857
DXS10101									-	0.44970	0.69204	0.02451
DXS10146										-	0.04598	0.83002
DXS10134											-	0.54668
DXS7423												-

A fizikai távolságot a gyártó által megadott pozíciók alapján számoltuk ki. Ez alapján az 1. és a 2. alkapcsoltsági csoport 57 Mb távolságra van egymástól. Az eredmények azt mutatják, hogy az általunk vizsgált mintákban az eldönthető esetek 38,7 %-ában történt rekombináció, ez 74 cM genetikai távolságot jelent, tehát a rekombináció gyakorisága a vártnál alacsonyabb. Ennek oka lehet, hogy az 1. kapcsoltsági alcsoport az X kromoszóma rövid karján helyezkedik el, míg a 2. kapcsoltsági alcsoport a hosszú karon, a centroméra közelében, és a centroméra környékén a rekombináció repressziója jellemző (heterokromatin régió).

A 2. és 3. alkapcsoltsági csoport távolsága 67 Mb, tehát itt is azt vártuk volna, hogy a rekombináció gyakorisága 50% körül legyen. Az informatív meiózisok 41,0%-ában

tapasztaltunk rekombinációt, ez 86 cM genetikai távolságnak felel meg, ami viszonylag megközelíti a várt eredményt.

A 3. és 4. alkapcsoltsági csoport távolsága 16 Mb. Az eldönthető esetek 36,1%-ában kaptunk rekombinációt, ez 64 cM genetikai távolságnak felel meg. Ez az eredmény magasabb a vártnál és nem egyezik meg korábbi kutatások eredményével, mely szerint a 3.- 4. alcs csoport közti genetikai távolság nagyjából 34 cM-nek felel meg [21, 23]. A kapcsoltsági alcs csoportok közti rekombinációs értékeket a **8. táblázat** tartalmazza.

8. táblázat. Rekombinációs gyakoriságok a kapcsoltsági alcs csoportok között

Kapcsoltsági csoport	Fizikai távolság	Genetikai távolság (számolt)	Rekombinációk száma/ informatív meiózis	Rekombinációs gyakoriság
1/2	57 Mb	74 cM	24/62	0,387
2/3	67 Mb	86 cM	25/61	0,410
3/4	16 Mb	64 cM	22/61	0,361

Általánosságban elmondható, hogy a rekombinációs események száma viszonylag alacsony, eredményeink nem bizonyították rekombinációs hotspot jelenlétét az 1.- 2. illetve a 2.-3. kapcsoltsági alcs csoport között az X kromoszómán. A 3.- 4. kapcsoltsági alcs csoport között megfigyelhető, a fizikai távolsághoz képest viszonylag magas, genetikai távolság rekombinációs hotspot jelentére utalhat, ezt az észrevételt egy másik tanulmány is alátámasztja [21]. A kapcsoltsági alcs csoporton belüli rekombináció igen ritka, tehát az egy alcs csoportba tartozó lokuszok valóban kapcsoltnak öröklődnek. A mutációs események száma is alacsonynak bizonyult.

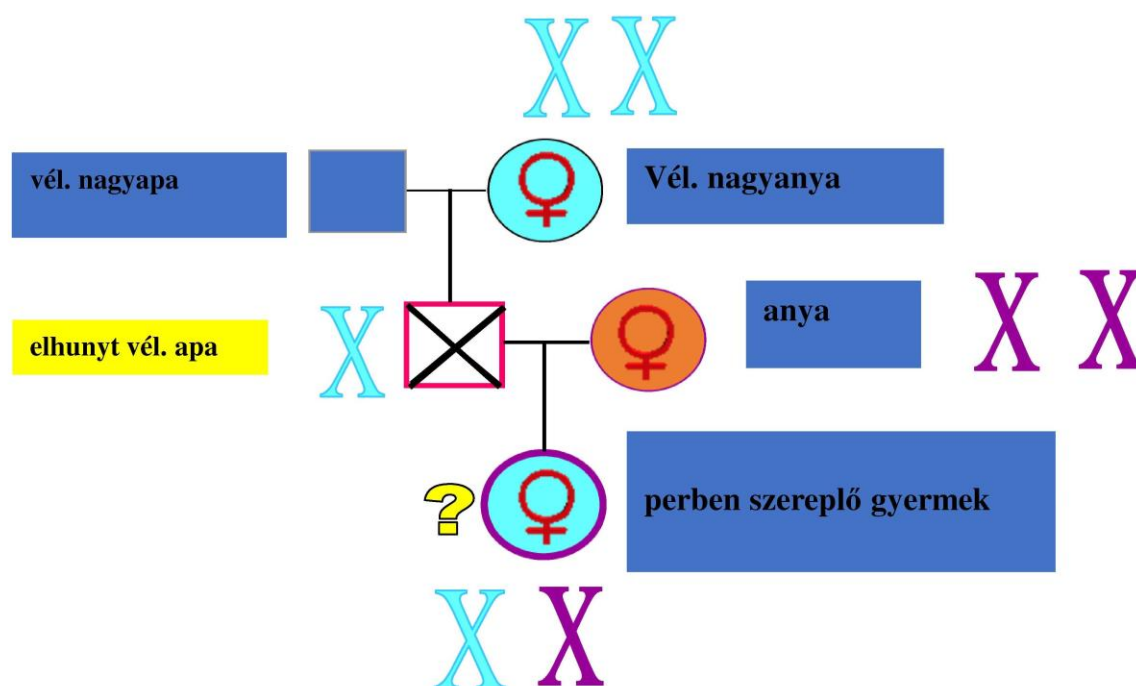
A linkage disequilibrium jelenségét sikerült bizonyítani, a 3. kapcsoltsági alcs csoporthoz tartozó lokuszok alléljai nagy gyakorisággal stabil allélkombinációkban, haplotípusként öröklődnek (ld. a **7. táblázatban** vastagított értékeket) egyik generációról a másikra [22].

5.1.3. X-kromoszómális STR markerek vizsgálata hiányos apasági ügyekben

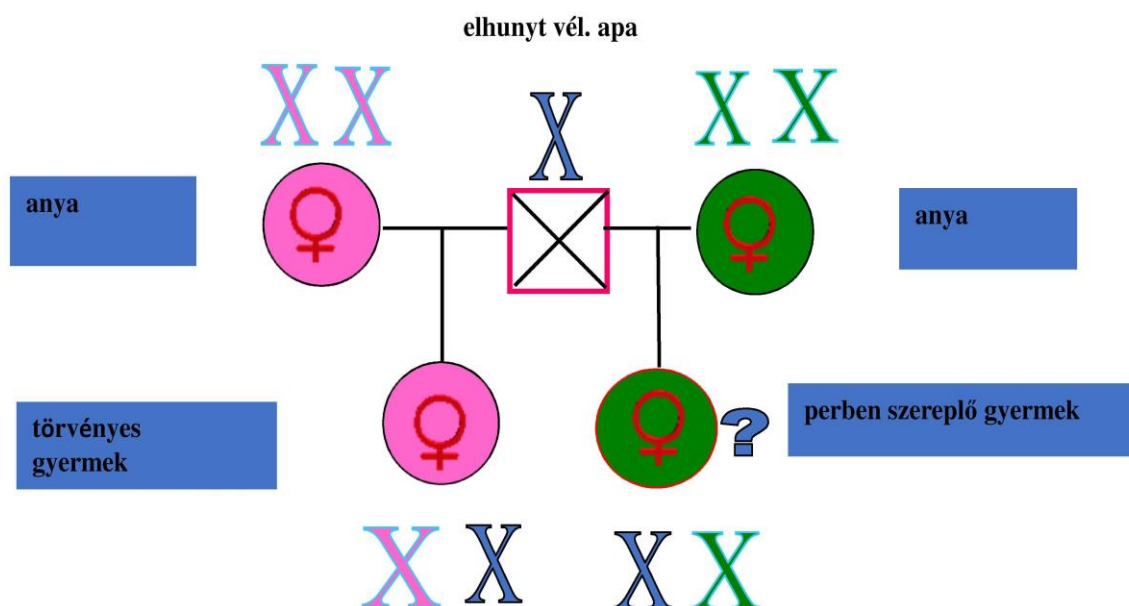
Az X-kromoszómális markereknek az igazságügyi genetikában leginkább származás-megállapítási ügyekben van jelentősége. A hiányos apasági ügyekben az X-STR lokuszok vizsgálata abban az esetben célravezető, ha a szülő-gyermek kapcsolat legalább egyik tagja

nő, mivel az édesapák egyetlen X-kromoszómája mindenképpen öröklődik leánygyermekükbe, a fiúgyermek pedig az édesanyjuktól öröklik az X-kromoszómájukat.

Az X-kromoszómás markerek vizsgálata pl. akkor jelenthet megoldást a gyakorlatban, ha a vélelmezett apa elhunyt, és a kérdéses gyermek leány. A lánygyermeket, annak édesanyját és a vélelmezett apa édesanyját vizsgálva az X-kromoszómás markerekkel úgy számolhatunk, mintha a gyermeket, édesanyját és a vélelmezett apát vizsgálnánk autoszómális markerekre. Mivel a vélelmezett apába az anyai X-kromoszóma kerül be a megtermékenyítéskor és ugyanaz az X-kromoszóma adódik tovább a lánygyermekébe, ezért az X-kromoszóma nyomon követése, azaz a vélelmezett nagymama X-kromoszómájának vizsgálata célravezető ilyen típusú hiányos apasági vizsgálatokban (25. ábra). Ha a vélelmezett nagymama nem vizsgálható, akkor a perben szereplő lánygyermeket, édesanyját, a vélelmezett apa törvényes lánygyermekét (ha létezik) és annak édesanyját vizsgáljuk; ebben az esetben a két lánygyermek apai X-kromoszómáját hasonlíthatjuk össze (26. ábra).



25. ábra. Pedigré: a vélelmezett nagymama az „apajelölt”

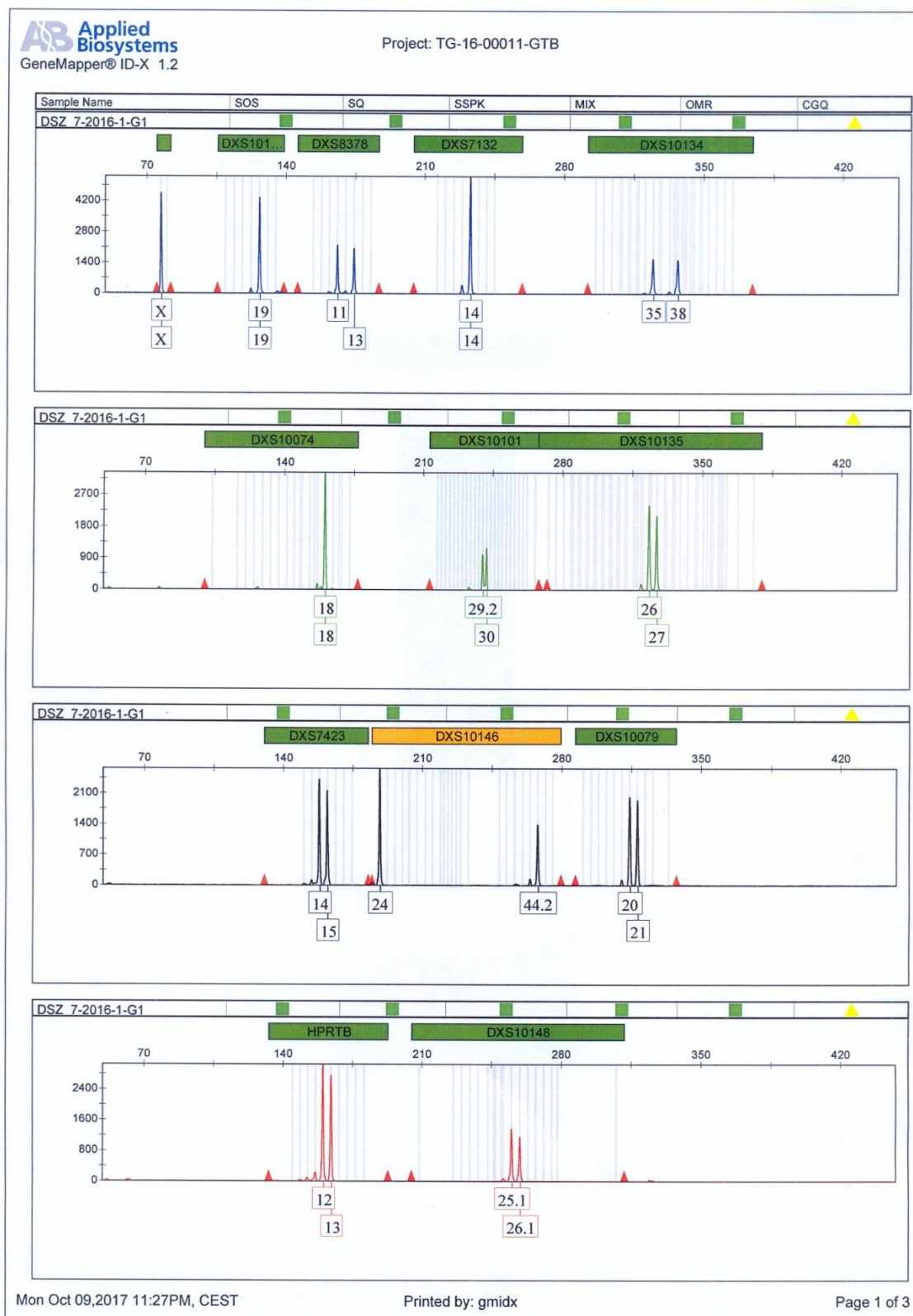


26. ábra. Pedigré: a vélelmezett lány féltestvér az „apajelölt”

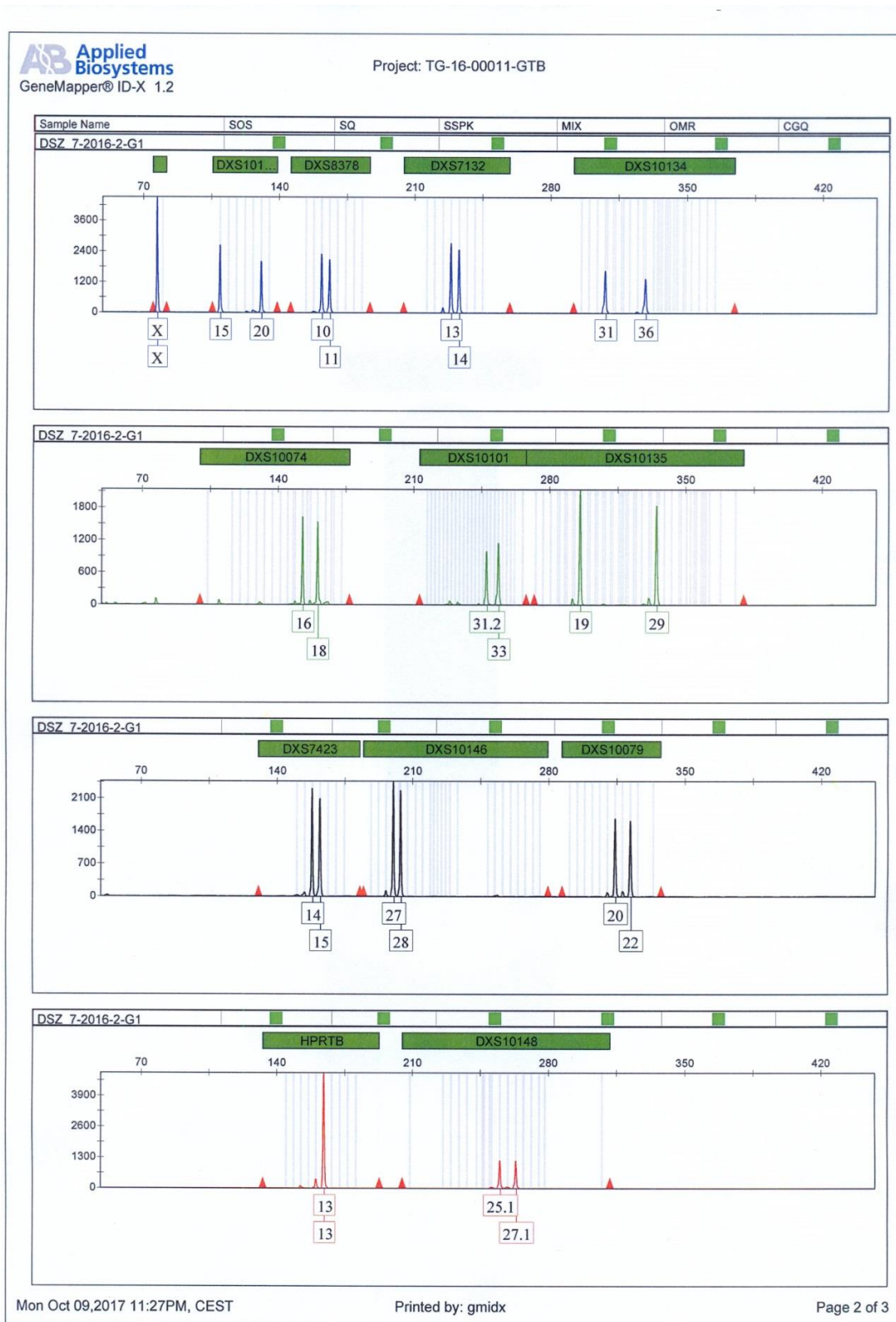
Standardizált vizsgálati vegyszercsomagok („kit”-ek) (Biotype, Qiagen) használatával mára lehetségessé vált X-STR lokuszok [15, 16, 17] vizsgálata. Az X-STR lokuszok négy, különböző kapcsoltsági régióban helyezkednek el, mindegyik kapcsoltsági alcsoporton belül 3-3 szorosan kapcsolt lokusz helyezkedik el, amelyek úgy viselkednek, mintha haplotípusok lennének. A lokusz-triókon belül a rekombináció valószínűsége csekély, míg közöttük gyakori a rekombináció (24. ábra).

Az Investigator Argus X-12 kit igazságügyi genetikai alkalmazhatóságát egy hiányos apasági ügy vizsgálatán keresztül mutatom be.

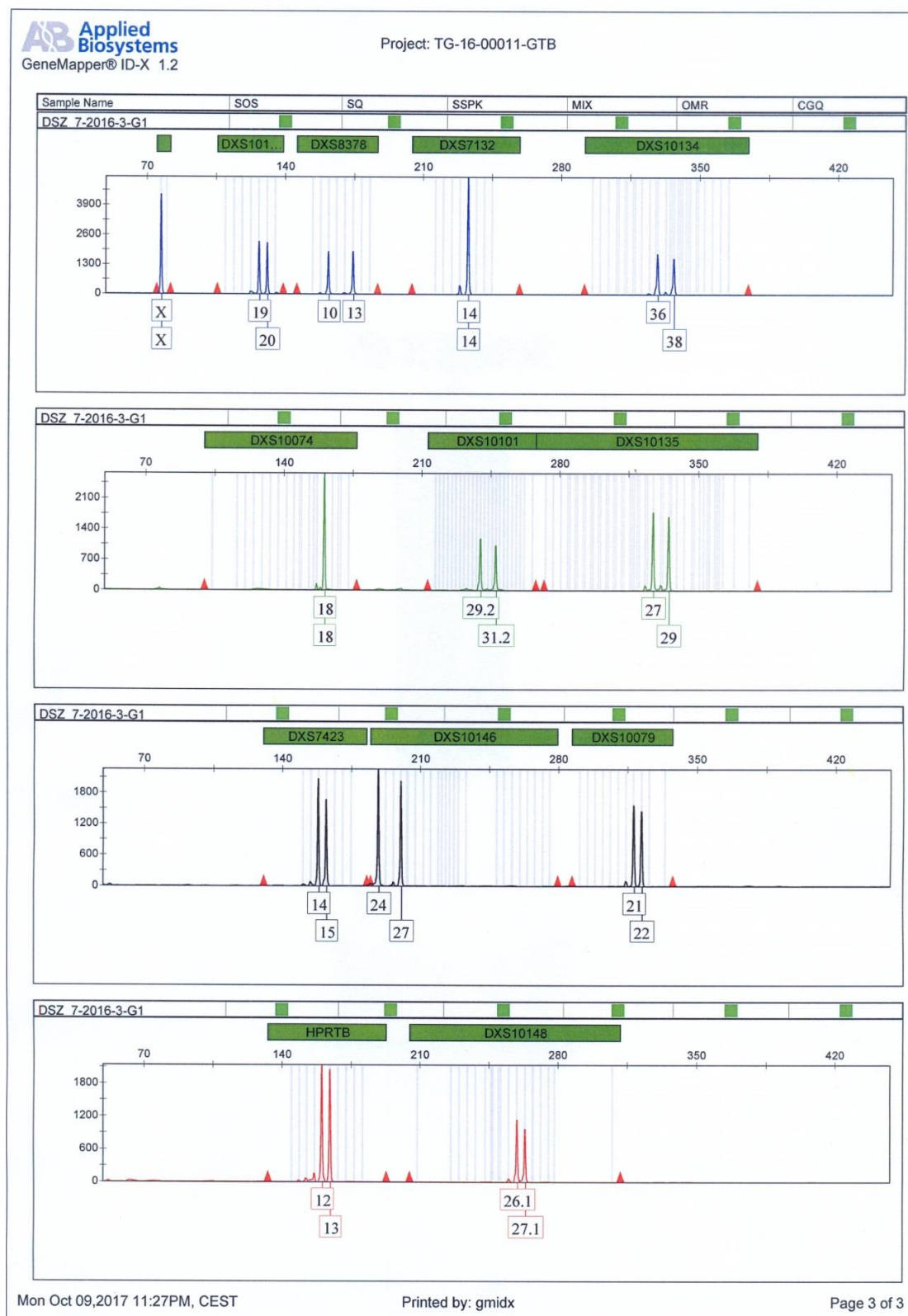
A lehetséges pedigrék közül az első alapján (25. ábra) három személy (anya, felperes leánygyermek és vélelmezett nagyanya) igazságügyi genetikai vizsgálatát végeztük el, melynek során a három vizsgált személy X-kromoszómális genotípusát 12 lokuszra nézve határoztuk meg (27A., 27B., és 27C. ábra), a gyermek és az anya genotípusának, valamint a vélelmezett nagyanya genotípusának összevetésével definiáltuk a gyermekek lehetséges apai alléljait.



27A. ábra. A vélelmezett nagymama X-kromoszomális DNS-profilja



27B. ábra. Az anya X-kromoszomális DNS-profilja



27C. ábra. A kiskorú gyermek X-kromoszomális DNS-profilja

A kiskorú felperes és anyja, illetve a vélelmezett nagymama X-kromoszómáin található allélok összevetésekor nem találtunk kizáró allélkombinációt, ezért a leszármazást az apasági valószínűség statisztikai kalkulációjával véleményyeztük (**28. ábra**).

Az X érték annak a valószínűsége, hogy a felperes gyermek biológiai apja a vélelmezett apa (felperes hipotézise). Az édesanya és a gyermek genotípusát összevetve megállapítjuk a biológiai apa genotípusát. A lokusztriókon lévő összes lehetséges allélkombinációt figyelembe kell venni. Az 1. és 4. kapcsoltsági alcsoportban mindhárom lokuszán heterozigóta volt a vélelmezett nagymama, így mindkét kapcsoltsági alcsoportban (cluster) összesen nyolcféle lehetséges haplotípust adhatott tovább (ezért $X=0,125=1/8$ az adott kapcsoltsági régióra nézve), amit a vélelmezett apa örökölhett és változatlanul (haplotípusként) tovább is adhatott lánygyermekének, amennyiben a gyermek tőle származik. A 2. kapcsoltsági alcsoportban a vélelmezett nagymama 2 lokuszon homozigóta volt, tehát a vélelmezett apa genotípusa (haplotípusa) kétféle lehet, ezért a felperesi hipotézis értéke $X=0,5=1/2$. A 3. kapcsoltsági alcsoportban egy lokuszon (DXS10103) volt homozigóta a vélelmezett nagymama (19-19), ezért a lokusztrió négyféle kombináció lehetséges ($X = 0,25=1/4$). **Az Y érték annak a valószínűsége, hogy a felperes gyermek biológiai apja nem a vélelmezett apa, hanem egy másik, a népességből (jelen esetben magyar) véletlenszerűen kiválasztott, a vélelmezett apával megegyező allélokkal (genotípussal) rendelkező potenciális donor (alperes hipotézise).** Az Y érték a magyar populációra kiszámított lokusztriókénti haplotípus gyakoriságok szorzatának felel meg. Az első három kapcsoltsági alcsoportban egyértelműen megállapítható, hogy milyen allél trióból állhat az apai genotípus. Viszont a 4. kapcsoltsági alcsoportban a DXS7423 lokuszban az anya és a gyermek alléljai megegyeznek (14-15), ebben az esetben nem lehet eldönteni, hogy melyik az apai allél, tehát a vélelmezett apa haplotípusa a lokusz trión kétféle lehet (24-38-**14** vagy 24-38-**15**), ezért Y értéke a két haplotípus gyakoriságának összege ($0,0046+0,0046=0,0092$). A haplotípus gyakorisági értékek (a **28. ábrán** az Y oszlop) saját publikációból származnak [15].

X/Y adja a **PI** (Paternity Index, apasági index) **értékét**, ebből az apasági valószínűség százalékos értéke **$W\% = PI/(1+PI)$** képlettel számítható ki. Az ismertetett esetben az apasági valószínűségi százalék értéke meghaladta a hazai és nemzetközi szakmai előírásban megadott minimum 99,99%-ot ($PI=10000$), ami „**az apaság gyakorlatilag bizonyított**” verbális kategóriának felel meg.

Adatok	Nevek		DSZ alszám				
Vélelmezett apai nagyanya	Napsugár		DSZ 7/2016/1		Bíróság:	Magna Hungaria Járásbíróság	
Anya	Dzsesszika		DSZ 7/2016/2		Ügyszám	2.P.20.179/2015/22.	
Gyermek	kk. Holda		DSZ 7/2016/3				
Rendszerek	Nagyanya	Anya	kk. gyermek	X	Y	W%	Acum. W%
DXS10148 DXS10135 DXS8378	25.1-26.1 26-27 11-13	25.1-27.1 19-29 10-11	26.1-27.1 27-29 10-13	0,125	0,0046	96,45%	96,450617%
DXS7132 DXS10079 DXS10074	14-14 20-21 18-18	13-14 20-22 16-18	14-14 21-22 18-18	0,5	0,0137	97,33%	99,899270%
DXS10103 HPRTB DXS10101	19-19 12-13 29.2-30	15-20 13-13 31.2-33	19-20 12-13 29.2-31.2	0,25	0,032	88,65%	99,987095%
DXS10146 DXS10134 DXS7423	24-44.2 35-38 14-15	27-28 31-36 14-15	24-27 36-38 14-15	0,125	0,0092	93,14%	99,999050%
							99,999050%
			"gyakorlatilag bizonyított"				

28. ábra. Az apaság valószínűsítése (anonimizált részlet a szakértői véleményből)

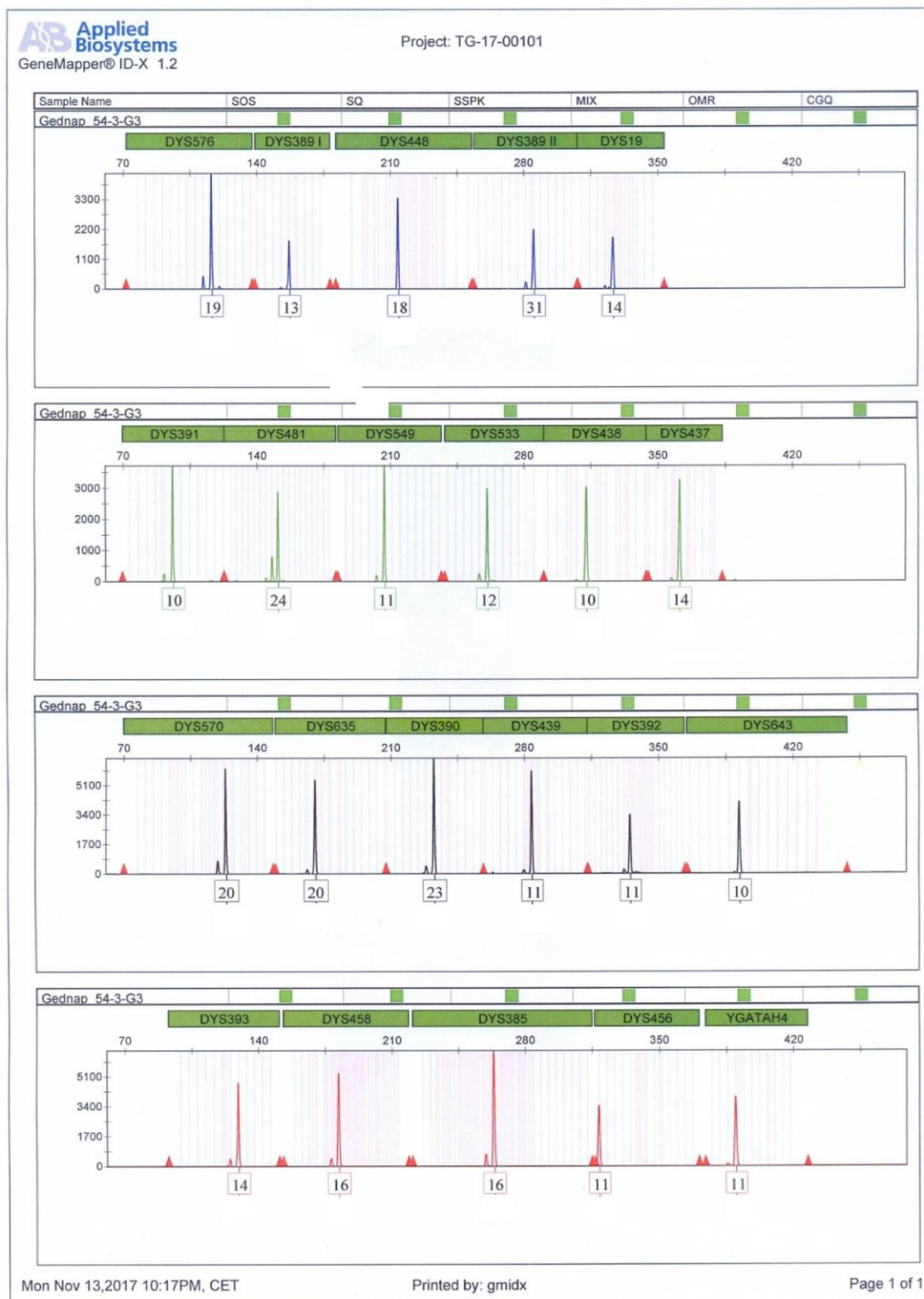
5.1.4. Y-STR lokuszok vizsgálata az igazságügyi alkalmazásban

Y-STR (haplotípus) és Y-SNP vizsgálatokat az elmúlt 17 év során 2045 mintán végeztük. Mind a haplotípus, mind a haplocsoport adatokat az Y Haplotípus Referencia Adatbázisba (Y Haplotype Reference Database: **YHRD**, www.yhrd.org) küldtük, amely az igazságügyi és populáció genetikusok által használt egyetlen nagy nemzetközi adatbázis. Az igazságügyi genetikusok a büntető vagy polgári peres eljárásban szakértői vélemény készítés során az Y-kromoszóma haplotípus gyakoriságát ezen adatbázis segítségével becslik. A haplotípus gyakoriság alapján számolható ki a profilegyezési valószínűség vagy az *a posteriori* apasági valószínűségi érték.

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 nyomán biztosított női hüvelyváladék, ahol értelemszerűen az áldozattól származó laphámsejtek jóval nagyobb arányban találhatók, mint a férfi hímivarsejtjei. Azonkívül erőszakos bűncselekmény során dulakodásból adódóan a női sértett körme alá kerülhetnek férfitől származó bőrsejtek, amelyeket „köröm kaparéknak” hívunk. A körömkaparéokban is túlnyomó részt női DNS található, amelynek következtében a csekély mennyiségű férfi autoszómális DNS vizsgálata sikertelenné válhat. Ezért célszerű, ilyen bűncselekményekből származó biológiai anyagmaradványok vizsgálatánál célzottan a csak férfitől származó Y-kromoszóma STR (haplotípus) tesztelésére összpontosítani.

Az apasági DNS vizsgálatokban csak akkor alkalmazhatók Y-kromoszómális lokuszok, ha a perben szereplő gyermek fiú és a vélelmezett apa nem vizsgálható (elhunyt vagy eltűnt). Ilyen esetekben a vélelmezett apa valamelyik apai ági férfirokonának (apja, nagyapja, testvére, nagybátyja, unokaöccse, unokatestvére stb.) vizsgálata célravezető. Ha a perben szereplő fiúgyermek és a vizsgált férfi azonos haplotípussal rendelkezik, eltekintve a ritkán előforduló mutációs eseményektől, akkor feltételezhetően azonos apai felmenőktől származnak. Ebben az esetben a közös haplotípus gyakoriságot az **YHRD** adatbázisból keressük ki. A haplotípus gyakorisági adat birtokában kiszámolható az *a posteriori* apasági valószínűségi érték (%).

A személyazonosításban vagy származásmegállapítási vizsgálatokban az Y-kromoszómális STR vizsgálatok hátránya az, hogy a család összes férfi tagja (felmenő, lemenő és oldalági férfi rokonok) azonos Y-kromoszómális haplotípussal rendelkezik, ezért az egyedi azonosítás hatékonysága vagy az apasági valószínűségi *a posteriori* értéke alacsonyabb, mint az autoszómális STR vizsgálatok esetében. Példaként látható egy férfi hiányos apasági ügyben készített anonimizált Y-STR profilja (**29. ábra**) és az ügghöz tartozó biostatistikai számítás (**30. ábra**).



29. ábra. A vizsgált férfi és a fiúgyermek Y-STR haplotípusa 23 Y-STR lokusz vizsgálatával¹⁷

¹⁷ Mivel a vizsgált férfi és a fiúgyermek haplotípusa azonos, ezért csak egy Y-STR haplotípus elektroferogramját mutatjuk be.

Nemzeti Szakértői és Kutató Központ

BISZI Genetikai Szakértői Intézet

Szakértői vélemény az apaság valószínűségére Y kromoszómális DNS lokuszok vizsgálata alapján

1. Melléklet

MIR azonosító: GSZLI.SZAK.VEL.MEFNY7

Adatok	Nevek			DSZ alszám		
Vél. apai nagyapa	Hun Attila			169/2017/1	Bíróság:	Magna Hungaria
kk. gyermek	kk. Kond Bendegúz			169/2017/2	Ügyszám:	10.P.30.500/2017/24
Lokuszok	Hun Attila	kk. Kond Bendegúz	X	Y	%	%Acum.
DYS576	19	19	1	0,000025	99,998%	99,99750%
DYS389I	13	13				
DYS448	18	18				
DYS389II	31	31				
DYS19	14	14				
DYS391	10	10				
DYS481	24	24				
DYS549	11	11				
DYS533	12	12				
DYS438	10	10				
DYS437	14	14				
DYS570	20	20				
DYS635	20	20				
DYS390	23	23				
DYS439	11	11				
DYS392	11	11				
DYS643	10	10				
DYS393	14	14				
DYS458	16	16				
DYS385	16-16	16-16				
DYS456	11	11				
Y GATA H4	11	11				
	"gyakorlatilag bizonyított"					
	Results are based upon Release R55 valid as per 2017-10-20 08:34:40 UTC.					99,99750%

Formanyomtatvány kiadás dátuma: 2017.11.17.
Nyomtatás dátuma: 2017.12.04

Kiadás száma: 2.

Oldalszám: 1/1

30. ábra. A vizsgált személyekre vonatkozó *a posteriori* valószínűség

5.1.5. A mtDNS vizsgálat az igazságügyben

Az mtDNS haplotípus vizsgálatát egy szakértői példán keresztül szeretném bemutatni. A Honvédelmi Minisztérium felkérte a volt ISZKI Budapesti Orvosszakértői Intézetet a II. világháborúban a keleti hadműveletek során elesett néhai Á. Gy. magyar királyi honvéd főhadnagy személyazonosítására. Az elhunytnak csak anyai ági rokona volt vizsgálható referenciaszemélyként. Az elhunyttól csontminta, a hozzátartozójától szájnnyálkahártya törlet vizsgálata történt.

Az mtDNS kontroll régió három polimorf szakaszának (HV1, HV2, HV3) PCR sokszorozását, tisztítását és szekvenálási reakcióját, valamint annak tisztítását követően az ABI3130 DNS analízáló készüléken az adott régiók szekvencia meghatározása megtörtént, mindkét minta tekintetében. Mindkét minta mtDNS szekvencia eltérései az rCRS referencia szekvenciától a következők voltak: A263G, -315.1C, -3152.C, T16093C, tehát a haplotípusuk azonos volt. A nemzetközi mitokondriális haplotípus adatbázisban (www.empop.org) 2015. szeptember 28-án végzett lekérdezés során a mintákból kimutatott haplotípussal megegyező haplotípus nem volt megtalálható, vagyis a kimutatott haplotípus nemcsak Európában, hanem a világ népességében is ritkának tekinthető. A kimutatott haplotípus európai népességben becsült gyakoriságának 95%-os konfidencia intervalluma $0,00-5,53 \times 10^{-4}$. A haplotípus gyakorisággal számolt *a posteriori* valószínűség 99,945%. Tehát az eredmény nagymértékben alátámasztja a feltételezést, hogy a két minta/személy anyai ágon közeli rokonságban áll egymással.

