

CARACTERITZACIÓ GENÈTICA DE LLINATGES

Jordi MESTRES DE LA MEANA

RESUM

Gràcies als avenços espectaculars dels darrers anys en les tècniques d'anàlisi genètica i a una reducció important dels seus costos, som la primera generació que pot tenir accés a informació genètica individual i utilitzar-la en l'estudi de la nostra genealogia. Aquesta nova disciplina s'anomena genealogia genètica i ofereix un suport biològic a la genealogia tradicional basada en la recollida d'evidències documentals. Aquest treball és una breu introducció a la utilització i interpretació dels resultats genètics per a la caracterització genètica de llinatges.

Paraules clau: genealogia genètica, ADN autosòmic, cromosoma Y, ADN mitocondrial, mutacions genètiques.

ABSTRACT

Thanks to the recent spectacular technical advances in genetic analysis and to an important reduction of its costs, we are the first generation that can have access to individual genetic information and use it for the study of our genealogy. This new discipline is referred to as genetic genealogy and offers biological support to the traditional genealogy based on collecting documental evidences. This work is a brief introduction to the use and interpretation of genetic results for the genetic characterisation of lineages.

Keywords: genetic genealogy, autosomic DNA, Y chromosome, mitochondrial DNA, genetic mutations.

La genealogia (del grec *genea*=generació, família i *logia*=estudi, tractat) és, com la seva etimologia indica, l'estudi de la descendència i/o ascendència d'una persona o família. Tradicionalment, la genealogia s'ha basat en la recopilació d'evidències documentals, principalment registres civils, parroquials i notariaus, que descriuen les relacions de filiació a cada generació. Fins fa ben poc, recopilar aquest tipus d'informació podia representar un treball d'anys de consulta presencial als arxius històrics i diocesans, sovint combinat amb alguna visita als arxius de parròquies remotes.

En els últims anys, els esforços globals d'indexació i digitalització de documentació civil i sacramental han facilitat enormement l'accés remot d'aquests documents des de qualsevol punt del planeta. Per exemple, l'Església de Jesucrist dels Sants dels Darrers Dies ofereix la consulta integrada, pública i gratuïta de registres civils i parroquials de tot el món a través d'una pàgina web especialment dedicada a l'estudi de la genealogia familiar (www.familysearch.org). Alguns arxius diocesans, com el de Tarragona, estan en procés de digitalitzar tots els seus llibres parroquials i els van penjant gradualment a la internet per a què tothom hi pugui accedir gratuïtament (www.ahat.cat). Al País Basc, l'accés als índexs dels llibres parroquials és públic i gratuït (dokuklik.snae.org), tot i que obtenir una còpia digital impresa de cada registre té un cost de 3 Euros per a l'usuari. Altres arxius diocesans, com el de Granada, admeten peticions de consulta dels seus llibres parroquials però a un cost de 30 Euros per expedient. En qualsevol cas, ja siguin d'accés públic i gratuït o privat i comercial, la possibilitat de que qualsevol persona pugui accedir a la informació genealògica continguda en aquests documents ha contribuït en gran mesura a incrementar significativament l'interès general pels estudis de genealogia. A aquest volum de documentació cal afegir les diverses opcions disponibles a internet per a publicar i comparar arbres genealògics (com per exemple, www.myheritage.com o www.geneanet.org), fet que permet detectar connexions remotes entre arbres de persones sense cap relació de parentesc aparent, quelcom impensable fa només uns anys.

