

Sukulaisuutta setvimässä

Hannu Korhonen, Orimattila

Ilmestynyt Wanha Tuomas -lehdessä 2019

Päivitetty 17.9.2019

Ihminen tuntee useimmiten lähimmät sukulaisensa: isoisä, tyttärentytär, serkku. Mutta mitä oikeastaan on sukulaisuus? Yhteinen sukunimi? Yhteinen perimä? Kasvaminen samassa perheessä? Onko sukulaisuus tunteen vai tiedon asia? Mitä perimä kertoo sukulaisuudesta? — Tämä on juttu, jonka olisin halunnut lukea ymmärtääkseni dna-sukututkimusta. Koska sellaista kattavaa yleistajuista ajantasalla olevaa artikkelia ei ollut ainakaan millään osaamallani kielellä, niin se piti kirjoittaa itse. Kirjan kokoisia esityksiä on sen sijaan paljon ja monenlaisia.

Otetaanpa esimerkiksi Kasse ja Lumi Korhonen. Sukuselvitysten mukaan he ovat sukulaisia, molemmat viidettätoista sukupolvea Rautalammin Korhosten kantaisästä Tuomas Pekanpoika Korhosesta lukien. Kasse on nelikymppinen toimitusjohtaja Jyväskylästä ja Lumi viisivuotias esikoululainen Kaliforniasta. Heidän yhteinen esi-isänsä on noin 1570 syntynyt Paavo Lemetinpoika Korhonen (sukukirjan taulu 3). Sukupolvissa laskettuna etäisyys on 12 sukupolvea ja ajassa 400–440 vuotta. Keskimääräinen sukupolven pituus on tästä laskien noin 33–36 vuotta, siis vähän enemmän kuin laajaan aineistoon perustuva 31 vuotta.

Pieni osa perimästä säilyy likimain muuttumattomana sukupolvesta toiseen. Sen avulla voidaan seurata **isälinjaa** eli isää, isänisää, isänisänisää ja niin edelleen tuhansien, jopa kymmenien tuhansien vuosien taakse; samoin **äitilinjaa** eli äitiä, äidinäitiä jne. Suurin osa perimästä kuitenkin muuttuu sukupolvien mukana, kun isänäitien ja äidinisien mukana tulee uutta perimäainesta ja vanhasta osa jää satunnaisuuden takia siirtymättä jälkeläisille. Isä- ja äitilinjaja luonnehtivan perimän osuutta lukuunottamatta yhteinen perimä puolittuu tällä tavalla keskimäärin sukupolvessa. Kahdentoista sukupolven jälkeen muusta kuin isä- ja äitilinjaja luonnehtivasta sukulaisuudesta olisi siten jäljellä vain $0,5^{12} \approx 0,00024 \approx 0,2$ promillea. Sattumasta riippuva vaihtelu on kuitenkin suurta.

Otetaan mukaan Matilda ja Hemmi Hassinen sekä Reijo Korhonen. Hemmi on Lumin kolmas serkku. Heillä on yhteinen esi-isä: Lumin isänisänäidinissä ja Hemmin äidinäidinäidinissä, siltavahti Jussi Korhonen s. 1874. Siis selvästi sukulaisia erilaisesta sukunimestä huolimatta. Matilda taas on Hemmin sisko, mutta adoptoitu kaukomailta. Yhteistä perimää heillä ei siten ole sen enempää kuin ihmislajin jäsenillä yleensäkkään, mutta Suomen lain mukaan he ovat tasavertaisia sisaruksia esimerkiksi perintökaaressa.

Onko myös Matilda siis Lumin sukulainen samoin kuin Hemmikin? Ainakin lähisuvun mielestä on, sillä näin sanottiin asiaa tiedustellessani: "sukulaisia mitä sukulaisia, kaikki yhtä lailla geeneistä piittaamatta" ja "ei tuo adoption kautta syntyvä sukulaisuus ole sen enempää ajatuksia herättävä kuin sukulaisuus, ystävyys muutenkaan".

Reijo on puolestaan aivan satunnainen Korhonen, Kaskisissa asuva kuusikymppinen merimies. Onko hän Lumin sukulainen, koska hänellä on sama sukunimi? Perinteinen kirkon- ja henkikirjoihin nojaava sukututkimus saattaisi vastata tähän, mutta todennäköisesti ei vastaa, sillä Korhosia on paljon ja sukunimikin on niin vanha, että sen alkuhistoria ulottuu kirjallisten lähteiden

takaiseen aikaan. Dna-sukututkimus vastaisi kysymykseen paremmin, mutta se edellyttäisi testien tekemistä.

Isälinjat

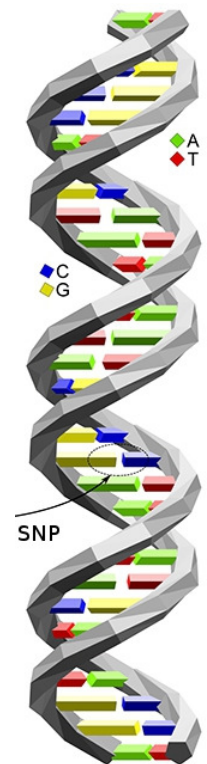
Dna-sukututkimuksen perustan ymmärtämiseen riittävät hyvin nykyisen koulubiologian tiedot. Tieto ihmisen perimästä sisältyy kaksoiskierteisen **deoksiribonukleiinihappomolekyylin** aminohappopareihin. **Dna**:ta on ihmisen kaikissa soluissa, jokaisen solun tumaan kromosomeissa ja muissa soluelimissä, kaikkiaan kolmisen miljardia rakenneyksikköä. Kromosomeja on 23 vastinkromosomiparia.

Vastinkromosomipareista viimeinen määrää sukupuolen. Äidillä on tässä parissa kaksi X-kromosomia, joten lapsi saa häneltä aina X-kromosomin. Isällä vastinkromosomit ovat erilaiset X ja Y. Lapsi saa näistä toisen. Jos saa X:n, lapsesta tulee tyttö, jos taas Y:n, niin lapsesta tulee poika. Sukupuolen määräävä geeni on Y-kromosomin lyhyemmässä haarassa. Mekanismi on herkkä eikä siksi aina kaikki mene pääsäännön mukaan. Useimmille ihmisille tällainen biologinen sukupuolen määräytyminen pätee kuitenkin.

