

# Geenisukututkimus: mitä, miksi ja miten?

Jaakko Häkkinen 18.12.2010

## Sisällys:

1. Käsitteet
2. Omiin testituloksiin tutustuminen
3. Geenisukututkimus testitulosten pohjalta

*Tämän lyhyen oppaan avulla toivon lukijan saavan vastauksia otsikon kysymyksiin. Ensimmäisessä luvussa selitän käytettyjä käsitteitä, joiden täsmällinen ymmärtäminen helpottaa esityksen mukana pysymistä. Toisessa osassa kerron, miten omista tiedoista saadaan mahdollisimman paljon irti ja mitä kaikkea kannattaa tehdä. Kolmannessa osassa esitän omaa isälinjatulostani esimerkkinä käyttäen, kuinka geenisukututkimusta lähdetään tekemään ja kuinka historiantakaiset esivanhemmat saadaan kiskottua esihistorian kollektiivisesta, sameasta puurosta todellisiksi yksilöiksi.*

## 1. Käsitteet

**Haploryhmä** (*haplogroup*) = SNP-tason (*single-nucleotide polymorphism*) perusteella luokitellut ihmiskunnan isälinjojen pääryhmät. Mutaatiot koskevat emäsjärjestystä: yleensä esimerkiksi T muuttuu C:ksi tai päinvastoin, ja G muuttuu A:ksi tai päinvastoin.

**Haplotyyppi** (*haplotype*) = STR-tason (*short tandem repeats*) perusteella tehty yksilöllinen luokitus. Mutaatiot koskevat toistojaksoja, eli 10 voi muuttua 11:ksi tai 9:ksi, tai jopa 12:ksi tai 8:ksi. Geenisukututkimuksessa tarkastellaan pääasiassa tätä tasoa.

**Haplotyypisukupolvi** (*haplotype generation*) = sukulinjan haplotyyppi joka eroaa edellisestä ja seuraavasta haplotyypistä yhden mutaation verran. Yhden mutaation syntyyn kuluva aika vaihtelee, joten kahden peräkkäisen haplotyypisukupolven välissä voi olla yksi todellinen sukupolvi tai kymmenen todellista sukupolvea.

**Sukulinja** (*genealogical lineage*) = geenisukututkimuksessa suora linja esi-isän haplotyypistä jälkeläisen haplotyyppiin, kattaen kaikki väliin jäävät haplotyypisukupolvet.

**Markkeri** (*marker*) = merkkitekijä. Geeni tai DNA:n jakso jolla on tietty arvo. Esimerkiksi markerin DYS390 arvo suomalaisilla miehillä on yleensä 24.

**Kaksosmutaatio** (*twin mutation*) = mutaatio, jossa kaksi vierekkäistä locusta muuttuu yhdessä: esimerkiksi DYS464a,b = 14 14 > 15 15. Käsitellään yhtenä mutaationa silloin kun sukulinjassa ei esiinny sellaisia välimuotoja, joissa vain toinen olisi muuttunut.

**Tuplamutaatio** (*double mutation*) = mutaatio, jossa yhden locuksen arvo muuttuu kerralla kaksi askelta: esimerkiksi DYS449 = 27 > 29. Käsitellään yhtenä mutaationa silloin kun sukulinjassa ei esiinny välimuotoja (eli arvoa 28).

**Takaisinmutaatio** (*back-mutation*) = mutaatio, jossa markkeri palaa siihen arvoon, joka sillä oli joskus sukulinjan aikaisemmassa vaiheessa: esimerkiksi  $DYS391 = 10 > 11 > 10$ .

**Itsenäinen rinnakkaismutaatio** (*independent parallel mutation*) = mutaatio, joka tapahtuu kahdessa tai useammassa eri sukulinjassa samanlaisena. Geneettisen etäisyyden perusteella erotetaan, onko kyseessä yhteinen ryhmämutaatio vai itsenäinen rinnakkaismutaatio.

**Ryhmämutaatio** (*group mutation*) = tietylle lähisukuisten haplotyyppien ryhmälle tai alaryhmälle ominainen mutaatio; vanhempi kuin yksilömutaatio.

**Yksilömutaatio** (*individual mutation*) = vain yhdelle ryhmän tai alaryhmän haplotyypille ominainen mutaatio; muiden ryhmien haplotyypeissä saattaa silti esiintyä sama mutaatio (ks. itsenäinen rinnakkaismutaatio).

**Geneettinen etäisyys, GD** (*genetic distance*) = laskennallinen etäisyys; yhden markkerin eroava arvo kahdessa sukulinjassa antaa näiden GD:ksi  $1/67$ .

**Todellinen geneettinen etäisyys, TGD** (*true genetic distance*) = selviää vasta kun sukulinjat juurutetaan haploryhmän sukupuuhun: mutaatiot lasketaan oksia seuraten ja rungon kautta kiertäen eikä ”lunnuntietä” kuten laskennallisessa GD:ssä. Kaksos- ja tuplamutaatiot lasketaan yhdeksi.

**Modaalihaplotyyppi, MoHT** (*modal haplotype*) = lähisukuisten haplotyyppien ryhmän kunkin markkerin yleisimmät arvot sisältävä laskennallinen haplotyyppi (joukon 10, 10, 11, 12, 13 modaali on 10); tällä karsitaan nuoret yksilömutaatiot ja tavoitellaan ryhmän perustajahaplotyyppiä juurruttamista varten.

**Mediaanihaplotyyppi, MeHT** (*median haplotype*) = lähisukuisten haplotyyppien ryhmän kunkin markkerin arvojen keskiluvun sisältävä laskennallinen haplotyyppi (joukon 10, 10, 11, 12, 13 keskiluku on 11); tällä karsitaan nuoret yksilömutaatiot ja tavoitellaan ryhmän perustajahaplotyyppiä juurruttamista varten.

**Rekonstruoitu perustajahaplotyyppi, FHT** (*founder haplotype*) = lähisukuisten haplotyyppien ryhmän rekonstruoitu yhteinen edeltäjä (*ancestor*); esimerkiksi eri markkerien ryhmämutaatioiden levinneisyserot saattavat paljastaa, että tietyn markkerin jokin arvo on todennäköisesti vanhempi kuin jonkin toisen markkerin jokin arvo (laajalevikkisempi mutaatio on vanhempi, ellei ole tapahtunut takaisinmutaatioita). Usein FHT on sama kuin modaali- tai mediaanihaplotyyppi, mutta se voi myös erota näistä (joukon 10, 10, 11, 12, 13 FHT:ssä voi olla esim. arvo 12 tai 13).

**Juurrutettu perustajahaplotyyppi, RFHT** (*rooted founder haplotype*) = RFHT saadaan selville vasta kun ryhmä juurutetaan koko haploryhmän sukupuuhun; RFHT on kaikkein totuudenmukaisin perustajahaplotyyppi, mutta ennen juurruttamista kannattaa selvittää MoHT, MeHT ja/tai FHT.

**Juurruttaminen** (*rooting*) = läheisten haplotyyppien (pieni keskinäinen GD/TGD) ryhmä yhdistetään haploryhmän sukupuuhun silloittamalla se jonkin jo juurrutetun ryhmän haplotyyppiin siten, että muodostuu aukoton sukulinja (jokaisen haplotyyppisukupolven välissä on vain yksi mutaatio). Jos TGD on pienimmilläänkin esimerkiksi  $2/67$ , joudutaan tyytymään rekonstruoituun silta-haplotyyppiin (jota ei siis ole vielä löytynyt yhdestäkään testatusta).