Malgrat tot, és important tenir sempre present que cap document, civil o parroquial, representa una prova biològica de filiació. De fet,

estudis recents realitzats en varis països amb individus que afirmaven tenir plena confiança de la seva relació paterna-filial van donar com a resultat que, sorprenentment, entre un 1% i un 2% dels casos, depenent del país, el pare legítim no era el pare biològic [1]. En aquests casos es produeix una transmissió del cromosoma Y d'un pare amb un cognom que no és del llinatge a un individu que portarà el cognom del llinatge. Les conseqüències d'aquestes introgressions genètiques per a la genealogia són tremendament importants. Prenem per exemple un arbre genealògic mínim que tingui unes 1000 persones. Si assumim una distribució similar d'homes i dones, l'arbre tindria aproximadament unes 500 relacions paterna-filials, fet que suposaria que caldria esperar que entre 5 i 10 d'elles el pare legítim que apareix en els documents no es correspondria amb l'autèntic pare biològic. Si extrapolem l'efecte a un arbre genealògic complet que cobreixi 15 generacions, és a dir, fins ben entrat el segle XVI (el límit del que normalment poden arribar a cobrir els registres parroquials que es conserven), caldria esperar que entre 164 i 328 de les relacions paterna-filials descrites en l'arbre serien legítimes però no biològiques.

En els darrers dos anys, els avenços tecnològics en la realització de proves d'ADN han permès reduir-ne els costos a un nivell assequible per a la població general. Això ha permès el desenvolupament d'una nova branca de la genealogia, l'anomenada genealogia genètica, que té com a principal objectiu l'estudi genètic de llinatges i relacions de parentesc a gran escala en base a la informació genètica de poblacions. El nostre ADN està format per 23 parells de cromosomes resultat de la combinació a l'atzar dels cromosomes dels nostres pares. Els primers 22 parells constitueixen el denominat ADN autosòmic, és a dir, no sexual, mentre que el parell 23 (XY per a home i XX per a dona) és el cromosoma sexual. L'ús d'aquesta informació genètica està cridat a canviar radicalment la nostra perspectiva de les genealogies de base purament documental. A continuació s'exposen les diverses proves d'ADN i la seva aplicabilitat en genealogia genètica.

Genètica de Parentesc

Les anàlisis de genètica de parentesc es basen en la comparació de l'ADN autosòmic entre individus. La Figura 1 il·lustra com utilitzar el grau de semblança autosòmica entre individus per a assignar el seu grau de parentesc. Comencem per l'individu situat a la part inferior esquerra marcat amb l'etiqueta IV*. Cada individu rep la meitat dels cromosomes del seu pare i l'altra meitat de la seva mare. Per tant, a nivell d'ADN autosòmic, l'individu IV* compartirà un 50% del seu ADN amb el del seu pare i un 50% amb el de la seva mare (III*). Alhora, els seus pares compartiran un 50% del seu ADN autosòmic amb els seus pares respectius, fet que es tradueix en que l'individu IV* compartirà un 25% del seu ADN autosòmic amb el de cada un dels seus quatre avis, entre ells, la seva àvia materna (II*). Si seguim aplicant la norma a la següent generació, obtindrem que l'individu IV* compartirà un 12.5% del seu ADN autosòmic amb cada un dels seus vuit besavis, un dels quals està representat per l'individu I. Arribats a aquest punt, cal tenir en compte que cada fill d'una parella rep el 50% dels cromosomes de cada un dels pares, però no el mateix 50%. És per aquesta raó que els fills tenen característiques similars però són essencialment diferents, i a nivell genètic es tradueix en el fet que els germans comparteixen el mateix percentatge d'ADN autosòmic entre ells que entre cada un d'ells i els seus pares (un 50%). Així, veiem que l'individu IV* compartirà un 12.5% del seu ADN autosòmic amb l'individu II, germà de la seva àvia (II*). Novament, aplicant la mateixa norma a la següent generació, s'obté que l'individu IV* compartirà un 6.25% del seu ADN autosòmic amb l'individu (III), cosí de la seva mare (III*), i un 3.125% amb l'individu IV, el seu cosí segon. Els individus IV i IV* estan connectats genealògicament per l'individu I. L'individu I és doncs l'ancestre comú més recent (ACMR) entre els individus IV i IV* i el vincle que connecta genealogia i genètica.

