PAMJAV HOROLMA – FEHÉR TIBOR
NÉMETH ENDRE – CSÁJI LÁSZLÓ KOPPÁNY

A populáció- és filogenetikai adatok őstörténeti értelmezésének lehetőségei és korlátai

A történeti kutatásokat régóta segítik a természettudományos vizsgálatok. Különösen nagy szükség van rájuk az írott forrásokban szűkölködő őstörténet esetében. A fizikai antropológia például a 19. század vége óta hathatósan támogatja a régészetet, napjainkban pedig már a természettudományok robbanásszerű fejlődésével számtalan új módszer használható fel. A genetika fejlődése is egyre újabb vizsgálati lehetőségeket kínál, s ami tegnap korszerű volt, mára szinte elavult. A természettudományos-technikai fejlődéssel párhuzamosan azonban szaporodnak a kérdések is. Kitől is kellene mintát venni? És mit is jelent a kinyert eredmény? Úgy tűnik, rögös út vezet még addig, amíg a genetikai eredmények tényleges történelmi felhasználásának módszere letisztul.

A DNS-alapú filogenetika fiatal tudományág. James D. Watson és Francis Crick 1953-ban fedezte fel a DNS szerkezetét. Kary Mullis 1983-ban dolgozta ki az úgynevezett polimeráz láncreakció technikát, amelynek segítségével mesterségesen másolható és sokszorosítható a DNS laboratóriumi körülmények között. E felfedezések jelzik azokat a főbb állomásokat, amelyek lehetővé tették, hogy a kutatók vizsgálják a mitokondriális DNS és az Y-kromoszóma szerkezetét, és a szerkezetükben bekövetkező változások alapján elinduljon a leginkább meghatározó emberi migrációk feltérképezése. Ez azért volt lehetséges, mert a mitokondriális DNS tisztán anyai vonalon, az Y-kromoszóma tisztán apai vonalon sok generáción keresztül öröklődik.

Genetika és történelem

A populáció- és filogenetika egyik alapvető tanulsága, hogy az emberiség egy családot alkot, tehát „mindenki rokona mindenkinek", ezért nem használja a genetikai „rokonság" fogalmát. Ehelyett a kapcsolatok és átfedések különböző modelljei segítségével a genetikai „távolság" fogalmát alkalmazza. Továbbá a túlságosan is érzelmi alapokon nyugvó nép/nemzet/etnikai csoport helyett a populáció (népesség) kategóriáját dolgozta ki, amely kísérlet egy átláthatóbb, neutrálisabb kategorizálás felé, és távolodás a kortárs közvéleményt általában érdeklő etnikai tartalomtól.

Míg egy nyelvi közösség régészeti hagyatéka, tárgyi kultúrája néhány nemzedék alatt akár gyökeresen megváltozhat, az adott populációt alkotó, egyenes ágon levezetett apai és anyai „leszármazási vonalak" nagy biztonsággal visszakövethetők a távoli múltba. Ez azonban nem feltétlenül jelent egy irányba mutató migrációs modelleket haplocsoportonként. Természetesen a folyamatos keveredés hatására egy nyelvi közösség apai és anyai vonalainak összetétele (a jellemző haplocsoportok egymáshoz viszonyított aránya) is jelentősen átalakulhat, de a különböző anyai és apai vonalú szálak nagy pontossággal meghatározhatók maradnak.

Ma már a történeti olvasatokat adó szakemberek előszeretettel nyúlnak a filogenetika eszköztárához, mert az évezredek óta tartó folyamatos keveredések, a házasodási szokások nemritkán elfednek olyan szempontokat a vizsgálódások elől, amelyek elemzéséhez a genetika fontos adatokat szolgáltathat. Ki gondolná például a külsődleges embertani jegyek alapján, hogy a mai finn népesség apai vonalainak több mint 60%-a közép-szibériai eredetűnek tartott haplocsoportokba tartozik? Ez az észrevétel már átvezet szűkebb kutatási területünkhöz, az ugor apai és anyai vonalak elemzéséhez, illetve az uráli nyelveket beszélő népességek eredetének genetikai alapokon nyugvó vizsgálatához.

