

**Honfoglalás kori, valamint magyar és székely  
populációk apai ági genetikai  
kapcsolatrendszerének vizsgálata**

Ph. D. értekezés

Kovácsné Csányi Bernadett

Témavezető: Prof. Dr. Raskó István

Biológia Doktori Iskola

MTA Szegedi Biológiai Központ  
Genetikai Intézet

Szegedi Tudományegyetem  
Természettudományi és  
Informatikai Kar

2009.

Szeged

# Tartalomjegyzék

<b>Előszó</b> .....	4
<b>I. Bevezetés</b> .....	5
I.1. Nyelvében él a nemzet.....	5
I.2. Populációk eredetének vizsgálatára alkalmas genetikai rendszerek.....	8
I.2.1. Az Y kromoszóma, mint a populáció eredetvizsgálatok egyik sajátos eszköze.....	9
I.3. Az Y kromoszóma szerkezete.....	10
I.3.1. A férfi-specifikus régió (MSY) heterokromatikus szakaszai .....	12
I.3.2. A férfi-specifikus régió (MSY) eukromatikus szakaszai és jellemző szekvencia típusai .....	12
I.4. Az Y kromoszóma polimorf markerei.....	14
I.5. Az Y-kromoszómális nevezéktan rendszer és a haplocsoportok filogenetikája....	16
I.6. A modern ember afrikai eredete és elterjedése a Földön az Y-kromoszómális vizsgálatok alapján .....	19
I.6.1. Európa paleolitikumi és neolitikumi leszármazási vonalai.....	22
I.7. A Tat (M46) marker által meghatározott leszármazási vonal.....	27
I.8. Magyar nyelvű populációk eddigi genográfiai vizsgálati eredményei a klasszikus, a mitokondriális és az Y-kromoszómális SNP markerek felhasználásával .....	29
I.9. Az archaikus DNS jellemzői .....	31
<b>II. Célkitűzések</b> .....	36
<b>III. Anyagok és módszerek</b> .....	37
III. 1. A vizsgált minták .....	37
III. 2. DNS extrakció.....	40
III. 2. 1. DNS extrakció modern mintákból .....	40
III. 2.1.1. DNS extrakció vérből .....	40
III. 2.1.1.1. DNS extrakció vérből kisózással.....	40
III. 2.1.1.2. DNS extrakció vérből Chelex módszerrel.....	41
III. 2.1.2. DNS extrakció hajszálból .....	42
III. 2. 2. DNS extrakció régészeti csontleletekből .....	42
III. 2.2.1. A csontminták előkészítése.....	42

III. 2.2.2. DNS extrakció a csontmintákból .....	43
III. 3. Az archaikus DNS amplifikációja .....	43
III. 3. 1. Mitokondriális DNS amplifikáció .....	43
III. 3. 2. Teljes genom amplifikáció .....	44
III. 3. 3. A Tat polimorfizmust hordozó DNS szakasz amplifikációja .....	46
III. 4. A szennyeződés megelőzése és az eredmények hitelesítése .....	47
III. 5. A modern DNS minták Y-kromoszómális analízise .....	48
III. 5. 1. Multiplex PCR (az M96, M89, M9, M45 SNP markerek amplifikálása) .....	51
III. 5. 2. Az M96, M89, M9, M45 SNP markerek genotípusának meghatározása .....	52
III. 5. 3. PCR-RFLP analízis .....	56
III. 5. 4. Statisztikai analízis .....	59
<b>IV. Eredmények .....</b>	<b>60</b>
IV. 1. A Tat polimorfizmus vizsgálata .....	60
IV. 2. A modern DNS minták Y-kromoszómális analízise .....	66
IV. 3. A modern minták Y-kromoszómális adatainak statisztikai vizsgálata .....	72
<b>V. Diskusszió .....</b>	<b>75</b>
V. 1. A Tat polimorfizmus vizsgálata .....	75
V.2. A ma élő magyar és székely populáció Y-kromoszómális genetikai összetételének vizsgálata .....	77
<b>VI. Következtetések .....</b>	<b>84</b>
<b>VII. Köszönetnyilvánítás .....</b>	<b>85</b>
<b>VIII. Irodalomjegyzék .....</b>	<b>86</b>
<b>Összefoglalás .....</b>	<b>98</b>
<b>Summary .....</b>	<b>103</b>

*„Ha ismered a múltadat, tudni fogod honnan jöttél.”*

Bob Marley, „Buffalo Soldier”

## **Előszó**

*DNS-ünk négy egyszerű betűbe rejtve igazi történeti dokumentumot hordoz magával, visszanyúlva az élet kezdetéig, az első önmagát reprodukáló molekuláig, majd amőbaszerű őseinken keresztül el egészen a jelenig. Több mint egymilliárd éve tartó evolúciós barkácsolás eredményei vagyunk, és génjeink biztosítják hozzá a szegecseket, varratokat, amelyek feltárják fejlődésünk teljes történetét. Nem maguk a gének hordozzák az üzenetet, sokkal inkább az eltérések, amelyeket két vagy több személy DNS-ének összehasonlításakor láthatunk. E különbségek képezik a gének történeti nyelvezetét (Wells 2003).*

# I. Bevezetés

## I.1. Nyelvében él a nemzet

A magyar nyelv az uráli nyelvcsalád tagja, a finnugor nyelvek közé tartozó ugor nyelvek egyike (Róna-Tas 1999, Szíj 2005). Legközelebbi rokonai a manysi és a hanti nyelv. A magyar nyelv beszélőinek száma megközelítőleg 15 millió (ezzel az uráli nyelvcsalád legnépesebb tagja), közülük körülbelül 10 millióan élnek Magyarországon és mintegy 3 millióan a szomszédos országok területén. A Kárpát-medencében élő magyarul beszélő populációk képezik az uráli nyelvet beszélők legnyugatibb előfordulási csoportját.

A magyar nép kialakulásának, történetének kérdéseit nem lehet a magyar nyelv történetének ismerete nélkül megoldani, de ez csak az egyik, bár igen fontos eszközünk (Róna-Tas 1996). A magyar nyelv története rögzítette történelmünk számos fontos eseményét, a magyar nyelv története a nyelvörző típus szerint alakult. A magyar nyelv több ezer éves története alatt számos nyelvvel érintkezett, az indoeurópai, különösen az iráni, majd a török, a szláv, a német és más nyelvek gazdagították a magyar nyelvet, de a magyar nyelv átalakulva is megmaradt magyar nyelvnek. Itt azonban fontos megjegyezni, hogy egy nyelv története korántsem esik egybe egy nép történetével, azonban a magyar őstörténet legkorábbi időszakairól nem maradtak fenn korabeli írásos források, így a koratörténet megrajzolása elsősorban a nyelvészet nyújtotta ismeretek alapján történik (Mende 2005).

A magyarság hasonlóan más népekhez több gyökerű volt. Kialakulásában egyaránt jelentős szerepet játszottak a finnugor nyelvet beszélők és a török nyelvű népek. A sztyeppe területén, majd később a Kárpát-medence elfoglalását követően számos további néptöredék is betagozódhatott a folyamatosan formálódó magyarságba (Mende 2005). A magyar nyelv nemcsak a közlések kicserélésére volt alkalmas, hanem a hagyományok őrzésére is ugyanúgy, ahogy az új és még újabb ismeretek befogadására és kifejezésére. Éppen a magyar nyelv kívülállása, eltérése a környezettől biztosította azt az etnikai tudatot, az önazonosság tudatát, a *mi* és *ti* elkülönülését, ami szükséges volt a magyar nép évezredek át tartó fennmaradásához. (Az etnikai-népi azonosság megőrzéséhez a nyelven kívül hozzájárulhatott még a magyarság társadalmi-gazdasági szerkezete, az idegenek befogadása, melynek következtében a magyarság erősebb és gazdagabb lett, ugyanakkor a Kárpát-medence lakosságának sajátossága is fontos

szerepet játszott- az elszlávosodó avarok a magyarokban felszabadítóikat, rokonaikat látták (Róna-Tas 1996)). A magyarság kialakulásának és korai, honfoglalás előtti történetének rekonstruálásában a nyelvi források mellett az írásos források, a régészet, a néprajz, a biológia (például a modern genetika) és az úgynevezett közvetett források (például Julianus barát útjáról szóló feljegyzés) segíthetnek.

Mai ismereteink szerint a finnugor nyelvcsalád a szamojéd nyelvekkel együtt az Urál hegység és az Ob folyó vidékén alakult ki az őskorban. A nyelvészek ezeket a nyelveket együttesen az uráli nyelvcsaládhoz sorolják, míg ezt az időszakot az uráli kornak nevezik. A Krisztus előtti 2. évezredben a finnugor nyelvek szétváltak. A finn-permi ágtól elváltak az ugor nyelvek, köztük a magyar is. Az ugor nyelvű csoportok (magyar, vogul (manysi), osztják (hanti)) valószínűleg az Uráltól keletre az Ob, Irtyisz, Iszim és a Tobol környékén élhettek (Mende 2005). A Krisztus előtti 1. évezred második felében (körülbelül Kr. e. 500-ban) két részre szakadt az ugor nyelvet beszélő közösség (Fodor 1996). A magyar nyelvet használók egyre délebbre kerültek, míg az obi-ugorok az Ob mentén észak felé húzódtak. Az északi ugor közösségek utódai a vogulok és az osztjákok máig ezen a területen élnek. A déli irányba forduló ősmagyar nyelvű csoport az indoeurópai és a török nyelvű népek közé kerülve számos más nyelvet beszélő csoporttal ötvöződött. A magyarság kialakulásában így egyre nagyobb szerepet kaptak a finnugor örökség mellett a sztyeppe nomád kultúrák is (Mende 2005). Őseink két évezrednyi sztyeppe vándorlást követően, a 9. század végén telepedtek le a Kárpát-medencében (Bálint 1996). A korai magyarok a letelepedés idején finnugor nyelvet beszéltek, de török módra éltek (Róna-Tas 1995, Sindbæk 1999, Bálint 2006). A magyarság a nyelvi háttere alapján az uráli nyelvcsaládhoz köthető, azonban a történetkutatás mai elképzelése szerint a nép etnikai összetevői jelentősen megváltozhattak az elmúlt ezer év alatt a más törzsi és népi egységekkel történő keveredés következtében (Bálint 2005).

A székelység korai történetéről keveset tudunk, s ez a kombinációknak tág teret ad. Kialakulásukra, eredetükre vonatkozóan alapvetően két fő elképzelés létezik (Bóna 1990, Kordé 1994, Kristó 2005). Az egyik elmélet szerint a székelyek a volgai bolgárok eszkil törzsének utódai, egy török népesség, amely a honfoglalás során a magyarsággal együtt érkezett a Kárpát-medencébe és telepedett meg a történelmi Magyarország keleti határán (Erdélyben) (Kristó 2005). A másik elmélet szerint a székelyek a kora Árpád-korban (a 12. században) a Dunántúlról a keleti határrészre kihelyezett magyarok voltak

(Bóna 1990). A székelyek határvédő szolgálatot teljesítettek, nemesi kiváltságokkal rendelkeztek, melynek következtében a középkortól (a 12. századtól) egy viszonylag zárt etnikai csoportot képeztek egészen a 19. századig (Kordé 1994). Lazulás a kötöttségeikben csak a 19. század második felében volt, amit aztán lezárt a trianoni béke és a terület elcsatolása, ami újabb, de már más politikai okokkal magyarázható izolációt eredményezett a székelység körében.

Kísérletsorozatunkban egyrészt arra törekedtünk, hogy megvizsgáljuk, hogy a honfoglaló magyarok Y-kromoszómális génkészletében tetten érhető-e az uráli nyelvű populációkkal való rokonság emléke, másrészt a ma élő magyarországi magyar és a korondi székely populáció apai ági leszármazási vonalainak részletes tanulmányozásával további információkat kívántunk szerezni a két populáció egymással és más európai populációkkal való genetikai kapcsolatrendszeréről.

## I.2. Populációk eredetének vizsgálatára alkalmas genetikai rendszerek

A humán genom közel 99%-a a sejtmag kromoszómaiban, míg a maradék 0,5-1%-a a mitokondriumokban található. A mitokondriális genom összesen 37 gént hordoz, a nukleáris genomban található gének száma körülbelül 30 000.

A nukleáris géneket mindkét szülő egyenlő mértékben hagyományozza ránk. Az ivaros szaporodás fő célja új genomok létrehozása minden elkövetkező generációban. Az új kombináció nem csupán a megtermékenyítéskor keletkező 50-50%-os keveredésnek köszönhető, hanem a már korábban lejátszódott apai/anyai ivarsejtképződésnek. Ezt a preszexuális keveredést genetikai rekombináció néven ismerjük. Evolúciós szempontból nézve, az apai és az anyai kromoszómák keveredéséhez hasonlóan, a rekombináció megnöveli az utódnemzedék genetikai diverzitását. A magas diverzitás pedig a legjobb háttér a fennmaradáshoz egy változásnak indult környezetben. Bár az evolúció oldaláról nézve a rekombináció egy igen hatékony diverzitásnövelő mechanizmus, mégis alaposan megkeseríti a molekuláris biológusok életét, mikor el akarják olvasni a humán genom történelemlékönyvét. Az idő múlásával ugyanis a polimorfizmusok számtalanszor rekombinálódnak és néhány száz vagy ezer generáció után a közös őst eredeti polimorfizmus-mintázata örökre elvesz. Ennek megfelelően pillanatnyilag reménytelenül bonyolult dolog a sejtmag génjeiből és géndarabjaiból kiolvasni az emberiség történetének különféle egyedi változatait (Sykes 2002, Wells 2003).

Populációk eredetének tanulmányozásához olyan markereket célszerű vizsgálni, melyeknél nem keveredik az anyai és az apai információ az egymást követő nemzedékekben. Ellentétben a nukleáris DNS döntő hányadával vannak olyan nem rekombináló genetikai egységek, amelyek csak anyai vagy csak apai ágon öröklődnek. Ilyen a mitokondriális DNS, amely a petesejtek mitokondriumai révén kizárólag anyai ágon öröklődik (Giles et al. 1980), vizsgálatával populációk anyai ági leszármazási vonalait követhetjük nyomon. A másik ilyen egység az apáról fiúra öröklődő Y kromoszóma körülbelül 95%-a (a pszeudoautoszómális régiók közötti közbülső rész), melynek analízisével populációk apai ági leszármazási vonalait deríthetjük fel.



### **I.2.1. Az Y kromoszóma, mint a populáció eredetvizsgálatok egyik sajátos eszköze**

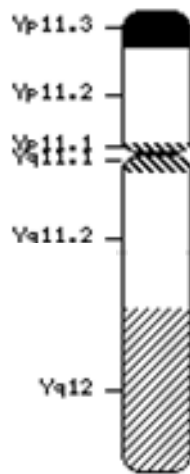
Az Y kromoszóma legfontosabb biológiai szerepe a nemiség meghatározása és a férfi fertilitás biztosítása. Sajátos jellemzői következtében, melyek egyedivé teszik a kromoszómák között, az Y kromoszóma hatékony eszköznek bizonyult a populációgenetikuskok számára a humán diverzitás tanulmányozásában és az apai ági leszármazási vonalak nyomon követésében. Az Y kromoszóma haploid, csak a férfiakban van jelen, apáról fiúra öröklődik és 95%-án nem történik rekombináció a meiózis során. A rekombináció hiányának jelentősége, hogy a haplotípusok, melyeket az Y kromoszómán található markerek allélikus állapotának kombinációi határoznak meg, általában érintetlenül adódnak tovább generációról generációra. Más szavakkal az Y kromoszóma nem rekombinálódó régiója egyetlen lókuszként öröklődik. Mivel csak az idővel halmozódó mutációk révén jöhet létre változás az autoszómális DNS-hez képest az Y kromoszómák egy viszonylag egyszerűbb genetikai történetet őriznek. Az Y kromoszóma további sajátos jellemzője, hogy 1:1 nemi arány esetén a teljes populációban az Y várható effektív populációmérete negyede bármelyik autoszómáéénak, harmada az X kromoszómáéénak és egyenlő méretű a mitokondriális DNS-ével. Ennek megfelelően az Y-kromoszómális genetikai változatosságot (csakúgy mint a mtDNS-ét) nagyobb mértékben érinti a genetikai sodródás hatása. A genetikai sodródás felgyorsítja a különböző populációkban az Y-kromoszómális leszármazási vonalak csoportjai közötti differenciációt és hozzájárul jelentős földrajzi halmozódásukhoz (Jobling és Tyler-Smith 2003) (5. ábra). Gyakoriságában az Y lényegesen nagyobb genetikai differenciákat mutat a populációk között, mint a legtöbb egyéb marker. Ahogy Lewontin (1972) genetikai analízise kimutatta, az emberi faj genetikai változatosságának zömét populáción belül találjuk, és csak töredékét, 10-15%-át a populációk között, azonban ugyanez az arány az Y kromoszóma elemzéseknél már 30-40%-nak adódott. Nagyobb genetikai kontraszt nagyobb felbontást tesz lehetővé, emiatt is alkalmas olyannyira az Y a migrációk nyomon követésére (Wells 2003). Az Y kromoszóma populációk közti nagyfokú diverzitását a társadalmak patrilokalitása is magyarázza. E szerint a társadalmak körülbelül 70%-ánál amikor két ember egybekel, többnyire a nők váltanak lakóhelyet, ők költöznek a férjükhöz. Más szóval a férfiak genetikusan kevésbé mozgathatók. Ennek következményeként egy jóval homogénebb

mitokondriális DNS elterjedési térkép alakult ki, míg az Y kromoszómák egymástól függetlenül divergálódtak a különböző populációkban (Seielstad et al. 1998). Ez az eredmény is mutatja, hogy az emberi kultúrának milyen hatalmas szerepe van/volt fajunk genetikai mintázatának kialakításában. A patrilokalitás mintegy ellenpontja a nemre specifikus génáramlás, mely az elmúlt 500 év során az európaiaknak Óceániában és az amerikai földrészen történő elterjedését kísérte. A gyarmatosító tevékenység Polinéziában (Hurles 1998), Grönlandon (Bosch et al. 2003) és Dél-Amerikában (Carvajal-Carmona et al. 2000, Carvalho-Silva et al. 2001) az „öshonos” mitokondriális DNS vonalak megtartása mellett az európai Y kromoszómák nagymértékű beáramlásával járt együtt.

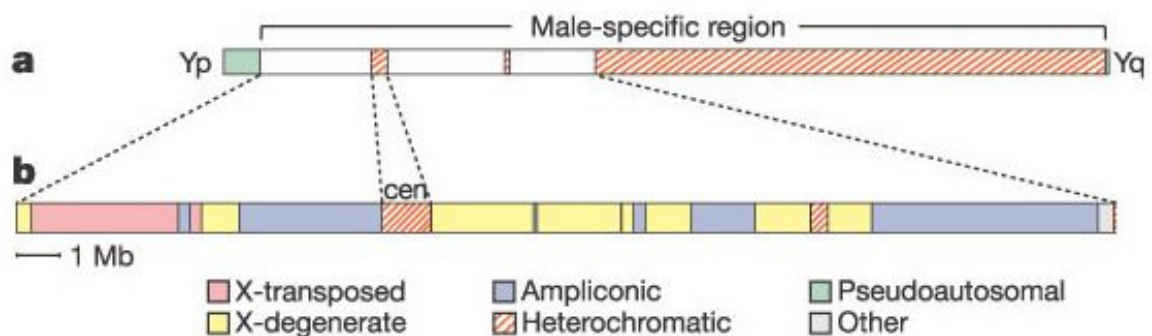
### **I.3. Az Y kromoszóma szerkezete**

Az Y kromoszóma körülbelül 60 millió bázispár (60 Mbp, 60 Mb (megabázispár)) hosszú (1. ábra). A kromoszóma hosszának 95%-án nem játszódik le rekombináció a meiózis során az X és az Y kromoszóma között. Ezt a szakaszt az Y kromoszóma nem rekombinálandó régiójának (**Non-Recombining region of Y - NRY**, **Non-Recombining Portion Y - NRPY**), illetve férfi-specifikus régiójának (**Male-Specific region of Y - MSY**) nevezik. Ezt a régiót mindkét oldalon a kromoszóma telomer szakaszain elhelyezkedő pszeudoautoszómális régiók (kevesebb, mint 3 megabázisnyi szakasz a kromoszóma körülbelül 60 megabázisnyi hosszából) szegélyezik, melyek rekombinációja az X kromoszómával a férfi meiózis gyakori és szabályszerű eseménye. Skaletsky és munkacsoportja (2003) határozta meg a férfi-specifikus régió 97%-ának szekvenciáját és jellemezte ezt a régiót. Ez a régió nagyjából két kiterjedt részre osztható: eukromatikus és heterokromatikus területre (2. ábra). Skaletsky és munkatársai a szatellita szekvenciákat tekintették a kromoszóma heterokromatikus és minden egyéb szekvenciát a kromoszóma eukromatikus állományának.

### Ideogram



1. ábra Az Y kromoszóma ideogramja – forrás: NCBI (National Center for Biotechnology Information) adatbázis (<http://www.ncbi.nlm.nih.gov>)



2. ábra Az Y kromoszóma férfi-specifikus régiója (Skaletsky et al. 2003)

**a:** Az egész kromoszóma sematikus ábrázolása a pseudoautoszómális és a heterokromatikus régiókkal együtt.

**b:** A férfi-specifikus régió egy 24 Mb hosszú szakaszának kinagyított képe a rövid kar pseudoautoszómális régiójának proximális határától a hosszú kar nagy méretű heterokromatikus régiójának proximális határáig az eukromatin különböző szekvencia típusainak és a heterokromatikus szekvenciák jelölésével.

### **I.3.1. A férfi-specifikus régió (MSY) heterokromatikus szakaszai**

Az Y kromoszóma férfi-specifikus szakasza három heterokromatikus régiót tartalmaz. Ezek közül az egyik a centromerikus heterokromatin, amely hozzávetőleg 1 Mb nagyságú (Tyler-Smith et al. 1993). A második egy sokkal nagyobb (nagyjából 40 Mb nagyságú – azonban hossza jelentős mértékben változhat, nagymértékű polimorfizmust mutat (Skaletsky et al. 2003)) heterokromatikus blokk, mely a hosszú kar disztális szakaszának nagy részét elfoglalja (Yq12) (1. ábra). A harmadik heterokromatikus blokk egy élesen elhatárolt sziget a hosszú kar proximális részén az eukromatikus szekvenciák között (Yq11.22). Mintegy 400 kilobázis hosszú, több, mint háromezer 125 bázispáros tandem ismétlődésből áll (Skaletsky et al. 2003).

### **I.3.2. A férfi-specifikus régió (MSY) eukromatikus szakaszai és jellemző szekvencia típusai**

Az Y kromoszóma férfi-specifikus szakaszán az eukromatikus DNS szekvenciák közel 23 Mb nagyságú területet foglalnak el, ebből 8 Mb a rövid karon, 14,5 Mb a hosszú karon található. Ezen a szakaszon Skaletsky és munkacsoportja (2003) 156 transzkripciós egységet azonosított, melyekből 78 összesen 27 különböző fehérjét vagy fehérje családot kódol. A 78 fehérje kódoló egységből körülbelül hatvan 9 különböző MSY-specifikus géncsalád tagja. A maradék 18 protein kódoló gén mindegyike egy kópiában van jelen a férfi-specifikus régióban. A 27 különböző fehérje kódoló génből vagy géncsaládból 12 általánosan, a legtöbb szövetben expresszálódik, míg 11 elsősorban vagy kizárólag a herékben fejeződik ki.

Csaknem az összes eukromatikus szekvencia besorolható 3 osztályba, ezek az X-ről áthelyeződött (X-transposed), az X-degenerálódott (X-degenerate) és az amplikon (Ampliconic) (az amplikonok ismétlődő, megsokszorozódott DNS szakaszok) szekvencia osztályok.

Az X-ről áthelyeződött szekvenciák 99%-ban azonosak az X kromoszóma hosszú karján (Xq21) található DNS szekvenciákkal. Jelenlétük a humán férfi-specifikus szakaszon az X kromoszómáról az Y kromoszómára történő transzpozíció következménye, mely körülbelül 3-4 millió évvel ezelőtt ment végbe, a humán és a csimpánz leszármazási vonalak elágazását követően. Később az MSY szakasz rövid

karján egy inverzió hatására kettéhasadt ez a blokk két nem folyamatos szegmentumra. Az X-ről áthelyezett szekvenciák nem vesznek részt az X-Y crossing overben a meiózis során, ez a tényező megkülönbözteti őket a pszeudoautoszómális régióktól. Az X-ről áthelyeződött szegmentumokban csak két gént azonosítottak, ugyanakkor nagy sűrűségben fordulnak elő közbeiktatott ismétlődő elemek (interspersed repeat elements).

Az X-degenerálódott szegmentumokon belül az egy kópiában előforduló gének és a pszeudogének 60-96%-os nukleotid szekvencia egyezést mutatnak az X-hez kapcsolt homológjaikkal. Valószínűleg azon ősi autoszómák maradványait képviselik, melyekből az X és az Y kromoszóma kialakult. A 12 általánosan kifejeződő MSY specifikus gén az X-degenerálódott régiókon helyezkedik el. A nemiség meghatározásában fontos SRY gén is ebben a régióban található meg (Skaletsky et al. 2003).

Az amplikon szegmentumok nagymértékben olyan szekvenciákból állnak, melyek jelentős, 99,9%-os egyezést mutatnak más, szintén a férfi-specifikus régió elhelyezkedő szekvenciákkal. Az amplikon régiókban figyelhető meg a legnagyobb génsűrűség. Kilenc különböző MSY-specifikus fehérje kódoló géncsalád van jelen, melyek kópiaszáma kettőtől 35-ig változik. Mind a kilenc géncsalád elsősorban vagy kizárólagosan a herékben fejeződik ki. Az amplikon régiók legjellegzetesebb strukturális sajátossága a kromoszóma hosszú karján a 8 tükörszimmetriát mutató palindrom, melyeknél a karok közötti nukleotid egyezés 99,94 - 99,997%. A 8 palindrom a férfi-specifikus régió eukromatikus szakaszának egynegyedét teszi ki. Ezek közül hat hordoz fehérje kódoló géneket, melyek specifikusan a herékben fejeződnek ki. Minden esetben egyező, vagy közel egyező génkópiák találhatóak a palindrom ellentétes karjain. A palindromokon kívül öt fordított ismétlődést (Inverted Repeat - IR) és különböző hosszú tandem tömböket (long tandem arrays) is találunk ezekben a régiókban (Skaletsky et al. 2003).

Az irodalom szerint (Rozen et al. 2003, Skaletsky et al. 2003) a férfi-specifikus régió eukromatikus szakaszának 30%-án történik Y-Y génkonverzió (nem reciprok rekombináció). A kromoszóma ezen szakaszában, amely magában foglalja a 8 palindromot és a fordított ismétlődések közül az IR2 és az IR3 nagy részét, a szekvencia párok közötti intrakromoszómális nukleotid egyezés  $\geq 99,9\%$ . A 9 géncsaládból, melyek elsősorban vagy kizárólagosan a herékben fejeződnek ki, 8 esetében található génkópia palindromokon vagy fordított ismétlődéseken, ahol a nukleotid egyezés  $\geq 99,9\%$ . Úgy

gondolják, hogy a génkonverzióknak fontos szerepe lehet a génfunkció megőrzésében crossing over hiányában.

#### **I.4. Az Y kromoszóma polimorf markerei**

Az Y kromoszómán található nagyszámú polimorfizmus két nagyobb csoportba sorolható be. Az egyik a biállélikus markerek, a másik a tandem ismétlődések, vagy multi-allélikus markerek csoportja (Jobling és Tyler-Smith 2000, Hammer és Zegura 2002).

A biállélikus markerek közé tartoznak a pontmutációk (**S**ingle **N**ucleotide **P**olymorphisms - SNP markerek), valamint az inszerciók és a deléciók (indels).

Az SNP markerek esetében a mutációs ráta alacsony ( $\sim 2 \times 10^{-8}$ /bázis/generáció (Nachman és Crowell 2000)), ennél fogva a korai demográfiai események vizsgálatát teszik lehetővé. Elnevezésük egy betű és egy szám segítségével történik. A betű a markert felfedező munkacsoportra utal (például a P betűs markereket alapvetően Hammer és munkacsoportja, míg az M betűs markereket főként Underhill és munkacsoportja írta le), míg a szám azt mutatja, hogy hanyadikként fedezte fel/dokumentálta az adott munkacsoport a markert.

Az inszerciók és deléciók polimorfizmusok esetén (például **Y Alu Polimorfizmus -YAP**) azok jelenlétét vagy hiányát detektálják a vad típushoz képest. A nagyobb átrendeződések, főleg az Y-specifikus géneket érintő deléciók (például az Azoospermia Faktor különböző régióiban elhelyezkedő, a spermatogenezisben fontos szerepet játszó génekben bekövetkező deléciók) károsan érinthetik a férfi fertilitást (Fernandes et al. 2002, Krausz et al. 2000, Vogt et al. 1996, Yen 1998).

Az inszerciók és deléciók egy része nem hat a férfi fertilitásra, ezek generációkon át fennmaradnak, hordozóikban nem hoznak létre detektálható fenotípust és eléggé gyakori előfordulást mutatnak a populáció(k)ban, ezek alapján polimorfizmusnak tekinthetők. Ilyen polimorfizmus például a 12f2 marker 2 kb nagyságú deléciója (Casanova et al. 1985), mely a J és a D2 haplocsoportba tartozó Y kromoszómákat jellemzi.

A markerek másik része tandem ismétlődő szekvenciákból épül fel. Az ismétlődő egységek hossza a szatelliták esetében ezer- több ezer bázispár, a miniszatelliták vagy változó számú tandem ismétlődések (**V**ariable **N**umber of **T**andem **R**epeats – VNTR markerek) esetében 10-100 bázispár, míg a mikroszatelliták vagy

rövid tandem ismétlődések (Short Tandem Repeats - STR markerek) esetében kevesebb, mint 10 bázispár, legtöbbször 2-6 bázispár (Nakamura et al. 1987, Charlesworth et al. 1994, Chambers és MacAvoy 2000).

A mikroszatellita markereket széles körben használják igazságügyi, származástani és populációgenetikai vizsgálatokra (de Knijff et al. 1997, Jobling et al. 1997, Kayser et al. 1997). Mutációs rátájuk a biallélikus markerekénél magasabb, ezért a filogenetikai vizsgálatokban az újabb keletű demográfiai események részleteinek tanulmányozására adnak lehetőséget. Kayser és munkatársai (2000) 15 különböző mikroszatellita lókuszon vizsgálta a mutációs ráta nagyságát. Di-, tri-és tetranukleotid mikroszatellitákat is tanulmányoztak apa-fiú párokon. A 15 lókuszra vonatkozóan az átlagos mutációs ráta  $2,8 \times 10^{-3}$  értékűnek adódott. Megfigyelték, hogy ez az érték nem különbözik szignifikánsan az autoszómális mikroszatelliták analízisének kapott átlagos mutációs ráták nagyságától ( $2,1 \times 10^{-3}$  (Brinkmann et al. 1998),  $2,7 \times 10^{-3}$  (Henke és Henke 1999),  $0,6 \times 10^{-3}$  (Sajantila et al. 1999)). A tetranukleotid ismétlődéseknél is hasonló eredmények születtek: Weber és Wong (1993) autoszómális tetranukleotid ismétlődések vizsgálatakor  $2,0 \times 10^{-3}$ /generáció nagyságú átlagos mutációs rátát állapított meg. Heyer és munkatársai (1997) az Y-kromoszómális tetranukleotid ismétlődések tanulmányozásakor az átlagos mutációs rátát az előbbivel azonos nagyságúnak találták. Az irodalom szerint (Kayser et al. 2000, Jobling és Tyler-Smith 2003) az Y-kromoszómális és az autoszómális mikroszatelliták jellemzői (például a mutációs rátára vonatkozó becslések) nem különböznek jelentősen egymástól, amely alapján feltételezhető, hogy a mikroszatelliták esetében a mutáció mechanizmusa független a rekombinációtól.