Järjestelmän herkkyydestä ja monimutkaisuudesta kertoo myös se, että kromosomit eivät pysy samanlaisina, vaan muuttuvat eli mutatoituvat satunnaisesti pienempien osalta kerrallaan. Yksittäistä muutoskohtaa nimitetään "snipiksi" (SNP), joka on lyhenne ilmauksesta "single nucleotide polymorphism", yhden emäksen monimuotoisuus. Sitä voitaisiin kutsua myös pistemutaatioksi, koska kyseessä on yhden nukleiinihappoparin muutos dna-ketjussa.

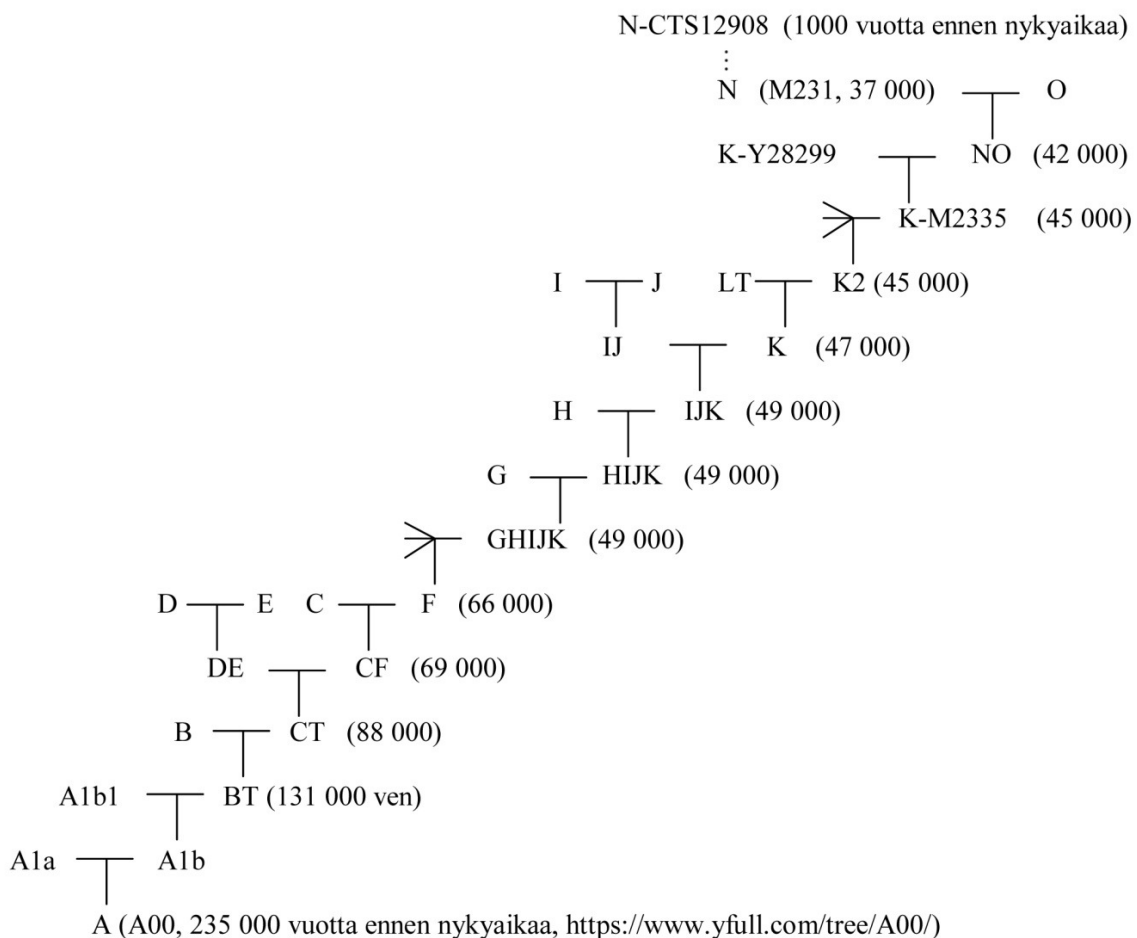
Muutoksia on tapahtunut vuosituhsien aikana paljon. Arvioidaan, että ihmisen perimässä tapahtuu keskimäärin kaksi Y-kromosomin mutaatiota sukupolvessa. Nämä saattavat periytyä miespuolisille jälkeläisille. Perimältään eroavia miesryhmiä eli **Y-haploryhmiä** on siis hyvin paljon. Nykyään (2018) tunnetaan lähes kuusi miljoonaa Y-kromosomin muunnelmaa. Nimitys "haploryhmä" tulee siitä, että sukusolujen muodostumisessa syntyvä kromosomisto on yksinkertainen eli haploidinen, koska sukusoluissa on jakautumisen seurauksena vain toinen vastinkromosomeista. Muiden solujen kromosomisto muodostuu vastinkromosomipareista ja on siis kaksinkertainen eli diploidinen.

Haploryhmät voidaan järjestää tavanomaista sukupuuta muistuttavaksi haplosukupuuksi. Se on yksi esimerkki biologisten ominaisuuksien muuttumista kuvaavasta evoluutiopuusta (phylogenetic tree), missä eliöt esitetään yhtäläisyyksiensä ja erojensa mukaan järjestettyinä kokonaisuuksina. Y-dna:n haplopuun päähaarat nimetään kirjaimilla A–T tai muutaman kirjaimen ryhmillä, alaryhmät kirjainten lisäksi numeroilla. Puu kasvaa ja sen rakenne muuttuu jatkuvasti sitä mukaa kuin uutta tietoa saadaan, vaikkakin pääryhmät alkavat jo olla vakiintuneita.



Suomalaisilla miehillä yleisin ryhmä on N (58 %). Näitä ovat useimmat Korhosetkin. N-ryhmä on itäistä perua, syntynyt noin 37 000 vuotta sitten jossain Intian suunnalla tai vieläkin idempänä.. Sitä määrittää kaksi mutaatiota M231 ja M232, joiden keskinäistä syntyjärjestystä ei tunneta. Kaikilla N-miehillä on molemmat näistä. Eurooppalaisten yleisin ryhmä on I. Suomalaisilla se on toiseksi yleisin (29 %). Se on puolestaan läntinen, syntynyt Balkanilla noin 30 000 vuotta sitten ja kiertänyt myöhemmän Länsi-Euroopan kautta Suomeen. Kaikkien haploryhmien yhteisen kantaisän, dna-Aatamin, arvellaan eläneen 235 000 vuotta sitten, vaikka arvioiden vaihteluväli onkin suuri: 160–300 000 vuotta.