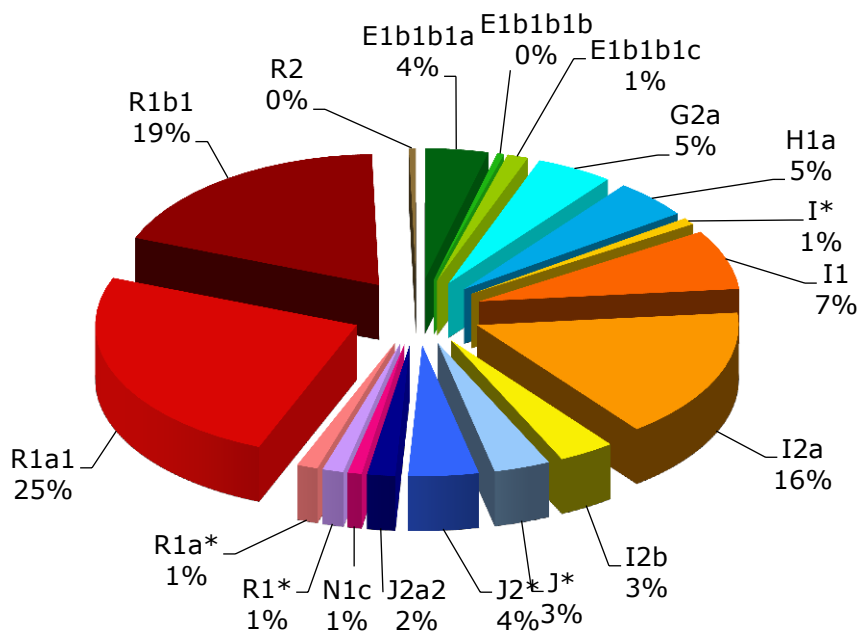
5.2. ÖSTÖRTÉNET: GÉNEKBE VÉSETT TÖRTÉNELEM

5.2.1. A férfi genetikai útja: Y-kromoszóma STR és SNP vizsgálata magyar és roma populációkban

5.2.1.1. Haplocsoport eloszlás a recens magyar populációban

Az első átfogó, komplex Y-kromoszóma STR és SNP vizsgálatot 215 magyarországi magyar férfi minta bevonásával 2009-ben publikáltuk. A tanulmányban 11 Y-STR és 49 Y-SNP lokuszt vizsgáltunk.

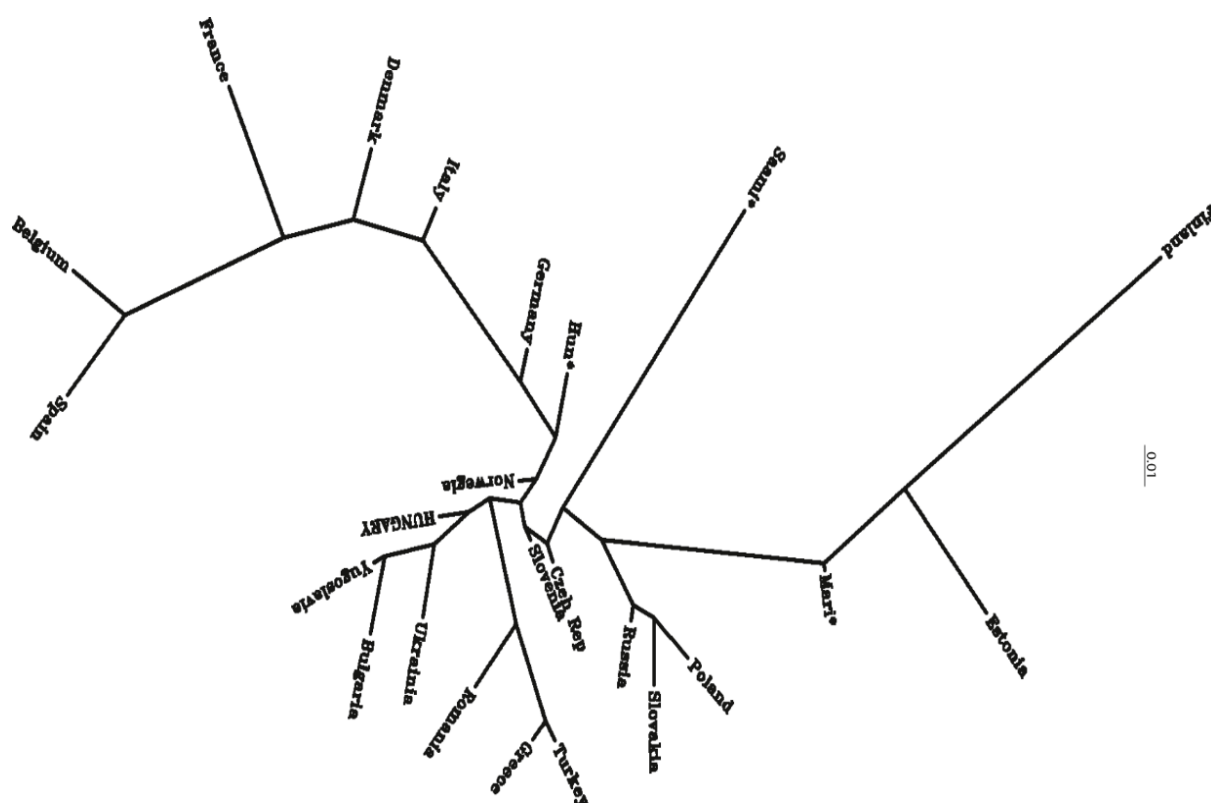
Magyarországi területi lefedettséggel összesen 215 (később 230-ra bővítve), nem rokon magyar férfi vizsgálatát végeztük el. Az Y-STR haplotípusok 77,4%-a, azaz 178 volt egyedi, a leggyakoribb haplotípus 9-szer fordult elő. A populáció haplotípus, illetve haplocsoport diverzitása 0,99686 és 0,86348 volt. A vizsgált minták összesen 18-féle haplocsoportba tartoztak, a leggyakoribb haplocsoportok az R1a1 (24,8%), R1b1 (18,7%), I2a (16,1%) és I1 (7,4%) (**31. ábra**). Ezeket a mintákat igazságügyi felhasználás céljából gyűjtöttük, ezért nem szelektáltuk őket etnikumokra nézve. Látható, hogy az indiai eredetű H1a haplocsoport előfordulása 5%, ami a roma kisebbség hazai jelenlétének tudható be.



31. ábra A magyar populáció Y-haplocsoport eloszlása

A haplocsoport eloszlását tekintve a legnagyobb 2 szelet az **R1a** és **R1b**, amelyek indoeurópai migrációs markernek tekinthetők [74], a magyar populáció 44%-át teszi ki. Az I haplocsoport az európai őslakosságra jellemző, utódjai a magyar lakosság 27%-át teszik ki. A J, G és E haplocsoport a neolitikumban, Mezopotámiából érkezett Európába, a J és G jelenlegi tudásunk szerint a Közel-Keleten (a G2a lehetséges, hogy a Kaukázusban), az E pedig Kelet-Afrikában keletkezett (a Közel-Keletről Afrikába visszavándorló ősapák között). Ezen ősapák leszármazottjai a recens magyar populáció 19%-át teszik ki [140].

Az Y-SNP adatainkat 23 másik európai populáció adataival hasonlítottuk össze (32. ábra). Az Fst genetikai távolságokat filogenetikai fa formájában ábrázoltuk.

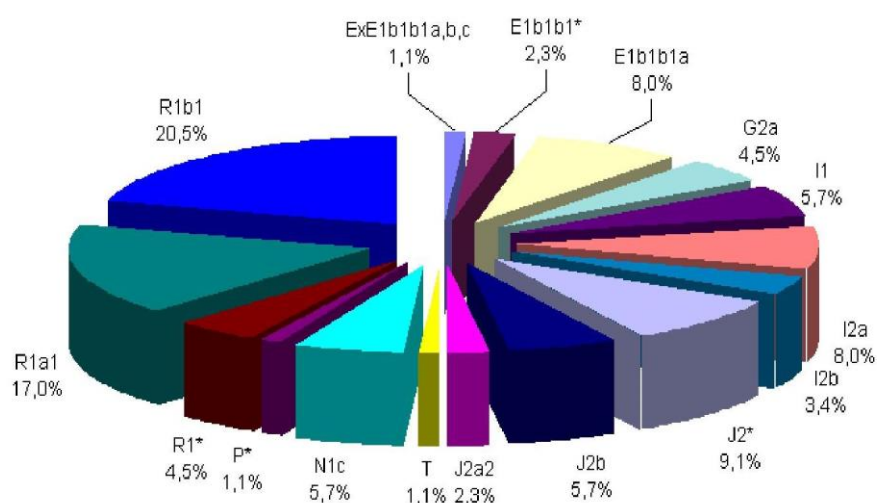


32. ábra. Gyökér nélküli filogenetikai fa

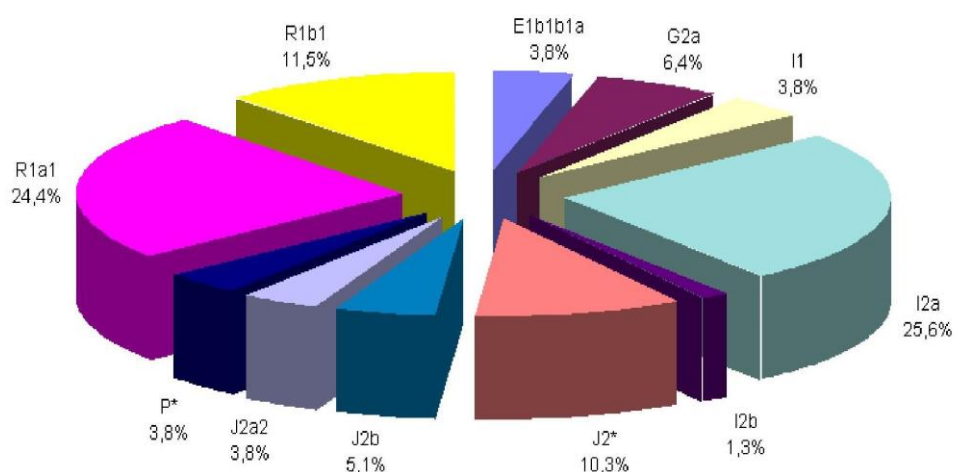
A filogenetikai fa klaszterei tükrözik a populációk történelmi és földrajzi hátterét, a magyar populáció a kelet-európai klaszterhez tartozik, néhány balkáni országgal együtt, míg az észak-európai országok populációi egy másik klaszterbe (Spanyolország, Belgium, Franciaország, Dánia, Olaszország és Németország) tartoznak, amely azok történelmi múltját tükrözheti. A finnek, észtek, marik és számik (lapp) a többiektől elkülönülnek. Az általunk vizsgált, valamint a publikált magyar populációs minták [141] a filogenetikai fán egymástól

távol helyezkednek el, ami valószínűleg a mintaszám ($n=36$) és a mintavétel különbsége miatt következett be. A legnagyobb genetikai távolság a finnek és magyarok között észlelhető, ami nem támasztja alá a nyelvi kapcsolatot. A recens-kárpát medencei magyar populáció genetikai összetétele apai ágon a környező népeiségek genetikai összetételétől nem igen különbözik.

Volt BSZKI-s kollégáknak (jelenleg NSZKK) köszönhetően kaptunk csíkszeredai székely és gyimesi csángó DNS mintákat. A két magyarul beszélő népcsoport Y-STR haplotípus vizsgálatát a volt BSZKI-s kollégák, Y-SNP haplocsoport vizsgálatát pedig mi végeztük. Mindkét populáció haplocsoport eloszlásai (33. és 34. ábra) közel azonosak a kárpát-medencei magyarokéval, viszont a székelyek között az N1c-Tat haplocsoport közel 6%-ban, a csángók között pedig a Q haplocsoport (P^*) kb. 4%-ban megjelent [142]. Egy hazai regionális populációban (bodrogi) vizsgálati eredményünk viszont az Nc1-Tat haplocsoport 6%-os jelenlétét mutatta [143]. Ebből az következik, hogy egyes zártabb régiókban az előbbi– kelet-ázsiai és dél-szibériai eredetű – haplocsoportok (N1c, Q) a vártnál nagyobb arányban is jelen vannak.



33. ábra. A csíkszeredai székely populáció haplocsoport eloszlása



34. ábra. A gyimesi csángó populáció haplocsoport eloszlása

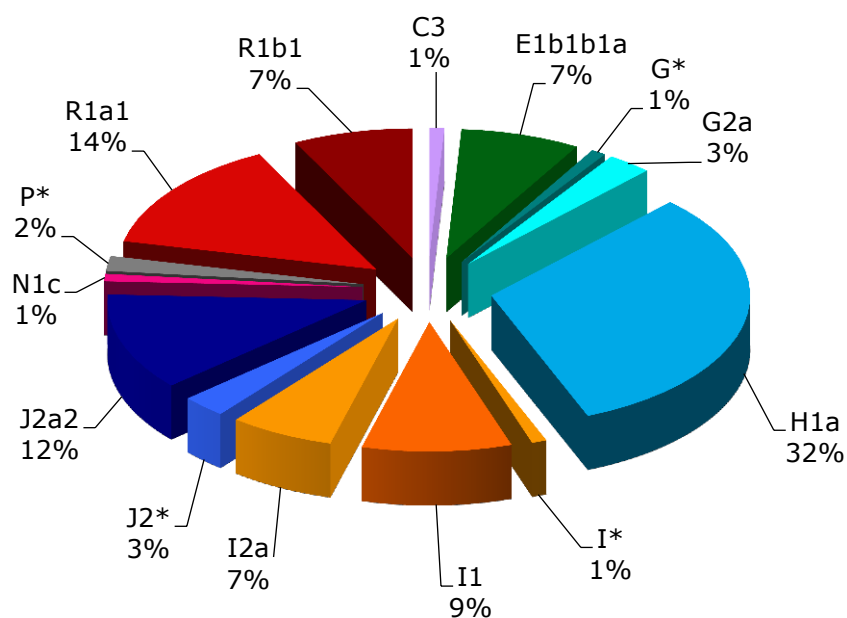
A fent említett 2 haplocsoport (**N1c**, **Q**) szinte teljesen hiányzott a kárpát-medencei magyarokból, ezenfelül a magyarországi magyarokból szintén hiányzó J2b haplocsoport 5-6%-ban jelen van mindkét magyar etnikumban [142]. A J2b Görögországban gyakoribb és Nyugat-Törökországban ritkább, ezért a dél-európai neolitikus kolonizáció során tengeri útvonalon érkezett markerként értelmezhető [144].

Összefoglalva megállapítható, hogy a kárpát-medencei magyar, székely és csángó populációk Y-kromoszómális haplocsoport eloszlásai közel azonosak egymással, valamint a közép- és kelet-európai országok populációiével, viszont elszigeteltebb populációkban (székely és csángó) jobban megőrződtek a dél-szibériai eredetű apai tulajdonságok (N1c és Q haplocsoport).

A vizsgált magyar férfiak (230 személy) kb. 20% -a a neolitikumban 8-10 ezer éve a Termékeny Félhold területéről érkező földművelők (E, G, J haplocsoport), 27%-a európai őslakók (I haplocsoport), 46%-a bronzkori keleti bevándorlók (R1a, R1b), azaz indoeurópaiak, 6%-a roma kisebbség (H, R2) és maradék 1%-a ugor eredetű nép (N1c) leszármazottja.

A roma populáció tekintetében nem a teljes adatállománnyal (mindegyik publikált), hanem csak kisebb volumenű populációs adattal prezentáljuk a férfiak genetikai útját. Kezdetben összesen 107, de később több, nem rokon magyar roma populációs mintát

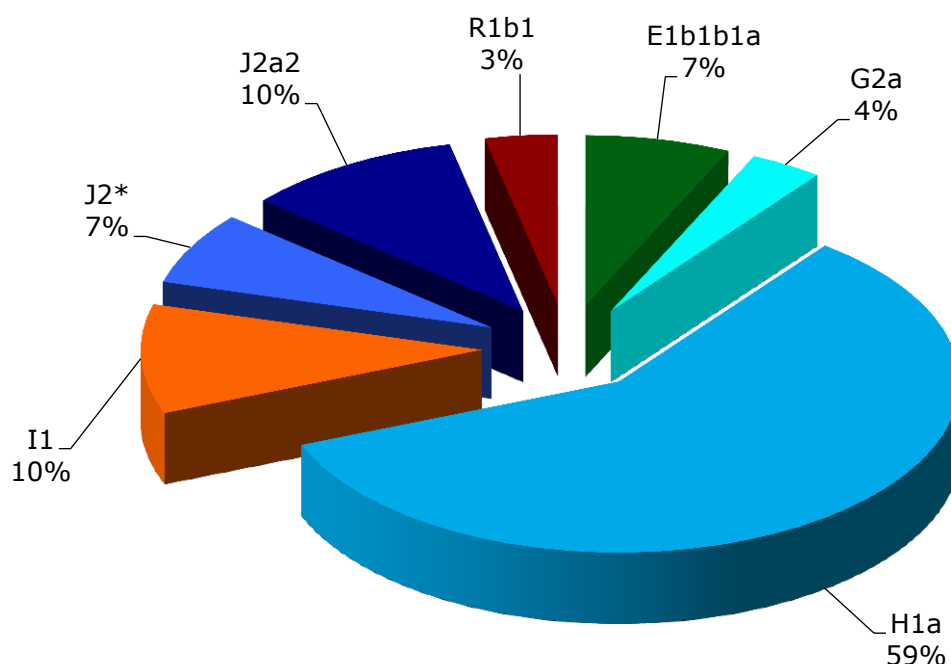
gyűjtöttük az ország egész területéről. A vizsgált személyek roma alpopulációkba tartozásáról (pl. oláh roma, romungro stb.) adatokkal nem rendelkezünk, így az eredmények a „kevert” magyarországi roma népességre vonatkoztathatók. Az Y-STR haplotípusok 36,4 %-a, azaz csupán 39 volt egyedi, a leggyakoribb 12-szer fordult elő. A populáció haplotípus-, illetve haplocsoport diverzitása 0,97337 és 0,84606. A vizsgált minták összesen 14-féle haplocsoportba tartoztak, leggyakoribb az **indiai eredetű H1a**, a vizsgált személyek mintegy harmada (31,8%) tartozik ebbe. Gyakori még az R1a1 (14,0%), J2a2 (12,1%), I1 (9,3%), R1b1 (7,5%) illetve E1b1b1a (7,5%) is (**35. ábra**). Az adatokból látható, hogy eléggé heterogén, nem pedig izolált populációról van szó, azonban az indiai őslakosokra jellemző H1a haplocsoport előfordulása a leggyakoribb (32%) [145].



35. ábra. A „kevert” hazai roma populáció haplocsoport eloszlása

Most nézzünk meg egy izoláltabb roma populációt, pl. a tiszavasvári roma férfiakat. A tiszavasvári roma populációból 29 nem rokon, oláh roma férfit vizsgáltunk. Az egyedi Y-STR haplotípusok aránya alacsony volt, mindössze 20,7%, azaz csupán 6 fő a 29-ből. A populáció több mint harmada, összesen 12 fő azonos haplotípuson osztozott. Ennek megfelelően a haplotípus (0,78876) és haplocsoport (0,62813) diverzitási értékek is alacsonyak (a legalacsonyabbak az összes vizsgált populációban). A populációban 7-féle haplocsoport volt megfigyelhető. Közülük az indiai eredetű H1a aránya igen magas, a tiszavasvári romák több, mint fele (58,6%) tartozik ide. A többi 6 haplocsoport ennél jóval alacsonyabb gyakorisággal (I1 – 10,3%, J2a2 – 10,3%, E1b1b1a – 6,9%, J2* – 6,9%, G2a – 3,4% és R1b1 – 3,4%) volt

megtalálható (**36. ábra**). Említést érdemel, hogy a Közép- és Kelet-Európában elterjedt R1a1 haplocsoport ebben a populációban nem volt megfigyelhető. Ezen leszármazási vonalak feldúsulása vagy eltűnése a nyilvánvalóan genetikai sodródásnak tudható be [146]. A bemutatott ábrákon nem láthatók, viszont a többi roma csoportokban elszórtan jelentkeztek az R2 és L alcsoportok is.

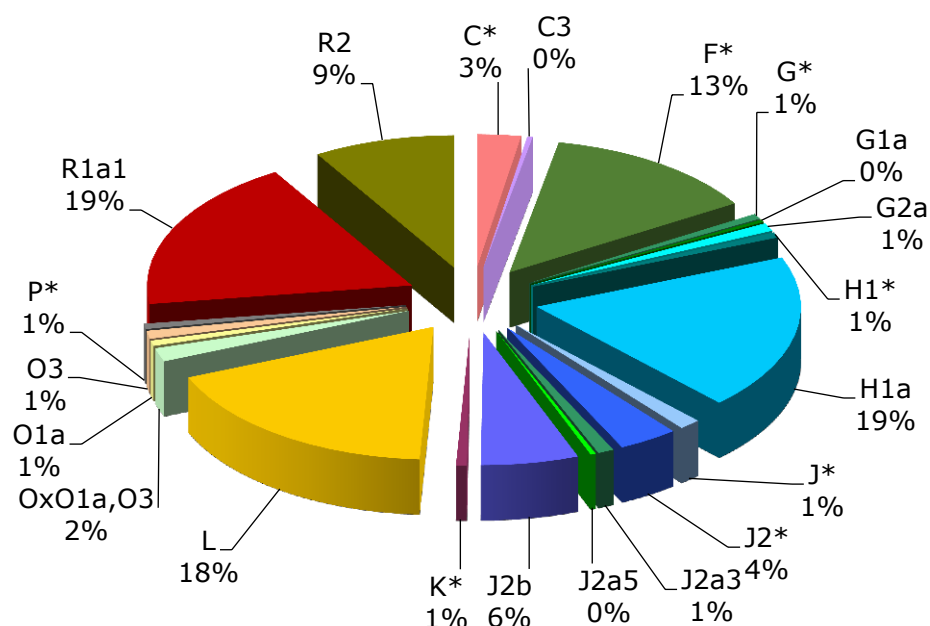


36. ábra. A tiszavasvári roma populáció haplocsoport eloszlása

A roma populációkkal való összehasonlítás céljából 301, nem rokon malajziai indiai férfi Y-SNP haplocsoportba sorolását végeztük el. A malajziai indiaiak Dél-Indiából érkező migránsok leszármazottjai a Malajzia brit gyarmatosításának idejéből. Az Y-STR haplotipizálás Malajziában történt [147]. Az Y-STR haplotípusok 84%-a (253 különböző haplotípus) egyedi volt, azaz csak egy személyben fordult elő, míg a leggyakoribb haplotípus 6-szor volt megfigyelhető az adott populációban. A haplotípus- illetve haplocsoport diverzitási érték 0,99898 és 0,86848.

A 301 Y-kromoszóma 21 különböző haplocsoportba tartozott. Leggyakoribb az indiai eredetű **H1a** (18,9%) illetve az Indiában is gyakori R1a1 (18,6%), a jellegzetesen Ázsiában előforduló **L** (17,9%) és **F** (13,3%), valamint a szintén Indiában gyakori **R2** (8,6%) (**37.**

ábra). Az összes többi – köztük a jellegzetesen ázsiai C és O – haplocsoport ennél kisebb gyakorisággal fordult elő [145].



37. ábra. A malajziai indiai populáció haplocsoport eloszlása

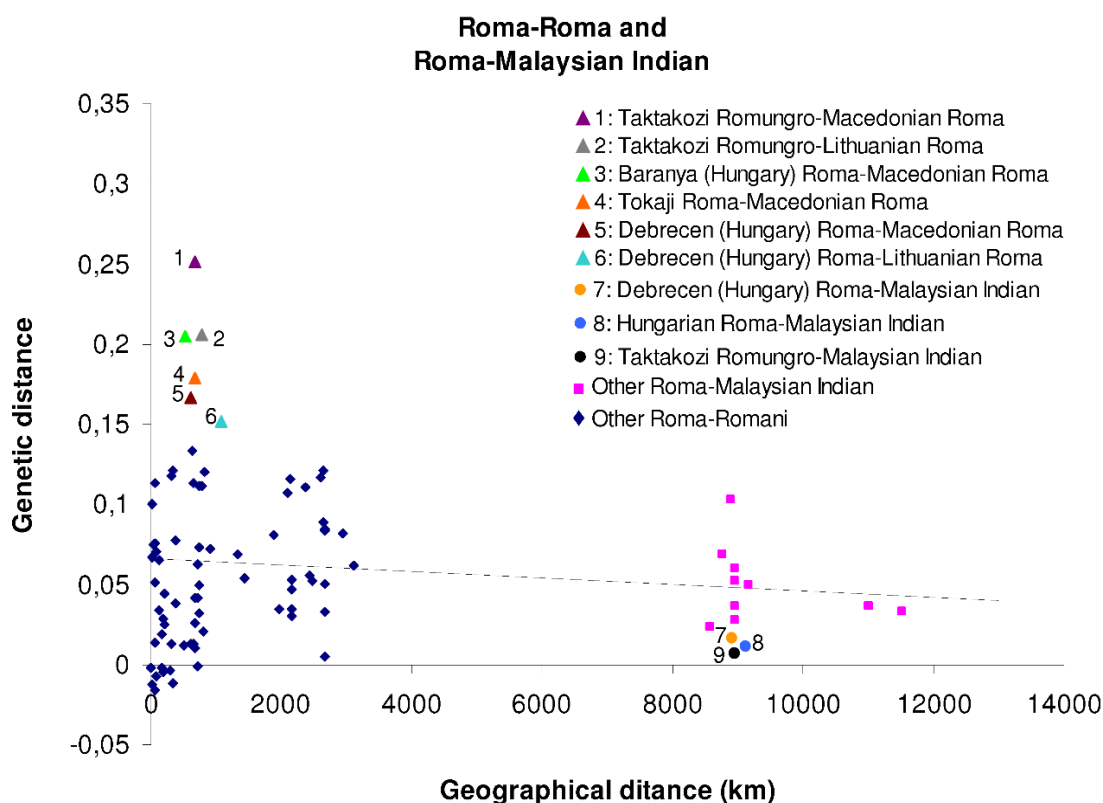
Összehasonlítva a roma populációkat egymással, a magyar és a malajziai indiai populációkkal az állapítható meg, hogy a roma férfiak az eredeti indiai génállományuk egy részét, a jelenleg csak indiai eredetű populációkban jelen lévő H1a, R2 és L haplocsoportokat, megőrizték. Génállományuk másik részét a Balkánon (a romák kb. 1000 éve érkeztek meg a Balkán-félszigetre) vagy a befogadó országok populációjától szerezték. Ezeket az átfogó vizsgálatokat elsőként végeztük el hazai szinten.

5.2.1.2. Filogeográfiai analízis roma populációkban

A vizsgált roma populációk haplotípus gyakoriságai (7 Y-STR lokusz: DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS92, DYS393) alapján AMOVA analízist végeztük Arlequin 2.0 szoftverrel [113]. A páronkénti genetikai távolságokat a következő 14 populációs mintára nézve számítottuk ki: baranyai roma [148], bolgár roma [149], debreceni roma [150], ibériai roma [151], litván roma [152], macedón roma [153], magyar roma, malajziai indiai [147], spanyol roma [152], 2 szlovákiai roma [154, 155], taktaközi romungro roma [145], tiszavasvári oláh roma [145] és tokaji oláh roma [145]. A filogeográfiai analízis

(38. ábra) keretében a földrajzi távolságok függvényében ábrázoltuk a vizsgált roma populációk haplotípus gyakoriságain alapuló Rst távolságokat.

Ez az ábrázolás is azt mutatja, hogy a roma populációs minták egy része (pl. taktaközi romungro, baranyai roma, tokaji roma, debreceni roma) genetikailag közelebb áll a tőlük földrajzilag távol elhelyezkedő malajziai indiai populációs mintához, mint a földrajzilag közelebb eső roma populációs minták egy részéhez.



38. ábra. Az Rst genetikai távolságok a földrajzi távolságok függvényében. Az ábrán látható, hogy a vizsgált roma populációk (7, 8, 9) Rst genetikai távolságok alapján 9000 km földrajzi távolság ellenére is közelebb állnak a malajziai indiai populációhoz; az európai roma populációk (1,2,3,4,5,6) pedig földrajzi közelségük ellenére is elkülönülnek egymástól az Rst genetikai távolságok alapján, utóbbi a genetikai sodródásnak tudható be. A sötétkék rombuszok az általunk vizsgált roma populációkat nem tartalmazó európai roma-roma összehasonlítások.

Az összehasonlításban a 41 populációs adatból minden roma mintacsoport összefüggésben állt a balkáni és anatóliai populációs mintákkal [145]. Ez a megfigyelés egybevág azzal a lehetőséggel, miszerint ezek a területek fontos földrajzi kapcsot képeztek a Közel-Kelet, Ázsia és Európa között a romák vándorlása során. A magyarországi referencia magyar mintacsoport szorosabb kapcsolatot mutat az indiai felsőbb kasztokkal, mint a balkáni

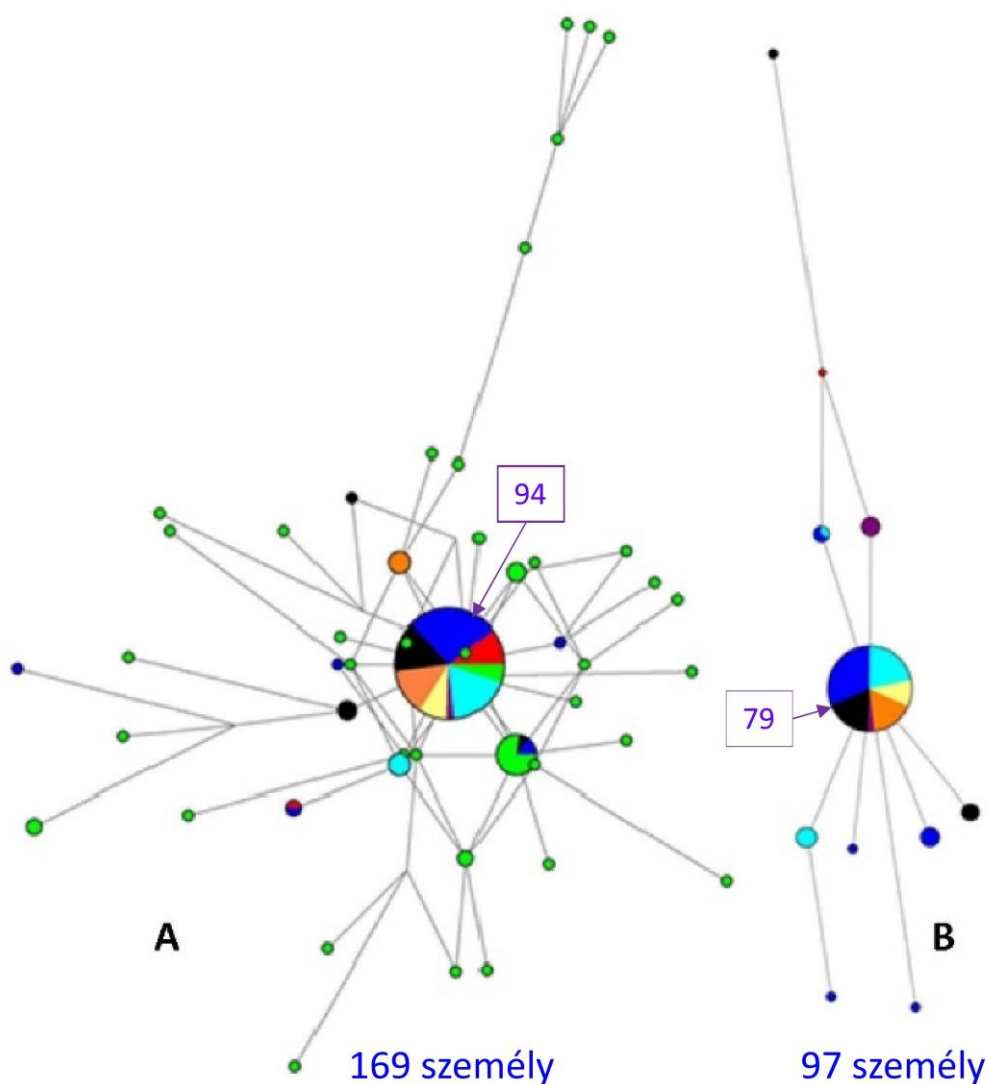
populációkkal, ami az R1a1-M198 haplocsoport magas gyakoriságával magyarázható mind a magyar mintacsoport, mind az indiai felső kasztok körében [140, 156]. Ez alátámasztja azt a feltételezést, hogy az R1a1-M198 haplocsoport közös ősrre vezethető vissza. A roma és indiai populációs csoportok közötti, meglehetősen alacsony genetikai távolság (R_{st}) szintén alátámasztja a feltételezett indiai eredetet.

A roma és nem-roma populációs minták között megfigyelt genetikai távolságok, tehát értékelhetők úgy, hogy közös eredetet, genetikai sodródást és génáramlást (gene flow) egyaránt tükröznek. Az utóbbi két folyamat növelhette a roma populációk közötti genetikai távolságot, míg a roma és nem-roma populációk közötti keveredés csökkenthette a genetikai távolságot.

5.2.1.3. Median Joining Network analízis

A Median Joining (MJ) Network vagy hálózat analízis lényege, hogy a vizsgált populációkban detektált apai leszármazási vonalakat, azaz minden egyes sikeres apai vonalat (N, R1a, R1b stb.) összevetjük más földrajzi régióban élő és/vagy más populációba tartozó, azonos ősapától származó személyekkel, megkeressük a közös haplotípusokat és kiszámítjuk a legközelebbi közös őst óta eltelt időt.

Az azonos haplocsoportba tartozó Y-kromoszómák közötti kapcsolat MJ hálózat formájában ábrázolható (39. ábra). A MJ network ábrázolást példaképpen az indiai eredetű H1a haplocsoporton mutatjuk be a vizsgált és publikált roma populációs adatokat bevonva. A MJ hálózatot 10 Y-STR (DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439) lokuszra vonatkozó haplotípusok alapján készítettük el a roma populációkban gyakori haplocsoportra, azaz H1a apai leszármazási vonalra. A DYS385 lokuszra vonatkozó adatokat, mivel ez a lokusz duplikálódott, nem használhattuk fel a MJ hálózatok készítéséhez. A körök mérete arányos az adott haplotípust hordozó Y-kromoszómák (férfiak) számával, míg a vonalak hossza az egyes haplotípusokat elválasztó mutációs lépések számával. Az egyes populációkat a következő színek jelölik: magyar-piros ●, magyar „kevert” roma - sötétkék ●, tiszavasvári oláh roma - narancssárga ●, tokaji oláh roma - sárga ●, taktaközi romungro roma - lila ●, szlovákiai romungro roma - fekete ●, ibériai roma - világoskék ●, malajziai indiai - zöld ●.



39. ábra. H1a-M82 kromoszómák MJ hálózata

A. A roma és a malajziai indiai populációk együttesen (a nagy színes kör teljesen azonos haplotípust jelent 94 vizsgált személyben)

B. Csak a roma populációk (a nagy színes kör a minden vizsgált roma populációban előforduló teljesen azonos haplotípust)

A volt ISZKI BOI DNS laboratóriumában vizsgált malajziai indiai, 5 különböző roma és az ibériai roma [151] populációs mintákból H1a-M82 haplocsoportba tartozó 169 db. Y-kromoszóma haplotípusa (10 Y-STR lokusz) alapján készült MJ hálózat látható a **39A. ábrán**. Az elkülönülő, központi haplotípus klaszterbe 94 db (55,6 %) H1a-M82 Y kromoszóma tartozik (**39A. ábra**). A csak a roma populációs mintákban fellelhető H1a-M82 Y-kromoszómákat ábrázolva (**39B. ábra**) csillag-szerű elrendezés látható, ahol a vizsgált személyek jelentős része szintén egy elkülönülő, központi haplotípus klaszterbe tartozik. Ez a

központi haplotípus klaszter 79 személyt tartalmaz, azaz a vizsgált H1a-M82 Y-kromoszómák 78%-át, a többi haplotípust egy vagy két mutációs lépés választja el a központi haplotípustól.

Kiszámítottuk a malajziai indiai populációs mintában a felhalmozódott STR variancia alapján a H1a-M82 haplocsoport korát (TMRCA=Time to Most Recent Common Ancestor), ami 8707 ± 1760 évnek adódott (6947 – 10476 év, 95% konfidencia intervallum mellett), ha feltételezzük, hogy az alapító (közös ős) a központi haplotípus volt (**39A. ábra**). Ugyanebben a haplocsoportban a csak a vizsgált roma populációs adatokra vonatkozó TMRCA 968 ± 336 év (**39B. ábra**). **A számolt korok összhangban vannak a indiai telepések indiai szubkontinensen paleolitikumban történt és a roma csoport Balkánon történt megjelenésének idejével [152, 157].**

Az indiai szubkontinens legkorábbi telepeseit tartják a mai törzsi csoportok őseinek [158]. Ezekre az őslakosokra többek között az F, H, C és O haplocsoportok a jellemzők [159]. Leggyakoribb a H*-M69 (H1a-M82 alcsoport) haplocsoport, ez a törzsi csoportokban 30%-os, az alsóbb kasztokban pedig 25%-os gyakoriságú [157, 160]. Ez a vonal Dél-Ázsián kívül igen ritkán fordul elő, bár egy alcsoportja, azaz a fent tárgyalt H1a-M82, feltűnően gyakori a roma népességben, ami alátámasztja a romák indiai eredetét.

A H1a-M82 Y kromoszómák MJ hálózata jól látható, hogy az összes vizsgált roma mintacsoportban megfigyelhető a közös Y-STR haplotípus (**39. ábra**), ami a H1a-M82 Y-kromoszómák több mint felét magában foglalja. Ez a klaszter egymással közeli rokonsági (leszármazási) viszonyban lévő Y-kromoszómák egy csoportját reprezentálja a roma populációkban, az ábrán látható minden egyes haplotípus egyazon központi haplotípus leszármazottja, függetlenül a mostani földrajzi helyzetétől. Ez is alátámasztja a közös indiai őstől való származást.

A fentiek alapján tehát megállapítható, hogy az európai roma férfiak 1000 éves európai jelenlétük ellenére még mindig hordozzák az indiai őseik genetikai nyomait apai ágon.