## I.5. Az Y-kromoszómális nevezéktan rendszer és a haplocsoportok filogenetikája

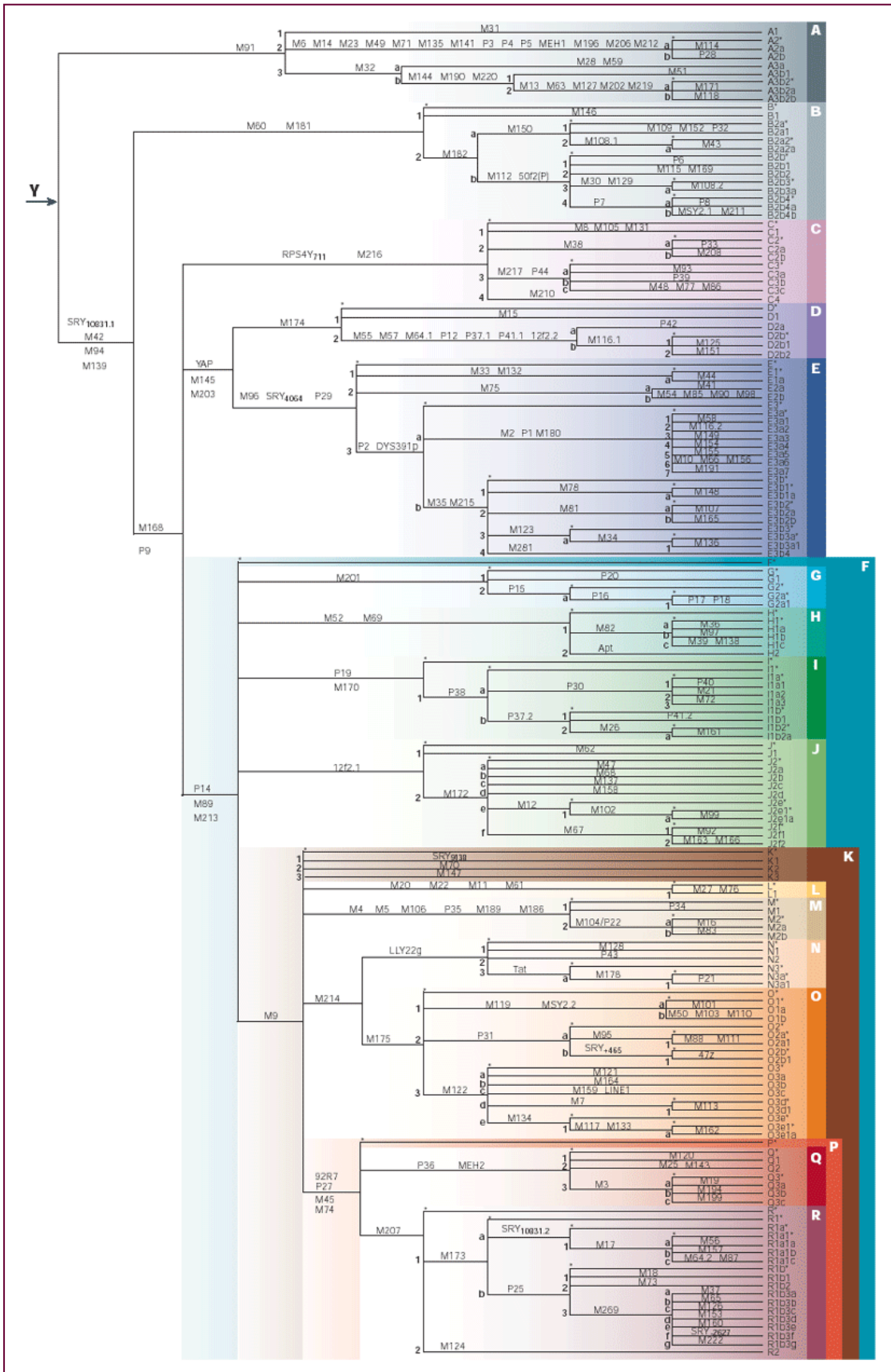
2002-ben az Y Chromosome Consortium (YCC) egy ma is használatban lévő és a korábbi, egymástól eltérő nomenklatúrákat egységesítő nevezéktan rendszert fejlesztett ki, valamint 245 marker genotípusának meghatározásával egyetlen, 153 Y-kromoszómális haplocsoportot átfogó, nagy felbontású, hierarchikus szerveződésű filogenetikai fát szerkesztett.

A haplocsoport kifejezés a biallélikus polimorfizmusok által meghatározott Y-kromoszómális leszármazási vonalakra utal. Egy minta haplocsoportba sorolása az Y-kromoszómán derivált/mutáns allélikus állapotban jelenlévő biallélikus markerek szerint történik. A főbb haplocsoportokat az ABC nagybetűivel jelölték. 18 nagyobb kládot különböztethetünk meg A-tól R-ig jelölve, az Y haplocsoport magában foglalja az összes nagy kládot (A-R-ig). A filogenetikai fa belső nóduszain találhatóak a potenciálisan parafiletikus csoportok (paracsoportok). Ide olyan kromoszómák/leszármazási vonalak tartoznak, melyek egy adott kládba besorolhatók, de annak egyik szubkládjába sem, mivel nem hordozzák az egyes szubkládokat meghatározó markerek derivált allélját. Ezen csoportokat \* jelöléssel különböztetjük meg a haplocsoportoktól. A szubkládok jelöléséhez egy váltakozó alfanumerikus rendszert használnak. Ebben a rendszerben addig váltják egymást a számok az ABC kisbetűivel, amíg a filogenetikai fa ágainak csúcsán található haplocsoportokhoz nem érünk. Ennek a rendszernek a használatával minden egyes haplocsoport elnevezése egyértelműen utal a filogenetikai fán elfoglalt helyére. Az olyan esetekben, amikor egy kládon belül a markereknek csak egy része kerül tipizálásra és egy adott minta a filogenetikai fa egy belső nóduszára sorolható, a haplocsoport elnevezés egy zárójeles rendszer segítségével történik. A zárójelen belül egy x (=excluding) jelzi, hogy melyik az a marker, illetve haplocsoport, ahova az adott kládon belül biztosan nem sorolható a mintánk. Például a P-M45\*(xM173) azt jelzi, hogy a tipizált minta hordozza az M45 marker derivált allélját, vagyis besorolható a P haplocsoportba, de a P klaszteren belül nem tartozik az M173 marker által jellemzett R1 kládba.

Ebben a tanulmányban a haplocsoportok elnevezésénél és ennek megfelelően a minták besorolásánál a 2002-es filogenetikai fának egy tovább finomított változatát vettük alapul, melyet Jobling és Tyler-Smith közöltek 2003-ban (YCC2003 Tree) (3. ábra).



A haplotípus kifejezés a mikroszatelliták (STR markerek) variációi által meghatározott Y-kromoszómális csoportokra utal. Az evolúcióval kapcsolatos vizsgálatokban az STR markereket a biallélikus markerekkel kombinációban használják (de Knijff et al. 1997), analízisükkel a haplocsoporton belüli diverzitás mértékét tudjuk tanulmányozni. A mikroszatellita variációk vizsgálata lehetőséget ad arra, hogy közelítőleg meghatározzuk az Y-kromoszómális leszármazási vonalak egyesülési idejét (mikor élt a legutolsó közös ős (**T**ime to **M**ost **R**ecent **C**ommon **A**ncessor - TMRCA). Ugyanakkor számos tényező bonyolítja a kalkulációt, melyek következményeként a becslések tartalmaznak némi bizonytalanságot (Jobling és Tyler-Smith 2003, Wells 2003). Ennek megfelelően az irodalomban a legutolsó közös ős korát a legvalószínűbb dátum mellett a 95%-os megbízhatósági tartománnyal együtt tüntetik fel. (95%-os konfidencia intervallum: 95% annak a valószínűsége, hogy a mért paraméter valódi előfordulási gyakorisága (átlaga, stb.) a populáción belül a mintában mért előfordulási gyakoriság konfidencia intervallumába esik, és 5% annak a valószínűsége, hogy nem esik ebbe a tartományba. Konfidencia intervallum: a mintában végzett mérés alapján a teljes populációra vonatkozó becslés pontossága.)



3. ábra Az Y-kromoszómális haplocsoportok filogenetikai fája (Jobling és Tyler-Smith 2003)

## **I.6. A modern ember afrikai eredete és elterjedése a Földön az Y-kromoszómális vizsgálatok alapján**

Az Y kromoszóma tanúsága szerint az összes ma élő férfi apai származásvonalainak a gyökerénél egyetlen ősapát találunk, az „Y-kromoszómális Ádámot”. A közös férfi ős korára vonatkozóan számos becslés született. Thomson és munkatársai (2000) szerint a férfi, akitől minden ma élő férfi Y kromoszómája levezethető, 59 ezer évvel ezelőtt élt (95% konfidencia intervallum: 40 000-140 000 év). Pritchard és munkacsoportja (1999) 46 000 (95% konfidencia intervallum: 16 000-126 000 év) évben határozta meg Ádám korát. A közös férfi ős korával kapcsolatos bizonytalanságok ellenére elmondható, hogy egyetlen 200 000 évnél idősebb leszármazási vonalat sem találtak az eddigi vizsgálatok alapján (Jobling és Tyler-Smith 2003). Fontos megjegyezni, hogy az a tény, hogy egy adott ősből levezethetjük az összes ma élő embert, nem azt jelenti, hogy rajta kívül nem éltek mások abban az időben. Egyszerűen az ő leszármazási vonala fennmaradt, míg a többieké kihalt. A közös ős kora nem mutat egyebet, mint az időben visszapörgetve azt a pontot, ahol már nem látunk genetikai diverzitást az Y kromoszóma vonalunkon. Nem reprezentálja fajunk eredetének dátumát. Bár nyilvánvalóan valós személy volt -minden ma élő férfi közös őse- az ő őseiről már lényegében semmit sem tudunk mondani (Wells 2003).

Peter Underhill és munkatársai (2000) 21, a világ minden sarkából származó, 1062 férfi mintáján összesen 166 biallélikus és egy triallélikus polimorfizmust vizsgáltak meg a markerek hierarchikus szerveződésének megfelelően. A szekvenciaváltozatok mintázatából leszármazási fát szerkesztettek a maximális takarékoság modell alapján (maximum parsimony). A diagram szerint Afrikában történt meg a legrégebbi hasadás az Y kromoszóma múltjában, vagyis a férfiak családfájának gyökere Afrikában eredt. A mitokondriális DNS analíziséből származó eredményekkel (Cann et al. 1987) összhangban az Y kromoszómális adatok is igazolták az anatómiailag modern ember afrikai eredetét. Illeszkedve ehhez a képhez, mind a mitokondriális DNS (Cann et al. 1987, Vigilant et al. 1991, Ingman et al. 2000), mind az Y kromoszómális biallélikus és mikroszatellita markerek (Seielstad et al. 1999, Hammer et al. 2001, Underhill et al. 2001), mind az autoszómális STR és SNP polimorfizmusok (Calafell et al. 1998) analízise szerint Afrikában találjuk a legnagyobb számú polimorfizmust, sokkal több változatot, mint bármely más kontinensen. Fajunk genetikai polimorfizmusának zömét

az afrikai népek hordozzák magukban, az európaiak, az ázsiaiak és az amerikai indiánok csupán egy apró szeletkét birtokolják az afrikai falvakban fellelhető rendkívüli változatosságnak. Ugyanakkor minél szélesebb genetikai varianciát figyelünk meg egy adott populációban, az minden bizonnyal annál idősebb is (Wells 2003).

A két legrégebbi haplocsoport, az A és a B klaszter, melyeknél az M168 marker vad típusú allélja fordul elő, szinte kizárólag csak Afrikában van jelen. A kontinensen széles körben elterjedtek, de a gyakoriságuk alacsony (Underhill et al. 2000).

Underhill és munkatársai (2000) 44 000 évben határozták meg az M168 marker korát, mely az Afrikából történő kivándorlást jelzi. Thomson és munkacsoportja (2000) hasonló dátumot kapott a marker korára vonatkozóan: 40 000 év (95% konfidencia intervallum: 31 000-79 000 év). Az M168 marker derivált allélját hordozó őstünket nevezhetnénk Eurázsiai Ádámnak is, hiszen ő minden ma élő nem afrikai férfi utolsó közös őse. Valószínűleg Északkelet-Afrikában, a jelenlegi Szudán vagy Etiópia területén élt. A fiai és unokái által véghezvitt utazás határozta meg az emberi történelem elkövetkező lépéseit (Wells 2003). Leszármazottai 3 nagyobb csoportba különülnek el (3. ábra, 4. ábra).

Az M130 marker által jellemzett Y-kromoszómális vonal (C haplocsoport) leszármazottainak jelenkori elterjedése az afrikai kivándorlás part menti útvonalának hipotézisét bizonyítja. E szerint őseink egy része körülbelül 50 000 évvel ezelőtt hagyta el az afrikai kontinenst és indult el Ázsia déli partjai mentén, India déli csücskét érintve, Délkelet-Ázsiába, majd tovább Ausztrália felé (4. ábra). Ennek a vándorlásnak a régészeti nyomai a tengerszint emelkedése következtében ma nagyrészt rejtve vannak (ötvenezer éve mintegy 100 méterrel alacsonyabb volt a tengerszint). M130-as kromoszómát nem találtak Afrikában, ami azt jelzi, hogy az M168-ból valahol Ausztrália felé, útközben keletkezhetett (4. ábra). Az M130-as marker ugyanakkor a kelet-ázsiai populációkban, köztük a kínaiakban is megtalálható, illetve magas gyakoriságban van jelen Északkelet-Ázsiában, főképp Mongóliában és Szibériában, jelezve, hogy jelentős északi irányú expanzió is lezajlott Délkelet-Ázsia felől (5. ábra). E mellett a genetikai adatok azt igazolják, hogy az M130 vonal az elmúlt 10 000 év során elterjedt Amerikában is. A rendelkezésre álló adatok a partvidéki vándorlás folytonosságára utalnak, mely szerint 50 000 évvel ezelőtt elindult Afrikából, kelet felé haladva elérte Indiát, Délkelet-Ázsiát, végül Ausztráliát, majd észak felé fordulva a

Sarkvidék és Amerika felé vette az irányt. A partvidéki marker ősrégi viszonyt tár fel a kiterjedt térségek lakói közt (Wells 2003).

A második csoportot az M89 marker jelöli (az F szuperhaplocsoportot jellemzi), amely meghatározza a nem afrikaiak fő Y kromoszóma vonalát. M89-es Y kromoszómák nincsenek sem Ausztráliában (az őslakosok körében!), sem Délkelet-Ázsiában, de elég magas gyakorisággal fordulnak elő Északkelet-Afrikában. Ez alapján arra következtethetünk, hogy az M89 kicsivel később alakult ki (körülbelül 45 000 éves), mint az M130, egy olyan populációban, amely hátramaradt Afrikában azután, hogy egy csapat a tengerpartot követve elindult Ausztrália felé. Ezek az M130 nélküli emberek kolonizálták elsőként a Közel-Keletet. Az M89 marker északkelet-afrikai és levantei jelenléte, valamint Levante felső paleolitikus lelőhelyeinek kora együtt azt bizonyítják, hogy a második kivándorlási hullám a Közel-Kelet irányába zajlott le (Wells 2003).

Az Y kromoszóma tanúsága szerint nagyon kevés ma élő európai tudja genetikai eredetét közvetlenül a 45 ezer évvel ezelőtti Levantéig visszavezetni. Ők lehetnek azoknak a kis létszámú közel-keleti bevándorlóknak a leszármazottai, akik elsőként hozták be felső paleolit markereket Európába.

Az M89 leszármazási vonalon számos különböző haplocsoport (G-R) alakult ki az azokat meghatározó mutációk/markerek megjelenésével (4. ábra).

Az M89 marker által meghatározott leszármazási vonal képviselői között körülbelül 40 000 évvel ezelőtt valahol a mai Irán vagy Közép-Ázsia déli területein tűnt fel az M9 marker, mely a K haplocsoport jellemzője. A markert hordozó populációkat együttesen eurázsiai klánnak hívjuk. Az eurázsiai vonalon hozzávetőleg 35 ezer éve jelent meg Közép-Ázsiában az M45 marker (P haplocsoport), mely két másik, nagyon fontos vándorlási útvonalhoz vezetett. Az egyik az Európa irányába visszakanyarodó M173-as vonal (R1 haplocsoport), a másik az M242 (Q haplocsoport) és M3 (Q3 haplocsoport) markerek által jellemzett útvonal Szibérián keresztül Amerikába (4. ábra).

A felső paleolitikumi anatómiailag modern európaiak részben Közép-Ázsiából (M173 marker), részben a Közel-Kelet területéről (M170 marker) eredeztethetők. A Közel-Kelet területéről a neolitikum idején is történt egy jelentős bevándorlás (M172, M201, M35), amely hatással volt az európai génállományra.

A harmadik csoportot részben az Y Alu Polimorfizmus vagy M1 néven ismert marker jelöli, körülbelül 50 000 évvel ezelőtt alakult ki Afrikában. Afrikán kívül ez a

vonala kettéhasad és lényegében ugyanazt az útvonalat járja be, mint a másik kettő (Wells 2003). Ez a polimorfizmus filogenetikailag egyenértékű az M145 és az M203 markerekkel, mind a három marker jelen van a D (M174 marker jellemzi) és az E (M96, SRY4064, P29 markerek határozzák meg) haplocsoportba tartozó Y kromoszómákon, azok közös eredetét jelzik (3. ábra). A D klaszter Ázsiában fordul elő, míg az E haplocsoport elsősorban Afrikában elterjedt, de származéka az E3b-M35 klád Európában is jelen van (5. ábra).

### **I.6.1. Európa paleolitikumi és neolitikumi leszármazási vonalai**

Semino és munkatársai közleménye alapján (2000a) az Európában élő férfiak többsége, több, mint 95%-uk besorolható 10 Y-kromoszómális leszármazási vonal valamelyikébe. Munkájuk során 25 európai és közel-keleti helység 1007 férfi lakójának Y kromoszómáján 22 bináris markert vizsgáltak meg, elsősorban a földművelés terjedésének bizonyítéka után kutatva. Eredményeik szerint a neolitikus közel-keleti markerek által meghatározott leszármazási vonalak a modern európaiak kisebb részében vannak csak jelen. Következtetéseik szerint, melyek összhangban vannak a mitokondriális DNS analíziséből származó eredményekkel (Richards et al. 1998, Torroni et al. 1998), az európai férfiak Y-kromoszómális génkészletének körülbelül 80%-a a paleolitikum idejére vezethető vissza, mintegy 20%-a pedig a neolitikus közel-keleti bevándorlóktól származik. Semino és munkatársai szerint az általuk vizsgált markerek földrajzi megoszlása és becsült kora két paleolitikus és egy neolitikus vándorlási eseményre derít fényt, melyek hatása kimutatható a modern európai génállományban.

A mai európai férfiak körülbelül 50%-ának Y kromoszómáján van jelen az M173 mutáció (R1 haplocsoport), mely körülbelül 30 000 éves. Ez a marker vezet vissza minket Közép-Ázsiába, mely az egyik kiindulópontja volt Európa modern emberek általi benépesülési folyamatának (Wells 2003) (4. ábra). Semino és munkacsoportja szerint (2000a) az M173 egy olyan ősi eurázsiai marker, amit az Európába érkező Homo sapiens vagy magával hozott, vagy érkezésekor alakulhatott ki. A marker korát hozzávetőleg 30 000 évben határozták meg.

Az M173 leszármazási vonal és a később ebből egy pontmutáció (G>-) révén létrejött M17 vonal (R1a1 haplocsoport) Európában egymással ellentétes földrajzi

megoszlást mutat. Az M17 leszármazási vonal Kelet-Európában gyakori, majd előfordulása nyugati irányba csökken. Az M173 (xM17) vonal (R1b haplocsoport) a legnagyobb gyakorisággal Nyugat-Európában van jelen, ugyanakkor előfordulása keleti irányba csökken. (Egy későbbi vizsgálat (Cruciani et al. 2002) kimutatta, hogy az M173 markert hordozó európai férfiak 2 csoportba oszthatók. Az egyik csoportba tartozó férfiak Y kromoszómáján az M173 marker mellett az M17 marker van jelen, míg a másik csoportba tartozóknál az M173 marker az M269 markerrel együtt fordul elő.)

Semino és munkatársai szerint a két leszármazási vonal kontinensen belül tapasztalható ellentétes földrajzi megoszlása úgy értelmezhető, mint a jelenlegi Ukrajna területéről (M17 vonal) és az Ibériai félszigetről (M173 (xM17) vonal) (mint jégkorszaki „menedékek”-ből (refúgiumokból) kiinduló elszigetelt népesség-magvak expanziójának bélyege az utolsó jégkorszak leghidegebb időszakát (Last Glacial Maximum - LGM) követően. (A jégkorszak alatt (13-20 ezer évvel ezelőtt) a különböző humán csoportok arra kényszerültek, hogy az észak-balkáni menedéket kivéve elhagyják Közép-Európa területét.)

A mai európai Y-kromoszómális génkészlet körülbelül 20%-át jellemzi az M170-es mutáció jelenléte, mely az I haplocsoportot határozza meg. Ez az egyetlen olyan nagyobb Y-kromoszómális haplocsoport, mely széles körben elterjedt Európában, de gyakorlatilag máshol nem fordul elő (Semino et al. 2000a, Rootsi et al. 2004). Ez a mutáció Európában alakulhatott ki a paleolitikum során (körülbelül 22 000 éve) a Közel-Keletről mintegy 20-25 ezer évvel ezelőtt érkező bevándorlók utódaiban. Az M170 vonal (I haplocsoport) az M89 marker által meghatározott leszármazási vonalak (F-R haplocsoport) közé tartozik.

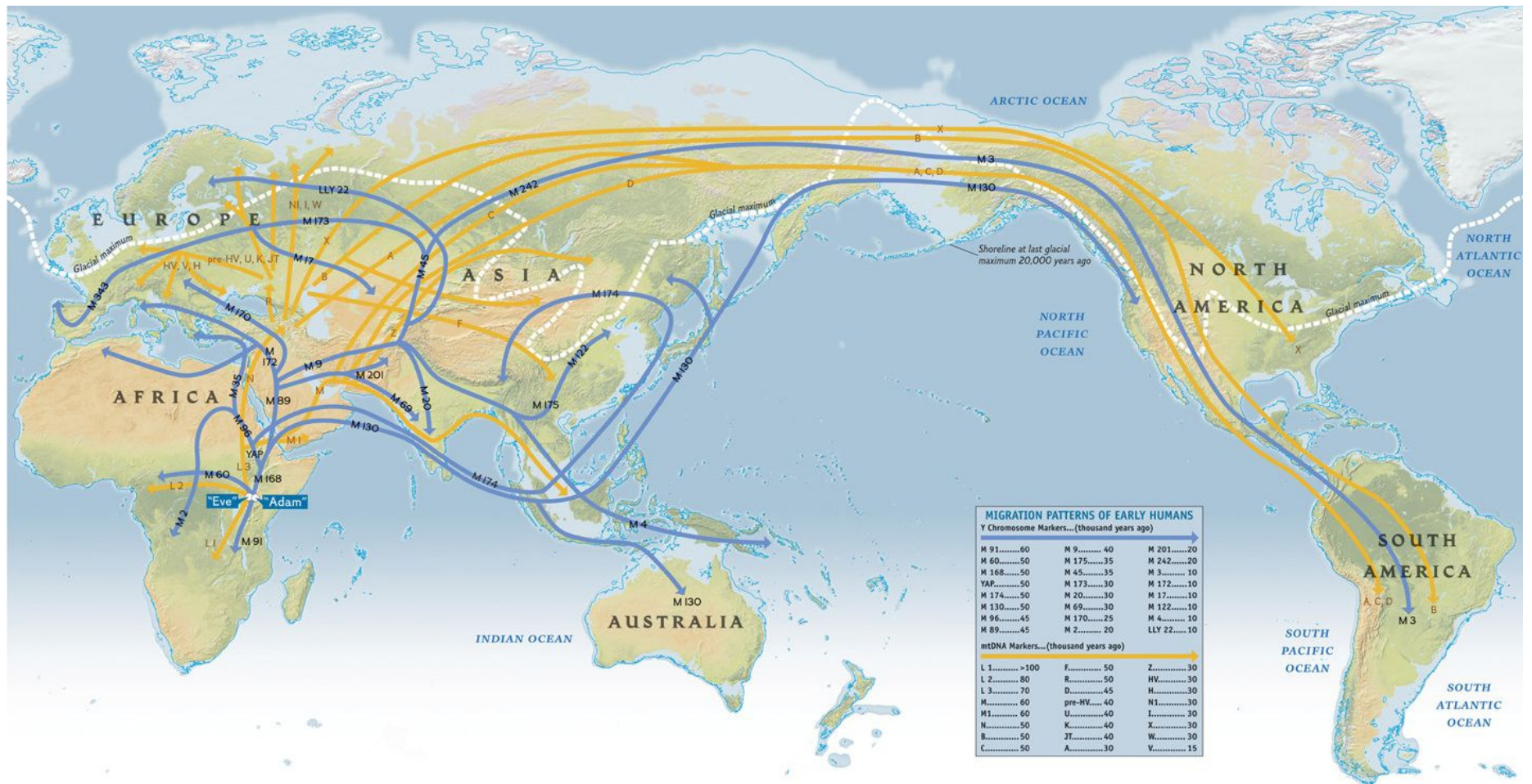
Az európai Y kromoszómák körülbelül 22%-a sorolható be a neolitikus közel-keleti markerek által meghatározott leszármazási vonalak valamelyikébe.

Az M35 marker (E3b haplocsoport), az M172 marker (J2 haplocsoport), az M201 marker (G haplocsoport) és az M89 marker (F haplocsoport) által jellemzett leszármazási vonalak előfordulási gyakorisága a Közel-Kelet felől Európa irányába csökken. Semino és munkatársai szerint az E3b, a G és a J2 (valójában az egész J) haplocsoport a neolitikum idején a Közel-Keletről érkező földművelő népek genetikai hatását tükrözi Európa férfi génállományán. Vizsgálataik alapján a neolitikus gazdálkodók genetikai hatása a mediterrán partvidék mentén kifejezettebb.

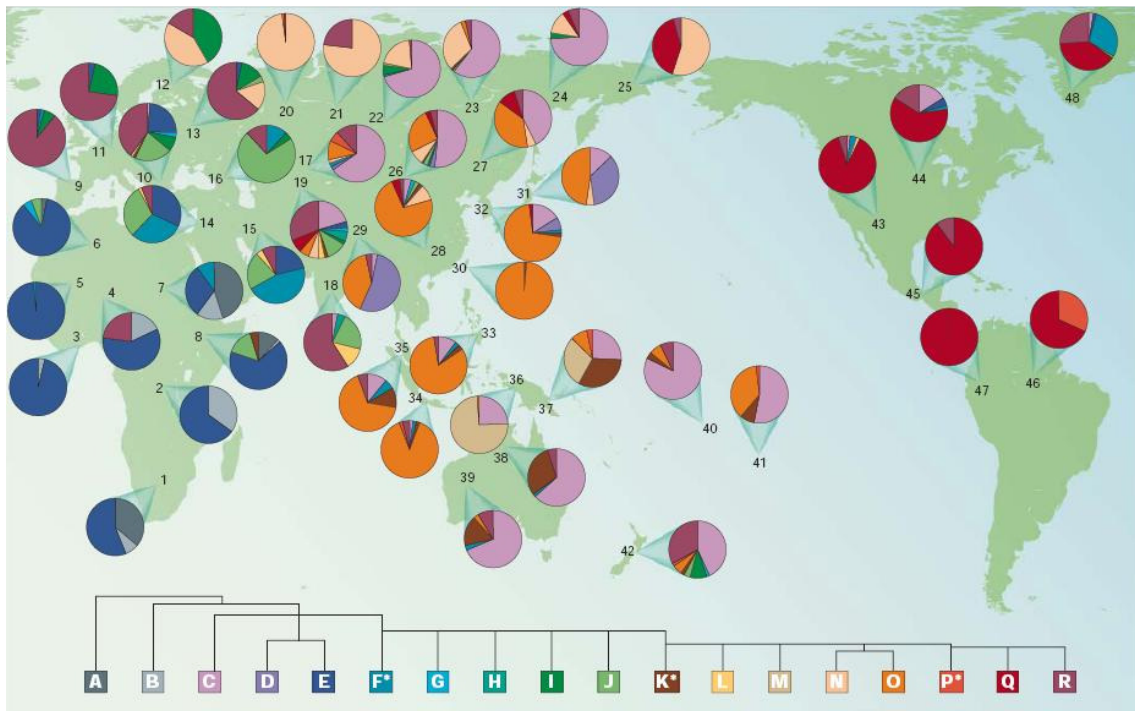
A J2 (az egész J), a G és az F haplocsoportot a M89 marker egyesíti (3. ábra), míg az E3b klaszter, illetve a csoportot meghatározó M35 marker az YAP<sup>+</sup> marker által meghatározott leszármazási vonalak közé tartozik. Az irodalom szerint (Semino et al. 2000a) a legtöbb európai férfi, akinek Y kromoszómáján jelen van az YAP, az E3b haplocsoportot meghatározó M35 mutációt is hordozza.

A fent említett leszármazási vonalakon kívül a 10 Y-kromoszómális leszármazási vonal között szerepelt a Tat vagy M46 mutáció által meghatározott leszármazási vonal (N3 haplocsoport) is. Semino és munkacsoportja az általuk vizsgált minták közül az uráli nyelvű számi, udmurt és mari populációban detektálta számottevő gyakorisággal a Tat marker jelenlétét. Ugyanakkor azt találták, hogy a szintén uráli nyelvet beszélő magyar populációban egyáltalán nincs jelen ez a polimorfizmus. Zerjal és munkatársai szerint (1997) a marker körülbelül 4000 éves és egy Észak-Európára korlátozódó újkeletű uráli migrációra utal.





4. ábra A Föld modern emberek általi benépesítésének folyamata (kék szín: Y-kromoszómális leszármazási vonalak elterjedése, sárga szín: a mitokondriális leszármazási vonalak elterjedése) Forrás: National Geographic Maps, Atlas of the Human Journey



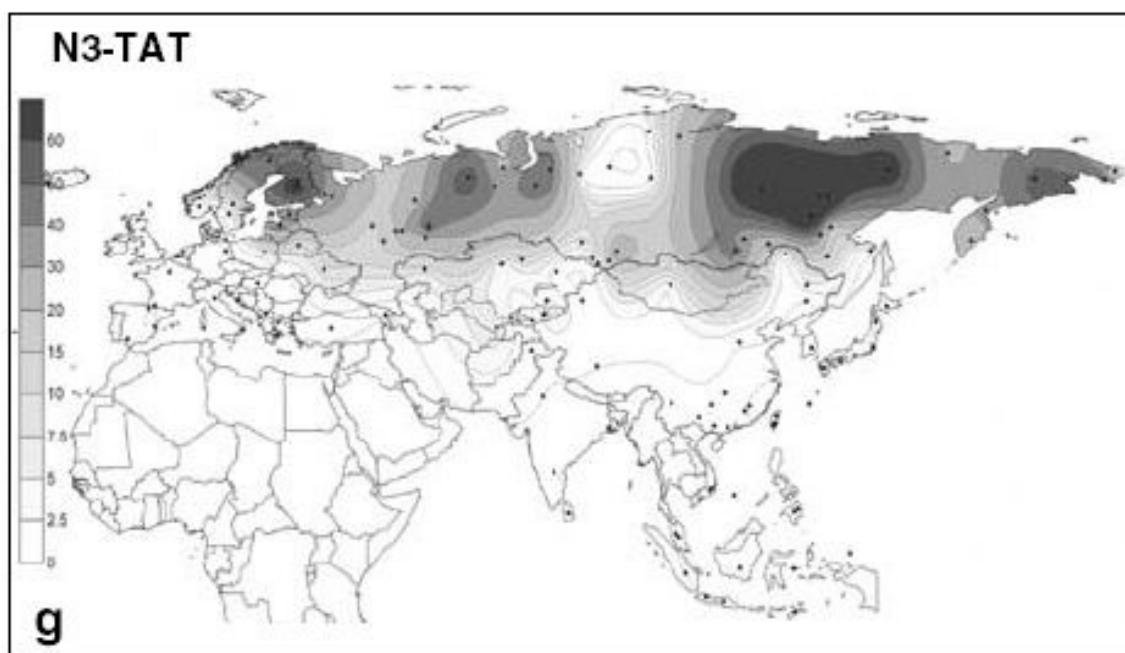
5. ábra A 18 nagyobb Y-kromoszómális haplocsoport földrajzi megoszlása (Jobling és Tyler-Smith 2003)

Az egyes kördiagrammok a különböző Y-kromoszómális haplocsoportok egy-egy populációban detektálható előfordulási gyakoriságát mutatják. Megfigyelhető, hogy míg a szomszédos populációk Y-kromoszómális összetétele, illetve gyakorisági mintázata egymáshoz hasonló, addig a világ különböző részein élő populációk között nagy mértékű különbségek vannak.

(A feltüntetett populációk: 1, !Kung; 2, Biaka Pygmies; 3, Bamileke; 4, Fali; 5, Senegalese; 6, Berbers; 7, Ethiopians; 8, Sudanese; 9, Basques; 10, Greeks; 11, Polish; 12, Saami; 13, Russians; 14, Lebanese; 15, Iranians; 16, Kazbegi (Georgia); 17, Kazaks; 18, Punjabis; 19, Uzbeks; 20, Forest Nentsi; 21, Khants; 22, Eastern Evenks; 23, Buryats; 24, Evens; 25, Eskimos; 26, Mongolians; 27, Evenks; 28, Northern Han; 29, Tibetans; 30, Taiwanese; 31, Japanese; 32, Koreans; 33, Filipinos; 34, Javanese; 35, Malaysians; 36, West New Guineans (highlands); 37, Papua New Guineans (coast); 38, Australians (Arnhem); 39, Australians (Sandy Desert); 40, Cook Islanders; 41, Tahitians; 42, Maori; 43, Navajos; 44, Cheyenne; 45, Mixtecs; 46, Makiritare; 47, Cayapa; 48, Greenland Inuit)

### I.7. A Tat (M46) marker által meghatározott leszármazási vonal

A Tat polimorfizmus pontmutáció, T>C tranzíció, az Y kromoszóma hosszú karjának proximális részén, az RBF5 lókuszon helyezkedik el. A markert 1997-ben Zerjal és munkacsoportja fedezte fel. A Tat C allél az N3 haplocsoportot határozza meg, amely széles körben elterjedt Észak-Eurázsia populációiban (6. ábra), Y-kromoszómás génkészletük jelentős részét képezve (Zerjal et al. 1997, 2001, Rootsi et al. 2000, 2007, Rosser et al. 2000, Semino et al. 2000a, Tambets et al. 2001, 2004, Lell et al. 2002). Ázsiában az északi területeken élő populációkban fordul elő számottevő gyakorisággal. Európában az északi és a keleti területeken élő népcsoportokban van jelen, ugyanakkor szinte egyáltalán nem található meg a kontinens nyugati és déli részein. Gyakoriságában egy éles kelet-nyugati irányú csökkenést mutat Skandináviában a finnugor- (pl. sámik- 47,2%) és a germán nyelvű populációk (norvégok, svédok 4-8%) között, valamint Litvánia (33%) és Lengyelország (3,2%) populációi között.