## 2. Omiin testituloksiin tutustuminen

Y-kromosomista tutkitussa isälinjassa on kaksi erillistä tasoa:

1. **Haploryhmän** taso. Tämä selvitetään luotettavasti SNP-mutaatioista, mutta nykyään sen näkee luotettavasti myös 67 markkerin haplotyypistä. Siksi FamilyTreeDNA ei testaa SNP-mutaatioita ilman eri maksua. Ennustettu haploryhmä on vain N, mutta haplotyypistä näemme että se on N1c1.
2. **Haplotyyppin** taso. Yhteensä testataan 67 STR-markkeria. Kaikki tarkemmat vertailut tehdään tällä tasolla.

FamilyTreeDNA:n sivuille kirjaututtuasi vasemman palkin kohdasta **My Account > User Preferences** pääset täyttämään varhaisimman suoran isälinjan esi-isän tietoja. Sivun lopulla on kohta **Displaying The Most Distant Known Ancestor**, jossa kentässä **Paternal Side** lukee esi-isäsi nimi, jonka olet antanut testiä tilatessasi. Perinteisen sukututkimuksen avulla edistyessäsi voit muuttaa tähän kohtaan aina varhaisemman nimen. Mainitse nimen perässä mielellään myös syntymävuosi ja -paikka, se hyödyttää muita tietokantojen käyttäjiä.

Vasemmassa palkissa on ryhmä **Y-DNA**, jonka valikoista näet suunnilleen kaiken olennaisen:

**Matches** = osumat. Täältä näet lähimmät osumasi; tässä keskitymme 67:n markkerin tarkkuuteen, ja useimmilla suomalaisilla on tällä tasolla osumia muutaman mutaatioaskeleen päässä (GD 0–6). Minulla on vain yksi osuma 67:llä markkerilla: Marttinen (GD 7). 37:llä markkerilla ei ole yhtään, ja 25:llä markkerilla on kolme: testatut nimiltään Oinonen, Johansson ja Wennerstierna (kaikkiin GD 2). 12:llä markkerilla osumia on jo satoja, joten tämä taso voidaan ohittaa.

**Haplotree** = ihmiskunnan Y-kromosomin haploryhmien sukupuu. Kun valitset **Your match**, näet N-haploryhmän sukupuun. Koska enempää SNP-mutaatioita ei ole tarkistettu, linja ei etene N1c1:een jota minä tosiasiaassa edustan.

**Ancestral origins** = esi-isien alkuperä. Täällä näet osumien lukumäärän ja prosentit maittain eri tarkkuuksilla.

**DYS values** = markkerien arvot. **Alleles**-rivillä olevat numerot ovat ne, joita verrataan muihin. Sama numero tarkoittaa että sen markkerin osalta vertailtavilla on sama perimä. Y-Search-tietokantaan syötetään myöhemmin nämä numerot oikeille kohdille.

**Print Certificate** jne. = tulosta todistus. Ei tarvitse tulostaa, mutta **Migration Map** (vaelluskartta) on hyödyllinen katsella. Siitä nähdään, että haploryhmä N on tullut Eurooppaan idästä Kiinan ja Siperian kautta.

Toinen tärkeä vasemman palkin ryhmä on **My Maps** (karttani):

**Plot ancestral locations** = täytä esivanhempien tiedot: nimi, syntymäpaikka ja -aika.

**Maps** = kartat. Paina **CLOSE** niin tietoruutu katoaa. Säädetään kartan vasempaan yläreunaan ensin **Y DNA 67 marker**, ja kartalle ilmestyy kaksi värikoodattua ”pisaraa”. Valkea olen minä, mikä käy ilmi tietokentästä jonka voi sulkea (**Close**). Harmaa on seitsemän askeleen päässä oleva Marttinen, jonka varhaisin esi-isä on Kangasniemen Marttisia 1500-luvulta. Kun vaihdetaan tilaksi **Y DNA 37 marker**, osumia ei näy yhtään, mutta kun vaihdetaan **Y DNA**

**25 marker**, näkyy kaksi keltaista (GD 2/67) pisaraa: Oinonen Karstulasta 1800-luvulta ja Ruotsissa Johansson, jonka varhaisin tunnettu esi-isä on Porkka 1600-luvulta.

Edelleen **Tools**-kohdasta pääsee YSearch-tietokantaan. Sitä kannattaa painaa hiiren oikealla painikkeella ja valita ”Avaa uuteen välilehteen”. Täytä täällä esi-isäsi nimi ja syntymäaika ja -paikka sekä haplotyyppin arvot. Huomaa jättää tyhjäksi yhdellä punaisella tähdellä (\*) merkityt kohdat (paitsi tietysti jos sinulla jostain syystä ne löytyvät haplotyyppistäsi).

Kohdassa **Join Projects** voi liittyä projekteihin. Kannattaa liittyä kaikkiin oman tuloksen kannalta olennaisiin projekteihin, jotta sinut löydetään. Jos satut löytämään oman sukunimesi kohdasta **Surname projects**, liity siihen. Etsi myös kohdasta **Dual geographical projects** F:n alta **Finland DNA** -projekti ja liity siihen. Etsi kohdasta **Y haplogroup projects** oma haploryhmäsi ja liity siihenkin. Kun liityt projekteihin, haplotyyppisi päivittyy automaattisesti projektin tietokantaan. Pääset katselemaan haplotyyppitaulukoita projektin sivulta yläpalkin kohdasta **Results**.

Huom! Oma testauspaketti kannattaa tilata liittymällä SuomiDNA-projektiin, silloin saa jonkinlaisen alennuksen.

## Haploryhmä

Suomen yleisin haploryhmä N1c on syntynyt Etelä-Siperian ja Keski-Aasian tienoilla (suunnilleen Venäjän, Mongolian ja Kiinan rajaseudulla). Sieltä se on levinnyt itään (aina Koreaan ja Beringinsalmelle asti) ja länteen (Fennoskandiaan ja Baltiaan, mutta sitä lännempänä ja etelämpänä se on jo todella harvinainen). Mitään N-haploryhmää ei tavata Amerikan intiaaneilta (poikkeuksena myöhään sinne päätnyt iberialainen ryhmä), joten sen leviäminen on ilmeisesti alle 10 000 vuoden ikäinen tapahtuma. N1c1:n (aikaisemmalta nimeltään N3) ja sen veljeslinjojen levinneisyyttä voi katsoa tämän geenitutkimuksen kartoista: [http://volgagermanbrit.us/documents/Rootsi\\_et\\_al.pdf](http://volgagermanbrit.us/documents/Rootsi_et_al.pdf).

N1 (nykyään N1a), N2 (nykyään N1b) ja N3 (nykyään N1c) ovat N:n jälkeläislinjoja, ja N ja O (yleinen Kiinassa ja Kaakkois-Aasiassa) ovat NO:n jälkeläislinjoja. Itä-Siperian turkinsukuisilla jakuuteilla isälinjaa N1c1 on jopa 90 %:lla miehistä, suomalaisilla 60 %:lla ja balteilla (latvialaisilla ja liettualaisilla) 40 %:lla. Se on yleinen myös kaikilla uralilaisilla (~ suomalais-ugrilaisilla) kansoilla paitsi samojedeilla ja unkarilaisilla. Skandinaviassa N1c1:tä on 10 % luokkaa vain pohjoisessa ja Ruotsin itärannikolla, muualla sitä on vähemmän. Itä-Suomessa N1c1 on Länsi-Suomea yleisempi, noin 70 %, ja Kainuussa jopa 80 %. Tämä johtuu perustajavaikutuksesta, joka on voimakas myöhään asutetuilla alueilla.