5.2.1.4. Az R1a haplocsoport keletkezése, terjedése és alcsoportjai

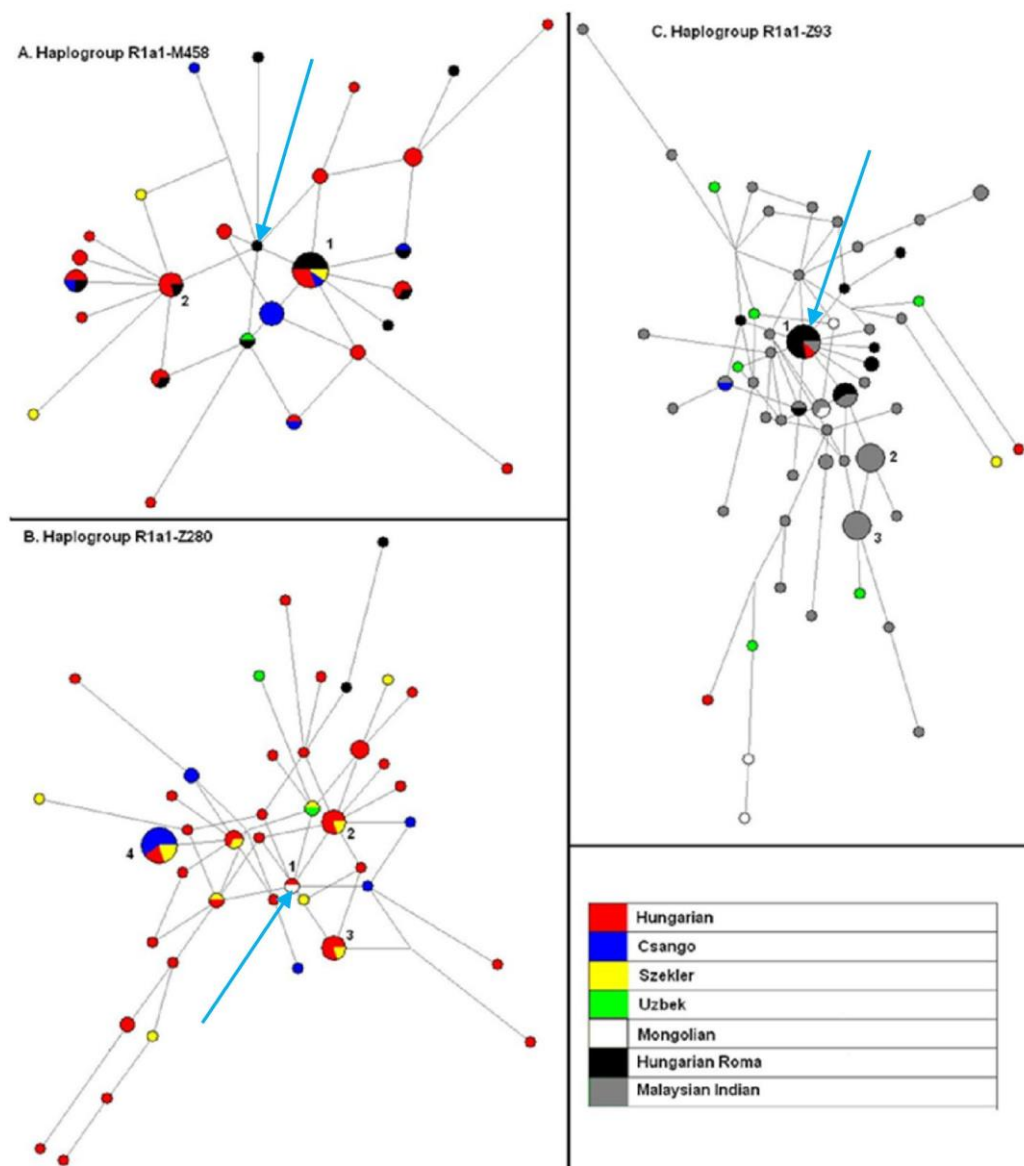
Az R főhaplocsoport jelenlegi tudásunk szerint Közép Ázsiában keletkezett kb. 19-28 ezer évvel ezelőtt. Az utódjaiból két fivér született, az R1 és R2. Az R1 fivér leszármazottjai egész Euráziát benépesítették, viszont az R2 fivér leszármazottjai manapság főleg Indiában élnek.

Az R1a-M198 haplocsoport az Y-kromoszómális haplocsoportok közül az egyik legjelentősebb. 2012-ig számos erőfeszítés történt a nagyszámú alcsoportjainak SNP-alapú azonosítására és azok migrációs útjának feltérképezésére. Az R1a kromoszómák őskori és későbbi történelmi eseményeinek mélyebb értelmezése az eurázsiai populációkban nem volt lehetséges az R1a haplocsoporton belüli új downstream SNP-k hiánya miatt. Kutatók azon vitatkoztak, hogy az R1a ősapja vajon hol születhetett. Az R1a-M198 ősapja születését a jelenlegi Ukrajna területén [65] lokalizálták és feltételezték, hogy onnan a *kurgán kultúrával* és az indoeurópai nyelvek terjedésével együtt terjedt el mind Európában, mind kelet felé [161]. Egyes kutatók azt sugallták, hogy az R1a haplocsoport indiai [156, 157] vagy közép-ázsiai eredetű [162]. 2010-ben jelent meg egy tanulmány [163], amelyben az R1a-M198 haplocsoporton belül sikerült két alcsoportot (R1a-M458, R1a-M434) azonosítani. Az M458 Kelet-Európában, az M434 kisebb gyakorisággal Pakisztánban fordult elő.

Kutató csoportunknak – 217 R1a mintát vizsgálva Európából és Ázsiából – nemzetközi szinten elsőként 2012-ben további 2 alcsoportot sikerült beazonosítania, nevezetesen az R1a-Z280 és R1a-Z93 alcsoportokat. Ez az egyik legjelentősebb tanulmány, amelynek sikerült különbséget tennie az ázsiai és európai R1a kromoszómák között [164]. A vizsgált minták R1a alcsoport eloszlását a **9. táblázat**, MJ hálózata a **40. ábra** mutatja.

9.táblázat. A vizsgált R1a minták alcsoport eloszlása

Populáció	(R1a)	M198*	M458	Z280	Z93	Színek a networkben
Magyar [287]	78	4	30	41	3	piros
Székely [88]	15	0	3	11	1	sárga
Csángó [106]	22	0	10	11	1	kék
Roma [256]	35	0	16	2	17	fekete
Malajziai indiai [301]	53	0	0	0	53	szürke
Üzbég [40]	9	0	1	2	6	zöld
Mongol [222]	5	0	0	1	4	fehér
Összes mintaszám: [N=1.300]	217	4	60	68	85	-



40. ábra. Median Joining network az R1a haplocsoporton belül

A körméretek arányosak a haplotípus frekvenciákkal, a legkisebb kör egy személynek felel meg. A hálózatba bevont személyek közös ősapjának haplotípusát kék nyíllal jelöltük.

A. 60 R1a1-M458 Y-kromoszóma haplotípus hálózata. **1. klaszter:** M458 modal (ősi) haplotípus, mint alapító haplotípus (3 magyar, 1 székely, 1 csángó és 5 roma). **2. klaszter:** kisebb M458 haplotípus klaszter (4 magyar és 1 roma).

B. 68 R1a1-Z280 Y-kromoszóma haplotípus hálózata. **1. klaszter:** Z280 medián haplotípus, mint alapító haplotípus (1 magyar és 1 mongol). **2. és 3. klaszter:** off-modal haplotípusok (4 magyar és 1 székely). **4. klaszter:** magyar nyelvűekre jellemző haplotípus (6 csángó, 2 magyar, 2 székely).

C. 85 R1a1-Z93 Y-kromoszóma haplotípus hálózata. **1. klaszter:** Z93 core/központi haplotípus, mint alapító haplotípus (7 roma, 1 magyar és 1 malajziai indiai). **2. és 3. klaszter:** a malajziai indiaiakra jellemző a haplotípusok.

Evolúciós mutációs rátával (6.9×10^{-4} /lokusz/25év) számolva [32] és a „modal” haplotípus klasztert véve alapító leszármazási vonalnak (**40A. ábra, 1. klaszter**) az elemzett populációk R1a-M458 vonalán felhalmozódott STR variációk kora 7306 ± 2321 év (95% CI, 4985 - 9627 év).

A vizsgálat legfontosabb megfigyelése az, hogy az új bináris markerek hatékony eszköznek mutatkoznak az R1a-M198 haplocsoporton belüli alcsoportba sorolására, mivel az elemzett R1a minták több mint 98%-a a három alcsoport egyikébe tartozik. Az R1a-M458 haplocsoport legmagasabb frekvenciája Kelet-Európára (Magyarországra és a romániai magyar etnikumokra) korlátozódik, továbbá az R1a-M458 gyakorlatilag hiányzik Ázsiából, ami egybeesik Underhill és mtsai, 2010 eredményével [163].

Az R1a-M458 hordozókra számolt 7,3 ezer évvel ezelőtti közös őstől való elválási idő összhangban van az Underhill és mtsai, 2010 által Európára számított korrall [163], ami 7,9 ezer év. A hivatkozott közlemény szerint az R1a-M458 időben társítható a neolitikus *lineáris kerámia kultúrával* (*Linear Pottery*) a mai Magyarország területén - ezt a megfigyelést az adataink is alátámasztják.

A **40B. ábrán** a legnépesebb **4. klaszter** az, amely érdekes lehet a magyarok etnikai eredetével kapcsolatban. Ebbe a klaszterbe a három magyar etnikumból összesen 10 haplotípus tartozik (6 csángó, 2 magyar, 2 székely), ami tükrözheti a magyarok érkezését a Kelet-európai-síkságról. Az R1a-Z280 haplotípus kora (TMRCA) 10283 ± 2574 év (95% CI, 7709-12857 év). Az R1a-Z280 haplotípusokra számolt diverzifikációs idő 10,3 ezer év (TMRCA), amely közel azonos az Underhill és mtsai, 2010 [163] által a kelet-európai R1a1a* (xM458) kromoszómákra számított idővel (11 ezer év).

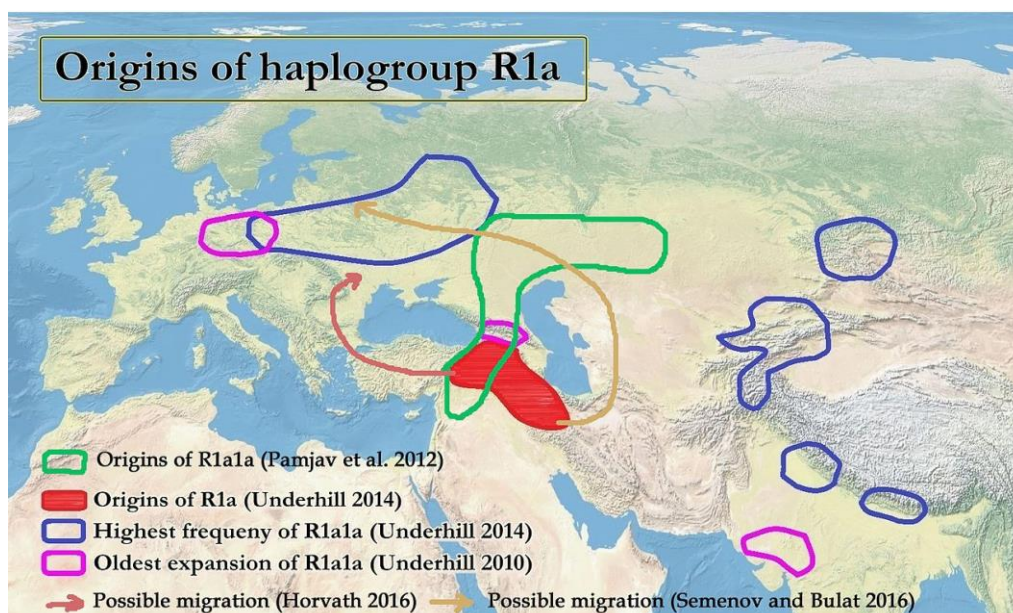
A 85 R1a-Z93 haplotípus Median Joining hálózata az **40C. ábra** mutatja. Az ábrának van egy csillagszerű struktúrája. Viszonylag nagyszámú egyén tartozik egy központi haplotípus-klaszterhez (**1. klaszter**: 9 személy), amelyen három populáció osztozik (7 roma, 1 magyar és 1 malajziai indiai); ez a leggyakoribb és egyben a roma csoport ún. „modal” haplotípusa. Az ábrán két másik nagy klaszter is kirajzolódik ki (**2. és 3. klaszter**), amelyek mindegyike hat malajziai indiait tartalmaz. Fontos megjegyezni, hogy a 3 magyar és 1 székely Z93 kromoszóma közül egyik sem képzett azonos klasztert roma vagy indiai Z93 haplotípusokkal, azt sugallva, hogy nem minden magyar Z93 kromoszóma származik romákkal való keveredésből. Az R1a-Z93 haplotípus becsült kora 10272 ± 2187 év (95% CI, 8085-12459 év). Tehát az általunk végzett vizsgálat szerint alacsonyabb (kb. 10,3 ezer év) az R1a-Z93 kromoszómák esetében a közös őstől való leválás óta eltelt idő, mint az indus-völgyi

populáció esetében az R1a1a * (xM458) kromoszómák alapján számított STR diverzifikációs idő (14 ezer év) [163].

Összefoglalva az a következtetés vontható le, hogy az R1a-M458 és az R1a-Z280 alcsoportok (ősapák) jellemzők a magyar nyelvű populációkra, míg az R1a-Z93 alcsoport a malajziai indiai és a magyar roma populációkra jellemző.

Belső- és Közép-Ázsia átfedési zóna lehetett az R1a-Z280 és az R1a-Z93 leszármazási vonalak tekintetében. Ezeknek az alcsoportoknak az eloszlási mintázata arra utal, hogy az R1a-M198 korai differenciálódási zónája feltehetően valahol az eurázsiai sztyeppéken, a Közel-Keleten vagy a Kaukázus térségében lehetett, mivel ezek a területek Dél-Ázsia és Kelet-Európa között fekszenek.

Később, 2014-ben Peter Underhill (Stanford Egyetem) irányításával 126 eurázsiai populációból 16244 férfi vizsgálatával, nemzetközi kooperációban átfogó tanulmányt készítettek, ami megerősítette eredményeinket, valamint további alcsoportokat írt le [92]. A kutatók az R1a ősapa születési helyét a Közel-Keleten, a mai Irán közelében lokalizálták, ahol a parahaplocsoportok (más szóval „vad” típusok) változatlanul megmaradtak ősi formájukban (pl. az R1a* magas előfordulását és a haplocsoport diverzitás magas szintjének lenyűgöző példáját találták). Valaki, aki eléggé kíváncsi volt az R1a ősapa születésének földrajzi lokalizációjára, készített egy összefoglaló rajzot, amelyet a **41. ábrán** mutatok be.



41. ábra. Az R1a haplocsoport (ősapa) születése¹⁸

¹⁸https://upload.wikimedia.org/wikipedia/commons/c/c4/R1a_origins_%28Underhill_2010%29_and_R1a1a_oldest_expansion_and_highest_frequency_%282014%29.jpg

Underhill és mtsai az R1a korát (TMRCA) az SNP alapú becsléssel (9,99 Mb szekvenálásával) kb. 5800 (95% CI: 4800–6800) évre kalkulálták [163].

5.2.1.5. Apai leszármazási vonalak összehasonlítása a magyar és a manysi népességben

A nyelvészek mind a hagyományos, mind a modern osztályozási próbálkozások alapján az obi-ugor nyelveket (manysi, hanti) tekintik a magyar nyelv legközelebbi rokonainak [165, 166]. Nemcsak a szókészlet, hanem az obi-ugor mondavilág szimbólumrendszere, a hőseik cselekményei is jelzik, hogy a hantik és a manysik (vagy elődeik) valaha lovas nép lehettek, akik a tajga övezetbe szorulásukkal térhettek át (vagy vissza) egyfajta tajgai gazdálkodásra, majd az orosz kolonizáció előre nyomulása miatt a vadászó életmódra. Nem elhanyagolható mennyiségű párhuzam található a manysi folklórkincs és a hazai népmesei motívumok között, például az égig érő fa, az ún. bűvármadár-motívum vagy a táltos paripa motívuma. Ez utóbbi kettő mintha ugyanannak kétféle megfogalmazása lenne: a sámán (táltos) segítőt kap és világok között közlekedik. Ezt már a korai folklorisztikától kezdve az egyik legősibb hiedelem-elemünknek tartják a kutatók, Róheim Gézától, Diószegi Vilmoson át Dömötör Tekláiig.¹⁹

A manysi nyelv egyike a legveszélyeztetettebb nyelveknek Euráziában. A magát manysinak vallók száma alig több mint tízezer. Azok száma, akik valóban rendszeresen használják is a manysi nyelvet a mindennapokban, még ennél is jóval kevesebb.

Kutatócsoportunk úgy érezte, talán az utolsó pillanatban vagyunk ahhoz, hogy az egyik legdinamikusabban fejlődő tudományág, a filogenetika, eszköztárával megvizsgálhassuk és feltérképezzük a manysik genetikai megoszlását, továbbá a mai magyarok és a manysik közötti lehetséges genetikai kapcsolatokat felvázoljuk.

2012-ben Vándor Anna kutató három hónapot töltött Nyugat-Szibériában és hatvan északi manysi férfitől vett mintát apai vonalú genetikai vizsgálathoz. A mintákat a volt ISZKI BOI DNS Laboratóriumában (jogutódja: NSZKK Genetikai Szakértői Intézet) vizsgáltuk; a manysi populációt finn, székely, magyar, csángó, mongol, burját és üzbég populációkkal

¹⁹ Íme egy folklórpélda: „Két öljükben vala hósínű hét ló. / Szétomlott ölhoz menj ki, kedvesem: / Valaha itt egy kanca csikót ellik vala, / Társai azt letapossák a föld alá, / Az omladék helyén áss le jó mélyen, / Megleled a csikót egy rőfnyi mélyen!» / A holt csikót is hazaviszi, / Nénje, Arany Kaltes kezébe kapta, /Egyik oldalát tóvízzel mosta, / Másik oldalát Ob-vízzel mosta: / Új életre hozta. / Amerre fordul, fény ragyog arra, / S mind a ketten / Felpattannak rája. / Két ég, két menny közzé / Szállnak, mint a szélvész.” (Részlet a „Szenttűzőzön legendája” c. manysi énekből)

használtuk össze. Az eredményeket Molecular Genetics and Genomics nemzetközi genetikai folyóiratban publikáltuk [167].

Az északi manysikra vonatkozó eredmények nem mutattak jelentős eltérést a korábban végzett kutatáshoz képest [168]. Alapos családtörténeti előkészítéssel azonban sikerült magyarázatot találni az európai és közel-keleti jellegű I és J haplocsoportok jelenlétére, amelyek a korábbi mintában [167] még csak szórványosan jelentek meg. A Vándor Anna által gyűjtött 60 fős mintában 86,7% volt az Észak-Euráziára jellemző N haplocsoport aránya, azon belül a többség az N-P43 (összesen 63,3%, korábban N1b-P43), a kisebbség az N-Tat (összesen 23,3%) alcsoportba tartozott. Minden vizsgált N-Tat (korábban N1c-Tat) haplocsoportú férfi L708 mutációval is rendelkezik. Más apai vonalú haplocsoport alacsony számmal fordult csak elő, így 6,7% volt a nyugat-európai és dél-kaukázusi jellegű R1b-M269 (esetünkben a kaukázusi kapcsolat valószínűsíthető), valamint 5% a fentebb jellemzett R1a-M198 aránya. Egyetlen minta a manysikkal szomszédos két népcsoportra jellemző, de például az amerikai őslakosságra (indiánokra) tipikusnak mondható Q-M242 csoportba tartozott (**10. táblázat**). A Q haplocsoport a Bering-szoros mindkét oldalán, az Újvilágban egészen Dél-Amerikáig markáns csoport (pl. a navahók 92,3%-a tartozik ebbe), de régészeti csontleletekből az Altáj-hegységből és Kínából is előkerült [95, 174].

Vizsgálataink eredménye szerint az N-Tat haplocsoporton belül elkülöníthető egy alcsoport, az L1034, amely a manysi és magyar/székely N-Tat minták között jelen van, míg az ugyancsak vizsgált burját, mongol és finn populációkból hiányzik [167].

Az N-L1034 alcsoportra egy genetikai alapú családfakutatással foglalkozó, amerikai DNS laboratórium, a FamilyTree DNA hívta fel a figyelmet. Az amerikai laboratóriumhoz beérkezett adatok alapján az N-L1034 alcsoport olyan területeken fordult elő elsősorban, ahol más tudományágak is számolnak ugor történelmi jelenléttel. Így a Kárpát-medencében, az Urál-vidéken, pontosabban Baskíriában és Tatárföldön, Görögországban és Kelet-Kazahsztánban. Fontos megemlíteni, hogy a legújabb régészeti kutatások szerint éppen Baskíria és Tatárföld területére esik Magna Hungaria [169]. Természetesen egyedi találatból nem szabad messze menő történelmi következtetéseket levonni, mégis említésre méltó, hogy az ugor migrációk esetleges kelet-kazahsztáni érintettségét a mi kutatásaink is alátámasztják. A szomszédos Kelet-Üzbegisztánból (Ferganai-medence) származó üzbég mintákból szintén volt egy L1034-es találat.

Az FTDNA szakemberei az N-L1034 alcsoportot a Kárpát-medencében a Csíki-medencében, a Felső-Tisza vidékén, a Sárreten (Bihar), 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 ugor kori 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, és nem asszimilálódtak, mint ahogyan a kunok vagy jászok. Számos kutató gondolta úgy idáig, hogy a székely egy a magyar törzsszövetséghez csatlakozott, eredetileg vélhetően török etnikumú népesség volt (ld. később). 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 szignifikánsan magasabb a modellünk szerint az ugor és finnugor etnogenezisben meghatározó szerepet játszó N-L1034 és R1a-Z280 haplocsoportok aránya az északi manysikhoz képest. Ez összhangban van az elvárásunkkal, amely szerint a nehezebben megközelíthető, *de facto* endogám területeken élők jobban megőrzik eredeti génállományukat.

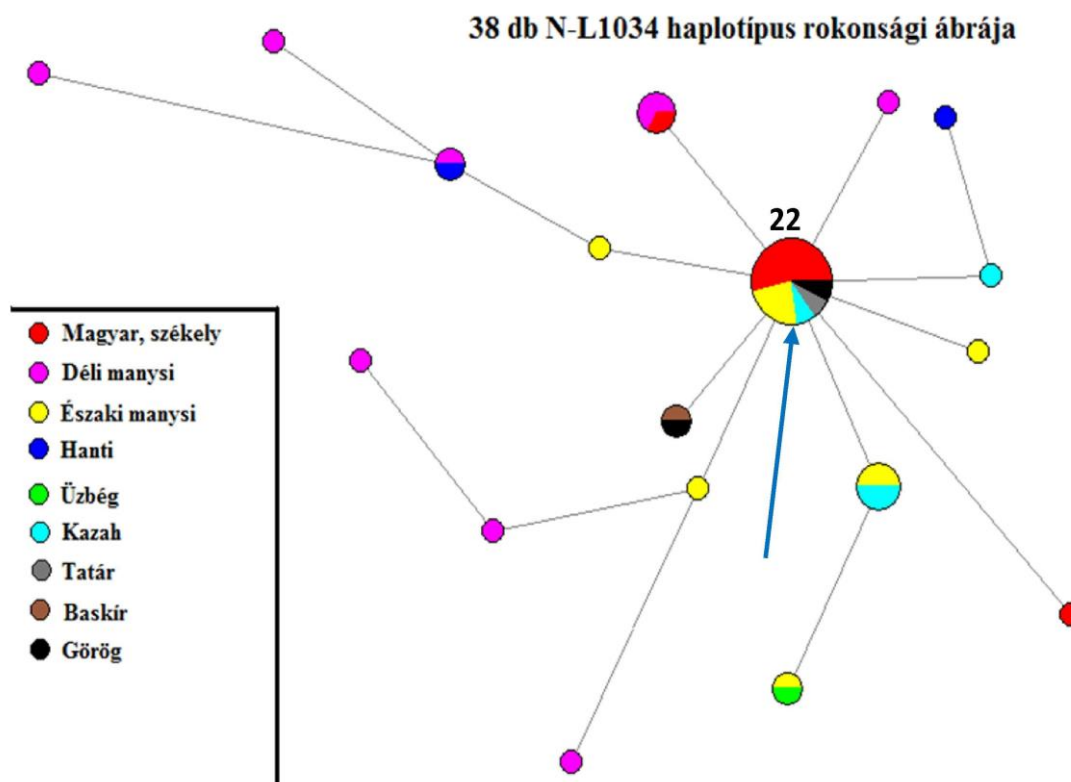
10. táblázat. Manysi populáció haplocsoport eloszlása

Haplocsoport	Északi manysi (n=60)	Déli manysi (n=36)	Feltételezett ²⁰ eredet
I1-L22	0,0%	2,8%	viking/orosz
I2-P37	0,0%	2,8%	szláv/orosz
N-P43 „A”	36,7%	25,0%	szamojéd
N-P43 „E”	26,7%	8,3%	szamojéd/permi?
N-L708*	8,3%	5,6%	más finnugor/szibériai
N-L1034	15,0%	27,8%	ugor
Q-M242	1,7%	0,0%	ket
R1a-M458	1,7%	5,6%	szláv/orosz
R1a-Z280	3,3%	19,4%	finnugor vagy szláv
R1b-P25*	6,7%	0%	keleti indoeurópai
R1b-P312	0,0%	2,8%	nyugati indoeurópai

A kondai és északi manysik között hat különböző R1a-Z280 haplotípus azonosítható, amelyből ötnek volt kárpát-medencei magyar egyezése, ami az R1a-Z280 haplocsoport fontosságára utalhat a magyar és ugor etnogenezisben. A magyarországi magyar adatbázisból az öt haplotípushoz tizenkét egyén tartozik, ami a minták 3,6%-át jelenti

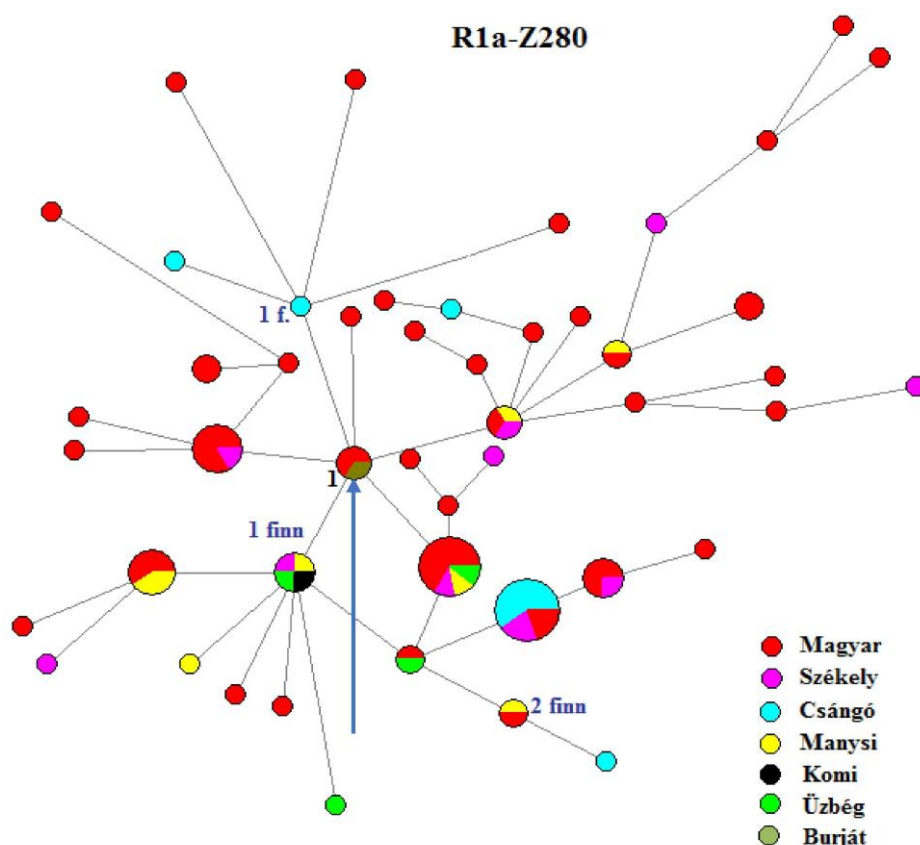
²⁰ A feltételezés természetesen az adott népek génállományára vonatkozó hatalmas szakirodalom alapján is csak egy lehetőség, hiszen a kapcsolat létrejöhetett más, például közvetett módon is. Az összehasonlításban is jól látszik, hogy nyelvek, nyelvcsaládok és mára eltűnt középkori népnevek is találhatók benne, ami nem a szerzők zavaros forráskezelése miatt van így, hanem a populációgenetikai szakirodalomban általánosan használt populáció-elnevezések kategóriái, amelyeket összefüggésbe lehet hozni az északi manysi és déli manysi haplocsoportokkal.

Az alábbiakban ismertetjük az összes vizsgált N1c-L1034 haplotípust tartalmazó MJ hálózatot (**42. ábra**), illetve az összes vizsgált finnugor és közép-ázsiai R1a-Z280 haplotípus rokonsági viszonyait mutató ábrát (**43. ábra**).



42. ábra. N-L1034 network analízis

A körméretek arányosak a haplotípus frekvenciákkal, a legkisebb kör egy személynek felel meg. A színskála a populációkat jelöli. A kék nyíl a közös ősapa haplotípusát jelenti. A nagy színes körben az 5 populációból 22 személy osztozik azonos haplotípuson.



43. ábra. R1a-Z280 network analízis

A körméretek arányosak a haplotípus frekvenciákkal, a legkisebb kör egy személynek felel meg. A színskála a populciókat jelöli. A kék nyíl a közös ősapa haplotípusát jelenti.

A **42. és 43. ábrák** alapján a következő észrevételek tehetők:

1. A magyar-manysi haplotípus egyezések, ha eltekintünk az N-L1034 alapító haplotípusában található egyezésektől, meglepő földrajzi eloszlásról tanúskodnak. Egyenletes eloszlás helyett a manysi egyezések 80%-a kondai manysik közül való, és azon belül is az egyező minták nagy része egyetlen régióból, Mezsdurecsenszkij körzetéből származik (**42. ábra**).
2. Ezen felül úgy tűnik, hogy a kondai manysik nem egyszerűen nagyobb arányban őrizték meg az eredeti ugor génállományt, de a diverzitás is magasabb a körükben, mint az északi manysik körében. Érdekes, hogy az alapító N-L1034 haplotípus viszont nem található meg a kondai manysik között.
3. A két kazah alapító N-L1034 haplotípus Kelet-Kazahsztánból származik.
4. A **43. ábrán** nem a részletek fontosak, hanem az egy-egy klaszteren (körön) belüli a színek együttes előfordulása, amely azonos apai eredetre utal (azonos haplotípusok).

A sárga szín manysi személyeket, a piros pedig kárpát-medencei magyarokat jelöl (lila: székely; kék: csángó). Az ábrán finn populációs adat nem szerepel, de az „1 finn” illetve „2 finn” felirattal jelölt haplotípusok a finn populációban is megtalálhatóak. Látható, hogy az R1a-Z280 csoporton belül van magyar-manysi egyezés (a sárga-piros kör haplotípusa ráadásul még két finn férfival is megegyezik, az ábrán „2 finn” felirattal jelölve), és van egy székely-manysi-komi-üzbég egyezés (lila-sárga-fekete-zöld kör, amely haplotípusa még egy finn férfival is azonos, az ábrán „1 finn” felirattal jelölve). Ezek alátámasztják, de archeogenetikai adatok hiányában nem bizonyítják az R1a-Z280 jelenlétét a korai finnugor nyelvű népesség génállományában.

Ha az ábrák alapján összevetjük a haplotípus egyezéseket, akkor több zavarba ejtő észrevételt fogalmazhatunk meg.

5. Magyar szemszögből nézve az R1a-Z280 haplocsoportban egyértelműen számos magyar-manysi haplotípus egyezés található, akárcsak az N1c-L1034 alcsoportban. Ezen felül az obi-ugor találattal rendelkező magyar R1a-Z280 haplotípusok halmaza sokkal diverzebb, mint a teljes magyar nyelvű N1c-L1034 haplotípusok halmaza. De meg kell jegyezni, hogy az R1a-Z280 minták száma jóval több, mint az N-L1034 mintáké.

6. A magyar-üzbég R1a-Z280 egyezések (a **43. ábrán** piros-zöld és piros-zöld-sárga-lila körök) Kelet-Üzbegisztánban, Taskentben és a Fergánai-medencében található. Az egyetlen üzbég N-L1034 (a **42. ábrán** sárga-zöld) szintén a Fergánai-medencéből származik [167]. Az N-M231 haplocsoport kb. 19-22 ezer éve jelent meg a mai Kína területén (Yunnan tartomány) az utolsó jégkorszak végén. A leszármazottjai elég hosszú utat tettek meg az elmúlt évezredekben Dél-Kínától Észak-Eurázia legnyugatibb csücskéig, Finnorszáig és Svédorszáig. Az N1c-Tat (N-Tat) haplocsoport valószínűleg a Bajkál-tó környékén keletkezhetett kb. 12 ezer éve [170].

Összefoglalva elmondható, hogy a mai magyar populációban nagyon alacsony az N-Tat haplocsoport aránya a többi finnugor nyelvet beszélő népességhez képest. Az N-Tat haplocsoport az egyik alcsoportja, az N-L1034 összeköti a magyar és manysi népséget, és potenciálisan elválasztja őket más népektől.

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.

Magyarországon nem sok regionális bontású genetikai felmérés történt. Más országokban készült vizsgálatok tapasztalatai alapján nem elhanyagolható különbségek lehetnek a különböző régiók között egy országon belül is. Mindenesetre FTDNA projektekből úgy tűnik – egyelőre alacsony mintaszám mellett – hogy a Bihar megyei Sárréten és az Őrség burgenlandi területein is 6-7% körüli az N-Tat haplocsoport aránya. Tehát elképzelhető, hogy a kutatók által nem vizsgált régiókban, a székelyhez hasonlóan, magas lehet az N-Tat haplocsoport aránya. Azt is érdemes figyelembe venni, hogy a magyarországi lakosság körében 1%-os gyakoriságú N-Tat csoport összesen nagyságrendileg 50 ezer férfit jelenthet, ami meghaladja a hanti és manysi lakosság teljes számát. Összességében tehát abszolút számban (lélekszámban) akár több lehet a magyar és székely N-L1034 férfi, mint a manysi N-L1034.

Úgy tűnik, hogy a recens magyar népesség génállományában az R1a-Z280 alcsoportban több ugor kori elem található, mint az N-Tat haplocsoportban. A jelenség magyarázatra szorul, de archeogenetikai adatok hiányában nem állítható több a jelenség felismerésénél.

Ha a további vizsgálatok is megerősítik az N-L1034 alcsoport és az ugor népek közötti szoros kapcsolatot, akkor az egy potenciális ugor migrációs markernek tekinthető. E marker földrajzi elterjedésének feltérképezése komoly támogatást jelenthetne a későbbiekben a magyar őstörténeti kutatásoknak. A rendelkezésre álló adatbázis alapján például felmerül a Kirgizisztántól nyugatra lévő Ferganai-medence mai népességének az ugor népekkel való kapcsolata, mert ott is megtalálható a rendkívül ritka N-L1034 alcsoport és az R1a-Z280 alcsoporton belül a magyar és a manysi populációkban megegyező haplotípus.

Míg az R1a-Z280 hatalmas területen elterjedt, a mai magyarokkal szomszédos népekben is jelentős az aránya, addig az N-L1034 tudományos jelentősége rendkívüli ritkaságában rejlik. Vélhetően az R1a-Z280 a magyar etnogenezisben játszott nagyobb szerepe ellenére is kevesebb tudományos bizonyító erővel rendelkezik a magyar vándorlás feltérképezésében, mint az N-L1034, amely a Kárpát-medencében etnospecifikusnak tűnik. Az N-L1034 valószínűleg meggyőzőbb érveket szolgáltathat a tekintetben is, hogy megtalálhatók és lokalizálhatók-e az Urál-vidéken Julianus barát által részben megtalált magyarok genetikai nyomai, esetleg a magyarok korábbi vándorlási útvonalai, állomásai.

A székelyek és az északi manysik körében is 60% fölötti az N-L1034 alcsoport aránya az N-Tat haplocsoporton belül. Ha feltételezzük, hogy az ugor etnogenezisben az N-Tat haplocsoport meghatározó szerepet játszott, akkor az is valószínű, hogy a korai ugor

közösségekben az N-L1034 is meghatározó szerepet játszhatott. Mivel az N-L1034 alcsoport relatíve fiatal alcsoport (az YFull 3900-4100 évesre teszi, ami megfelel a finnugor közösség nyelvészek által feltételezett felbomlási idejének), feltételezhető, hogy az ugor nyelvi kapcsolatok kialakulásért felelős társadalmi folyamatok egy jól körülhatárolható területen és viszonylag kis létszámú közösségekben zajlottak le.

5.2.1.6. A székelyek eredetével kapcsolatos új nézőpontok

A székelyek eredete élénken foglalkoztatja a közvéleményt és a magyar őstörténet kutatóit. A székelyek feltételezett eredetéről számos színvonalas tanulmány jelent meg, a kérdés mégis a mai napig megválaszolatlan. A kutatók évtizedek óta közel azonos hipotetikus modellek egyike vagy másika mellett foglalnak állást a székelyek eredetével kapcsolatban. Valóban nem túlzás azt mondani, hogy – áttörést eredményező adatok hiányában – „a székelyek származásának kutatása jó ideje helyben jár, a történeti források értelmezési lehetőségei kimerültek” [171].

A székelyek Y haplocsoport eloszlása is igen változatos [142], főleg a dél-szibériai eredetű haplocsoportok (N1c, Q) különböztetik meg őket a környező népektől. A székelyek esetében feltérképeztünk két kelet-eurázsiai eredetű haplocsoportot, az N-L1034 alcsoportot az N-Tat haplocsoporton belül, illetve a Q-M25 alcsoportot a Q-M242 haplocsoporton belül.

Az N-Tat és a Q-M242 haplocsoportok aránya ugyan meglehetősen alacsony a mai székely népességben, de különböző arányban (1-6%) kimutatható az összes eddig vizsgált, nagyobb lélekszámú székely mintában, így a korondi [172], nagyrészt csíkszeredai [142] és a bukovinai székely (Family Tree DNA) mintákban [36].

A Q haplocsoport a rendelkezésre álló adatok alapján valószínűleg kapcsolatba hozható az avarokkal és a hunokkal is [173]. Mivel a hunok és az avarok is felmerültek egyes szerzőknél a székelyek eredetével kapcsolatban, ezért a Q haplocsoport és ezen belül a székelyekben megtalálható Q-M25 alcsoport megérdemel egy részletesebb kitérőt. A szakirodalom az őskőkorszaki Dél-Szibériát tartja a Q-M242 csoport őshazájának (22-30 ezer éve), majd egyes csoportjai innen népesítették be Amerikát. Az Euráziában maradt Q három jelentősebb alcsoportra oszlik, amelyből a leggyakoribb az Amerikát is meghódító Q-M346-os alcsoport [174]. Ez Skandináviában is fellelhető, de a jenyiszeji ketek között domináns [175]. Ugyancsak jelentős a Q-M378-as alcsoport, amely az afgánok és hazarák, valamint az askenáziak között jellemző [174]. A legritkább a mi szempontunkból fontos Q-M25-ös alcsoport, amely a türkmének és a balkárok (az észak-kaukázusi török eredetű nép) mellett a

székelyeknél van jelen nagyobb mértékben [142, 174, 176]. A publikált N1c és Q haplocsoport adatokat a **11. táblázat** tartalmazza.