6. ábra A Tat leszármazási vonal előfordulási gyakorisága (Rootsi et al. 2007)

A marker előfordulása az egyes populációkban független attól, hogy az adott populáció milyen nyelvet beszél. Jelen van számos uráli-, indoeurópai (keleti szláv és balti), csukcs-kamcsatkai és altáji nyelvet beszélő populációban.

A marker korára és kialakulási helyére vonatkozóan különböző becslések születtek az irodalomban. Zerjal és munkatársai (1997) vizsgálata alapján ez a pontmutáció körülbelül 2000-4000 évvel ezelőtt keletkezett Mongólia/Kína területén. A szerzők a Tat C allél jelenlétét Európában a legnagyobb gyakorisággal az uráli nyelvű populációkban detektálták, amely alapján feltételezhető, hogy ezek a populációk fontos szerepet játszottak a mutáció egészen Finnország területéig történő nyugati irányú elterjesztésében. Véleményük szerint a marker előfordulása az észak-európai populációkban jelentős ázsiai genetikai hatás következménye.

Lahermo és munkatársainak (1999) számításai alapján ez a pontmutáció egy viszonylag újkeletű esemény, kora körülbelül 4440 év (95%-os konfidencia intervallum: 3140-6200 év).

Az előbbiekkal ellentétben Rootsi és munkacsoportja (2000) a Tat C allélhoz kapcsolódó mikroszatellita diverzitás értékét Kelet-Európában magasabbnak találta, e szerint valószínűbbnek tartották, hogy a mutáció Kelet-Európában jelent meg és innen terjedt el Észak-Euráziában. Ugyanakkor Rootsi és munkatársai egy későbbi vizsgálat (2007) eredménye alapján arra a következtetésre jutottak, miszerint a marker valószínűleg a mai Észak-Kína területéről származik, majd ezt követően terjedt el Szibérián keresztül Kelet-Európába. (Északkelet-Európában az N3 vonal másodlagosan terjedt el.) Néhány észak-európai populációban (például a finneknél) magas STR varianciát tapasztaltak. Véleményük szerint a magas variancia érték ezekben a populációkban nem az alapító vonal hosszú idejű in situ differenciációjának eredménye, hanem sokkal inkább a különböző N3 alapító típusok keveredésének következménye. Ebben a tanulmányban a szerzők a marker korát  $11,8 \pm 6,8$  ezer évben határozták meg az Észak-Kínában megfigyelhető N3-STR variáció alapján.

## **I.8. Magyar nyelvű populációk eddigi genográfiai vizsgálati eredményei a klasszikus, a mitokondriális és az Y-kromoszómális SNP markerek felhasználásával**

Guglielmino és munkatársai (2000) 24 klasszikus genetikai marker (vércsoportok, vörösvérsejt enzimek, HLA-A, HLA-B gének) nyolc magyar népcsoport (őrségyiek, palócok, matyó népcsoport, székelyek, csángók, kiskunok és nagykunok, jászok), egy budapesti kevert minta, egy délnyugat-és egy észak-magyarországi populáció, valamint 10 nem magyar populáció esetében megfigyelt gyakorisági értékeit hasonlították össze. Eredményeik alapján a ma élő magyarok genetikailag közeli rokonságban vannak más európai populációkkal, ugyanakkor a szerzők szerint feltételezhető, hogy az egyes magyar népcsoportok jobban megőrizték eredeti génkészletüket, mint a nem specifikált (egyik népcsoporthoz sem tartozó) magyar populáció. Guglielmino és munkacsoportja (1990) szerint az uráli gének a teljes magyar génkészlet körülbelül 13%-át adják. Itt mindenképpen fontos megjegyezni, hogy ezen tanulmányok tárgyi tévedéseit a történettudomány területéről újabban számos kritika érte (Bálint 2006).

Az anyai ágon öröklődő mitokondriális DNS markerek vizsgálata (Lahermo et al. 2000, Bogácsi-Szabó et al. 2005, Tömöry et al. 2007) szintén a modern magyar populációk, köztük a székelyek és a csángók más európai népeiségekhez való genetikai hasonlóságát igazolta. Mindamellet a X-XI. századból származó magyar és a modern eurázsiai populációk mitokondriális DNS variációinak statisztikai analízise alapján az ősi magyar populáció az ázsiai és az európai populációk közé, legközelebb a török-, az ukrán- és a finnugor komi populációhoz térképeződik (Tömöry et al. 2007).

Az előbbieken felsorolt vizsgálatok azt mutatják, hogy a ma élő magyar populációk genetikailag nem különülnek el az őket körülvevő indoeurópai nyelvű populációktól.

Rosser és munkatársai (2000) 36 magyar férfi mintáján 11 Y-kromoszómális biallélikus polimorfizmust tipizáltak. Az előzőekhez hasonlóan arra a következtetésre jutottak, hogy a modern magyarok genetikailag az őket körülvevő populációkhoz hasonlóak. Munkájuk során összesen 47 európai és Európához közel élő populációt analizáltak. Vizsgálataik alapján az európai populációk apai ági rokonságát elsődlegesen a földrajzi közelség és nem a nyelvi rokonság határozza meg.

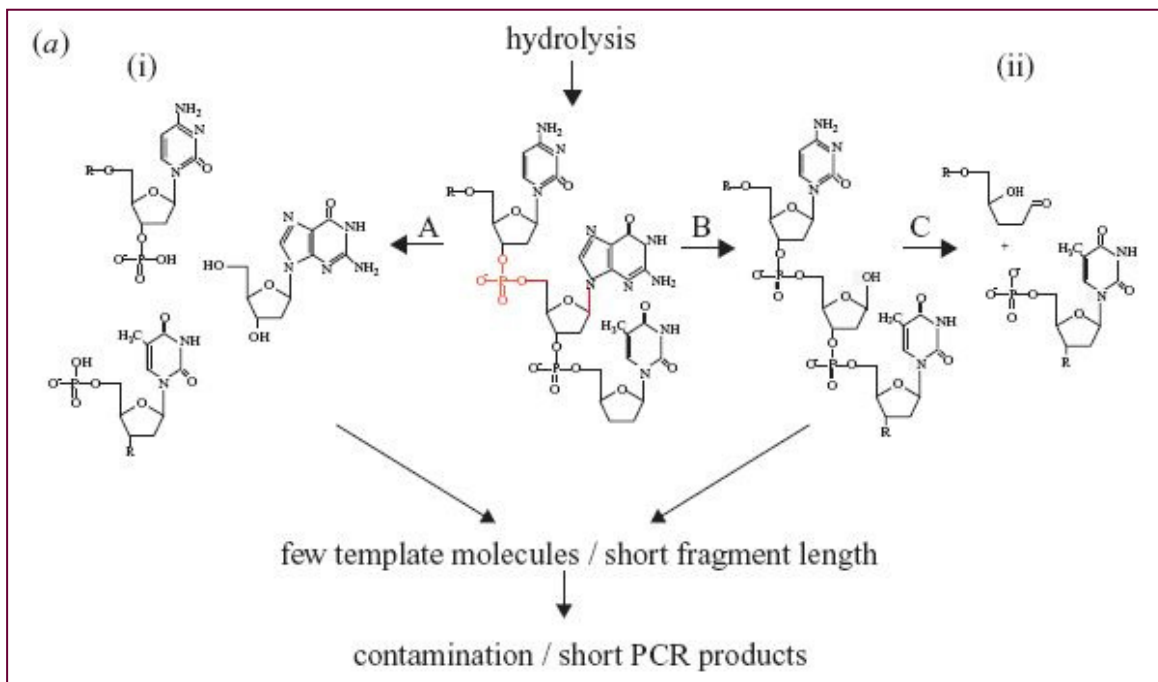
Semino és munkacsoportja (2000a) 25 európai és közel-keleti helység 1007 férfi lakójának Y kromoszómáján 22 Y-kromoszómális bináris markert vizsgált meg. Eredményeik alapján az Európában élő férfiak többsége, több, mint 95%-uk besorolható 10 Y-kromoszómális leszármazási vonal valamelyikébe, ahogy azt már az előzőekben említettük. Az általuk vizsgált elsősorban palóc férfiak (45 fő) mintáiban legnagyobb gyakorisággal (60%) az Eu19 haplotípust detektálták, mely a 2002-es egységesített nomenklátúra alapján az R1a1 haplocsoportnak felel meg (az M17 marker jellemzi). Az adatok statisztikai vizsgálata alapján az általuk tipizált “magyar” csoport közel térképeződik a lengyel és az ukrán populációhoz.

A Tat (M46) marker által meghatározott leszármazási vonal különös érdeklődésre tart számot a finnugor nyelvű populációk vizsgálatokor (Zerjal et al. 1997). Az irodalom alapján a Tat C allél minden eddig vizsgált uráli nyelvű populációban megtalálható, ugyanakkor a modern magyar populáció kivételt képez ez alól, gyakorlatilag nem fordul elő ez az Y variáns sem a csángó, sem a palóc, sem más magyarul beszélő populáció génkészletében (Zerjal et al. 1997, Lahermo et al. 1999, Rootsi et al. 2000, Rosser et al. 2000, Semino et al. 2000a, 2000b, Tambets et al. 2001, 2004).

Ezen megfigyelések alapján felvetődik a kérdés, hogy vajon a honfoglalás kori populációban jelen volt-e ez a mutáció vagy sem.

## I.9. Az archaikus DNS jellemzői

A nukleinsavak postmortem instabilitása az archaikus DNS-t vizsgáló kutatások központi módszertani problémája (Willerslev és Cooper 2005). A metabolikusan aktív szövetekben a DNS molekulákat érő károsodások a repair folyamatokon keresztül gyorsan és hatékonyan javításra kerülnek (Lindahl 1993). Ezzel szemben az inaktív sejtekben a DNS károsodások idővel halmozódnak, melynek következtében a régészeti minták vagy már nem tartalmaznak amplifikálható endogén DNS-t vagy degradálódott, rövid, 100-500 bp hosszúságú, DNS fragmentumok vannak jelen a mintákban (Pääbo 1989, Handt et al. 1994, Höss et al. 1996). A postmortem DNS degradálódásához nagymértékben hozzájárul a mikroorganizmusok, a talajban élő gerinctelenek és a celluláris nukleázok aktivitása, valamint az oxidáció és a hidrolízis módosító hatása (Binladen et al. 2006). A postmortem DNS módosulásokat a következő ábrák (7. a, b, c) mutatják.

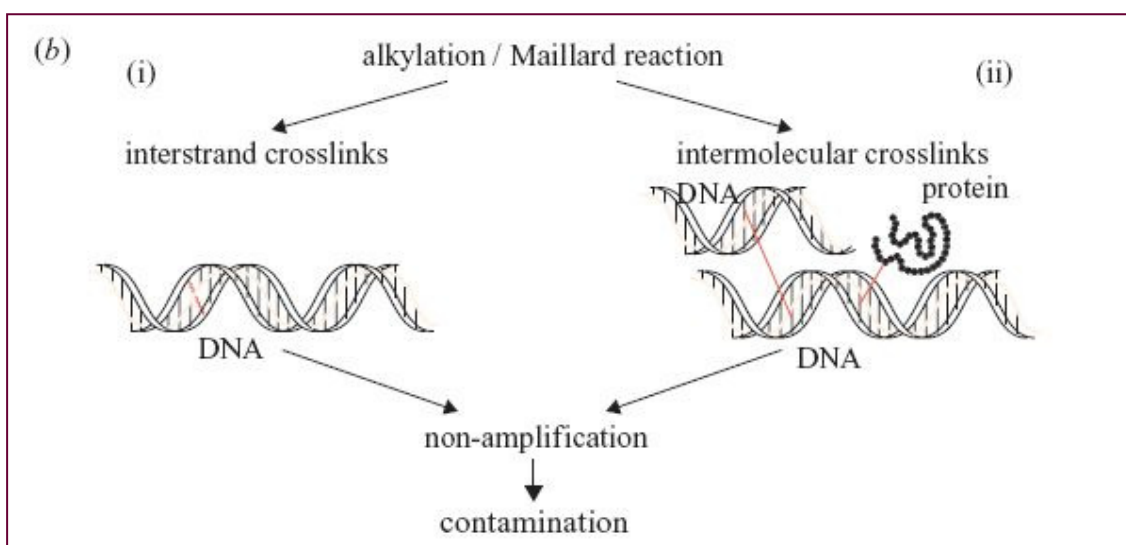


7. ábra A DNS postmortem módosulásai a fosszilis maradványokban (Willerslev és Cooper 2005)

a, Száltörések kialakulása a hidrolitikus károsodás során. (i)(A) A foszfodiészter kötés hidrolitikus hasítása a cukor-foszfát láncban. (ii) (B) A glikozidos kötés hidrolitikus hasításával abázikus hely (AP hely) képződik (az ábrán depurináció történik). (C) A cukor-foszfát lánc töreése  $\beta$ -elimináció következtében. (A PCR rövidítés jelentése **P**olymerase **C**hain **R**eaction – magyar elnevezése: polimeráz láncreakció.)

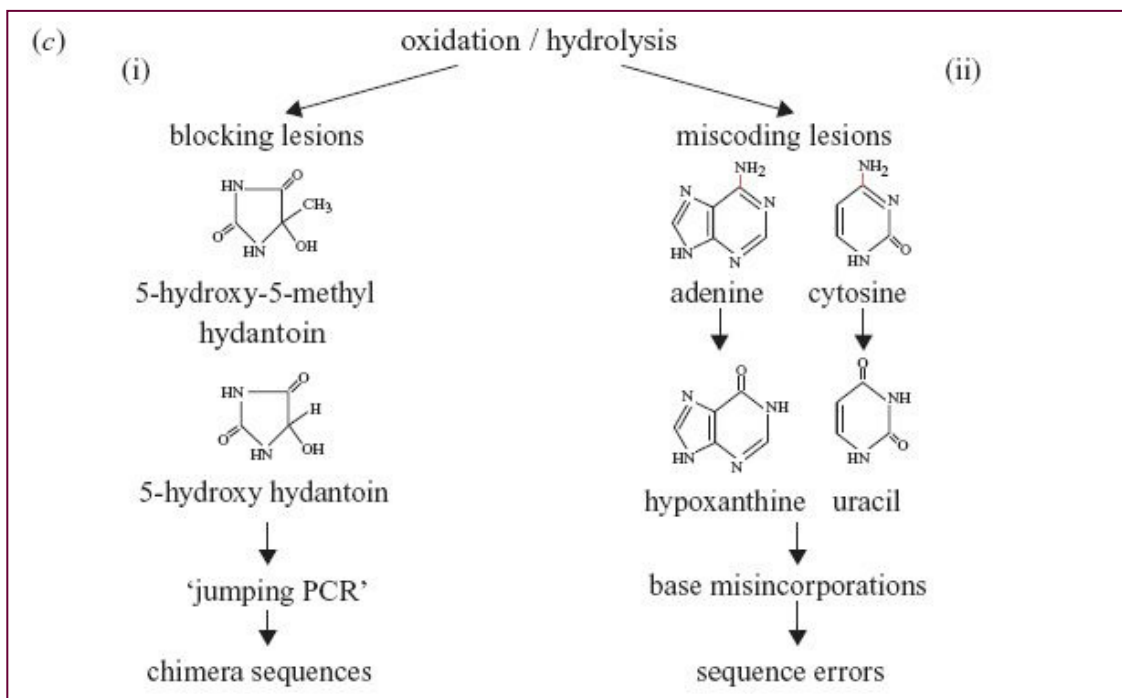
Úgy gondolják, hogy nagymértékben a száltörések felelősek azért, hogy a fosszilis maradványok esetében a DNS nagy része elvész, illetve csak rövid fragmentumok amplifikálhatók.

Száltörések keletkezhetnek a mikroorganizmusok károsító hatásának, illetve a postmortem sejtekben felszabaduló nukleázok aktivitásának eredményeként is (Pääbo et al. 2004).



b, A keresztkötések kialakulásának különböző típusai. (i) A DNS szálak között alkiláció révén kialakuló keresztkötések. (ii) A DNS és más biomolekulák között a Maillard reakció eredményeként kialakuló keresztkötések. A Maillard termékek a DNS-ben lévő cukrok, valamint a fehérjék és a nukleinsavak elsődleges aminocsoportjai közötti kondenzációs reakció során képződnek (Pääbo et al. 2004). A keresztkötések megakadályozhatják az endogén templát molekulák amplifikációját, így módon megnövelik a szennyeződés veszélyét.





c, A bázisok oxidatív és hidrolitikus módosulásai (i) blokkolhatják a PCR-t vagy (ii) hibás bázis beépülését eredményezhetik. Néhány oxidatív károsodás olyan léziókat eredményez, melyek gátolják a polimeráz enzim működését és a 'jumping' PCR jelenségen keresztül elősegítik a kiméra szekvenciák képződését. Ebben az esetben a sérülés helyénél mintegy „leugrik” a DNS polimeráz enzim a templátról és egy másik templátról folytatja a szintézist, amelynek eredményeképpen az eredetitől eltérő termék keletkezik (Pääbo et al. 1990). (A száltörések, az apurinált helyek és az UV fény által indukált DNS károsodások szintén kiváltó okai lehetnek a 'jumping PCR' jelenségének.) (Az ábrán feltüntetett hydantoin származékok a timin és a citozin bázisok oxidatív károsodása során keletkeznek.)

Az oxidatív károsodás során a jelenlevő szabadgyökök (peroxid-, hidrogén-peroxid és hidroxil gyökök) bázisvesztéshez (AP hely), bázismódosuláshoz, illetve a dezoxiribóz molekula fragmentálódásához vezethetnek (Burger et al. 1999, Pääbo et al. 2004).

A hidrolitikus károsodás a bázisok aminocsoportjainak elvesztésével járhat. Ennek során az adenin hypoxantinná, a citozin uracillá, az 5-metil-citozin timinné, a guanin xantinná alakul át, melynek következtében az amplifikáció során hibás bázis épül be az újonnan szintetizálódó DNS szálakba.

Az előbb felsorolt hatások következtében az archaikus DNS általában korlátozott számban, illetve erősen fragmentált és nagymértékben károsodott formában van jelen a leletekben. Archaikus leletek vizsgálata esetén nem elhanyagolható az a tény, hogy a mtDNS több száz, vagy ezer példányban van jelen egyetlen sejtben. Ezért ősi leletekben nagyobb az esélye annak, hogy épen maradt, molekuláris biológiai módszerekkel vizsgálható mitokondriális DNS szakaszokat találjunk (Handt et al. 1994). Ezzel

szemben a fossziliákból történő Y tipizálás esetén problémát jelenthet, hogy az Y kromoszóma egyetlen példányban található meg a férfiak minden egyes sejtjében. Általában nukleáris DNS esetén a mitokondriális DNS vizsgálatokhoz képest rövidebb DNS fragmentumok amplifikálhatók (Binladen et al. 2006). Ennek oka lehet egyszerűen a templát mennyisége, hiszen a mitokondriális DNS nagyobb kópiaszáma megnöveli a viszonylag hosszabb intakt molekulák jelenlétének valószínűségét. Ugyanakkor az is előfordulhat, hogy a nukleoszóma mag (146 bázispárnyi nukleáris DNS hiszton fehérje magra tekeredik fel a kromoszómális szerveződés során) mintegy „védelmi (megőrző) egységet” („preservation unit”) képez és a száltörések a nukleoszómákat összekötő linker régiókon jönnek létre. Mindazonáltal a kromoszómális fehérjék (például a hisztonok) DNS-hez való közelsége megnöveli a fehérje-DNS keresztkötések kialakulásának esélyét, melyek gátolhatják a polimeráz enzim működését az amplifikáció során (Binladen et al. 2006).

Ugyanakkor a környezeti tényezők jelentősen befolyásolják a DNS degradációjának mértékét. A régészeti leletekben a DNS megőrzésének kedvez a mikroorganizmusok hiánya, illetve az UV sugárzás és a radioizotópok hiánya, a szárazság, a DNS molekulák kötődése ásványi felülethez, pl. hidroxapatitához, az alacsony hőmérséklet, a neutrális vagy kissé alkalikus pH érték (Burger et al. 1999).

A DNS megőrzésében a környezeti tényezők nagyobb szerepet játszanak, mint az idő, vagyis nincs általános összefüggés a lelet kora és az endogén DNS degradáltságának mértéke között (Höss et al. 1996, Poinar et al. 1996).

A régészeti leletek gyakran nagy mennyiségű gátlóanyagot tartalmaznak, melyek az archaikus DNS-sel együtt tisztulnak és megakadályozzák annak amplifikálását. Ezek a gátlóanyagok származhatnak a minta környezetéből, ide tartoznak a humuszsavak, a fulvosav, a hidroxapatit, a tannin és a kontamináló, nem endogén DNS (Kalmár et al. 2000). A biológiai anyag (a csont és a fog szöveteinek) degradálódása során keletkező inhibitorok az I-es típusú kollagén és a Maillard-termékek (Scholz et al. 1998).

Különbféle anyagok, módszerek léteznek, melyek alkalmazásával a postmortem módosulások ellenére is lehetőségünk nyílik az archaikus DNS felsokszorozására, megbízható vizsgálatára.

Az N-fenaciltiazólium bromiddal (PTB) (Vasan et al. 1996) történő kezelés során a biomolekulák közötti keresztkötések (Maillard termékek) felhasadnak, ezáltal megnő az

esélye az olyan archaikus DNS molekulák amplifikációjának, melyek egyébként addig nem voltak elérhetőek a PCR számára.

Az uracil -N-glikoziláz (UNG) enzim a citozin deaminációs termékeit távolítja el a DNS-ből (Pääbo et al. 2004).

A 3'-5' exonukleáz (más néven proofreading) aktivitással rendelkező high fidelity polimeráz enzimek, mint a Pfu és a TaqHiFi polimerázok, használata lehetővé teszi, hogy nagyobb hűséggel történjen a templát DNS átírása, ily módon megnövelik az amplifikáció hatékonyságát (Willerslev és Cooper 2005).

## II. Célkitűzések

Kísérletsorozatunk egyik célja, hogy megvizsgáljuk, hogy a Tat marker derivált C allélja, amely széles körben elterjedt az uráli nyelvű populációkban, ugyanakkor az eddigi vizsgálatok szerint gyakorlatilag hiányzik a ma élő magyar nyelvű populációkból, jelen volt-e a honfoglalás kori populáció génkészletében. A kérdés tisztázásához elsősorban X. századi mintákon tanulmányoztuk a Tat polimorfizmus allélikus állapotát. A minták Y-kromoszómális analíziséhez olyan módszert kellett bevezetnünk, amely amellett, hogy megnöveli a vizsgálatok hatékonyságát, alkalmazásával ismételtető és hiteles eredményeket kapunk.

Munkánk másik célja, hogy a ma élő magyar és székely populáció apai ági genetikai diverzitását részleteiben vizsgáljuk, további adatokat nyerve a Tat C allél előfordulására vonatkozóan is. Választ kerestünk arra is, hogy a magyar és a székely populáció Y-kromoszómális genetikai összetételét tekintve mennyire heterogén, illetve hogyan viszonyul egymáshoz és más európai populációkhoz. Ennek érdekében 100, elsősorban az Alföld területéről származó magyar és 97 korondi székely férfi mintáján 22 biallélikus markert tipizáltunk, melyek az Y kromoszóma nem rekombinálódó régióján lokalizálódnak. A kapott eredményeket összehasonlítottuk más európai populációk Semino és munkatársainak vizsgálatából (2000a) származó adataival.

### **III. Anyagok és módszerek**

#### **III. 1. A vizsgált minták**

Munkánk során 100 magyar, 97 székely és 7 (IX-)X. századi csontlelet DNS mintáját vizsgáltuk meg.

A vizsgálatba bevont apai ágon nem rokon magyar férfiak Magyarország különböző területein születtek, de a leginkább reprezentált régió az Alföld területe volt (90/100). A székely minták apai ágon nem rokon, önkéntes donoroktól származnak, akik Erdély területén, Korondon születtek (8. ábra). 94 magyar és 97 székely férfi vérből izolált DNS mintáját a Szegedi Tudományegyetem Igazságügyi Orvostani Intézete bocsátotta rendelkezésünkre. 6 magyar férfi esetében hajszálból izoláltunk DNS-t a laboratóriumban.

A 7 csontlelet 6 különböző, jól dokumentált ásatásból származik a Kárpát-medence területéről (8. ábra). Hat lelet egyértelműen a X. századra datálható. A hatból négy lelet (B1/3c, 13/1, 13/7, 14/B) mellett klasszikus, gazdag mellékletanyagot tártak fel az ásatás során, amely egyértelműen a X. századi magyar temetkezési szokásokat jeleníti meg (Szőke 1962, Révész 1996, Mesterházy 1997). A négy lelet a szegényes mellékletű T2/41 mintával együtt 4 különböző 10-11. századi temetőből származik. A szarvas-kákapusztai temetőt elsősorban a 9. században használták, de 2 sír a 10. századra datálható; ezek egyike a 378/B lelet sírja. A csekeji 9-11. századi temető területén feltárt 48/B lelet esetében nem egyértelmű a datálás; csak a leletanyag alapján a sír pontos korát nem lehet megállapítani, származhat a IX. és a X. századból is.

A csontminták régészeti és embertani adatait az 1. táblázat tartalmazza. A leletek nemét az antropológiai vizsgálat során férfinak határozták meg.



8. ábra A vizsgált minták származási helyét szemléltető térkép

A számok (1-6) a csontminták temetkezési helyét mutatják (1. táblázat)

A vizsgálatba bevont magyar férfiak (100 fő) Magyarország különböző területein születtek, de a leginkább reprezentált régió az Alföld területe volt (90/100). A székely minták (97 fő) Korondról, Erdély területéről származnak.

1. táblázat A vizsgált csontminták régészeti és embertani jellemzői

Minta <sup>a</sup>	Temetkezési hely <sup>b</sup>	Sírszám	Csont, melyből a mintavétel történt	Antropológiai jellemzők		
				Nem és elhalálozási kor	Taxonómia	Mellékletanyag típusa
B1/3c	Örménykút (1)	3c	jobb femur	férfi, 61-67 éves	Europo-Mongolid	klasszikus
13/1	Szabadkígyós-Pálliget (2)	1	bal femur	férfi, 45-50 éves	-	klasszikus
13/7	Szabadkígyós-Pálliget (2)	7	jobb femur	férfi, 25-30 éves	Europid (Cromagnoid)	klasszikus
T2/41	Mözs-Szárazdomb (3)	41	bal femur	férfi, 24-28 éves	-	szegényes
14/B	Harta (4)	17	jobb femur	férfi, 53-59 éves	-	klasszikus
378/B	Szarvas-Kákapusztá (5)	16	jobb femur	férfi, maturus	-	szegényes
48/B	Csekej (6)	209	bal femur	férfi, középkorú	-	szegényes

<sup>a</sup>Az *Annals of Human Genetics* folyóiratban megjelent Y-chromosome analysis of ancient Hungarian and two modern Hungarian-speaking populations from the Carpathian Basin című publikációban (Csányi et al. 2008) az első 4 csontminta került közlésre. Ebben a minták elnevezése: 13/7-anc4, T2/41-anc19, B1/3c-anc21, 13/1-anc28, a korábban megjelent, a mitokondriális DNS vizsgálatok eredményét publikáló tanulmányban (Tömöry et al. 2007) használtaknak felel meg.

<sup>b</sup>A zárójelben feltüntetett számok az ábrán jelölt temetkezési helyekre utalnak.

## **III. 2. DNS extrakció**

### **III. 2. 1. DNS extrakció modern mintákból**

#### **III. 2.1.1. DNS extrakció vérből**

A vérminták esetében a DNS extrakció kisózással (Miller et al. 1988.) vagy a Walsh és munkatársai által kidolgozott Chelex-alapú módszerrel történt (Walsh et al. 1991). A vérmintákból történő DNS izolálást a Szegedi Tudományegyetem Igazságügyi Orvostani Intézetének munkatársai végezték.

##### **III. 2.1.1.1. DNS extrakció vérből kisózással**

Az EDTA-s vérből a DNS kinyerése NaCl-os extrakcióval, Miller és munkatársai módszere alapján történt.

###### *1. Vörösvértestek hemolízise:*

Az EDTA-s vért 0,05 M KCl oldattal 3x mossuk. A vér és a KCl oldat aránya 1:4. A mosás után hemoglobintól mentes leukocitákat és vörösvértest membránt tartalmazó üledékhez jutunk.

###### *2. Leukociták és leukocita sejtmagok lízise:*

A folyamat során a DNS izolálása Proteináz K enzimmel történik. A proteináz K egyrészt a peptideket bontja szét, másrészt a nukleázok működését akadályozza meg. A Proteináz K működéséhez szükséges az SDS oldat, mely a membrán kettős lipid rétegét emulgeálja, ezáltal aktiválja az enzim működését. Az emésztés folyamata lízispuffer jelenlétében megy végbe, összetétele:

10mM TRIS/HCL pH:8,0

10mM NaCl

10mM EDTA-Na

A leukociták lízisét egy éjszakán át 37 °C-os vízfürdőn 250 µl Proteináz K enzimet (2mg/ml) és 100 µl 20% SDS-t tartalmazó 2 ml lízispufferben történő inkubációval végezzük.



### 3. Fehérje mentesítés (kicsapás)

Az inkubációt követően az oldatban levő fehérjét 1 ml 6M-os NaCl oldattal precipitáljuk, majd 20 percig 3600 fordulaton centrifugáljuk. A DNS a felülúszóban marad.

### 4. DNS extrakció

A felülúszót egy másik centrifugacsőbe leszívjuk. Ehhez 96 %-os hideg etanol adunk. A csövet néhányszor óvatosan megdöntögetjük, a DNS az etanolban összezsapódva látható.

A DNS-t kiemelve mossuk 70%-os etanolban, majd szobahőmérsékleten szárítjuk és TE pufferban oldjuk (10mM TRIS, 1mM EDTA).

## III. 2.1.1.2. DNS extrakció vérből Chelex módszerrel

A vérfoltmintákból a DNS kinyerését Walsh és munkatársai által leírt Chelex módszerrel végezzük.

A Chelex 100 (Biorad) iminodiacetát csoportokat tartalmazó sztírol gyanta, mely a többértékű fémionokkal kelátokat képez. A fémionok (vér esetében az Fe ion) magas hőfokon katalizálják a DNS széttöredezését, a Chelex 100 ezeket a fémionokat köti le.

1. A vérfoltból egy kb. 3x3 mm-es darabot Eppendorf csőbe teszünk és 1 ml steril ioncserélt vízben szobahőmérsékleten 30 percig inkubáljuk. Közben többször keverjük (vortexelés), ezzel a hordozó anyagról a hemoglobin leoldódását elősegítjük.
2. Az inkubációt követően 5 percig 14 000 x g fordulaton centrifugáljuk.
3. A felülúszót óvatosan leszívjuk az alsó 50 µl kivételével.
4. 150 µl 5%-os Chelex oldatot adunk hozzá.
5. 30 perces 56 °C-os vízfürdőn inkubáljuk.
6. Nagy sebességgel vortexeljük 5-10 másodpercig.
7. 8 percig forrásban lévő vízben inkubáljuk
8. Nagy sebességgel vortexeljük 5-10 másodpercig.
9. 10000-15000 x g-vel 2-3 percig centrifugáljuk.

10. A felülúszó 5 µl-ét adjuk templátként az 25 µl végtérfogatú PCR-hez.
11. A maradék mintát -20°C-on tároljuk.

### **III. 2.1.2. DNS extrakció hajszálból**

A hajszálból történő Chelex-alapú DNS extrakciót szintén Walsh és munkatársai dolgozták ki.

1. A hajszálat tisztítjuk steril, deionizált vízben történő többszöri öblítéses, áztatásos mosással.
2. A hajszál hajhagymát tartalmazó végéből egy körülbelül 1 cm hosszú mintát veszünk.
3. A hajszál darabot 200 µl 5%-os Chelex-100 oldatba tesszük.
4. 56°C-on egy éjszakán át (legalább 6-8 óráig) inkubáljuk.
5. Nagy sebességgel vortexeljük 5-10 másodpercig.
6. Ellenőrizzük, hogy a hajszáldarab teljesen elmerüljön a Chelex-oldatban, majd 8 percig forrásban lévő vízben inkubáljuk.
7. Nagy sebességgel vortexeljük 5-10 másodpercig.
8. 10 000-15 000 x g-vel 2-3 percig centrifugáljuk.
9. A felülúszót egy új csőbe pipettázzuk és 5 µl-t teszünk a 25 µl végtérfogatú PCR-hez.
10. A maradék mintát -20°C-on tároljuk.