Uusi mutaatio erottaa ryhmästä uuden alaryhmän, jos mutaation kantajan miespuolisia jälkeläisiä jää eloon. Toisiaan seuraavat mutaatiot ja niiden määrittelemät haploryhmät muodostavat ketjuja, joilla on yhteinen juuri: kanta-Aatami (Y-haploryhmä A). Ajoitukset perustuvat tietoon keskimääräisestä mutatoitumisnopeudesta ja paikka-arviot tietoihin alueellisista haploryhmätihentymistä¹. Tässä on kaksi epävarmuutta luovaa asiaa. Yhtäältä kaikkia mutaatioita ei ole havaittu eikä välttämättä ehkä koskaan havaitakaan, kun mutaation kantajia ei ole enää elossa. Mutaatiomääristä lasketut aikaväliarviot jäävät silloin liian lyhyiksi. Samoin käy, jos mutaatiota seuraa takaisinmutaatio täsmälleen samassa dna-ketjun kohdassa. Tällä tavalla saattaa jäädä havaitsematta useammankin peräkkäisen, vastakkaisiin suuntiin tapahtuneen mutaation ketju.



Y-haplopuun pääryhmät 2018. Tässä on esitetty vain N-ryhmään johtava linja. Muut ryhmät ovat jakaantuneet ja jatkuneet nykyaikaa kohti vastaavalla tavalla.

Pääryhmien luokittelu on varsin nuorta. Esimerkiksi haploryhmä N muodostettiin vuonna 2008 yhdistämällä useampi aikaisempi ryhmä. Siinä luokituksessa suomalaiset kuuluivat alahaploryhmään N1c1a². Vuonna 2017 N-ryhmän alaryhmät nimettiin uudestaan. Tällöin N1c1a (M178) sai nykyisin käytössä olevan tunnuksensa N1a1a³. Verkossa on edelleen paljon materiaalia, jossa viitataan vanhempaan tunnukseen N1c.

Haplosukupuun oksien tunnistamisessa pistemutaation tunnus onkin siksi ehkä varmempi kuin itse ryhmän tunnus, jonka merkitys varmistuu vasta, kun on selvillä minkä ajankohdan tiedosta on kyse. Pistemutaatioidenkin nimeämisessä on päällekkäisyyttä, sillä yritykset nimeävät niitä oman luokituksensa mukaan, ainakin niitä, jotka yritys on löytänyt ensimmäisenä. Niinpä esimerkiksi Sotkamon Korhosten nuorempaa haaraa luonnehtivasta mutaatiosta käytetään tunnuksia N-FCG72591 ja N-Y67221⁴. Tilannetta helpottaa se, että yritykset kyllä usein mainitsevat omissa haplopuissaan myös käytössä olevat rinnakkaiset nimet.

Äitilinjat

Äitilinjan määrittäminen ei voi perustua sukupuolikromosomiin X, koska nainen on saanut toisen niistä äidiltään, mutta toisen isältään. Vain sattumasta riippuu, kumpi periytyy tyttärelle. Onneksi perintötekijöitä on myös muualla kuin sukupuolikromosomeissa, nimittäin solun energiatuotannosta vastaavissa soluelimissä eli **mitokondrioissa**. Perimäainesta niissä on hyvin vähän, vajaat 17 000 emäsparia kaikista kolmesta miljardista eli noin kuusi miljoonasosaa, mutta se riittää, koska se kaikki tulee äidiltä munasolun mukana. Se ei siis periä tavalliseen tapaan eli satunnaisesti molemmilta vanhemmilta.

Äidin mitokondriot tulevat munasolun mukana myös miehille, joten äitilinjan testi voidaan tehdä miehenkin antamasta näytteestä. Samoin kuin sukupuolen määräytymisessä on tässäkin pieni poikkeus, sillä – vaikkakin hyvin harvoin – myös isän tai siis tarkemmin isänäidin mitokondrio-dna:ta saattaa poikkeuksellisesti vuotaa jälkeläiselle siittiösolun hännän mukana. Normaalistihan häntä tuhoutuu välittömästi hedelmöitymisen jälkeen eivätkä isän mitokondriot siis siirry jälkeläiselle. Mutatoitumismekanismikin on erilainen siinä mielessä, että mutaatio tapahtuu yhden solun yhden mitokondrion dna-renkaassa ja yleistyy, jos on yleistykseen, vasta jälkeläispolvissa. Siirtymävaiheessa puhutaan heteroplasmasta.

Arviot mitokondrio-dna:n mutaationopeudesta vaihtelevat hyvin paljon menetelmästä ja testatusta ihmisryhmästä riippuen. Mitokondrio-dna on joka tapauksessa paljon vakaampi kuin Y-dna. Mutatoitumisnopeuden vaihteluväli on yksi mutaatio 15–60 sukupolvea kohti, keskimäärin yksi mutaatio 1000 vuodessa. Tästä syystä mitokondrio-dna:sta ei voi tehdä muutaman sadan vuoden aikaa koskevia sukulaisuuspäätelmiä. Suuret erot arvioissa johtuvat osaltaan myös siitä, että vaikeina aikoina – geneettisissä pullonkauloissa – syntyneet mutaatiot eivät ole päässeet periytymään, vaan ovat hävinneet syntymättä jääneiden tai jälkeläisiä jättämättä kuolleiden lasten tai lastenlasten mukana. Kaikille ryhmille yhteinen esiäiti, dna-Eeva, on tämän perusteella arvioiden elänyt 100–200 000 vuotta sitten.

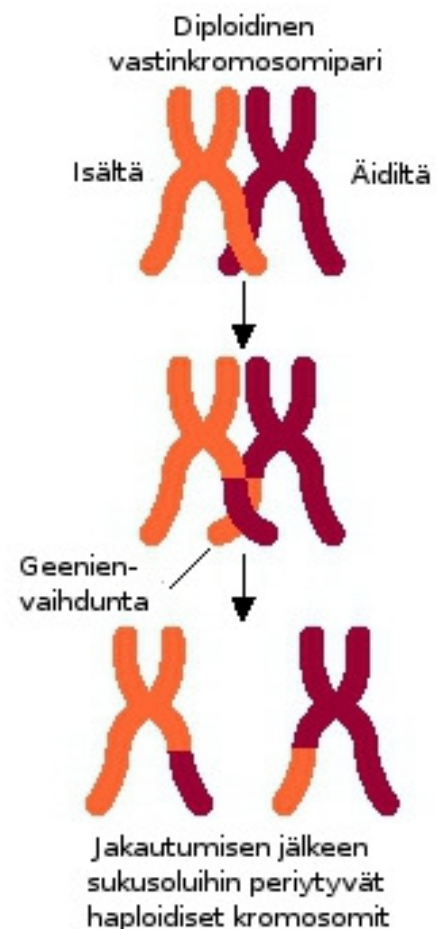
Periaatteet ovat muuten samat kuin isälinjan kohdalla. Mutaatioiden kautta syntyneet perimänmuutokset antavat mahdollisuuden ryhmitellä ihmiset haploryhmiin myös mitokondrio-dna:n perusteella. Ryhmät nimetään vastaavalla tavalla yksittäisin kirjaimin A–Z tai kirjain- ja numeroryhmin⁵. Yleisin ryhmä suomalaisilla on H (40 %). Se on yleisin myös koko Euroopassa (44 %). Toiseksi yleisin Suomessa on U (28 %). Sitä on muualla Euroopassa hyvin vähän, mutta puolet saamelaisista (53 %) kuuluu tähän ryhmään. Vakautensa ansiosta mitokondrio-dna on tärkeämpi antropologiassa ja biomaantieteessä kuin sukututkimuksessa.