11. táblázat. Az N-Tat (első 3 oszlop), Q-M25 és Q346 mintákra publikált adatok

Nép (mintaszám)	N-Tat N-L1034 alcsoport	N-Tat (N-M2019 és N-M2119) „jakutokra jellemző” alcsoportok	N-Tat egyéb alcsoportok	Q-M25	Q-M346	Forrás
Magyar (332)	1	0	2	0	0	Bíró et al. 2015 (saját) [142]
Székely (95)	4	1	1	1	0	Bíró et al. 2015 [142] (saját)
Csángó (95)	0	0	0	3	0	Bíró et al. 2015 [142] (saját)
Manysi (60)	9	0	5	0	1	Fehér et al. 2015 [167] (saját)
Szerb (179)	0	1	0	0	0	Mirabal et al. 2010 [177]
Román (178)	0	0	1	0	0	Martinez-Cruz et al. 2012 [178]
Ukrán (154)	0	0	2	0	0	Mielnik-Sikorska et al. 2012 [179]
Horvát (457)	0	0	1	0	2	Barac et al. 2003 [180]
Balkár (135) kaukázusi nép	0	0	0	5	0	Balanovsky et al. 2011 [181]

A Family Tree DNA projekteken ezen haplocsoportok népcsoport szerinti bontásban az alábbi arányban vannak jelen (2015.10.14.). A mintaszám (zárójelben, dőlt számmal), az adott országból tesztelt összes személyt jelzi, így nem biztos, hogy etnikailag mind ahhoz a csoportoz tartoznak (**12. táblázat**).

12. táblázat. Az FTDNA projektben szereplő N1c és Q minták

Nép (db minta)	N-L1034	N-Tat „jakutokra jellemző” (N-M2019 és N-M2119)	N-Tat egyéb	Q-M25	Q-M346
Magyar* (296)	3	0	6	0	3
Székely (129)	1	1	1	6	0
Lengyel (3650)	0	0	>30	2	3
Német (3530)	0	0	12	1	1
Ukrán (395)	0	0	16	0	1
Orosz (850)	0	0	>40	1	2
Szlovák (95)	0	0	1	0	0
Cseh (627)	0	0	9	0	0
Baskír (62)	3	0	1	0	0
Csuvas** (18)	0	0	3	0	0
Bolgár (170)	0	0	0	0	0
Román (427)	0	0	2	0	0
Horvát (171)	0	1	0	0	1
Görög (213)	2	0	0	0	1
Balkár*** (122)	0	0	0	13	1

*kivéve a székelyek

** Csuvasok: a Volgától Szibériáig terjedő területen élő (türk) nép

***Balkárok: Észak Kaukázusban élő török eredetű nép

Érdekes, hogy az N-L1034 mellett még egy N-Tat alcsoport azonosítható, ha rendkívül alacsony arányban is, a székelyekben. A haplotípusa alapján „jakut” típusnak elnevezett alcsoport más magyaroknál nem, csak a székelyeknél fordul elő. A jakut N1c az M2019 vagy M2119 SNP-vel azonosítható, legfőbb jellegzetessége a DYS438=11 (a többi N1c-ben 10), valamint a DYS392=15 vagy 16 (14 helyett). Ezt elméletileg okozhatja adott esetben a viszonylag friss, például 17-18. századi bekerülés a népességbe (például a tatárok erdélyi betörései alkalmával), ami a székelyek izolált elhelyezkedése miatt egyelőre még nem eredményezte a kárpát-medencei magyarságban való nagyobb arányú elterjedést. Ehhez

azonban ki kellene mutatni e haplocsoport nagyobb arányú jelenlétét Krímben, az Észak-Kaukázusban vagy a délkelet-európai sztyeppén is, hiszen a krími tatárok ezekről a területekről származó harcosokkal csaptak le többször is Erdélyre. A jakutok a mongol terjeszkedés miatt hagyták el a Bajkál-tó mellékét, közöttük található meg legmagasabb arányban az N-Tat haplocsoport [182]. Az Y-kromoszómális adatok alapján a jakutok esetében kimutatható egy szűk közösségből (alapító) induló viszonylag késői népességrobbanás – palacknyak (bottleneck) effektus. Ez azt eredményezte, hogy az alapító N-Tat haplotípustól néhány stepwise mutációban eltérő speciális haplotípusok jellemzik a jakut Y-kromoszómális DNS állományt (genetikai sodródás). Ezeknek a haplotípusoknak, elsősorban a DYS392 lokuszban detektálható 15-ös vagy 16-os allél a közös jellemzője. Ráadásul a közép-európai térségben csak kevés ilyen minta ismert (1-1 szerb és horvát), de ezek jelenléte nem lehet meglepő, hiszen tudjuk, hogy a Balkán és Közép-Európa népei intenzív keveredésen mentek át az évszázadok során.

A Q haplocsoportból az Amerikát meghódító Q-M346 alcsoport ritkán, de előfordul Kelet-Európában; ezzel szemben az egész világon ritka Q-M25 alcsoportnak Q-L712 alcsoportja jellemzően csak néhány népnél fordul elő Kelet-Európában [183]. A székelyek között (a publikált és az FTDNA mintában is), valamint az észak-kaukázusi török nyelvű balkároknál jellemző, de előfordul elszórtan a lengyelek, illetve oroszországi népek körében is.

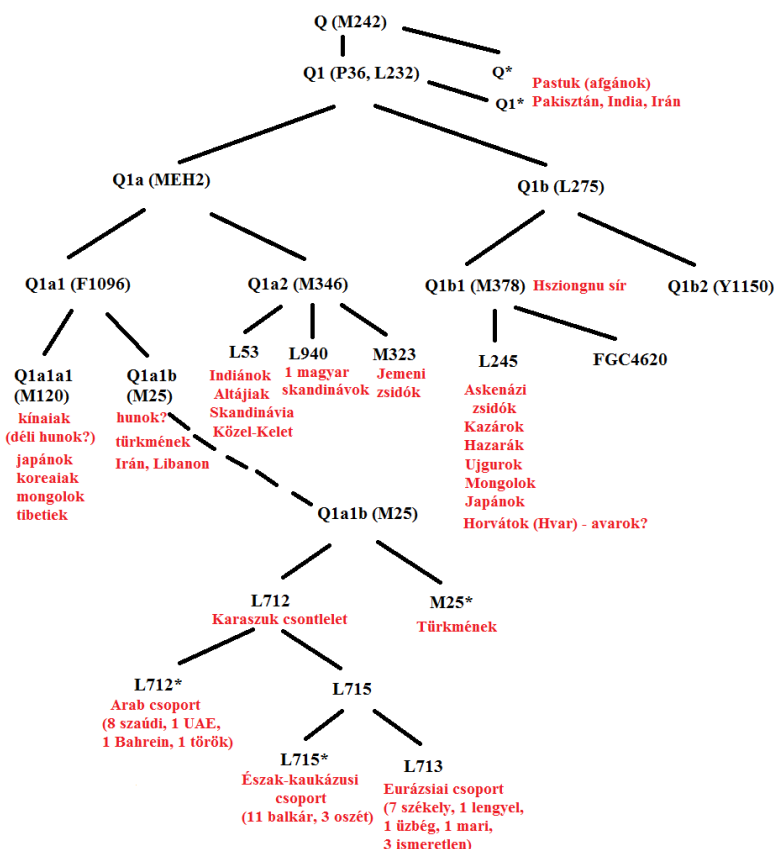
A székely alcsoport (L712+ L713+) közös őse 1700 évvel ezelőttire becsülhető pedigre mutációs rátával számolva [183].

Fontos kiemelni, hogy a Q-M25-nek az L712 alcsoportját megtalálták egy i.e. 1400-900 közé datálható, *Karaszuk kultúra* (ld. a **19. ábrát**) területén talált csontmintában Dél-Szibériában [74], tehát az L712-es alcsoport legalább 3500 éves (az YFull szerint 14200 éves). A *Karaszuk-kultúra* képviseli az *Okunevo* (i.e. 2500-2000; **19. ábra**), *Szintasztá* (i.e. 2100-1800; **19. ábra**) és az *Andronovo* (i.e. 1700-1500; **19. ábra**) kultúrák folytatását és átalakulását. Ennek a kultúrának a hatása az Altaj-Sayan régiótól az Aral-tó területén át Kínáig és Nyugat-Szibériáig vezethető vissza [74]. A Q csoport helyi elem lehet, miközben a karaszuk mintákban a *Szintasta-Andronovo-kultúrákkal* folyamatosságot mutató R1a-Z93 jelenléte is kimutatható, azaz a genetikai adatok összhangban vannak a feltételezett történelmi modellel. Másrészt a *Karaszuk-kultúra* kapcsán felmerül elvi lehetőségként, hogy eredhet-e végső soron a székely és balkár Q-L712-es alcsoport ebből vagy valamelyik a *Karaszuk-kultúrához* közvetett módon kapcsolódó kultúrából.

Meg kell említeni, hogy Barac és mtsai, 2003 [183] horvát mintájában Hvar és Korcula szigetén meglepően nagy arányban fordult elő a Q haplocsoport, melynek többsége azonban haplotípus alapján a Q-M378 alcsoportba tartozott, ez pedig a magyarokból és székelyekből is hiányzik. Az alcsoport jelenléte vagy az avarok 6. századi Dalmáciába érkezésével vagy az Észak-Kínából induló ősi selyem út vándoraival magyarázható [180]. Így életszerűnek tűnik az a szakirodalmi feltevés, hogy a Q-M378 alcsoportban az avarok nyomait találták meg a dalmát szigetvilágban. Hasonló logika mentén a dalmáciai és a székely jakut típusú N-Tat kapcsán is felmerülhetne az avar eredet lehetősége.

Csontleletekből kinyert adatok alapján úgy tűnik, hogy a Q haplocsoport jelen volt az ázsiai hunok génállományában [174]. Az elmúlt években kínai kutatók publikáltak a hunok ázsiai elődjének tekintett hsziangnu (Xiongnu) sírokból származó DNS mintákra vonatkozó adatokat. Az egyik legkiemelkedőbb a Hszincsiang Ujgur Autonóm Körzetben lévő Hami városához közeli „fekete Gouliang” temető, amely az ázsiai hun uralkodó nyári rezidenciájának helyén található. Itt az i.e. 2–1. században élt 12 férfiből mind a 12 a Q haplocsoporthoz tartozott, azon belül 6 fő Q-M346, 4 fő Q-M378 és 2 fő meghatározhatatlan Q alcsoportba tartozott [184]. Ez alapján a hun arisztokrácia az egyik csoportja a Q haplocsoportba tartozhatott.

Huang és mtsaival közös 2017. évi tanulmányunk [174] alapján mind a Q-M25, mind a Q-M346 alcsoport gyakori türk nyelvű populációkban, időbecslések szerint keletkezésük körülbelül 3-5 ezer évvel ezelőttre tehető. A Q-M25 alcsoport Közép-Ázsiából Nyugat-Ázsiába, majd onnan Közép-Európába, a Q-M346 alcsoport pedig Dél-Szibériából Eurázsia többi része és a Comore-szigetek felé terjedhetett [174]. A szegedi kollégákkal együtt publikált archeo mtDNS vizsgálat alapján feltételezhető, hogy a magyarok valószínűleg Közép-Ázsiából és Dél-Szibériából származhatnak körülbelül 4 ezer évvel ezelőttről [185], ez összhangban van a fenti tanulmányban szereplő időbecslésekkel is. A Q haplocsoport filogenetikai fáját a **44. ábra** mutatja.



44. ábra. A Q haplocsoport főbb alcsoportjai és mai elterjedésük
A szaggatott vonal a Q1a1b (M25) haplocsoport további felosztását jelöli.

A kelet-ázsiai haplocsoportokon túl két olyan, haplotípusok alapján azonosított, alcsoport található a székelyekben, amely a környező magyar és román mintavételekből hiányzik. A J2-M67 és a G2a-P15 haplocsoportok egyes alcsoportjai szintén a Kaukázusra mutatnak, mint végső genetikai forrásra. Ezek közül a G2a-P15 a leggyakoribb, ennek forrása a Kaukázus vidéke. A J2 haplocsoport hatalmas területen terjedt el, de az M67-es alcsoportja jelentős arányban kizárólag a Kaukázus térségben található meg [95, 104, 181].

Mindent egybevetve, a mai genetikai mintázat alapján a székelyek eredetével kapcsolatos kérdés nehezen megválaszolható. Bár a Q haplocsoport valószínűleg fontos szerepet játszott a belső-ázsiai hun és avar közösségekben, egyikben sem található meg a székelyekre jellemző Q-M25 alcsoport. **Azaz a székelyeknek a belső-ázsiai hunokhoz vagy avarokhoz való speciális viszonyát a jelenlegi genetikai adatok nem támasztják alá.**

5.2.1.7. Az Y-kromoszómális haplocsoportok összesített eloszlása a mai magyar népességben

A következő táblázatban összesítettük mind az általunk vizsgált, mind a FTDNA nyilvános adatbázisban szereplő magyar vezetéknévvel rendelkező személyek Y-SNP adatait, illusztrálva, hogy a különböző helyekről származó, különböző számú mintavételek mennyire befolyásolják a leszármazási vonalak eloszlását (**13. táblázat**).

13. táblázat. Összesített Y-haplocsoport adatok magyar nyelvű népességekben

Haplocsoport	Magyar-ország	Székely	Csángó	Összes publikált magyar	FTDNA Magyar	Összes ismert magyar	Neolit Kárpát-medence	A haplocsoportot legnagyobb mértékben hordozó nem magyar etnikai csoportok és térségek
Mintaszám	332	95	95	522	429	952	42	
C-F3393	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	11,9%	paleolit európai, ausztrál abo
C-M217	0,0%	0,0%	0,0%	0,0%	0,2%	0,1%	0,0%	mongolok, kazahok
E*(xM35)	0,0%	1,1%	0,0%	0,2%	0,0%	0,1%	0,0%	Afrika
E1b-M35	6,6%	10,5%	4,2%	6,9%	6,8%	6,8%	4,8%	Észak-Afrika, Balkán
F* & H2	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	26,2%	mára szinte eltűnt
G2-P15	5,7%	4,2%	6,3%	5,5%	7,0%	6,2%	35,7%	korai földművelő, Kaukázus
H-M82	3,3%	0,0%	0,0%	2,1%	0,0%	1,2%	0,0%	indiaiak, romák
I1-M253	8,1%	6,3%	3,2%	6,9%	8,9%	7,8%	2,4%	germánok, vikingek
I2-CTS10228 ²¹	13,2%	7,4%	28,4%	14,9%	14,2%	14,6%	0,0%	délszlávok, Balkán
I2 (xCTS10228) ²²	4,2%	3,2%	1,1%	3,4%	2,1%	2,8%	14,3%	szászok, kőkori népesség
J1-M267	2,7%	0,0%	0,0%	1,7%	0,5%	1,2%	0,0%	sémi népek, dagesztániak
J2a-M410 ²³	5,7%	13,7%	14,7%	8,8%	4,4%	6,8%	4,8%	Közel-Kelet, mediterránok
J2b-M12	0,6%	4,2%	4,2%	1,9%	2,8%	2,3%		Balkán, albánok, arománok
L-M20	0,0%	0,0%	0,0%	0,0%	0,2%	0,1%	0,0%	Indus-völgy, drúzok, Irán
N* (xP43, Tat)	0,0%	0,0%	0,0%	0,0%	0,5%	0,2%	0,0%	Kelet-Ázsia
N-Tat (xL1034)	0,6%	2,1%	0,0%	0,8%	1,6%	1,2%	0,0%	finnek, jakutok, litvánok
N-L1034	0,3%	4,2%	0,0%	1,0%	0,9%	0,9%	0,0%	obi-ugorok
N-P43	0,0%	0,0%	0,0%	0,0%	0,2%	0,1%	0,0%	szamojédek, obi-ugorok
Q-M242	0,3%	1,1%	3,2%	1,0%	2,1%	1,5%	0,0%	indiánok, hunok, türkmének
R1a-M198	28,5%	16,8%	22,1%	25,2%	27,3%	26,2%	0,0%	szlávok, finnugorok, türkök
R1b-Z2105 ²⁴	3,9%	5,3%	4,2%	4,2%	3,7%	4,0%	0,0%	örmények, gödörsíros kultúra
R1b-M412	15,6%	18,9%	8,4%	14,9%	15,9%	15,3%	0,0%	kelták, germánok, újlatinok
R2-M124	0,3%	0,0%	0,0%	0,2%	0,0%	0,1%	0,0%	dél-ázsiaiak, Irán
T-M70	0,3%	1,1%	0,0%	0,4%	0,7%	0,5%	0,0%	Közel-Kelet, mediterránok

A táblázat első oszlopában láthatók a haplocsoportok, ezt követően a volt Igazságügyi Szakértői és Kutató Intézetek (ISZKI) Budapesti Orvosszakértői Intézet (BOI) DNS Laboratóriuma (jogutódja: NSZKK Genetikai Szakértői Intézet) által vizsgált, 2014-ben

²¹ I2-P37-ként publikálva

²² I* és I2-M223 összege a publikált adatokban

²³ J2 (xJ2b-M12) publikált adatok összesítve

²⁴ R1b-P25*-ként publikálva

publikált²⁵ három, összesen 522 fős mintacsoport (magyarországi magyar, székely-magyar, csángó-magyar) eredményei, majd e három csoport összesített haplocsoport-megoszlása (ld. „összes publikált magyar” oszlop). A hatodik oszlopban pedig a Family Tree DNA magyar genetikai mintavételezési és elemzési projektben 2016. február 7-én megtalálható 429 fős, egymástól független, kizárólag magyar vezetéknévű minta²⁶ szerepel. A 7. oszlopban az összesített magyar adatok szerepelnek. A 8. oszlopban összehasonlításképpen feltüntettük 42 ismert i.e. 5800-4200 közötti időszakból Magyarország területéről származó újkőkori (neolit) földművelő minta adatait, melyek több publikációból származnak [81, 82, 83].

A két összefoglaló adatsort tartalmazó 5. és 6. oszlop összehasonlítása alapján látható, hogy a két egymástól független 400-500 fős mintán elvégzett kutatás a főbb haplocsoportok tekintetében hasonló, de természetesen nem teljesen azonos eredményt ad. Mindkét esetben egyértelműen az R1a-M198 csoport a leggyakoribb férfi leszármazási vonal (25-27%). Ezen belül elenyésző a törökségi és indoiráni nyelvű népekre jellemző R1a-Z93-as alcsoport aránya (a publikált mintában 1,3%, az FTDNA mintában 0,7%, a táblázatban nem jelölve), tehát a többség az ún. „európai” térségre jellemző alcsoportokba tartozik (R1a-Z283 és alcsoportjai, a táblázatban nem jelölve), amelyek elsősorban a szláv és finnugor nyelvű népességre jellemzők. A második leggyakoribb csoport 15-16%-os gyakorisággal a „nyugat-európai” indoeurópai nyelveket beszélőkre jellemző R1b-M412 és ennek alcsoportjai (az U106 főként a germán nyelveket beszélők között gyakori, a P312 pedig a kelták utódainak tartott írek, skótok, walesiek és az újlatin nyelveket beszélők között), ehhez az Eurázsia sztyeppei térségére jellemző, 4%-nyi „keleti” R1b-Z2105 (L23) alcsoportot hozzászámítva közel 20%-os az R1b-M269 csoport aránya. A harmadik leggyakoribb, 14-15%-os előfordulású csoport a „dinári-kárpáti” típusnak nevezett I2-CTS10228 (I2-P37-ként publikálva), amely a Balkántól Nyugat-Ukrajnáig gyakori [85, 86, 87, 88, 94, 95, 140, 142, 164]. Minden egyéb haplocsoport részesedése ennél alacsonyabb a magyar férfiak génállományában. A két fő oszlop összevetésével jól látható, hogy például az indiaiakra jellemző H-M82 és R2-M124 csoportok [145, 146] az etnikai identitást is figyelembe vevő 6. oszlopban eltűntek a mintából. Számottevő különbség még a két oszlop között a J2 csoportok esetében figyelhető meg, amely aránya az 5. oszlopban feltüntetett mintában a székelyek és csángók relatíve nagyobb súlya miatt kétszer annyi, mint az FTDNA mintában. Ez átvezet minket a reprezentativitás

²⁵ A korábbi mintavételek ma már nem megfelelő részletességű vizsgálata miatt egy újabb, viszonylag széles körű mintavételezéssel járó adatbázis alapul vételével (Bíró et al. 2014). A csoportok elnevezése a mintavétel körzeteire utal, azt kellett követnünk összefoglaló táblázatunkban is.

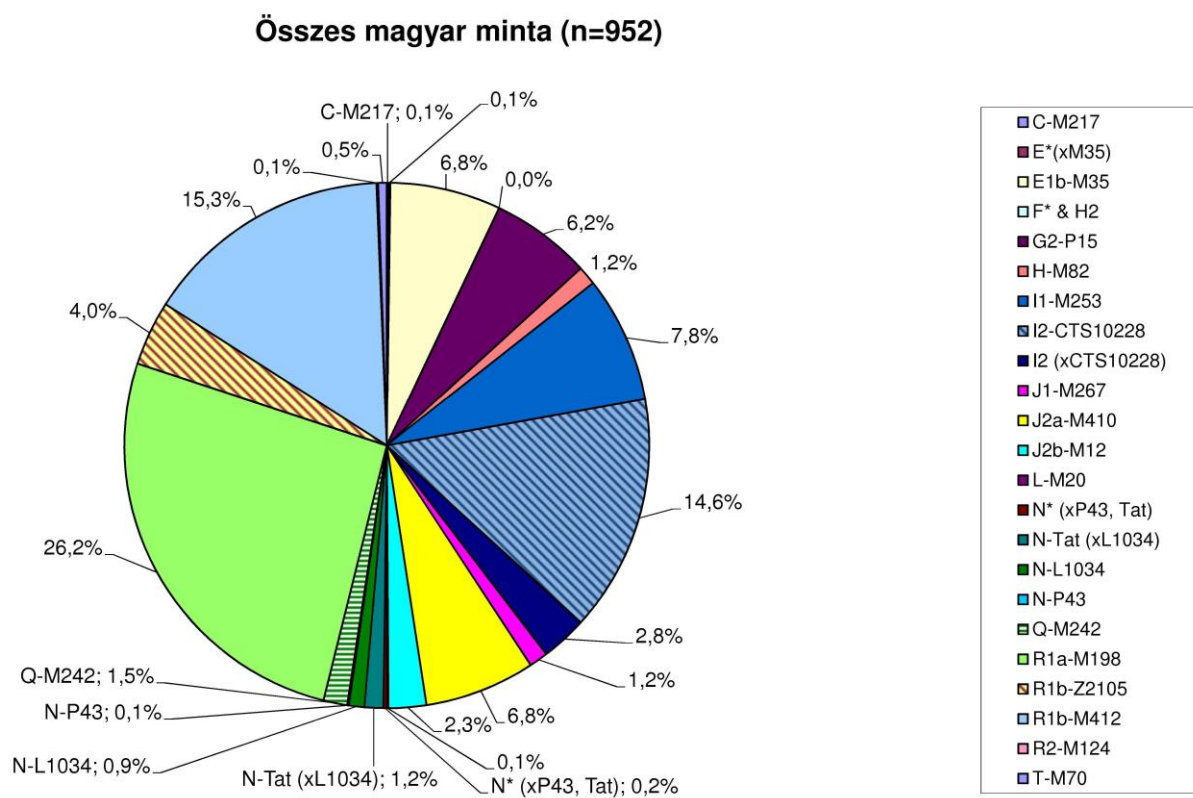
²⁶ Az általunk bemutatott ezen adatok a Kárpát-medence teljes területéről származnak, de a mintavételből csak a magyar vezetéknévű és apai ági eredetű genetikai mintáit tartalmazzák (tehát a szláv, sváb, roma, askenázi stb. származású mintaadók nem szerepelnek a táblázatban.)

kérdéskörére és problematikájára, illetve a területi különbségek és arány-eltolódások jelenségére. Az obi-ugorokra jellemző N-L1034 csoport [167] mindkét mintában 1% gyakoriságú (a székely mintában azonban a magyarországi magyar mintának három és félszerese).

Feltűnő a Kárpát-medence férfiági genetikai folytonosságának hiánya a 6-8 ezer évvel ezelőtti korai földművelő mintákban látható Y-haplocsoport arányokhoz képest: a ma leggyakoribb három csoport (összesen a mai magyar férfiak 60%-ának őse) teljesen hiányzik a neolit minták közül, miközben a neolit minták 38%-át kitevő két csoport (F*/H2; C-F3393) mára teljesen eltűnt a magyarokból, de legalábbis egy ezrelék alatti az előfordulása. Egyedül a G2a-P15 tipikusan neolit földműves csoport, valamint a mezolit vadász-gyűjtögető európai népesség leszármazottjának tekinthető I2* csoport van ma jelen a magyarokban, de ezek súlya is ötödére csökkent a génállományban. A ma a neolitikuminál nagyobb súlyú E1b-M35, I1-M253 és J2-M172 csoportok a neolitikum után más népekkel is érkezhettek a Kárpát-medencébe, így nem feltétlenül a korai földművelők örökségét jelentik.

A viszonylag jól feltérképezett kárpát-medencei neolitikum után kiemelten fontos lenne a bronzkori, vaskori (kelták, szkíták), római kori, hun és avar kori, valamint a honfoglaló csontminták Y-DNS vizsgálata is. A fentiek mellett fontos megjegyezni, hogy a magyar népesség génállománya regionálisan jelentősen eltérhet (ld. a Bodrogi közti populációra vonatkozó vizsgálati eredményeinket, ahol 6%-ban detektálható az N1c [143]), ezért a jövő egyik célja a szakszerű csont-DNS vizsgálatok mellett a különböző magyar etnikai csoportok és tájegységek genetikai összehasonlítása lehet.

Kördiagramokon szemléltetjük (45. ábra) az általunk publikált és az FTDNA minták együttes, közel 1000 férfira vonatkozó adatait. Az ábrán a 0,1%-os gyakoriságú haplocsoportok közül nem mindegyik látható. Látható, hogy ázsiai eredetű haplocsoportok jelenléte elenyésző (N és Q haplocsoportok összesen kb. 4%), tehát a recens magyar népesség apaiági leszármazási vonalai főleg európai népességekben előforduló haplocsoportokból állnak.



45. ábra. Az összes magyar minta Y-haplocsoport eloszlása

5.3. Az mtDNS VIZSGÁLATI EREDMÉNYEK

Mitokondriális DNS vizsgálatot rutinszerűen nem végeztünk, ezért számottevő eredmény értelemszerűen nem született. A témában egy utolsó szerzős és két társszerzős cikkem jelent meg: a mitokondriális haplocsoportok gyakoriságán alapuló klaszterezési módszer kidolgozása, amelyet 174 eurázsiai populációs adaton teszteltünk [186], a szegedi egyetemen dolgozó kollégákkal közösen a honfoglaló temetőkből származó csontminták mitokondriális hipervariábilis régió szekvenciáján alapuló genetikai adatok prezentálása [185], valamint 17 recens európai és közel-keleti populációból származó teljes mtDNS genom szekvencia adatai [187].

5.3.1. 174 eurázsiai populáció mtDNS elemzése iteratív rangkorrelációs módszer segítségével

Ebben a tanulmányban 174 eurázsiai, észak-afrikai és amerikai populáció 27-dimenziós mtDNS haplocsoport eloszlását elemeztük, köztük számos történelmi és történelem előtti mtDNS adatot is. Ennek a munkának a legjelentősebb fejleménye volt a recens és az ősi populációk haplocsoport eloszlásának leírása bizonyos hipotetikus ősi magpopulációk összetételeként, amelyek hosszú időn keresztül közvetlenül vagy közvetve meghatározták Eurázsia migrációs folyamatait. Ezen magpopulációk azonosítására új iteratív algoritmust fejlesztettünk ki, amely meghatározza a vizsgált 27 mtDNS haplocsoport klasztereit, amelyek a populációk meghatározott részhalmazai között erős rangkorrelációval rendelkeznek. E tanulmány alapján a jelenlegi eurázsiai populációk három ősi magpopuláció összetételének tekinthetők az anyai leszármazási vonalak esetében. Azt igyekeztük bizonyítani, hogy a recens és ősi populációk egyidejű elemzése egy új, iteratív rangkorrelációs algoritmussal (IRC: Iterative Rank Correlation) és súlyozott SOC (Self Organising Cloud) tanulási technikák segítségével rávilágíthat a legfontosabb és legmeghatározóbb múltbeli migrációs folyamatokra. Ez a technika lehetővé tette számunkra, hogy földrajzilag, történelmileg és nyelvileg jól értelmezhető klasztereket határozzunk meg, amelyek nagyon specifikus, alig osztályozható struktúrával rendelkeznek. A módszert kétdimenziós, ún. *stepping stone modell* segítségével validáltuk.

A tanulmány egy módszertani cikk volt, ezért terjedelme miatt nem részletezzük, hanem csak a következtetéseket ismertetjük [186]. A módszer, melyet a honfoglalók genetikai összetételének elemzésére alkalmaztunk, röviden a következő fejezetből ismerhető meg.

A tanulmány egyik fő célja az volt, hogy a recens és az ősi mtDNS haplocsoportok eloszlásait, mint bizonyos ősi magpopulációk történelmi vándorlásának következményeit, mutassuk ki. 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 összetétele a kezdeti populációkra jellemző haplocsoportok magas rangkorrelációját eredményezi. Az iteratív rangkorrelációs algoritmust alkalmazva 3 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 nagy számú ősi mtDNS minta bevonása az elemzésbe felbecsülhetetlen értékű eszközt jelentett az ókori népmozgások és a recens populációk élőhelyének összehasonlításában. Az *Andronovo*- és *kurgán*-kultúrákhoz köthető mtDNS adatok recens populációs eloszlásokkal való összehasonlítása alapján az a következtetés vonható le, hogy a recens kelet- és közép-európai populációk egy olyan összetett populációnak tekinthetők, amely a közel-keleti neolitikus forrásból származó nyugati szubsztrátum, illetve egy *Andronovo* területéről eredő szibériai és nyugati komponens keveredéséből származik. A recens és ősi mtDNS haplocsoport eloszlások szoros kapcsolata ezeken a területeken azt is valószínűsíti, hogy az uráli és altáj nyelveket beszélő nők jelentős szerepet játszhattak az *Andronovo*- és a *kurgán*- kultúrában.

A haplocsoportok szisztematikus közös terjedése az ősi magpopulációkban elégséges, de nem feltétlenül szükséges feltétele a haplocsoportok magas rangkorrelációs értékeinek, ezért a történelmi, genetikai, nyelvi kontrollok mindig szükségesek az invalid indikátorok szűréséhez. Reméljük, hogy ezek a módszerek és eredmények elégséges ösztönzést jelentenek későbbi tanulmányokra, mint például genetikai, nyelvi, zenei, stb. korrelációk keresésére.

5.3.2. A karosi honfoglaló temetőből származó minták vizsgálata

A korai magyar korszakból Északkelet-Magyarországon, a Tisza és a Bodrog folyók között, Karos falunál találták a leggazdagabb síremlékekkel tarkított temetőket [188].

A homokdűnék között három nagyon hasonló, jó állapotban lévő temetkezési helyet tártak fel, ahova első generációs honfoglaló magyarok temetkeztek a IX. század utolsó évétől a X. század közepéig. Ezeket a temetőket a tizedik század közepén hirtelen hátrahagyták, valószínűleg a társadalom átszerveződése és a kereszténység terjedése miatt.

Ebben a tanulmányban a Karos III. temetőből származó 19 vizsgált csontminta közül 17-ből megállapított anyai ági haplotípus eredményeket mutatjuk be. A mitokondriális DNS vizsgálat eredménye a szegedi egyetemen dolgozó kollégáknak köszönhető.

A honfoglaló magyarok genetikai összetételének és a recens populációkhoz való viszonyának feltárásához archeogenetikai módszereket alkalmaztunk. A honfoglaló magyarok a 895-ös évektől kezdődően a hét törzs szövetségeként költöztek be az eurázsiai Pontusz-sztyeppéről a Kárpát-medencébe, de további történetük még nem teljesen feltárt. 17 mtDNS haplotípust és 4 Y-kromoszóma haplocsoportot mutattunk ki, amelyek az első generációs honfoglalók temetőjének genetikai összetételét ábrázolják. A **14. táblázat** mutatja a honfoglalók mtDNS haplotípusait és haplocsoportjait. Két mintát (a 2. és 7. sírból) a rossz DNS-minőség miatt kivettünk az adatkészletből. A 17 csontminta 13 különböző haplotípussal rendelkezett, melyek 7 fő haplocsoportba sorolhatók. A leggyakoribb haplocsoport a B volt, amely az A haplocsoporttal együtt a karosi mintapopuláció mintegy 30% -át tette ki. Ez a 2 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, ami 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 [189].

Különösen érdekes a vezérnek tekinthető személy (11. minta) X2f-es anyai haplotípusa, mivel ez valószínűleg dél-kaukázusi eredetű (ebben a régióban a legmagasabb gyakorisága), viszont ritka Kelet-Európában és Közép-Ázsiában, továbbá gyakorlatilag hiányzik a finnugor és a török nyelvű népekből a Volga-Urál régióban [190]. A 10. minta a H5 haplocsoportba tartozik, amely az X2f haplocsoportba hasonló eloszlást mutat, tehát a Kaukázusban a leggyakoribb, a H csoport több, mint 20% -át képezi egyes etnikai csoportoknak, de nem fordul elő a volga-uráli finnugor és a közép-ázsiai populációkban [189]. A H5 Európában leggyakrabban a szlovákokban és a franciákban fordul elő, de egy újkőkori, Kárpát-medencéből származó *LBK (vonaldíszes kerámia kultúra)* mintában is megtalálható [81, 83], így nem feltétlenül érkezett a honfoglalókkal. Az U2 haplocsoport alacsony (0,5-2%) gyakorisággal fordul elő Eurázsia egészében, de a karosi temetőben két független U2-es egyed található, ami azt sugallja, hogy ez a haplocsoport átlagosan gyakoribb ebben a mintacsoportban. Az U2 gyakorisága ma a Volga-Urál vidéken és az Észak-Kaukázus kisebb etnikai csoportjaiban a lemagasabb, de az U2 és az U5 már a mezolitikum óta a legrégebbi európai vonalakhoz is tartozik. Ugyanez igaz a J és T haplocsoportokra is, amelyek Európában a korai neolitikum óta jelen vannak. A T1a haplocsoporttal rendelkező három

minta azt jelentheti, hogy ez a haplocsoport a honfoglaló magyaroknál átlagosan gyakoribb volt, és az is figyelemreméltó, hogy a T1a haplocsoport ma nagyobb gyakorisággal fordul elő székelyek körében, mint a környező populációkban, ahol kb. 2,25% az előfordulása [191].

14. táblázat. A karosi honfoglaló temető mtDNS genetikai összetétele

Minta (sír no.)	HVR-I mutációk	HVR-II és kódoló régiós mutációk	Haplocsoport	Haplocsoport (%)
1.	183C 189C 217C	263G 7028T 9bp del 11719A 14766T	B4	100%
3.	362C	239C 263G	H6	100%
4.	069T 092C 126C 261T	228A 263G 295T 7028T 11719A 12612G 14766T	J1c7	100%
5.	183C 189C 217C	263G 7028T 9bp del 11719A 14766T	B4	100%
6.	189C	263G 7028T 9bp del 11719A 14766T	B4'5	100%
8.	051G 189C 362C	263G 7028T 11467G 11719A 14766T	U2	95,89%
9.	051G	263G 7028T 11467G 11719A 14766T	U2	100%
10.	rCRS	263G	H2a2	100%
11.	189C 223T 278T	263G 6371T 7028T 11719A 12705T 14766T	X	100%
12.	183C 189C 223T 290T 319A	235G 263G 4248C 7028T 11719A 12705T 14766T	A	100%
13.	189C	263G 7028T 9bp del 11719A 14766T	B4'5	100%
14.	126C 163G 186T 189C 294T	195G 263G 7028T 11719A 13368A 14766T	T1a	100%
15.	069T 126C 362C	263G 295T 7028T 11719A 12612G 14766T	J	98,15%
16.	256T 270T	263G 7028T 11467G 11719A 14766T	U5a	100%
17.	362C	239C 263G	H6	100%
18.	126C 163G 186T 189C 294T	214G 263G 7028T 11719A 13368A 14766T	T1a10a	100%
19.	126C 163G 186T 189C 294T	214G 263G 7028T 11719A 13368A 14766T	T1a10a	100%

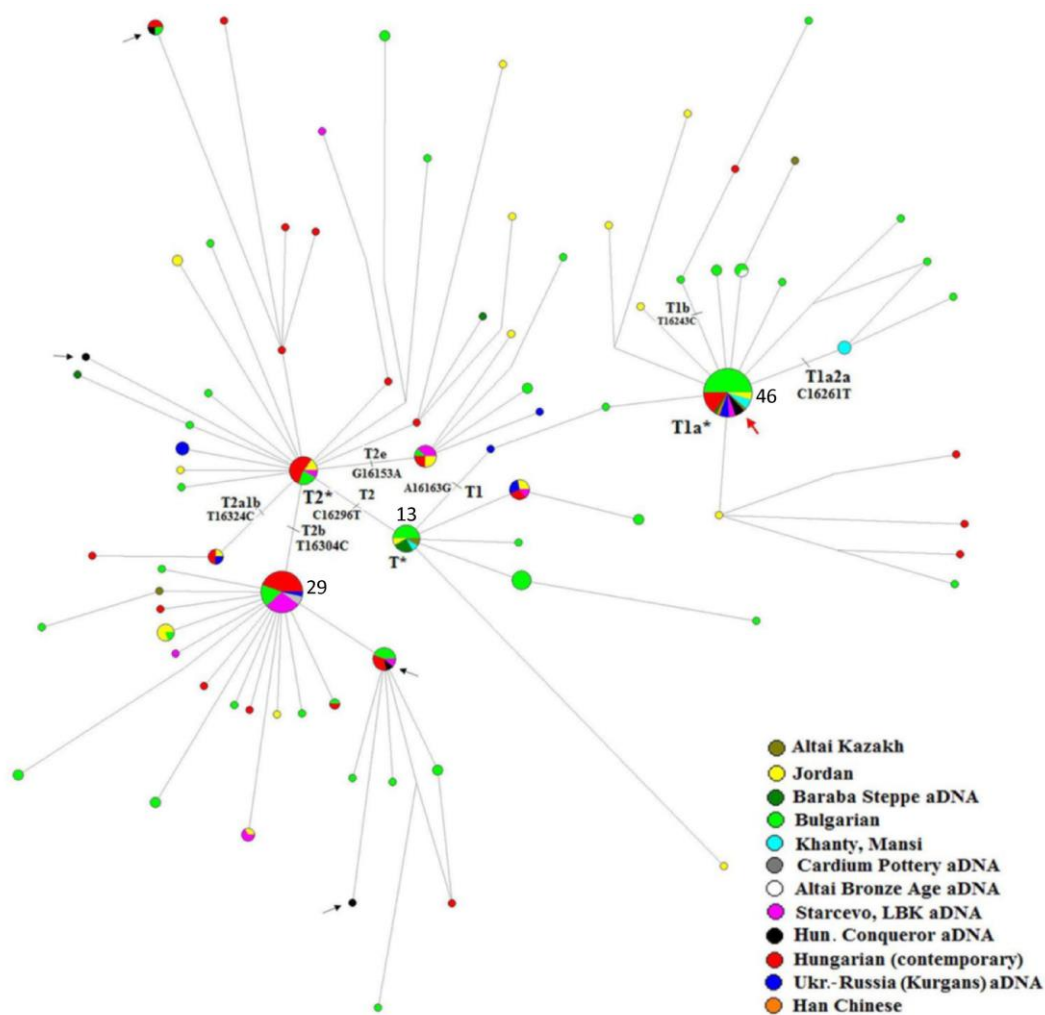
Az elmúlt 10 évben az ősi DNS (aDNS) kutatás jelentősen megváltoztatta tudásunkat a korai európai népekről [74, 75, 76], de a bronzkort követő adatok sporadikusak, ráadásul a történelmileg dokumentált migrációk nagymértékben átalakították a kontinentst. Eredményeink új adatokat szolgáltatnak az utóbbi vándorlások egyikéről, amely elősegítheti a bronzkor utáni genetikai változások megértését. Adataink azt mutatják, hogy a magyarok nemcsak a kontinensen már jelen lévő haplocsoportokat birtokolták, hanem új, ázsiai génekkel frissítették az európai génállományt.