### **III. 2. 2. DNS extrakció régészeti csontleletekből**

#### **III. 2.2.1. A csontminták előkészítése**

A csont felszíni szennyeződéseit steril, deionizált vízzel és hígított háztartási hypoval eltávolítjuk, majd 1 J/cm<sup>2</sup> UV-C fényel 30 percig besugarazzuk steril fülke alatt. A combcsont diafizisének felületét körben UV-C fényel dekontaminált csiszolókoronggal lecsiszoljuk legalább 2-3 mm mélységig körülbelül 3-4 cm széles hengerfelületen, majd 30 percre ismét besugarazzuk 1 J/cm<sup>2</sup> UV-C fényel. Steril

fúrófejjel mintát veszünk a csont felületéről. A csontport előbb steril Petri csészében gyűjtjük össze, majd steril Eppendorf csőbe tesszük át.

### **III. 2.2.2. DNS extrakció a csontmintákból**

A csontminták esetében a DNS extrakció elsősorban olyan metódus alkalmazásával történt, amely egyesíti a Kalmár és munkatársai által közölt (Kalmár et al. 2000), illetve a Qiagen által javasolt protokollt.

1. 450-700 mg csontport 1,6 ml extrakciós pufferben (0.1 M EDTA, 0.5% N-laurylsarcosine-Na só, 100 mg/ml proteináz K) oldunk fel, majd vortexelés után 37°C-on, folyamatos vertikális forgatással egy éjszakán át inkubálunk.
2. A mintát 13 000 rpm-en 5-10 percig centrifugáljuk, majd 350 µl felülúszót egy 1.5 ml-es Eppendorf csőbe teszünk.
3. A felülúszóhoz adunk 350 µl 4M NH<sub>4</sub>-acetátot és 700 µl 96%-os etanolt és vortexeljük.
4. A DNS-t -70°C-on 7-10 percig vagy -20°C-on egy éjszakán át (vagy 1-2 órát) precipitáljuk.
5. Ettől a ponttól a Qiagen DNeasy Tissue Kit-jét használtuk, követve a gyártó utasításait (Qiagen Cat. no. 69504). Ennek során az elegyet membránon (DNeasy Mini spin column) szűrjük, majd két mosási lépést követően eluáljuk a DNS-t. A javasolt protokolltól az utolsó lépésben tértünk csak el, melynek során a DNS-t a javasolt 100-200 µl helyett 40-60 µl oldatban vettük fel, annak érdekében, hogy növeljük a végső DNS koncentrációt az eluátumban.

## **III. 3. Az archaikus DNS amplifikációja**

### **III. 3. 1. Mitokondriális DNS amplifikáció**

Mivel a mitokondriális DNS megmaradási valószínűsége sokkal nagyobb (Handt et al. 1994), minden egyes csontminta esetében először a mitokondriális DNS

hipervariábilis I régiójának egy 239 bp hosszúságú szakaszát (16182-16420 nukleotidpozíció) amplifikáltuk. (A számozás a Cambridge-i referencia szekvenciára utal (Andrews et al. 1999.)

A reakcióelegy 40 µl végtérfogatban tartalmaz 5-10 µl csont DNS extraktumot, 1,5 U AmpliTaq Gold DNS Polimerázt (Applied Biosystems), 1x AmpliTaq Gold puffert, 3 mM MgCl<sub>2</sub>-ot, 200 µM-t az egyes dNTP-kből (Fermentas), 25 pmol-t az egyes primerekből (L16182-H16401 (2. táblázat)) és 100 µg/ml BSA-t (New England Biolabs).

A PCR lépései:

1. 94 °C 6 perc	} 30 ciklus
2. 93 °C 30 másodperc	
3. 56 °C 40 másodperc	
4. 72 °C 40 másodperc	
5. 72 °C 5 perc	

Az amplifikált termék 10 µl-ét megfuttattuk 8%-os natív poliakrilamid gélen, majd etidium-bromidos festés után ellenőriztük a reakció sikerességét UV transzilluminátor segítségével.

Primer neve	Szekvencia	Referencia
L16182	5'-AACCCCCTCCCATGCTTA-3'	Kalmár et al. 2000
H16401	5'-TGATTTACGGAGGATGGTG-3'	Wilson et al. 1995.

2. táblázat A mitokondriális DNS amplifikációjához használt primerek

### III. 3. 2. Teljes genom amplifikáció

A sikeres és szennyeződéstől mentes mitokondriális DNS amplifikációt követően próbáltunk meg Y-kromoszómális DNS-t kinyerni a csontmintákból. A specifikus Y-kromoszómális analízis előtt minden esetben szükség volt a megmaradt teljes töredékes genom random primerekkel történő dúsítására. A rendelkezésünkre álló archaikus DNS mennyiségének felszaporításához az ún. PEP (**P**rimere **E**xtension **P**reamplification - PEP) módszernek (Zhang et al. 1992) egy tovább fejlesztett

változatát használtuk. Ez az ún. miPEP (**modified improved PEP**) módszer (Hanson és Ballantyne 2005), mely 2 polimeráz enzim, Taq DNS Polimeráz és Tgo DNS Polimeráz, elegyét (Expand High Fidelity Enzyme Mix, Roche Diagnostics) használja nagy mennyiségben (10,5 U/25 µl). A két enzim közül az egyik (Tgo DNS Polimeráz) 3'-5' exonukleáz (más néven proofreading) aktivitással rendelkezik. Kísérleteinkben a Fermentas cég High Fidelity PCR Enzyme Mix termékét (#K0192) használtuk, melyben a proofreading aktivitással rendelkező Pfu DNS Polimeráz egy rendkívül processzív Taq DNS Polimerázzal együtt van jelen.

Reakciókörülmények: Az elegy tartalmaz 25 µl-es végtérfogatban: 40 µM PEP primert (5'-NNNNNNNNNNNNNNNN-3'), 0,8 mM  $\Sigma$ dNTP-t (Fermentas), 2,5 mM MgCl<sub>2</sub>-ot, 1x High Fidelity PCR puffert, 10,5 U High Fidelity Enzyme Mix-et (Fermentas), 100 µg/ml BSA-t (New England Biolabs) és 5-10 µl csont DNS izolátumot. A reakcióelegy összeállítása gyári, nukleázmentes víz (Fermentas) alkalmazásával történt.

A PCR körülmények:

1. 94 °C 1 perc	} 50 ciklus
2. 37 °C 2 perc	
3. 38-54°C 1°C/10 másodperc	
4. 55 °C 4 perc	

Ez a módszer azáltal, hogy nagyobb hősséggel írja át a jelenlévő templát DNS-t, lehetővé teszi, hogy a lókuszt specifikus amplifikáció során a DNS szakaszhoz hatékonyabban tudjon hibridizálni az amplifikációhoz szükséges primerpár. Ily módon megnő a vizsgálat érzékenysége.

Ezzel a módszerrel amellet, hogy nagyobb pontossággal végezhetjük a DNS tipizálást, az alkalmazásával a specifikus reakciók során olyan mennyiségben keletkezik az archaikus DNS esetében is akár Y-kromoszómális PCR termék, hogy a detektáláshoz nem feltétlenül van szükség rendkívül érzékeny, kis mennyiségű DNS kimutatására is alkalmas rendszerre (pl. ALF-ra - **A**utomated **L**aser **F**luorescent **D**N**A** **S**equencer).

A további Y-kromoszómális vizsgálatokhoz az miPEP PCR termék 4-10 µl-ét használtuk.

### III. 3. 3. A Tat polimorfizmust hordozó DNS szakasz amplifikációja

A Tat polimorfizmus vizsgálatakor egy 112 bp hosszú szakaszt amplifikálunk, majd restrikciós emésztéssel határozzuk meg az allélikus állapotot.

A PCR elegy 25 µl végtérfogatban tartalmaz 200 µM-t az egyes dNTP-kből (Fermentas), 1µM-t az egyes primerekből (5. táblázat), 3 mM MgCl<sub>2</sub>-ot (Applied Biosystems), 1U AmpliTaq Gold DNS Polimerázt (Applied Biosystems), 1x AmpliTaq Gold puffert, 100 µg/ml BSA-t (New England Biolabs) és 5-10 µl miPEP PCR terméket. A reverz primer az 5'-végén Cy5 fluoreszcens festékkel jelölt.

A PCR lépései:

1. 94 °C 6 perc	} 38 ciklus
2. 93 °C 30 másodperc	
3. 56 °C 1 perc	
4. 72 °C 45 másodperc	
5. 72 °C 5 perc	

A PCR termék 10 µl-ét kezeltük HpyCH4 IV restrikciós enzimmel (New England Biolabs), mely a C allél jelenléte esetén hasít, egy 27 bp és egy 85 bp hosszú hasítási terméket adva (5. táblázat).

Az amplifikáció sikerességét és a restrikciós emésztés eredményét részben az ALF (Automated Laser Fluorescent DNA Sequencer) rendszeren (ALFexpress II DNA Analyser System, Amersham Biosciences), 4 %-os denaturáló poliakrilamid gélen ellenőriztük. Az ALF egy számítógépes fragment analízáló és szekvenáló rendszer, mely a Cy5 festékkel jelölt DNS fragmentumok lézerefény hatására keletkező fluoreszcenciáját érzékeli. Kis mennyiségű (fentomolnyi) DNS jelenlétében is nagy pontossággal működik. Mivel a kísérletünkben az amplifikációhoz használt reverz primer az 5'-végén Cy5 festékkel volt jelölve, a keletkező 112 bp hosszú PCR terméket, valamint restrikciós hasítás esetén a reverz szál 5'-végi, 85 bp hosszú szakaszát tudtuk az ALF készülékkel detektálni (12. ábra).

A jelenlévő fragmentumok hosszát az ALFwin Fragment Analyser 1.00 szoftver alkalmazásával határoztuk meg. A pontos méretmeghatározás érdekében minden futtatás során megfelelő molekulaméret standardokat használtunk. Egy 100 bp és egy 150 bp (ALFexpress Sizer 100 bp és 150 bp) hosszú belső standardot futtattunk a mintákkal együtt, illetve egy külső standardot a mintáktól külön, amely 50-500 bp között tartalmaz jelölt fragmentumokat, melyek hossza 50 bp-onként emelkedik

(ALFexpress Sizer 50-500 bp) (12. ábra). Az egyértelmű tipizálást segítette, hogy minden egyes vizsgálat során a tipizálni kívánt mintákkal párhuzamosan ismert genotípusú kontroll mintákat, a vad típusú T allélt és a C allélt hordozó mintát is használtunk. Az analízis során negatív kontrollként női DNS-t tartalmazó mintát is vizsgáltunk.

Az miPEP módszer alkalmazásával a jobb megtartású minták esetében lehetőség nyílt arra, hogy a PCR, illetve a restrikciós emésztés eredményét egyszerűbben és rövidebb idő alatt ellenőrizzük a fragmentumokat 8%-os poliakrilamid gélen történő elválasztásával és etidium-bromid festéssel (13. ábra).

Az ALF rendszeren történő detektálás esetén 4 µl, a 8 %-os poliakrilamid gélen 10 µl mintát futtattunk meg.

#### **III. 4. A szennyeződés megelőzése és az eredmények hitelesítése**

Annak érdekében, hogy a csontminták vizsgálata során a nem endogén DNS-sel történő szennyeződés veszélyét kizárjuk, szigorú óvintézkedéseket tettünk.

Az összes felhasznált eszközt és a munkaterületeket (a steril fülke és a PCR box) a munka megkezdése előtt, illetve a munka befejezésével háztartási hypoval sterilre tisztítottuk és legalább 60 percen keresztül  $1 \text{ J/cm}^2$  UV-C fénnel sugaraztuk be a modern DNS eliminálása érdekében. A DNS extrakcióhoz használt vegyszereket használat előtt átszűrtük  $0,22 \mu\text{m}$  pórusátmérőjű szűrőn (Millipore), majd (a Proteináz K kivételével) 30 percen keresztül  $1 \text{ J/cm}^2$  UV-C fénnel besugaraztuk. Pipettázáshoz a keresztszennyeződést kizáró, szűrős, steril Universal Fit Filter Tips tippeket (Corning Incorporated) használtuk. A PCR és Eppendorf csöveket autoklávoztuk, majd 30 percen keresztül  $1 \text{ J/cm}^2$  UV-C fénnel besugaraztuk. A minta előkészítését, a csont porítását, az extrakciót és a PCR összeállítását az archaikus DNS-sel történő munkának fenntartott, elkülönített laboratóriumi helyiségekben végeztük védőfelszereléssel (védőruházat, szájmazsk, plexi arcmaszk, hajháló, steril kesztyű) használva.

A minta előkészítése, a csont porítása és a DNS extrakció 254 nm germicid lámpával felszerelt lamináris fülke alatt történt (DEFI). Az amplifikációs reakciót beépített ultraibolya csövekkel ellátott PCR fülkékben állítottuk össze (UVC/T Cleaner box, BioSan Ltd.). A levegőt belső UV csővel felszerelt levegő cirkulátorokkal (UV Air

Flow Cleaner-Recirculator UVR-M, BioSan Ltd) mentesítettük a szennyeződéstől.

Az esetleges szennyeződés kiszűrése érdekében negatív kontroll mintákat használtunk. Egy csontport nem tartalmazó extrakciós kontrollt, a DNS extrakció tisztaságának ellenőrzésére, valamint egy templát nélküli amplifikációs kontrollt, mellyel a PCR reagensek és a PCR elegy összeállításának tisztaságát monitoroztuk. Mind a Régészeti Intézet (régész, antropológus), mind pedig csoportunk férfi tagjainál, akik kapcsolatba kerül(het)tek a csontmintákkal, meghatároztuk a Tat polimorfizmus allélikus állapotát.

Az eredmények hitelesítése érdekében minden egyes csontmintát legalább két független kísérletben csoportunk két kutatónöje vizsgált meg. Ennek során a munka minden fázisát (minta előkészítése, csont porítása, DNS extrakció, amplifikáció, a PCR-t követő analízis) legalább kétszer megismételtük.

Minden egyes csontminta esetében az MTA Régészeti Intézetének archeogenetikai laboratóriumában is sor került független mintavételre és porításra. A porított minta egy részét a Régészeti Intézet a rendelkezésünkre bocsátotta. A vizsgálat további lépéseit (DNS extrakció, amplifikáció, PCR-t követő analízis) csoportunk két kutatónöje, egymástól függetlenül végezte a laboratóriumunkban.

A mintavételek a femur diafizisének különböző területeiről történtek.

A Tat marker allélikus állapotára kapott eredményt csak akkor fogadtuk el, ha az legalább két kísérletsorozat során meg tudtuk ismételni és az egyidejűleg használt negatív kontrollok (extrakciós kontroll, amplifikációs kontroll és a női kontroll) esetén nem kaptunk PCR terméket.

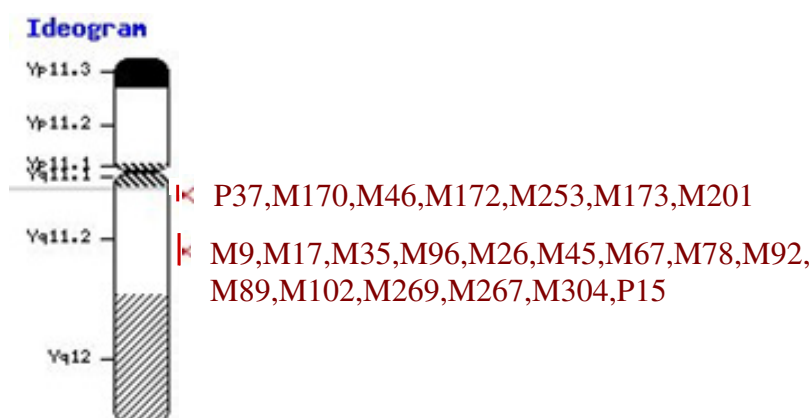
### **III. 5. A modern DNS minták Y-kromoszómális analízise**

A ma élő magyar és székely populációban, összesen 197 egyénnél, 22 (M96, M89, M9, M45, M35, M78, M170, M253, P37, M26, M201, P15, M304, M267, M172, M102, M67, M92, Tat, M173, M17, M269) Y-kromoszómális biallélikus markert vizsgáltunk meg az Y-kromoszómális filogenetikai fán elfoglalt hierarchikus szerveződésük szerint (Y Chromosome Consortium 2002, Jobling és Tyler-Smith, 2003). A vizsgált markerek lokalizációját az Y kromoszóma mentén a 3. táblázat és a 9. ábra mutatja.



A MARKER NEVE	A MARKER HELYE AZ Y KROMOSZÓMÁN	A MARKER AZONOSÍTÓJA (REFSNP ID)
P37	13 001 692 bp	-
M170	13 357 186 bp	<a href="#">rs2032597</a>
M46 (Tat)	13 431 977 bp	<a href="#">rs34442126</a>
M172	13 479 028 bp	<a href="#">rs2032604</a>
M253	13 532 101 bp	<a href="#">rs9341296</a>
M173	13 535 818 bp	<a href="#">rs2032624</a>
M201	13 536 923 bp	<a href="#">rs2032636</a>
M9	20 189 645 bp	<a href="#">rs3900</a>
M17	20 192 556 bp	<a href="#">rs3908</a>
M35	20 201 091 bp	-
M96	20 238 386 bp	<a href="#">rs9306841</a>
M26	20 325 209 bp	<a href="#">rs2032629</a>
M45	20 327 175 bp	<a href="#">rs2032631</a>
M67	20 338 197 bp	<a href="#">rs2032628</a>
M78	20 352 691 bp	-
M92	20 363 411 bp	<a href="#">rs2032648</a>
M89	20 376 701 bp	<a href="#">rs2032652</a>
M102	20 385 500 bp	<a href="#">rs2032608</a>
M269	21 148 755 bp	<a href="#">rs9786153</a>
M267	21 151 206 bp	<a href="#">rs9341313</a>
M304	21 159 241 bp	<a href="#">rs13447352</a>
P15	21 653 414 bp	-

3. táblázat A vizsgált markerek lokalizációja az NCBI referencia Y kromoszóma mentén, valamint a markerek azonosítója az NCBI adatbázisban



9. ábra A vizsgált markerek lokalizációjának szimbólikus jelölése az Y kromoszóma mentén az NCBI adatbázis alapján

Minden minta esetében először az M96, az M89, az M9 és az M45 SNP markereket analizáltuk. A markereket multiplex reakcióban vizsgáltuk, tipizálásukhoz a SNaPshot technikát használtuk. A 4 SNP marker vizsgálatával minden egyes mintát el tudtunk helyezni a filogenetikai fa valamelyik nagyobb ágán. Ily módon, a markerek hierarchikus szerveződésének megfelelően, tudtuk azt, hogy az egyes mintáknál milyen további markereket kell megvizsgáljunk az egyértelmű haplocsoportba sorolás érdekében. Összesen további, a filogenetikai fa megfelelő részén elhelyezkedő, 18 marker (M35, M78, M170, M253, P37, M26, M201, P15, M304, M267, M172, M102, M67, M92, Tat, M173, M17, M269) allélikus állapotát határoztuk meg PCR-RFLP vizsgálattal (az RFLP rövidítés jelentése **R**estriction **F**ragment **L**ength **P**olymorphism – magyar elnevezése: restrikciós fragmenthossz polimorfizmus) (5. táblázat). Azokban az esetekben, amikor a DNS szakaszon sem a vad, sem a derivált allél jelenlétében nem találtunk a vizsgálni kívánt markert magában foglaló restrikciós enzim felismerő helyet, a dCAPS (**d**erived **C**leaved **A**mplified **P**olymorphic **S**equences) módszert (Neff et al. 2002) használtuk. A dCAPS Finder 2.0 program alkalmazásával a PCR-ben olyan forward vagy reverz primert használtunk, melynek szekvenciája egy nukleotidnyi eltérést hordoz a templát DNS szakaszhoz képest. Ezáltal a PCR terméken létrehoztunk egy, a kérdéses markert hordozó restrikciós enzim felismerő helyet (5. táblázat).

Annak érdekében, hogy a jövőben a fent említett polimorfizmusokat az archaikus minták esetében is analizálni tudjuk, az amplifikációhoz olyan primereket terveztünk vagy választottunk az irodalomból, melyek rövid (170 bp alatti) PCR terméket adnak (a vizsgált markerek többségénél mi terveztük meg az amplifikációhoz szükséges primereket) (4. és 5. táblázat).

Minden egyes marker analízisének használtunk olyan kontroll mintákat, melyek közül az egyik a vad típusú allélt, míg a másik a marker derivált allélját hordozza. A haplocsoportok elnevezésénél és ennek megfelelően a minták besorolásánál a Jobling és Tyler-Smith szerzők által 2003-ban közölt filogenetikai fát vettük alapul (3. ábra).

### III.5.1. Multiplex PCR (az M96, M89, M9, M45 SNP markerek amplifikálása)

A multiplex reakció kivitelezése során a Belgrader és munkatársai (1996), illetve a Paracchini és munkacsoportja (2002) által közölt protokoll részben módosított változatát használtuk. Az analízis során két lépésből álló PCR-t végzünk (10. ábra). Az első reakcióban olyan primerekkel amplifikálunk, melyeknek a 3' része (19-22 nukleotid) specifikus arra a DNS szakaszra, amely a kérdéses markert hordozza. A primerek közös 5'-végi szakaszának szekvenciája megegyezik a 22 bp (ZIP1 primer) vagy 21 bp (ZIP2 primer) hosszú ZIP kód primerek szekvenciájával, ugyanakkor nem komplementer a templát DNS szekvenciával, így ahhoz nem hibridizál. Ezen hibrid primerek alkalmazásával az első multiplex PCR után a különböző SNP markert hordozó amplifikátumok 5'-végi szekvenciája azonos lesz. Ennek következtében a második reakcióban a ZIP kód primerekkel nagyjából ugyanolyan mennyiségben tudjuk felszorzozni a különböző polimorfizmusokat tartalmazó DNS fragmentumokat.

Az archaikus DNS töredezettségét figyelembe véve az első reakció primereit úgy terveztük meg, hogy a keletkező PCR termékek ne legyenek hosszúak. Ily módon egy 106 bp (M45 marker esetén), egy 121 bp (M9), egy 137 bp (M89) és egy 157 bp (M96) hosszúságú specifikus terméket kapunk. Ezek a PCR termékek ugyanakkor a primerek közös 5'-végi ZIP szekvenciái miatt rendre 43 nukleotiddal hosszabbak és így a DNS fragmentumok tényleges hossza: 149 bp, 164 bp, 180 bp, 200 bp (4. táblázat).

Az első PCR körülményei: A reakcióelegy 12,5 µl végtérfogatban tartalmaz 2,5 µl vérből vagy hajból izolált DNS templátot, 1,5 U AmpliTaq Gold DNS Polimeráz (Applied Biosystems), 1x AmpliTaq Gold puffert (Applied Biosystems), 4 mM MgCl<sub>2</sub>-ot, 400 µM-t az egyes dNTP-ekből és 0,32 µM-t az egyes hibrid primerekből (4. táblázat).

A PCR lépései:

1. 94 °C 9 perc	} 15 ciklus
2. 94 °C 30 másodperc	
3. 63 °C 30 másodperc	
4. 72 °C 1 perc	
5. 72 °C 3 perc	

A második PCR körülményei: A reakcióelegy 12,5 µl végtérfogatban tartalmaz 1 µM-t a ZIP kód primerekből (4. táblázat), 4 mM MgCl<sub>2</sub>-ot, 1,5 U AmpliTaq Gold

DNS Polimerázt (Applied Biosystems) és 1x AmpliTaq Gold puffert (Applied Biosystems). Ezt az elegyet adjuk az előző reakció 12,5 µl-hez.

A PCR lépései:

1. 94 °C 9 perc	} 30 ciklus
2. 94 °C 30 másodperc	
3. 59 °C 30 másodperc	
4. 72 °C 1 perc	
5. 72 °C 3 perc	

A PCR termék 4 µl-ét futtattuk meg 8%-os natív poliakrilamid gélen, majd etidium-bromid festés után ellenőriztük a reakció sikerességét UV transzilluminátor segítségével. A PCR termékek méreteit GelBase gél dokumentációs rendszer felhasználásával határoztuk meg (UVP).

A reakciók során olyan negatív kontrollokat alkalmaztunk, ahol templátként női DNS-t, illetve vizet tettünk a csövekbe. Pozitív kontrollnak olyan mintát használtunk, melynél előzetesen direkt szekvenálással mind a 4 marker esetében meghatároztuk az allélikus állapotot.

### **III.5.2. Az M96, M89, M9, M45 SNP markerek genotípusának meghatározása**

A multiplex PCR-t követően előkészítettük a PCR termékeket a SNaPshot reakcióhoz. Ennek során a 4 PCR terméket magában foglaló elegy 3 µl-hez Exonukleáz I és Shrimp Alkaline Foszfátáz enzimet tartalmazó, 1,2 µl ExoSAP-IT (USB) reagenst adtunk, melynek segítségével 37 °C-on 15 perces reakcióban eltávolítottuk az elegyből a fel nem használt primereket és a be nem épült nukleotidokat. Ezt követően a reagenst inaktiváltuk 80 °C-on 15 perces reakcióban. A megtisztított multiplex PCR elegy 3 µl-ét használtuk a SNaPshot reakcióban, melyet a gyártó utasításai szerint végeztünk (Applied Biosystems Protocol P/N 4323357) ABI PRISM SNaPshot ddNTP Primer Extension Kit (Applied Biosystems) alkalmazásával. A Kit különböző fluoreszcens festékekkel jelölt dideoxi nukleotidokat tartalmaz. A reakcióhoz a 4 különböző SNP marker egyértelmű tipizálása érdekében 4 eltérő hosszúságú miniszekvenáló primert terveztünk (4. táblázat) a gyártó javaslatait követve. A primerek úgy hibridizálnak az egyes DNS templátokhoz, hogy a 3'-végi szakaszuk közvetlenül a vizsgálni kívánt SNP marker előtt, upstream irányban végződik. A SNaPshot reakció során a jelöletlen

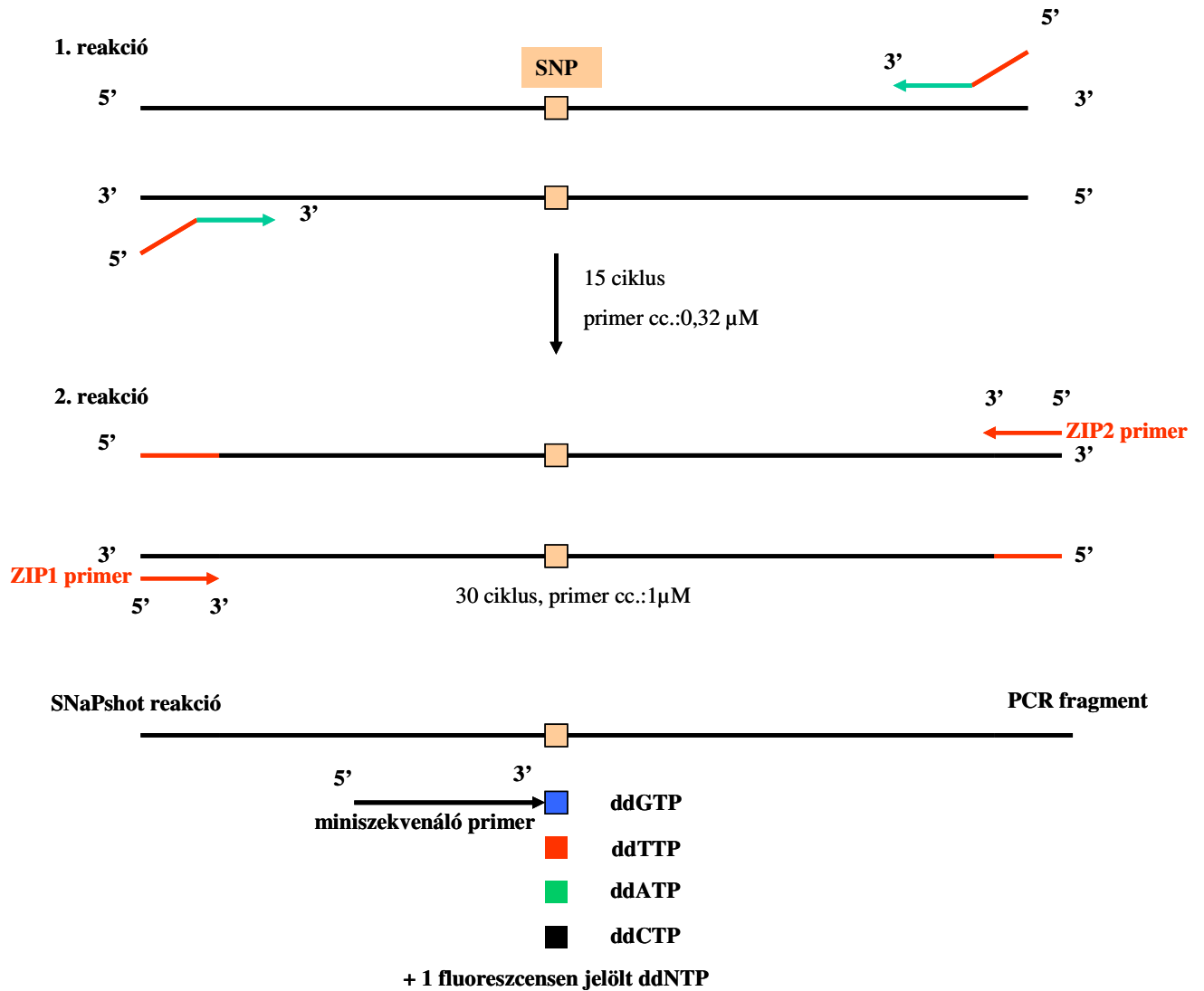
oligonukleotid primerek kötődnek a velük komplementer templát DNS szakaszhoz fluoreszcensen jelölt dideoxi dNTP-k és AmpliTaq DNS Polimeráz jelenlétében. Mivel csak dideoxi nukleotidok vannak a reakcióelegyben a primerek 3'-végi szakaszához egyetlen, a kérdéses SNP markerrel komplementer nukleotid tud kapcsolódni, majd az extenzió terminálódik és a keletkező fragment (miniszekvenáló primer + 1 ddNTP= SNaPshot termék) fluoreszcens jelölést kap. A reakció után Shrimp Alkaline Foszfátáz kezeléssel eltávolítottuk a be nem épült dideoxi nukleotidokat, majd a fluoreszcensen jelölt dideoxi terminált fragmentumokat (0,25 µl-t) kapilláris elektroforézissel ABI PRISM 310 Genetic Analyzer (Applied Biosystems) rendszeren elválasztottuk. A genotípus meghatározásához a Genotyper szoftver 3.7-es verzióját (Applied Biosystems) használtuk.

Mivel a 4 SNP marker vizsgálatához 4 különböző hosszúságú primert terveztünk a minták elektroforetikus képén a SNaPshot termékek hossza alapján tudjuk azonosítani, hogy melyik termék melyik markerhez tartozik, míg a nukleotidok eltérő fluoreszcens jelölése alapján, a görbe színéből, az adott marker genotípusát tudjuk leolvasni (17. ábra). Minden futtatásnál megfelelő molekulaméret standardot (GeneScan-120 LIZ Size Standard, Applied Biosystems) használtunk.

A miniszekvenáló primerek koncentrációja a reakcióban rendre a következő volt: 0,1 µM az M96, 0,6 µM az M89, 0,1 µM az M9 és 0,3 µM az M45 marker esetében. A primerek hossza között legalább 5 nukleotid különbség van (4. táblázat) annak érdekében, hogy a SNaPshot termékeket a futtatás során egyértelműen el tudjuk választani. A fluoreszcens festékek ugyanis a DNS fragmentumok mobilitásbeli eltolódását okozzák az elektroforézis során és így a SNaPshot termékek tényleges hossza és a futásnál kapott hosszuk néhány nukleotiddal eltér. A termékek elválasztásánál tapasztaltunk ilyen eltéréseket, azonban ezek egyáltalán nem zavarták a minták genotipizálását. A primerek hosszát a másodlagos struktúrák elkerülése végett 5'-végi nem homológ polinukleotidok (poliT)(M89, M45 marker), illetve az M96 markerhez tartozó primer esetében 'taat' szakasz hozzáadásával alakítottuk ki.

A fent leírt standard eljárást (Multiplex PCR+Genotípus meghatározása) sikerrel alkalmaztuk a modern minták vizsgálatokor, illetve részben, degradált DNS esetében is. A T2/41-es csontminta analízisekor a 4 különböző SNP marker közül az M45 és az M9 markert hordozó, 2 rövidebb (a specifikus fragmentumok hossza: 106 bp és 121 bp) DNS fragmentumot tudtuk a multiplex reakcióban amplifikálni. A csontminta vizsgálata

abban különbözött a leírt stratégiától, hogy a multiplex reakcióegyhez templátként miPEP amplifikációs terméket (4 µl) adtunk, majd a multiplex PCR termék 10 µl-ét futtattuk meg poliakrilamid gélen.