A honfoglalást megelőző, ún. korai magyar történelem kutatása meglehetős információ- és adathiánnyal küszködik. Ezért a különböző tudományterületek kutatói nagy reményeket fűztek a magyar népesség anyai és apai leszármazási vonalainak filogenetikai feltérképezéséhez. Hamar kiderült azonban, hogy a mai magyar népesség anyai és apai vonalainak átfogó képe

A mintavétel problémái

A genetikai mintavételek általában néhány tucat személyt és legfeljebb néhány helyszínt érintenek, s a szervezők a saját maguk által alkotott kategóriákkal jelölik meg a mintáikat. Az így kapott mintákat vagy a helységnévvel, vagy a térség nevével, esetleg az etnikai névvel jelölik meg adatbázisukban. Ez a többi adatbázis adataival való összehasonlítást megnehezíti, hiszen szinte lehetetlen megtalálni a nemzetközi adatbázisok közös nevezőjét: kis- vagy nagytérségi, etnikai, földrajzi vagy nemzeti keretekben (fogalmakkal) történik-e a minták „címkézése".

További nehézség, hogy a genetikai laboratóriumok gyors technológiai fejlődésének köszönhetően egyre részletgazdagabban látunk bele az adott DNS-lánc szakaszaiba, egyre több szegmens által jutunk információhoz. Emiatt a korábbi adatok sokszor nehezen értelmezhetők a később vizsgált, részletesebb (nagyobb felbontású) minták elemzése során. Egy adott populáció ráadásul nem feltétlenül reprezentálható néhány tucat vagy néhány száz mintával, és nagy nehézséget jelent, hogy milyen szempontok alapján történjen a reprezentativitásra törekvő mintavétel.

További nehézség, hogy a minták címkézése a kettős és többes identitás problematikájával nemigen tud mit kezdeni. Egyelőre a bevett eljárás az, hogy aki egy adott etnikai csoport tagjának vagy egy területi egység lakójának vallja magát, azt lehet az adott csoport genetikai feltérképezése érdekében mintavételre felkérni, és a genetikai mintára az általa megadott etnikai/földrajzi/nemzeti önazonosságra utaló jelölést alkalmazzák csupán. Ez a gyakorlat azonban a háttéradatok hiányában a minták összehasonlítását megnehezíti.

nem különbözik jelentősen más középeurópai népekétől. Egyre többen kérdőjelezték meg azt is, hogy a populációk nyelvi és genetikai viszonyai között egyáltalán összefüggés lehet.

A részletesebb vizsgálatok a nyelvek terjedésének, a beszélt nyelvek társadalmi és kulturális hátterének sokféle változási folyamatát, modelljét tárták fel. Bizonyos esetekben valóban igen nehezen mutatható ki valamiféle genetikai kapcsolat egy adott nyelvcsaládhoz tartozó beszélők között. Ilyenek például a török nyelveket beszélő népek. A török nyelvek expanziójának hátterében a török nyelvek társadalmilag magas presztízse (is) állhatott, amely nagyfokú nyelvi asszimilációt eredményezett.

A genetikai adatok tehát újabb kérdéseket vetnek fel, amelyek megválaszolására a társadalomtudományok és a természettudományok képviselőinek párbeszédére, közös gondolkodására van szükség. A Pamjav Horolma által 2012-ben szervezett interdiszciplináris kutatócsoportunk tagjai között populáció- és igazságügyi genetikus szakember (Pamjav Horolma) mellett a modellezés módszereit ismerő matematikus (Németh Endre), a filogenetikai adatbázisok nemzetközi kapcsolatrendszerét jól ismerő szakember (Fehér Tibor), továbbá az etnicitás kérdésével és a közösségkonstrukciós folyamatokkal foglalkozó szociálantropológus (Csáji László Koppány) is helyet kapott.