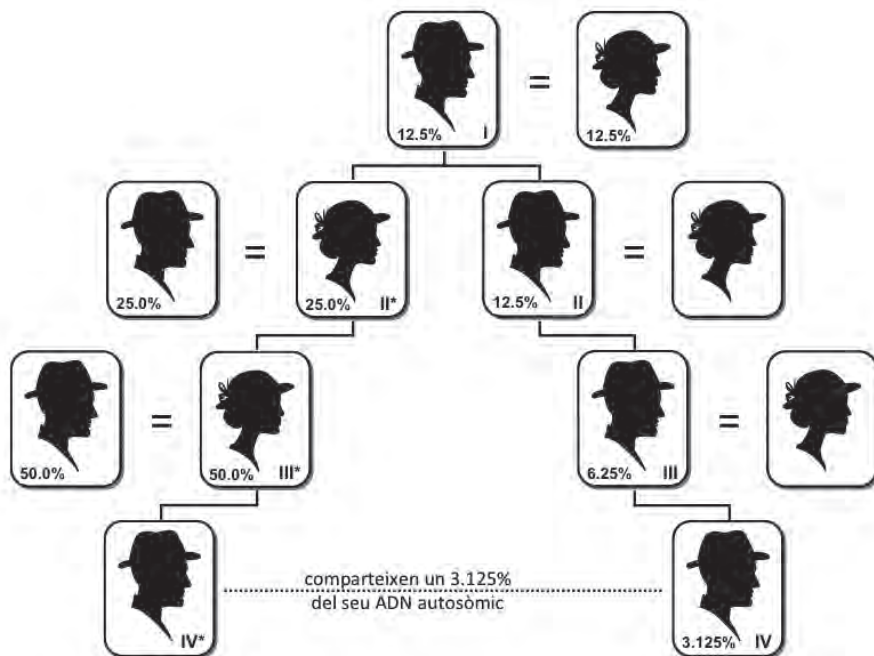


Figura 1. Relació entre el percentatge d'ADN autosòmic compartit i el grau de parentesc familiar entre individus.

Cal remarcar que els percentatges de la Figura 1 són valors teòrics i que, a la pràctica, els valors que s'obtenen dels tests genètics es trobaran dins d'uns marges raonables. Per exemple, tot i que teòricament els cosins primers, segons, i tercers haurien de compartir un 12.5%, un 3.125%, i un 0.78% del seu ADN autosòmic, respectivament, aquest valors poden variar entre un 7.31% i un 13.8% pels cosins primers, entre un 2.85% i un 5.04% pels cosins segons, i entre un 0.3% i un 2.0% pels cosins tercers. La resolució genealògica estimada d'aquests tests d'ADN autosòmic és limitada i com a molt estirar pot arribar a detectar vincles de parentesc fins a unes 7-8 generacions, és a dir, fins a cosins cinquens i sisens que compartirien un 0.05% i un 0.01% del seu ADN autosòmic, respectivament. Però, per altra part, té la capacitat de cobrir totes les branques de la genealogia d'una persona, fet que els fa enormement

valuosos. La presa de material genètic per aquest tipus de proves és senzilla i gens invasiva. Requereix una simple mostra de la nostra saliva, com és el cas de l'empresa 23andMe (www.23andme.com), o un raspap de l'interior de les galtes, com és el cas de l'empresa FamilyTreeDNA (www.familytreedna.com), totes dues americanes. Actualment, el seu cost està entre \$79 (FTDNA) i \$167 (23andMe) i els resultats solen estar disponibles entre quatre a sis setmanes després d'haver enviat la mostra. El preu inclou l'emmagatzement de la mostra i dels resultats genètics en una base de dades, la comparació permanent del teu ADN autosòmic amb el de les noves mostres que entren a la base de dades cada setmana, i l'accés a un entorn personal on pots veure i analitzar visualment els resultats obtinguts i, eventualment, contactar amb els parents genètics identificats.