Mikäli kuulut johonkin toiseen haploryhmään (seuraavaksi yleisimmät Suomessa ovat I1, R1a1 ja R1b), löydät niistä perustietoa toisesta kirjoituksestani: <http://www.mv.helsinki.fi/home/jphakkin/Isalinjat.html>

### 3. Geenisukututkimus testitulosten pohjalta

Käytän tässä oppaassa esimerkkinä oman isälinjani selvittelyä. Korostan, että isälinjan (tai minkä tahansa muun genomin osan) testauttaminen itsessään ei vielä anna paljonkaan tietoa – varsinainen geenisukututkimus oikeastaan vasta alkaa tulosten valmistumisesta. On omasta vaivannäöstäsi kiinni, kuinka paljon saat irti testituloksestasi.

#### Lähiosumien kerääminen

Kun saat sähköpostiisi ilmoituksen tulosten valmistumisesta, kirjaudu sisään FamilyTreeDNA:n kotisivuilla (tai missä yrityksessä nyt geenitestisi oletkaan teettäneet). Seuraavaksi mennään selämään SuomiDNA-projektin tietokantaa. Ensin valitaan yläpalkista **Projects**, ja sitten etsitään **Dual Geographical Projects** -ryhmästä F:n alta Finland (suora osoite löytyy kirjoituksen lopusta). Projekti kuuluu tähän ryhmään, koska se sisältää sekä isä- että äitilinjat; pelkät isälinjaprojektit löytyvät ryhmästä **Y-DNA Geographical Projects** ja pelkät äitilinjaprojektit ryhmästä **MT Geographical Projects**. Projektin kuvaussivulla painetaan linkkiä **Project Website**. Sitten yläpalkista kohdasta **Y-DNA Results** valitaan **Colorized**.

Näyttöön aukeaa pitkä taulukko suomalaisten haplotyypeistä. Valitaan **Markers**-kohdasta **Y-DNA37**, koska halutaan tarkastella vain tarkimpia (37 ja 67 markkeria) haplotyyppijä. Sitten avataan itse selainohjelman (Firefox, Explorer tms.) – ei siis sivun vaan selaimen! – ylävalikosta **Muokkaa > Etsi**. Sivun alareunaan ilmestyy hakukenttä, johon kirjoitetaan esi-isän sukunimi, tässä tapauksessa **Häkkinen**. Hakukenttä muuttuu punaiseksi, eli sivulla ei ole tällaista nimeä. Rullataan sivun loppuun ja aivan oikeaan reunaan, mistä nähdään että sivuja on kaikkiaan kolme. Siirrytään toiselle sivulle ja toistetaan haku. Häkkinen löytyy, mutta lähellä ei näy aiemmin mainittuja lähiosumia, joten haetaan ne erikseen hakutoiminnolla.

Marttisia on paljon, ja vasta seitsemäs on Kangasniemeltä kuten osumani. Kahdeksas Marttinen on myös Kangasniemeltä ja nimen ja vuosiluvun perusteella vasta hän on se minun osumani. Kirjataan hänen arvonsa aiemmin perustettuun haplotyyppitaulukkoon ja merkitään eroavat kohdat esimerkiksi eri väreillä; näin on helppo laskea silmämääräisesti geneettinen etäisyys. Excel-  
taulukko alkaa siis kehittyä tämän näköiseksi:

FTDNA	YSearch	Nimi	Paikka	Ryhmä	DYS393	DYD390	DYS19	DYS391	DYS385a
187856	63KTH	Häkkinen	Eno PKA	8a1a*	14	24	14	11	11
156107	S4MHM	Marttinen	Kangasniemi	8a1a	14	24	14	11	11

\* = Tämä selviää vasta juurruttamisen jälkeen. Jokaiselle markkerille on tietysti oma sarake, mutta tässä esitän tilanpuutteen vuoksi vain viisi markkeria.

Tämä Marttinen eroaa juuri 7 kohdan verran minusta. Tässä vaiheessa nähdään jo, että minua ja Marttista yhdistää  $DYS464a,b,c,d = 15\ 15\ 15\ 15$ , kun taas useimmilla ympärillemme listatuilla haplotyypeillä tässä kohdassa on 13 13 14 14.

Seuraavaksi etsitään ja taulukoidaan samalla tavoin Oinonen ja Porkka, joilla samoin on kyseisessä kohdassa 15 15 15 15. He löytyvät peräkkäisiltä riveiltä vähän Häkkisen yläpuolelta, mikä on hyvä merkki, koska tietokannassa samoilla arvoilla alkavat haplotyyppit ovat lähekkäin. Nyt voidaan jo päätellä, että mainittu harvinainen lukusarja luonnehtii juuri minun lähiosumiani. Selaan samaa sivua ja löydän muitakin haplotyyppijä, joilla on tämä arvo. Osa tosin poikkeaa selvästi

omasta haplotyyppistäni, mutta pikaisen silmäilyn perusteella löydän vielä neljä haplotyyppiä, jotka muistuttavat omaani muutenkin läheisesti: Kemppaiset Hyrynsalmelta ja Vaalasta, Lampinen Mikkelistä ja yksi nimetön osuma.

Kun olen kerännyt haplotyyppitaulukkoon kaikkien lähisukuisten osumieni arvot, minulla on koossa kahdeksasta haplotyyppistä koostuva ryhmä. Lisään taulukkoon yhden rivin tämän ryhmän yläpuolelle, ja merkitseän siihen *modaalihaplotyyppin* arvot lihavoituna. Tämä merkitsee sitä, että katson jokaisen markkerin kohdalta yleisimmän arvon – tällä tavoin yksilölliset mutaatiot karsiutuvat pois, joten modaalihaplotyyppi muistuttaa suurella todennäköisyydellä ryhmän yhteistä esi-isää (tämä varmistuu myöhemmin). Modaalihaplotyyppin edustus on hyvin selvä, ja se on itse asiassa melkein täysin samanlainen kuin Oinosen ja Marttisen haplotyyppit, mutta CDYa,b on modaalihaplotyyppissä 36 36, kun se Oinosella on 37 37, ja DYS439 on 10 kun se Marttisella on 11. Oma etäisyyteni ryhmämme modaalihaplotyyppiin on 6/67, mutta kun kaksosmutaatiot lasketaan yhtenä, TGD on vain 4/67.

## YSearch-tietokanta

YSearch-tietokannassa voi löytää osumia GD 6:een asti, joten se ei auta vielä oman haplotyyppini kohdalla: osumia ei löydy, koska ero lähimpään eli Marttiseenkin on 7. Nyt kun olen kuitenkin laskenut ryhmäni modaalihaplotyyppin, syötän sen YSearchiin. Valitaan ylävalikosta **Create a new user** ja syötetään arvot. HUOMIO! Jätä yhdellä punaisella tähdellä (\*) merkityt kohdat tyhjiksi (-); muuten järjestys on sama kuin FamilyTreeDNA:n tietokannoissa. Lisäksi on valittava salasana, jolla uutta tunnusta pääsee muokkaamaan, ja annettava yhteyshenkilön nimi ja sähköpostiosoite. Kohtaan **Last name** kannattaa kirjoittaa selvennys tyyliin **Modal (Häkkinen, Oinonen, Marttinen, Lampinen, Porkka, Kemppainen)**. Vahvistamisen jälkeen saat tunnuksen, jossa on viisi kirjainta tai numeroa (minulla **D8UYB**); lisää tämä haplotyyppitaulukon soveltuvaan sarakkeeseen modaalihaplotyyppin eteen.