A magyarság genetikai képe

Van azonban olyan eset, amikor többékevésbé egyértelműen beazonosítható egy vagy több genetikai szegmens, amely egyértelműen összeköti az egy nyelvcsaládhoz tartozó népességeket (populációkat), annak ellenére, hogy emellett más haplocsoportok és -típusok sokkal szélesebb körű kapcsolatrendszerre is utalhatnak. Érdemes megvizsgálni, hogy az azonosított „közös" genetikai haplocsoportoknak lehet-e őstörténeti értelmezést elősegítő jelentőségük. Például a ma valamely finnugor nyelvet beszélőket az N-Taton belül az N-L708, továbbá az R1a-Z280 alcsoport mindenképpen összekapcsolja filogenetikai szempontból. Az N-L708 közép-szibériai, az R1a-Z280 a szakirodalom szerint kelet-európai eredetű. Az eddigi genetikai adatok összhangban vannak azokkal a társadalom- és nyelvtörténeti modellekkel, amelyek a korai uráli nyelvű közösségeket az uráli embertani (fizikai antropológiai) típussal hozták összefüggésbe, amely a Szaján-hegység és az Urál közötti bonyolult embertani átmenetek gyűjtőelnevezése.

A finnugor (nyelvű) népekben egyértelműen kimutatható közép-szibériai komponens azonban más megvilágításba helyezheti az uráli és az altáji

Fogalomtár

populáció: a populáció egy adott időben a vizsgálat szempontjából azonos csoportba tartozónak tekintett emberek jól körülhatárolható köre. Az elnevezéseket és csoportképző szempontokat földrajzi, etnikai, nyelvi vagy más kategóriák szerint az adott kutatás határozza meg, azonban emiatt a különböző kutatásokban vizsgált populációk gyakran nehezen összehasonlíthatók (inkompatibilitás). A legtágabb kör maga az emberi populáció (emberiség), amelynek egyedei térben és időben együtt élnek.

populációgenetika: a vizsgált populációkban – földrajzi vagy más szempontból csoportot alkotó – genetikai markerek megoszlásának, sokféleségének és az azokat befolyásoló tényezőknek a tanulmányozása, valamint a kapott eredmények matematikai-statisztikai elemzése (az összehasonlításban elméletileg bármilyen genetikai marker vizsgálható).

filogenetika: a vizsgált populációk anyai és apai leszármazási vonalainak, ezek sokféleségének, összetételének feltérképezése és az azok közötti kapcsolatok rekonstrukciója, egymással való összehasonlítása egy filogenetikai fastruktúra (leszármazási ágrajz) formájában. A vizsgált genetikai markerek inkább a hosszú időtávú genetikai kapcsolatok feltérképezésére adnak lehetőséget.

haplocsoport: a haplocsoportallélok (single nucleotide polymorphism: SNP-k) kombinációja egy bizonyos kromoszóma régióján belül, amelyek szorosan kapcsolódnak és együttesen öröklődnek a következő gene-

nyelvek terjedésére vonatkozó ismereteinket, valamint az uráli és kelet-szibériai jukagirok közötti nyelvi kapcsolatok kutatását. Ezek az adatok nem átírják az eddigi elképzeléseket, hanem olyan új adatokat adnak a kutatásokhoz, amelyeket a továbbiakban figyelembe lehet venni érvként vagy ellenérvként a tudományos vitában, illetve egyes hipotéziseket megerősíthetnek, másokat pedig gyengíthetnek. A genetikai adatok nem a „megoldást" nyújtják, hanem új szempontokat adhatnak az eddigi nyelvészeti, régészeti, történettudományi, etnológiai adatok elemzéséhez.

Nem ritka, hogy nyelvileg egységes populációkban jelentős regionális különbségek mutathatók ki haplocsoporteloszlás tekintetében. Jelenleg nincs regionális bontású magyar filogenetikai adatbázis, így távol vagyunk attól, hogy pontos képet alkossunk a mai magyar népességről. Mindenesetre a rendelkezésre álló adatok alapján úgy tűnik, hogy a mai magyar apai és anyai ági génállomány nem tér el jelentősen a szomszédos kelet-közép-európai népekétől, néhány haplocsoport jelenléte és aránya mégis elgondolkodtató. A három leggyakoribb haplocsoport a magyar népességben az R1a-M198 (28,5%), az R1b-M412 (15,6%) és az I2-CTS10228 (13,2%), amelyek teljesen hiányoztak a 6–8 ezer évvel ezelőtti Kárpát-medencei mintákból, ami nem támasztja alá azokat az elképzeléseket, hogy a mai magyar népesség genetikai értelemben őshonos volna a Kárpát-medencében.