A T haplocsoport median joining (MJ) hálózata a **46. ábrán** látható. A T* haplocsoport alapító haplotípusa másik 12 mintáéval (2 késői krotovói és 2 fjodorovói minta a Barabasztyeppéről, 1 hanti, 1 altáj kazah, 6 bolgár és 1 jordán) egyezik meg. Mind a három karosi honfoglaló T haplotípusú minta a T1a* alapító haplotípusával volt azonos (ld. a piros nyíl a **46. ábrán**), az adatbázisban pedig másik 43 ősi és recens mintával kaptunk találatot; eredetük szerint 23 bolgár, 7 székely és csángó, 2 jordán, 2 manysi, 1 hanti, 1 altáj kazah, 1 han kínai, 3 jamnája-kurgán aDNS, 1 baraba-sztyepei késő bronzkori aDNS, 1 starčevoi neolitikus aDNS, és 1 dunántúli lineáris fazekasságból (*Linear Pottery*) származó aDNS. Ez a megállapítás megerősíti a magyar honfoglalók és a mai magyarok, valamint a bolgárok közötti genetikai kapcsolatot, akik szintén ugyanabból a sztyepei régióból származnak, és filogeográfiai kapcsolat feltételezhető a karosi temető népessége (KAR minták), továbbá az ősi és recens népességek között Dél-Szibériában.

A T2b alapító haplotípus 29 mintát tartalmazott, köztük 13 recens magyar, 5 bolgár, 4 neolitikumi starčevoi minta, 4 lineáris fazekasság minta, 2 benyomásos kerámia (*Cardium Pottery*) neolitikus minta és 1 eneolitikus -késő újkőkori (rézkori) kurgán minta Ukrajnából. A korábban Tömöry és mtsai. [192] által publikált négy honfoglaló magyar aDNS a T2b haplocsoporthoz tartozott (a **46. ábrán** fekete nyilak). Mivel a T1a jelentős dél-szibériai jelenléttel rendelkezik (recens és bronzkori), míg a T2b nem található meg Ázsiában, ráadásul ezek a T2b csoportba tartozó minták azonosak mind a dunai, mind a mediterrán neolitikus aDNS mintákkal, ezért úgy véljük, hogy a karosi honfoglaló T1a minták inkább dél-szibériai anyai gyökerekkel rendelkeznek, viszont a korábban publikált T2 minták a X. századi Kárpát-medencéből valószínűleg a neolitikus európai eredetű haplotípusokat reprezentálják.

A **47. ábrán** a B haplocsoport MJ hálózata látható. A B4 haplocsoport alapító haplotípusa négy mintával azonos: 2 karosi honfoglaló (piros nyíl a **47. ábrán**), 1 han kínai Liaoningból és 1 üzbég Taskentből. A két karosi B4'5 minta (kék nyíllal a **47. ábrán**) két mutáció távolságra helyezkedik el az alapító haplotípustól. Egy korábban publikált [192], honfoglaló B-haplotípusa (fekete nyíl a **47. ábrán**) egy altáj kazahkal azonos a B4c1b alcsoporton belül. Egy recens magyar minta egy mutációs távolságra volt ettől a honfoglalótól.

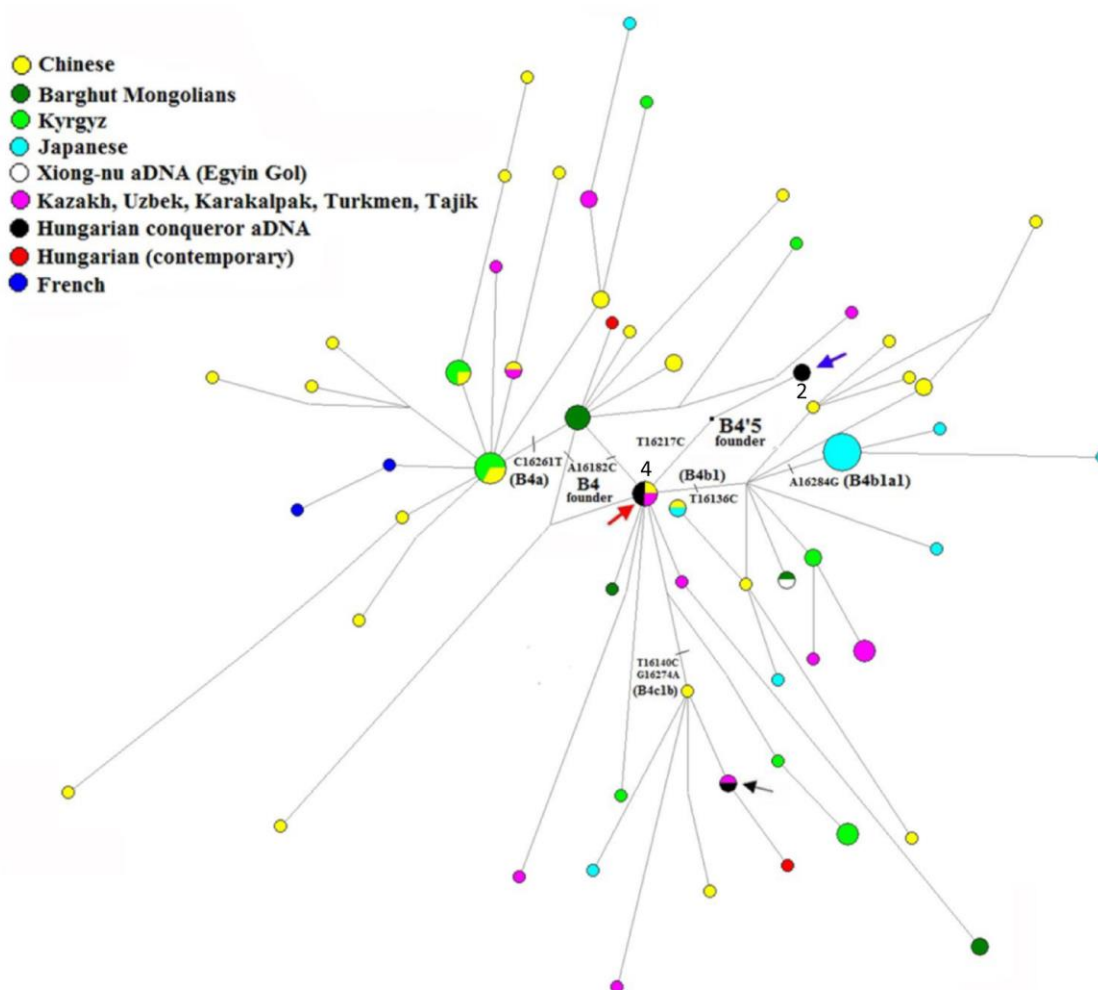
Mivel Liaoning történelmileg a tunguz nyelvű mandzsuk által lakott terület volt, így figyelembe véve az üzbég, az altáj kazah és a honfoglaló találatot (haplotípus egyezés), azt feltételezzük, hogy a honfoglalók B haplocsoport (haplotípus) egyezése valószínűleg az altáji (türk) etnikum jele lehetett.



46. ábra. A T haplocsoport Median Joining hálózata

A körméretek arányosak a haplotípus frekvenciákkal, a legkisebb kör egy személynek felel meg. A színek a populációkat jelentik az ábra alján lévő színskála szerint. A piros nyíl a KAR egyéneket mutatja, a fekete nyilak pedig az AH2 egyéneket mutatják. A hálózat elkészítéséhez használt különböző populációk publikált forrásból származnak [185].

AH2: más szerzők által vizsgált honfoglaló minták [192]; KAR: karosi honfoglaló minták (Bodrogköz).



47. ábra. A B haplocsoport MJ hálózata.

A körméretek arányosak a haplotípus frekvenciákkal, a legkisebb kör egy személynek felel meg. A színek a populációkat jelölik az ábra tetején lévő színskála szerint. A piros és a kék nyilak karosi mintákat jelentenek, a fekete nyíl mutatja az AH2-t. AH2: más szerzők által vizsgált honfoglaló minták [192]; KAR: Karos honfoglaló minták (Bodrogköz).

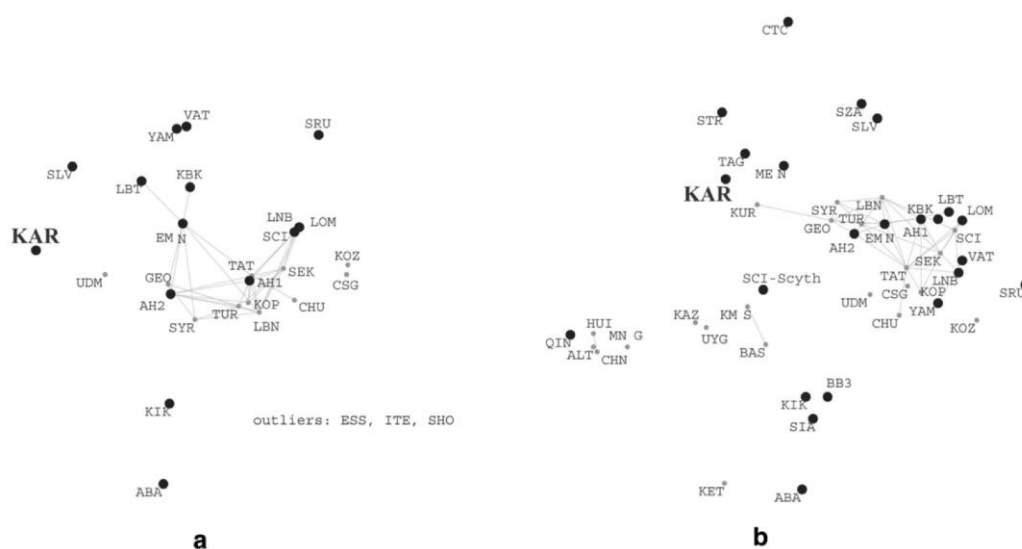
A karosi honfoglalók anyai ági genetikai összetétele több forrásból származó, kevert eredetet mutat. A mintacsoport mintegy 30-40% -a egyértelműen Kelet- és Közép-Ázsiából (East-Middle Asia) származik, legalábbis szélesebb időskálán belül. A B haplocsoport MJ hálózatainak adatai azt jelzik, hogy ez a komponens származhatott a mai altáj (török) nyelvet beszélő etnikai csoportok egy közös génállományából. Annak ellenére, hogy az ázsiai haplocsoportok jelentős hányada jelen volt a honfoglalókban, elsősorban európai antropológiai jellegekkel rendelkeztek, ugyanez igaz a Karos III temetőre is [193]. Ez arra utalhat, hogy a honfoglalók generációkkal a honfoglalás előtt keveredhettek ázsiaiakkal a

kelet-európai Pontuszi-sztyeppéken. A T haplocsoport MJ hálózata azt is jelzi, hogy ellentétben Tömöry és mtsai, 2007-ben publikált adataival [192], genetikai folytonosság mutatkozik a honfoglaló és a recens magyarok között.

A korábban publikált (az előző fejezetben tárgyalt) IRC és a SOC módszerekkel történő, 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 (KAR) álltak legközelebb (**48. ábra**), amelyek a nyugati és a szibériai génáramlás metszéspontjába estek. A SIA és a BB3 populációk földrajzi helyzete illeszkedik a történelmi várakozásokba, a nyugati és szibériai CHgC-k kiegyensúlyozott összetételével már rendelkeztek [194]. Ez arra enged következtetni, hogy a honfoglalók génállományának alapkomponense, 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 populációk rövidítését és haplocsoport eloszlásait ld. a csatolt eredeti cikkben [186]). A karosi minták a 2. klaszterbe kerültek, melyet a nyugati CHgC-k dominanciája jellemez, de jelentős arányú szibériai CHgC-t is tartalmaz. Azonban a karosi minták szibériai komponenseik magas aránya miatt a 2. klaszteren belül viszonylag izoláltak, és genetikai távolságok alapján (euklideszi távolság) megközelítik a 3. klasztert, ahol a nyugati és szibériai komponensek aránya a leginkább kiegyensúlyozott. A 2. klaszter nyugati elemei sokkal hasonlóbba 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.

Figyelemre méltó megfigyelés, hogy a 2. klaszter tartalmaz minden ősi és recens populációt: 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 2 mintacsoport viszonylag magas, 5-7,4%-ban közép-ázsiai és belső-ázsiai apai haplocsoportokat is tartalmaz [142], ami megkülönbözteti őket a nyugat-európai populációktól. Az is említést érdemel, hogy a 35 ősi populáció közül 14 populáció 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 tartalmazza (LBT: LBK Transdanubia; STR: Starčevo; SZA: Szakálhát, Hungary; LOM: Lombard early medieval,

Hungary-Italy; SLV: Medieval Slavic, Slovakia; AH1: Ancient Hungarian, 10th cent; AH2: Hungarians 900 AD; KAR: Ancient Hungarian Karos) [185]. Ennek a váratlan eredménynek az esetleges magyarázata a közel-keleti neolitikus (MEN) populációban keresendő, ami nem messze helyezkedik el a KAR mintáitól az MDS térképen (ez azt is jelentheti, hogy mindezen populációk jelentős genetikai behatást kaphattak a korai neolitikumban a Termékeny Félhold területéről). Nyelvi szempontból meg kell jegyezni, hogy nagyon kevés indoeurópai nyelvet beszélő népesség (SCI = skót és KUR = kurd) került be a 2. és 3. klaszter egységes csoportjába a recens populációk közül, a többség a türk, mongol, uráli és kartvéli nyelvi csoportokhoz tartozott.



48. ábra. a). A 2. klaszter MDS-térképének Hg eloszlása. **b).** A 2. és 3. klaszterek MDS-térképének kombinált Hg eloszlása. A vonalak, amelyek összekapcsolják a populációkat a legkisebb távolságokat jelenti. Az ősi populációkat nagy fekete pontok, míg a recens populációkat kis szürke pontok jelölik. A populációk rövidítése a kapcsolódó cikkek gyűjteményében található [185].

Összefoglalva: az új algoritmus segítségével, amellyel összehasonlítottuk a birtokunkban lévő mitokondriális DNS-haplocsoport adatokat más ókori és recens eurázsiai adatokkal, kiderült, hogy a honfoglalók génállományának jelentős része valószínűleg egy régen konszolidált, közép-ázsiai és dél-szibériai génállományból származhatott, amely még mindig detektálható a recens magyar népességben. A honfoglaló magyarok nyugatra vándorlásuk során különböző európai eredetű populációkkal való keveredésből szerezhették egy másik genetikai rétegüket, és ezek közül az egyik fontos elem a kaukázusi régióra mutat. Azok a recens populációk,

amelyek a honfoglaló magyarokhoz genetikailag legközelebb állnak, ma indoeurópai nyelveket beszélnek.

Eredményeinkkel hozzájárulhattunk ahhoz, hogy megértsük Európa benépesítését oly módon, hogy ősi mtDNS adatokat szolgáltatunk egy genetikailag eddig kevésbé vizsgált humán migrációs időszakból.

5.3.3. Az anyai ági mtDNS haplocsoportok a mai magyar népességben

A **15. táblázat** az mtDNS haplocsoportok eloszlását mutatja a recens magyar populáció és a honfoglaló mintacsoport összehasonlításával. Az összesített recens magyar mtDNS haplocsoport eloszlását a **49. ábrán** látható.

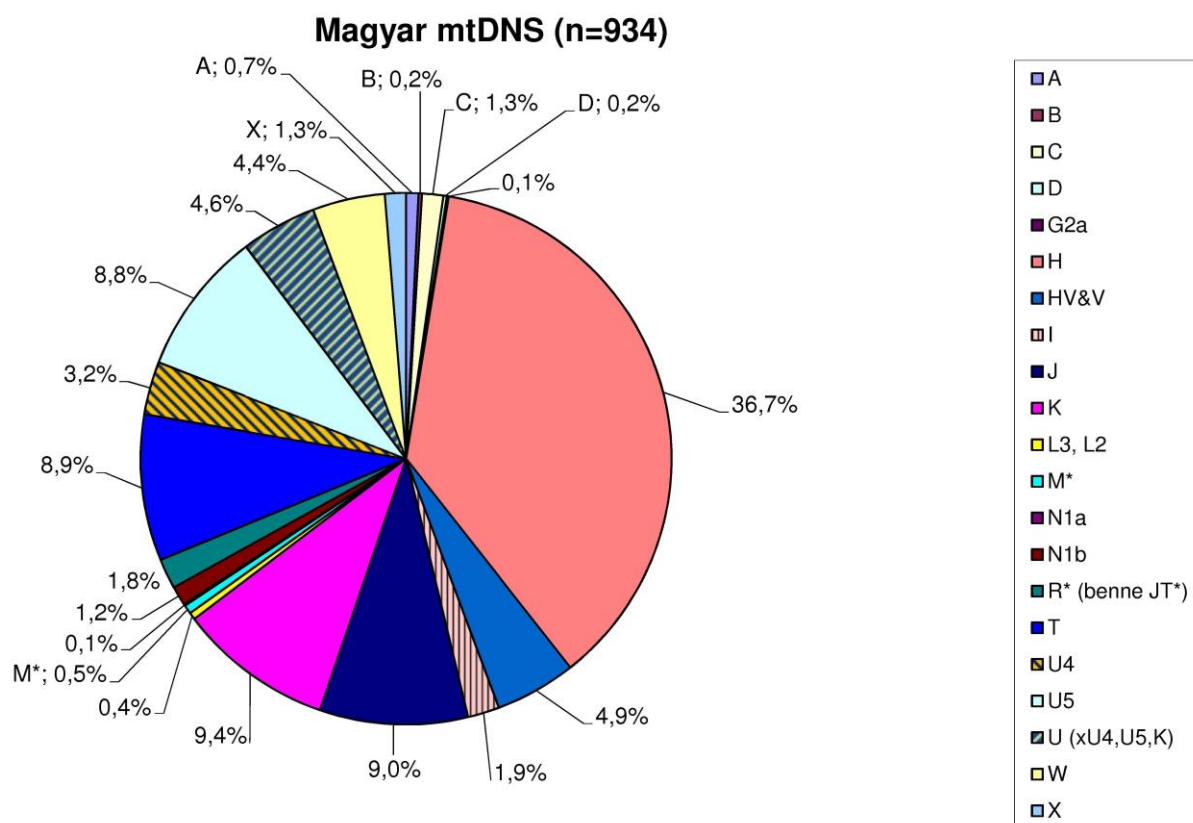
15. táblázat. Magyar populációk mtDNS haplocsoport összetétele

Haplocsoport	Magyarországi	Székely	Csángó	Összes élő magyar	Honfoglaló Magyar	Neolit	A haplocsoportot legnagyobb mértékben hordozó nem magyar etnikai csoportok és térségek
Mintasza	442	259	233	934	49	85	
A	0,2%	0,8%	1,7%	0,7%	2,0%	0,0%	Eszkimók, Szibéria
B	0,2%	0,4%	0,0%	0,2%	2,0%	0,0%	DK-Ázsia, Szibéria
C	0,0%	4,2%	0,4%	1,3%	2,0%	1,2%	Szibéria
D	0,5%	0,0%	0,0%	0,2%	0,0%	1,2%	Kelet-Ázsia
F	0,0%	0,0%	0,0%	0,0%	2,0%	0,0%	Kelet-Ázsia
G2a	0,0%	0,4%	0,0%	0,1%	0,0%	0,0%	Csukcsok, Szibéria
H	43,7%	34,0%	26,6%	36,7%	30,6%	22,4%	Nyugat-Eurázsia
HV&V	6,8%	4,6%	1,7%	4,9%	6,1%	3,5%	Nyugat-Eurázsia
I	2,5%	0,8%	2,1%	1,9%	2,0%	0,0%	Nyugat-Eurázsia
J	7,9%	7,7%	12,4%	9,0%	2,0%	7,1%	Nyugat-Eurázsia
K	5,4%	6,2%	20,6%	9,4%	4,1%	18,8%	Nyugat-Eurázsia
L3, L2	0,9%	0,0%	0,0%	0,4%	0,0%	0,0%	Afrika
M*	0,7%	0,0%	0,9%	0,5%	2,0%	0,0%	India, Pakisztán
N1a	0,2%	0,0%	0,0%	0,1%	6,1%	10,6%	Közel-Kelet
N1b	0,2%	0,4%	3,9%	1,2%	2,0%	0,0%	Nyugat-Eurázsia
R* (benne JT*)	0,9%	3,5%	1,7%	1,8%	4,1%	1,2%	Közel-Kelet
T	9,7%	11,6%	4,3%	8,9%	8,2%	20,0%	Nyugat-Eurázsia
U4	4,1%	2,7%	2,1%	3,2%	4,1%	1,2%	kőkorszaki európai
U5	7,9%	8,5%	10,7%	8,8%	4,1%	2,4%	kőkorszaki európai
U (xU4,U5,K)	3,8%	6,2%	4,3%	4,6%	12,2%	3,5%	Nyugat-Eurázsia
W	3,2%	6,2%	4,7%	4,4%	0,0%	2,4%	Nyugat-Eurázsia
X	0,9%	1,5%	1,7%	1,3%	4,1%	4,7%	Európa, Szibéria
Y1	0,2%	0,4%	0,0%	0,2%	0,0%	0,0%	Amur-vidék, Szibéria

A honfoglaló minták forrása több publikáció [192, 195]. A mai Magyarország területén található újkőkori (neolit) földművelő minták az i.e. 5800-4200 közötti időszakból

származnak, ahogy erre több publikáció [81,82 83] is utal. A táblázat székely és csángó mintái szintén tudományos publikációkból származnak [191, 192]. A magyarországi minták forrása egyrészt a Family Tree DNA Hungarian mtDNA projektje, publikációk [13, 192], valamint egy 25 fős, egyik szerzőnk által gyűjtött, eddig nem publikált minta. Fontos felhívni a figyelmet arra, hogy a 10 milliós magyarországi népesség mintaszáma az összesen 1 milliónál kevesebb székely és csángó összegével megegyező, ami az összmagyar mintát erősen torzíthatja (arányeltolódást eredményez az összesítésben). Ezért a 2. oszlopot érdemes összehasonlítani külön is a 3. és 4. oszloppal. Különösen szembetűnő a K haplocsoport arányeltolódása, amely Brandstätter és mtsai (2007) szerint [191] csángó mintában meghatározó, ez valószínűleg egyetlen „alapító” csángó őspanyára és genetikai sodródásra vezethető vissza. A csángók körében a J csoport is sokkal gyakoribb, mint a másik két mintában. A székely minta elsősorban a C és T haplocsoportok adatai alapján különbözik az összmagyartól. Hangsúlyozni kell továbbá, hogy az anyai ági leszármazás - a vezetéknevek alapján viszonylag jól visszakövethető és rendszerezhető Y-haplocsoporttal szemben - nem ad lehetőséget az őspanyák pontosabb vizsgálatára (így a magyarországi mintában gyakorlatilag a Kárpát-medence minden etnikai csoportja képviseltetheti magát).

A neolit kori földművelő mintáknál látható, hogy a K és T csoportok aránya a H csoportéhoz mérhető, a honfoglalás korára azonban jelentős eltolódás figyelhető meg a H csoport irányába. Különösen érdekes az N1a csoport, amely a neolit kori 10%-ról a 6%-os honfoglaló adaton át mára gyakorlatilag eltűnt a magyar génállományból. Az mindenesetre leszögezhető, hogy a mai legmeghatározóbb anyai ági haplocsoportok már a neolitikumban, azaz 6-8 ezer éve jelen voltak a Kárpát-medencében, miközben az apai Y-kromoszómális tulajdonságok jelentős mértékben lecserélődtek azóta. Ugyanakkor további, teljes mtDNS genomon történő összehasonlító vizsgálatokat igényelne annak felmérése, hogy a mai magyar női ágak azonos haplocsoporton belül milyen mértékben jelentik a neolit földművelők közvetlen továbbélését, és milyen mértékben a más területekről történt későbbi rátelepülést (a H, J, K, T csoportok egész Nyugat-Euráziában széles körben elterjedtek, így több migrációs hullám is hozhatta őket a Kárpát-medencébe).



49. ábra. Az összesített magyar minta mtDNS haplocsoport eloszlása
Az ábrán csak az ázsiai eredetű haplocsoportok nevét tüntettük fel százalékok mellett.

5.3.4. A teljes mtDNS genom vizsgálata európai és közel keleti populációkban

Az európai őstörténet megértését forradalmasította az új generációs szekvenáló technológiák (NGS) elérhetősége, amelyek lehetővé tették a modern és az ősemberi genomok szekvenciaváltozásának elfogulatlan jellemzését. Az ősi (ancient) genomi DNS (aDNA) adatok egyértelmű diszkontinuitást mutatnak a paleolitikus vadászó-gyűjtögetők és a neolitikus gazdálkodók között [72, 73, 74, 75, 76]. A genetikai diverzitási mintázat alapján kis létszámú populációkat lehet feltételezni a paleolitikumban, nyugati és skandináviai csoportok közötti regionális különbségekkel. Ezt a képet tovább finomítja az ókori jamnája pásztorok DNS-ének tanulmányozása a Pontuszi-Kaszipi sztyeppéről, amely nyilvánvaló forrása a bronzkori Európába és Ázsiába irányuló migrációnak (ld. 19. ábrát) [74, 75, 80], valamint az indoeurópai nyelvek vitatott származási területe is [161]. Ezek a tanulmányok kimutattak egy mintegy 4,5 ezer évvel ezelőtti keletről kiindult genetikai hatást Európára, majd feltárták a bronzkori Európában kialakult genetikai struktúrát [74, 75].

Az európai őstörténetre vonatkozó hipotéziseket nemcsak az autoszómális DNS vizsgálatok befolyásolták nagymértékben, hanem az uniparentálisan öröklődő markerek vizsgálata is, utóbbiak a nemek szempontjából „elfogult” (sajátos) demográfiai folyamatokról szolgáltatnak információt. Az aDNS elemzések azt mutatják, hogy a mai leggyakoribb Y-kromoszóma haplocsoport (R1b-M269) Európában 4,5 ezer évvel ezelőttig nagyon ritka volt [196], miközben jelen volt minden, a *Jamnája-kultúrát* reprezentáló mintában [74, 80] (ld. **a 19. ábrát**). Ez eredetileg a Pontuszi-Kaszipi sztyeppéről származó férfiak jelentős behatását sugallta; azonban az Európa-szerte általánosan elterjedt R1b alcsoport (R1b-L11) a Jamnája minták között még nem volt megtalálható [187]. Az Y-kromoszómális adatokkal ellentétben az ősi mtDNS minták adatai rávilágították az anyai vonalak késői jégkorszakban, azaz cirka 14,5 ezer évvel ezelőtt történt elterjedésének fázisaira [197, 198], és az autoszómális DNS adatokkal együtt megerősítették a paleolitikumi vadászó-gyűjtögetők és a neolitikumi földművelők genetikai diszkontinuitását, viszont nem mutatták a bronzkori sztyepei eredetű expanzió semmilyen látható jelét [199].

Az MSY (Y-kromoszóma férfi-specifikus régiója) újra-szekvenálási eredményeink [85] és más tanulmány is [200] hangsúlyozza a bronzkori átmenet fontosságát. Az eredményünkre támaszkodó demográfiai rekonstrukciók alátámasztanak egy expanziót, amely kb. 2100-4200 évvel ezelőtt kezdődött. A három nagy érintett haplocsoport – beleértve az R1b-t – legközelebbi közös őstől (TMRCA) való leválásának idejét 3500 és 7300 év közöttire becsüljük [85]. Számos mtDNS vizsgálat történt recens európai mintákon, leginkább a mitokondriális genom hipervariábilis régiójára összpontosítva, abból kiindulva, hogy az európai kontinens egyetlen egységnek tekintendő [201, 202]. Olyan filogeográfiai megközelítés is szerephez jutott, mely az érdeklődésre számot tartó specifikus leszármazási vonalakra összpontosított [203, 204, 205]. Általános következtetésként fogalmazódik meg, hogy a mai mtDNS variáció reprezentálja az utolsó jégkorszak utáni újbóli expanziót. Azonban hiányoztak a populációkra jellemző teljes mtDNS genom vizsgálatok Európában, így a férfiak és a nők demográfiai történeteinek szisztematikus összehasonlítására még nem adódott lehetőség.

Ezért nemzetközi kooperációban egy populáció alapú vizsgálatot végeztünk, amely 17 populációból származó, populációnként 20-20 – összesen – európai és közel-keleti egyén mintájában a teljes mtDNS genom újra szekvenálására irányult. A projektben vizsgált magyar minták egy részét – a felmenői legalább 3 generáción keresztül ugyanazon földrajzi régióban éltek – a csíkszeredai búcsújárás alatt gyűjtöttem, a donorokat mintavétel után egy kupica pálinkával jutalmaztam. Összehasonlítható adatként már rendelkezésre álltak ugyan ezen

populációk Y-kromoszómális (MSY régió) szekvenciái [85], 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.

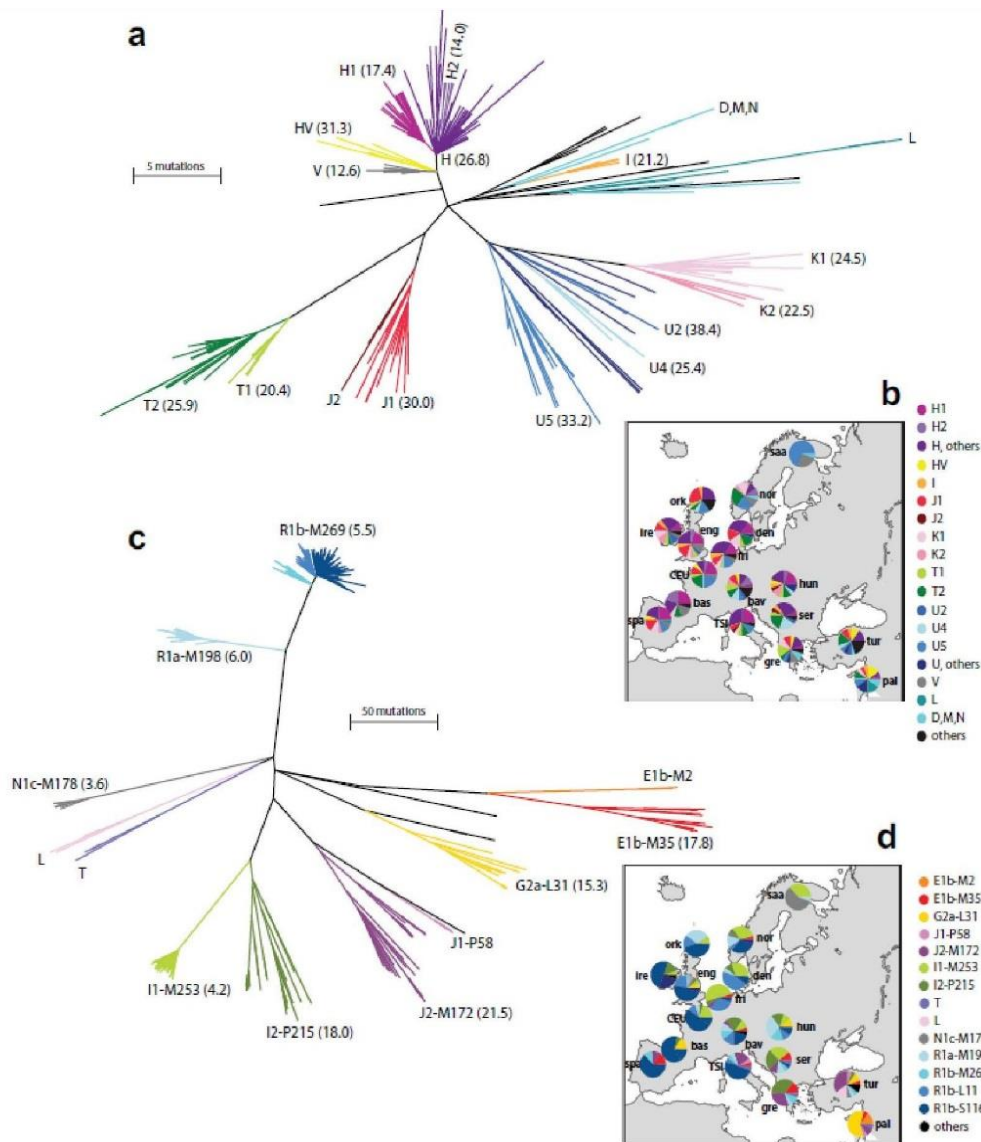
Az mtDNS kódoló régión alapuló maximális parszimónia (MP) fa (50a. ábra) a legmegfelelőbb megbízható filogenetikai következtetések tételéhez a viszonylag alacsony számú rekurzív mutáció miatt. A szekvencia alapú haplocsoport gyakoriságok és azok földrajzi eloszlása (50b. ábra) nagyjából összhangban van a korábban publikált adatokkal [206]. 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% sok kisebb haplocsoportból áll; ezek közé tartoznak (a palesztin és a spanyol mintákban előforduló) például az L2 és L3 vonalak, amelyek jellemzően a Szaharától délre eső Afrika területén fordulnak elő [207].

A haplocsoportok eloszlásának vizuális vizsgálata (50b. ábra) azt mutatja, hogy a számi/lapp populáció egyértelműen kívül áll (alacsonyabb diverzitása miatt), az összképükben az U5 és a V haplocsoport dominál, egyetlen D-haplocsoportba (főként Észak- és Kelet-Ázsiában fordul elő) tartozó minta mutatható ki [208]. A mitokondriális genom egyes szegmenseinek elemzése alapján tett korábbi megfigyelésekkel összhangban a többi populációban nem mutatható ki egyértelmű földrajzi mintázat [206].

A mitokondriális genomok filogeográfiai analizését összehasonlítottuk az MSY analizisével (50c. és 50d. ábrák), amely 3,7 Mb Y-DNS újra szekvenálásán alapult [85]. A haplocsoportok földrajzi eloszlása az MSY alapján sokkal inkább korlátozott (50d. ábra), mint az mtDNS alapján (50b. ábra): például az R1b különösen nagy gyakorisággal van jelen észak-nyugaton, az R1a Észak- és Közép-Európában, a J2 pedig Dél-Európában. A két MP fának nagyon eltérő a skálája, mivel az mtDNS-ben sokkal kisebb számú nukleotidot elemeztünk, mint az MSY-ben (1: 250 arány). Viszont mindkét fa deep-rooting (mélyen gyökerező) kládokat (mtDNS: U, K és T2 haplocsoportok; MSY: E1b-M35, G2a-L31, I2-P215, J2-M172, L és T), valamint csillagszerű mintázatokat mutatott, ami a populációk expanzióját jelenti. Az mtDNS esetében ezek a haplocsoportok a mintánk 51,7% -át reprezentáló H, J1, T1 és V; míg az MSY esetében az I1-M253, N1c-M178, R1a-M198 és az R1b-M269 (összesen 64%). A fő különbség a belső ágak relatív hosszúsága, ami azt mutatja, hogy az mtDNS vonalak (anyai) expanziói sokkal ősbbe, mint az MSY vonalaké (apai). Ezt

támasztja alá a csillagszerű haplocsoportokra vonatkozó **(50a. és 50c. ábra, 16. táblázat)** TMRCA becslés is (a teljes mitokondriális genom alapján), amelyek mindegyike az MSY esetében legfeljebb 6 ezer év (post-neolitikus) [85], míg az mtDNS-esetében több mint 13 ezer év (paleolitikus).

Bár ezek az eredmények összhangban vannak a korábban megfigyelt kontinens szintű különbségekkel [201, 202], itt is az látható, hogy a különbség egyes populációk nagy részére igaz, ugyanakkor rámutat a földrajzi mintázat hiányára is Európában. A legősibb közel-keleti (Near and Middle East) mtDNS expanziók eurázsiai korai benépesítéséhez közelebbi időszakra datálhatók (40-50 ezer év). Az európai anyai és apai vonalak expanziójának időbeli eltérése jól kimutatható az indiai szubkontinensen is, egy jelenlegi tanulmány [209] szerint az mtDNS vonalak expanziói a preholocén korszak folyamatait tükrözik, míg az MSY expanziója többnyire az elmúlt 10 ezer évben történtekét (a bronzkor idején Közép-Ázsiából kiindulva).



50. ábra. Az európai mtDNS és MSY vonalak filogenetikai és földrajzi eloszlása

(a) A kódoló régió újra szekvenálásán alapuló maximum parszimónia mtDNS fa [85]. Az ágak hossza arányos a haplotípusok közötti molekuláris divergenciával, a színek a haplocsoportokat jelölik. A becült TMRCA ezer években zárójelben a haplocsoport neve után olvasható (ld. még 16. táblázat).