10. ábra Az M96, M89, M9, M45 SNP markerek tipizálásához használt 2 lépéses PCR és a SNaPshot reakció sémája

4. táblázat Az M96, M89, M9 és az M45 SNP markerek PCR/SNaPshot tipizálásához használt primerek

		Primerek		
		PCR	SNaPshot reakció	
Marker/Mutáció	Szekvencia*	a PCR termék hossza <sup>a</sup>	Szekvencia	Hosszúság (mer)
M96 G→C	5'- ggagcagcgtatcccgtagac <b>GTTGCCCTCTCACAGAGCAC</b> -3' 5'-cgctgccaaactaccgcacatg <b>GAAGAGATTCACCCACCCAC</b> -3'	200 bp (157 bp)	5'-taatGAAAACAGGTCTCTCATAATA- 3'	25
M89 C→T	5'- ggagcagcgtatcccgtagac <b>TCCTATGAGGTGCCATGAAA</b> -3' 5'- cgctgccaaactaccgcacatg <b>GGATCACCAGCAAAGGTAGC</b> -3'	180 bp (137 bp)	5'- ttttttttttttttCTCAGGCAAAGTGAGAGAT-3'	35
M9 C→G	5'- ggagcagcgtatcccgtagac <b>TGCAAAGAAACGGCCTAAG</b> -3' 5'- cgctgccaaactaccgcacatg <b>GCATAATGAAGTAAGCGCTACC</b> -3'	164 bp (121 bp)	5'-ACGGCCTAAGATGGTTGAAT -3'	20
M45 G→A	5'- ggagcagcgtatcccgtagac <b>GTGGACTTTACGAACCAACCT</b> -3' 5'- cgctgccaaactaccgcacatg <b>CCTGGACCTCAGAAGGAGCT</b> -3'	149 bp (106 bp)	5'- ttttttttCCTCAGAAGGAGCTTTTTTGC -3'	30

\*A genomiális szekvenciát félkövér betűvel, a zip szekvenciát kisbetűvel jelöltük.

<sup>a</sup> A zárójelben feltüntetett számok a specifikus PCR termék hosszát mutatják.

A zip kód primerek szekvenciáját Belgrader és munkatársai (1996) közzétették.

Az M9 és az M45 markerek esetében a PCR-hez használt forward és reverz primerek, az M96 marker esetében a reverz primer lókuszt specifikus 3'-részét ez a tanulmány közli.

Az M89 marker vizsgálatához használt PCR primereket Paracchini és munkatársai (2002), az M96 marker amplifikálásához használt forward primer szekvenciáját Underhill és munkatársai közzétették (2000).

### III. 5. 3. PCR-RFLP analízis

Az M35, M78, M170, M253, P37, M26, M201, P15, M304, M267, M172, M102, M67, M92, Tat, M173, M17, M269 biallélikus markerek allélikus állapotát PCR-RFLP vizsgálattal határoztuk meg.

Az amplifikációs reakcióelegy 25 µl végtérfogatban tartalmazott 1 U AmpliTaq Gold DNS Polimerázt (Applied Biosystems), 1x AmpliTaq Gold DNS puffert (Applied Biosystems), 3 mM MgCl<sub>2</sub>-ot, 200 µM-t az egyes dNTP-ekből, 1-1 µM primert (5. táblázat) és 5 µl DNS templátot.

A PCR lépései:

- |  |             |
|--|-------------|
| 1. 94 °C 6 perc  | } 33 ciklus |
| 2. 93 °C 20 másodperc  |             |
| 3. primer specifikus annealing hőmérséklet (50-67 °C) 20 másodperc |             |
| 4. 72 °C 30 másodperc  |             |
| 5. 72 °C 5 perc  |             |

A fent leírt standard amplifikációs körülményeket néhány marker esetében módosítani kellett annak érdekében, hogy a PCR-t specifikusabbá tegyék. Az M170 és az M267 markerek esetében a 33 ciklust megelőzően 10 magasabb annealing hőmérsékleten végbemenő ciklusra volt szükség. A 10 ciklus a következő lépésekből állt: denaturáció: 93 °C 20 másodperc, annealing: M170 marker-64 °C, M267 marker-54 °C 20 másodperc, elongáció: 72 °C 30 másodperc. A P37 polimorfizmus amplifikálásához 0,5 µM-t használtunk az egyes primerekből és 3 µl DNS templátot.

A PCR termék 3 µl-ét használtuk a megfelelő restrikciós enzimmel (5. táblázat) végzett hasítási reakcióban. Az egyes minták genotípusának megállapításához a restrikciós fragmentumokat 8%-os poliakrilamid gélen elválasztottuk, majd etidium-bromid festést követően a fragmentumok méretét GelBase gél dokumentációs rendszer (UVP) felhasználásával meghatároztuk.



5. táblázat M35, M78, M170, M253, P37, M26, M201, P15, M304, M267, M172, M102, M67, M92, Tat, M173, M17 és az M269 biállélikus markerek tipizálásához használt PCR/RFLP primerek és paraméterek

Markerek			Primerek			RFLP		
Név	Mut.	Referencia	Szekvencia <sup>a</sup>	Referencia	Annealing hőmérséklet	Enzim	vad (bp)	derivált (bp)
M35	G>C	Underhill et al. 2000	5'-gagagggcatggtcccttc 5'-ttcggagtctctgcctgtg	ez a tanulmány "	57 °C	Bts I	88	60+28
M78	C>T	Underhill et al. 2000	5'-tcgacatgaacacaaattgatacactt 5'-aaagcaagtactatgaccagc	Onofri et al. 2006 ez a tanulmány	58 °C	Aci I	45+46	91
M170	A>C	Underhill et al. 2000	5'-catattctgtgcattatacaaat 5'-acacaacccacactgaaaaa <u>ca</u>	ez a tanulmány "	60 °C	Nla III	57+20	77
M253	C>T	Cinnioğlu et al. 2004	5'-tccatttagctgatctgttcc 5'-caaggactcattcaatgaaga	ez a tanulmány "	58°C	Hinc II	117+34	151
P37	T>C	Karafet et al. 2001	5'-gcatagtgataggggtggattgg 5'-cctggcaaaggcgggaatcc	ez a tanulmány "	67°C	HpyCH4 III	95	29+66
M26	G>A	Underhill et al. 2000	5'-gacagctggcttaccagtgg 5'-gaatttcattagccattcagtgttctc <u>gg</u>	ez a tanulmány "	59°C	Aci I	81+29	110
M201	G>T	Underhill et al. 2001	5'-tatgcattgttgagtatatgtcaaat 5'-tccaactaagtacattacga <u>aga</u>	Onofri et al. 2006 ez a tanulmány	55°C	Bbs I	129+34	163
P15	C>T	Hammer et al. 2000	5'-cctcacatgaatagagccaa 5'-acttcatctgccttc <u>agc</u>	ez a tanulmány "	52°C	Mlu I	41+21	62
M304	A>C	Cinnioğlu et al. 2004	5'-tgtaacaaacagtatgtggga 5'-ataccaaaatcaccagttg	ez a tanulmány "	56°C	NmuC I	81	54+27
M267	T>G	Cinnioğlu et al. 2004	5'-tataccaagtctggatagcggga 5'-ctccacacaaaataactgaac <u>ct</u>	ez a tanulmány "	50°C	Mnl I	71	41+30
M172	T>G	Underhill et al. 2000	5'-ctgcctctcagtatcaacag 5'-taataattgaagacctttg <u>ag</u> t	ez a tanulmány "	58°C	Hinf I	91	68+23
M102	G>C	Underhill et al. 2000	5'-agataaaattcacatagtca 5'-tccttaactctaggggt	ez a tanulmány "	54°C	NmuC I	106	24+82

5. táblázat *folytatás*

Markerek			Primerek			RFLP		
Név	Mut.	Referencia	Szekvencia <sup>a</sup>	Referencia	Annealing hőmérséklet	Enzim	vad (bp)	derivált (bp)
M67	A>T	Underhill et al. 2000	5'-aaagctctatcttcaagtagctgtccta 5'-cacttgctcgtggaccctctacat	Onofri et al. 2006 ez a tanulmány	59°C	Nde I	130	104+26
M92	T>C	Underhill et al. 2000	5'-ttggacttaaggtggcttg 5'-ttcagaaactggtttgtgtcc	Underhill et al. 2000 ez a tanulmány	54°C	HpyCH4 IV	161	30+131
Tat	T>C	Zerjal et al. 1997	5'-gactctgagttagacttgtga 5'-gaaggtgccgtaaaagtgtgaa	Zerjal et al. 1997 "	56°C	HpyCH4 IV	112	27+85
M173	A>C	Underhill et al. 2000	5'-ttctacaattcaagggcatttggaaac 5'-gctgcagtttccagatcct	Onofri et al. 2006 ez a tanulmány	67°C	Nla IV	98	25+73
M17	G>-	Underhill et al. 1997	5'-agagttgtggttgctggtgttacggg 5'-tgatgtagagacatctgaaaccac	ez a tanulmány "	56°C	Mlu I	117	24+92
M269	T>C	Cruciani et al. 2002	5'-ggggaatgatcagggtttgg 5'-actctttgtgtgccttctgag	ez a tanulmány "	57°C	ScrF I	88	27+61

<sup>a</sup>A target DNS szekvenciával nem komplementer nukleotidot aláhúztuk és félkövér betűvel jelöltük.

### III. 5. 4. Statisztikai analízis

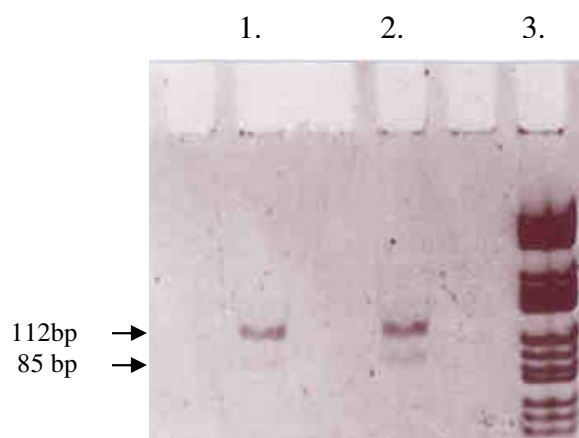
Az E3b-M35 (E3b\*, E3b1), G-M201, I-M170 (I\*, I1a, I1b\*), J1-M267, J2-M172 (J2\*, J2e1, J2f\*, J2f1), N3-Tat, R1a1-M17 és az R1b3-M269 haplocsoportok gyakorisági értékei alapján (18. ábra) az Arlequin szoftver 2.000 (Schneider et al. 2000) verziójának alkalmazásával meghatároztuk az általunk vizsgált magyar és székely populáció genetikai diverzitásának mértékét ( $H$ ) (Nei, 1987), valamint páronkénti genetikai távolságot ( $F_{st}$ ) (Reynolds et al. 1983) számoltunk a magyar és a székely populáció, illetve a Semino és munkatársai által analizált (2000a) további 20 európai populáció között (vagyis minden egyes populáció gyakorisági értékeit összehasonlítottuk minden más populáció gyakorisági értékeivel). A populációk közötti páronkénti genetikai távolság értékeket mátrix formában kaptuk meg, melynek adataiból többdimenziós skálázási (Multidimensional Scaling - MDS) eljárással 2 dimenziós térben ábrázoltuk az egyes populációk közötti távolságviszonyokat. Az MDS analízist az SPSS szoftver 5.0 verziójának (SPSS, Inc.) felhasználásával hajtottuk végre.

## IV. Eredmények

### IV. 1. A Tat polimorfizmus vizsgálata

A munkánk során 100 magyar, 97 székely és 7 csontminta esetében határoztuk meg a Tat marker allélikus állapotát. A tipizálás során egy 112 bp hosszú DNS szakaszt amplifikáltunk, mely a 28. pozícióban hordozza a Tat markert. A DNS fragmentum HpyCH4 IV restrikciós enzimmel történő kezelését (5. táblázat) követően határoztuk meg a minta genotípusát. A HpyCH4 IV enzim a C allél jelenléte esetén hasít, egy 27 bp és egy 85 bp hosszú emésztési terméket ad.

Az általunk analizált modern magyar és székely férfiak közül csupán egyetlen székely férfi esetében mutattuk ki a Tat C allélt (1/197). A mutáns C allélt hordozó modern minta segítségével sikerült beállítanunk úgy az emésztési reakciót, hogy ugyanolyan mennyiségű templátok esetében egyszerre több (8) modern mintát tudunk együtt emészteni úgy, hogy még biztonsággal detektálni tudtuk, ha a vad típusú minták között akár csak egyetlen, mutáns allélt hordozó minta is jelen volt (11. ábra).



11. ábra Hasítási kép HpyCH4 IV restrikciós enzimmel (C allél jelenléte esetén hasít) végzett reakciót követően- a nyolcas mintacsoport esetében is látszik a 85 bp hosszú hasítási termék

1. nyolcas mintacsoport; 2. hetes mintacsoport; 3. pBR322 DNS marker

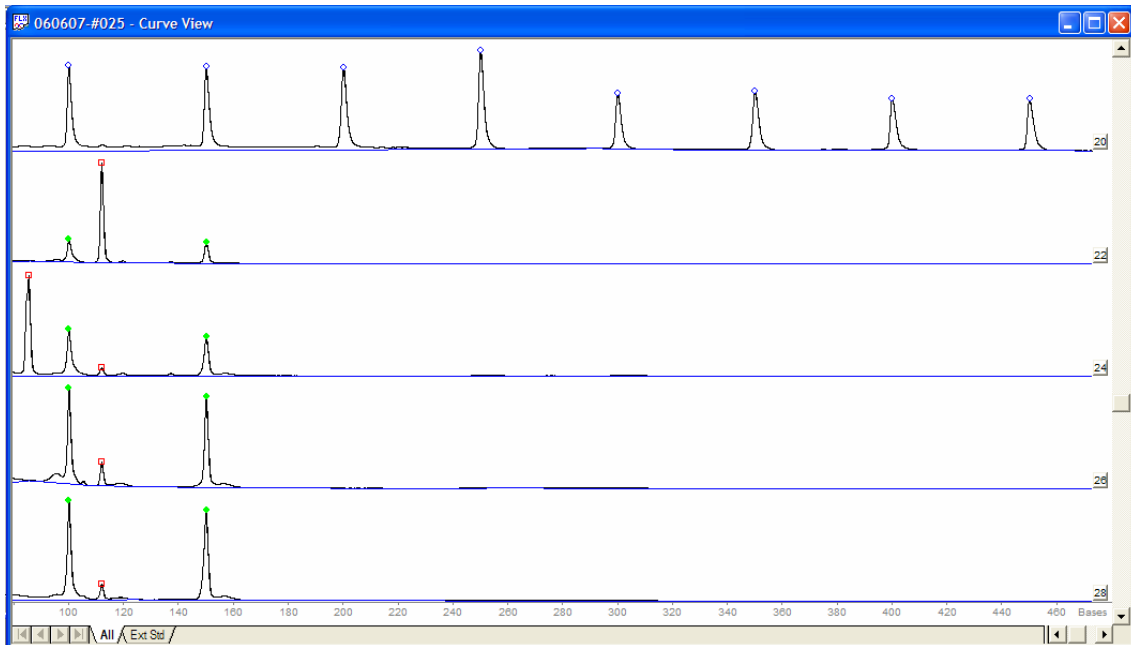
A 7 sikeresen tipizált (IX-)X. századi csontminta közül 2 esetben is igazoltuk a mutáns C allél jelenlétét (2/7), míg 5 leletnél a vad típusú T allél előfordulását (6. táblázat).

minta	Tat T/C allél	lelőhely
B1/3c	C	Örménykút (Békés megye)
13/1	C	Szabadkígyós-Pálliget (Békés)
13/7	T	Szabadkígyós-Pálliget (Békés)
378/B	T	Szarvas-Kákapusztá (Békés)
T2/41	T	Mözs-Szárazdomb (Tolna)
14/B	T	Harta (Bács-Kiskun)
48/B	T	Csekej-Szlovákia

6. táblázat A vizsgált minták allélikus állapota és lelőhelye

A specifikus Y-kromoszómális analízis előtt minden esetben szükség volt a megmaradt teljes töredékes genom random primerekkel történő dúsítására. A rendelkezésünkre álló archaikus DNS mennyiségének felszaporításához az ún. PEP (**P**rimers **E**xtension **P**reamplification) módszernek egy tovább fejlesztett változatát (**m**odified **i**mproved PEP - miPEP) alkalmaztuk. Az utóbbi módszer 2 polimeráz enzim elegyét használja, melyek közül az egyik 3'-5' exonukleáz (más néven proofreading) aktivitással rendelkezik. Ily módon lehetővé teszi, hogy nagyobb hűséggel történjen a templát DNS átírása.

A Tat polimorfizmust hordozó 112 bp-os DNS szakasz amplifikációjának sikerességét és a restrikciós emésztés eredményét részben az ALF rendszeren (12. ábra), részben, a jobb megtartású minták esetében, a fragmentumok 8%-os poliakrilamid gélen történő elválasztásával és etidium-bromid festéssel ellenőriztük (13. ábra).



12. ábra A Tat polimorfizmus archaikus DNS-en történő PCR-RFLP vizsgálata ALF rendszer segítségével

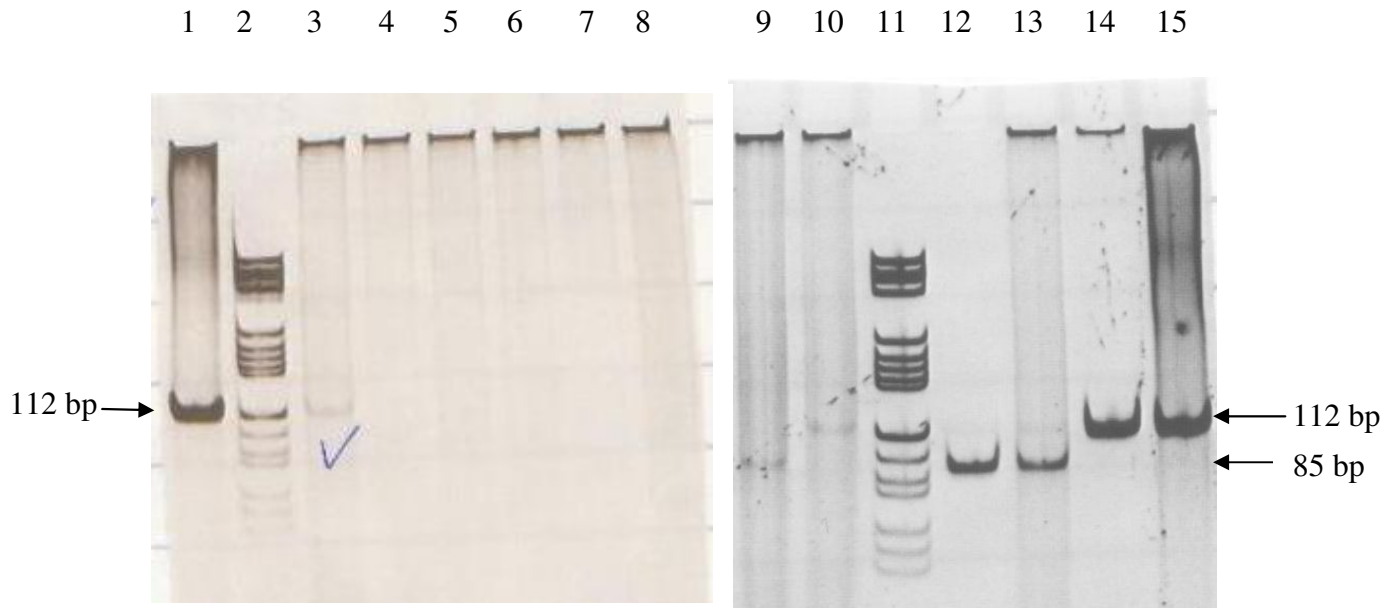
A csúcsok a Cy5 festékkel jelölt DNS fragmentumok fluoreszcencia intenzitását reprezentálják (a PCR reverz primerét jelöltük Cy5 festékkel). A PCR és a restriktív emésztés során kapott termékek mérete az X tengely mentén olvasható le.

20. sáv: ALFexpress Sizer 50-500 bp (100 bp-tól 450 bp-ig)

22, 26. sáv: a 112 bp hosszú PCR termékek a 13/1 (22. sáv) és a T2/41 (26. sáv) csontminta esetén és a 100 bp és 150 bp hosszú belső molekulaméret standardok

24, 28. sáv: a HpyCH4 IV enzimmel végzett restriktív hasítás eredménye: a 85 bp hosszú emésztési termék jelenléte alapján (24. sáv) a 13/1 minta hordozza a Tat C allélt, míg a termék hiánya a 28. sávban arra utal, hogy a T2/41 minta Y kromoszómáján a T allél fordul elő.

A 21, 23, 25, 27 és a 29. sávokba nem futtattunk mintát annak érdekében, hogy elkerüljük a minták szomszédos sávokból történő szennyeződését.



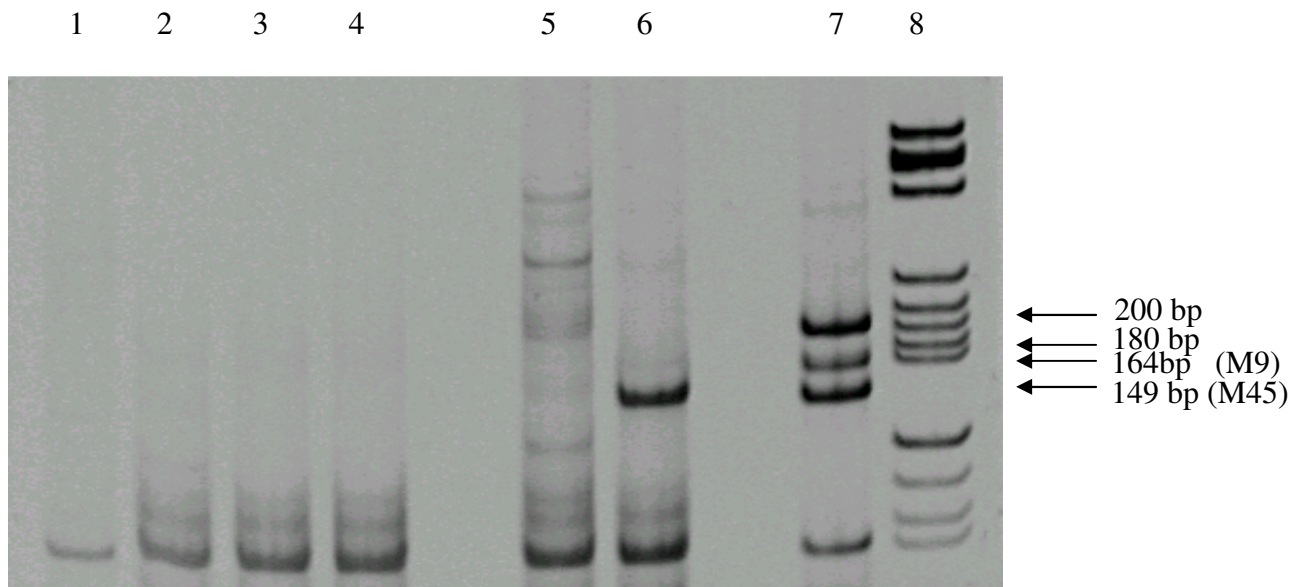
13. ábra A Tat polimorfizmus archaikus DNS-en történő PCR-RFLP vizsgálata 8%-os poliakrilamid gélen és etidium-bromid festéssel

- 1. férfi pozitív kontroll
- 2. pBR322 DNS marker
- 3. B1/3c csontminta
- 4-8. negatív kontrollok

- 9. B1/3c (Tat C allél)
- 10. 14/B (Tat T allél)
- 11. pBR322 DNS marker
- 12-13. férfi pozitív kontroll (C)
- 14-15. férfi pozitív kontroll (T)

A T2/41 csontminta esetében a PCR-RFLP vizsgálat eredményét az M9 és az M45 SNP marker multiplex vizsgálatával is sikerült megerősítenünk. Az analízis során a modern mintákon sikerrel beállított multiplex PCR alkalmazásával az M96, M89, M9 és az M45 markerek közül az M45 és az M9 markert hordozó, 2 rövidebb (a specifikus fragmentumok hossza: 106 bp és 121 bp) DNS fragmentumot tudtuk amplifikálni, melyek közül a legrövidebb, az M45 markert hordozó PRC termék (és ennek megfelelően a SNaPshot termék is) erősebb jelet adott (14., 15. ábra). (A PCR termékek a primerek közös 5'-végi ZIP szekvenciái miatt rendre 43 nukleotiddal hosszabbak, így a DNS fragmentumok tényleges hossza: 149 bp (M45), 164 bp (M9)).

Vizsgálataink hitelességét támasztja alá az amplifikáció hatékonysága és az amplifikált termék hossza közötti inverz kapcsolat, mely az archaikus DNS károsodásának, degradációjának következménye (Pääbo et al. 2004, Handt et al. 1994). Az M89 és az M96 lókuszt nem sikerült amplifikálnunk, valószínűleg a magasabb molekulaméret tartomány miatt.

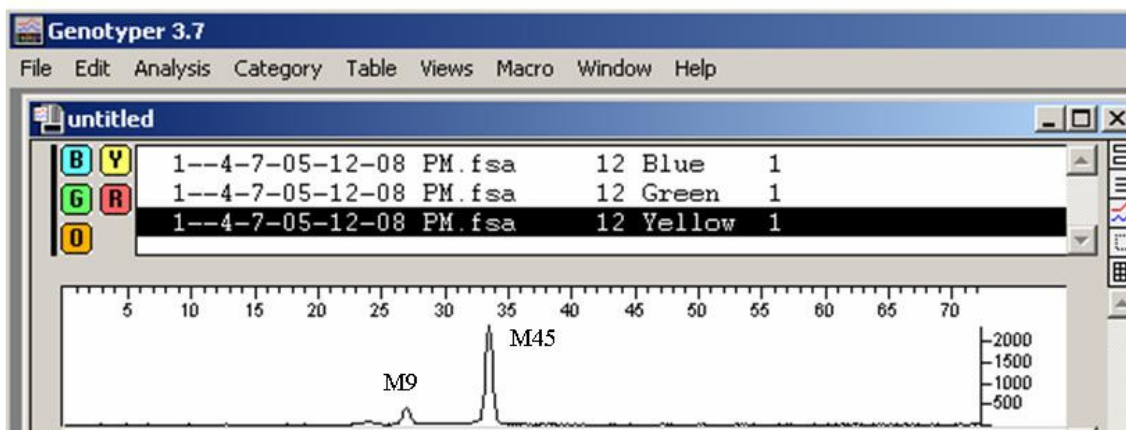


14. ábra Az M9 és az M45 marker multiplex amplifikálása csontmintán  
A poliakrilamid gélelektroforézis a T2/41 csontminta esetében csak a legrövidebb, az M45 markernek megfelelő fragmentum jelenlétét mutatta egyértelműen.

- |                                    |                          |
|------------------------------------|--------------------------|
| 1.női DNS kontroll                 | 5.Csont B1/3c            |
| 2.neg.kontroll (víztemplát)        | 6.Csont T2/41            |
| 3.extrakciós kontroll –Csont B1/3c | 7.férfi pozitív kontroll |
| 4.extrakciós kontroll –Csont T2/41 | 8.pBR322 DNS marker      |

A kapilláris rendszeren történő genotípus meghatározás ugyanakkor 2 jelet adott (15. ábra). E szerint az M9 markert tartalmazó DNS szakaszt is sikerült amplifikálnunk. A tipizálás szerint mindkét marker esetében egy-egy C allél van jelen, mely a markerek vad típusát jelöli. A markerek hierarchikus szerveződésének (YCC 2002, Jobling és Tyler-Smith 2003) megfelelően a csontmintából izolált Y kromoszómán nem lehet jelen a Tat C allél, hiszen akkor az M9 marker derivált G allélját kellett volna kimutatnunk. (A Tat C alléllal rendelkező Y kromoszómák egy részhalmazát képezik azoknak az Y kromoszómáknak, ahol jelen van az M9 marker derivált G allélja, vagyis a Tat C allél egy olyan genetikai háttérrel jelent meg, ahol az M9 marker derivált allélja volt jelen.)



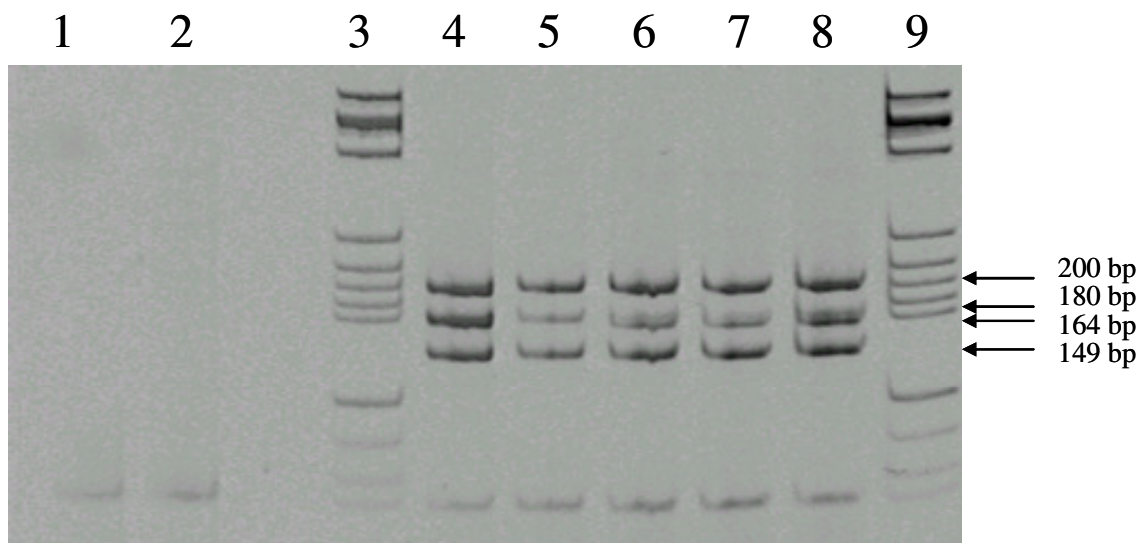


15. ábra Az M9 és az M45 marker multiplex tipizálása a T2/41 csontminta esetében

A két csontminta esetében is detektált Tat C allél egyben eredményeink hitelességét is igazolja, hiszen sem a Régészeti Intézet munkatársai (régész, antropológus), sem csoportunk férfi tagjai, akik kapcsolatba kerültek a csontmintákkal, nem hordozzák a Tat C mutációt.

## IV. 2. A modern DNS minták Y-kromoszómális analízise

A ma élő magyar és székely populációban, összesen 197 egyénnél, 22 Y-kromoszómális biallélikus markert vizsgáltunk meg (M96, M89, M9, M45, M35, M78, M170, M253, P37, M26, M201, P15, M304, M267, M172, M102, M67, M92, Tat, M173, M17, M269). A vizsgálatba bevont magyar férfiak (100 fő) Magyarország különböző területein születtek, de a leginkább reprezentált régió az Alföld területe volt (90/100). A székely minták (97 fő) Korondról, Erdély területéről származnak (8. ábra). Minden minta esetében először az M96, az M89, az M9 és az M45 SNP markereket analizáltuk, melyek gyors és hatékony vizsgálatához sikerült beállítanunk egy két lépésből álló multiplex reakciót. Az archaikus DNS töredezettségét figyelembe véve az első reakció primereit úgy terveztük meg, hogy a keletkező PCR termékek ne legyenek hosszúak. Ily módon egy 106 bp (M45-ös marker esetén), egy 121 bp (M9), egy 137 bp (M89) és egy 157 bp (M96) hosszúságú specifikus terméket kaptunk. Ezek a PCR termékek ugyanakkor a primerek közös 5'-végi ZIP szekvenciái miatt rendre 43 nukleotiddal hosszabbak és így a DNS fragmentumok tényleges hossza: 149 bp, 164 bp, 180 bp, 200 bp (16. ábra). A poliakrilamid gélen a 164 bp és a 180 bp hosszú termék közel fut egymáshoz, de egymástól jól elkülöníthetők.



16. ábra Az M96, M89, M9 és az M45 SNP marker multiplex amplifikálása

1. negatív kontroll (víztemplát)
2. negatív kontroll (női DNS)
- 3, 9. pBR322 DNS marker
- 4-8. ma élő magyar és székely férfiak mintái

A genotípus meghatározásához a SNaPshot technikát használtuk, amely a kapilláris gélelektroforézis rendszer alkalmazásával együtt lehetővé tette, hogy a 4 különböző marker esetében egyidejűleg, ugyanakkor egyértelműen azonosítsuk az egyes markerek alléljeit (17. ábra).



negatív  
kontroll

■ **ddGTP**
■ **ddATP**  
■ **ddTTP**
■ **ddCTP**

17. ábra 2 magyar minta genotípusának meghatározása ABI Prism 310 Genetic Analyzer rendszeren

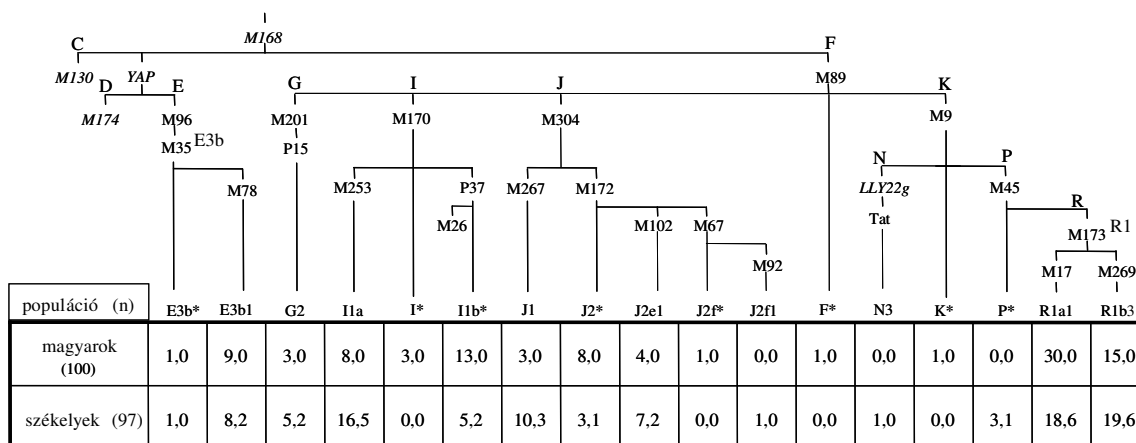
A negatív kontroll futtatásakor csak a molekulaméret standard (GeneScan-120 LIZ Size Standard) narancs színnel látható fragmentumai voltak jelen.