Az európai népek többségében is előforduló haplocsoportok részaránya meghaladja a 90%-ot a mai magyar nyelvű népességben (ezek közül ugyanakkor számos csoport nem korlátozódik Európára, hanem inkább a „nyugat-eurázsiai" jelzővel lehetne illetni). A magyarság filogenetikai szempontból hidat képez a nyugat-európai (germán és kelta), a kelet-európai (balto-szláv, Volga-vidéki) és a balkáni genetikai jellemzők között, amihez az északkelet-ázsiai és dél-szibériai csoportok mérhető jelenléte társul (különösen a székelyekben). Leegyszerűsítve ez azt jelenti, hogy a magyarokban több a mai germán és kelta utódnépek által lakott területeken gyakori haplocsoport (R1b-P312, R1b-U106, I1-M253), mint a környező szláv népek körében, de több a szlávokban előforduló csoport, mint az osztrákoknál (R1a-M458, R1a-Z280, I2-CTS10228). Mindehhez adódik néhány százaléknyi uráli és belső-ázsiai népcsoportokra jellemző haplocsoport is (N-Tat, Q-M242, R1a-Z93, C-M217).

Ez az összetétel olyan nagymértékű haplocsoport-változatosságot eredményez, amely Európa egyik legsokszínűbb népességévé teszi a mai magyar populációt ebből a szempontból. A jelenség hátterében nem annyira etnospecifikus, inkább a régióra jellemző demográfiai folyamatok állhatnak. Megjegyezzük, hogy a mai románokra és szerbekre is hasonló haplocsoport-változatosság jellemző. Ennél is nagyobb genetikai változatosságot mutatnak az Ázsia középső és délnyugati részén elhelyezkedő népek, például a törökök, a perzsák, a tádzsikok vagy az üzbégek.

A filogenetikai adatok alapján egyértelmű, hogy a nyelvileg viszonylagos egységet képező magyarság genetikai történetét tekintve többgyökerű, nagyon sokszálú múlttal rendelkezik. A genetikai változatossághoz, különböző haplocsoportok integrálásához vezető folyamatok (asszimiláció, elitdominancia, lingua franca modell stb.) feltárása a jövő feladata, és csak a történeti és társadalomtudományokkal közösen történhet.

Manysi minták

Kutatócsoportunk első, az adatgyűjtéstől a kiértékelésig tartó esettanulmányát a nyelvészek által széles körben elfogadott finnugor/uráli nyelvrokonság egyik alaptétele indukálta: eszerint a magyar nyelv legközelebbi nyelvrokona a manysi és a hanti nyelv. Ezért elsőként a mai magyar génállomány nagy valószínűséggel ún. ugor kori elemeit kívántuk beazonosítani és arányosítani a teljes népesség megoszlásához. Ennek érdekében a nemzetközi kutatásokból (főleg finn és orosz adatgyűjtésből) származó mintákat a manysik között végzett új adatgyűjtéssel egészítettük ki, és ezek eredményét összegeztük.