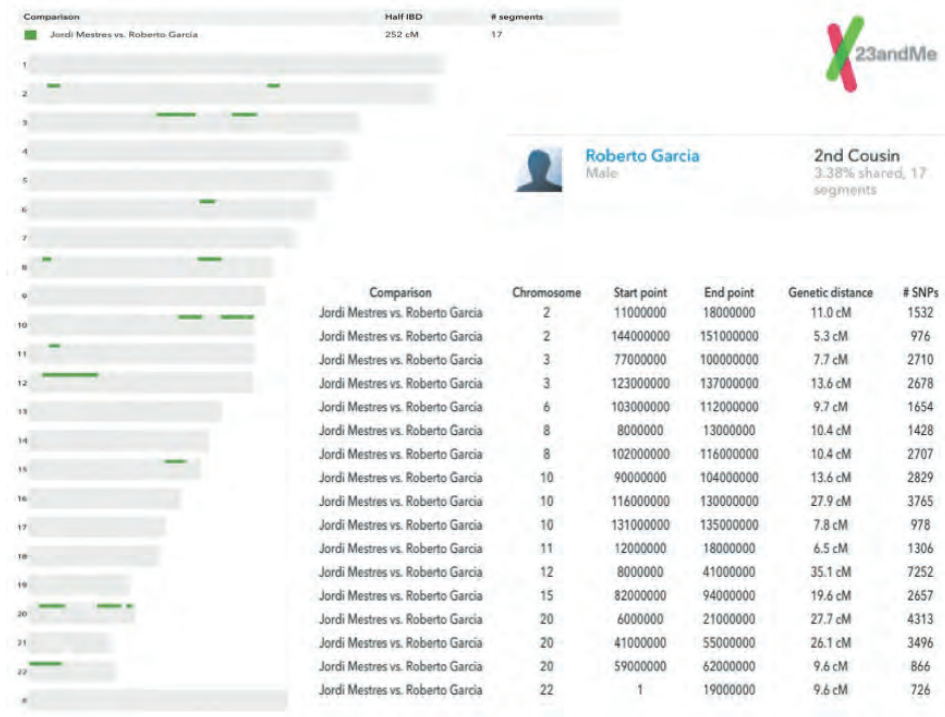


Figura 2. Comparació de l'ADN autosòmic de Jordi Mestres amb el de Roberto García.

La Figura 2 mostra un exemple dels resultats obtinguts amb el test d'ADN autosòmic de l'empresa 23andMe. Tal i com es pot observar, en Jordi Mestres i en Roberto García comparteixen un 3,38% dels seus ADNs autosòmics, valor pròxim al 3,125% esperat teòricament per a una relació genealògica de cosins segons (Figura 1) i certament dins de l'interval esperat (2,85%–5,04%) per aquest grau de parentesc. Aquest percentatge d'ADN autosòmic compartit està repartit en 17 segments de 10 cromosomes. La posició exacta de cada segment compartit s'il·lustra amb una línia verda sobre la llargada total (en gris) de cada cromosoma i es quantifica en una taula que l'usuari es pot descarregar i que, per a cada segment, conté dades sobre el cromosoma on es troba el segment, les posicions de començament i final del segment dins el cromosoma, la distància genètica (en centiMorgans, cM), i el nombre de SNPs. El segment més llarg compartit cobreix pràcticament una quarta part del cromosoma 12. La longitud acumulada per tots els segments de cromosoma compartit és de 252 cM, valor pròxim al 212,5 cM esperat teòricament per a cosins segons. A aquest nivell, la genètica ens dona informació sobre el grau de parentesc (cosins segons) però no és capaç de dir-nos quina de les quatre parelles de besavis respectius (quarta generació) és comuna en les genealogies respectives. Per arribar a identificar quins besavis són els responsables de la relació genètica observada, cal que ambdues persones coneguin bé la seva genealogia. En aquest cas, els besavis comuns són en Francisco Meana (* Argüero, 1876) i l'Ernesta Quintes (* Villaverde, 1884) de les branques asturianes respectives. Els 252 cM dels 17 segments compartits en 10 cromosomes són doncs un romanent genètic de persones que van viure fa 140 anys. És en aquest sentit que podem dir perfectament que els nostres avantpassats estan, d'alguna manera, en nosaltres.