Valitse tämän jälkeen ylävalikosta **Search for genetic matches**; tunnuksesi on nyt jo valmiiksi valittuna, joten paina **Search**. Määritellään hakuasetukset: **Show users that tested at least 37 of the markers that I did**, ja **Allow: a maximum genetic distance of 6**. Ensimmäisellä rivillä on vertailuhaplotyyppimme (modaali), ja sen alla ovat osumat. GD on Marttiseen 1, Oinoseen 2 (tosin YSearchiin ei ole päivitetty kuin 37 markkeria), Lampiseen 6 ja Häkkiseen 6. Muita osumia ovat Ojanperä (esi-isä Kuitu), Gustason (Yli-Kokko), Mononen, Lammi (Hyytinen); esi-isien nimet, syntymäajat ja -paikat sekä haplotyyppin arvot näkee kun napsauttaa kyseistä tunnusta. Näillä kaikilla uusilla osumilla GD oman ryhmäni modaalihaplotyyppiin on 4–6, ja niitä kaikkia yhdistää DYS464a,b,c,d = 13 13 15 15. Valitsemalla haplotyyppien valintaruudun ja painamalla **Compare** näkee haplotyyppit listattuina allekkain, mikä mahdollistaa eri markkerien arvojen vertailun: uudet osumat ovat muuten hyvin samanlaisia kuin oma ryhmäni, mutta mainitussa kohdassa niissä on lukusarjan 15 15 15 15 sijasta 13 13 15 15.

Palataan vielä taaksepäin hakuasetussivulle. Valitaan tällä kertaa **Allow: maximum genetic distance of 1 per marker compared above 37 markers**. Nyt osumia tulee enemmän; aloitetaan niistä osumista joihin on pienin GD. Keinäseen GD on 7, ja hänelläkin on arvot 13 13 15 15, samoin kuin Jouttelalla, johon GD on 8. Löytyy muitakin, joihin etäisyys on alle 10 ja joilla esiintyvät mainitut arvot. Kun näiden haplotyyppien arvot on lisätty haplotyyppitaulukkoon, käy ilmi että 13 13 -ryhmä jakautuu kahteen haaraan: ensimmäisen haaran modaalihaplotyyppin erottaa minun ryh-

mäni modaalihaplotyypistä ainoastaan mainittu ero (DYS464a,b,c,d = 13 13 15 15 vs. 15 15 15 15), kun taas toisen haaran modaalihaplotyypissä on lisäksi DYS449 = 30 (vs. 31) ja Y-GATA 4H = 12 (vs. 11). Siltahaplotyypikin haarojen väliltä näyttää löytyvän: yhdessä haplotyypissä on arvot 31 ja 12.

Tässä vaiheessa meillä on jo kolme lähisukuista, eri modaalihaplotyyppien ympärille syntyneitä ryhmiä; minun ryhmäni modaalihaplotyypistä 13 13 -ryhmän ensimmäisen haaran modaalihaplotyypin GD on 2 (TGD on 1, koska kyseessä on kaksosmutaatio), ja siitä saman ryhmän toisen haaran modaalihaplotyypin GD ja TGD on 2. Tässä vaiheessa ei vielä tiedetä, mikä on näiden ryhmien polveutumisjärjestys, eli mikä niistä on se vanhin, jonka pohjalta muut ovat syntyneet. Sen verran voidaan jo kuitenkin sanoa, että minun ryhmäni ympäriltä ei löydy muita kuin 13 13 -ryhmän ensimmäinen haara, joten minun ryhmäni sijoittuu N1c-sukupuussa latvaan tai oksan kärkeen. Todennäköisesti vanhin ryhmistä on toisen ryhmän ykköshaara, koska kakkoshaarassa on tapahtunut sellaisia mutaatioita, joita ei löydy minun ryhmästäni eikä toisen ryhmän ykköshaarasta.

Tässä vaiheessa en vielä vie YSearchiin uusia modaalihaplotyyppejä, mutta haluan silti nähdä, löytyykö jommankumman 13 13 -ryhmän haaran takaa vielä uusia ryhmiä. Siksi valitsen jälleen **Search for genetic matches**, ja sitten **Click here to enter any sequence and search by Haplotype**. Sitten syötetään jälleen arvot, ensin kakkoshaarasta: 37 ensimmäistä markkeria riittää, ja alhaalta valitaan **Show users that tested at least 37 of the markers that I did** ja **Allow: a maximum genetic distance of 6**, ja sitten painetaan **Search**.

Aiemmin nähtyjen tulosten lisäksi tulee uusi osuma Norjasta, joka näyttäisi olevan kakkoshaaran jälkeläinen (Sydhagen, GD 6: 12 12 15 15), mutta ei mitään läheistä toisenlaista haplotyyppeä. Palataan jälleen taaksepäin ja muutetaan kaksi eroavaa kohtaa ykköshaaraa vastaaviksi: Y-GATA 4H = 11 ja DYS449 = 31. Muut säädöt pidetään ennallaan. Nytkään ei löydy läheisiä osumia, jotka johtaisivat uusiin ryhmiin.

YSearchissa ei kuitenkaan ole kaikkia tuloksia, vaan osa löytyy pelkästään FamilyTreeDNA:n projektilistoista. Nyt eteenpäin – tai oikeastaan taaksepäin – pääsee ainoastaan juurruttamalla nämä ryhmät N1c-haploryhmän sukupuuhun.

## Juurruttaminen

Ryhmä eli sukupuun oksa on juurrutettava koko haploryhmän sukupuuhun. Tämä tapahtuu vertaamalla ryhmän perustajahaplotyypin (tässä vaiheessa modaalihaplotyypin) geneettistä etäisyyttä koko N1c-haploryhmän perustajahaplotyypin. Jos kahden ryhmän perustajahaplotyyppien välillä on silta, niin yleensä se ryhmä, jonka GD koko haploryhmän perustajahaplotyypin on suurempi, polveutuu siitä ryhmästä, jonka GD on pienempi. Harvinaisessa poikkeustilanteessa voi käydä niin, että nuorempi ryhmä onkin kehittynyt takaisinmutaatioiden kautta lähemmäs N1c-perustajahaplotyyppeä ja siksi näyttää vanhemmalta. Tämä selviää kuitenkin kun juurruttamista jatketaan eteenpäin: jos kahden läheisen ryhmän juurruttaminen sukupuuhun onnistuu vain yhden siltan kautta, joka kulkee näistä nuoremman näköiseen ryhmään, on oletettava vanhemman näköisessä klusterissa tapahtuneen takaisinmutaatioita.

Normaalisti ryhmän juurruttaminen haploryhmän sukupuuhun vaatii paljon työtä, mutta nyt geenisukututkijat voivat hyödyntää minun pohjatyötäni haploryhmän N1c kohdalla. Olen juurruttanut oman ryhmäni ja muut suomalaismiesten yleisimmät ryhmät haploryhmämme sukupuuhun etsimällä siltahaplotyyppejä, jotka eroavat toisistaan vain yhden mutaation verran – tai mikäli tällaista

ei tietokannoista löydy, laskemalla modaali- tai mediaanihaplotyyppin tai rekonstruoimalla perustajahaplotyyppin.