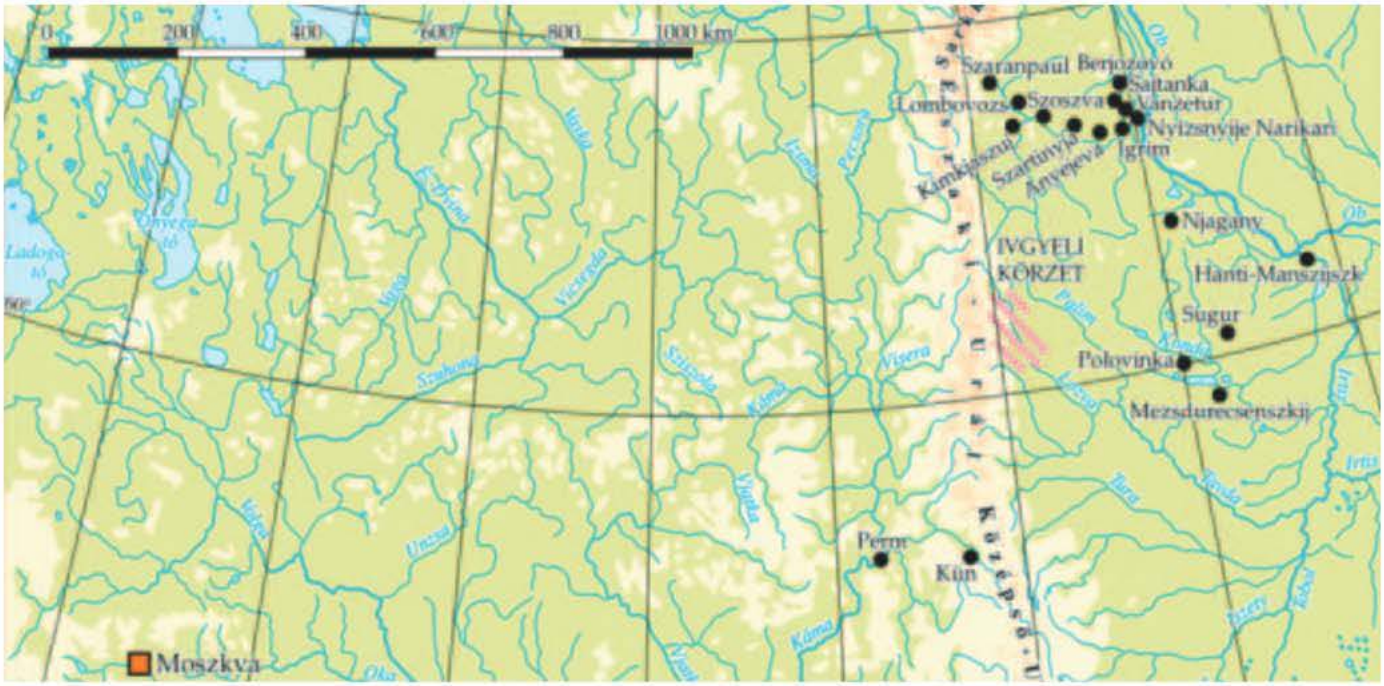
Elsőként értelmeznünk kellett azt a jelenséget, hogy a magyar az egyetlen finnugor nyelvű nép, amelyben az N-Tat haplocsoport aránya viszonylag alacsony. Kutatócsoportunk ezért részletesebb elemzést kívánt végezni a mai magyarok és obi-ugorok közötti apai és anyai vonalú kapcsolatok terén. Két mintavételi körutazást szerveztünk Szibériába, amelynek során az általunk felkért kutató 96 manysi családtól gyűjtött genetikai mintát, a mintákat adó személyek társadalmi kapcsolataira, családi hátterére vonatkozó adatokkal is kiegészítve.

rációkba. A haplocsoport-meghatározással – a DNS egy bizonyos pontján a mutációik jelenléte (SNP) alapján – az emberek egy-egy közös ősre (egy adott genetikai leszármazási vonal) vezethetők vissza a molekuláris evolúció tekintetében. Mivel egy haplocsoportba a hasonló haplotípussal rendelkező személyek tartoznak, ezért általában a haplocsoportot a meglévő haplotípusok alapján lehet vizsgálni, változásait megbecsülni.

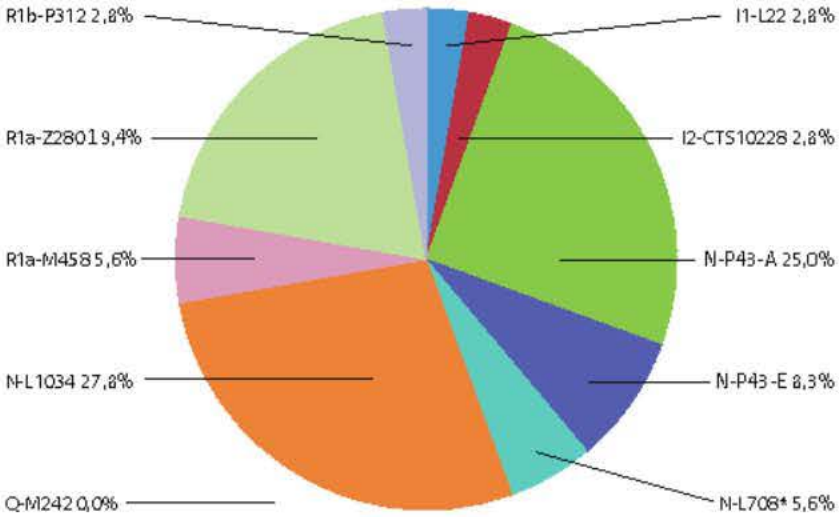
migrációs marker: a haplocsoportok keletkezése általában egy adott földrajzi régióhoz köthető, ezért azok nyomon követésével elméletileg feltérképezhetők az embercsoportok migrációs útvonalai. Történeti szálakat egyes genetikai markerek (haplocsoportok) leszármazása és a mutációk elterjedése terén kereshetünk, nyelvi vagy népességtörténeti állításokat a genetika nem alapozhat erre.