(b) A kördiagramok az mtDNS haplocsoportok frekvenciáit ábrázolják 17 Európából és a Közel-Keletről származó populációban (színeket ld. az (a) pontnál) [85]. A populációk rövidítései: **bas**: spanyol baszkok; **bav**: Bajorország; **CEU**: északi és nyugat-európai származású lakosok Utahból, CEPH gyűjtemény (HapMap projekt, Franciaország); **den**: Dánia; **eng**: Anglia **fri**: Frízia (Hollandia); **gre**: Görögország; **hun**: Magyarország; **ire**: Írország; **nor**: Norvégia; **ork**: Orkney (Egyesült Királyság); **pal**: Palesztina; **saa**: számi (Finnország); **ser**: Szerbia; **spa**: Közép Spanyolország; **TSI**: Toszkána (Olaszország); **tur**: Törökország.

(c) Y-kromoszóma MSY régiójának (3,7 Mb) szekvenálásán alapuló maximum parszimónia fa [85]. Az ágak hossza arányos a haplotípusok közötti molekuláris divergenciával, a színek a haplocsoportokat jelölik. A becült TMRCA zárójelben található a haplocsoport neve után.

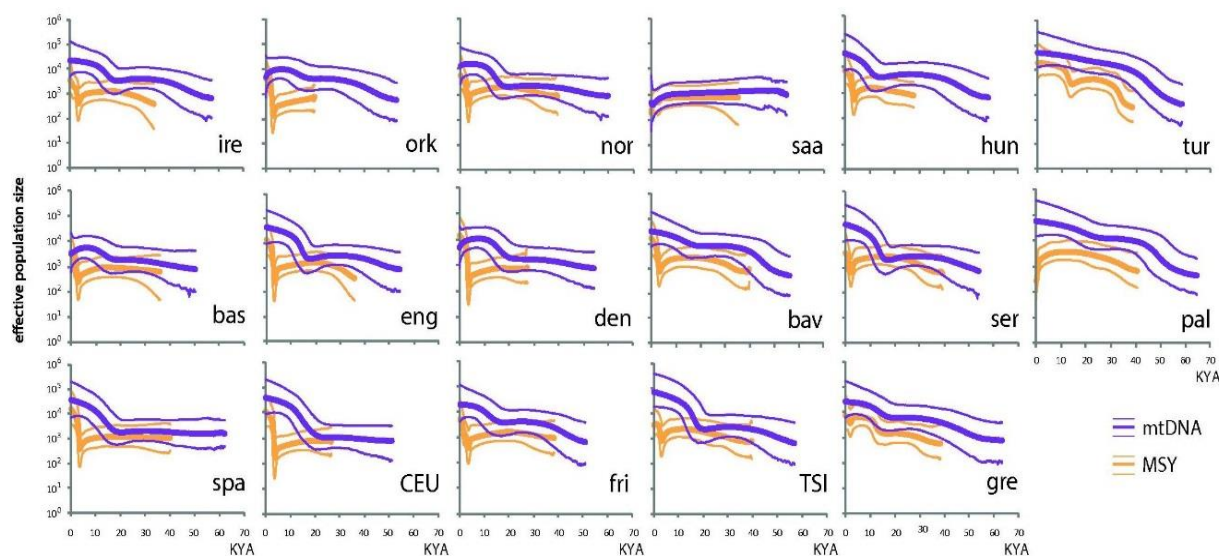
(d) A kördiagramok az Y-haplocsoportok frekvenciáit ábrázolják 17 Európából és a Közel-Keletről származó populációban (színeket ld. a (c), populációk rövidítéseit a (b) pontnál) [85].

16. táblázat. A legnagyobb mitokondriális haplocsoportok kora (TMRCA) Európában

Hg	N	TMRCA/K YA	95% HPD interval
H	116	34.1	25.2-45.0
H1	39	33.1	24.3-43.2
H2	11	11.4	5.2-20.0
HV	10	43.8	28.8-59.4
J	31	39.7	26.3-55.0
J1	28	37.8	24.2-53.6
I	8	23.3	15.1-32.2
K	25	39.2	28.9-50.6
K1	17	33.0	23.7-43.5
K2	8	30.3	18.7-42.6
T	46	38.6	27.0-52.0
T1	14	22.0	11.5-34.5
T2	32	28.3	19.6-38.3
U	61	72.2	60.5-85.0
U2	6	67.7	50.3-83.1
U4	9	30.4	17.1-45.3
U5	37	45.2	31.7-59.9
V	18	19.9	10.9-31.4

Hg: haplocsoport; N: a szekvenciák száma; HPD: a legmagasabb utólagos denzitás; KYA: ezer évvel ezelőtt.

A demográfiai történeteket az mtDNS szekvenciákon alapuló Bayesian Skyline Plots (BSPs) alkalmazásával, mint populációkat rekonstruáltuk, nem pedig mint a leszármazási vonalakat (**51. ábra**). Ahogy azt a „szokatlan” haplocsoport összetétel alapján vártuk, a számi populáció itt is kívülálló a tényleges populáció méretének folyamatos csökkenése miatt, ami 5 ezer évnél egyre markánsabbá válik. Minden más populáció 13-20 ezer évvel ezelőtti paleolitikus expanzió nyomait viseli. A török és a palesztin minták eltérnek a többségtől a populáció sokkal régebbi (úgy 40 ezer éve) expanziója miatt. Ezek a mintázatok ugyanezen populációcsoporton belül éles ellentétben állnak az MSY [85] BSP mintázataival, amely a legtöbb populáció esetében (13 populáció 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. Valamennyi összehasonlításban a tényleges populáció méretének jelenlegi becslése magasabb az mtDNS, mint az MSY esetében, kivéve a baszkokat és a dánokat.



51. ábra. Bayesian Skyline Plots (BSP) az mtDNS és az MSY esetében.

A vastag vonalak jelzik a tényleges populáció méret medián értékét, a vékony vonalak pedig a 95% -os posterior denzitási intervallumokat (mtDNS: lila, MSY: narancssárga). A populációk rövidítései: **bas**: spanyol baszkok; **bav**: Bajorország; **CEU**: északi és nyugat-európai származású lakosok Utahból, CEPH gyűjtemény (HapMap projekt, Franciaország); **den**: Dánia; **eng**: Anglia **fri**: Frízia (Hollandia); **gre**: Görögország; **hun**: Magyarország; **ire**: Írország; **nor**: Norvégia; **ork**: Orkney (Egyesült Királyság); **pal**: Palesztina; **saa**: számi (Finnország); **ser**: Szerbia; **spa**: Közép Spanyolország; **TSI**: Toszkána (Olaszország); **tur**: Törökország. Az MSY-grafikonok a Battini és mtsai. (2015) cikkében [85] publikáltak módosított verziói.

Minden egyes populációra diverzitási indexet számoltuk (az adatokat ld. Battini és mtsai (2015) cikkében) [187]. Összehasonlítva a BSP megfigyeléseivel: a számi csoportban korlátozott számú haplocsoport található, ez a populáció mutatja a legalacsonyabb értéket minden diverzitási mérésben. A legmagasabb értékeket a palesztin és a török minták adták, ami ismét összhangban van a BSP-mintázatokban megfigyelt ősi népességnövekedéssel.

Mind a D , mind az F_S esetében negatív értékek figyelhetők meg a számit kivéve az összes populációra, ami a népesség növekedését jelezheti; azonban mindkét érték csak tizenkét esetben szignifikáns a fennmaradó populációkban. Első pillantásra a déli populációkban nagyobb diverzitás mutatkozik, mint az északi populációkban (50b. ábra). Ennek formális vizsgálata érdekében korreláció-elemzést végeztünk a genetikai diverzitás (polimorf helyek száma, illetve nukleotid diverzitás) és a földrajzi szélesség, hosszúság, illetve a jégkorszaki menedékhelyektől (a franko-kantábriai, ill. a közel-keleti) szárazföldön mérhető távolságok mérőszámai között. Az összes populáció bevonásával mindkét mérőszám

statisztikailag szignifikáns korrelációt mutat a földrajzi szélességgel és a közel-keleti menedékhelytől való távolsággal, de nem a földrajzi hosszúsággal. Ezek a korrelációk elvesznek, ha eltávolítjuk a fent leírt „kilógó” populációkat (számi, török és palesztin), bizonyítva, hogy Európa legnagyobb részén az mtDNS-nek nincs diverzitási mintázata.

A 17 populációs adatunkban is szereplő számi kilógó státusza egyértelmű, mivel nemcsak az egymással szorosan összefüggő mtDNS szekvenciák száma magas köztük, hanem a máshol ritkán előforduló mtDNS és MSY haplocsoportoké is (**50b. ábra és 50d. ábra**), illetve több egyénben is detektálhatók azonos MSY szekvenciák és Y-STR haplotípusok (saját publikált adat) [86]. Ezek a jellemzők megegyeznek a BSP-ben megfigyelt tényleges populációméret növekedésének hiányával (**51. ábra**). A populáció alapú genom-szintű SNP-analízis [210] és egy személy teljes genom szekvenálása [211] azt is megmutatja, hogy a számi populáció genetikailag differenciálódott az európaiakhoz képest és kelet-ázsiai eredetű genetikai komponenseket hordoz. Adataink összhangban vannak az aDNS adatokkal is [197, 198, 199] megerősítve a uniparentális genetikai folyamatok szerepét a mai európai demográfiai változásokban. A recens mtDNS diverzitás mintázatai nem utalnak vissza a bronzkori expanzióra, miközben a recens európai MSY diverzitás nagy részét a bronzkori expanzió alakította ki [85]. Ezzel ellentétben a recens MSY adatok nem mutatják egyértelmű jelét a neolitikus átmenetnek, melyet viszont az ősi mitokondriális és autoszómális adatok sugallnak [196, 199]. Ez a jelenség a mintavételi hatásoknak, valamint a genetikai sodródásnak tudható be, ami fontos szerepet játszott az uniparentális markerek megfigyelt mintázatának kialakításában.

Számos előrelépés történt az európai kontinens prehisztorikus benépesülésének megértésében, ennek során az első klasszikus genetikai adatokat a földművelők által közvetített démikus diffúzió javára értelmezték [212]. A recens és ősi DNS adatok ma már elérhető sokasága rávilágított egyes, korábban nem várt, múltbeli migrációs és expanziós eseményekre is. Ezeket a nemek szerinti külön vizsgálat és a populáció alapú újra szekvenálási megközelítés is alátámasztja. A jövőt illetően azonban még mindig sok a kutatnivaló, egyrészt a minták méretének és földrajzi lefedettségének növelése, másrészt az ősi és a recens adatok teljes körű integrálása tekintetében.

5.4. A ZENÉBE VÉSETT GÉNEK, AVAGY A ZENEGENETIKA (MUSIC-GENETICS)

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

A kutatómunka során Juhász Zoltán barátommal (KFKI, mesterséges intelligencia kutató) való véletlen találkozásnak köszönhetően sejtettük meg a népzében 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 Organising Map), amelyekre már megvoltak a genetikában általánosan használt MDS módszerrel készült ábrák.

Nagy örömünkre zenei MDS módszerrel ugyan-olyan populációs klasztereket kaptunk, mint a nem-metrikus MDS módszerrel, amely szaknyelven a módszer validálását jelentette.

Később megszületett a közös zenegenetikai cikk, azonban nagy „problémát” okozott, hogy találjunk olyan folyóiratot, amelynek hatóköre illeszkedne a cikk témájához. Mivel a cikk tartalma – főleg zenegenetikai adatok - nemzetközi szinten teljesen új interdiszciplináris terület létrejöttét jelentette, bírálata pedig populációgenetikai, matematikai-statisztikai, valamint zenetudományi ismereteket követelt, nem találtunk megfelelő folyóiratot, nem is beszélve arról, hogy mellékletként nem DNS szekvenciákat, hanem népzenei kottákat csatoltunk a cikkhez, ami akkor elég szokatlan társításnak tűnt. Végül az 1908-ban, a világon elsőként alapított genetikai szaklap, a *Molecular Genetics and Genomics* jóvoltából sikerült a publikációt 2012-ben megjelentetni. Azóta már tucatnyi cikk és PhD disszertáció is született ebben a témában, bizonyítva és megerősítve eme eredeti magyar kutatási irányzatnak helyességét, ami jó érzéssel tölt el.

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.

Az emberi történelem szorosan összefügg a népességek migrációjának történelmével. A genetikai diverzitás mintázata szolgáltatja a populáció történetére vonatkozó információkat, mivel minden fontosabb demográfiai esemény nyomot hagyott hátra a populációk genetikai diverzitásában, és mivel ezek a demográfiai szignálok átadódnak generációról generációra, így a jelenkor embereinek genomja az egyének demográfiai múltjának lenyomatát is őrzi.

Az evolúciós történelmi tanulmányok során nagy szerep jut az mtDNS és az Y-kromoszóma MSY régiójának elemzésének, mivel az emberi genom ezen régiói nem „köttyvasztódnak” össze rekombinációval, hanem egyik generációról a másikra generációra változás nélkül adódnak át, „szerzetesi tisztasággal” megtartva egy-egy populáció anyai és apai leszármazási vonalát. Populációk osztoznak közös mtDNS vagy az Y-kromoszóma leszármazási vonalakon, amelyek utalnak a közös eredetre vagy a génáramlási folyamatokra (gene flow vagy admixture).

A populációk közötti genetikai hasonlóság származtathatók a közös eredetből vagy a közelmúltbeli, földrajzi közelségből eredő keveredésből. A populációk közötti genetikai távolságok általában korrelálnak a földrajzi távolságokkal, az *ún. távolságtól függő elszigeteltségi modell* szerint [118], de ez nem releváns minden esetben.

A fejlesztett 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 [213, 214]. 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 rendelkezésre állt. 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, dakotai, komi, hanti, horvát-szerb (Balkán csoport),

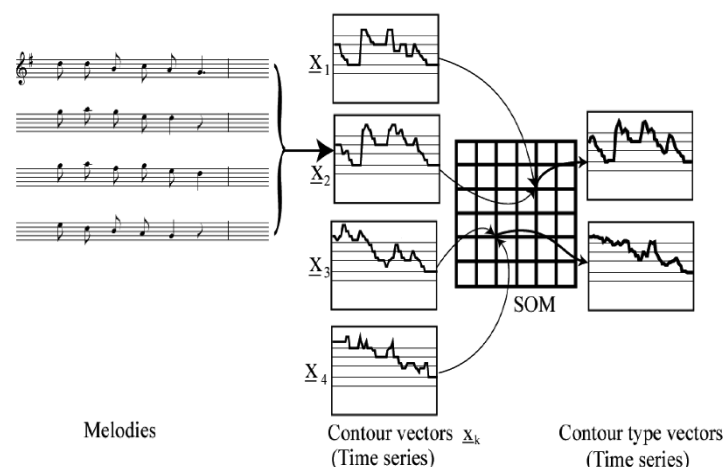
kurd, orosz (a pszkovi területről), navahó és lengyel (két kisebb régió: Warmia és Kasubia). A zenei adatbázis vizsgálatával párhuzamosan, számítottuk a megfelelő genetikai távolságmátrixokat is, a széles körben használt szoftvereket alkalmazva [113]. A fent felsorolt populációk publikált adatai alapján a következő különbségekre kellett tekintettel lenni: (1) a genetikai adatok nem voltak megtalálhatók a luxemburgi, a karacsáj és a dakota indián népesség esetében, (2) az Y-kromoszómális és mtDNS adatokat külön-külön használtunk a horvát, tatár, mari, udmurt/votják és csuvas (a Volga-régió helyett), illetve az ír és skót populációk esetében, (3) a három magyar nyelvet beszélő csoport genetikai adatait (magyar, székely és csángó) szintén külön-külön kezeltük, (4) egy lengyel populáció Y-kromoszóma és mtDNS adatait használtuk összehasonlításokra, míg három olyan lengyel népdal adatbázist, melyek Lengyelország különböző területeiről származtak.

Az Y-kromoszómális és mtDNS haplocsoportok frekvenciáin alapuló genetikai távolságokat (Fst) az Arlequin 2.0 szoftverbe beépített AMOVA programmal számítottuk [112]. Multidimenziós skálázási (MDS) ábrák készültek ViSta 7.9.2.4 szoftverrel.

Az Y-kromoszómális haplocsoportokat a következő csoportokba soroltuk be: C, D, E, F, I, J2, K, L, N1c, PxR1a és R1a.

Az anyai vonalak összehasonlítására az mtDNS haplocsoportok frekvenciáit használtuk a fent említett programcsomagok segítségével. A haplocsoportokat a következő csoportokba soroltuk: A, B, C, D, E, F, G, H, I, J, K, L, M (9Q, G, E, D, N, P, Q, R, S, T, U9K, V, W, X, Y és Z).

A népzene elemzésére kidolgoztunk egy rendszert, amely a mesterséges intelligencia egy általánosan alkalmazott típusán alapul a dallam kontúrok automatikus osztályozására - az úgynevezett önszerveződő térképet (SOM=Self Organization Map) [215]. Rendszerünk működése a **52. ábrán** látható. Az első lépés a D-dimenziós dallamvonal-vektorok (D-elemű hangmagasság-idősor) szerkesztése a kották digitális kódjaiból [216]. Mivel D mindegyik dallamra ugyanaz volt, a kontúrok összehasonlíthatók egymással egy közös Euklideszi távolság függvény segítségével ugyanabban a D-dimenziós dallamtérben, függetlenül azok tényleges egyéni hosszától.



52. ábra. D-dimenziós dallamvonal-vektorok szerkesztése a kották digitális kódjaiból

A dallamvonal-vektorokban, a dallamok fő ritmikai jellemzői 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 népdal kultúrában, a dallamoknak van egy jól definiált alaphangja (tonika), mely a záróhanggal azonos; ezért ez a technika zeneileg releváns a leggyakrabban vizsgált esetekben. Ugyanakkor ez nem igaz általánosságban, főleg a nyugati népdalok esetében, ezért ezeknek az adatbázisoknak automatikus átvétele részben szakértők által, részben további algoritmusok által korrigálható.

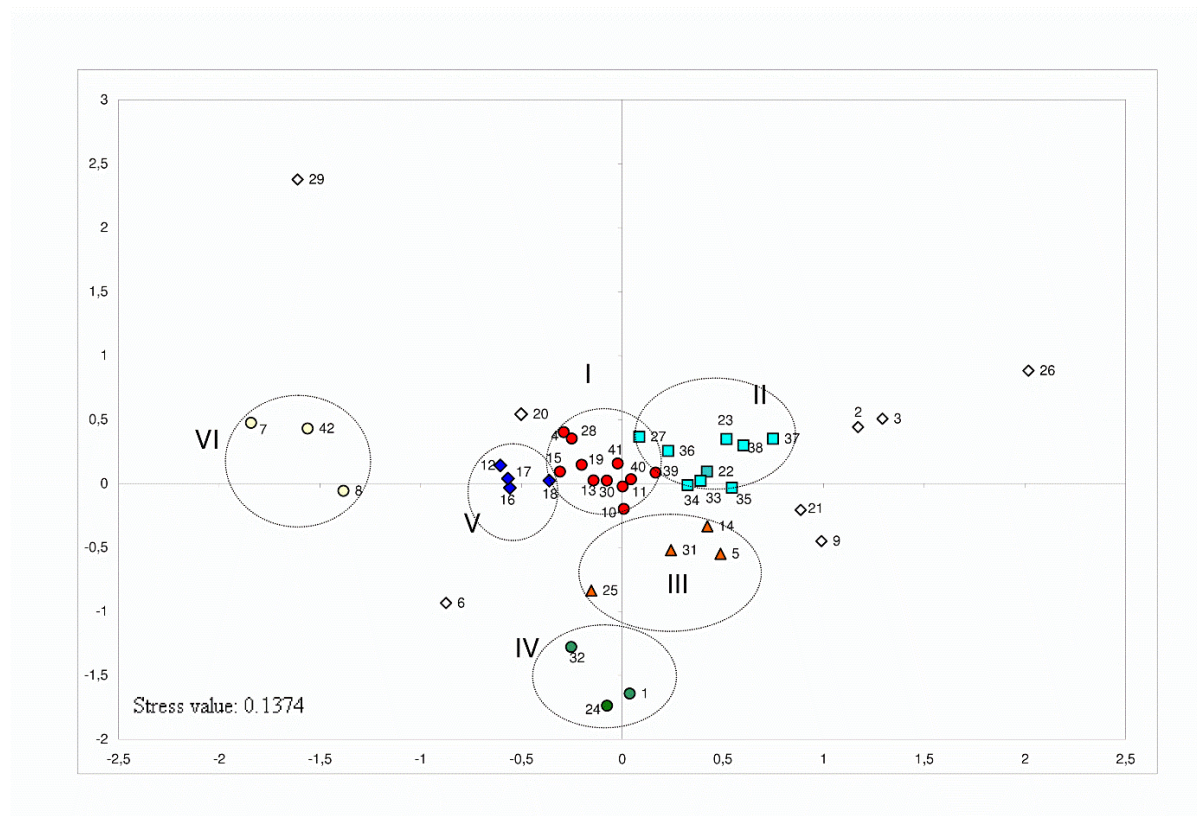
A betanítás után a SOM egy adott kultúra dallamainak kontúr vektoraival, a térkép különböző rácshelyeihez tartozó megtanult dallamvonaltípus-vektorokkal (DVT-k) képezi le az adott kultúrát jellemző legfontosabb kontúrokat [216]. Más szóval, egy jól betanított SOM DVT-készlete egy olyan zenei nyelvet 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 hasonlónak 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.

Az A és B kultúrák közötti általános összefüggés mértékét azok DVT-éinek átfedései jellemzik. Pontosabban, az i -edik és az j -edik kultúrák kapcsolatának mértékét, a $d_{i,j} = 1 - \frac{n_{i,j}}{N_i}$ képlettel számítjuk, ahol $n_{i,j}$ az i -edik kultúrában a j -edik kultúra DVT-i által aktivált DVT-k száma, N_i pedig az j -edik kultúra kontúrtípusainak teljes száma. Mivel a kapcsolat nem szimmetrikus ($d_{i,j} \neq d_{j,i}$), definíciónk nem felel meg a matematikai értelemben vett távolságfüggvény kritériumainak, de megállapíthatjuk, hogy minél közelebb van két kultúra egymáshoz, annál kisebbek a megfelelő $d_{i,j}$ és $d_{j,i}$ értékek. Így bármely zenei kultúra viszonya a többihez egy 31 dimenziós vektorral jellemezhető, mely a fent definiált $d_{i,j}$ értékeket tartalmazza. Annak érdekében, hogy világos képet kapjunk a nem szimmetrikus adatok által leírt, nagyon bonyolult kapcsolatrendszeréről, a többdimenziós skálázás (MDS) módszer speciális verzióját alkalmaztuk [217]. Ahogy korábban szó volt róla, ez a technika jól ismert a genetikában. Ezért vagyunk képesek feltárni a zenei kultúrák közötti kapcsolatrendszerek titkát ugyanazon módszer alkalmazásával, amelyet a filogenetikai vizsgálatok már régóta használtak.

5.4.1. A populációk genetikai kapcsolatai

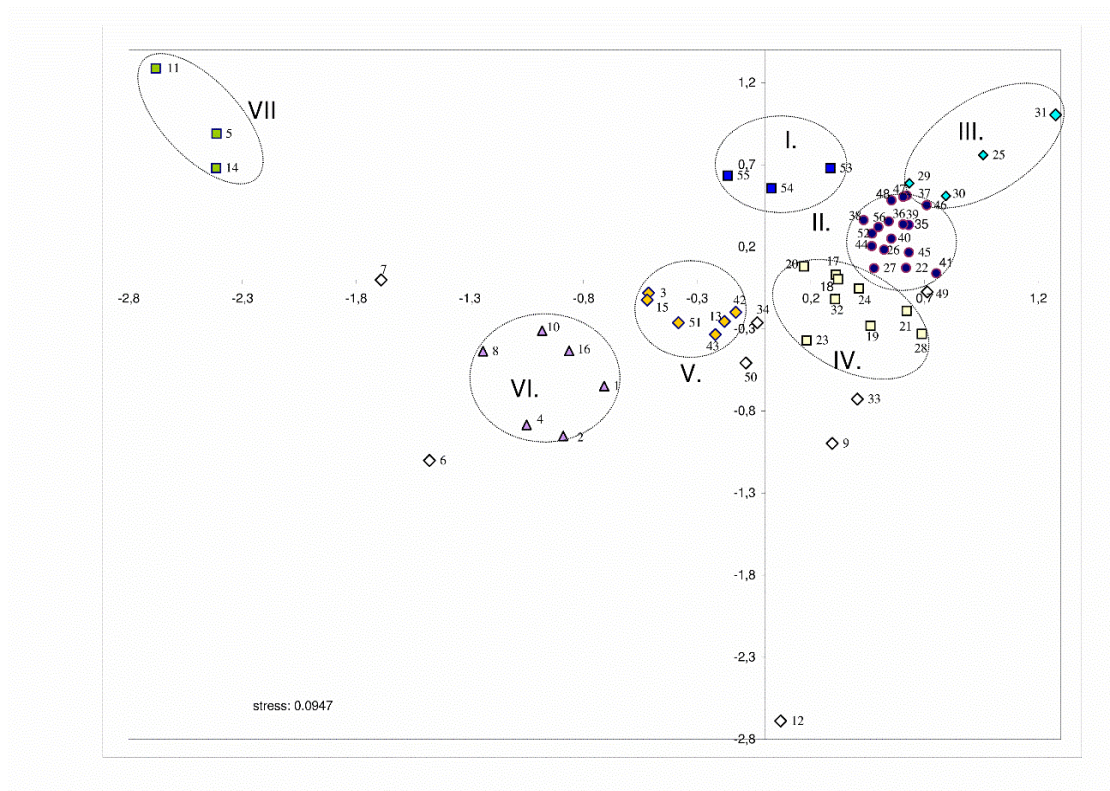
A férfiak által közvetített genetikai örökség feltárására az Y-kromoszómális haplocsoportok gyakoriságai alapján számolt genetikai távolságokat (Fst) hasonlítottuk össze 42 populáció esetében és azokat nem-metrikus MDS térképen (53. ábra) jelenítettük meg (az adatok/cikkek hivatkozását ld. Pamjav és mtsai, 2012 cikkében [218]). Az összehasonlított populációk közül a legtöbb az ábra közepén található négy klaszter valamelyikébe sorolható, két kisebb klaszter pedig az ábra szélén helyezkedik el, azonban néhány populáció kívül állónak tűnik. Az egy klaszterbe tartozó populációkat úgy határoztuk meg, hogy a tagok közötti genetikai távolság kisebb legyen 0,05-nél, valamint a lehető legtöbb populáció kerüljön egy-egy klaszterbe. Ilyen elv alapján az I. klaszterbe 3 magyar (magyar, székely és csángó), két kelet-európai (bolgár és szerb), két tatár (krími és kazanyi), két kurd, egy szicíliai és egy görög populáció került be. A legtöbb nyugat-európai (német, holland, spanyol, norvég és olasz) és három közép-európai (cseh, horvát és szlovák) populáció a II. klaszterbe került. A III. klaszter 3 populációt tartalmazott a Volga-régióból (udmurt, csuvas és manysi) és egy orosz. A IV. klaszterbe három finnugor populáció (mari, hanti és finn) csoportosult. A 2 azeri, a török és a csecsen populáció az V. klaszterbe került. A VI. klaszterbe 3 ázsiai populáció (evenki, belső- és külső-mongol) jutott.



53. ábra. Az Y haplocsoportok F_{st} genetikai távolságai alapján szerkesztett MDS térkép

1 mari; 2 skót; 3 ír; 4 bolgár; 5 udmurt; 6 kínai; 7 evenki (összevont); 8 belső mongol; 9 kirgiz; 10 kazáni tatár; 11 krimi tatár; 12 azeri; 13 kurd; 14 orosz; 15 szicíliai; 16 azeri2; 17 csecsen; 18 török; 19 görög; 20 román; 21 lengyel; 22 német; 23 spanyol; 24 hanti; 25 manyisi; 26 navahó; 27 horvát (összevont); 28 szerb; 29 ajnu; 30 kurd2; 31 csuvas; 32 finn; 33 norvég; 34 szlovák; 35 cseh; 36 francia; 37 holland; 38 olasz; 39 magyar; 40 székely; 41 csángó; 42 külső-mongol populáció (saját adat). Ha egy populáció többször szerepel, akkor különböző régióból vagy más publikációból származnak az adatok.

Hogy a fenti megállapításokat a nők által közvetített genetikai örökség kontextusába helyezzük, kiszámítottuk az mtDNS haplocsoportok genetikai távolságait a publikált forrásokból származó 56 populáció adatait használva és azokat ismét MDS-diagramként (**54. ábra**) mutattuk be (az adatok/cikkek hivatkozását ld. Pamjav és mtsai, 2012 cikkében [218]). A populációk klaszterbe sorolása azonos elveken alapult, mint a férfiaknál, vagyis a közeli rokonnak tekinthető populációk genetikai távolságainak kisebbnek kellett lenniük 0,05-nél.



54. ábra. Az mtDNS haplocsoportok F_{st} genetikai távolságai alapján szerkesztett MDS térkép

1 külső-mongol; 2 kínai; 3 kirgiz; 4 belső-mongol; 5 evenki (Köves Tunguszka); 6 evenki2 (Iengra); 7 evenki3 (nyukhaza); 8 evenki4 (jakutul beszélő); 9 krími tatár; 10 kirgiz2; 11 evenki5; 12 navahó; 13 evenki6; 14 manysi; 15 kirgiz3; 16 külső-mongol2; 17 azeri; 18 tatár; 19 csuvas; 20 komi-permják; 21 komi-zürjén; 22 mari; 23 udmurt; 24 bolgár; 25 francia; 26 német; 27 finn; 28 olasz (toszkán); 29 szicíliai; 30 norvég; 31 spanyol (közép); 32 török; 33 kurd (Irán); 34 kurd (Türkmenisztán); 35 szlovák; 36 cseh; 37 görög; 38 román; 39 lengyel; 40 orosz; 41 finn2; 42 kondai manysi; 43 manysi (Lyamin); 44 norvég2; 45 orosz2; 46 német2; 47 ír; 48 skót; 49 spanyol2 (Kantábria); 50 hanti; 51 manysi3; 52 horvát; 53 magyar; 54 székely; 55 csángó; 56 holland populáció. Ha egy populáció többször szerepel, akkor különböző régióból vagy más publikációból származnak az adatok.

Itt az I. klaszterben három magyar nyelvű népcsoport (magyar, székely és csángó) csoportosult össze. A legtöbb európai populáció egy meglehetősen kompakt klasztert alkot az ábrán (II. Klaszter). A finn-ugor nyelvű populációk közül 2 is (finn és mari) a II. klaszterbe került 14 más európai populációval együtt. Négy másik európai populáció (szicíliai, norvég, francia és spanyol) együttesen képezte a III. klasztert. A IV. klaszter 3 volga-vidéki (csuvas, udmurt és tatár), 2 komi (Urál), azeri (Kaukázus), török (Anatólia) és 2 európai populációt (olasz és bolgár) tartalmaz. Az V. klaszterbe 3 manysi (Urál), 2 kirgiz (Közép-Ázsia) és egy evenki (Szibéria) populáció tartozik, míg a VI. klaszter 6 ázsiai populációból (3 mongol, evenki, kirgiz és kínai populáció) állt. A VII. klaszter tartalmazta a 2 evenki (Szibéria) és a manysi populációt az Urálból. Az összehasonlított 56 populáció közül a maradék nem

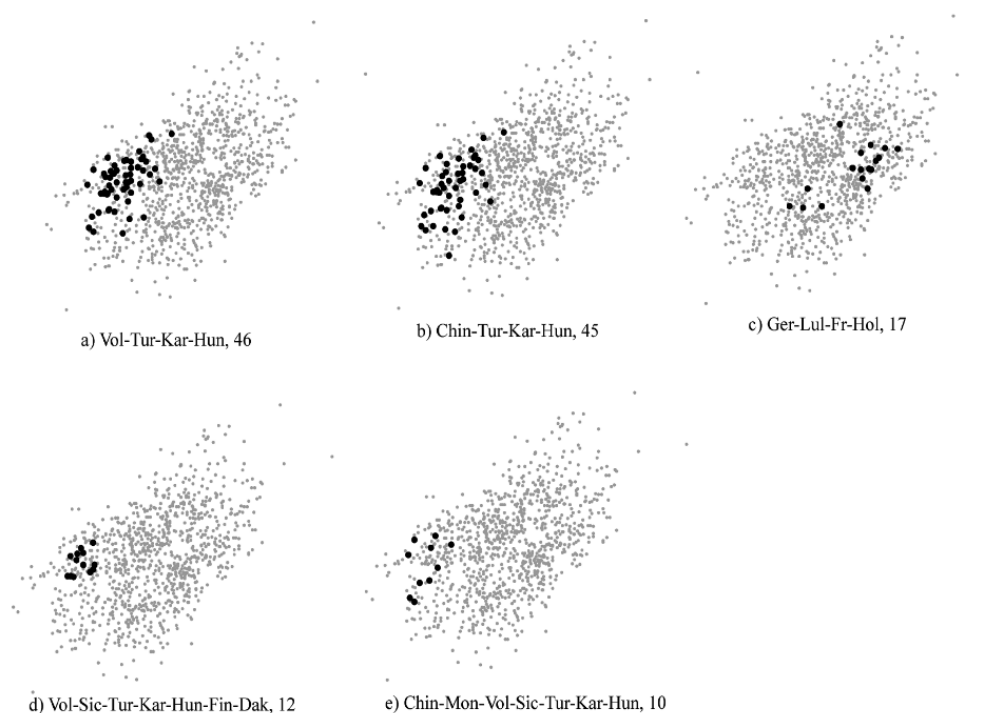
arra utal, hogy az algoritmus megtalált egy optimális elrendezést, de a bonyolult kapcsolatrendszer kétdimenziós térképezése nem tükrözheti pontosan a tényleges viszonyokat. A gráf szerkezete mégis valóban tükrözi a következő főbb tulajdonságokat: a kínai-mongol-volga (IV) csoport nagyon leegyszerűsített képe utalhat a nagy ámbitus (hangterjedelem) és a pentaton skála dominanciájára. A csoport ereszkedő kontúrjai szoros kapcsolatban állnak a dakota, anatóliai török, karacsáj és szicíliai (V) népzenevel, bár az utóbbiak inkább diatonikusak.

A török zene másik rétege, amely diatonikus dallamokat tartalmaz alacsony ámbitusokkal, szoros román és spanyol kapcsolatokkal rendelkezik. Másrésről a plagális melódiákat tartalmazó spanyol korpusz további jelentős rétege szoros francia és finn kapcsolatokat mutat (ld. **Fr** és **Fin** csomópontokat az **55. ábrán**).

A közép-európai csoport (III: magyar-szlovák-nagy-lengyel korpusz) szoros kapcsolatban áll szinte valamennyi említett zenei családdal, köszönhetően a magyar népzene különösen kiterjedt kapcsolatainak. A magyar csomópont központi helyét a közeli kínai, volga-vidéki, karacsáj, szicíliai, török, román, spanyol, lengyel, norvég, finn, ír-skót-angol és szlovák zenei kapcsolatok határozzák meg, ám ez az optimális pozíció még mindig meglehetősen hosszú élekkel kapcsolódik a rokon kultúrákhoz a kétdimenziós megközelítés korlátai miatt.

Idáig igyekeztünk arra törekedni, hogy azonosítsuk, a zenei kultúrák mely csoportjai hordoznak jelentős mennyiségű közös dallamot. A komplementer megközelítés az, hogy találjuk meg a 31 kultúra egy közös részhalmazában 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 az **56. ábra** mutatja, ahol az egyesített DVT vektorok a sík (plane) pontjaihoz vannak hozzárendelve, és az eredményként kapott pontrendszert 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 részhalmazait kerestük, melyek 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 típusokat jelölő fekete pontok a dallam térkép jól körülhatárolt területén helyezkednek el (**56a. és 56b. ábra**), következésképpen egy jól definiált zenei stílushoz tartozó folyamatos variációképződés eredményei.



56. ábra. A dallamok SOM térképe

A keleti népzene hasonló dallamai a zenei térkép bal oldalán, a nyugati népzene hasonló dallamai pedig a zenei térkép jobb oldalán csoportosulnak.

Összehasonlításként négy nyugati kultúra közös dallamtípusait is bemutatjuk (**56c. ábra**). Ezek a típusok az egyesített felhő teljesen más területein helyezkednek el, így teljesen más zenei formákat jelentenek.

A nyugati típusok maximális közös eleme (17) azt mutatja, hogy a nyugat-európai zenei mag sokkal kisebb, mint a fent említett csoportoké.

A legtöbb közös DVT-t tartalmazó kultúrák hétagú csoportjainak keresése során ugyancsak a 4-elemű keresés alapkultúrái találhatók meg, kiegészülve a szicíliai, finn és dakota, ill. a mongol, volga-vidéki és szicíliai kultúrákkal (**56d. ábra**). A fenti kultúrák genetikai távolságadatait a **17. táblázatban** foglaljuk össze, ahol a szoros kapcsolatokat félkövér betűkkel kiemeltük.

17. táblázat. Genetikai távolságok (Fst) a női és férfi leszármazási vonalakra

	Fst values (females)											
	Chinese	Mongolian	Mari	Udmurt	Tatar	Chuvash	Sicilian	Turkish	Hungarian	Szekler	Csango	Finnish
Chinese	-	0.012	0.195	0.137	0.150	0.169	0.214	0.139	0.189	0.158	0.153	0.169
Mongolian	0.322	-	0.147	0.098	0.104	0.117	0.175	0.090	0.159	0.127	0.127	0.124
Mari	0.346	0.458	-	0.053	0.007	0.012	0.024	0.017	0.054	0.064	0.095	0.003
Udmurt	0.272	0.394	0.165	-	0.025	0.036	0.109	0.025	0.086	0.081	0.110	0.047
Tatar	0.187	0.303	0.223	0.059	-	0.003	0.040	-0.002	0.051	0.052	0.068	0.004
Chuvash	0.197	0.372	0.133	-0.015	0.001	-	0.079	0.005	0.099	0.094	0.099	0.013
Sicilian	0.191	0.246	0.309	0.189	0.056	0.136	-	0.054	0.023	0.040	0.065	0.022
Turkish	0.164	0.247	0.299	0.198	0.049	0.126	0.004	-	0.060	0.049	0.060	0.012
Hungarian	0.222	0.314	0.291	0.105	0.011	0.071	0.074	0.080	-	0.007	0.043	0.045
Szekler	0.201	0.279	0.260	0.091	0.001	0.053	0.022	0.031	0.011	-	0.028	0.049
Csango	0.272	0.318	0.325	0.144	0.012	0.096	0.053	0.055	0.005	0.010	-	0.064
Finnish	0.324	0.407	0.048	0.174	0.147	0.123	0.255	0.247	0.226	0.200	0.235	-

A kínai, mongol, volga-régiói (mari, udmurt, tatár, csuvas), török, 3 magyar (magyar, székely és csángó), szicíliai, és finn populációk közötti genetikai távolságok találhatóak összesítve a **17. táblázatban**. A vastag betűvel szedett Fst távolságok kisebbek 0,05-nél.