Olen verrannut kolmen lähisukuisen ryhmän modaalihaplotyyppijä FamilyTreeDNA:n Suomi-projektin, N-projektin ja monien muiden maaprojektien (Russia, Poland, Latvia, Lithuania, Scandinavia) haplotyyppitaulukoihin etsien mahdollisimman vähän kyseisistä modaalihaplotyyypeistä eroavia haplotyyppijä. (Vertailtava haplotyyppi kannattaa sijoittaa sivun alareunaan, niin silmät eivät hypi erehdyksessä väärälle riville.) Kuten edellä arvelinkin, toisen ryhmän ykköshaara on se, josta johtaa jatkumo taaksepäin ja josta siis minun ryhmäni ja toisen ryhmän kakköshaara ovat kehittyneet:

## 2. ryhmä, haara 1 (MoHT) [ryhmä **8a1**]

14 24 14 11 11 13 11 12 10 14 14 16 17 10 10 11 12 26 14 19 31 13 13 15 15  
11 11 18 18 14 15 18 18 36 36 12 10 11 08 15 17 08 08 10 08 11 10 12 21 22  
14 10 12 12 17 07 13 20 21 15 12 11 10 11 11 12 11

Kohti sukupuun tyveä päästään tämän haplotyyppin kautta:

## Kuokkala/Limberg, Multia [ryhmä **8a**]

14 24 14 11 11 13 11 12 11 14 14 16 17 10 10 11 12 25 14 19 31 13 13 15 15  
11 11 18 18 14 15 18 18 37 37 12 10 11 08 15 17 08 08 10 08 12 10 12 21 22  
14 10 12 12 16 07 13 20 21 15 14 11 10 11 11 12 11

Muuttuneet arvot 11, 12, 16 ja 14 ovat yksilömutaatioita (koska Kuokkalan lähimmillä osumilla ei ole tapahtunut niitä), joten ne voidaan ohittaa. Pari 37 37 edustaa kaksosmutaatiota, mutta tämä markkeri on niin epävakaa, että se on harvoin sama edes lähisukuisilla haplotyypeillä, joten sekin voidaan ohittaa. Jos Kuokkalalla olisi riittävän lähisukuisia osumia, heidän ryhmänsä modaalihaplotyyppissä näitä muutoksia ei todennäköisesti näkyisi. Eli siltahaplotyyppi ei ole Kuokkala itse vaan hänen oletetun ryhmänsä modaalihaplotyyppi (vaikka valitettavasti mitään ryhmää ei olekaan säilynyt). Tärkeintä on nyt, että arvoa 25 vastaa vertailuryhmämme arvo 26.

Seuraava siltahaplotyyppi on:

## Marttinen, Juva [ryhmä **8**]

14 24 14 11 11 13 11 12 10 14 14 16 17 10 10 11 12 25 14 19 30 13 13 15 15  
11 11 18 18 15 15 20 18 37 37 12 10 11 08 15 17 08 08 10 08 11 10 12 21 22  
14 10 12 12 19 07 13 20 21 15 12 11 10 11 11 12 11

Tässä voidaan jälleen ohittaa yksilömutaatiot 15, 20, 37 37 ja 19. Yksilömutaatioita tulee aina, kun haplotyyppit tai ryhmät ovat eronneet satoja vuosia sitten. Olennainen on arvo 25, joka on sama kuin Kuokkalalla, ja lisäksi arvo 30, joka vie meitä jälleen yhden askeleen lähemmäksi N1c-sukuun tyveä.

Tällä tavoin haplotyyppijä (tai useammin modaalihaplotyyppijä joista yksilömutaatiot on karsittu) siltana käyttäen saadaan juurrutettua käsittelemämme ryhmät N1c-haploryhmän perustajahaplotyyppiin, joka näyttää seuraavalta:

393	390	19	391	385a	385b	426	388	439	389I
<b>14</b>	<b>23</b>	<b>14</b>	<b>10</b>	<b>12</b>	<b>13</b>	<b>11</b>	<b>12</b>	<b>10</b>	<b>13</b>
392	389b (II)	458	459a	459b	455	454	447	437	448
<b>14</b>	<b>16 (29)</b>	<b>17</b>	<b>10</b>	<b>10</b>	<b>11</b>	<b>12</b>	<b>25</b>	<b>14</b>	<b>19</b>
449	464a	464b	464c	464d	460	Ygata4H	YCAIIa	YCAIIb	456
<b>28</b>	<b>14</b>	<b>14</b>	<b>15</b>	<b>15</b>	<b>11</b>	<b>11</b>	<b>18</b>	<b>20</b>	<b>14</b>
607	576	570	CDYa	CDYb	442	438 #	# 531	578	395S1a
<b>15</b>	<b>16</b>	<b>17</b>	<b>36</b>	<b>36</b>	<b>13</b>	<b>10</b>	<b>11</b>	<b>8</b>	<b>15</b>
395S1b	590	537	641	472	406S1	511	425*	413a	413b
<b>17</b>	<b>8</b>	<b>8</b>	<b>10</b>	<b>8</b>	<b>11</b>	<b>10</b>	<b>12</b>	<b>21</b>	<b>22</b>
557	594	436	490	534	450	444**	481	520	446***
<b>14</b>	<b>10</b>	<b>12</b>	<b>12</b>	<b>17</b>	<b>7</b>	<b>13</b>	<b>20</b>	<b>21</b>	<b>14</b>
617	568	487	572	640	492	565			
<b>12</b>	<b>11</b>	<b>10</b>	<b>11</b>	<b>11</b>	<b>12</b>	<b>11</b>			

\*, \*\* ja \*\*\* ovat yleensä näillä kohdilla, mutta YSearchin *compare*-toiminnossa ne sijoittuvat tässä järjestyksessä peräkkäin #-merkkien väliin.

Perustajahaplotyyppi on rekonstruoitu kahdella tavalla vertailemalla: (a) *sisäisesti*, eli vertaillen erilaisia N1c-haplotyyppisiä, painottaen niitä joiden GD on lähimpänä muodostuvaa perustajahaplotyyppiä, sekä (b) *ulkoisesti*, eli vertailemalla veljeshaploryhmiin N1a ja N1b sekä isähaploryhmään N1. Nyrkkisääntönä voidaan pitää, että jos N1c-haploryhmän perustajahaplotyyppin arvoksi on ehdolla kaksi vaihtoehtoa, valitaan se joka on sama tai lähempänä kuin veljes- ja isähaploryhmillä. Automaattisesti ei kuitenkaan voida valita näiden läheisten haploryhmien arvoja N1c-perustajahaplotyyppiin, vaan ainoastaan silloin jos arvo löytyy myös jostain vanhasta N1c-ryhmästä.