Genètica de Llinatge

Les anàlisis de genètica de llinatge es basen en la identificació de marcadors genètics individuals extrets de la regió no recombinant del

cromosoma Y (ADN-Y), en el cas dels homes, o de l'ADN mitocondrial (ADNmt), en el cas de les dones. Aquests marcadors genètics tenen el seu origen en el mecanisme evolutiu que ha establert la biologia: la generació d'errors, coneguts també com mutacions. Hi ha uns 10 trilions de cèl·lules en el nostre cos que es repliquen constantment. Per tant, és molt probable que de tant en tant es produeixin errors. Però aquestes mutacions rarament es fan evidents. L'excepció són les cèl·lules sexuals (òvuls i esperma), perquè la mutació es transmetrà de pare a fill. Les mutacions en les cèl·lules sexuals són doncs les responsables que el cromosoma Y dels fills mascles sigui lleugerament diferent al del seu pare. Aquestes petites diferències genètiques actuen com a signatura de llinatge i ens permeten saber la distància en la relació genealògica de dos individus. Hi ha dos tipus de mutacions d'especial interès genealògic: els anomenats polimorfismes per canvi d'un sol nucleòtid (SNP) i els micro satèl·lits curts que es repeteixen en tàndem (STR) [2].

Els SNP són mutacions que afecten una sola base de la seqüència d'ADN. L'aparició d'un SNP és un fet extremadament rar. És tant rar, que si passa en una persona en un lloc determinat del seu cromosoma en qualsevol moment de la història, la probabilitat que torni a passar just al mateix lloc és zero. Per exemple, una mutació que es va produir en l'ADN d'un avantpassat nostre per línia masculina estricta que va viure fa 4.000 anys, no apareixerà en el mateix lloc en cap altre home en tota la història. Per la mateixa raó, un cop s'ha produït la mutació i aquesta és transmesa a la següent generació, no es desfarà mai més. Això fa que comparant els SNPs dels homes, podem establir la seva relació genealògica mil·lenària i entendre els patrons migratoris de la humanitat. En aquest sentit, cada home que fa una nova aportació de la informació del seu cromosoma Y permet una agrupació més detallada dels individus en haplogrups i ajuda a expandir el nostre coneixement de l'arbre genètic humà (www.isogg.org/tree/).

Els STR, en canvi, són segments de seqüència d'ADN que es repeteixen varies vegades. En aquest cas, la mutació afecta al nombre de vegades que un determinat segment es repeteix. Al contrari del que succeeix amb els SNPs, les mutacions STR són més comuns.

S'estima que es produeix una mutació STR cada 3-5 generacions, és a dir, cada 90-150 anys. Per tant, comparant els STRs dels homes podem establir la seva relació genealògica en un marc de temps d'interès per la genealogia. En aquest sentit, els estudis de genètica de llinatge amb STRs permeten traçar la línia paterna estricta durant segles. Si dos individus mostren 10 valors diferents dels 111 que s'identifiquen en un test de Y-STR111, s'estima que hi ha un 99% de probabilitat que tinguin un ancestre en comú fa 29 generacions, és a dir, fa aproximadament uns 900 anys.