A vizsgált populációk közötti genetikai kapcsolatok feltárásához Fst genetikai távolságokat alkalmaztunk és a távolságmátrix általános szerkezetét MDS-diagramokon ábrázoltuk. A populációk klaszterezése, illetve a populációk genetikai struktúrájának hasonlósága gyakran azok közös eredetét vagy a génáramlást tükrözi. Mind az Y-kromoszómális, mind az mtDNS MDS ábrákat a genetikai adatok függvényében közel azonos számú populációból állítottuk elő. A vizsgált populációk Y-kromoszómális és mtDNS MDS-diagramjai hat és hét különálló klasztert alkottak az egymáshoz viszonyított közeli Fst távolságok alapján (**53. és 54. ábra**). Az egy-egy klasztert képező populációk többségének genetikai közelsége magyarázható azok földrajzi közelségével és/vagy közös történelmi múltjukkal.

A populációk közötti MDS alapú genetikai kapcsolatok részletes tárgyalása mellőzhető, mivel egyrészt ennek ismerete a populációgenetikusok számára evidens, másrészt az eredeti cikkben bővebben olvasható [218]. Ezért a következő alfejezetben csak a zene és a genetika közötti kapcsolatról lesz szó.

5.4.3. Zenei és genetikai korreláció

A **55. ábrán** látható vonalak (élek) zenei és genetikai hasonlóságokra utalnak. Folytonos fekete vonalak jelzik az egyidejűleg szoros zenei és genetikai viszonyt a megfelelő populációk között; míg a hosszan szaggatott vonalak összekapcsolják azokat a populációkat, amelyek szoros zenei kapcsolattal rendelkeznek szoros genetikai kapcsolat nélkül. A röviden szaggatott vonalas kapcsolatok jelzik azokat az eseteket, ahol nem található genetikai adatok a kapcsolódó népekről. A dakota, a karacsáj és a luxemburgi populációs adatok nem álltak rendelkezésre sem mitokondriális, sem férfi vonalakon. Annak érdekében, hogy minél részletesebb képet kapjunk a zenei kapcsolatokról, nem zártuk ki ezeket a kultúrákat a zenei elemzésből, mivel alapvetően azokra a kultúrákra koncentráltunk, amelyekről egyszerre álltak rendelkezésre zenei és genetikai adatok, ezért a hiányos adatok nem befolyásolják az elemzés pontosságát.

Bizonyos esetekben több genetikai adat vonatkozott ugyanarra a zenei gyűjteményre. Ezt illusztrálják például a mari, csuvas, tatár és udmurt genetikai távolsági adatok, amelyeknek egységes népzenei gyűjteményét az ún. "volgai" tartalmazza. Például a "volgai" zenei kultúra felhasználásával a török népesség genetikai távolságai 0,3; 0,2; 0,049 és 0,13 a mari, udmurt, tatár és csuvas Y-kromoszómális adatokra nézve. Ezen adatok közül három távolság is nagyobb 0,05-nél, csak a török-tatár genetikai távolság (0,049) teljesíti kritériumunkat; ezért azt mondhatjuk, hogy a török és tatár zenei kultúra szoros kapcsolata a török és tatár népek szoros genetikai kapcsolatára is visszavezethető. Ráadásul a török és a tatár populációk mitokondriális genetikai távolsága csupán 0,002, ezért ez utóbbi adat elfogadható az adott populációk közötti jellegzetes távolságként (ld. a **17. táblázatot**). Ennek a megfontolásnak köszönhetően az egyidejűleg szoros zenei és a genetikai kapcsolatokat a **55. ábrán** a török és a volgai csomópontokat is összekötő folytonos fekete vonal jelzi.

Általánosságban azt mondhatjuk, hogy két populáció zenei és genetikai jellemzői közötti kapcsolat korábbi fizikai és biológiai kapcsolatot jelez; ezért a zenei kapcsolat visszavezethető az adott népcsoportok őseinek feltételezett kulturális kölcsönhatására vagy együttműködésére. Ezért esetünkben a szoros genetikai összefüggés fogalmát az alábbiak szerint adjuk meg: egy zenei kultúrához tartozó populációpár közötti szoros genetikai összefüggést úgy határozzuk meg, hogy a megfelelő populációpárok között az mtDNS illetve az Y-kromoszómális haplocsoport eloszlások alapján számított genetikai távolságok közül legalább az egyiknek kisebbnek kell lennie a kritikus küszöbértéknél, azaz 0,05-nél. Mivel a genetikai távolság definíció szerint 0 és 1 közötti értéket vehet fel, a 0,05-ös küszöbérték

kiválasztása meglehetősen szigorú követelmény egy szoros genetikai kapcsolat definiálásra. Ez a követelmény az Y-kromoszómális és a mitokondriális adatoknak csupán 37% -ára teljesül. Összeszámoltuk azon esetek számát, ahol szoros kapcsolatot észleltünk:

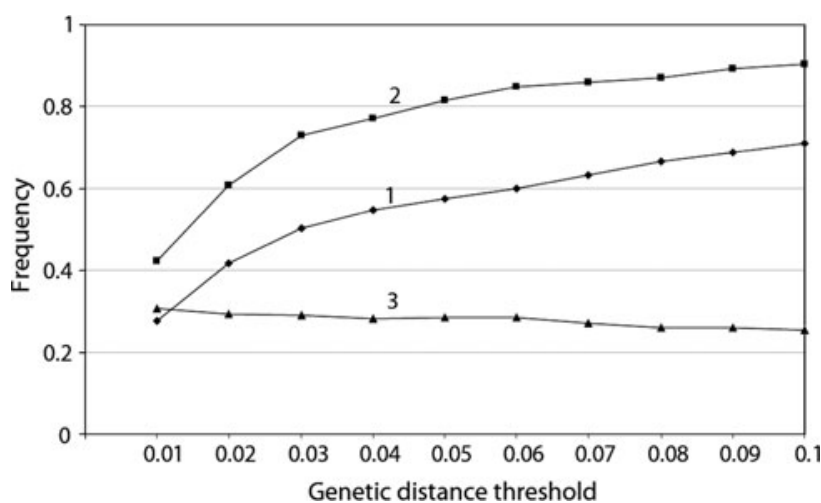
1. genetikai (G);
2. zenei (M);
3. mind a zenei, mind a genetikai (X) távolságok tekintetében.

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 valószínűsége szerint a populációk szoros genetikai kapcsolata nem feltétlenül megfelelő mutató a hasonló zenei kultúrák 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 57. ábra mutatja a zenei és a genetikai kapcsolatok korrelációjának függését a genetikai távolság küszöbértékének függvényében. Ezen az ábrán a 2. görbe mutatja a $p(g | m) = X / M$ feltételes valószínűség változását a genetikai távolság küszöbértékének függvényében. Más szavakkal, a 2. görbe mutatja a közeli genetikai kapcsolatok gyakoriságának változását a szorosan kapcsolódó zenei kultúrák körében.



57. ábra. A zenei és a genetikai kapcsolatok korrelációja

A genetikai és zenei kapcsolatok teljes függetlensége esetén ez a görbe (2. görbe) lényegében megegyezne azzal a görbével, amely a szoros genetikai kapcsolatok arányát mutatja a vizsgált 31 kultúra egészében. Azonban, az utóbbi frekvenciát mutató 1. görbe szignifikánsan és következetesen alacsonyabban fut a 2. görbénél. Ez a megfigyelés igazolja, hogy zenei-genetikai korreláció, melyet a mi történetesen 0,05-ös genetikai küszöbértékkel határoztunk meg, valójában független a küszöbérték választásától.

A 3. görbe a $p(m|g) = X/G$ valószínűségét mutatja a genetikai távolság küszöbértékének függvényeként (a zenei kapcsolatok gyakorisága a genetikailag szoros kultúrákon belül). A fenti két görbével ellentétben ez a görbe (3. görbe) enyhén csökkenő tendenciát mutat, ezért azt mondhatjuk, hogy a szoros genetikai kapcsolatok alapján nem lehet megjósolni szoros zenei kapcsolatokat, és ez szintén független a genetikai küszöbértéktől. Összehasonlítva az előző ábrákat (55., 56. és 57. ábrák) látható, hogy a genetikai és a zenei klaszterek között csak részleges megfelelés (átfedés) található. Például a közeli bolgár-azeri, szerb-kurd, magyar-tatár, ír-holland, csuvas-udmurt és finn-mari zenei kapcsolatok valóban közös genetikai klasztereket alkotó kultúrákat kapcsolnak össze (55. ábra), de a török-magyar, mongol-csuvas, finn-skót, stb. zenei kapcsolatok egyáltalán nem illeszkednek bele ebbe a rendszerbe. Ez a tény az aszimmetrikus feltételek egy természetes következménye, amelyeket a fent említett aszimmetrikus feltételes valószínűségi értékek is kifejeznek. Ugyanakkor a szoros zenei kapcsolatok alapján szoros genetikai kapcsolat jósolható, amelyet a 55. ábrán látható folytonos fekete vonalak dominanciája mutat.

Ez a zenei elemzés mutatja azt, hogy kölcsönösen intenzív érintkezés van e kultúrákon belül, de a nagy földrajzi távolságok kérdésessé 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. Ezeknek az adatoknak a fényében 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 55. ábrán látható összekötő vonalak a három kultúrában rendkívül sok egymást átfedő dallam típusra utalnak. Ezen túlmenően a példák azt is mutatják, hogy ezek az átfedések egyaránt jelentős mennyiségű közös zenei típust fednek le, amelyek mindhárom kultúrában egyszerre jelentkezők. 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 az 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 (DVT-k kisebb-nagyobb csoportjai), 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. Az ezekre az izgalmas kérdésekre adott válasz az **56. ábrán** látható, ahol kimutatható, hogy a teljesen közös melódiák legnagyobb részhalma a volga-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, számos különböző DVT-t bizonyos közös tulajdonságokkal, például a legtöbbjüknek ereszkedő kontúrja van egy oktáv hangterjedelemmel.

A 31 kultúra nagyobb részhalma 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 hét elemes csoportjaiban a legelterjedtebb dallamtípusok száma 12 és 10 között van. A dakotai és karacsáj genetikai adatok hiánya mellett is az előbbi csoport szoros zenei kapcsolatai teljesen korrelálnak az ismert genetikai adatokkal (**17. táblázat**).

Sajnálatos, hogy abban az időben a karacsáj genetikai adatok nem álltak rendelkezésre, azonban a szomszédos csecsen népekre számolt genetikai távolságok nagy része szoros genetikai kapcsolatot mutat a magyar, török, szicíliai és tatár (volga) populációkkal. Figyelembe véve a második szeptettet, nagy genetikai távolságok mutatkoznak a nagyon kompakt volgai-szicíliai-török-magyar csoport tagjaitól. Így az **56. ábrán** látható **b** és **e**

csoportok biztosan nem teljesen koherensek genetikai értelemben; remélhetőleg ez az anomália néhány jövőbeli elemzéssel magyarázható lesz.

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 mutatnak (56c. ábra). 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 (56. ábra).

Alig lehet elképzelni másfajta magyarázatot az ilyen közös zenei stílusra, mint egy közös "zenei ősn nyelv" 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 magyar népzene kapcsolatairól szóló korábbi tanulmányok arra a következtetésre jutottak, hogy a magyar, valamint mari, csuvas [219], anatóliai [220, 221, 222, 223] és mongol [224] népzenei dallamok legfontosabb közös része a nagy kiterjedésű, néha kvintváltó dallamok csoportja. Eredményeink megerősítik, hogy a független tanulmányok által kimutatott összefüggések nem véletlenszerűek, mivel a "keleti" kultúrák közös magjának létezéséből levezethetők. A fenti eredmények szerint a jelenlegi kultúrában megőrződött zenei csoportok nagy része közös genetikai gyökerekkel rendelkezik; ennek következtében a zenei kultúra fejlődése a nagyon régi időkben élt közös őseinknek tulajdonítható.

Összefoglalva, 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 visszavezethetők a vizsgált kultúrák 82% -ában genetikai alapokra, míg a populációk szoros genetikai kapcsolatai nem feltétlenül megfelelő indikátorai hasonló zenei kultúráknak (csupán 28%-ban).

A kapott eredmény könnyen elfogadható a közelmúltban ugyanazon a területen élők esetében, mint a nyugati 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 legszélesebb közös (ősi) zenei stílust, amely számos különböző kultúrában egyszerre jelenik meg, hasonló dallam formák 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 [225, 226, 227], 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) [228], valamint az emberek és állatok kreativitásával [228, 229] kapcsolatban is hasonló kutatásba fogtak. Reméljük, hogy egy újonnan keletkező interdiszciplináris tudomány területnek a magjait ültettük el, amik idővel ki is hajtanak...

6. IRODALOMJEGYZÉK

1. Jeffreys AJ, Wilson V, Thein SL. Hypervariable minisatellite regions in human DNA. *Nature* 314:67-72 (1985a).
2. Jeffreys AJ, Wilson V, Thein SL (1985b) Individual-specific "fingerprints" of human DNA. *Nature* 316:76-79 (1985b).
3. Butler JM. Forensic DNA typing. Biology technology, and genetics of STR markers. Second edition. Elsevier Academic Press, ISBN 0-12-147952-8 (2005).
4. Butler JM. Fundamentals of forensic DNA typing. First Edition. Elsevier Academic Press, ISBN 978-0-12-374999-4 (2010).
5. Valenzuela RK, Henderson MS, Walsh MH, Garrison NA, Kelch JT, Cohen-Barak O, Erickson DT, John Meaney F, Bruce Walsh J, Cheng KC, Ito S, Wakamatsu K, Frudakis T, Thomas M, Brilliant MH. Predicting phenotype from genotype: normal pigmentation. *J Forensic Sci.*55(2):315-22 (2010).
6. Walsh S, Liu F, Ballantyne KN, van Oven M, Lao O, Kayser M. IrisPlex: a sensitive DNA tool for accurate prediction of blue and brown eye colour in the absence of ancestry information. *Forensic Sci Int Genet.* 5(3):170-80 (2011).
7. Walsh S, Chaitanya L, Breslin K, Muralidharan C, Bronikowska A, Pospiech E, Koller J, Kovatsi L, Wollstein A, Branicki W, Liu F, Kayser M. Global skin colour prediction from DNA. *Hum Genet.* 136(7):847-863 (2017).
8. Pádár Zs. Kutya eredetű anyagmaradványok igazságügyi genetikai vizsgálata. Szent István Egyetem, Állatorvos-tudományi Doktori Iskola (2006).
9. Cavalli-Sforza LL, Menozzi P, Piazza A. The History and Geography of Human Genes. Princeton University Press, ISBN 0691029059 (1994).
10. Evett IW, Weir BS. Interpreting DNA evidence. Sinauer Associates, Inc. Publishers, Sunderland, Massachusetts (1998).
11. Gusmão L, Butler JM, Carracedo A, Gill P, Kayser M, Mayr WR, Morling N, Prinz M, Roewer L, Tyler-Smith C, Schneider PM. DNA Commission of International Society of Forensic Genetics (ISFG): an update of the recommendations on the use of Y-STRs in forensic analysis. *Forensic Sci Int* 157:187-197 (2006).
12. Gill P, Brinkmann B, d'Aloja E, Andersen J, Bär W, Carracedo A, Dupuy B, Eriksen B, Jangblad M, Johnsson V, Kloosterman AD, Lincoln P, Morling N, Rand S, Sabatier M, Scheithauer R, Schneider P, Vide MC. Considerations from the European DNA profiling group (EDNAP) concerning STR nomenclature. *Forensic Sci Int* 87:185-192 (1997).

13. Egyed B. Mitokondriális DNS és mikroszatellita polimorfizmusok igazságügyi genetikai aspektusú vizsgálata a magyar népességben, ELTE TTK Biológia doktori Iskola (2007).
14. Holland MM, Parsons TJ. Mitochondrial DNA sequence analysis – Validation and use for forensic casework. *Forensic Sci Rev* 11:21-50 (1999).
15. Horváth G, Zalán A, Kis Z, Pamjav H. A genetic study of 12 X-STR loci in the Hungarian population. *Forensic Sci Int Genet* 6:e46-e47 (2012).
16. 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):74-7 (2007).
17. 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.* 175(1):73-8 (2008).
18. Tillmar AO. Population genetic analysis of 12 X-STRs in Swedish population. *Forensic Sci Int Genet.* 6(2):e80-1. doi: 10.1016/j.fsigen (2011).
19. Szibor R, Krawczak M, Hering S, Edelmann J, Kuhlisch E, Krause D. Use of X-linked markers for forensic purposes. *Int. J. Legal Med.* 117 67–74 (2003).
20. Szibor R. X-chromosomal markers: past, present and future. *Forensic Sci.Int. Genet.* 1 93–99 (2007).
21. Tillmar AO, Mostad P, Egeland T, Lindblom B, Holmlund G, Montelius K. Analysis of linkage and linkage disequilibrium for eight X-STR markers. *Forensic Sci Int Genet.* (1):37-41. doi: 10.1016/j.fsigen (2008).
22. Pamjav H, Kugler R, Zalán A, Völgyi A, Straky Z, Endrédy P, Kozma Z. X chromosomal recombination study in three-generation families in Hungary. *Forensic Sci Int Genet.* 6(3):e95-6. doi: 10.1016/j.fsigen (2012).
23. Inturri S, Menegon A, Amoroso, Torre C, Robino C. Linkage and linkage disequilibrium analysis of X-STRs in Italian families, *Forensic Sci. Int. Genet.* 152–154 (2011).
24. Nothnagel M, Szibor R, Vollrath O, Augustin C, Edelmann J, Geppert M, Alves C, Gusmão L, Vennemann M, Hou Y, Immel UD, Inturri S, Luo H, Lutz-Bonengel S, Robino C, Roewer L, Rolf B, Sanft J, Shin KJ, Sim JE, Wiegand P, Winkler C, Krawczak M, Hering S. Collaborative genetic mapping of 12 forensic short tandem repeat (STR) loci on the human X chromosome. *Forensic Sci Int Genet.* 6(6):778-84 (2012).
25. Kosambi DD, The estimation of map distances from recombination values, *Ann. Eugen.* 12, 172–175 (1944).

26. Kling D, Dell'Amico B, Haddeland PJ, Tillmar AO. 21. Population genetic analysis of 12 X-STRs in a Somali population sample. *Forensic Sci Int Genet.* 11:e7-8. doi: 10.1016/j.fsigen (2014).
27. Mathias N, Bayes M, Tyler-Smith C. Highly informative compound haplotypes for the human Y chromosome. *Hum Mol Genet* 3:115-123 (1994).
28. Jobling MA, Pandya A, Tyler-Smith C. The Y chromosome in forensic analysis and paternity testing. *Int J Legal Med* 110:118-124 (1997).
29. Kayser M, Caglia A, Corach D, Fretwell N, Gehrig C, Graziosi G, Heidorn F, Herrmann S, Herzog B, Hidding M, Honda K, Jobling M, Krawczak M, Leim K, Meuser S, Meyer E, Oesterreich W, Pandya A, Parson W, Penacino G, Perez-Lezaun A, Piccinini A, Prinz M, Schmitt C, Schneider PM, Szibor R, Teifel-Greding J, Weichhold G, Knijff P de, Roewer L. Evaluation of Y-chromosomal STRs: a multicenter study. *Int J Legal Med* 110:125-133 (1997a).
30. Kayser M, de Knijff P, Dieltjes P, Krawczak M, Nagy M, Zerjal T, Pandya A, Tyler-Smith C, Roewer L. Applications of microsatellite-based Y chromosome haplotyping. *Electrophoresis* 18: 1602-1607 (1997b).
31. Jobling MA, Tyler-Smith C. The human Y chromosome: an evolutionary marker comes of age. *Nature* 4:598–612 (2003).
32. Zhivotovsky LA, Underhill PA, Cinnioğlu C, Kayser M, Morar B, Kivisild T, Scozzari R, Cruciani F, Destro-Bisol G, Spedini G, Chambers GK, Herrera RJ, Yong KK, Gresham D, Tournev I, Feldman MW, Kalaydjieva. The effective mutation rate at Y chromosome short tandem repeats, with application to human population-divergence time. *Am J Hum Genet* 74:50-61 (2004).
33. Peter de Knijff. Messages through bottlenecks: on the combined use of slow and fast evolving polymorphic markers on the human Y chromosome. *Am. J. Hum. Genet.* 67:1055-1061 (2000).
34. Karafet TM, Mendez FL, Meilerman MB, Underhill PA, Zegura SL, Hammer MF. New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. *Genome Res* 18:830-8 (2008).
35. 1000 Genomes Project Consortium, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA, Abecasis GR. A global reference for human genetic variation. *Nature.* 526(7571):68-74 (2015).
36. <https://www.familytreedna.com/projects.aspx>
37. https://isogg.org/tree/ISOGG_All_Papers.html

38. Anderson S, Bankier AT, Barrel BG, de Bruijn MHL, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJH, Staden R, Young IG Sequence and organization of the human mitochondrial genome. *Nature* 290:457-65 (1981).
39. Andrews RM, Kubacka I, Chinnery PE, Lightowlers RN, Turnbull DM, Howell N. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet* 23:147 (1999).
40. Behar DM, van Oven M, Rosset S, Metspalu M, Loogväli EL, Silva NM, Kivisild T, Torroni A, Villems R. A "Copernican" reassessment of the human mitochondrial DNA tree from its root. *Am J Hum Genet.* 90(4):675-84 (2012).
41. Satoh M, Kuroiwa T. Organization of multiple nucleoids and DNA molecules in mitochondria of a human cell. *Exp Cell Res* 196:137-140 (1991).
42. Cann RL, Stoneking M, Wilson AC. Mitochondrial DNA and human evolution. *Nature.* 325(6099):31-6 (1987).
43. Parsons TJ, Muniec DS, Sullivan K, Woodyatt N, Alliston-Greiner N, Wilson MR, Berry DL, Holland KA, Weedn VW, Gill P, Holland MM. A high observed substitution rate in the human mitochondrial DNA control region. *Nat Genet* 15:363 (1997).
44. Sutovsky P, Moreno RD, Ramalho-Santos J, Dominko T, Simerly C, Schatten G. Ubiquitin tag for sperm mitochondria. *Nature* 402:371-2 (1999).
45. Kivisild T, Villems R, Jorde LB, Bamshad M, Kumar S, Hedrick P, Dowling T, Stoneking M, Parsons TJ, Irwin JA Questioning evidence for recombination in human mitochondrial DNA. *Science* 288:1931a (2000).
46. Parson W, Gusmão L, Hares DR, Irwin JA, Mayr WR, Morling N, Pokorak E, Prinz M, Salas A, Schneider PM, Parsons TJ. DNA Commission of the International Society for Forensic Genetics: revised and extended guidelines for mitochondrial DNA typing. *Forensic Sci Int Genet.* 13:134-42 (2014).
47. Gill P, Ivanov PL, Kimpton C, Piercy R, Benson N, Tully G, Evett I, Hagelberg E, Sullivan K. Identification of the remains of the Romanov family by DNA analysis. *Nat Genet* 6:130 (1994).
48. Bendall KE, Sykes BC. Length heteroplasmy in the first hypervariable segment of the human mitochondrial DNA control region. *Am J Hum Genet* 57:248(1995).
49. Wallace DC, Brown MD, Lott MT. Mitochondrial DNA variation in human evolution and disease. *Gene* 238:211-30 (1999).
50. Krings M, Stone A, Schmitz RW, Krainitzki H, Stoneking M, Paabo S. Neandertal DNA sequences and the origin of modern humans. *Cell* 90:19-30 (1997).

51. Handt O, Richards M, Trommsdorf M, Kilger C, Simanainen J, Georgiev O, Bauer K, Stone A, Hedges R, Schaffner W. et al. Molecular genetic analyses of the Tyrolean Ice Man. *Science*. 264:1775–1778 (1994).
52. Excoffier L, Langaney A. Origin and differentiation of human mitochondrial DNA. *Am J Hum Genet* 44:73-85 (1989).
53. Watson E, Forster P, Richards M, Bandelt HJ. Mitochondrial footprints of human expansions in Africa. *Am J Hum Genet* 61:691-704 (1997).
54. Parson W, Dür A. EMPOP--a forensic mtDNA database. *Forensic Sci Int Genet*. 88-92 (2007).
55. Zimmermann B, Rock AW, Dur A, Parson W. Improved visibility of character conflicts in quasi-median networks with the EMPOP NETWORK software. *Croat Med J*. 55(2):115-20 (2014).
56. Stringer CB. Modern human origins: progress and prospects. *Philos Trans R Soc Lond B Biol Sci*. 357:563-579 (2002).
57. van Oven M, Kayser M. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum Mutat* 30(2):E386-E394 (2009).
58. Goedbloed M, Vermeulen M, Fang RN, Lembring M, Wollstein A, Ballantyne K, Lao O, Brauer S, Krüger C, Roewer L, Lessig R, Ploski R, Dobosz T, Henke L, Henke J, Furtado MR, Kayser M. Comprehensive mutation analysis of 17 Y-chromosomal short tandem repeat polymorphisms included in the AmpFISTR Yfiler PCR amplification kit. *Int J Legal Med*. 123(6):471-82 (2009).
59. <https://genographic.nationalgeographic.com/>
60. Juric I, Aeschbacher S, Coop G. The Strength of Selection against Neanderthal Introgression. *PLoS Genet*. 12(11) (2016).
61. Vernot B, Tucci S, Kelso J, Schraiber JG, Wolf AB, Gittelman RM, Dannemann M, Grote S, McCoy RC, Norton H, Scheinfeldt LB, Merriwether DA, Koki G, Friedlaender JS, Wakefield J, Pääbo S, Akey JM. Excavating Neandertal and Denisovan DNA from the genomes of Melanesian individuals. *Science*. 352(6282):235-9 (2016).
62. International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. *Nature*. 431(7011):931-45 (2014).
63. Mendizabal I, Lao O, Marigorta UM, Wollstein A, Gusmão L, Ferak V, Ioana M, Jordanova A, Kaneva R, Kouvatsi A, Kučinskas V, Makukh H, Metspalu A, Netea MG, de Pablo R, Pamjav H, Radojkovic D, Rolleston SJ, Sertic J, Macek M Jr, Comas D, Kayser M.

Reconstructing the population history of European Romani from genome-wide data. *Curr Biol*.22(24):2342-9 (2012).

64. <https://www.ncbi.nlm.nih.gov/grc/human/data?asm=GRCh38>
65. Semino O, Passarino G, Oefner PJ, Lin AA, Arbuzova S, Beckman LE, De Benedictis G, Francalacci P, Kouvatsi A, Limborska S, Marcikiae M, Mika A, Mika B, Primorac D, Santachiara-Benerecetti AS, Cavalli-Sforza LL, Underhill PA. The genetic legacy of Paleolithic Homo sapiens sapiens in extant Europeans: a Y chromosome perspective. *Science*. 90(5494):1155-9 (2000).
66. Bandelt HJ, Forster P, Röhl A. Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* 16: 37-48 (1999).
67. <https://yhrd.org/pages/resources/ysnps>
68. <http://phylotree.org/Y/tree/index.htm>
69. <https://www.yfull.com/tree/>
70. Caspermeyer J. New Genetic Evidence Resolves Origins of Modern Japanese. *Mol Biol Evol*. 32(7):1913 (2015).
71. Nakagome S, Sato T, Ishida H, Hanihara T, Yamaguchi T, Kimura R, Mano S, Oota H; Asian DNA Repository Consortium. Model-based verification of hypotheses on the origin of modern Japanese revisited by Bayesian inference based on genome-wide SNP data. *Mol Biol Evol*. 32(6):1533-43 (2015).
72. Seguin-Orlando A, Korneliussen TS, Sikora M, Malaspinas AS, Manica A, Moltke I, Albrechtsen A, Ko A, Margaryan A, Moiseyev V, Goebel T, Westaway M, Lambert D, Khartanovich V, Wall JD, Nigst PR, Foley RA, Lahr MM, Nielsen R, Orlando L, Willerslev E. Paleogenomics. Genomic structure in Europeans dating back at least 36,200 years. *Science*. 346(6213):1113-8 (2014).
73. Fu Q, Hajdinjak M, Moldovan OT, Constantin S, Mallick S, Skoglund P, Patterson N, Rohland N, Lazaridis I, Nickel B, Viola B, Prüfer K, Meyer M, Kelso J, Reich D, Pääbo S. An early modern human from Romania with a recent Neanderthal ancestor. *Nature*. 524(7564):216-9 (2015).
74. Allentoft ME, Sikora M, Sjögren KG, Rasmussen S, Rasmussen M, Stenderup J, Damgaard PB, Schroeder H, Ahlström T, Vinner L, Malaspinas AS, Margaryan A, Higham T, Chivall D, Lynnerup N, Harvig L, Baron J, Della Casa P, Dąbrowski P, Duffy PR, Ebel AV, Epimakhov A, Frei K, Furmanek M, Gralak T, Gromov A, Gronkiewicz S, Grupe G, Hajdu T, Jarysz R, Khartanovich V, Khokhlov A, Kiss V, Kolář J, Kriiska A, Lasak I, Longhi C, McGlynn G, Merkevicus A, Merkyte I, Metspalu M, Mkrtychyan R, Moiseyev V, Paja L, Pálfi G, Pokutta

- D, Pospieszny Ł, Price TD, Saag L, Sablin M, Shishlina N, Smrčka V, Soenov VI, Szeverényi V, Tóth G, Trifanova SV, Varul L, Vicze M, Yepiskoposyan L, Zhitenev V, Orlando L, Sicheritz-Pontén T, Brunak S, Nielsen R, Kristiansen K, Willerslev E. Population genomics of Bronze Age Eurasia. *Nature*. 522(7555):167-72 (2015).
75. Haak W, Lazaridis I, Patterson N, Rohland N, Mallick S, Llamas B, Brandt G, Nordenfelt S, Harney E, Stewardson K, Fu Q, Mittnik A, Bánffy E, Economou C, Francken M, Friederich S, Pena RG, Hallgren F, Khartanovich V, Khokhlov A, Kunst M, Kuznetsov P, Meller H, Mochalov O, Moiseyev V, Nicklisch N, Pichler SL, Risch R, Rojo Guerra MA, Roth C, Szécsényi-Nagy A, Wahl J, Meyer M, Krause J, Brown D, Anthony D, Cooper A, Alt KW, Reich D. Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature*. 522(7555):207-11 (2015).
76. Lazaridis I, Patterson N, Mittnik A, Renaud G, Mallick S, Kirsanow K, Sudmant PH, Schraiber JG, Castellano S, Lipson M, Berger B, Economou C, Bollongino R, Fu Q, Bos KI, Nordenfelt S, Li H, de Filippo C, Prüfer K, Sawyer S, Posth C, Haak W, Hallgren F, Fornander E, Rohland N, Delsate D, Francken M, Guinet JM, Wahl J, Ayodo G, Babiker HA, Bailliet G, Balanovska E, Balanovsky O, Barrantes R, Bedoya G, Ben-Ami H, Bene J, Berrada F, Bravi CM, Brisighelli F, Busby GB, Cali F, Churnosov M, Cole DE, Corach D, Damba L, van Driem G, Dryomov S, Dugoujon JM, Fedorova SA, Gallego Romero I, Gubina M, Hammer M, Henn BM, Hervig T, Hodoglugil U, Jha AR, Karachanak-Yankova S, Khusainova R, Khusnutdinova E, Kittles R, Kivisild T, Klitz W, Kučinskas V, Kushniarevich A, Laredj L, Litvinov S, Loukidis T, Mahley RW, Melegh B, Metspalu E, Molina J, Mountain J, Näkkäläjärvi K, Nesheva D, Nyambo T, Osipova L, Parik J, Platonov F, Posukh O, Romano V, Rothhammer F, Rudan I, Ruizbakiev R, Sahakyan H, Sajantila A, Salas A, Starikovskaya EB, Tarekegn A, Toncheva D, Turdikulova S, Uktveryte I, Utevska O, Vasquez R, Villena M, Voevoda M, Winkler CA, Yepiskoposyan L, Zalloua P, Zemunik T, Cooper A, Capelli C, Thomas MG, Ruiz-Linares A, Tishkoff SA, Singh L, Thangaraj K, Villems R, Comas D, Sukernik R, Metspalu M, Meyer M, Eichler EE, Burger J, Slatkin M, Pääbo S, Kelso J, Reich D, Krause J. Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature*. 513(7518):409-13 (2014).
77. Rootsi S, Magri C, Kivisild T, Benuzzi G, Help H, Bermisheva M, Kutuev I, Barac L, Pericić M, Balanovsky O, Pshenichnov A, Dion D, Grobei M, Zhivotovsky LA, Battaglia V, Achilli A, Al-Zahery N, Parik J, King R, Cinnioğlu C, Khusnutdinova E, Rudan P, Balanovska E, Scheffrahn W, Simonescu M, Brehm A, Goncalves R, Rosa A, Moisan JP, Chaventre A, Ferak V, Füredi S, Oefner PJ, Shen P, Beckman L, Mikerezi I, Terzić R, Primorac D,

- Cambon-Thomsen A, Krumina A, Torroni A, Underhill PA, Santachiara-Benerecetti AS, Villems R, Semino O. Phylogeography of Y-chromosome haplogroup I reveals distinct domains of prehistoric gene flow in Europe. *Am J Hum Genet.* 75(1):128-37(2004).
78. Jones ER, Gonzalez-Fortes G, Connell S, Siska V, Eriksson A, Martiniano R, McLaughlin RL, Gallego Llorente M, Cassidy LM, Gamba C, Meshveliani T, Bar-Yosef O, Müller W, Belfer-Cohen A, Matskevich Z, Jakeli N, Higham TF, Currat M, Lordkipanidze D, Hofreiter M, Manica A, Pinhasi R, Bradley DG. Upper Palaeolithic genomes reveal deep roots of modern Eurasians. *Nat Commun.* 6:8912 (2015).
79. Chekunova EM, Yartseva NV, Chekunov MK, Mazurkevich AN. The first results of genetic typing of local population and ancient humans in Upper Dvina region, in A. Mazurkevich, M. Polkovnikova and E. Dolbunova (eds.), *Archaeology of lake settlement IV-II mill. BC*, pp. 290-294 (2014).
80. Mathieson I, Lazaridis I, Rohland N, Mallick S, Patterson N, Roodenberg SA, Harney E, Stewardson K, Fernandes D, Novak M, Sirak K, Gamba C, Jones ER, Llamas B, Dryomov S, Pickrell J, Arsuaga JL, de Castro JM, Carbonell E, Gerritsen F, Khokhlov A, Kuznetsov P, Lozano M, Meller H, Mochalov O, Moiseyev V, Guerra MA, Roodenberg J, Vergès JM, Krause J, Cooper A, Alt KW, Brown D, Anthony D, Lalueza-Fox C, Haak W, Pinhasi R, Reich D. Genome-wide patterns of selection in 230 ancient Eurasians. *Nature.* 28(7583):499-503 (2015).
81. Szécsényi-Nagy A, Brandt G, Haak W, Keerl V, Jakucs J, Möller-Rieker S, Köhler K, Mende BG, Oross K, Marton T, Oszás A, Kiss V, Fecher M, Pálfi G, Molnár E, Sebők K, Czene A, Paluch T, Šlaus M, Novak M, Pećina-Šlaus N, Ósz B, Voicsek V, Somogyi K, Tóth G, Kromer B, Bánffy E, Alt KW. Tracing the genetic origin of Europe's first farmers reveals insights into their social organization. *Proc Biol Sci.* 282(1805) (2015).
82. Gamba C, Jones ER, Teasdale MD, McLaughlin RL, Gonzalez-Fortes G, Mattiangeli V, Domboróczki L, Kővári I, Pap I, Anders A, Whittle A, Dani J, Raczky P, Higham TF, Hofreiter M, Bradley DG, Pinhasi R. Genome flux and stasis in a five millennium transect of European prehistory. *Nat Commun.* 5:5257 (2014).
83. Szécsényi-Nagy Anna. Molecular genetic investigation of the Neolithic population history in the western Carpathian Basin, PhD dissertation (2015).
84. Lacan M, Keyser C, Ricaut FX, Brucato N, Duranthon F, Guilaine J, Crubézy E, Ludes B. Ancient DNA reveals male diffusion through the Neolithic Mediterranean route. *Proc Natl Acad Sci U S A.* 108(24):9788-91 (2011).

85. Batini C, Hallast P, Zadik D, Delser PM, Benazzo A, Ghirotto S, Arroyo-Pardo E, Cavalleri GL, de Knijff P, Dupuy BM, Eriksen HA, King TE, López de Munain A, López-Parra AM, Loutradis A, Milasin J, Novelletto A, Pamjav H, Sajantila A, Tolun A, Winney B, Jobling MA. Large-scale recent expansion of European patrilineages shown by population resequencing. *Nat Commun.* 6:7152 (2015).
86. 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):661-73 (2015).
87. <https://yhrd.org/tools/branch/R1b-L23>
88. <https://yhrd.org/tools/branch/R1b-M412>
89. Myres NM, Rootsi S, Lin AA, Järve M, King RJ, Kutuev I, Cabrera VM, Khusnutdinova EK, Pshenichnov A, Yunusbayev B, Balanovsky O, Balanovska E, Rudan P, Baldovic M, Herrera RJ, Chiaroni J, Di Cristofaro J, Villems R, Kivisild T, Underhill PA. A major Y-chromosome haplogroup R1b Holocene era founder effect in Central and Western Europe. *Eur J Hum Genet.* 19(1): 95–101 (2011).
90. Hans de Beule. The “Vikings” of I-L38, A Network Analysis of I-L38 Branches (2018).
91. Haak W, Brandt G, de Jong HN, Meyer C, Ganslmeier R, Heyd V, Hawkesworth C, Pike AWG, Meller H, Alt KW. Ancient DNA, Strontium isotopes, and osteological analyses shed light on social and kinship organization of the Later Stone Age. *Proc Natl Acad Sci U S A* 105, 18226–31 (2008).
92. Underhill PA, Poznik GD, Rootsi S, Järve M, Lin AA, Wang J, Passarelli B, Kanbar J, Myres NM, King RJ, Di Cristofaro J, Sahakyan H, Behar DM, Kushniarevich A, Sarac J, Saric T, Rudan P, Pathak AK, Chaubey G, Grugni V, Semino O, Yepiskoposyan L, Bahmanimehr A, Farjadian S, Balanovsky O, Khusnutdinova EK, Herrera RJ, Chiaroni J, Bustamante CD, Quake SR, Kivisild T, Villems R. The phylogenetic and geographic structure of Y-chromosome haplogroup R1a. *Eur J Hum Genet.* 23(1):124-31 (2015).
93. Günther T, Valdiosera C, Malmström H, Ureña I, Rodríguez-Varela R, Sverrisdóttir ÓO, Daskalaki EA, Skoglund P, Naidoo T, Svensson EM, Bermúdez de Castro JM, Carbonell E, Dunn M, Stora J, Iriarte E, Arsuaga JL, Carretero JM, Götherström A, Jakobsson M. Ancient genomes link early farmers from Atapuerca in Spain to modern-day Basques. *Proc Natl Acad Sci USA.* 112(38):11917-22 (2015).