## Sukupu

Seuraava sukupuu perustuu haplotyyppisukupolvien jatkumoon: jokainen vaihe eroaa yhden mutaation osalta edellisestä ja seuraavasta vaiheesta. Jotta välttyttäisiin haarojen nimitysten tarpeettomalta paisumiselta (tyyliin **1a1a2a1b1a2c**), olen päättänyt pitää numeroinnin aluksi juoksevana. Toisin sanoen koko haploryhmän N1c perustajahaplotyyppi on **0**, sen jälkeläinen on **1**, sen jälkeläinen on **2**, sen jälkeläinen on **3** ja niin edelleen. Jos jälkeläisiä on useampia, otetaan käyttöön pienet kirjaimet: **3:n** jälkeläisiä ovat esimerkiksi **3a**, **3b** ja **4**. Merkitsemätöntä päähaaraa seuraa se jälkeläinen, joka on synnyttänyt eniten jälkeläisiä. Jos taas **3a:lla** on jälkeläisiä, niitä merkitään **3a1**, **3a2** ja niin edelleen. **3a** ja **4** ovat siis veljeksiä, ja ne onkin alla olevassa taulukossa sisennetty kohdakkain.

Kunkin haaran TGD (*true genetic distance* = todellinen geneettinen etäisyys, erotukseksi laskennallisesta GD:stä) perustajahaplotyyppistä on helposti laskettavissa: ensimmäinen numero merkitsee suoraan erottavien mutaatioiden määrää, ja siihen lisätään yksi per jokainen seuraava kirjain tai numero. Niinpä **5a1b:n** TGD perustajahaplotyyppiin olisi  $5 + 3 = 8$ , ja **8a:n** TGD perustajahaplotyyppiin on  $8 + 1 = 9$ . Yhtäaikaaisesti tapahtuvat kaksosmutaatiot (kuten DYS464a,b = 14 14 > 13 13) ja tuplamutaatiot (kuten DYS449 = 28 > 30) käsitellään yhtenä mutaationa.

## N1c-sukupu:

### 0. Perustajahaplotyyppi

0a. (DYS449 = 28 > 27 | DYS464a,b = 14 14 > 13 13)

0b. (DYS391 = 10 > 11 | DYS449 = 28 > 29)

#### 1. (DYS385a = 12 > 11)

1a. (DYS531 = 11 > 12 | DYS576 = 16 > 15)

1a1. (DYS390 = 23 > 24)

#### 2. (DYS391 = 10 > 11)

2a. (DYS442 = 13 > 14)

2a1. (DYS570 = 18 > 20)

2a2. (DYS391 = 11 > 10)

2a3. (DYS19 = 14 > 15) Balttilainen

2a3a. (DYS385b = 13 > 14) Keskieurooppalainen

2a3b. (DYS389-I = 13 > 14 | DYS456 = 14 > 13) Itäeurooppalainen

#### 3. (DYS389-I = 13 > 14)

3a. (DYS393 = 14 > 13 | DYS19 = 14 > 12 | DYS459a,b = 10 10 > 9 9 |

DYS447 = 25 > 26 | DYS442 = 13 > 14) Iberialainen

3b. (DYS448 = 19 > 20 | DYS449 = 28 > 29 | DYS464a,b = 14 14 > 15 15 |

DYS607 = 15 > 16 | DYS444 = 13 > 14) Turkkilainen

3c. (DYS439 = 10 > 11) Skandinaavinen

3c1. (DYS19 = 14 > 13)

#### 4. (DYS442 = 13 > 12)

#### 5. (DYS449 = 28 > 30)

#### 6. (DYS390 = 23 > 24) Suomalainen

#### 7. (DYS464a,b = 14 14 > 13 13; DYS576 = 16 > 18)

7a. (DYS464c,d = 15 15 > 14 14) [tai 9c.]

7a1. (DYS570 = 18 > 19) [tai 9c1. jne.]

7a1a. (DYS439 = 10 > 11)

7a1b. (DYS19 = 14 > 15)

7a1c. (DYS393 = 14 > 15)

#### 8. (YCAIIb = 20 > 18)

8a. (DYS449 = 30 > 31)

8a1. (DYS447 = 25 > 26)

8a1a. (DYS464a,b = 13 13 > 15 15)

8a1b. (Y-GATA 4H = 11 > 12 | DYF406S1 = 11 > 12)

8a1b1. (DYS449 = 31 > 30)

8a1b1a. (DYS19 = 14 > 15 | DYS389-b = 16 > 17 | DYS458 = 17 > 16)

#### 9. (DYS464c,d = 15 15 > 14 14)

9a. (DYS458 = 17 > 18)

9b. (DYS449 = 30 > 31)

[9c. = 7a. (YCAIIb = 18 > 20)]

[9c1. = 7a1. (DYS570 = 18 > 19) jne.]

Ryhmä **0 (0a, 0b)** tavataan harvinaisena Ukrainasta ja Suomesta. Ehkäpä Keski-Aasiassa ja Etelä-Siperiassa esiintyy myös tätä ryhmää, jos haploryhmä N1c on syntynyt siellä. Myös ryhmän **1** jälkeläisiä tavataan harvinaisina lähinnä Suomesta.

Ryhmä **2** edustaa ensimmäistä selvästi ekspansivista vaihetta, ja erityisesti sen eteläinen haara **2a3** jälkeläisineen vastaa käytännössä kaikesta Baltian ja sen eteläpuolisen Euroopan N1c:stä (lukuun ottamatta suppealevikkisiä turkkilaista ja iberialaista haaraa). Balttilaista **2a3**:a tavataan Baltian ja Puolan lisäksi Valko-Venäjältä, Ukrainasta ja Venäjältä. Sen jälkeläislinjoista erottuu läntinen **2a3a** (Saksa, Tšekki, Slovakia, Belgia, Puola, Baltia ja Venäjä; harvinaisena Suomessa) ja itäinen **2a3b** (Puola, Baltia, Ukraina ja Saksa). Pohjoisemmat jälkeläishaarat **2a1** ja **2a2** esiintyvät harvinaisina Suomessa.

Ryhmä **3** on laaja-alainen: sitä tavataan Unkarista, Ukrainasta, Venäjältä, Suomesta ja Norjasta. Sen jälkeläislinjoihin kuuluvat iberialainen **3a** (Espanja, Portugali ja Etelä-Amerikka), turkkilainen **3b** (turkkilaiset ja turkinkuiset kazakhstanilaiset) ja skandinaavinen **3c** (Ruotsi ja Norja). Ryhmä **3** edustaa siis N1c:n laajinta eurooppalaista ulottuvuutta. Sen jälkeläisryhmien sijoittuminen **2a3**:n pohjois- ja eteläpuolelle viittaa siihen, että ryhmä **3** syntyi jossakin Baltian tai Valko-Venäjän tienoilla ja levisi sitten etelään ja pohjoiseen kun ryhmä **2a3** jälkeläislinjoineen syrjäytti sen vanhoilla keskusalueilla.

Myös myyttinen itäskandinaavinen (”esiruotsalainen”) viikinki Rørikr, joka perimätiedon mukaan kutsuttiin hallitsemaan Rus-valtakuntaa (”esi-Venäjä”) 800-luvulla ja tunnettiin siellä nimellä Rurik tai Rjurik, kuului väitettyjen jälkeläistensä geenitestauksen perusteella N1c-haploryhmään ja sukupuuni ryhmään **3**. Hän ei kuitenkaan kuulu skandinaaviseen ryhmään **3c** vaan toiseen ryhmään, jossa DYS458 ja DYS459a,b ovat 18 9 9. Suomenruotsalaisista (jos nimestä ja asuinpaikasta voi päätellä) löytyvät lähimmät osumat rekonstruoidulle Rurikin haplotyypille, mutta Skandinaviassakin on melko läheisiä osumia. Kyseiset suomenruotsalaiset ovat vain GD 2/37 ja 3/37 päässä Rurikista, joten he saattaisivat olla jopa hänen suoria jälkeläisiään.