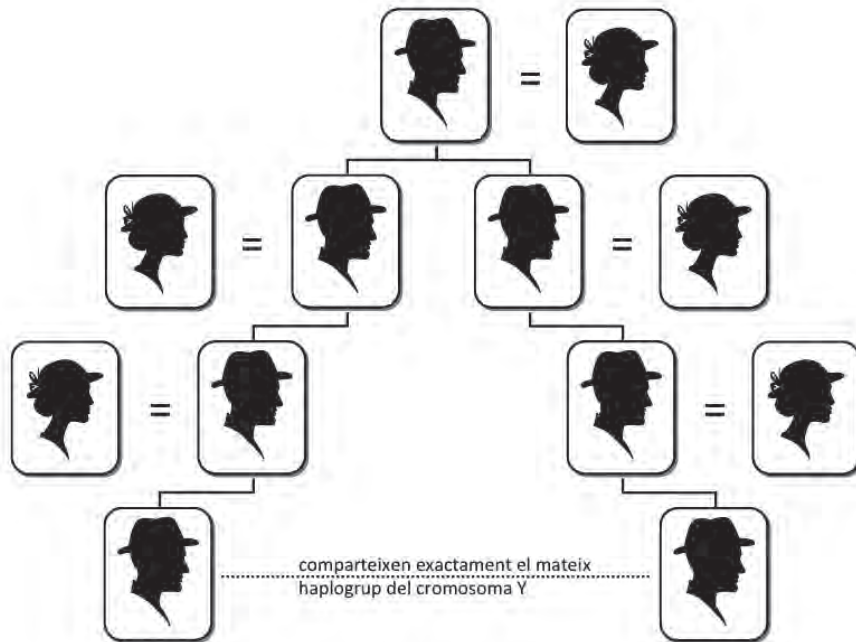


Figura 3. Relació entre l'haplogrup del cromosoma Y entre homes de diferents branques genealògiques que pertanyen a un mateix llinatge genètic.

La Figura 3 il·lustra un cas en que dos homes tenen el seu ACMR de la seva línia patrilinial a una distància de 4 generacions y, per tant, haurien de tenir els seus SNPs i STRs pràcticament idèntics. En

comparació, els SNP i STR dels homes IV i IV* de la Figura 1 provenen de línies patrilineals diferents i per tant els marcadors genètics dels seus cromosomes Y seran diferents, malgrat ser també cosins segons.

Per aquest tipus de tests es recomana l'empresa americana FamilyTreeDNA (www.familytreedna.com) que, com hem vist abans, requereix un simple raspall intern de galtes. Depenent del nombre de marcadors que volem que s'identifiquin, el cost dels tests de cromosoma Y oscil·la entre \$169 (per un test de 37 STRs), \$359 (per un test de 111 STRs), fins a \$575 (pel test anomenat BigY, que cobreix més d'un miler de SNPs coneguts i milions de llocs addicionals on podrien trobar-se nous marcadors).¹ El cost dels tests de cromosoma mitocondrial, en canvi, és més barat i oscil·la entre els \$99 i \$199 (si s'opta per la seqüència completa). Com és habitual, el preu inclou l'emmagatzemament de la mostra i dels resultats genètics en una base de dades, la comparació permanent del cromosoma sexual testejat amb el de les noves mostres que entren a la base de dades cada setmana, i l'accés a un entorn personal on pots veure i analitzar visualment els resultats obtinguts i, eventualment, contactar amb els parents genètics identificats. Tal i com mostra la Figura 4, un aspecte interessant afegit és l'ecosistema de projectes que s'han iniciat a partir de les dades genètiques generades a FamilyTreeDNA, molts d'ells gestionats per acadèmics i professionals en el camp de la genètica. Hi ha tot tipus de projectes, des d'aquells centrats en un cognom concret (com Campbell) a aquells que intenten cobrir tot un haplogrup (com el R1a1a). Des de l'entorn particular a FamilyTreeDNA es pot optar a unir-se a un o més projectes. La contribució de les dades pròpies ajuda a tot el projecte en general i permet aprofundir en el coneixement de l'origen del teu cromosoma Y en concret.