Y-DNS-vonalak: a sajátos DNS-szakaszokon a mutációk mindig „valahol" keletkeznek, és onnan bizonyos földrajzi régió(k)ban terjednek el. Az Y-DNS-vonalak elemzése kizárólag az apai ágon (uniparentális) öröklődő genetikai markerek vizsgálatára alkalmas. A filogenetika ezzel foglalkozó ága ezek különböző populációkban való megjelenésének mérésével és a változások nyomon követésével, a humán migrációs útvonalak feltérképezésével, a populációk közötti genetikai kapcsolatok rekonstrukciójával foglalkozik.

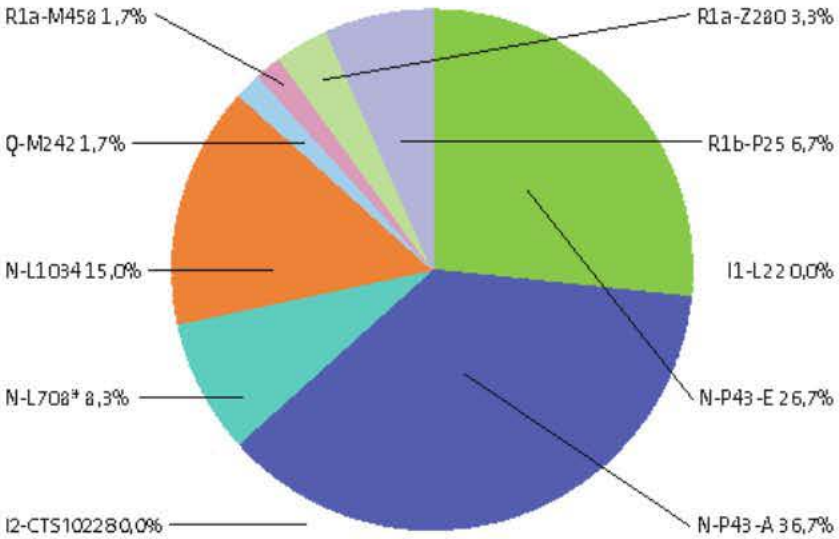
mtDNS-vonalak: az Y-DNS-vonalak „női" párja, szintén sajátos DNS-szakaszokon lévő mutációk elterjedésével foglalkozik. Az mtDNS (mitokondriális DNS) vonalak kizárólag az anyai ágon öröklődő genetikai markerek vizsgálatára alkalmasak. A filogenetika ezzel foglalkozó ága ezek különböző populációkban való megjelenésének mérésével és a változások nyomon követésével, a humán migrációs útvonalak feltérképezésével, a populációk közötti genetikai kapcsolatok rekonstrukciójával foglalkozik.