Kokonaisperimä

Dna kertoo perimästä paljon muutakin kuin isä- ja äitilinjat. Itse asiassa sukupuolikromosomeissa on perimästä vain kahdeksisen prosenttia, Y-kromosomissa tästä vajaa neljännes. Valtaosa (98 %) on siis muualla kuin Y-kromosomissa tai mitokondrioissa. Tilannetta monimutkaistaa se, että kromosomit eivät aina siirry jälkeläisille muuttumattomina, vaan geenit saattavat vaihtua vastinkromosomista toiseen tekijäinvaihdunnan (crossing-over) kautta. Tapahtumasta käytetään myös nimitystä uudelleenjärjestäytyminen (recombination) Tämä tarkoittaa sitä, että isältä ja äidiltä saatu perimäaines voi sekoittua. Geenit säilyvät mutaatioita lukuunottamatta sellaisinaan, mutta osa voi vaihtaa paikkaa vastinkromosomista toiseen. Geenien uudelleenjärjestäytyminen on harvinaisempaa miehillä – keskimäärin 27 geeninvaihduntaa lasta kohti – kuin naisilla, joilla on keskimäärin 41 geeninvaihduntaa lasta kohti.

Muun kuin isä- ja äitilinjoja koskevan sukulaisuuden selvittäminen muuttuukin näin ollen tilastolliseksi. Tutkitaan kuinka paljon mitokondrioiden ja muiden kromosomien kuin Y:n sisältämästä perimästä eli **autosomaalisesta** dna:sta on yhteistä eri henkilöiden antamissa näytteissä. Mittayksikkönä on **senttimorgan** (cM). Se kuvaa kahden kohdan välille sattuvan geeninvaihdunnan todennäköisyyttä. Ketjun pituus on yksi morgan, jos sen päiden välille sattuu yksi geeninvaihduntakohta.

Yksi senttimorgan tarkoittaa siis sellaista dna-ketjun osaa, jolle geeninvaihduntakohta sattuu yhden prosentin todennäköisyydellä. Ihmisen kromosomien pituus on muutamasta kymmenestä muutamaan sataan senttimorganiin eli muutamasta morganin kymmenyksestä muutamaan morganiin. Pisimmissä kromosomeissa saattaa siis tapahtua useampikin geeninvaihdunta. Kyse ei



Isältä ja äidiltä tuleva autosomaalinen perimä sekoittuu, kun geenit vaihtuvat vastinkromosomista toiseen.

ole ketjun todellisesta pituudesta, vaan siis ketjun pituudesta geenivaihdunnan todennäköisyydellä mitattuna.

Perimältään samankaltaisten ihmisten eli **geneettisten sukulaisten** sukulaisuutta voidaan luonnehtia yhtenevien dna-jaksojen määrällä ja pituudella. Senttimorgan sopii siten myös sukulaisuuden mitan yksiköksi. Lukema ennustaa sukulaisuuden läheisyyttä. Täyssisarusten näytteiden välillä lukema on keskimäärin 2700 senttimorgania, vaikka vaihtelu on suurta (2200–3400), serkuilla 900 (500–1300) ja pikkuserkuilla 200 (50–500)⁶. Laaja vaihteluväli johtuu kahdesta syystä. Ensinnäkään perimä ei tule tasan puoliksi kummaltakin vanhemmalta, vaan sattumalla on suuri vaikutus siihen, kuinka suuri osa autosomaalisesta perimästä tulee kummankin vanhemman esivanhempien linjoista. Toiseksi vanhemmilla voi yhteisen sukuhistoriansa ansiosta olla yhteistä perimää, jota tulee silloin jälkeläisille kummaltakin.

Tämä näkyy erityisen selvästi savolaisissa suvuissa. Kantaväestö oli pieni. Lisäksi vaimo haettiin usein läheltä. Sukulaisavioliitot olivat tavallisia ja toivat siten yhteistä perimäainesta molemmilta vanhemmilta. Kahden henkilön yhteinen sukulaisuus saattaaakin juontua useista yhteisistä, jopa kymmenistä, sukulinjoista. Asianlaita on toinen Hämeessä ja Varsinais-Suomessa sekä varsinkin Amerikassa, jonka väestöpohjaan sukulaisuusarviot perustuvat. Karkea sormituntuma on, että amerikkalaisten dna-analyysyjä tekevien yritysten serkkuenusteet "1–2" (serkku tai pikkuserkku) tai "3–4" (kaukaisemmat serkut) pitävät kohtuullisen hyvin paikkansa Länsi-, mutta eivät Itä-Suomessa. Ennustettu serkku saattaa paljastua siellä usein kolmanneksi tai neljänneksi serkuksi tai kaukaisemmaksi sukulaiseksi.

Valaiseva esimerkki on oman kokonaisperimätestini tulos. Ensimmäisenä listalla oli ja on vieläkin Ari Korhonen, yhteinen perimä 241 senttimorgania, ennuste "toinen tai kolmas serkku". Kuitenkin Ari on Korhosten dna-sukupuun nilsiäläisessä haarassa, mikä osoittaa, että meillä on yhteinen esi-isä ehkä tuhannen vuoden takana. Yhteisen perimän täytyy siis tulla pääasiassa muualta kuin isälinjan kautta. Sukulaisuutemme on vielä tarkemmin selvittämättä. Toinen esimerkki on Vaajasalmen Korhosiin kuuluva Petri Korhonen. Yhteinen esi-isämme on Risto Rekonpoika Korhonen (1700–1776, sukukirjan taulu 15) kuuden sukupolven takana. Kokonaisperimätesti näyttää meidän yhteiseksi perimäksemme 183 cM ja ennustaa Petrin olevan 2.–3. serkkuni.