1. Si s'opta pel test BigY, per un cost adicional de 50 Euros, es recomana enviar els resultats a analitzar a YTree (www.ytree.org).

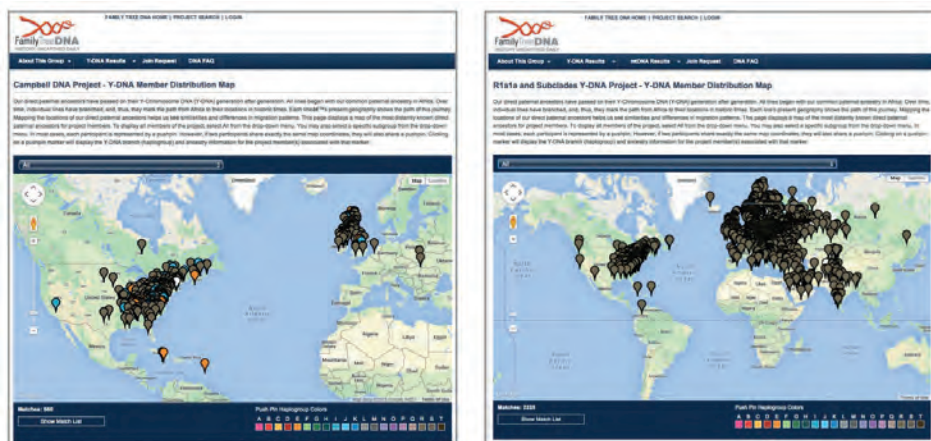


Figura 4. Exemples de pàgines de projectes centrats en mostres de cromosoma Y d'homes que porten el cognom Campbell (a l'esquerra) o que pertanyen a l'haplogrup R1a1a (a la dreta); cada marca representa un individu col·locat al lloc d'on se sap prové el seu ancestre més remot. Fixeu-vos que no tots els individus que es diuen Campbell pertanyen al mateix haplogrup: la majoria de les marques són grises, però n'hi ha de blaves i de taronges que indiquen possibles casos d'introgressió genètica.

Discussió i Conclusions

La caracterització genètica de genealogies complementa la condició legítima de les evidències documentals amb la confirmació biològica dels resultats genètics. Tradicionalment, es tendeix a confondre el primer cognom d'una persona amb el llinatge al que pertany. Però sovint cal tenir ben present que la transmissió del cognom no és sempre sinònim de transmissió de cromosoma Y (pels homes) i que un individu amb un cognom determinat pot tenir de fet el cromosoma Y propi del llinatge d'un altre cognom. Un cas interessantíssim per a l'aplicació de la genealogia genètica és el de la transmissió cognàtica (a través de la dona) del cognom, una pràctica habitual als llinatges de les comarques gironines fins ben entrat el segle XVIII, on els cognoms no estaven associats a la persona sinó a la casa a la que pertanyien. Així

queda reflectit en els llibres parroquials d'òbit de l'època amb frases com “*no se li coneix nom de casa*” quan es trobava algú mort al carrer i que no coneixien al poble.

Es tendeix a creure que aquest efecte era solament típic de les famílies nobles o benestants. Absolutament fals. Cal remarcar que això era un fet habitual a totes les cases, indistintament de la seva riquesa i condició social: quan un home passava a viure a casa de la dona, l'home i tota la seva descendència adoptava el cognom de la dona, el nom de la casa. Per exemple, el 1552, Bernat de Vinyals, Cavaller de Flaçà, casà amb la pubilla de can Ribot, filla de Joan de Ribot, Cavaller de Foixà; els seus fills portaran el cognom de Ribot. Més endavant, el 1590, en Jaume Suardell, Ciutadà Honrat de Barcelona, casaria amb la pubilla de can Garriga, pagesos de Juià; tots els seus fills portaran el cognom Garriga. I ja dins dels segle XVIII, concretament el 1719, en Miquel Calvó, pagès, es va casar amb la pubilla de can Mestres de Celrà; tots els seus fills portaran el cognom Mestres. Són exemples representatius on, independentment del nivell social de l'home, la transmissió del cognom es produeix per via cognatícia: un cavaller amb un filla de cavaller, un ciutadà honrat amb una filla de pagès, i un pagès amb una filla de pagès. En tots aquests casos la dona, pubilla propietària de la casa, imposarà el *nom de casa* a la següent generació. Compte doncs amb les genealogies gironines!