Ryhmät **4** ja **5** edustavat pullonkaulavaihetta: puuttuvaa lenkkiä **4** edustaa tiettävästi vain yksi puolalainen, ja puuttuvaa lenkkiä **5** yksi venäläinen. Näiden ryhmien levinneisyys on saattanut aiemmin ulottua Suomeenkin, koska ylivertaisesti suurin osa suomalaismiehistä on näiden ryhmien jälkeläisiä ja koska ryhmää **3** tavataan Suomesta muttei esimerkiksi Puolasta. Lisäksi puolalainen ei edusta ryhmä **4**:n päähaaraa vaan jotain sen sivuhaaraa. Mitään muuta siltaa ryhmästä **3** suomalaiseseen ryhmään **6** ei kuitenkaan ole nykyaineiston pohjalta löydettävissä.

Ryhmä **6** rajoittuu Suomenlahden pohjoispuolelle: Suomen lisäksi sitä ja sen jälkeläisryhmiä tavataan vain Skandinaviasta ja Venäjältä. Ryhmät **7**, **8** ja **9** ovat kaikki lisääntyneet Suomessa suuressa määrin. Ryhmää **7** tavataan skandinaaveilta mutta myös venäläisiltä ja saamelaisilta, joten se lienee vanhempi ja on ehkä syntynyt jossakin Laatokan lähistöllä. Ryhmiä **8** ja **9** tavataan skandinaaveilta muttei venäläisiltä eikä toistaiseksi saamelaisiltakaan, joten ne ovat ehkä syntyneet vasta Suomessa. Minun ryhmäni on **8a1a**.

Suomeen ei ole tullut eteläisempiä N1c-haaroja, vaan kaikki tänne päätyneet N1c-ryhmät näyttävät syntyneen täällä tai tulleen idän suunnalta – tosin virolaisten laajempi testaaminen 67 markkerilla voisi auttaa löytämään puuttuvia lenkkejä: nyt Virosta on vain yksi ainoa N1c-haplotyyppi. Aitoa balttilaista haaraa **2a3** tavataan Suomesta vain muutama osuma; mutaatio DYS19 = 14 > 15 on tapahtunut itsenäisesti myös eräissä suomalaisissa alaryhmissä.

Selvittääksesi mihin ryhmään haplotyyppisi kuuluu, sinun tarvitsee vain seurata sukupuuta tyvestä alkaen: jos sinulla on kohdassa DYS385a arvo 11, etene ryhmään **1**, ja niin edelleen.

## Esi-isien paikannus

Kartoittamalla lähimpänä esi-isiä olevien haplotyyppien levinneisyys maakuntatasolla päästään toivottavasti sukulinjan alueellisen leviämisen jäljille. Tässä kannattaa käyttää apuna karttoja, tai niiden puuttuessa Google Maps -palvelua sekä Wikipedian sivustoa *Suomen maakunnat*, missä niihin laskettavat kunnat on listattu. Jos esimerkiksi osuma on Loimaalta, katson karttapalvelusta sen sijaitsevan lounaisessa Suomessa. Kokeilen maakunnista ensin Varsinais-Suomea, ja sieltä Loimaa löytyykin. Entisen Suomen Karjalan (Laatokan Karjala ja Karjalankannas) merkitsen vain KAR (sen kunnat löytyvät eri sivulta), ja muiden maakuntien lyhenteet ovat läpinäkyviä. Lappi voidaan tarvittaessa jakaa viiteen seutukuntaan.

Otan mukaan myös veljesharoja, koska ne voivat auttaa täsmentämään yhteisen esi-isän paikannusta ja ajoitusta. Listaani haplotyyppitaulukostani löytyvät sijaintitiedot (kaikista niitä ei ole annettu):

**8a1a:** Eno PKA, Karstula KSU, Kangasniemi ESA, Mikkeli ESA, Hyrynsalmi KAI, Vaala KAI.

**2 = ESA, KAI**

**8a1b:** Ilmajoki EPM, Oulu PPM, Rantasalmi ESA, Lappajärvi EPM, Liperi PKA, Ylitornio LAP, Peräseinäjoki EPM, Suonenjoki PSA, Saarijärvi KSU, Kokemäki SAT, Laukaa KSU, Kangasniemi ESA, Kangasniemi ESA, Rautalampi PSA, Kangasniemi ESA, Nilsia PSA, Leppävirta PSA, Evijärvi EPM, Joutseno EKA, Kemi LAP, Kuopio PSA, Jääski KAR.

**5 = PSA | 4 = ESA, EPM**

**8a1:** Loimaa VSU, Räisälä KAR, Haapavesi PPM, Hiitola KAR, Suonenjoki PSA, Piikkiö VSU, Tuusniemi PSA, Kerimäki ESA, Raahe PPM, Jämsä KSU, Liperi PKA, Joroinen ESA, Laukaa KSU, Längelmäki PIR

**2 = VSU, KAR, ESA, PSA, KSU, PPM**

## Esi-isien ajoitus

Ajoittamisessa on tärkeintä etsiä mahdollisimman kauas kunkin ryhmän perustajahaplotyyppistä kehittyneet haplotyyppit. Tämä johtuu siitä, että mutaatiot saattavat jäädä tapahtumatta pitkälläkin aikavälillä, joten erottavien mutaatioiden vähyys ei automaattisesti kerro nuoruudesta. Sen sijaan mutaatioiden ilmenemisellä on tietty enimmäisnopeus, joten erottavien mutaatioiden suuri lukumäärä kertoo automaattisesti vanhuudesta.

Omassa ryhmässäni **8a1a** eniten perustajahaplotyyppistä eroamme minä (TGD 4/67 kun kaksosmutaatiot lasketaan yhtenä) ja Vaalan Kemppainen (TGD 5/67). Todellinen geneettinen etäisyytemme on siis 9/67. TMRCA-laskurissa laitetaan nyt kohtaan **Number of markers** arvo 67 ja kohtaan **Number of markers that match** arvo 58 (eli 67 – 9). Sitten kohtaan **Mutation rate** sijoitetaan oikea 67:n markkerin kerroin 0.0028 ja painetaan **Create list**.

Vasemmassa sarakkeessa näkyy kunkin *transitional eventin* (TE) yksittäinen todennäköisyys ja oikeassa sarakkeessa kumuloituva eli kerrostuva todennäköisyys. Nähdään, että suurin yksittäinen todennäköisyys on 0.02298, ja tämä suurin yksittäinen todennäköisyys on sillä vaihtoehdolla, että viimeinen yhteinen esi-isämme eli 51–52 TE:tä sitten. Tämä jaetaan kahdella, jolloin saadaan 26 sukupolvea, mikä sukupolven ollessa 25 vuotta antaa ryhmämme iäksi  $26 \times 25 = 650$  vuotta. Oikeasta sarakkeesta nähdään, että 50 % todennäköisyydellä (0.503) yhteinen esi-isämme eli 55

TE:tä eli 23 sukupolvea eli 575 vuotta sitten, 70 % (0.705) todennäköisyydellä 65 TE:tä eli 33 sukupolvea eli 825 vuotta sitten ja 90 % (0.901) todennäköisyydellä joka tapauksessa 81 TE:tä eli 42 sukupolvea eli 1050 vuotta sitten.