A 4 SNP marker tipizálása alapján minden egyes modern mintát el tudtunk helyezni a filogenetikai fa valamelyik nagyobb ágán, illetve tudtuk azt, a markerek hierarchikus szerveződésének megfelelően, hogy milyen további markereket kell még vizsgálnunk a pontos kategorizálás érdekében. Összesen további 18 marker (M35, M78, M170, M253, P37, M26, M201, P15, M304, M267, M172, M102, M67, M92, Tat, M173, M17, M269) allélikus állapotát határoztuk meg PCR-RFLP vizsgálattal (5. táblázat).

Annak érdekében, hogy a jövőben a fentebb felsorolt Y-kromoszómális markereket az archaikus mintákon is vizsgálni tudjuk, minden marker analíziséhez, melyet a modern mintákon tanulmányoztunk, olyan primerpárokat terveztünk, vagy választottunk az irodalomból, melyek rövid (170 bp alatti) PCR terméket adnak. A

markerek amplifikálásához szükséges PCR körülményeket a modern mintákon optimalizáltuk, így tulajdonképpen minden szükséges lépést megtettünk azért, hogy ősi DNS-en is vizsgáljuk azokat.

A vizsgált markerek segítségével a modern magyar és székely mintákat az E, F\*, G, I, J, K\*, N3, P\* és az R1 Y-kromoszómális haplocsoportokba soroltuk be (18. ábra). A haplocsoportok meghatározásánál és ennek megfelelően a minták kategorizálásánál a Jobling és Tyler-Smith szerzők által 2003-ban közölt filogenetikai fát vettük alapul (3. ábra). A 18. ábra a 3. ábrán látható, 2003-ban közölt filogenetikai fa egyszerűsített változatát mutatja, melyen szinte kizárólag a vizsgált markerek genealógiáját és az általuk meghatározott Y-kromoszómális haplocsoportokat tüntettük fel. 4 olyan marker szerepel az egyszerűsített ábrán, melyek a 2003-ban közölt filogenetikai fán nincsenek jelölve. A C haplocsoportot meghatározó M130 marker a 2003-as fán szereplő RPS4Y<sub>711</sub> markerrel azonos, ugyanazon marker két különböző elnevezéssel szerepel az irodalomban. Az M253, az M304 és az M267 SNP markereket Cinnioğlu és munkacsoportja 2004-ben közölte. Az I1a klasztert jellemző M253 marker filogenetikailag egyenértékű a 2003-ban közölt fán szereplő, ugyanazon haplocsoportot meghatározó P30 markerrel. Az M304 marker a 2003-as fán jelölt 12f2 markerrel együtt a J haplocsoportot határozza meg. A J1 haplocsoportot már nem a 2003-as filogenetikai fán jelölt M62 polimorfizmus, hanem a 2004-ben felfedezett M267 marker határozza meg (Cinnioğlu et al. 2004) (az M62 marker a J1 haplocsoport egyik szubklaszterét jelöli).



18. ábra A vizsgált Y-kromoszómális haplocsoportok genealógiája és előfordulási gyakorisága (%) a magyar és a székely populációban

A dőlt betűvel feltüntetett markereket nem tipizáltuk, de az allélikus állapotukra következtetni lehet a YCC 2003-as filogenetikai fa alapján (Jobling és Tyler-Smith 2003).

A megfigyelt haplocsoportok általánosan jellemzik a különböző európai populációkat, egyetlen kivételt képez ez alól a székely populációban felbukkanó P-M45\*(xM173) klaszter (3,1%), mely Közép-Ázsiában gyakori és csak nagyon ritkán fordul elő Európában (Wells et al. 2001).

Mindkét vizsgált populációban az M173 marker derivált C allélja által jellemzett R1 haplocsoport fordul elő a legnagyobb gyakorisággal (magyarok: 45%, székelyek: 38,2%). Ez a marker körülbelül 30 000 évvel ezelőtt, a felső paleolitikum idején terjedt el Európában Közép-Ázsia irányából (Semino et al. 2000a, Underhill et al. 2001, Wells et al. 2001). A mai európai férfiak körülbelül 50%-ának Y kromoszómáján van jelen ez az SNP marker (Semino et al. 2000a). Az irodalom szerint (Cruciani et al. 2002) a markerrel rendelkező európai férfiak alapvetően 2 szubklád valamelyikébe sorolhatók be. Az egyik az M17 marker által jellemzett R1a1 haplocsoport, a másik az M269 mutáció által meghatározott R1b3 klaszter.

Az R1a1-M17 haplocsoport a magyar populációban közel kétszer olyan gyakran fordul elő, mint a székely mintában (30,0% vs. 18,6%), ugyanakkor a magyar Y kromoszómáknak csak 15%-a sorolható be az R1b3-M269 leszármazási vonalba. A székelyeknél a két haplocsoport gyakorisága közel egyenlő (R1a1-M17: 18,6%, R1b3-M269: 19,6%).

A *P-M45\*(xM173)* haplocsoport jelenléte a székely populációban meglepő, mivel ez a klád szinte egyáltalán nem fordul elő Európában (Semino et al. 2000a, Wells et al. 2001).

A második leggyakoribb haplocsoport mindkét tipizált populációban az *I-M170* klaszter, a magyar populáció 24%-át, a székely populáció 21,7%-át jellemzi. Bár ez a haplocsoport széles körben elterjedt egész Európában, gyakorlatilag máshol nem fordul elő. A haplocsoportot meghatározó M170 mutáció valószínűleg Európában alakult ki az utolsó jégkorszak leghidegebb időszakát (LGM) megelőzően, körülbelül 22 000 éve (Semino et al. 2000a). A főbb szubkládjai közül az *Ia-M253* Észak-Európában fordul elő a leggyakrabban, míg az *Ib-P37* Kelet-Európában és a Balkánon elterjedt (Rootsi et al. 2004). Ez a két leszármazási vonal ellentétes megoszlási mintázatot mutat a 2 vizsgált népcsoportban. Míg a székely Y kromoszómák 16,5%-a sorolható be az *Ia-M253* haplocsoportba, addig a magyar populációban ez a klaszter csak 8%-os gyakorisággal fordul elő. Az *Ib-P37* haplocsoport a magyar minták 13%-át jellemzi, a székely populációban azonban csak 5,2%-os gyakoriságot mutat.

A jelen vizsgálatban az irodalmi adatokhoz képest a *J (I2f2/M304)* haplocsoport meglepően gyakori előfordulását tapasztaltuk a magyar (16%) és a székely (21,6%) populációban. A *J-I2f2 (M304)* klaszter az *E3b-M35* és a *G-M201* haplocsoportokkal együtt a neolitikum idején terjedt el először Európában a Közel-Kelet irányából (Semino et al. 2000a). Előfordulási gyakoriságuk a Közel-Kelet felől Európa irányába csökken. Európában a mediterrán partvidék mentén fordulnak elő számottevő gyakorisággal (Semino et al. 2000a, 2004).

A fentebb említett haplocsoportokon kívül 5%-nál nagyobb gyakorisággal van jelen a magyar és a székely populációban az *E3b-M35* klaszter (magyarok: 10%, székelyek: 9,2%), valamint a székely populációban a *G2-P15* haplocsoport (5,2%). Az *M35* marker által meghatározott *E3b* kládba tartozó Y kromoszómák többsége (89%; 90%) az *M78* marker által jellemzett *E3b1* szubklaszterbe sorolható.

### IV. 3. A modern minták Y-kromoszómális adatainak statisztikai vizsgálata

A modern minták vizsgálatával kapott eredményeinket statisztikai alapon is kiértékeltek. Az E3b-M35, G-M201, I-M170, J1-M267, J2-M172, N3-Tat, R1a1-M17 és az R1b3-M269 haplocsoportok gyakorisági értékei alapján meghatároztuk az általunk vizsgált magyar és székely populáció genetikai diverzitásának mértékét (Nei-féle géndiverzitás ( $H$ )), valamint páronkénti genetikai távolságot (páronkénti  $F_{st}$  (fixációs index) értékek) számoltunk a magyar és a székely populáció, illetve további 20 európai populáció között.

A genetikai távolság értéke 0 és 1 közé eső szám. A 0-hoz közel eső értékek alacsony szintű differenciáltságot tükröznek a vizsgált egységek között, míg az 1-hez közeli értékek a genetikai differenciáltság magas szintjét mutatják. Ha értéke 0,05-nél kisebb, a differenciáltságot csekély mértékűnek ítélnéljük, ekkor a vizsgált populációk között nincs szignifikáns különbség. Minél nagyobb a genetikai távolság értéke, annál „távolabbi rokonok” a vizsgált populációk, annál régebben váltak el a fejlődés során.

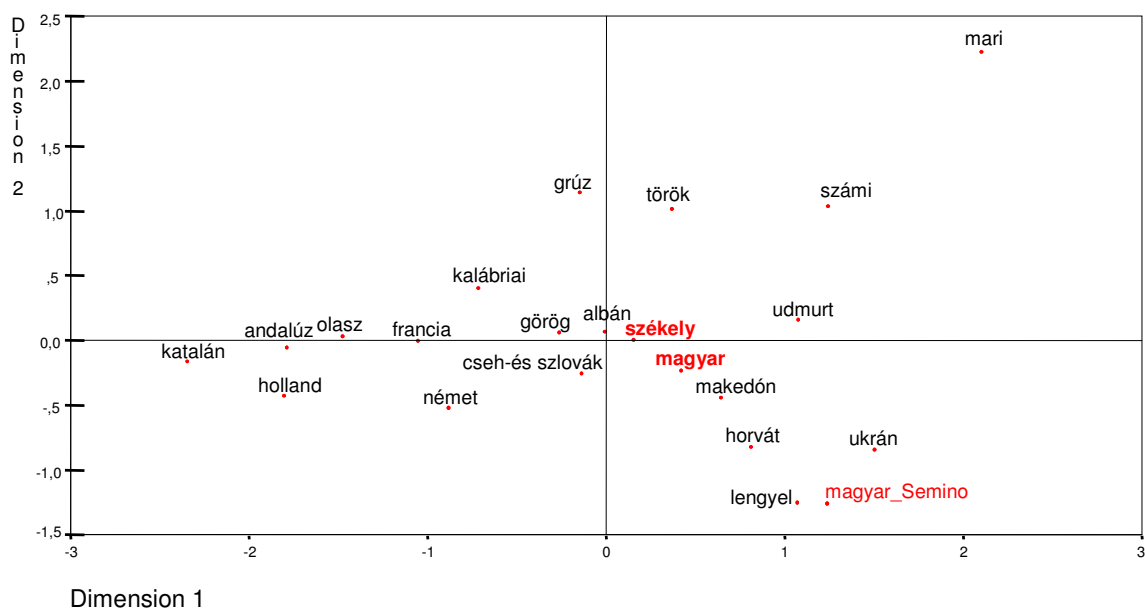
A páronkénti genetikai távolság értékek ( $F_{st}$ ) szerint a magyar és a székely populáció közeli rokonságban áll egymással ( $F_{st} < 0,05$ ;  $F_{st} = 0,00348$ ), valamint más közép-európai (pl. cseh és szlovák) és balkáni populációkkal (7. táblázat, 19. ábra). A populációk közötti távolságviszonyokat Multidimensional Scaling analízist használva egy 2 dimenziós térben ábrázoltuk, ahol az egyes populációknak megfelelő pontok közötti lineáris távolság egyenesen arányos a populációk közötti genetikai távolsággal. A populáción belüli diverzitási értékek ( $H$ ) alapján az általunk vizsgált magyar és székely populáció genetikailag heterogén összetételű (magyarok  $H = 0,8089$ , székelyek  $H = 0,8533$ ).



	<b>MAGYAR</b>	<b>SZÉKELY</b>
székely	0,00348	0,00000
makedón	-0,02375	-0,00178
horvát	0,02161	0,04260
cseh és szlovák	0,02206	0,01622
albán	0,02714	0,01152
ukrán	0,04514	0,09266
lengyel	0,04727	0,09274
görög	0,04821	0,02836
magyar_Semino	0,06189	0,10830
udmurt	0,07366	0,08250
török	0,08853	0,07070
német	0,09492	0,05992
kalábriai	0,09856	0,04930
francia	0,11197	0,06300
grúz	0,12494	0,09795
számi	0,13803	0,12275
olasz	0,17406	0,12687
andalúz	0,20057	0,14650
holland	0,20148	0,15481
katalán	0,26739	0,21399
mari	0,29971	0,27993

7. táblázat A magyar és a székely populáció további 20 európai populációhoz viszonyított genetikai ( $F_{st}$ ) távolság értékei

A szürke színnel kiemelt értékek 0,05-nél kisebbek, ennek megfelelően a kérdéses populációk között nincs szignifikáns különbség.



19. ábra 2 dimenziós MDS grafikon az általunk vizsgált magyar és székely populáció, valamint a Semino és munkatársai által (2000a) tipizált további 20 európai populáció páronkénti (Fst) genetikai távolság értékei alapján

## V. Diszkusszió

### V. 1. A Tat polimorfizmus vizsgálata

A Tat polimorfizmus egy olyan T>C pontmutáció az Y kromoszóma RBF5 lókuszán (Zerjal et al. 1997), mely különös érdeklődésre tart számot a finnugor nyelvű populációk vizsgálatakor. A Tat C allél az N3 haplocsoportot határozza meg, amely széles körben elterjedt Európa északi és keleti területein, ugyanakkor gyakorlatilag nem fordul elő a kontinens nyugati és déli részein élő populációkban. Ázsiában az északi területeken élő népcsoportokban fordul elő számottevő gyakorisággal (Zerjal et al. 1997, Rootsi et al. 2000, Rosser et al. 2000, Semino et al. 2000a, Zerjal et al. 2001, Tambets et al. 2004). Az irodalom alapján ez az Y-kromoszómális leszármazási vonal jelen van a legtöbb uráli nyelvű populációban, köztük a magyar nép legközelebbi nyelvi rokonainál is (Rédei 2005), a vogul és az osztják népcsoportban, ugyanakkor nem jellemzi a modern magyar populációkat, ezen belül a csángó és a palóc népcsoportot sem (Lahermo et al. 1999, Rootsi et al. 2000, Semino 2000b, Tambets et al. 2001, 2004).

Az irodalomban sokféle magyarázat született arra vonatkozóan, hogy vajon miért nincs jelen a Tat C allél az uráli nyelvű modern magyar populációban (Rootsi et al. 2000, Semino et al. 2000b). Az egyik felvetés szerint a vogulok és az osztjások, valamint több más szibériai populáció azután tettek szert erre az Y-kromoszómális leszármazási vonalra, amikor a magyarok ősei már maguk mögött hagyták a szibériai erdőseget és elindultak az eurázsiai sztyeppe felé. Egy másik lehetőség, hogy a magyarok és a szibériai ugor nyelvű populációk csak nyelvükben rokonok, genetikailag mindig is eltértek egymástól. Egy harmadik lehetőség szerint jelen volt a Tat C allél az ősi magyar populáció génekészletében, majd a vándorlások során, vagy a letelepedést követően veszett el a genetikai sodródás következtében.

Ezen megfigyelések alapján felvetődik a kérdés, hogy vajon a 9. század végén a Kárpát-medencében letelepült (Kovács 2005, Langó 2005) és uráli nyelvet beszélő honfoglalás kori populációban jelen volt-e a Tat C allél vagy sem.

Ennek a kérdésnek a tisztázásához (IX-) X. századi mintákon vizsgáltuk a Tat marker allélikus állapotát. A vizsgált leletek 6 különböző, a Kárpát-medence területén fekvő lelőhelyről származnak. A sikeres és szennyeződéstől mentes mitokondriális DNS amplifikációt követően próbáltunk meg Y-kromoszómális DNS-t kinyerni a

csontmintákból. A specifikus Y-kromoszómális analízis előtt minden esetben szükség volt a megmaradt teljes töredékes genom random primerekkel történő dúsítására. A 7 sikeresen tipizált 9-10. századi csontminta közül 2 esetben igazoltuk a mutáns C allél jelenlétét, míg 5 leletnél a vad típusú T allél előfordulását. A Tat mutációt hordozó 2 X. századi minta egy régióból, Békés megyéből származik. A 13/1 minta a szabadkígyospálligeti, a B1/3c lelet az örménykúti temető leletanyagát képezi. Mindkét lelet mellett klasszikus, gazdag mellékletanyagot tártak fel az ásatás során (13/1 minta: ló, kengyel, szablya, karperec; B1/3c minta: ló, kengyel, tegez). A T2/41 csontminta esetében kapott eredményt, mely szerint a vad típusú Tat T allél van jelen a mintában, sikerült megerősítenünk az M9 marker vad típusú C alléljának tipizálásával (YCC 2002, Jobling és Tyler-Smith 2003). A tény, hogy 7 csontminta közül kettő is hordozza a Tat C allélt, több mint érdekes, tekintve, hogy a 197 modern minta közül csak egyetlen székely férfi Y kromoszómáján detektáltuk az allél jelenlétét.

A Tat mutáció ritka előfordulása az általunk tanulmányozott modern magyar és székely populációban összhangban van a korábbi vizsgálatok eredményeivel (Lahermo et al. 1999, Rootsi et al. 2000, Rosser et al. 2000, Semino et al. 2000a, 2000b, Tambets et al. 2004) és további adatot szolgáltat ahhoz a felismeréshez, miszerint a Tat C allél gyakorlatilag hiányzik a modern magyarul beszélő populációkból.

A Tat mutáció által meghatározott N3 leszármazási vonal előfordulása a székely csoportban talán valamilyen múltbeli esemény genetikai hatását tükrözi. Természetesen néhány vizsgált minta és egy marker alapján nem vonhatunk le messzemenő következtetéseket, mégis elmondható, hogy a Tat mutáció jelenléte az ősi magyar populációban utalhat keveredésre vagy közös apai ági felmenőre más uráli nyelvű populációkkal.

Az eredményeink szerint a Tat C allél jelen volt a honfoglalás kori populációban, illetve a Tat polimorfizmus egy megfelelő marker az ősi magyar populáció Y-kromoszómális genetikai összetételének analíziséhez. A C allél jelenléte egyben alátámasztja eredményeink hitelességét is, hiszen sem a Régészeti Intézet munkatársai (régész, antropológus), sem csoportunk férfi tagjai, akik kapcsolatba kerültek a csontmintákkal, nem hordozzák ezt a mutációt.

A jövőben szeretnénk megvizsgálni, hogy a mutáció lokális jelenlétű-e, vagy előfordul más X-XI. századi temetőből származó csontleletek Y kromoszómáján is. Arra a kérdésre is szeretnénk választ kapni, hogy mikor került a Kárpát-medencébe ez a

mutáció. Éppen ezért olyan temetők leletanyagának tanulmányozásába kezdtünk, melyek használatában megfigyelhető a történeti kontinuitás. Ilyen temető pl. a csekeji lelőhely, ahova a 9. századtól a 11. századig temetkeztek. Ezt a temetőt az elképzelések szerint olyan szláv népesség használta, melyre hatással volt a X. századi honfoglalás kori magyar kultúra. A szarvas-kákapusztai temetőt olyan alapnépesség használta elsősorban a 9. században (de 2 sír a 10. századra datálható), melynek emlékényaga nem kapcsolódik a X. századi honfoglaló magyarokra jellemző kultúrához. Eddigi vizsgálataink során mindkét temetőből 1-1 leletet vizsgáltunk meg (48/B minta -Csekej, 378/B minta -Szarvas-Kákapusztai), egyikük sem hordozza a Tat mutációt. A 378/B lelet sírját tartják a szarvasi-kákapusztai temető egyik X. századi sírjának. A csekeji 48/B lelet a IX., illetve a X. századból is származhat; a leletanyag alapján nem egyértelmű a sír pontos kora.

## **V. 2. A ma élő magyar és székely populáció Y-kromoszómális genetikai összetételének vizsgálata**

A ma élő magyar és székely populációban, összesen 197 egyénnél, 22 Y kromoszómális biallélikus markert vizsgáltunk meg (18. ábra). A megfigyelt haplocsoportok általánosan jellemzik a különböző európai populációkat, egyetlen kivételt képez ez alól a székely populációban felbukkanó P-M45\*(xM173) klaszter (3,1%), mely Közép-Ázsiában gyakori és csak nagyon ritkán fordul elő Európában (Wells et al. 2001).

Az *R1a1-M17 haplocsoport* a leggyakoribb Y-kromoszómális klaszter a vizsgált magyar populációban (30%), míg a székelyek esetében 18,6%-os gyakorisággal van jelen.

Az M17 marker korát 10-15 000 évben határozták meg, valószínűleg az utolsó jégkorszak idején az ukrainai refúgiumban alakult ki. A haplocsoport gyakori (30-60%) Kelet-Európában, Közép-Ázsiában és Északnyugat-Indiában (Semino et al. 2000a, Wells et al 2001, 2003, Passarino et al. 2001, Kivisild et al. 2003). Európában a klaszter gyakorisága és varianciája nyugat felől keleti irányba nő, a legmagasabb értékeket a finnugor és a szláv nyelvű populációkban veszi fel (Peričić et al. 2005 és az ott idézett referenciák). Az irodalmi adatok szerint a haplocsoport elterjedésében legalább három

jelentős génáramlási esemény játszott szerepet: 1. Európa rekolonizációja az utolsó jégkorszak leghidegebb időszakát (LGM) követően a mai Ukrajna területén lévő refúgiumból (Semino et al. 2000a), ahol valószínűleg ez a mutáció kialakult; 2. vándorlások az északi Pontus-vidék területéről Kr. e. 3000 és 1000 között; 3. a szláv migráció Kr. u. 5–7. század között (Brather 2001, Fusek 2005) között (Peričić et al. 2005, Brather 2001, Fusek 2004).

Az R1a1-M17 haplocsoport előfordulási gyakorisága a vizsgált magyar (30%) és székely (18,6%) populációban a szomszédos területeken élő (például: csehek és szlovákok, Horvátország szárazföldi részén élő horvátok, bosnyákok, románok, szerbek), illetve néhány uráli nyelvű populáció (például: észtek, komik, mordvinok) esetében megfigyelt gyakorisági értékekhez (Tambets et al. 2004, Peričić et al. 2005 és a benne feltüntetett referenciák) hasonló.

Az R1b3-M269 haplocsoport az R1a1-M17 klaszterrel ellentétben Nyugat-Európában fordul elő a legnagyobb gyakorisággal (40-80%). Varianciája a legmagasabb értékeket Nyugat-Európában és Kis-Ázsiában veszi fel (Semino et al. 2000a, Peričić et al. 2005). Az irodalom szerint a haplocsoport elterjedési mintázata az európai kontinens Ibéria, illetve Kis-Ázsia területéről való benépesítését tükrözi a felső paleolitikum és a holocén idején (Cinnioğlu et al. 2004).

Az általunk vizsgált magyar férfiak 15%-a, a székely férfiak 19,6%-a tartozik ebbe a klaszterbe. A haplocsoport hasonló gyakorisággal van jelen néhány szláv (Horvátország szárazföldi részén élő horvátok, szlovének, lengyelek, bolgárok) és egyes uráli (komik, osztjások, mordvinok) nyelvű, valamint a román és a török populációban (Tambets et al. 2004, Peričić et al. 2005).

A közép-ázsiai P-M45\*(xM173) haplocsoport jelenléte a székely mintában szokatlan, mivel ez a klaszter szinte egyáltalán nem fordul elő Európában (Wells et al. 2001). A haplocsoport előfordulása valószínűleg némi ázsiai genetikai hatást tükröz, mely a székely populáció génkészletét akár az Erdélyben való letelepedés előtt akár azt követően érthette.

Az I-M170 haplocsoport az egyetlen olyan Y-kromoszómális klaszter, amely szinte kizárólag csak az európai kontinensen van jelen (Semino et al. 2000a, Rootsi et al. 2004). Mivel ez a klád gyakorlatilag más területeken nem fordul elő, így a Közel-Keleten sem, feltételezhető, hogy a haplocsoportot meghatározó M170 marker

Európában alakult ki az utolsó jégkorszak leghidegebb időszakát (LGM) megelőzően (Semino et al. 2000a).

Semino és munkacsoportja (2000a), illetve Barac és munkatársai (2003) feltevése szerint a Balkán északi része az LGM ideje alatt refúgiumként szolgálhatott és az M170 marker által meghatározott leszármazási vonalnak menedéket nyújthatott. Rootsi és munkatársai (2004) szerint a főbb szubkládok, az I1a, az I1b és az I1c a felső paleolitikum/mezolitikum periódusban váltak el az I\* leszármazási vonaltól. Az I1a-M253 szubklád Észak-Európában gyakori, Skandináviában az I haplocsoportba tartozó Y kromoszómák 88-100 %-a ebbe a szubkládba sorolható, majd gyakorisága gyorsan csökken mind a Kelet-európai-síkság, mind pedig az atlanti peremterület felé. A mikoszatellita diverzitás alapján az I1a és a kevésbé gyakori I1c szubkládok korai szétterjedése a franko-kantábriai refúgiumból történhetett. Az I1b-P37 leszármazási vonal az I haplocsoport leggyakoribb szubkládjá Kelet-Európában és a Balkánon (Rootsi et al. 2004). Peričić és munkatársai (2005) szerint az I1b\* (xM26) leszármazási vonalak Délkelet-Európából terjedhettek el Közép-, Kelet- és Dél-Európa felé valószínűleg a Younger Dryas holocén átmenet és a korai neolitikum közötti időszakban. (A Younger Dryas stadiális (rövidebb időtartamú lehűlés) 12 800 évvel ezelőtt kezdődött és 11 600 évvel ezelőtt végződött.) Közép- és Kelet-Európában az I1a és az I1b szubkládok átfedő gyakorisági grádiens mutatókat mutatnak (Rootsi et al. 2004).

Vizsgálatunkban az I haplocsoport közel azonos gyakorisággal fordul elő a magyar (24%) és a székely (21,7%) populációban. Ugyanakkor az I1a-M253 és az I1b\*(xM26) szubklád ellentétes megoszlási mintázatot mutat a két tipizált népcsoportban. Míg ugyanis az I1b szubklád előfordulási gyakorisága több mint kétszeres a magyar férfiak esetében (13%) a székely mintához (5,2%) viszonyítva, addig az I1a kétszeres gyakorisággal fordul elő a székely populációban (16,5%) a magyar csoporthoz (8%) képest. A megfigyelt gyakorisági értékek a Közép- és Kelet-Európában jellemző értéktartományon belül vannak (Rootsi et al. 2004, Peričić et al. 2005). Elképzelhető, hogy az I1a, illetve a fentebb említett R1b3 haplocsoport magasabb gyakorisága a székely populációban, legalább is részben, a 12. századtól Erdély területére betelepülő germán eredetű szászok (Makkai 1990, Kristó 2002) genetikai hatását tükrözi.

A *J*, az *E* és a *G* haplocsoportok európai elterjedését a neolitikum során a Közel-Keletről érkező földművelő népek expanziójához köthették (Semino et al. 2000a,

Underhill et al. 2000). A J és az E haplocsoport újabb filogeográfiai analízisének eredményei azt mutatják, hogy a neolitikum során és azt követően még számos migrációs esemény járult hozzá ezeknek a leszármazási vonalaknak az elterjedéséhez (Cruciani et al. 2004, Di Giacomo et al. 2004, Semino et al. 2004).

A jelen vizsgálatban az irodalmi adatokhoz képest a *J haplocsoport* meglepően gyakori előfordulását tapasztaltuk a magyar (16%) és a székely (21,6%) populációban.

A J klaszter, melyet a 12f2 polimorfizmus, illetve az azzal filogenetikailag egyenértékű M304 marker (Cinnioğlu et al. 2004) jellemez, a Közel-Keleten alakult ki (Semino et al. 2004). Gyakorisága a Közel-Kelet irányából Európa felé csökken. Európában a déli-délkeleti (a mediterrán területek középső régiójáig) területeken fordul elő nagyobb gyakorisággal (>20%). Az irodalmi adatok szerint ez a haplocsoport ritkán fordul elő a Balkán északi részén, illetve Közép-Európában (Di Giacomo et al. 2004 és az ott feltüntetett referenciák). A J haplocsoporton belül két nagyobb kládot különböztethetünk meg, ezek a J1-M267 és a J2-M172 (Cinnioğlu et al. 2004).

A *J1-M267 klád* nagy gyakorisággal van jelen a Közel-Keleten, Észak-Afrikában és Etiópiában (10-30%), míg Európában ritkán fordul elő (0-6%). Elterjedésében fontos szerepet tulajdonítanak a főként a 7. századtól történő arab diffúzióknak (Nebel et al. 2002, Semino et al. 2004).

A J1-M267 leszármazási vonal 3%-os gyakorisággal fordul elő a magyar populációban. Ez a gyakorisági érték beleillik a Közép-és Kelet-Európában tapasztalt értéktartományba (0-6%) (Semino et al. 2000a, 2004, Di Giacomo et al. 2004), ugyanakkor a székelyeknél detektált 10,3% -os gyakoriság érték ennél számottevően magasabb.

A *J2-M172 haplocsoport* sokkal elterjedtebb Európában, mint a J1-M267 klád. Szubkládjai közül a J2e-M12 és a J2f-M67, valamint az előbbieket származékai Európában és Ázsiában is jelen vannak (Semino et al. 2004). A J2e-M12 klasztert szinte teljes mértékben a J2e1-M102 szubklaszter reprezentálja, amely a legnagyobb gyakorisággal a Balkán déli részén és Északkelet-Olaszországban fordul elő. Valószínűleg a Balkán déli részéről terjedt el Európa többi területére (Semino et al. 2004).

A J2f-M67\* klád széles körben elterjedt, a Kaukázus területén élő populációkban fordul elő a leggyakrabban. A J2f1-M92 klaszter Anatólia és Dél-Olaszország között jelez rokonkapcsolatot; elsősorban a mediterrán területek északi részén van jelen,



Törökországtól nyugatra (Di Giacomo et al. 2004, Semino et al. 2004). A J2f-M67\* és a J2f1-M92 leszármazási vonalak a Boszporuszon keresztül érkezhettek Anatóliából Európába, valamint a tengerjáró neolitikus populációk révén, melyek talán elérték Dél-Olaszországot (Semino et al. 2004).

A J2-M172\* leszármazási vonal csökkenő gyakorisági grádienszt mutat a Közel-Kelet felől Nyugat-Európa irányába. Habár a kontinentális Európában és Indiában a legtöbb M172 markert hordozó Y kromoszóma a J2-M172\* klaszterbe sorolható, rokonságuk mértéke és közös történetük ezidáig bizonytalan maradt (Semino et al. 2004).

Az irodalom alapján (Semino et al. 2004) a J2-M172 (xM12) haplocsoport megoszlási mintázata a klaszter levantei/anatóliai folyosón keresztül történő európai elterjedését mutatja.

Semino és munkatársai (2004), valamint Di Giacomo és munkacsoportja (2004) szerint Európában a J haplocsoporton belül megfigyelhető diverzitás valószínűleg Törökország és az Égeikum területére vezethető vissza.

Vizsgálatainkban a J2-M172 szubkládok közül a J2e1-M102 nagyobb gyakoriságot mutat a székely populációban (7,2%, a magyaroknál 4%), míg a J2-M172\* Y kromoszómák gyakoribb előfordulást mutatnak a magyar populációban (8%, székelyeknél 3,1%). A J2f-M67\* szubkládba egy magyar, a J2f1-M92 szubklaszterbe egy székely férfi mintáját tudtuk besorolni.

A J2-M172 haplocsoport teljes gyakorisága a vizsgált magyar (13%) és székely (11,3%) populációban nagyságrendjében hasonló a csehek és szlovákok (Semino et al. 2000a), illetve a románok, bolgárok (Di Giacomo et al. 2004) és az ukránok (Semino et al. 2004) esetében tapasztalt értékekhez.

A korábbi vizsgálatok (Rosser et al. 2000, Semino et al. 2000 a,b, 2004) azt mutatták, hogy a *J haplocsoport* ritkán (2-3%) fordul elő a magyar nyelvű populációkban, habár a Rootsi és munkatársai (2000) által tipizált magyar csoportban 14%-os gyakorisággal volt jelen. Valószínűleg a mintavételben való eltérés eredményezte, hogy az általunk nyert gyakorisági adatok eltérnek a korábbi irodalmi értékektől. Az általunk vizsgált magyar minták Magyarország különböző területeiről származnak, de a leginkább reprezentált régió az Alföld területe volt (100 vizsgált férfiból 90 születési helye az Alföldön található). Ugyanakkor Semino és munkatársai (2000a) Budapestről, Egerből és Mátradereszkéről gyűjtöttek mintákat; Semino professzor leírása szerint elsősorban palócokat vizsgáltak.