A 2012-es (északi) és a 2014-es (kondai) manysi mintavételi helyszínek



Az Y-DNS-haplocsoportok eloszlása a mai kondai manysi népességben



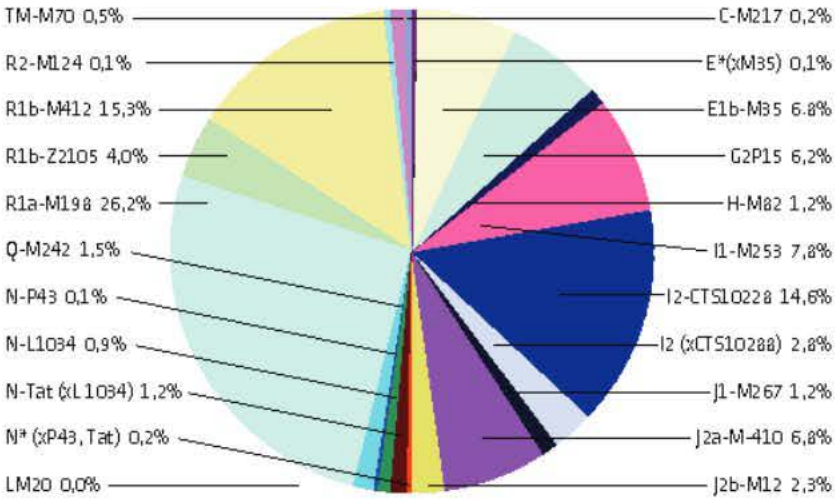
Az Y-DNS haplocsoportok eloszlása a mai északi manysi népességben

Ezeket a mintákat azonos bázisszinten a jelenlegi legnagyobb magyar, székely és csángó adatbázissal hasonlítottuk össze. Nem egyszerűen az így kapott „genetikai térképre" voltunk kíváncsiak, hanem azt szerettük volna megállapítani, hogy van-e olyan szegmens, amelyben sajátos kapcsolat feltételezhető a mai magyarok és manysik között. Azt találtuk, hogy arányaiban a kondai manysik között lényegesen nagyobb a magyar-manysi haplotípusegyezések száma, mint az északi manysik között, és ez a kapcsolat elsősorban az R1a-Z280 haplocsoportban és az N-Tat egyik alcsoportjában (N-L1034) ölt testet. (Lásd a térképet)

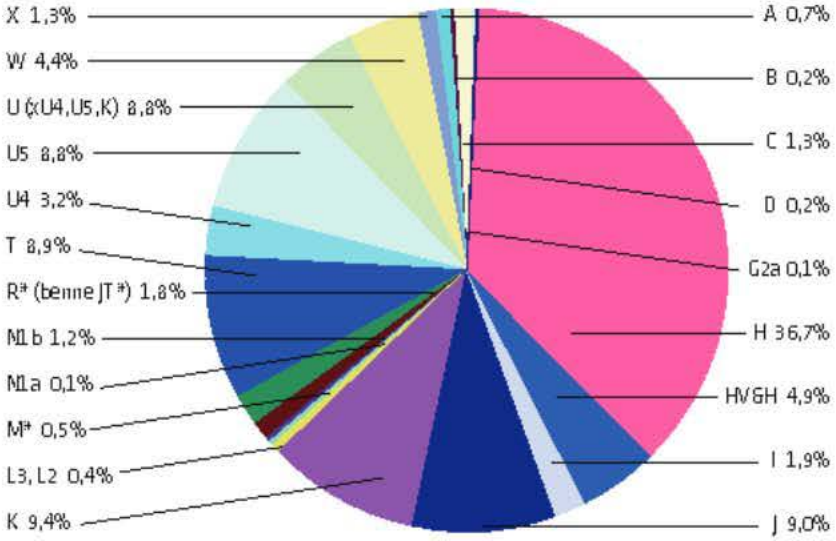
A manysik (külső – orosz – elnevezéssel a vogulok) az Északi-Urálban és Nyugat-Szibériában, az Ob folyó alsó folyása és mellékfolyói mentén élő nép, a hantik mellett a magyarok legközelebbi nyelvrokonai. A – főleg nyelvészeti és történeti néprajzi adatokon nyugvó – feltételezések szerint a magyarok ősei és az obi-ugorok (hantik és manysik) ősei mintegy 2500-3000 évvel ezelőtt váltak szét, ezt követően a magyarság inkább a sztyeppei lovasnomád, míg az obi-ugorok ősei inkább a tajgai halász-vadász népekkel álltak szoros kapcsolatban. A nyelvrokonság azonban nem feltétlenül jelent társadalomtörténeti átfedést, vérségi kapcsolatot, ezért utána akartunk járni annak, hogy vajon milyen genetikai szálak

fűzik össze a mai manysi és magyar népességet.