Kokonaisperimätesti antaa siis tietoja suuresta määrästä sukulaisia, joista monet ovat jo hyvin kaukaisia. Se kertoo yhteisen perimän määrän, vaikka ei sukulaisuuden laatua. Esimerkiksi isälinjan osumissani on kaksi amerikkalaista Hietapelto-nimistä veljestä, joista minulla ei ollut mitään tietoa. Kokonaisperimätestin osumissa näkyi myös heidän lapsiaan, senttimorganmäärät 83–140, ennusteet 2.–5. serkku. Ennuste ei ole ihan huono, sillä meillä on yhteinen esi-isä kahdeksan sukupolven päässä Vaajasalmen Korholassa. Sieltä muutti aivan 1700-luvun alussa sukulaisemme Iisveden toiselle puolelle Lieteenmäkeen, niin kuin kerrottiin buurisotajutussa (WT 2016 s. 8) ja seuraavassa sukupolvessa Pyhäjärvelle. Pohjanmaalla sukunimi vaihtui useampaankin kertaan talon mukaan ja yksi jälkeläinen muutti 1800-luvun lopulla Yhdysvaltoihin. Rautalammin Korhosten geenit kulkivat kuitenkin pojalta pojalle ja elossa olevien mieskantaisten sukulaisten tyttärilläkin on siis yhteistä perimää minun kanssani niin paljon, että se näkyy testeissä.

Dna-analyysit

Suurin osa ihmisen perimästä saatiin analysoiduksi vuonna 2001, koko perimä vasta vuonna 2010. Maailmalla on satoja DNA-analyysijä tekeviä laboratorioita. Näistä vain kymmenkunnan päätoimiala on suvustaan kiinnostuneiden yksityisasiakkaiden palveleminen, useimmat amerikkalaisia kaupallisia yrityksiä. Vain muutamien maiden kansalaisten perimää on pyritty selvittämään järjestelmällisesti, parhaana esimerkkinä Islanti, missä toimiva yritys julkisti 14 000 islantilaisen perimään perustuvan tutkimuksen syksyllä 2017. Suomessa kaupallisia DNA-laboratorioita ei ole, vaan esimerkiksi muinais-DNA:n määritykset tehdään Saksassa.

Perimää tutkittaessa kohdistetaan huomio yleensä vain pieneen osaan perimää, tiettyihin geeneihin tai pienempiin dna-jaksoihin, joista käytetään nimitystä **merkkijakso** eli markkeri. Kuhunkin tarkoitukseen soveltuvien merkkijaksojen valinta on ensisijaisesti tilastollinen ongelma. On löydettävä sellaiset kohdat, joiden perusteella voidaan tehdä luotettavasti kyseessä olevaa asiaa koskevia päätelmiä. Niinpä esimerkiksi syöpätutkimuksessa käytetään useimmiten yhden emäksen monimuotoisuuksia eli SNP-merkkijaksoja. Oikeuslääketieteessä ja perinnöllisyystutkimuksessa käytetään taas tavallisimmin STR-merkkijaksoja. Nimi tulee englanninkielisestä ilmauksesta *short tandem repeat*, mikä tarkoittaa 2–6 vierekkäisen emäksen toistumista.

Sukututkimuksessa käytetään molempia. Isälinjan perusanalyysi perustuu tavallisimmin 37, 67 tai 111 STR-merkkijaksoon. Jo 12 merkkijakson testi saattaa antaa karkean arvion haploryhmästä, mutta tulos on sitä tarkempi, mitä useampaa merkkijaksoa käytetään. Usein suositetaan perusanalyysiksi 67 merkkijakson testiä. Sen perusteella esimerkiksi Family Tree DNA -yritys antaa ennusteen haploryhmästä, minun tapauksessani pääryhmän N, jota luonnehtii mutaatio M231. Siinä on vielä mukana suomalaisten lisäksi baltteja, jakuutteja, balkanilaisia sekä jopa kiinalaisia ja vietnamilaisia.

Aikaisemmin – ennen vuotta 2014 – sama yritys on voinut antaa samantasoisien testin perusteella arvion, joka sijoittuu ylemmäksi haplopuuhun, esimerkiksi ryhmään N1c1a eli nykyisessä järjestelmässä N1a1a, jota luonnehtii mutaatio M178. Joskus tämän tason tiedosta on käytetty nimitystä haplotyyppi, vaikka sanalla on useita muitakin merkityksiä. Tulosta tarkennetaan SNP-testeillä. Niitä voi tilata erillisinä tai erikokoisina yhdistelminä. Laajassa testissä, esimerkiksi Family Tree DNA -yrityksen BigY-testissä, käytetään viittä sataa STR- ja tuhansia SNP-merkkijaksoja. Tämä antaa perimästä tarkan kuvan, jonka perusteella tutkittava voidaan useimmiten sijoittaa haplosukupuun uloimmille oksille.

Tulosten vertailu

Hyödyllisintä eivät kuitenkaan ole dna-analyysin tulokset sinänsä eli yksittäisten merkkijaksojen arvot, vaan tulostietokannat, joiden perusteella perimää vertaillaan muiden ihmisten perimään. Näihin tietokantoihin sisältyvien tietojen yhdistämiseen perustuvat edellä kuvatut haploryhmät ja -puut. Siksi sukulaisten löytymiseen tarvitaan kaikkien testejä teettäneiden myötävaikutusta. Heidän

on annettava perimä- ja yhteystietonsa ainakin samankaltaisen perimän omaavien ihmisten eli lähiosumien (match) saataville. Vain omaan käyttöön rajatut tiedot eivät hyödytä ketään.

Osa tiedosta on julkista. Sitä voi kuka tahansa katsella⁷. Analyysitulokset näkyvät merkkijakso merkkijaksolta. Näytteen antajan nimi ei näy julkisella sivustolla, vaan vain vanhin tunnettu esi-isä, jos testin teettäjä on sen merkinnyt tietokantaan (kuva). Tässä merkkijaksot ovat STR-merkkijaksoja, mutta haploryhmä on varmistettu SNP-merkkijaksojen avulla. Ensimmäisestä merkkijaksosta DYS393 käytetään joskus myös nimeä DYS395. Se on Y-kromosomiin sisältyvä neljän emäksen jakso, suhteellisen vakaa, mutatoituu keskimäärin kerran 1300 vuodessa.

Kit Number	Paternal Ancestor Name	Haplogroup	DYS393	DYS390	DYS391	DYS385	DYS388	DYS426	DYS439	DYS389I	DYS392	DYS389II	DYS458	DYS459	DYS455	DYS447	DYS448	DYS449	DYS464			
115186	Tuomas Pekanpoika Korhonen, 1530-1597, Rautalampi, FI	N-CTS12908	14	24	14	11	11-13	11	12	11	14	14	29	16	10-10	11	12	24	14	18	31	13-13-14-14

Alkua erään Rautalammin Korhosen perimän analyysituloksesta. Henkilö itse ei näy julkisesti, vaan vain hänen näytteensä numero (kit number).