A *J haplocsoport* nagyobb gyakorisága az általunk vizsgált populációkban (magyarok:16%, székelyek: 21,6%) egyrészt utalhat anatóliai, illetve dél-balkáni genetikai hatásra, ugyanakkor, főként a székely népcsoport esetében, genetikai sodródás következménye is lehet, hiszen a székelyek zárt közössége politikai okok miatt sokáig fennmaradt (Makkai 1990, Kristó 2002).

Más részről különböző történelmi események is hozzájárulhattak a *J haplocsoport* magasabb gyakoriságához. Ilyen esemény lehetett például az Ottomán Birodalom terjeszkedése, illetve ennek hatására a balkáni régióból Magyarország területére menekülő népek letelepedése (Pállfy 2000).

Elképzelhető, hogy az ősi magyar populációnak is szerepe lehetett a *J haplocsoport* Kárpát-medencében való elterjesztésében. A mitokondriális DNS vizsgálataink szerint (Tömöry et al. 2007) mindazok a populációk, melyek kapcsolatba kerültek az uráli régióból a Kárpát-medence területére vándorló korai magyar népcsoporttal, nyomot hagytak azok génkészletében. Az ősi magyar és a modern eurázsiai populációk mtDNS variációinak statisztikai analízise azt mutatta, hogy az ősi magyar populáció a legkisebb genetikai távolságra helyezkedik el nemcsak néhány közép-ázsiai és európai népcsoporttól, hanem néhány közel-keleti populációtól (szír, palesztin, török) is. A Közel-Kelet területén a *J haplocsoport* nagy gyakorisággal van jelen (Semino et al. 2000a, 2004, Cinnioğlu et al. 2004, Di Giacomo et al. 2004, Flores et al. 2005).

Ugyanakkor annak a lehetősége, hogy a *J klaszter* jelen volt az ősi magyar populációban és detektálható nyomot hagyott a modern magyar és székely apai ági génkészletben természetesen további vizsgálatokat igényel. Az analízis jelen állapotában más magyarázatok, mint a genetikai sodródás, vagy egy újabb keletű génáramlás, valószínűbbnek/elfogadhatóbbnak tűnnek.

Az általunk vizsgált, illetve a Semino és munkatársai által tipizált (2000a) populációk adataiból nyert genetikai távolság értékeket többdimenziós skálázási (Multidimensional Scaling) eljárással kétdimenziós térben ábrázoltuk. Az MDS grafikonon (19. ábra) az egyes populációknak megfelelő pontok közötti lineáris távolság egyenesen arányos a populációk közötti genetikai távolsággal. A statisztikai vizsgálat szerint az általunk vizsgált magyar és székely populáció genetikailag közeli rokonságban áll egymással, illetve más közép-európai és balkáni populációkkal. Ugyanakkor a 2. dimenzió egyértelműen elkülöníti egymástól az általunk analizált és a Semino és munkatársai által tipizált magyar csoportot. Az utóbbi közel térképeződik a

lengyel és az ukrán populációhoz. Egy lehetséges magyarázat lehet erre az eltérésre, hogy Semino és munkatársai elsősorban észak-magyarországi mintákat vizsgáltak, ahol a 19. században jelentős volt a szomszédos északi területek más anyanyelvű népességeivel való diffúzió (Kósa 1998). Valószínűleg ez az esemény felelős az általuk tapasztalt kiugróan magas R1a1-M17 haplocsoport gyakoriságért is (60%-míg az általunk vizsgált mintánál 30%).

Eredményeink alapján kimondható, hogy Semino és munkacsoportjának adatai (2000a) nem reprezentatívak a jelenkori magyar populáció vonatkozásában.

Az *E3b-M35 haplocsoport* 10%-os gyakorisággal fordul elő a vizsgált magyar és 9,2%-os gyakorisággal a székely populációban (*E3b\** + *E3b1* szubklaszterek együttes gyakorisága – 18. ábra). Mind a két népcsoportban csaknem az összes minta (~90%) az M78 marker által meghatározott *E3b1* szubklaszterbe sorolható. Az *E* haplocsoport főleg Afrikában elterjedt, de származéka az *E3b-M35* klád Európában is jelen van, ahol a neolitikum idején, illetve azt követően terjedt el. Mind az M35, mind pedig az M78 marker valószínűleg Kelet-Afrikában alakult ki (Semino et al. 2004, Cruciani et al. 2004).

Az *E3b1-M78 haplocsoport* széles körben elterjedt, jelen van Észak-és Kelet-Afrikában, a Közel-Keleten és Európában is, ahol a leggyakrabban előforduló *E3b* szubklasztert reprezentálja (Semino et al. 2004, Cruciani et al. 2004). Ez a leszármazási vonal a legnagyobb gyakorisággal Dél- és Délkelet-Európában fordul elő, majd csökkenő gyakorisági grádiens mutat Nyugat-, Közép-és Kelet-Európa irányába (Peričić et al. 2005).

Gyakorisága az általunk vizsgált magyar (9%) és székely (8,2%) populációban illeszkedik a Közép-és Kelet-Európában megfigyelt értékekhez (Peričić et al. 2005).

Az irodalom alapján (Peričić et al. 2005) az *E3b1-M78 haplocsoport* elterjedésének egyik legfontosabb iránya (~7300 évvel ezelőtt) Európa déli, délkeleti része felől vezetett a Vardar–Morava–Duna vonalon keresztül a kontinentális területek irányába. Cruciani és munkatársai szerint (2004) az *E3b* leszármazási vonalak jelenlegi európai megoszlási mintázatának kialakításában fontos szerepe volt a körülbelül 7 800 éve a Balkán területéről Nyugat-Európa és Közép-Európa déli része felé irányuló expanciónak.

## VI. Következtetések

Az eredményeink szerint a Tat C allél, amely széles körben elterjedt az uráli nyelvű populációkban, jelen volt az ősi magyar populáció génkészletében a Kárpát-medence területén történő letelepedéskor. Ugyanakkor eredményeink összhangban vannak a korábbi irodalmi adatokkal, miszerint ez a mutáció gyakorlatilag hiányzik a ma élő magyarul beszélő populációk génkészletéből; egyetlen székely férfi Y kromoszómáján tudtuk csak a C allélt detektálni.

A modern magyar és székely populáció Y-kromoszómális mintázata alapvetően jól illeszkedik az Európában megfigyelhető genetikai jellemzőkhöz. Más európai populációkhoz hasonlóan a magyar és a székely népcsoportban is nagyobb részt a kontinensen a paleolitikum óta jelen lévő korai leszármazási vonalak fordulnak elő, míg a neolitikus és az azt követő vándorlási események hatása kisebb mértékben érzékelhető génkészletükben.

A korábbi vizsgálatok eredményeivel összhangban az általunk vizsgált magyarul beszélő populációk genetikailag közeli rokonságban vannak a földrajzi szomszédokkal. A magyar és a székely populáció közel térképeződik néhány egyéb közép-európai (például csehek és szlovákok), de főként a balkáni populációkhoz.

Ugyanakkor a jelenlegi vizsgálatban a P-M45\*(xM173) haplocsoport jelenlétét detektáltuk a székely populációban. Ez a klaszter csak nagyon ritkán fordul elő Európában. Jelenléte a székely mintában a közép-ázsiai populációkkal való genetikai kapcsolatra utalhat.

A magyar és a székely populációban az irodalmi adatokhoz képest magasabb gyakoriságban detektáltuk a J haplocsoport előfordulását. A J haplocsoport nagyobb gyakorisága az általunk vizsgált populációkban egyrészt anatóliai és dél-balkáni genetikai hatás következménye lehet. Más részről a történelmi adatok és az ősi magyar populáció anyai leszármazási vonalainak összehasonlító vizsgálata alapján elképzelhető, hogy a korai magyar populációnak is szerepe lehetett a J haplocsoport Kárpát-medencében történő elterjedésében.

## VII. Köszönetnyilvánítás

Hálás köszönettel tartozom szüleimnek és testvéremnek, akik minden helyzetben mellettem álltak és támogatták tanulmányaimat. Hálás vagyok férjemnek, kisfiamnak és a férjem családjának türelmükért, kitartásukért, odaadó támogatásukért. Köszönettel tartozom barátaimnak lelkes támogatásukért.

Köszönetet mondok témavezetőmnek, Dr. Raskó István professzornak, aki kitüntetett bizalmával és csoportjába fogadott, majd segítette szakmai fejlődésemet, munkámat.

Köszönetem fejezem ki csoportunk minden tagjának nélkülözhetetlen tanácsaikért, segítőkészségükért és a jó hangulatban eltöltött munkanapokért. Külön köszönöm Czibula Ágnesnek a primer tervező és szekvencia analizáló számítógépes programok használatában nyújtott segítségét. Köszönöm Bogácsi-Szabó Erikának, hogy mind szakmailag, mind barátilag minden esetben számíthattam rá.

Köszönettel tartozom Langó Péternek és Dr. Mende Balázs Gusztávnak a régészeti és embertani kérdésekben nyújtott segítségükért.

Köszönetet mondok Dr. Csete Klárának a vizsgálatokhoz szükséges modern minták rendelkezésemre bocsátásáért.

Köszönetet mondok Dr. Németh Istvánnak az ALF rendszer működésének elsajátításában, illetve Dr. Zsolnai Attilának a modern minták futtatásában nyújtott hasznos segítségéért. Köszönettel tartozom Dr. Cserpán Imrének és Dr. Pénzes Zsoltnak hasznos szakmai tanácsaikért.

Hálás vagyok Dr. C. Stephen Downes professzornak, hogy Coleraine-ben, a University of Ulster, School of Biomedical Sciences, Cancer and Ageing munkacsoportjában dolgozhattam, valamint Eleanore K. Conant-nak hasznos szakmai tanácsaiért.

Köszönetem fejezem ki a Genetikai Intézet valamennyi dolgozójának tanácsaikért és segítőkészségükért.

## VIII. Irodalomjegyzék

Andrews, R.M., Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N. (1999) Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet* **23**(2):147.

Barač, L., Peričić, M., Klarić, I. M., Rootsi, S., Janićijević, B., Kivisild, T., Parik, J., Rudan, I., Villems, R. & Rudan, P. (2003) Y chromosomal heritage of Croatian population and its island isolates. *Eur J Hum Genet* **11**:535–542.

Bálint, Cs. (1996) A kora középkori kelet-európai steppe régészete és a 9–10. századi magyarok. (The Archaeology of the Early central European Steppe and the Hungarians of the 9<sup>th</sup> and 10<sup>th</sup> Centuries.) *MT* **103**: 937–947.; 1041. Akadémiai Kiadó.

Bálint, Cs. (2005) Ki volt „magyar” a honfoglalás korban és Szent István korában? In: *Mi a magyar?* (eds. I. Romsics & M. Szegedy-Maszák), pp. 37–56. Budapest: Rubicon Kiadó.

Bálint, Cs. (2006) Az etnosz a korai középkorban. (A kutatás lehetőségei és gátjai.) (The *Ethnos* in the Early Middle Ages. Possibilities and Limits of Research.) *Sz* **140**:277–347. Magyar Történelmi Társulat.

Belgrader, P., Marino, M.M., Lubin, M. & Barany, F. (1996) A multiplex PCR-ligase detection reaction assay for human identity testing. *Genome Sci Technol* **1**:77-87.

Binladen J., Wiuf C, Gilbert MT, Bunce M, Barnett R, Larson G, Greenwood AD, Haile J, Ho SY, Hansen AJ, Willerslev E. (2006) Assessing the fidelity of ancient DNA sequences amplified from nuclear genes. *Genetics* **172**(2):733-741.

Bogácsi-Szabó, E., Kalmár, T., Csányi, B., Tömöry, Gy., Czibula, Á., Priskin, K., Horváth, F., Downes, C.S. & Raskó, I. (2005) Mitochondrial DNA of ancient Cumanians: culturally Asian steppe nomadic immigrants with substantially more western Eurasian mitochondrial DNA lineages. *Hum Biol* **77**:639-662.

Bóna, I. (1990) Die ungarische Landnahme. Siebenbürger und der Ostrand der Großen Ungarischen Tiefebene zwischen 895 und 1003. In: *Kurze Geschichte Siebenbürgens*. (ed. B. Köpeczi), pp. 109–136.; 697–700. Budapest: Akademia Verlag.

Bosch, E., Calafell F, Rosser ZH, Nørby S, Lynnerup N, Hurler ME, Jobling MA. (2003) High level of male-biased Scandinavian admixture in Greenlandic Inuit shown by Y-chromosomal analysis. *Hum Genet* **112**(4):353-63.

Brather, S. (2001) *Archäologie der westlichen Slawen*. Reallexikon der Germanischen Altertumskunde – Ergänzungsbände, 30. Berlin–New York. Walter de Gruyter Verlag.

Brinkmann B., Klintschar M, Neuhuber F, Hühne J, Rolf B. (1998) Mutation rate in human microsatellites: influence of the structure and length of the tandem repeat. *Am J Hum Genet* **62**(6):1408-15.

Burger, J., Hummel S, Hermann B, Henke W. (1999) DNA preservation: a microsatellite-DNA study on ancient skeletal remains. *Electrophoresis* **20**(8):1722-8.

Calafell F, Shuster A, Speed WC, Kidd JR, Kidd KK. (1998) Short tandem repeat polymorphism evolution in humans. *Eur J Hum Genet* **6**(1):38-49.

Cann, R.L., Stoneking, M., Wilson, A.C. (1987) Mitochondrial DNA and human evolution. *Nature* **325**:31-6.

Carvajal-Carmona L.G., Soto ID, Pineda N, Ortíz-Barrientos D, Duque C, Ospina-Duque J, McCarthy M, Montoya P, Alvarez VM, Bedoya G, Ruiz-Linares A. (2000) Strong Amerind/white sex bias and a possible Sephardic contribution among the founders of a population in northwest Colombia. *Am J Hum Genet* **67**(5):1287-95.

Carvalho-Silva D.R., Santos FR, Rocha J, Pena SD. (2001) The phylogeography of Brazilian Y-chromosome lineages. *Am J Hum Genet* **68**(1):281-6.

Casanova M, Leroy P, Boucekkine C, Weissenbach J, Bishop C, Fellous M, Purrello M, Fiori G, Siniscalco M. (1985) A human Y-linked DNA polymorphism and its potential for estimating genetic and evolutionary distance. *Science* **230**(4732):1403-6.

Chambers, G.K. & MacAvoy, E.S. (2000) Microsatellites: consensus and controversy. *Comp Biochem Physiol B Biochem Mol Biol* **126**(4):455-76.

Charlesworth B., Sniegowski P, Stephan W. (1994) The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature* **371**(6494):215-20.

Cinniöglu, C., King, R., Kivisild, T., Kalfoglu, E., Atasoy, S., Cavalleri, G.L., Lillie, A.S., Roseman, C.C., Lin, A.A., Prince, K., Oefner, P.J., Shen, P., Semino, O., Cavalli-Sforza, L.L. & Underhill, P.A. (2004) Excavating Y-chromosome haplotype strata in Anatolia. *Hum Genet* **114**:127–148.

Cruciani, F., Santolamazza, P., Shen, P., Macaulay, V., Moral, P., Olckers, A., Modiano, D., Holmes, S., Destro-Bisol, G., Coia, V., Wallace, D.C., Oefner, P.J., Torroni, A., Cavalli-Sforza, L.L., Scozzari, R. & Underhill, P.A. (2002) A back migration from Asia to sub Saharan Africa is supported by high-resolution analysis of human Y-chromosome haplotypes. *Am J Hum Genet* **70**:1197–1214.

Cruciani, F., La Fratta, R., Santolamazza, P., Sellitto, D., Pascone, R., Moral, P., Watson, E., Guida, V., Colomb, E.B., Zaharova, B., Lavinha, J., Vona, G., Aman, R., Calì, F., Akar, N., Richards, M., Torroni, A., Novelletto, A. & Scozzari, R. (2004) Phylogeographic analysis of haplogroup E3b (E-M215) Y chromosomes reveals multiple migratory events within and out of Africa. *Am J Hum Genet* **74**:1014–1022.

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. (2008) Y-chromosome analysis of ancient Hungarian and two modern Hungarian-speaking populations from the Carpathian Basin. *Ann Hum Genet* **72**:519-34. Epub 2008 Mar 27.

de Knijff P., 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, Roewer L, et al. (1997) Chromosome Y microsatellites: population genetic and evolutionary aspects. *Int J Legal Med* **110**(3):134-49.

Di Giacomo, F., Luca, F., Popa, L.O., Akar, N., Anagnou, N., Banyko, J., Brdicka, R., Barbujani, G., Papola, F., Ciavarella, G., Cucci, F., Di Stasi, L., Gavrila, L., Kerimova, M.G., Kovatchev, D., Kozlov, A.I., Loutradis, A., Mandarino, V., Mammi, C., Michalodimitrakis, E.N., Paoli, G., Pappa, K.I., Pedicini, G., Terrenato, L., Tofanelli, S., Malaspina, P. & Novelletto, A. (2004) Y chromosomal haplogroup J as a signature of the post-Neolithic colonization of Europe. *Hum Genet* **115**:357–371.

Fernandes S., Huellen K, Goncalves J, Dukal H, Zeisler J, Rajpert De Meyts E, Skakkebaek NE, Habermann B, Krause W, Sousa M, Barros A, Vogt PH. (2002) High frequency of DAZ1/DAZ2 gene deletions in patients with severe oligozoospermia. *Mol Hum Reprod* **8**(3):286-98.

Flores, C., Maca-Meyer, N., Larruga, J.M., Cabrera, V.M., Karadsheh, N. & Gonzalez, A.M. (2005) Isolates in a corridor of migrations: a high-resolution analysis of Y-chromosome variation in Jordan. *J Hum Genet* **50**:435-441.

Fodor, I. (1996) The Hungarian Conquest. In: *The Ancient Hungarians*. (eds. I. Fodor, L. Révész, M. Wolf & I.M. Nepper), pp. 13-18. Budapest: Hungarian National Museum.

Fusek, G. (2004) „Slawen” oder Slawen? Eine polemische Auseinandersetzung über eine wertvolle Monographie. *SlovArch* **52**:161–181. Arheologicky Ústav SAV.

Giles R.E., Blanc H., Cann H.M., Wallace D.C. (1980) Maternal inheritance of human mitochondrial DNA. *Proc Natl Acad Sci USA* **77**. 6715-9.

Guglielmino, C.R., Piazza, A., Menozzi, P. & Cavalli-Sforza, L.L. (1990) Uralic genes in Europe. *Am J Phys Anthropol* **83**:57-68.

Guglielmino, C.R., De Silvestri, A. & Béres, J. (2000) Probable ancestors of Hungarian ethnic groups: an admixture analysis. *Ann Hum Genet* **64**:145-159.

Hammer, M.F., Redd, A.J., Wood, E.T., Bonner, M.R., Jarjanazi, H., Karafet, T., Santachiara-Benerecetti, S., Oppenheim, A., Jobling, M.A., Jenkins, T., Ostrer, H. & Bonnè-Tamir, B. (2000) Jewish and Middle Eastern non-Jewish populations share a common pool of Y-chromosome biallelic haplotypes. *Proc Natl Acad Sci U S A* **97**:6769-6774.



- Hammer M.F., Karafet TM, Redd AJ, Jarjanazi H, Santachiara-Benerecetti S, Soodyall H, Zegura SL. (2001) Hierarchical patterns of global human Y-chromosome diversity. *Mol Biol Evol* **18**(7):1189-203.
- Hammer, M.F. & Zegura, S.L. (2002) The Human Y chromosome haplogroup tree: Nomenclature and Phylogeography of Its Major Divisions. *Annu Rev Anthropol* **31**: 303-321.
- Handt, O., Höss, M., Krings, M., Pääbo, S. (1994) Ancient DNA: methodological challenges. *Experientia* **50**: 524-529.
- Handt, O., Richards, M., Trommsdorff, M., Kilger, C., Simanainen, J., Georgiev, O., Bauer, K., Stone, A., Hedges, R., Schaffner, W., Utermann, G., Sykes, B. & Pääbo, S. (1994) Molecular genetic analyses of the Tyrolean Ice Man. *Science* **264**:1775-1778.
- Hanson, E.K., & Ballantyne, J. (2005) Whole genome amplification strategy for forensic genetic analysis using single or few cell equivalents of genomic DNA. *Anal Biochem* **346**:246-257.
- Henke, J. & Henke, L. (1999) Mutation rate in human microsatellites. *Am J Hum Genet.* **64**(5):1473-4.
- Heyer E., Puymirat J, Dieltjes P, Bakker E, de Knijff P. (1997) Estimating Y chromosome specific microsatellite mutation frequencies using deep rooting pedigrees. *Hum Mol Genet* **6**(5):799-803.
- Höss M., Jaruga P, Zastawny TH, Dizdaroglu M, Pääbo S. (1996) DNA damage and DNA sequence retrieval from ancient tissues. *Nucleic Acids Res* **24**(7):1304-7.
- Hurles M.E., Irvén C., Nicholson J., Taylor P.G., Santos F.R., Loughlin J., Jobling M.A., Sykes B.C. (1998) European Y-chromosomal lineages in Polynesians: a contrast to the population structure revealed by mtDNA. *Am J Hum Genet* **63**(6):1793-806.
- Ingman, M., Kaessmann, H., Pääbo S., Gyllensten, U. (2000) Mitochondrial genome variation and the origin of modern humans. *Nature* **408**:708-713.
- Jobling MA, Pandya A, Tyler-Smith C. (1997) The Y chromosome in forensic analysis and paternity testing. *Int J Legal Med* **110** (3):118-24.
- Jobling, M.A. & Tyler-Smith, C. (2000) New uses for new haplotypes the human Y chromosome, disease and selection. *Trends Genet* **16**(8):356-62.
- Jobling, M.A. & Tyler-Smith, C. (2003) The human Y chromosome: an evolutionary marker comes of age. *Nat Rev Genet* **4**:598-612.

Kalmár, T., Bachrati, Cs.Z., Marcsik, A. & Raskó, I. (2000) A simple and efficient method for PCR amplifiable DNA extraction from ancient bones. *Nucleic Acids Res* **28**:E67.

Karafet, T., Xu, L., Du, R., Wang, W., Feng, S., Wells, R.S., Redd, A.J., Zegura, S.L. & Hammer, M.F. (2001) Paternal population history of East Asia: sources, patterns, and microevolutionary processes. *Am J Hum Genet* **69**:615-628.

Kayser M, Caglià A, Corach D, Fretwell N, Gehrig C, Graziosi G, Heidorn F, Herrmann S, Herzog B, Hidding M, Honda K, Jobling M, Krawczak M, Leim K, Meuser S, Meyer E, Oesterreich W, Pandya A, Parson W, Penacino G, Perez-Lezaun A, Piccinini A, Prinz M, Schmitt C, Roewer L, et al. (1997) Evaluation of Y-chromosomal STRs: a multicenter study. *Int J Legal Med* **110**(3):125-33, 141-9.

Kayser M., Roewer L, Hedman M, Henke L, Henke J, Brauer S, Krüger C, Krawczak M, Nagy M, Dobosz T, Szibor R, de Knijff P, Stoneking M, Sajantila A. (2000) Characteristics and frequency of germline mutations at microsatellite loci from the human Y chromosome, as revealed by direct observation in father/son pairs. *Am J Hum Genet* **66**(5):1580-8.

Kivisild, T., Rootsi, S., Metspalu, M., Mastana, S., Kaldma, K., Parik, J., Metspalu, E., Adojaan, M., Tolk, H.V., Stepanov, V., Gölge, M., Usanga, E., Papiha, S.S., Cinnioğlu, C., King, R., Cavalli-Sforza, L.L., Underhill, P.A. & Villems, R. (2003) The genetic heritage of the earliest settlers persists both in Indian tribal and caste populations. *Am J Hum Genet* **72**:313-332.

Kordé, Z. (1994) Székelyek. In: *Korai magyar történeti lexikon (9–14. század)* (ed. Gy. Kristó), pp. 623–625. Budapest: Akadémiai Kiadó.

Kósa, L. (1998) *'Ki népei vagytok?'*. Magyar néprajz. Budapest.

Kovács, L. (2005) Remarks on the archaeological remains of the 9<sup>th</sup>–10<sup>th</sup> century Hungarians. In: *Research on the prehistory of the Hungarians: A review*. (ed. B.G. Mende) *VAH* **18**:351–368. Hungarian Academy of Sciences Archaeological Institute.

Kristó, Gy. (2002) *A székelyek eredete*. Budapest: Balassi Kiadó.

Kristó, Gy. (2005) *A korai Erdély*. Szeged. Szegedi Tudományegyetem Középkorász Műhelye.

Krausz C, Quintana-Murci L, McElreavey K. (2000) Prognostic value of Y deletion analysis: what is the clinical prognostic value of Y chromosome microdeletion analysis? *Hum Reprod* **15**(7):1431-4.

- Lahermo, P., Savontaus, M.L., Sistonen, P., Béres, J., de Knijff, P., Aula, P. & Sajantila, A. (1999) Y chromosomal polymorphisms reveal founding lineages in the Finns and the Saami. *Eur J Hum Genet* **7**:447-458.
- Lahermo, P., Laitinen, V., Sistonen, P., Béres, J., Karcagi, V. & Savontaus, M.L. (2000) MtDNA polymorphism in the Hungarians: comparison to three other Finno-Ugric-speaking populations. *Hereditas* **132**:35-42.
- Langó, P. (2005) Archaeological research on the conquering Hungarians: A review. In: *Research on the prehistory of the Hungarians: A review* (ed. B.G. Mende) *VAH* **18**: 175–340. Hungarian Academy of Sciences Archaeological Institute.
- Lell, J.T., Sukernik RI, Starikovskaya YB, Su B, Jin L, Schurr TG, Underhill PA, Wallace DC. (2002) The dual origin and Siberian affinities of Native American Y chromosomes. *Am J Hum Genet* **70**(1):192-206.
- Lewontin, R.C. (1972) The apportionment of human diversity. *Evolutionary Biology* **6**: 381-398.
- Lindahl, T. (1993) Instability and decay of the primary structure of DNA. *Nature* **362**: 709-715.
- Makkai, L. (1990) Siebenbürgen im mittelalterlichen Königreich Ungarn. In: *Kurze Geschichte Siebenbürgens* (ed. B. Köpeczi), pp. 137-240. Budapest: Akademia Verlag.
- Mende, B. G. (ed.) (2005) Research on the Prehistory of the Hungarians: A Review. In: Mende Balázs G. (ed.): *Varia Archaeologica Hungarica*. 18. Hungarian Academy of Sciences Archaeological Institute.
- Mesterházy, K. (1997) A honfoglaló magyarok tárgyi emlékei. *Életünk*, pp. 30–67.
- Miller, S.A., Dykes, D.D. & Polesky, H.F. (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* **16**:1215.
- Nakamura Y., Leppert M, O'Connell P, Wolff R, Holm T, Culver M, Martin C, Fujimoto E, Hoff M, Kumlin E, et al. (1987) Variable number of tandem repeat (VNTR) markers for human gene mapping. *Science* **235**(4796):1616-22.
- Nebel, A., Landau-Tasser, E., Filon, D., Oppenheim, A. & Faerman, M. (2002) Genetic evidence for the expansion of Arabian tribes into the Southern Levant and North Africa. *Am J Hum Genet* **70**:1594–1596.
- Neff, M.M., Turk, E. & Kalishman, M. (2002) Web-based primer design for single nucleotide polymorphism analysis. *Trends Genet* **18**:613-615.
- Nei, M. (1987) *Molecular evolutionary genetics*. Columbia University Press, New York, NY, USA.

Onofri, V., Alessandrini, F., Turchi, C., Pesaresi, M., Buscemi, L. & Tagliabracci, A. (2006) Development of multiplex PCRs for evolutionary and forensic applications of 37 human Y chromosome SNPs. *Forensic Sci Int* **157**:23-35.

Pääbo S. (1989) Ancient DNA; extraction, characterization, molecular cloning and enzymatic amplification. *Proc Natl Acad Sci U S A* **86**: 1939-1943.

Pääbo S., Irwin DM, Wilson AC. (1990) DNA damage promotes jumping between templates during enzymatic amplification. *J Biol Chem* **265**(8):4718-21.

Pääbo S., Poinar H., Serre D., Jaenicke-Despres V., Hebler J., Rohland N., Kuch M., Krause J., Vigilant L., Hofreiter M. (2004) Genetic analyses from ancient DNA. *Annu Rev Genet* **38**:645-79.

Pálffy, G. (2000) *A tizenhatodik század története*. Budapest: Pannonica Kiadó.

Paracchini, S., Arredi, B., Chalk, R. & Tyler-Smith, C. (2002) Hierarchical high-throughput SNP genotyping of the human Y chromosome using MALDI-TOF mass spectrometry. *Nucleic Acids Res* **30**:E27.

Passarino, G., Semino, O., Magri, C., Al-Zahery, N., Benuzzi, G., Quintana-Murci, L., Andellnovic, S., Bulic-Jakus, F., Liu, A., Arslan, A. & Santachiara-Benerecetti, A.S. (2001) The 49a,f haplotype 11 is a new marker of the EU19 lineage that traces migrations from northern regions of the Black Sea. *Hum Immunol* **62**:922-932.

Peričić, M., Lauc, L.B., Klarić, I.M., Rootsi, S., Janičević, B., Rudan, I., Terzić, R., Čolak, I., Kvesić, A., Popović, D., Šijački, A., Behluli, I., Đorđević, D., Efremovska, L., Bajec, Đ.D., Stefanović, B.D., Villems, R. & Rudan, P. (2005) High-resolution phylogenetic analysis of southeastern Europe traces major episodes of paternal gene flow among Slavic populations. *Mol Biol Evol* **22**:1964-1975.

Poinar H.N., Höss M, Bada JL, Pääbo S. (1996) Amino acid racemization and the preservation of ancient DNA. *Science* **272**(5263):864-866.

Pritchard J.K., Seielstad MT, Perez-Lezaun A, Feldman MW. (1999) Population growth of human Y chromosomes: a study of Y chromosome microsatellites. *Mol Biol Evol* **16** (12):1791-8.

Rédei, K. (2005) *Őstörténetünk kérdései. A nyelvészeti dilettantizmus kritikája*. Budapest: Balassi Kiadó.

Révész, L. (1996) *A karosi honfoglalás kori temetők. Adatok a Felső-Tisza-vidék X. századi történetéhez*. (Die Gräberfelder von Karos aus der Landnahmezeit. Archäologische Angaben zur Geschichte des Oberen Theißgebietes im 10. Jahrhundert.) Magyarország honfoglalás kori és kora Árpád-kori sírleletei, 1. Miskolc. Herman Ottó Museum.

Reynolds, J., Weir, B.S. & Cockerham, C.C. (1983) Estimation of the coancestry coefficient: basis for a short-term genetic distance. *Genetics* **105**:767-779.

Richards M.B., Macaulay VA, Bandelt HJ, Sykes BC. (1998) Phylogeography of mitochondrial DNA in western Europe. *Ann Hum Genet* **62**(Pt 3):241-60.

Róna-Tas, A. (1995) *A magyarság korai története*. Magyar Őstörténeti Könyvtár, 9. Szeged. University of Szeged.

Róna-Tas, A. (1996) *A honfoglaló magyar nép. Bevezetés a korai magyar történelem ismeretébe*. Budapest.

Róna-Tas, A. (1999) *Hungarians and Europe in the Early Middle Ages*. Budapest: Central Europe University Press.

Rootsi, S., Kivisild, T., Tambets, K., Adojaan, M., Parik, J., Reidla, M., Metspalu, E., Laos, S., Tolk, H.V. & Villems, R. (2000) On the phylogeographic context of sex-specific genetic markers of Finno-Ugric populations. In: *The roots of peoples and languages of Northern Eurasia II and III* (ed. A. Kiinnap), pp. 148-164. University of Tartu, Division of Uralic Languages/Societas Historiae Fenno-Ugricae, Tartu.

Rootsi, S., Magri, C., Kivisild, T., Benuzzi, G., Help, H., Bermisheva, M., Kutuev, I., Barac, L., Peričić, M., Balanovsky, O., Pshenichnov, A., Dion, D., Grobei, M., Zhivotovsky, L.A., Battaglia, V., Achilli, A., Al-Zahery, N., Parik, J., King, R., Cinnioglu, C., Khusnutdinova, E., Rudan, P., Balanovska, E., Scheffrahn, W., Simonescu, M., Brehm, A., Goncalves, R., Rosa, A., Moisan, J.P., Chaventre, A., Ferak, V., Füredi, S., Oefner, P.J., Shen, P., Beckman, L., Mikerezi, I., Terzić, R., Primorac, D., Cambon-Thomsen, A., Krumina, A., Torroni, A., Underhill, P.A., Santachiara-Benerecetti, A.S., Villems, R. & Semino, O. (2004) Phylogeography of Y-Chromosome haplogroup I reveals distinct domains of prehistoric gene flow in Europe. *Am J Hum Genet* **75**:128-137.

Rootsi, S., Zhivotovsky LA, Baldovic M, Kayser M, Kutuev IA, Khusainova R, Bermisheva MA, Gubina M, Fedorova SA, Ilumäe AM, Khusnutdinova EK, Voevoda MI, Osipova LP, Stoneking M, Lin AA, Ferak V, Parik J, Kivisild T, Underhill PA, Villems R. (2007) A counter-clockwise northern route of the Y-chromosome haplogroup N from Southeast Asia towards Europe. *Eur J Hum Genet* **15**(2):204-11.