Míg az első tanulmányok azt jelezték, hogy alig van N-Tat a mai magyar génállományban (kb. 1%), mi azt találtuk, hogy jelentős regionális különbségek vannak a mai magyar populációkban az N-Tat haplocsoport előfordulása tekintetében. Még nagyon távol vagyunk az ideális, regionális bontású magyar adatbázistól, de azt már most látjuk, hogy például a csíkszeredai székelyekben az N-Tat haplocsoport aránya 6,3%. Továbbá az N-Tat haplocsoporton belül három sajátos alcsoport is megtalálható a mai magyarokban. Ebből a legnépesebb az N-L1034 alcsoport, amely potenciális ugor migrációs markernek tűnik: a rendelkezésre álló adatok alapján ugyanis az N-L1034 összeköti a mai manysi és magyar népességet, egyúttal elválasztja más, „nem ugor" népektől. Az N-L1034 rendkívül ritka, így komoly támogatást nyújthat az ugor migrációk feltérképezésében. Európában a Kárpát-medencén kívül – ennél kisebb arányban – elsősorban az Urál-vidéken mutatható ki: az Orenburgtól keletre élő déli baskírok között, továbbá az uráli átjárónál és Tatársztánban. Az N-L1034-es alcsoportnak ezenkívül két földrajzi régióban van szignifikáns előfordulása: a Balkánon és Közép-Ázsiában.



Az Y-DNS haplocsoportok eloszlása a mai magyar népességben
Összes magar minta (n=952)



Az anyai ági mtDNS-haplocsoportok eloszlása a mai magyar népességben
Magyar mtDNS (n=934)

Az anyai vonalak

A magyar és manysi anyai leszármazási vonalak összevetésével azonosítottuk az N-L1034 valószínűsített anyai vonalú, közép-szibériai „párját", a C és G2 haplocsoportot. A C és G2a haplocsoport Európában nagyobb arányban egyedül a baskírokban, a komi-permjákokban és a magyar adatbázison belül a csíkszeredai székelyekben fordul elő.

A két, anyai vonalú mitokondriális haplocsoport, a C és G2a maximumát szintén az Orenburgtól keletre élő baskírokban mutathattuk ki Európában, de a Bajkál környékén a leggyakoribb. A Szeged–Kiskundorozsma, Hosszúhát lelőhelyen feltárt 13 darab 10. századi mintából három haplotípusnak volt egyszerre mai obi-ugor és magyar párja is. Így ez az első eset, amikor különböző genetikai kutatóműhelyek eredményei egyértelműen átfedésben vannak a magyar őstörténet vonatkozásában a sírleletekből nyert hajdani és a mai népességből kapott „recens" mintákkal. Az archeogenetikai adatok alapján úgy tűnik, hogy az eurázsiai sztyeppeöv középső területén a bronzkori (pl. a késő krotovói és a fjodorovói) korszakban biológiai folytonosság áll fenn az anyai vonalak alapján, és a népességekben az ugorság jelentős szerepvállalását kell feltételeznünk.

Fel kell hívnunk azonban a figyelmet a populáció- és filogenetikai adatok társadalomtörténeti olvasatának nehézségeire és korlátaira is. A populáció- és filogenetika ugyanis önmagában nem alkalmas arra, hogy társadalom- és nyelvtörténeti kérdésekben eldöntse a már felmerült problémákat. Azonban a társadalomtudományok számára olyan adatokat tud szolgáltatni, amelyek interdiszciplináris szemlélettel árnyaltabb elemzéseket és értelmezéseket tehetnek lehetővé. Ehhez azonban érdemes átgondolni a genetikai mintavétel módszerétől a kiértékelésig a filogenetika eddigi elméleti keretének alkalmazhatóságát. Az elnevezésektől (populációbesorolási kategóriáktól) a reprezentativitásig, az etnicitás és a genetika viszonyától a genetikai adatgyűjtés és -kezelés módszertanáig sok más metodológiai és etikai kérdést is újra kell értelmeznünk. Szükséges lenne továbbá a tudományközi párbeszéd s a társadalom- és történeti tudományokkal való szorosabb együttműködés az adatok értelmezése terén. ◆