Joillakin yrityksillä on maittain tai suvuittain järjestäytyneitä ryhmiä. Tällainen on esimerkiksi *Family Tree DNA* -yrityksen *Finland DNA* -projekti⁸. Se palvelee myös suomeksi ja sen kautta tilatuista testeistä saa pienen alennuksen. Siellä on tätä kirjoittaessani (kesäkuussa 2018) noin kahdentoista ja puolen tuhannen ihmisen tiedot. Korhosillekin on perustettu oma dna-ryhmä⁹ alkusyksystä 2018.

Maallikon ei tarvitse kuitenkaan itse yrittää tulkita dna-testin tulosta eikä sijoittaa itseään haplosukupuuhun. Ryhmätiedon saa testituloksen mukana. Tarjolla on myös palveluja, jotka analysoivat tarkemmin dna-testin raakadataa ja näyttävät oman haploryhmän sijainnin koko ihmiskunnan perimäkartassa. Sieltä Y-haplopuun N-haaran (M231) uloimmilta oksilta löytyvät Korhosetkin.

Kaikki Korhoset eivät kuulu N-ryhmään, vaan esimerkiksi Sievin Korhosten haploryhmä on I1. Tähän on ilmeinen ja tiedossa oleva selitys. Kalajoen pitäjän Korttesjärvellä oli 1600-luvun alussa Lauri Korhosen asuttama talo. Se autioitui 1600-luvun puolivälissä ja uudeksi asukkaaksi tuli vasta parikymmenvuotiaana Sipi Simonpoika, s. n. 1648. Hän sai talon mukaan sukunimekseen Korhonen. Hänestä alkoi Sievin Korhosten suku, joka on siis ollut Korhosia jo runsaat kolmesataa vuotta, mutta ei ole perimältään samaa kantaa muiden Korhosten kanssa.¹⁰

Yritysten tarjoamissa palveluissa ja tekemissä testeissä on eroja. Esimerkiksi *MyHeritage* tekee vain kokonaisperimätestin (Ancestry Composition), kun sen sijaan *Family Tree DNA* tarjoaa kokonaisperimätestinsä (Family Finder) lisäksi eritasoisia isälinjatestejä (Y12, Y25, ... Y500) ja äitilinjatestin (mt-dna). Viimeaikoina käyttöön on tullut mahdollisuus siirtää dna-tuloksia toisten yritysten analysoitavaksi ja ohjelmistoja, joilla voidaan vertailla eri yritysten tekemien dna-

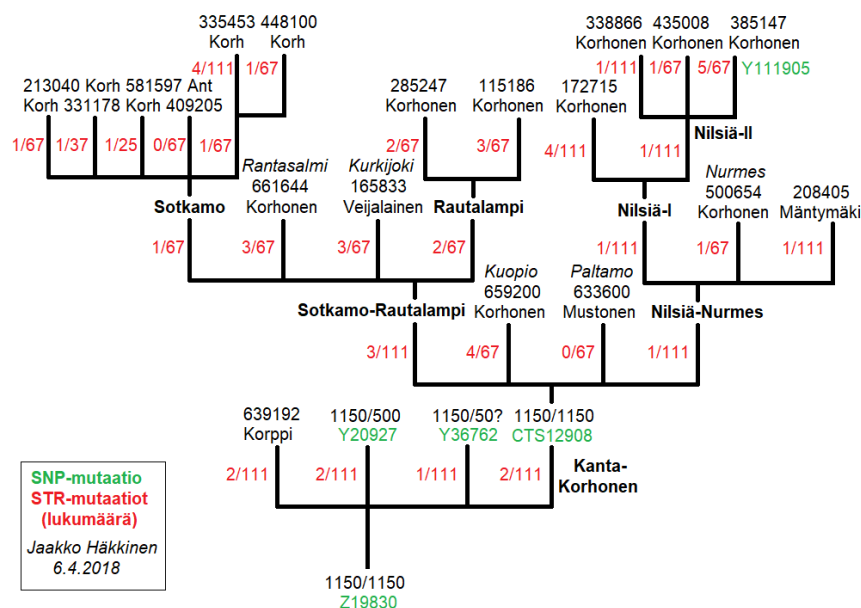
analyysien tuloksia. Siirtämisestä on se hyöty, että silloin tavoitetaan myös niitä sukulaisia, jotka ovat tilanneet testinsä toiselta yritykseltä ja joiden tiedot ovat vain sen tietokannassa.

Kannattaa siis jo ennen ensimmäisen testin tilaamista miettiä, halutaanko teettää vain kokonaisperimätesti vai halutaanko myös seurata isä- tai äitilinjoja. Halutaanko varata mahdollisuus jatkaa perusanalyyysiä pitemmälle? Onko sukutaustassa sellaisia kohtia, joihin kirkonkirjat eivät anna vastausta? Halutaanko varmistua oikeista biologisista vanhemmista? Halutaanko löytää sellaisia sukulaisia, joita perinteinen sukututkimus ei ole tavoittanut? Halutaanko päästä kauemmas menneisyyteen kuin asiakirjat antavat myöten? Halutaanko muuten vain etsiä kaukaisempia sukulaisia? Tai halutaanko tietoa periytyvistä sairauksista tai sairauksille altistavista geneistä?

Korhosten perimä

Kun puhutaan Korhosista, niin tarkastellaan erityisesti niitä sukulinjoja, joiden mukana sukunimi on periytnyt eli isälinjoja. Kaikkien Korhosten keskinäistä sukulaisuutta ei voi selvittää arkistotietojen perusteella, sillä sukunimemme on niin vanha, että sen synty sijoittuu kirjallisten lähteiden takaiseen aikaan. Useiden Korhos-sukujen yhteinen haploryhmä on N1a1a1a1a2a1a1a1a1a1a1 (N-CTS12908), näiden joukossa Rautalammin Korhosetkin.