94. Rootsi S, Myres NM, Lin AA, Järve M, King RJ, Kutuev I, Cabrera VM, Khusnutdinova EK, Varendi K, Sahakyan H, Behar DM, Khusainova R, Balanovsky O, Balanovska E, Rudan P, Yepiskoposyan L, Bahmanimehr A, Farjadian S, Kushniarevich A, Herrera RJ, Grugni V, Battaglia V, Nici C, Crobu F, Karachanak S, Kashani BH, Houshmand M, Sanati MH, Toncheva D, Lisa A, Semino O, Chiaroni J, Di Cristofaro J, Villems R, Kivisild T, Underhill PA. Distinguishing the co-ancestries of haplogroup G Y-chromosomes in the populations of Europe and the Caucasus. *Eur J Hum Genet.* 20(12): 1275–1282 (2012).
95. Kivisild T. The study of human Y chromosome variation through ancient DNA. *Hum Genet.* 136(5):529-546 (2017).
96. <https://www.worldfamilies.net/mtdnahaplogroups>
97. Comey CT, Koons BW, Presley KW, Smerick JB, Sobieralski CA, Stanley DM. DNA Extraction Strategies for Amplified Fragment Length Polymorphism Analysis. *J Forensic Sci* 39: 1254-1269 (1194).
98. Applied Biosystems. Quantifiler Kits (Quantifiler Human DNA Quantification Kit and Quantifiler Y Male DNA Quantification Kit) User's Manual (2006.04.).
99. Underhill PA, Shen P, Lin AA, Passarino G, Yang WH, Karuffman E, Bonn -Tamir B, Bertranpetit J, Francalacci P, Ibrahim M, Jenkins T, Kidd JR, Mehdi SQ, Seielstad MT, Wells RS, Piazza A, Davis RW, Feldman MW, Cavalli-Sforza LL, Oerner PJ. Y chromosome sequence variation and the history of human populations. *Nat Genet* 26:358–361(2000).
100. Underhill PA, Passarino G, Lin AA, Shen P, Miraz n Lahr M, Foley RA, Oefner PJ, Cavalli-Sforza LL. The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations. *Ann Hum Genet* 65:43–62 (2001).
101. Hammer MF, Karafet T, Rasanayagam A, Wood ET, Altheide TK, Jenkins T, Griffiths RC, Templeton AR, Zegura SL. Out of Africa and back again: nested cladistic analysis of human Y chromosome variation. *Mol Biol Evol* 15:427–441 (1998).
102. Hammer MF, Karafet TM, Redd AJ, Jarjanazi H, Santachiara-Benerecetti S, Soodyall H, Zegura SL. Hierarchical patterns of global human Y-chromosome diversity. *Mol Biol Evol* 18:1189–1203 (2001).
103. Karafet TM, Xu L, Du R, Wang W, Feng S, Wells RS, Redd AJ, Zegura SL, Hammer MF. Paternal population history of East Asia: sources, patterns, and microevolutionary processes. *Am J Hum Genet* 69:615–628 (2001).
104. Cinni glu C, King R, Kivisild T, Kalfo glu E, Atasoy S, Cavalleri GL, Lillie AS, Roseman CC, Lin AA, Prince K, Oefner PJ, Shen P, Semino O, Cavalli-Sforza LL, Underhill PA. Excavating Y-chromosome haplotype strata in Anatolia. *Hum Genet* 114:127–148 (2004).

105. Y Chromosome Consortium. A Nomenclature System for the Tree of Human Y-Chromosomal Binary Haplogroups. *Genome Res* 12:339-348 (2002).
106. Athey TW, Nordtvedt K. Resolving the placement of haplogroup I-M223 in the Y chromosome phylogenetic tree. *J Genet Geneal* 1:54–55 (2005).
107. <http://www.internationalgenome.org/>
108. Applied Biosystems. TaqMan® SNP Genotyping Assays Protocol (2006.05.).
109. Botstein D, White RL, Skolnick M, Davis RW. Construction of a genetic linkage map in man using restriction fragment length polymorphism. *Am J Hum Genet* 32:314-331(1980).
110. Desmarais D, Zhong Y, Chakraborty R, Perreault C, Busque L. Development of a highly polymorphic STR marker for identity testing purposes at the human androgen receptor gene (HUMARA). *J Forensic Sci* 43:1046–1049 (1998).
111. Krüger J, Fuhrmann W, Lichte KH, Steffens C. Zur Verwendung der sauren Erythrocytenphosphatase bei der Vaterschaftsbegutachtung. *Dtsch Z Gerichtl Med* 64:127–146 (1968).
112. Kishida T, Wang W, Fukuda M, Tamaki Y. Duplex PCR of the Y-27H39 and HPRT loci with reference to Japanese population data on the HPRT locus. *Jpn J Legal Med* 51: 67–69 (1997).
113. Schneider S, Roessli D, Excoffier L. Arlequin: a software for population genetics data analysis. Ver 2.000. Geneva: Genetics and Biometry Lab, Department of Anthropology, University of Geneva (2000).
114. Dent EA, von Holdt BM, STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour*, 4: 359-361 (2012).
115. Evanno G, Regnaut S and Goudet J, Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol*, 14: 2611-2620 (2005).
116. Chakraborty R. Sample size requirements for addressing the population genetic issues of forensic use of DNA typing. *Hum Biology*, 64: 141-159 (1992).
117. Kruskal JB, Wish M. Multidimensional Scaling, Series: Quantitative Applications in the Social Sciences. First edition. SAGE Publications ISBN 0-8039-0940-3 (1978).
118. Cavalli-Sforza LL, Bodmer WF. The genetics of human populations. San Francisco: Freeman. ISBN 0-486-40693-8 (1971).
119. Nei M. Analysis of gene diversity in subdivided populations. *Procl Natl Acad Sci USA* 70:3321-3323 (1973).
120. Wright S. The genetic structure of populations, *Ann Eugen* 15:323-354 (1951).

121. Weir BS, Cockerham CC. Genetic Data Analysis II. Sinauer Associates, Inc. Publishers, Sunderland, Massachusetts (1984).
122. Excoffier L, Smouse P, Quattro JM. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131:479-491 (1973).
123. Michalakis Y, Excoffier L. A generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. *Genetics* 142:1061-1064 (1996).
124. Ota T, Kimura M. A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. *Genet Res* 22:201-204 (1973).
125. Felsenstein J. PHYLIP - Phylogeny Inference Package (Version 3.2). *Cladistics* 5: 164-166 (1989).
126. Young FW. How to use ViSta. The Visual Statistics Project. Chapel Hill, NC, USA; Valencia, Spain (2001).
127. Drummond AJ, Suchard MA, Xie D & Rambaut A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol* **29**,1969–1973 (2012).
128. Soares P, Ermini L, Thomson N, Mormina M, Rito T, Röhl A, Salas A, Oppenheimer S, Macaulay V, Richards MB. Correcting for purifying selection: an improved human mitochondrial molecular clock. *Am J Hum Genet* 84, 740–759 (2009).
129. Fenner JN. Cross-cultural estimation of the human generation interval for use in genetics-based population divergence studies. *Am J Phys Anthropol* 128, 415–423 (2005).
130. Moorjani P, Sankararaman S, Fu Q, Przeworski M, Patterson N, Reich D. A genetic method for dating ancient genomes provides a direct estimate of human generation interval in the last 45,000 years. *Proc Natl Acad Sci USA* 113, 5652–5657 (2016).
131. Schneider PM. Expansion of the European Standard Set of DNA Database Loci—the Current Situation. www.promega.com (2009).
132. Prainsack B and Toom V. 'Performing the Union: The Prüm Decision and the European dream', *Studies in History and Philosophy of Biological and Biomedical Sciences* 44(1): 71-79 (2013).
133. McCartney C, Wilson TJ & Williams R. Transnational Exchange of Forensic DNA: Viability, Legitimacy, and Acceptability. *European Journal on Criminal Policy and Research*, 17(3), 305-322 (2011).
134. 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.* 19:18-21 (2015).

135. 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):543-4 (2011).
136. 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):555-6 (2011).
137. Ott J. *Analysis of Human Genetic Linkage*, 3rd. Ed., The John Hopkins University Press, Baltimore (1999).
138. Krawczak M. Kinship testing with X-chromosomal markers: mathematical and statistical issues. *Forensic Sci. Int. Genet.* 111-114 (2007).
139. Chen DP, Tseng CP, Tsai SH, Wang MC, Lu SC, Wu TL, Chang PY, Sun CF. Use of X-linked short tandem repeats loci to confirm mutations in parentage caseworks. *Clin Chim Acta.* 408(1-2):29-33 (2009).
140. Völgyi A, Zalán A, Szvetnik E, Pamjav H. Hungarian population data for 11 Y-STR and 49 Y-SNP markers. *Forensic Sci Int Genet.* 3(2):e27-8 (2009).
141. Rosser ZH, Zerjal T, Hurles ME, Adojaan M, Alavantic D, Amorim A, Amos W, Armenteros M, Arroyo E, Barbujani G, Beckman G, Beckman L, Bertranpetit J, Bosch E, Bradley DG, Brede G, Cooper G, Côrte-Real HB, de Knijff P, Decorte R, Dubrova YE, Evgrafov O, Gilissen A, Glisic S, Gölge M, Hill EW, Jeziorowska A, Kalaydjieva L, Kayser M, Kivisild T, Kravchenko SA, Krumina A, Kucinskas V, Lavinha J, Livshits LA, Malaspina P, Maria S, McElreavey K, Meitinger TA, Mikelsaar AV, Mitchell RJ, Nafa K, Nicholson J, Nørby S, Pandya A, Parik J, Patsalis PC, Pereira L, Peterlin B, Pielberg G, Prata MJ, Previderé C, Roewer L, Rootsi S, Rubinsztein DC, Saillard J, Santos FR, Stefanescu G, Sykes BC, Tolun A, Vilems R, Tyler-Smith C, Jobling MA. Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. *Am J Hum Genet* 67:1526-1543 (2000).
142. Bíró A, Fehér T, Bárány G, Pamjav H. Testing Central and Inner Asian admixture among contemporary Hungarians. *Forensic Sci Int Genet.* 15:121-6 (2015).
143. Pamjav H, Fóthi Á, Fehér T, Fóthi E. A study of the Bodrogeköz population in north-eastern Hungary by Y chromosomal haplotypes and haplogroups. *Mol Genet Genomics.* 292(4):883-894 (2017).
144. King RJ., Ozcan SS, Carter T, Kalfoğlu E, Atasoy S, Triantaphyllidis C, Kouvatsi A, Zhivotovsky L, Michalodimitrakis M, Underhill PA. Differential Y-chromosome Anatolian influences on the Greek and Cretan Neolithic. *Annals of Human Genetics* 72, 205–14 (2008).
145. Pamjav H, Zalán A, Béres J, Nagy M, Chang YM. Genetic structure of the paternal lineage of the Roma people. *Forensic Sci Int Genet.* 6(1): e46-7 (2012).

146. Zalán A, Béres J, Pamjav H. Paternal genetic history of the Vlax Roma. *Forensic Sci Int Genet.* 5(2):109-13 (2011).
147. Chang YM, Peruma R, Keat PY, Kuehn DLC. Haplotype diversity of 16 Ychromosomal STRs in three main ethnic populations (Malays, Chinese and Indians) in Malaysia, *Forensic Sci. Int.* 167(1): 70-76 (2007).
148. Füredi S, Woller J, Pádár Zs, Angyal M. Y-STR haplotyping in two Hungarian populations. *Int J Legal Med* 113:38–42 (1999).
149. Zaharova B, Andonova S, Gilissen A, Cassiman J, Decorte R, Kremensky I. Y-chromosomal STR haplotypes in three major population groups in Bulgaria. *Forensic Sci Int* 124:182–186 (2001).
150. Füredi S, Egyed B, Csikai M, Osztrovics A, Woller J, Padar Z. Y-STR haplotyping in seven Hungarian (speaking) populations. IV. International Forensic Y-User Workshop, Berlin, November 18–20 (2004).
151. Gusmão A, Gusmão L, Gomes V, Alves C, Calafell F, Amorim A, Prata MJ. A Perspective on the History of the Iberian Gypsies Provided by Phylogeographic Analysis of Y-Chromosome Lineages. *Annals of Human Genetics* 72:215–227 (2008).
152. Gresham D, Morar B, Underhill PA, Passarino G, Lin AA, Wise C, Angelicheva D, Calafell F, Oefner PJ, Shen P, Tournev I, de Pablo R, Kucinskas V, Perez-Lezaun A, Marushiakova E, Popov V, Kalaydjieva L. Origins and Divergence of the Roma (Gypsies). *Am J Hum Genet* 69:1314–1331(2001).
153. Perić M, Klarić IM, Lauc LB, Jančićjević B, Dordević D, Efremovska L, Rudan P. Population genetics of 8 Y chromosome STR loci in Macedonians and Macedonian Romani (Gypsy). *Forensic Sci Int* 154:257–261(2005).
154. 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* 169:19-26 (2007).
155. Petrejčikova E, Sotak M, Bernasovska J, Bernasovsky I, Sovicova A, Bozikova A, Boronova I, Svickova P, Gabrikova D, Macekova S. Y-Haplogroup frequencies in the Slovak Romany population. *Anthropological Science* 117:89-94(2010).
156. Sharma S, Rai E, Sharma P, Jena M, Singh S, Darvishi K, Bhat AK, Bhanwer AJ, Tiwari PK, Bamezai RN. The Indian origin of paternal haplogroup R1a1* substantiates the autochthonous origin of Brahmins and the caste system. *Journal of Human Genetics* 54:47–55 (2009).

157. Sengupta S, Zhivotovsky LA, King R, Mehdi SQ, Edmonds CA, Chow CE, Lin AA, Mitra M, Sil SK, Ramesh A, Usha Rani MV, Thakur CM, Cavalli-Sforza LL, Majumder PP, Underhill PA. Polarity and Temporality of High-Resolution Y-Chromosome Distributions in India identify Both Indigenous and Exogenous Expansions and Reveal Minor Genetic Influence of Central Asian Pastoralists. *The American Journal of Human Genetics* 78: 202-221(2006).
158. Majumder PP. Ethnic populations of India as seen from an evolutionary perspective. *J Biosci* 26(4 Suppl):533-545 (2001).
159. Cordaux R, Aunger R, Bentley G, Nasidze I, Sirajuddin SM, Stoneking M. Independent Origins of Indian Caste and Tribal Paternal Lineages. *Current Biology* 14:231–235(2004).
160. Thanseem I, Thangaraj K, Chaubey G, Singh VK, Bhaskar LV, Reddy BM, Reddy AG, Singh L. Genetic affinities among the lower castes and tribal groups of India: inference from Y chromosome and mitochondrial DNA. *BMC Genetics* 7: 42-52 (2006).
161. Gimbutus M. Proto-Indo-European culture: the Kurgan culture during the 5th to the 3rd millennia BC. In: Cardona G, Koenigswald HM, Senn A, editors. *Indo-European and Indo-Europeans*. Philadelphia: University of Pennsylvania Press. p155–198 (1970).
162. Wells RS, Yuldasheva N, Ruzibakiev R, Underhill PA, Evseeva I, Blue-Smith J, Jin L, Su B, Pitchappan R, Shanmugalakshmi S, Balakrishnan K, Read M, Pearson NM, Zerjal T, Webster MT, Zholoshvili I, Jamarjashvili E, Gambarov S, Nikbin B, Dostiev A, Aknazarov O, Zalloua P, Tsoy I, Kitaev M, Mirrakhimov M, Chariev A, Bodmer WF. The Eurasian Heartland: A continental perspective on Y-chromosome diversity. *Proc Natl Acad Sci USA* 98:10244-9 (2001).
163. Underhill PA, Myres NM, Rootsi S, Metspalu M, Zhivotovsky LA, King RJ, Lin AA, Chow CE, Semino O, Battaglia V, Kutuev I, Järve M, Chaubey G, Ayub Q, Mohyuddin A, Mehdi SQ, Sengupta S, Rogaev EI, Khusnutdinova EK, Pshenichnov A, Balanovsky O, Balanovska E, Jeran N, Augustin DH, Baldovic M, Herrera RJ, Thangaraj K, Singh V, Singh L, Majumder P, Rudan P, Primorac D, Villems R, Kivisild T. Separating the post-Glacial coancestry of European and Asian Y chromosomes within haplogroup R1a. *Eur J Hum Genet* 18:479-484 (2010).
164. 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):611-5 (2012).
165. Hajdú Péter. Preuráli nyelvi kapcsolatok. Magyar őstörténeti tanulmányok. Akadémiai Kiadó, Budapest (1977).
166. Hajdú-Domonkos. Uráli nyelvrokonaink. Tankönyv Kiadó, Budapest (1980).

167. Fehér T, Németh E, Vándor A, Kornienko IV, Csáji LK, Pamjav H. Y-SNP L1034: limited genetic link between Mansi and Hungarian-speaking populations. *Mol Genet Genomics*. 290(1):377-86 (2015).
168. Pimenoff VN, Comas D, Palo JU, Vershubsky G, Kozlov A, Sajantila A. Northwest Siberian Khanty and Mansi in the junction of West and East Eurasian gene pools as revealed by uniparental markers. *Eur J Hum Genet* 16(10):1254–1264 (2008).
169. Türk Attila. A magyar őstörténet és a szaltovói régészeti kultúrkör. Doktori értekezés, Szeged (2011).
170. Shi H, Qi X, Zhong H, Peng Y, Zhang X, Ma RZ, Su B. Genetic evidence of an East Asian origin and paleolithic northward migration of Y-chromosome haplogroup N. *PLoS One*. 8(6):e66102 (2013).
171. Klima L. A székelyek nevééről. <http://finnugor.elte.hu/?q=szekely> (2009).
172. Csányi B, Bogácsi-Szabó E, Tömöry G, Czibula A, Priskin K, Csősz A, Mende B, Langó P, Csete K, Zsolnai A, Conant EK, Downes CS, Raskó I. Y-chromosome analysis of ancient Hungarian and two modern Hungarian-speaking populations from the Carpathian Basin. *Ann Hum Genet*. 72(Pt 4):519-34 2008.
173. Keyser-Tracqui C, Crubézy E, Ludes B. Nuclear and mitochondrial DNA analysis of a 2,000-year-old necropolis in the Egyin Gol Valley of Mongolia. *Am J Hum Genet*. 73(2):247-60 (2003).
174. 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*. doi: 10.1007/s00438-017-1363-8 (2017).
175. Flegontov P, Changmai P, Zidkova A, Logacheva MD, Altınışık NE, Flegontova O, Gelfand MS, Gerasimov ES, Khrameeva EE, Konovalova OP, Neretina T, Nikolsky YV, Starostin G, Stepanova VV, Travinsky IV, Tříska M, Tříska P, Tatarinova TV. Genomic study of the Ket: a Paleo-Eskimo-related ethnic group with significant ancient North Eurasian ancestry. *Sci Rep*. 6:20768 (2016).
176. Grugni V, Battaglia V, Hooshiar Kashani B, Parolo S, Al-Zahery N, Achilli A, Olivieri A, Gandini F, Houshmand M, Sanati MH, Torroni A, Semino O. Ancient migratory events in the Middle East: new clues from the Y-chromosome variation of modern Iranians. *PLoS One*. 7(7):e41252 (2012).
177. Mirabal S, Varljen T, Gayden T, Regueiro M, Vujovic S, Popovic D, Djuric M, Stojkovic O and Herrera R J. Human Y-chromosome short tandem repeats: A tale of acculturation and

- migrations as mechanisms for the diffusion of agriculture in the Balkan Peninsula. *Am. J. Phys. Anthropol.* 142:380-390 (2010).
178. Martinez-Cruz B, Ioana M, Calafell F, Arauna LR, Sanz P, Ionescu R, Boengiu S, Kalaydjieva L, Pamjav H, Makukh H, Plantinga T, van der Meer JWM, Comas D, Netea MG, the Genographic Consortium. Y-Chromosome Analysis in Individuals Bearing the Basarab Name of the First Dynasty of Wallachian Kings. *PLoS One.* 7(7): e41803 (2012).
 179. Mielnik-Sikorska M, Daca P, Woźniak M, Malyarchuk BA, Bednarek J, Dobosz T, Grzybowski T. Genetic data from Y chromosome STR and SNP loci in Ukrainian population. *Forensic Sci Int Genet.* 7(1):200-3 (2013).
 180. Barać L, Perić M, Klarić IM, Rootsi S, Janićijević B, Kivisild T, Parik J, Rudan I, Villems R, Rudan P. Y chromosomal heritage of Croatian population and its island isolates. *Eur J Hum Genet.* 11(7):535-42 (2003).
 181. Balanovsky O, Dibirova K, Dybo A, Mudrak O, Frolova S, Pocheshkhova E, Haber M, Platt D, Schurr T, Haak W, Kuznetsova M, Radzhabov M, Balaganskaya O, Romanov A, Zakharova T, Soria Hernanz DF, Zalloua P, Koshel S, Ruhlen M, Renfrew C, Wells RS, Tyler-Smith C, Balanovska E; Genographic Consortium. Parallel evolution of genes and languages in the Caucasus region. *Mol Biol Evol.* 28(10):2905-20 (2011).
 182. Khar'kov VN, Stepanov VA, Medvedev OF, Spiridonova MG, Maksimova NR, Nogovitsyna AN, Puzyrev VP. The origin of Yakuts: analysis of Y-chromosome haplotypes. *Mol Biol (Mosk).* 42(2):226-37. Russian (2008).
 183. <https://www.familytreedna.com/groups/ydna-q-l-712/about/background>
 184. Kang LL, Jin T, Wu F, Ao X, Wen S, Wang C, Huang Y, Li X Li H. Y chromosomes of ancient Hunnu people and its implication on the phylogeny of East Asian linguistic families, ASHG 2013 Annual Meeting Evolutionary and Population Evolutionary and Population Genetics No Genetics, No. 2041(2013).
 185. 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):201-214 (2017).
 186. Juhász Z, Fehér T, Németh E, Pamjav H. mtDNA analysis of 174 Eurasian populations using a new iterative rank correlation method. *Mol Genet Genomics.* 291(1):493-509 (2016).
 187. Batini C, Hallast P, Vågane ÅJ, Zadik D, Eriksen 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):12086 (2017).

188. Révész L. A karosi honfoglaláskori temetők. Herman Ottó Múzeum és Magyar Nemzeti Múzeum Miskolc (1996).
189. Roostalu U, Kutuev I, Loogvali EL, Metspalu E, Tambets K, Reidla M, Khusnutdinova EK, Usanga E, Kivisild T, Villems R. Origin and expansion of haplogroup H the dominant human mitochondrial DNA lineage in west Eurasia: the Near Eastern and Caucasian perspective. *Mol Biol Evol* 24:436–448 (2007).
190. Reidla M, Kivisild T, Metspalu E, Kaldma K, Tambets K, Tolk H, Parik J, Loogväli E-L, Derenko M, Malyarchuk B. Origin and Diffusion of mtDNA Haplogroup X. *Am J Hum Genet* 73:1178–1190 (2003).
191. Brandstatter A, Egyed B, Zimmermann B, Duftner N, Padar Z, Parson W. Migration rates and genetic structure of two Hungarian ethnic groups in Transylvania Romania. *Ann Hum Genet* 71:791–803 (2007).
192. Tömöry G, Csányi B, Bogácsi-Szabó E, Kalmár T, Czibula Á, Csősz A, Priskin K, Mende B, Langó P, Downes CS, Raskó I. Comparison of maternal lineage and biogeographic analyses of ancient and modern Hungarian populations. *Am J Phys Anthropol* 134:354–368 (2007).
193. Kustár Á. A karosi honfoglaláskori temetők in: Kovács L Révész L (eds) Herman Ottó Múzeum és Magyar Nemzeti Múzeum Miskolc pp 395–402.(1996).
194. Molodin VI, Pilipenko AS, Romaschenko AG, Zhuravlev AA, Trapezov RO, Chikisheva TA. Human migrations in the southern region of the West Siberian Plain during the Bronze Age: Archaeological palaeogenetic and anthropological data in: *Population Dynamics in Prehistory and Early History: New Approaches Using Stable Isotopes and Genetics* pp 93–112 (Eds: Kaiser E, Burger J and Schier W (2012)).
195. Csősz-Mende. Archeogenetikai vizsgálatok Szeged-Kiskundorozsma, Hosszúhát lelőhely 10. századi népességén. Szeged: Móra Ferenc múzeum évkönyve. *Studia Archeologica* XII., 2011
196. Fu Q, Posth C, Hajdinjak M, Petr M, Mallick S, Fernandes D, Furtwängler A, Haak W, Meyer M, Mittnik A, Nickel B, Peltzer A, Rohland N, Slon V, Talamo S, Lazaridis I, Lipson M, Mathieson I, Schiffels S, Skoglund P, Derevianko AP, Drozdov N, Slavinsky V, Tsybankov A, Cremonesi RG, Mallegni F, Gély B, Vacca E, Morales MR, Straus LG, Neugebauer-Maresch C, Teschler-Nicola M, Constantin S, Moldovan OT, Benazzi S, Peresani M, Coppola D, Lari M, Ricci S, Ronchitelli A, Valentin F, Thevenet C, Wehrberger K, Grigorescu D, Rougier H, Crevecoeur I, Flas D, Semal P, Mannino MA, Cupillard C, Bocherens H, Conard NJ, Harvati K, Moiseyev V, Drucker DG, Svoboda J, Richards MP, Caramelli D, Pinhasi R, Kelso J, Patterson N, Krause J, Pääbo S, Reich D. The genetic history of Ice Age Europe, *Nature*, 534(7606):200-5 (2016).

197. Posth C, Renaud G, Mittnik A, Drucker DG, Rougier H, Cupillard C, Valentin F, Thevenet C, Furtwängler A, Wißing C, Francken M, Malina M, Bolus M, Lari M, Gigli E, Capecchi G, Crevecoeur I, Beauval C, Flas D, Germonpré M, van der Plicht J, Cottiaux R, Gély B, Ronchitelli A, Wehrberger K, Grigorescu D, Svoboda J, Semal P, Caramelli D, Bocherens H, Harvati K, Conard NJ, Haak W, Powell A, Krause J. Pleistocene mitochondrial genomes suggest a single major dispersal of Non-Africans and a late glacial population turnover in Europe. *Curr Biol* **26**, 827–833 (2016).
198. Richards MB, Soares P, Torroni A. Palaeogenomics: mitogenomes and migrations in Europe's past. *Curr Biol* **26**, R243–246 (2016).
199. Brandt G, Haak W, Adler CJ, Roth C, Szécsényi-Nagy A, Karimnia S, Möller-Rieker S, Meller H, Ganslmeier R, Friederich S, Dresely V, Nicklisch N, Pickrell JK, Sirocko F, Reich D, Cooper A, Alt KW; Genographic Consortium. Ancient DNA reveals key stages in the formation of central European mitochondrial genetic diversity. *Science* **342**, 257–261 (2013).
200. Poznik GD, Xue Y, Mendez FL, Willems TF, Massaia A, Wilson Sayres MA, Ayub Q, McCarthy SA, Narechania A, Kashin S, Chen Y, Banerjee R, Rodriguez-Flores JL, Cerezo M, Shao H, Gymrek M, Malhotra A, Louzada S, Desalle R, Ritchie GR, Cerveira E, Fitzgerald TW, Garrison E, Marcketta A, Mittelman D, Romanovitch M, Zhang C, Zheng-Bradley X, Abecasis GR, McCarroll SA, Flicek P, Underhill PA, Coin L, Zerbino DR, Yang F, Lee C, Clarke L, Auton A, Erlich Y, Handsaker RE; 1000 Genomes Project Consortium, Bustamante CD, Tyler-Smith C. Punctuated bursts in human male demography inferred from 1,244 worldwide Y-chromosome sequences. *Nat Genet* **48**, 593–599 (2016).
201. Karmin M, Saag L, Vicente M, Wilson Sayres MA, Järve M, Talas UG, Rootsi S, Ilumäe AM, Mägi R, Mitt M, Pagani L, Puurand T, Faltyskova Z, Clemente F, Cardona A, Metspalu E, Sahakyan H, Yunusbayev B, Hudjashov G, DeGiorgio M, Loogväli EL, Eichstaedt C, Eelmets M, Chaubey G, Tambets K, Litvinov S, Mormina M, Xue Y, Ayub Q, Zoraqi G, Korneliussen TS, Akhatova F, Lachance J, Tishkoff S, Momynaliev K, Ricaut FX, Kusuma P, Razafindrazaka H, Pierron D, Cox MP, Sultana GN, Willerslev R, Muller C, Westaway M, Lambert D, Skaro V, Kovačević L, Turdikulova S, Dalimova D, Khusainova R, Trofimova N, Akhmetova V, Khidiyatova I, Lichman DV, Isakova J, Pocheshkhova E, Sabitov Z, Barashkov NA, Nymadawa P, Mihailov E, Seng JW, Evseeva I, Migliano AB, Abdullah S, Andriadze G, Primorac D, Atramentova L, Utevska O, Yepiskoposyan L, Marjanovic D, Kushniarevich A, Behar DM, Gilissen C, Vissers L, Veltman JA, Balanovska E, Derenko M, Malyarchuk B, Metspalu A, Fedorova S, Eriksson A, Manica A, Mendez FL, Karafet TM, Veeramah KR, Bradman N, Hammer MF, Osipova LP, Balanovsky O, Khusnutdinova EK,

- Johnsen K, Remm M, Thomas MG, Tyler-Smith C, Underhill PA, Willerslev E, Nielsen R, Metspalu M, Villems R, Kivisild T. A recent bottleneck of Y chromosome diversity coincides with a global change in culture. *Genome Res* 25, 459–466 (2015).
202. Lippold S, Xu H, Ko A, Li M, Renaud G, Butthof A, Schröder R, Stoneking M. Human paternal and maternal demographic histories: insights from high-resolution Y chromosome and mtDNA sequences. *Investig Genet* 5, 13 (2014).
203. Behar DM, Harmant C, Manry J, van Oven M, Haak W, Martinez-Cruz B, Salaberria J, Oyharçabal B, Bauduer F, Comas D, Quintana-Murci L; Genographic Consortium. The Basque paradigm: genetic evidence of a maternal continuity in the Franco-Cantabrian region since pre-Neolithic times. *Am J Hum Genet* 90, 486–493 (2012).
204. Olivieri A, Pala M, Gandini F, Hooshier Kashani B, Perego UA, Woodward SR, Grugni V, Battaglia V, Semino O, Achilli A, Richards MB, Torroni. A Mitogenomes from two uncommon haplogroups mark late glacial/postglacial expansions from the near east and neolithic dispersals within Europe. *PLoS ONE* 8, e70492 (2013).
205. Pereira JB, Costa MD, Vieira D, Pala M, Bamford L, Harich N, Cherni L, Alshamali F, Hatina J, Rychkov S, Stefanescu G, King T, Torroni A, Soares P, Pereira L, Richards MB. Reconciling evidence from ancient and contemporary genomes: a major source for the European Neolithic within Mediterranean Europe. *Proc Biol Sci* 284, 20161976 (2017).
206. Richards M, Macaulay V, Hickey E, Vega E, Sykes B, Guida V, Rengo C, Sellitto D, Cruciani F, Kivisild T, Villems R, Thomas M, Rychkov S, Rychkov O, Rychkov Y, Gölge M, Dimitrov D, Hill E, Bradley D, Romano V, Cali F, Vona G, Demaine A, Papiha S, Triantaphyllidis C, Stefanescu G, Hatina J, Belledi M, Di Rienzo A, Novelletto A, Oppenheim A, Nørby S, Al-Zaheri N, Santachiara-Benerecetti S, Scozari R, Torroni A, Bandelt HJ. Tracing European founder lineages in the near eastern mtDNA pool. *Am J Hum Genet* 67, 1251–1276 (2000).
207. Rosa A. & Brehem A. African human mtDNA phylogeography at a glance. *J Anthropol Sci* 89, 25–58 (2011).
208. Stoneking M & Delfin F. The human genetic history of East Asia: weaving a complex tapestry. *Curr Biol* 20, R188–193 (2010).
209. Silva M, Oliveira M, Vieira D, Brandão A, Rito T, Pereira JB, Fraser RM, Hudson B, Gandini F, Edwards C, Pala M, Koch J, Wilson JF, Pereira L, Richards MB, Soares P. A genetic chronology for the Indian Subcontinent points to heavily sex-biased dispersals. *BMC Evol Biol* 17, 88 (2017).

210. Huyghe JR, Fransen E, Hannula S, Van Laer L, Van Eyken E, Mäki-Torkko E, Aikio P, Sorri M, Huentelman MJ, Van Camp G. A genome-wide analysis of population structure in the Finnish Saami with implications for genetic association studies. *Eur J Hum Genet* 19, 347–352 (2011).
211. Mallick S, Li H, Lipson M, Mathieson I, Gymrek M, Racimo F, Zhao M, Chennagiri N, Nordenfelt S, Tandon A, Skoglund P, Lazaridis I, Sankararaman S, Fu Q, Rohland N, Renaud G, Erlich Y, Willems T, Gallo C, Spence JP, Song YS, Poletti G, Balloux F, van Driem G, de Knijff P, Romero IG, Jha AR, Behar DM, Bravi CM, Capelli C, Hervig T, Moreno-Estrada A, Posukh OL, Balanovska E, Balanovsky O, Karachanak-Yankova S, Sahakyan H, Toncheva D, Yepiskoposyan L, Tyler-Smith C, Xue Y, Abdullah MS, Ruiz-Linares A, Beall CM, Di Rienzo A, Jeong C, Starikovskaya EB, Metspalu E, Parik J, Villems R, Henn BM, Hodoglugil U, Mahley R, Sajantila A, Stamatoyannopoulos G, Wee JT, Khusainova R, Khusnutdinova E, Litvinov S, Ayodo G, Comas D, Hammer MF, Kivisild T, Klitz W, Winkler CA, Labuda D, Bamshad M, Jorde LB, Tishkoff SA, Watkins WS, Metspalu M, Dryomov S, Sukernik R, Singh L, Thangaraj K, Pääbo S, Kelso J, Patterson N, Reich D. The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. *Nature* 538, 201–206 (2016).
212. Menozzi P, Piazza A & Cavalli-Sforza LL. Synthetic maps of human gene frequencies in Europeans. *Science* 201, 786–792 (1978).
213. Juhász Z. A model of variation in the music of a Hungarian ethnic group. *J New Music Res* 29:159–172 (2000).
214. Juhász Z, Sipos J. A comparative analysis of Eurasian folksong corpora, using self organising maps. *J Interdiscip Music Stud* 4:1–16 (2009).
215. Kohonen T. Self-organising maps. Springer, Berlin (1995).
216. Juhász Z. A systematic comparison of different European folk music traditions using self-organizing maps. *J New Music Res* 35:95–112 (2006).
217. Borg I, Groenen P. Modern multidimensional scaling: theory and applications, 2nd edn. Springer, New York (2005).
218. Pamjav H, Juhász Z, Zalán A, Németh E, Damdin B. A comparative phylogenetic study of genetics and folk music. *Mol Genet Genomics*. 287(4):337-49 (2012).
219. Kodály Z. Folk music of Hungary. Corvina, Budapest (1971).
220. Bartók B. On collecting folk songs in Turkey. *Tempo, NewSer*, No. 13, Bartok Number (Autumn, 1949), pp 15–19 (1949).
221. Sipos J. In the wake of Bartók in Anatolia. European Folklore Institute, Budapest (Bibliotheca Traditionis Europaeae 2) (2000).

222. Sipos J. Egy most felfedezett belső-mongóliai kvintváltó stílus és magyar vonatkozásai. *Ethnographia* 112, pp 1–80 (2001).
223. Sipos J. Comparative analysis of Hungarian and Turkic folk music. Edition of TIKA (Türk İ's, birliğ i ve Kalkınma İ dairesi Bas ,kanlığı ı) and the Hungarian Embassy, Ankara (2006).
224. Sipos J. Similar musical structure in Turkish, Mongolian, Tungus and Hungarian folk music. In: Berta A ´ (ed) *Historical and linguistic interaction between Inner-Asia and Europe*. Szeged, pp 305–317 (1997).
225. Brown S, Savage PE, Ko AM, Stoneking M, Ko YC, Loo JH, Trejaut JA. Correlations in the population structure of music, genes and language. *Proc Biol Sci.* 281(1774):20132072 (2014).
226. Savage PE. Measuring the cultural evolution of music: With case studies of British-American and Japanese folk, art, and popular music. Ph.D. dissertation (2016).
227. Le Bomin S, Lecointre G, Heyer E. The Evolution of Musical Diversity: The Key Role of Vertical Transmission. *PLoS One.* 11(3):e0151570 (2016).
228. Wiggins GA, Tyack P, Scharff C, Rohrmeier M. The evolutionary roots of creativity: mechanisms and motivations. *Philos Trans R Soc Lond B Biol Sci.* 370(1664):20140099 (2015).
229. Windram HF, Charlston T, Howe CJ. A phylogenetic analysis of Orlando Gibbons's Prelude in G. *Early Music*, Volume 42, Issue 4, 1 November 2014, Pages 515–528.

7. KÖZLEMÉNYEK LISTÁJA

7.1. 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, Vågane 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 Bodrogek population in north-eastern Hungary by Y chromosomal haplotypes and haplogroups. *Mol Genet Genomics*. 292:(4) pp. 883-894. (2017) (IF: 2.979).
4. Neparáczi 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, Endrédy 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).**

7.2. 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).