Rosser, Z.H., Zerjal, T., Hurler, M.E., Adojaan, M., Alavantic, D., Amorim, A., Amos, W., Armenteros, M., Arroyo, E., Barbujani, G., Beckman, G., Beckman, L., Bertranpetit, J., Bosch, E., Bradley, D.G., Brede, G., Cooper, G., Côté-Real, H.B.S.M., de Knijff, P., Decorte, R., Dubrova, Y.E., Evgrafov, O., Gilissen, A., Glisic, S., Gölge, M., Hill, E.W., Jeziorowska, A., Kalaydjieva, L., Kayser, M., Kivisild, T., Kravchenko, S.A., Krumina, A., Kučinskas, V., Lavinha, J., Livshits, L.A., Malaspina, P., Maria, S., McElreavey, K., Meitinger, T.A., Mikelsaar, A.V., Mitchell, R.J., Nafa, K., Nicholson, J., Nørby, S., Pandya, A., Parik, J., Patsalis, P.C., Pereira, L., Peterlin, B., Pielberg, G., Prata, M.J., Previderé, C., Roewer, L., Rootsi, S., Rubinsztein, D.C., Saillard, J., Santos, F.R., Stefanescu, G., Sykes, B.C., Tolun, A., Villems, R., Tyler-Smith, C. & Jobling,

M.A. (2000) Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. *Am J Hum Genet* **67**:1526–1543.

Rozen S, Skaletsky H, Marszalek JD, Minx PJ, Cordum HS, Waterston RH, Wilson RK, Page DC. (2003) Abundant gene conversion between arms of palindromes in human and ape Y chromosomes. *Nature* **423**(6942):873-6.

Sajantila A, Lukka M, Syvänen AC. (1999) Experimentally observed germline mutations at human micro- and minisatellite loci. *Eur J Hum Genet* **7**(2):263-6.

Schneider, S., Roessli, D. & Excoffier, L. (2000) *Arlequin ver. 2.000: A software for population genetics data analysis*. Genetics and Biometry Laboratory, University of Geneva, Switzerland.

Scholz M., Giddings I, Pusch CM. (1998) A polymerase chain reaction inhibitor of ancient hard and soft tissue DNA extracts is determined as human collagen type I. *Anal Biochem* **259**: 283-286.

Seielstad M.T., Minch E, Cavalli-Sforza LL. (1998) Genetic evidence for a higher female migration rate in humans. *Nat Genet* **20**(3):278-80.

Seielstad M, Bekele E, Ibrahim M, Touré A, Traoré M. (1999) A view of modern human origins from Y chromosome microsatellite variation. *Genome Res* **9**(6):558-67.

Semino, O., Passarino, G., Oefner, P.J., Lin, A.A., Arbuzova, S., Beckman, L.E., De Benedictis, G., Francalacci, P., Kouvatsi, A., Limborska, S., Marcikiae, M., Mika, A., Mika, B., Primorac, D., Santachiara-Benerecetti, A.S., Cavalli-Sforza, L.L. & Underhill, P.A. (2000a) The genetic legacy of Paleolithic *Homo sapiens sapiens* in extant Europeans: a Y chromosome perspective. *Science* **290**:1155–1159.

Semino, O., Passarino, G., Quintana-Murci, L., Liu, A., Beres, J., Czeizel, A. & Santachiara-Benerecetti, A.S. (2000b) MtDNA and Y chromosome polymorphisms in Hungary: inferences from the palaeolithic, neolithic and Uralic influences on the modern Hungarian gene pool. *Eur J Hum Genet* **8**:339–346.

Semino, O., Magri, C., Benuzzi, G., Lin, A.A., Al-Zahery, N., Battaglia, V., Maccioni, L., Triantaphyllidis, C., Shen, P., Oefner, P.J., Zhivotovsky, L.A., King, R., Torroni, A., Cavalli-Sforza, L.L., Underhill, P.A & Santachiara-Benerecetti, A.S. (2004) Origin, diffusion, and differentiation of Y-chromosome haplogroups E and J: inferences on the neolithization of Europe and later migratory events in the Mediterranean area. *Am J Hum Genet* **74**:1023-1034.

Sindbæk, S. M. (1999) A Magyar Occurrence. Process, practice and ethnicity between Europe and the Steppes. *ActaArch* **70**:149–164.

Szűj, E. (2005) Research on the prehistory of the Hungarians and Finno-Ugric studies. In: *Research on the prehistory of the Hungarians: A review*. (ed. B.G. Mende) *VAH* **18**:115–156. Hungarian Academy of Sciences Archaeological Institute.

Skaletsky, H., Kuroda-Kawaguchi T, Minx PJ, Cordum HS, Hillier L, Brown LG, Repping S, Pyntikova T, Ali J, Bieri T, Chinwalla A, Delehaunty A, Delehaunty K, Du H, Fewell G, Fulton L, Fulton R, Graves T, Hou SF, Latrielle P, Leonard S, Mardis E, Maupin R, McPherson J, Miner T, Nash W, Nguyen C, Ozersky P, Pepin K, Rock S, Rohlfing T, Scott K, Schultz B, Strong C, Tin-Wollam A, Yang SP, Waterston RH, Wilson RK, Rozen S, Page DC. (2003) The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature*. **423**(6942):825-37.

Sykes, B. (2002) *Éva hét leánya*. Európa Könyvkiadó.

Szőke, B. (1962) A honfoglaló és kora Árpád-kori magyarság régészeti emlékei. *RégTan* 1. Budapest: Akadémiai Kiadó.

Tambets, K., Rootsi, S., Kivisild, T. & Villems, R. (2001) The concepts of Richard Indreko about the origin of the Finno-Ugric speakers and the population genetics of the extant North-East European populations. *TRAMES* 5 (55/50), 1, 59-74.

Tambets, K., Rootsi, S., Kivisild, T., Help, H., Serk, P., Loogväli, E. L., Tolk, H.V., Reidla, M., Metspalu, E., Pliss, L., Balanovsky, O., Pshenichnov, A., Balanovska, E., Gubina, M., Zhadanov, S., Osipova, L., Damba, L., Voevoda, M., Kutuev, I., Bermisheva, M., Khusnutdinova, E., Gusar, V., Grechanina, E., Parik, J., Pennarun, E., Richard, C., Chaventre, A., Moisan, J.P., Barac, L., Peričić, M., Rudan, P., Terzić, R., Mikerezi, I., Krumina, A., Baumanis, V., Koziel, S., Rickards, O., De Stefano, G.F., Anagnou, N., Pappa, K.I., Michalodimitrakis, E., Ferák, V., Füredi, S., Komel, R., Beckman, L. & Villems, R. (2004) The western and eastern roots of the Saami--the story of genetic "outliers" told by mitochondrial DNA and Y chromosomes. *Am J Hum Genet* **74**:661–682.

Thomson, R., Pritchard JK, Shen P, Oefner PJ, Feldman MW. (2000) Recent common ancestry of human Y chromosomes: evidence from DNA sequence data. *Proc Natl Acad Sci U S A* **97**(13):7360-5.

Torroni A, Bandelt HJ, D'Urbano L, Lahermo P, Moral P, Sellitto D, Rengo C, Forster P, Savontaus ML, Bonn -Tamir B, Scozzari R. mtDNA analysis reveals a major late Paleolithic population expansion from southwestern to northeastern Europe. *Am J Hum Genet* **62**(5):1137-52.

T m ry, Gy., Cs nyi, B., Bog csi-Szab , E., Kalm r, T., Czibula,  ., Cs sz, A., Priskin, K., Mende, B., Lang , P., Downes, C.S. & Rask , I. (2007) Comparison of maternal lineage and biogeographic analysis of ancient and modern Hungarian populations. *Am J Phys Anthropol* **134**:354-368.

Tyler-Smith C., Oakey RJ, Larin Z, Fisher RB, Crocker M, Affara NA, Ferguson-Smith MA, Muenke M, Zuffardi O, Jobling MA. (1993) Localization of DNA sequences required for human centromere function through an analysis of rearranged Y chromosomes. *Nat Genet* **5**(4):368-75.

- Underhill, P.A., Jin, L., Lin, A.A., Mehdi, S.Q., Jenkins, T., Vollrath, D., Davis, R.W., Cavalli-Sforza, L.L. & Oefner, P.J. (1997) Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography. *Genome Res* **7**:996-1005.
- Underhill, P.A., Shen, P., Lin, A.A., Jin, L., Passarino, G., Yang, W.H., Kauffman, E., Bonne-Tamir, B., Bertranpetit, J., Francalacci, P., Ibrahim, M., Jenkins, T., Kidd, J.R., Mehdi, S.Q., Seielstad, M.T., Wells, R.S., Piazza, A., Davis, R.W., Feldman, M.W., Cavalli-Sforza, L.L. & Oefner, P.J. (2000) Y chromosome sequence variation and the history of human populations. *Nat Genet* **26**:358–361.
- Underhill, P.A., Passarino, G., Lin, A.A., Shen, P., Mirazon, Lahr, M., Foley R.A., Oefner, P.J. & Cavalli-Sforza, L.L. (2001) The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations. *Ann Hum Genet* **65**:43–62.
- Vasan S., Zhang X, Zhang X, Kapurniotu A, Bernhagen J, Teichberg S, Basgen J, Wagle D, Shih D, Terlecky I, Bucala R, Cerami A, Egan J, Ulrich P. (1996) An agent cleaving glucose-derived protein crosslinks in vitro and in vivo. *Nature* **382**(6588):275-8.
- Vigilant, L., Stoneking, M., Harpending, H., Hawkes, K., Wilson, A.C. (1991) African populations and the evolution of human mitochondrial DNA. *Science* **253**: 1503-7.
- Vogt P.H., Edelman A, Kirsch S, Henegariu O, Hirschmann P, Kiesewetter F, Köhn FM, Schill WB, Farah S, Ramos C, Hartmann M, Hartschuh W, Meschede D, Behre HM, Castel A, Nieschlag E, Weidner W, Gröne HJ, Jung A, Engel W, Haidl G. (1996) Human Y chromosome azoospermia factors (AZF) mapped to different subregions in Yq11. *Hum Mol Genet* **5**(7):933-43.
- Walsh, P.S., Metzger, D.A. & Higuchi, R. (1991) Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *Biotechniques* **10**:506-513.
- Weber, J.L. & Wong, C. (1993) Mutation of human short tandem repeats. *Hum Mol Genet* **2**(8):1123-8.
- Wells, R.S., Yuldasheva, N., Ruzibakiev, R., Underhill, P.A., Evseeva, I., Blue-Smith, J., Jin, L., Su, B., Pitchappan, R., Shanmugalakshmi, S., Balakrishnan, K., Read, M., Pearson, N.M., Zerjal, T., Webster, M.T., Zholoshvili, I., Jamarjashvili, E., Gambarov, S., Nikbin, B., Dostiev, A., Aknazarov, O., Zalloua, P., Tsoy, I., Kitaev, M., Mirrakhimov, M., Chariev, A. & Bodmer, W.F. (2001) The Eurasian heartland: a continental perspective on Y-chromosome diversity. *Proc Natl Acad Sci USA* **98**:10244–10249.
- Wells, S. (2003) *Az ember útja. Egy genetikai Odüsszeia*. Akkord Kiadó.
- Willerslev, E. & Cooper, E. (2005) Ancient DNA. *Proc Biol Sci* **272**(1558):3-16.



Wilson, M.R., Polansky D, Butler J, DiZinno JA, Replogle J, Budowle B. (1995) Extraction, PCR amplification and sequencing of mitochondrial DNA from human hair shafts. *Biotechniques* **18**(4):662-9.

Y Chromosome Consortium (2002) A nomenclature system for the tree of human Y-chromosomal binary haplogroups. *Genome Res* **12**:339–348.

Yen, P.H. (1998) A long-range restriction map of deletion interval 6 of the human Y chromosome: a region frequently deleted in azoospermic males. *Genomics* **54**(1):5-12.

Zerjal, T., Dashnyam, B., Pandya, A., Kayser, M., Roewer, L., Santos, F.R., Schiefenhover, W., Fretwell, N., Jobling, M.A., Harihara, S., Shimizu, K., Semjidmaa, D., Sajantila, A., Salo, P., Crawford, M.H., Ginter, E.K., Evgrafov, O.V. & Tyler-Smith, C. (1997) Genetic relationships of Asians and Northern Europeans, revealed by Y-chromosomal DNA analysis. *Am J Hum Genet* **60**:1174–1183.

Zerjal, T., Beckman, L., Beckman, G., Mikelsaar, A.V., Krumina, A., Kučinskas, V., Hurles, M.E. & Tyler-Smith, C. (2001) Geographical, linguistic, and cultural influences on genetic diversity: Y-chromosomal distribution in Northern European populations. *Mol Biol Evol* **18**:1077-1087.

Zhang L., Cui X, Schmitt K, Hubert R, Navidi W, Arnheim N. (1992) Whole genome amplification from a single cell: implications for genetic analysis. *Proc Natl Acad Sci U S A* **89**(13):5847-51.

## Összefoglalás

A magyar nyelv az uráli nyelvcsalád tagja, a finnugor nyelvek közé tartozó ugor nyelvek egyike.

Őseink körülbelül Krisztus előtt 500-ban váltak el a legközelebbi nyelvrokonaink, a vogulok és az osztjások őseitől, majd két évezrednyi sztyeppei vándorlást követően, a 9. század végén telepedtek le a Kárpát-medencében. Az irodalom szerint a korai magyarok a letelepedés idején finnugor nyelvet beszéltek, de török módra éltek.

A magyarság a nyelvi háttere alapján az uráli nyelvcsaládhoz köthető, azonban a történetkutatás mai elképzelése szerint a nép etnikai összetevői jelentősen megváltozhattak az elmúlt ezer év alatt a más törzsi és népi egységekkel történő keveredés következtében.

Az Y kromoszóma legfontosabb biológiai szerepe a nemiség meghatározása és a férfi fertilitás biztosítása. Sajátos jellemzői következtében, melyek egyedivé teszik a kromoszómák között, az Y kromoszóma hatékony eszköznek bizonyult a populációgenetikuskok számára a humán diverzitás tanulmányozásában és az apai ági leszármazási vonalak nyomon követésében. Az Y kromoszóma haploid, csak a férfiakban van jelen, apáról fiúra öröklődik és 95%-án nem történik rekombináció a meiózis során. A biallélikus markerek közé tartozó pontmutációk, illetve inszerció és delécións polimorfizmusok határozzák meg az Y-kromoszómális haplocsoportokat, melyek jellegzetes földrajzi eloszlást mutatnak. Egy minta haplocsoportba sorolása az Y kromoszómán derivált/mutáns allélikus állapotban jelenlévő biallélikus markerek szerint történik.

A Tat polimorfizmus egy olyan T>C pontmutáció az Y kromoszóma RBF5 lókusán, mely különös érdeklődésre tart számot a finnugor nyelvű populációk vizsgálatakor. A Tat C allél az N3 haplocsoportot határozza meg, amely széles körben elterjedt Európa északi és keleti területein, ugyanakkor gyakorlatilag nem fordul elő a kontinens nyugati és déli részein élő populációkban. Ázsiában az északi területeken élő népcsoportokban fordul elő számottevő gyakorisággal. Az irodalom alapján a Tat C allél minden eddig vizsgált uráli nyelvű populációban megtalálható, ugyanakkor nem jellemzi a modern magyar népcsoportot. Ezen megfigyelések alapján felvetődik a kérdés, hogy vajon a honfoglalás kori populációban jelen volt-e ez a mutáció vagy sem.

Ennek a kérdésnek a tisztázásához (IX-)X. századi mintákon vizsgáltuk a Tat marker allélikus állapotát. A vizsgált leletek 6 különböző, a Kárpát-medence területén fekvő lelőhelyről származnak.

A sikeres és szennyeződéstől mentes mitokondriális DNS amplifikációt követően próbáltunk meg Y-kromoszómális DNS-t kinyerni a csontmintákból. A specifikus Y-kromoszómális analízis előtt minden esetben szükség volt a megmaradt teljes töredékes genom random primerekkel történő dúsítására. A rendelkezésünkre álló archaikus DNS mennyiségének felszaporításához a PEP (**P**rimers **E**xtension **P**reamplification) módszernek egy tovább fejlesztett változatát (**m**odified **i**mproved PEP - miPEP) alkalmaztuk. Az utóbbi módszer 2 polimeráz enzim elegyét használja, melyek közül az egyik 3'-5' exonukleáz (más néven proofreading) aktivitással rendelkezik. Ily módon lehetővé teszi, hogy nagyobb hűséggel történjen a templát DNS átírása.

A Tat polimorfizmus vizsgálatokor egy 112 bp-os szakaszt amplifikálunk, majd ezt követően restrikciós emésztéssel (HpyCH4 IV enzim) határozzuk meg az allélikus állapotot. A 7 sikeresen tipizált (9-) 10. századi csontminta közül 2 esetben igazoltuk a mutáns C allél jelenlétét (13/1, B1/3c minták), míg 5 leletnél a vad típusú T allél előfordulását. A Tat mutációt hordozó 2 X. századi minta egy régióból, Békés megyéből származik. A 13/1 minta a szabadkígyós-pálligeti, a B1/3c lelet az örménykúti temető leletanyagát képezi. Mindkét lelet mellett klasszikus, gazdag mellékletanyagot tártak fel az ásatás során. A T2/41 csontminta esetében kapott eredményt, mely szerint a vad típusú Tat T allél van jelen a mintában, sikerült megerősítenünk az M9 marker vad típusú C alléljának tipizálásával.

A tény, hogy 7 csontminta közül kettő is hordozza a Tat C allélt, több mint érdekes, tekintve, hogy a 197 modern minta közül csak egyetlen székely férfi Y kromoszómáján detektáltuk az allél jelenlétét. A Tat mutáció ritka előfordulása az általunk tanulmányozott modern magyar és székely populációban összhangban van a korábbi vizsgálatok eredményeivel és további adatot szolgáltat ahhoz a felismeréshez, miszerint a Tat C allél gyakorlatilag hiányzik a modern magyarul beszélő populációkból. A Tat mutáció által meghatározott N3 leszármazási vonal előfordulása a székely csoportban talán valamilyen múltbeli esemény genetikai hatását tükrözi.

Természetesen néhány vizsgált minta és egy marker alapján nem vonhatunk le messzemenő következtetéseket, mégis elmondható, hogy a Tat mutáció jelenléte az ősi

magyar populációban utalhat keveredésre vagy közös apai ági felmenőre más uráli nyelvű populációkkal.

Az eredményeink szerint a Tat C allél, amely széles körben elterjedt az uráli nyelvű populációkban, jelen volt az ősi magyar populáció génkészletében a Kárpát-medence területén történő letelepedéskor, illetve a Tat polimorfizmus egy megfelelő marker az ősi magyar populáció Y-kromoszómális genetikai összetételének analíziséhez.

A C allél jelenléte egyben alátámasztja eredményeink hitelességét is, hiszen sem a Régészeti Intézet munkatársai (régész, antropológus), sem csoportunk férfi tagjai, akik kapcsolatba kerültek a csontmintákkal, nem hordozzák ezt a mutációt.

Annak érdekében, hogy a jövőben további Y-kromoszómális markereket tudjunk vizsgálni az archaikus mintákon, minden marker (összesen további 21) analíziséhez, melyet a modern mintákon tanulmányoztunk, olyan primerpárokat terveztünk, vagy választottunk az irodalomból, melyek rövid (170 bp alatti) PCR terméket adnak. A markerek amplifikálásához szükséges PCR körülményeket a modern mintákon optimalizáltuk, így tulajdonképpen minden szükséges lépést megtettünk azért, hogy ősi DNS-en is vizsgáljuk azokat.

A ma élő magyar és székely populációban, összesen 197 egyénnél, 22 (M96, M89, M9, M45, M35, M78, M170, M253, P37, M26, M201, P15, M304, M267, M172, M102, M67, M92, Tat, M173, M17, M269) Y-kromoszómális biallélikus markert vizsgáltunk meg az Y-kromoszómális filogenetikai fán elfoglalt hierarchikus szerveződésük szerint (Y Chromosome Consortium 2002, Jobling és Tyler-Smith, 2003). A vizsgálatba bevont magyar férfiak (100 fő) Magyarország különböző területein születtek, de a leginkább reprezentált régió az Alföld területe volt (90/100). A székely minták (97 fő) Korondról, Erdély területéről származnak.

Minden minta esetében először az M96, az M89, az M9 és az M45 SNP markereket analizáltuk. A markereket multiplex reakcióban vizsgáltuk, tipizálásukhoz a SNaPshot technikát használtuk. A 4 SNP marker vizsgálatával minden egyes mintát el tudtunk helyezni a filogenetikai fa valamelyik nagyobb ágán. Ily módon, a markerek hierarchikus szerveződésének megfelelően, tudtuk azt, hogy az egyes mintáknál milyen további markereket kell megvizsgáljunk az egyértelmű haplocsoportba sorolás érdekében. Összesen további, a filogenetikai fa megfelelő részén elhelyezkedő, 18 marker (M35, M78, M170, M253, P37, M26, M201, P15, M304, M267, M172, M102,

M67, M92, Tat, M173, M17, M269) allélikus állapotát határoztuk meg PCR-RFLP vizsgálattal.

Az M96, az M89, az M9 és az M45 SNP markerek multiplex vizsgálatát részben a T2/41-es jelű csontmintán is sikerült elvégeznünk. Az analízis során az M45 és az M9 markert hordozó, 2 rövidebb (a specifikus fragmentumok hossza: 106 bp és 121 bp) DNS fragmentumot tudtuk amplifikálni. A kapilláris rendszeren történő genotípus meghatározás szerint mind az M45, mind az M9 marker esetében egy-egy C allél van jelen, mely a markerek vad típusát jelöli. Ily módon, a markerek hierarchikus szerveződése miatt, a csontmintában nem lehet jelen a Tat mutáció, hiszen akkor az M9-es marker mutáns/derivált G allélját kellett volna kimutatnunk. Ez az eredmény összhangban van a Tat marker PCR-RFLP vizsgálata során kapott eredményünkkel.

A vizsgált markerek segítségével a modern magyar és székely mintákat az E, F\*, G, I, J, K\*, N3, P\* és az R1 Y-kromoszómális haplocsoportokba soroltuk be.

A megfigyelt haplocsoportok általánosan jellemzik a különböző európai populációkat, egyetlen kivételt képez ez alól a székely populációban felbukkanó (3,1%) P-M45\*(xM173) klaszter. A közép-ázsiai P-M45\*(xM173) haplocsoport jelenléte a székely mintában szokatlan, mivel ez a klaszter szinte egyáltalán nem fordul elő Európában. Elképzelhető, hogy a haplocsoport jelenléte némi ázsiai genetikai hatást tükröz.

A jelen vizsgálatban az irodalmi adatokhoz képest a J haplocsoport meglepően gyakori előfordulását tapasztaltuk a magyar (16%) és a székely (21,6%) populációban.

A J haplocsoport nagyobb gyakorisága az általunk vizsgált populációkban egyrészt utalhat anatóliai, illetve dél-balkáni genetikai hatásra, ugyanakkor, főként a székely népcsoport esetében, genetikai sodródás következménye is lehet, hiszen a székelyek zárt közössége politikai okok miatt sokáig fennmaradt. Más részről különböző történelmi események is hozzájárulhattak a J haplocsoport magasabb gyakoriságához. Ilyen esemény lehetett például az Ottomán Birodalom terjeszkedése, illetve ennek hatására a balkáni régióból Magyarország területére menekülő népek letelepedése.

A történelmi adatok és az ősi magyar populáció anyai leszármazási vonalainak összehasonlító vizsgálata alapján elképzelhető, hogy a korai magyar populációnak is szerepe lehetett a J haplocsoport Kárpát-medencében történő elterjedésében.

A modern minták vizsgálatával kapott eredményeinket statisztikai alapon is kiértékeltek. Az E3b-M35, G-M201, I-M170, J1-M267, J2-M172, N3-Tat, R1a1-M17

és az R1b3-M269 haplocsoportok gyakorisági értékei alapján az Arlequin 2.000 program felhasználásával meghatároztuk az általunk vizsgált magyar és székely populáció genetikai diverzitásának mértékét, valamint páronkénti genetikai távolságot számoltunk az általunk vizsgált magyar és székely populáció, illetve a Semino és munkatársai által tipizált (2000a) további 20 európai populáció között. A populációk közötti távolságviszonyokat többdimenziós skálázási (**Multidimensional Scaling - MDS**) eljárással 2 dimenziós térben ábrázoltuk.

A statisztikai vizsgálat szerint az általunk vizsgált magyar és székely populáció genetikailag heterogén összetételű, közeli rokonságban áll egymással, illetve más közép-európai és balkáni populációkkal. Ugyanakkor a 2. dimenzió egyértelműen elkülöníti egymástól az általunk analizált és a Semino és munkatársai által tipizált magyar csoportot. Az utóbbi közel térképeződik a lengyel és az ukrán populációhoz. Egy lehetséges magyarázat lehet erre az eltérésre, hogy Semino és munkatársai elsősorban észak-magyarországi mintákat vizsgáltak, ahol a 19. században jelentős volt a szomszédos északi területek más anyanyelvű népességeivel való diffúzió.

Eredményeink alapján kimondható, hogy Semino és munkacsoportjának adatai (2000a) nem reprezentatívak a jelenkori magyar populáció vonatkozásában.

## Summary

The Hungarian population belongs linguistically to the Finno-Ugric branch of the Uralic family.

The Hungarian nation traces its history to the early Magyars (ancient Hungarians), who settled in the Carpathian Basin at the end of the 9<sup>th</sup> century after two millenia of migration from the steppe zone. At the time of settlement the ancient Hungarians were a people who spoke a Finno-Ugric language, but had a Turkic way of life. Our ancestors appear to have separated from their closest linguistic relatives, the Voguls and the Ostyaks who now live in the forested areas beside the river Ob in western Siberia, around 500 BC.

Though Hungarians linguistically belong to the Finno-Ugric language family, the ethnic components of the ancient Hungarians' population, according to recent hypothesis, could have changed significantly during the last 1000 years due to admixture with other tribes and peoples.

The main biological importance of Y chromosome is its role in sex determination and male fertility. The singular characteristics of the Y chromosome, which include haploid state, male specificity, paternal inheritance and absence of recombination through most of its length (95%), make this chromosome a powerful tool for tracing and comparing paternal lineages and studying genetic diversity of human populations. Biallelic markers include point mutations, insertions and deletions. These binary polymorphisms define the Y-chromosomal haplogroups, which show geographical clustering. Haplogroup assignment of a sample is based on derived states of biallelic markers.

One specific Y-chromosomal base substitution, the T→C point mutation in the RBF5 locus, called the Tat C allele, is of interest in the Finno-Ugric context. It defines the Y chromosome haplogroup N3, which is present in northern and eastern Europe, but is virtually absent in the west and south. It is also frequent in northern Asia. The C allele of the Tat polymorphism is widespread in all Uralic-speaking populations studied so far, except that it is absent or extremely rare among modern Hungarian-speaking populations. These data raise the question of whether the ancient Hungarians, who spoke a Uralic language, possessed this polymorphism or not. To answer this question we attempted to screen for the Tat polymorphism in ancient DNA from skeletal remains

from the (IX)-X<sup>th</sup> century. The samples were derived from 6 different well-documented archaeological excavations from the Carpathian Basin.

After successful and contamination-free mitochondrial DNA amplification from the archaeological samples, Y-chromosomal DNA recovery was attempted. In order to enhance the efficiency random primers and the modified improved PEP (**P**rimers **E**xtension **P**reamplification) method were used. In this way we tried to increase the effective number of starting templates. This whole genome amplification method utilizes a mixture of two polymerases among which one possesses a 3'-5' exonuclease or "proofreading" activity, allowing for a more accurate amplification of the genomic sequence.

Screening for Tat polymorphism was carried out by amplifying a 112 bp fragment and a subsequent digestion with HpyCH4 IV restriction enzyme. Out of the 7 successfully typed 9-10<sup>th</sup> century skeletal remains two possess (samples 13/1, B1/3c) the mutation, while the others carry the ancestral T allele. The samples 13/1 and B1/3c were from two burial sites (13/1-Szabadkígyós-Pálliget, B1/3c-Örménykút) in the same county (Békés), they were excavated with rich grave goods, unambiguously typical of burial practices of Hungarian conquerors. In case of the sample T2/41 the presence of the ancestral Tat T allele was confirmed by typing the ancestral state (C allele) of the marker M9 (C→G). The fact that two of seven ancient samples possessed the Tat C allele, is more than intriguing, considering that from the 197 modern Hungarian-speaking males only one had this polymorphism. This latter finding is consistent with previous studies and gives further evidence for the fact that this polymorphism is practically absent in modern Hungarian-speaking populations. The single observation of this particular lineage in the Szekler group may reflect some past contribution.

On the other hand, the low number of ancient samples has to be taken into consideration. However the occurrence of this genetic variation in ancient Magyars could be the result of admixture or sharing a common male ancestor with other Uralic-speaking populations.

Our data suggest that the Tat C allele, which is widespread in Uralic-speaking populations, was present in the ancient Magyar population when they crossed the Carpathians and settled in the Carpathian Basin and the Tat polymorphism is an appropriate marker to investigate the ancient Hungarian gene pool.



No one from the Archaeological Institute (excavator, anthropologist) who has handled the bones and nobody from our group has the Tat C mutation, which strongly supports the authenticity of the ancient DNA.

In order to additional Y-chromosomal markers will be examined on ancient remains in future studies the primer pairs for all studied markers in case of recent samples were redesigned or chosen from literature, resulting in shorter PCR products (below 170 bp), to serve the special needs of amplification of ancient DNA. We optimized and checked the PCR conditions of these primer pairs, so they are ready to use on ancient DNA.

In case of the modern Hungarian and Szekler populations (in 197 individuals) twenty-two biallelic markers (M96, M89, M9, M45, M35, M78, M170, M253, P37, M26, M201, P15, M304, M267, M172, M102, M67, M92, Tat, M173, M17, M269) were examined in hierarchical order in agreement with the Y-chromosome phylogeny (Y Chromosome Consortium 2002, Jobling & Tyler-Smith 2003). All the Hungarian males selected for the analysis had a birthplace in different parts of Hungary, but the most represented area (90 samples) was the Great Hungarian Plain. The Szekler sample consists of 97 Szekler volunteer donors born and living in Corund, Transylvania, Romania.

All samples were surveyed for the M96, M89, M9 and M45 markers. These markers were typed by multiplex PCR followed by SNaPshot technique. Additional genotyping of samples was restricted to markers on the appropriate branch of the YCC haplogroup tree (Jobling & Tyler-Smith 2003). Markers M35, M78, M170, M253, P37, M26, M201, P15, M304, M267, M172, M102, M67, M92, Tat, M173, M17 and M269 were analysed by the PCR-RFLP assay.

In case of the ancient sample T2/41 out of four Y-chromosomal DNA fragments, containing the four basal markers (M96, M89, M9 and M45), the two shorter fragments (106 bp (marker M45) and 121 bp (marker M9) long specific PCR products) were successfully amplified. In case of these markers the ancestral C allele was typed. The ancestral state at marker M9 indicates that this sample could not carry the derived Tat C allele in accordance with Y-chromosome phylogeny (YCC 2002, Jobling & Tyler-Smith 2003). This finding confirms the result of the PCR-RFLP analysis of this sample.

The markers used allowed the classification of the modern samples into the E, F\*, G, I, J, K\*, N3, P\* and R1 haplogroups.

The modern Hungarian and Szekler populations share similar components described for other Europeans, except for P-M45\*(xM173) in Szekler samples. The presence of central-Asian haplogroup P-M45\*(xM173) in Szeklers is unusual for a European population, since it is almost absent in continental Europe and presumably reflects some Asian contribution.

In the present study, haplogroup J was unexpectedly common in the Hungarian-speaking populations (Hungarians: 16%, Szeklers: 21.6%).

The elevated frequency of haplogroup J may reflect Anatolian and southern Balkan contributions to the gene pools of Hungarians and Szeklers, however in the Szekler samples it might be partially attributed to genetic drift, since they lived in relative isolation for a long time. On the other hand, the elevated frequency of J in both groups could also be due to a range of historical events.

One is the expansion of the Ottoman Empire from the 16<sup>th</sup> century AD; refugees from the Balkan area fled to Hungarian territory. Historical data and the comparative analyses of maternal lineages of ancient Hungarian population suggest that the earlier migrations of the Magyars may also have contributed to the presence of this lineage in the Carpathian Basin.

Statistical analyses of Y chromosome data of two modern Hungarian-speaking populations were also performed. Based on haplogroup frequencies population genetic diversity and pairwise  $F_{st}$  genetic distances (between the examined Hungarian and Szekler groups and other European populations studied by Semino *et al.* (2000a)) were calculated by using the Arlequin package, version 2.000. Pairwise values of  $F_{st}$  were visualized in two dimensions by the use of multidimensional scaling (MDS) analysis.

According to statistical analyses the Hungarian and Szekler populations are genetically heterogeneous, closely related, and close also to other populations from Central Europe and the Balkans. However, the second dimension clearly separates our samples from those Hungarians studied by Semino *et al.* (2000a). The latter group has closer genetic relationships with Polish and Ukrainian populations. A possible explanation might be that Hungarian samples analyzed by Semino *et al.* (2000a) primarily originated from subjects from northern regions in Hungary, where in the 19<sup>th</sup> century extensive admixture occurred with inhabitants of neighbouring northern regions.

According to our results the Y chromosome data of Semino *et al.* (2000a) are not representing the modern Hungarian population.