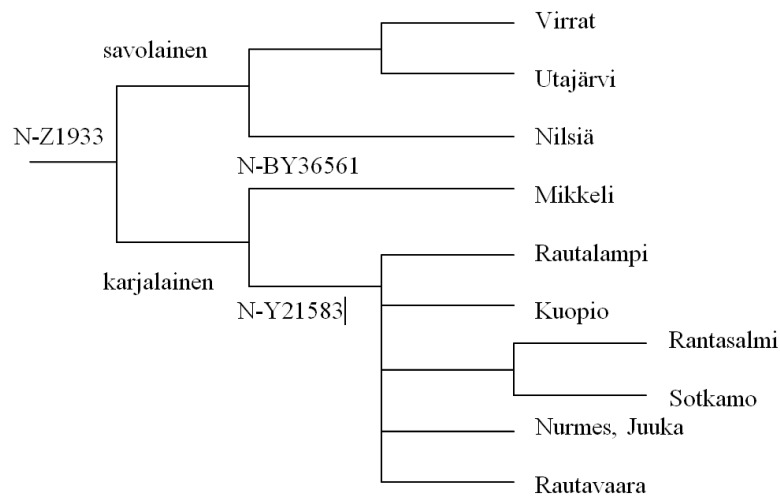
Jaakko Häkkinen¹¹ nimittää kanta-Korhoseksi mutaation CTS12908 ensimmäistä kantajaa. Tähän haploryhmään kuuluu Rautalammin ja Sotkamon Korhosten ohella Korhosia ainakin Kuopiosta, Nilsistä, Nurmekselta, Rantasalmelta ja Sulkavalta. Mutaation on arveltu syntyneen noin tuhat vuotta sitten, 95 prosentin luottamusväli 750–1300 vuotta¹². Jo kanta-Korhonen on siis saattanut olla sukunimeltään Korhonen, sillä itäsuomalaisista sukunimijärjestelmää pidetään noin tuhat vuoden ikäisenä. Seuraava Rautalammin Korhosiin johtava mutaatio syntyi noin 700 vuotta sitten, vastaava luottamusväli 400–1050 vuotta.



Korhosten sukuhaarat Jaakko Häkkinen mukaan. Ryhmiä luonnehtivat nykyiset paikannimet viittaavat paikkakuntiin, joissa näytteidenantajan vanhin tunnettu esi-isä on asunut. Sukunimien yläpuolella olevat luvut ovat dna-näytteiden tunnusnumeroita (kit numbers).

Noin 1900 vuotta sitten kanta-Korhoseen johtavasta linjasta on eronnut ryhmä, johon kuuluvien Korhosten esi-isiä on asunut Mikkelin seudulla, ja noin 2200 vuotta sitten toinen haara, johon kuuluvia Korhosia on monella suunnalla Savossa ja Pohjanmaalla: Nilsissä, Virroilla, Utajärvellä ja Kauhajoella (kuva). Ihmetystä herättää, miten nämä sukulaiset ovat saaneet saman sukunimen, kun sukunimemme tuskin on 2000 vuoden ikäinen ja kun tämän jälkimmäisen, savolaiseksi haaraksi kutsutun ryhmän arvellaan tulleen Suomeen meren yli Virosta¹³ ja meidän karjalaisen ryhmämme Karjalan kannaksen kautta¹⁴.

Nimityksiä "karjalainen haara" ja "savolainen haara" ei nykyään käytä enää edes haarojen alkuperäinen nimeäjä, mutta niiden käyttö ei ole aivan perusteetonta, koska karjalainen haara on sillä tavalla itäisempi, että sekä Suomi- että Karjala-DNA-projektiin liittyneistä ihmisistä kuuluu karjalaiseen haaraan enemmän itäsuomalaisia kuin länsisuomalaisia ja vastaavasti savolaiseen haaraan enemmän länsisuomalaisia kuin itäsuomalaisia¹⁵.

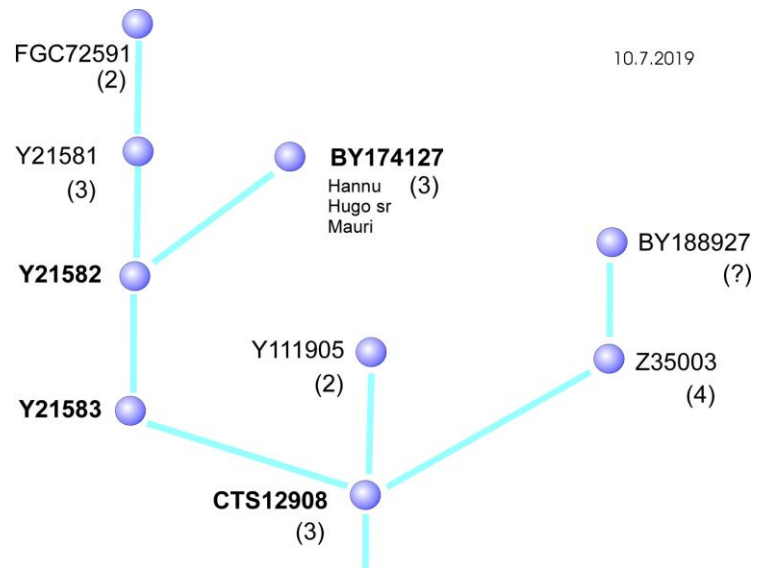


Korhosten Y-dna-linjoja Martti Korhosen mukaan. Tarkempia kaavioita on tekijän omalla kotisivulla¹⁶.

Korhosten perimää on kartoitettu tarkimmin sotkamolaisten osalta. Rautalammin Korhosista on toistaiseksi vain muutamia näytteitä: Ilkka Korhonen, joka on Sonkarinsaaren haaraa, sekä Petri ja Hannu Korhonen sekä edellä jo mainitut Hietapelot, jotka ovat kaikki vaajasalmelaisia. Hannun ja Petrin yhteinen esi-isä on Risto Rekonpoika Korhonen (1700–1776, sukukirjan taulu 15) kuuden sukupolven päässä, geneettinen etäisyys 2/67, mikä tarkoittaa sitä, että 67 merkkijakson testissä vain kahden merkkijakson arvot eroavat toisistaan.