Enrere ha quedat aquella època en què el genealogista s'inventava genealogies mítiques sense cap tipus de suport documental amb l'únic objectiu d'afalagar la vanitat del client. D'aquesta pràctica que tant de mal ha fet a la genealogia en queda la frase “*mentir com un genealogista*”. Avui dia, fer la genealogia d'un llinatge implica realitzar una investigació acurada i perfectament documentada de les relacions entre pares i fills a cada generació, detallant-ne el lloc i la data de cada esdeveniment personal (naixement, baptisme, matrimoni, testament, òbit). En aquest context, la genealogia genètica complementa la genealogia documental aportant una signatura biològica de llinatge. Ja existeixen iniciatives per crear els primers atlas genètics d'un nombre reduït de cognoms que permetran determinar tant si persones amb un mateix cognom

pertanyen a llinatges diferents com si persones de cognoms diferents provenen de fet del mateix llinatge (<http://www.cognoms.upf.edu>) [3]. La incorporació d'informació genètica de llinatges a les genealogies és doncs altament recomanable ja que aporta una marca de llinatge irrefutable i perpètua per a les generacions futures.

Recomanació Final

Per tal d'incloure informació genètica en la descripció d'una genealogia, es proposa afegir a cada generació l'haplogrup del cromosoma Y testejat del pare i del cromosoma mitocondrial testejat de la mare segons el següent format:

Nom_Pare 1erCognom_Pare 2onCognom_Pare, * Lloc_de_naixement dia.mes.any, † Lloc_de_defunció dia.mes.any, **DN** Haplogrup_adnY-SNP_més_profund; = Lloc_de_matrimoni dia.mes.any amb Nom_Mare 1erCognom_Mare 2onCognom_Mare, * Lloc_de_naixement dia.mes.any, † Lloc_de_defunció dia.mes.any, **DN** Haplogrup_adnMt

Exemple:

Generós Mestres i Teixidor, * Sant Daniel (Girona) 15.5.1940, † Girona 14.12.2010, **DN** R1a-Y5992; = Girona 10.6.1966 amb Dolors Aurora López i López, * Mecina Bombarón (Granada) 15.9.1943, **DN** H1t1a

Amb aquesta informació queden genèticament caracteritzades, a perpetuïtat, les línies estrictament masculines i femenines descendents d'aquest matrimoni. En l'exemple mostrat, tots els seus descendents per línia masculina estricta hauran de portar un cromosoma Y que pertany a l'haplogrup R1a i és positiu per al SNP Y5992, mentre que totes les seves descendents per línia femenina estricta hauran de portar un cromosoma mitocondrial que pertany a l'haplogrup H1t1a.

Referències

- [1] King, T.E., Jobling, M. A. *Founders, Drift, and Infidelity: The Relationship between Y Chromosome Diversity and Patrilineal Surnames*. Mol. Biol. Evol. 2009; 26: 1093-1102.
- [2] The Y Chromosome Consortium. *A Nomenclature System for the Tree of Human Y-Chromosomal Binary Haplogroups*. Genome Res. 2002; 12: 339-348.
- [3] Solé-Morata, N., Bertranpetit, J., Comas, D., Calafell, F. *Y-Chromosome Diversity in Catalan Surname Samples: Insights into Surname Origin and Frequency*. Eur. J. Hum. Genet. 2015; 23: 1549-1557.