Hannun ja Ilkan yhteinen esi-isä on jutun alussa mainittu Paavo Lemetinpoika Korhonen (1570–1639, sukukirjan taulu 3) kymmenen sukupolven päässä, geneettinen etäisyys 7/67. Hannun haploryhmä on Y21582, joka on ryhmän CTS12908 alaryhmän Y21583 alaryhmä (kuva). Petri kuuluisi luultavasti samaan ryhmään, mutta kun hänestä ei ole tehty yhtä tarkkaa analyysiä, niin nykyinen arvio on CTS12908. Ilkasta ja Hietapelloista ei ole tehty muuta kuin 67 merkkijakson testi, minkä mukaan heidän haploryhmäennusteensa ovat N-M178 ja N-M231. **Alkuperäisen jutun kirjoittamisen jälkeen on saatu BigY-tulokset Hugo Hietapelto seniorista ja Mauri Knuutilasta, molemmat Rautalammin Korhosten vaajasalmelaista päähaaraa..**

Haplopuu voi vielä haarautua Rautalammin Korhostenkin osalta monella tavalla, sillä esimerkiksi Y21582:n jälkeen on tapahtunut 11 mutaatiota, joita ei ole havaittu toistaiseksi kuin Hannun näytteestä eikä siis nimetty vielä. Mutaatio nimetään ja uusi oksa geenisukupuuun syntyy vasta, kun löytyy mutaation toinenkin kantaja. Y21583:n jälkeen syntyneistä mutaatioista on nimetty Sotkamon Korhosten ryhmää määrittävä Y21581. Se on syntynyt noin 600 vuotta sitten. Sotkamon ja Rautalammin Korhosten isälinjat ovat eronneet toisistaan siis 1400-luvulla. Geneettinen etäisyys eli mutaatioerojen määrä Hannun ja joidenkin sotkamolaisten välillä ei ole kuitenkaan sen suurempi kuin Hannun ja Ilkan välillä tai Hannun ja Hietapeltojen välillä eli 7/67.



Pian vuoden 2019 sukulehden ilmestymisen jälkeen valmistui Mauri Knuutilan BigY-tulos. Sen seurauksena Rautalammin Korhoset saivat oman haploryhmän BY174127. Mauri on Hannun ja Hugo Hietapelton tapaan Vaajasalmen päähaaraa. Maurin alahaarasta kertoi Esko Kalliokangas Wanha Tuomas -lehdessä 2006 s. 14–16.

Mitä sitten tulee jutun alun kysymyksiin sukulaisuudesta, niin geenisukututkimuksen anti on kahtalainen. Yhtäältä se avulla voidaan varmistaa muista lähteistä saatuja sukulaisuustietoja ja etsiä tuntemattomia esivanhempia. Voidaan myös seurata isä- ja äitilinjoja kirjallisten lähteiden antamia tietoja kauemmaksi. Yksittäisten henkilöiden nimiä ei silloin tietenkään saada eikä muitakaan henkilötietoja, mutta esimerkiksi se selviää, mistä päin ja sunnilleen milloin esivanhempien suku on tullut Suomeen.

Toisaalta muiden esivanhempien kuin suoran isä- tai äitilinjan jäljet näkyvät kokonaisperimässä. Jos pienetkin samanlaisten perintötekijöiden hippuset otetaan huomioon, niin yhtäkkiä voi saada satoja, tuhansia, jopa kymmeniä tuhansia geneettisiä sukulaisia, osumia (matches). Heistä useimmat ovat tosin niin kaukaisia, että sukulaisuuden jäljittäminen muilla keinoin on työlästä, monesti jopa mahdotonta. Suomessa on paljon sukuja, joiden perimässä on jälkiä ruotsalaisista, saksalaisista, englantilaisista ja monista muista. Voi siellä näkyä merkkejä kaukaisemmistakin ihmisryhmistä, esimerkiksi Afrikasta tai Aasiasta. Oma perimäni on poikkeuksellisesti sataprosenttisesti suomalainen.

Suuret kiitokset jutun syntymisestä kuuluvat Mikkelin Martti Korhoselle, jolta olen saanut paljon neuvoja ja asiatietoja, sekä Lahden seudun sukututkijoiden dna-illoille, jotka ovat avanneet silmäni huomaamaan, miten laajaa geneettinen sukututkimus nykyisellään on sekä taloudellisesti että harrastajamäärien puolesta.

 Lisää tietoa

- ¹ The Journey of Your Past | National Geographic (video) osoitteessa <https://www.youtube.com/watch?v=RGtaq3PiIoU>
- ² *N1C1 Y-DNA Project* osoitteessa <https://www.familyreedna.com/groups/n-1c-1/about/background> tai *Haplogroup N1c* osoitteessa https://www.eupedia.com/europe/Haplogroup_N1c_Y-DNA.shtml.
- ³ *Y-DNA Haplogrou N and its Subclades – 2018* osoitteissa https://isogg.org/tree/ISOGG_HapgrpN.html tai <https://docs.google.com/spreadsheets/d/1GKL9OIEtDEXg8ikSV1tuRSiWqeql0UshJrdaMMcCv6I/edit#gid=0>, viitattu 25.6.2018.
- ⁴ Tässä yhteydessä NCG viittaa amerikkalaiseen yritykseen *Full Genomes Corporation* <https://www.fullgenomes.com/> ja Y venäläiseen yritykseen *YFULL* <https://www.yfull.com/>
- ⁵ *Haplogroup U5 (mtDNA)* sivulla https://www.eupedia.com/europe/Haplogroup_U5_mtDNA.shtml tai *PhyloTree.org - mtDNA tree Build 17 (18 Feb 2016)* osoitteessa <http://www.phyloree.org/tree/index.htm>, viitattu 25.6.2018.
- ⁶ *Autosomal DNA* osoitteessa https://isogg.org/wiki/Portal:Autosomal_DNA ja *The Shared cM Project 3.0 tool v4* osoitteessa <https://dnapainter.com/tools/sharedcmv4>, viitattu 25.6.2018.
- ⁷ Esimerkiksi *Finland DNA Project - Y-DNA Colorized Chart* osoitteessa <https://www.familyreedna.com/public/Finland?iframe=ycolored>, viitattu 22.3.2018.
- ⁸ *Finland DNA – Suomi DNA* julkinen ryhmä osoitteessa <https://www.familyreedna.com/groups/finland/about/background>, viitattu 26.6.2018.
- ⁹ *Korhonen*. Korhosten DNA-tuloksiin keskittyvä ryhmä osoitteessa <https://www.familyreedna.com/groups/korhonen/activity-feed>, viitattu 8.9.2018.
- ¹⁰ Mikkililäisen Martti Korhosen sähköpostiviesti kirjoittajalle 9.9.2018.
- ¹¹ *Muinainen Suomi* osoitteessa <http://muinainensuomi.foorumi.eu/viewtopic.php?f=12&t=2239&start=180>.
- ¹² *YFull YTree v6.02* osoitteessa om/, viitattu 27.6.2018.
- ¹³ Korhonen, M. (2018) *Korhosten karjalaiset juuret*, s. 35. Omakustanne.
- ¹⁴ Häkkinen, J. *DNA kertoo Korhosista*. Wanha Tuomas 2014 s. 32–48.
- ¹⁵ Ahti Kurrin suullinen ilmoitus Lahden seudun sukututkijoiden DNA-illassa 9.10.2018.
- ¹⁶ *Korhosten DNA-sukututkimustietoja ja -testejä* osoitteessa <http://www.123kotisivu.fi/Korhoant/76104614>, viitattu 15.11.2018.