Lasketaan vastaavasti veljesryhmämme **8a1b** TMRCA. Rantasalmen Hiltunen eroaa ryhmänsä perustajahaplotyypistä TGD 7/67 verran ja Tornion Pietari TGD 4/67 verran. Possakka eroaisi Pietaria enemmän, mutta hän kuuluu samaan alaryhmään kuin Hiltunen, joten hän ei kelpaa: vertailtavien haplotyyppien pitää olla kehittynyt eri suuntiin. Tämän ryhmän suurin sisäinen ero on siis TGD 11/67, mikä antaa hieman vanhemman ajoituksen: suurin yksittäinen todennäköisyys on 64 TE:tä eli 32 sukupolvea eli 800 vuotta sitten. Tämä veljesryhmämme voi siis olla jopa 150 vuotta omaa ryhmääni vanhempi, mikä voi osaltaan selittää myös levinneisyyksien eron.

Kun sitten lasketaan isäryhmän **8a1** ikää, ei luonnollisestikaan voida rajoittua tuon ryhmän haplotyyppisiin (mikä antaisi aivan liian nuoren ajoituksen), vaan sen kahden jälkeläisryhmän kaivimmaksi mutatoituneita haplotyyppisiä verrataan sen juurutettuun perustajahaplotyyppiin. Näin vertailussa ryhmän **8a1a** Kemppainen eroaa ryhmän **8a1** RFHT:stä TGD 6/67 verran ja ryhmän **8a1b** Hiltunen eroaa siitä TGD 9/67 verran. Näiden kahden ääripään TGD on siis peräti 15/67, mikä antaa todennäköisimmän yksittäisen ajoituksen 90 TE:tä eli 45 sukupolvea eli 1125 vuotta sitten. Esi-isä **8a1** näyttäisi siis eläneen suunnilleen viikinkiajalla, vuosivälillä 800–1050 jaa.

Tämän kauemmas taaksepäin ei päästä ennen kuin muistakin **8a:n** jälkeläishaaroista kuin **8a1**:stä on kerätty haplotyyppisiä. Mutta tämä aikasyvyys riittää jo pitkälle esihistoriaan.

## Tulosten yhdistäminen tunnettuun historiaan ja esihistoriaan

Alustavasti voidaan sanoa, että esi-isä **8a1** näyttää eläneen jossain päin eteläistä Suomea viikinkiajalla. Jos pohjaksi otetaan se, mitä tiedetään suomalaisten heimojen varhaisista väestöliikkeistä, niin hänet voidaan sijoittaa viikinkiaikaiseen Lounais-Suomeen. Hänen jälkeläisiään (ryhmää **8a1**) on jäänyt tänne (VSU, PIR), mutta heitä on myös levinnyt yhtäältä pohjoiseen Perämeren rannoille (PPM) ja toisaalta itään (KAR).

Juuri viikinkiaikana länsisuomalaisia (muinaishämäläisiä) ehti Perämeren rannoille asti, missä skandinaavit tuntevat heidät kveeneinä. Viikinkiajan alussa he vaikuttavat myös Laatokan rannalla Vuoksen suuseudulla, missä heidän ja alkuväestön yhteisvaikutuksesta syntyy kukoistava Muinais-Karjala; muinaisvenäläisissä lähteissä Käkisalmi tunnetaan nimellä Korela, joka liittyy Karjalanimitykseen.

**8a1:n** jälkeläisryhmistä vanhempi, **8a1b**, syntyy suunnilleen ristiretkiajalla (karkeasti välillä 1050–1300 jaa.) todennäköisesti Karjalassa (yksi karjalainen osuma). 1100-luvun tienoilla Mikkelin alue karjalaistuu arkeologisesti, ja tästä tapahtumasta syntyy savolainen heimo. Savolaisten laajamittainen levittäytyminen Etelä-Savosta pohjoispuolisiin suuntiin alkoi 1400-luvulla, ja savolaisekspansio selittääkin hyvin ryhmän **8a1b** levinneisyyden: sitä tavataan eniten Pohjois- ja Etelä-Savosta sekä Etelä-Pohjanmaalta; jälkimmäinen esiintymäalue liittyy savolaiskiilaan, joka tunkeutui Etelä- ja Keski-Pohjanmaan väliin. Keuruu ja Evijärvi kuuluvatkin savolaiskiilan alueeseen.

**8a1:n** jälkeläisryhmistä nuorempi, oma ryhmäni **8a1a**, syntyy keskiajalla (noin aikavälillä 1300–1500 jaa.). Sitä ei tavata Karjalasta, joten se lienee syntynyt vasta Etelä-Savossa, minne sen esimuoto **8a1** levisi joskus 1100-luvun tienoilla alueen karjalaistuessa. Myös **8a1a** ehti osallistua savolaisekspansioon, ja se levisi Keski-Suomeen, Pohjois-Karjalaan ja Kainuuseen.

Näin olen saanut isälinjastani jo paljon täsmällisemmän ja ennen kaikkea elävämmän kuvan. Esivanhempien vaiheista saa vieläkin elävämmän käsityksen, kun tutustuu näitä tuloksia vastaaviin kohtiin Suomen historiassa ja esihistoriassa: millaista oli elämä viikinkiajan Varsinais-Suomessa, missä sokkeloinen saaristo suojasi tehokkaasti skandinaavisilta kauppiastryöväreiltä eli viikingeiltä? Millaista oli elämä kukoistavassa Muinais-Karjalassa Laatokan länsirannalla, josta tehtiin sotaretkiä hämäläisiä vastaan vahvan mutta kristinuskoon painostavan eteläisen naapurin Novgorodin tuke- mana? Millaista oli elämä keskiajan Etelä-Savossa, osana Ruotsin valtakuntaa ja aivan silloisen rauhattoman itärajan tuntumassa?

Vaikka historiallisia dokumentteja esivanhemmistamme ei noin varhaisilta ajoilta olekaan, mahdollistaa geenisukututkimus kuitenkin yksilön tavoittamisen. Emme ehkä tavoita jokaista suku- polvea, mutta ainakin ne jotka vastaavat haplotyyppisukupolvia: ne joissa on tapahtunut jokin mu- taatio, jonka syntyajan ja -paikan voimme edellä esitetyllä tavalla saada selville.

## Sukunimet

Koska sukunimi kulkee yleensä isälinjan mukana, on selvitettävä sukunimen alkuperä ja levinnei- syys: niiden perusteella saatetaan saada lisävalaistusta isälinjan historiaan. Sukunimen Häkkinen taustalla näyttää olevan karjalainen sana *häkki* 'härkä, sonni' (Mikkonen & Paikkala: *Sukunimet*, Otava 2000). Kuusiosaisen *Karjalan kielen sanakirjan* verkkoversion perusteella tätä sanaa ei juuri- kaan tavata Vienasta (siniset murteet, ks. yläpalkista **Kartta**), vaan ainoastaan Etelä-Karjalasta ja Aunuksen Karjalasta (punaiset ja vihreät murteet). Sukunimen alkuperää on siis etsittävä Laatokan ympäristöstä.

Kaksiosaisen *Suomalaisen sukunimikartaston* (Juhani Pöyhönen, SKS 1998 ja 2003) perus- teella Häkkisiä on Karjalassa Laatokan pohjoispuolella ja Viipurin tienoilla, ja Suomessa suurin keskittymä on Etelä-Savossa. Tämä sopii hyvin yhteen edellä saadun kuvan kanssa, vaikka Karja- lasta Etelä-Savoon siirtynyttä esi-isääni (**8a1**:n jälkeen mutta ennen **8a1a**:ta) ei vielä voikaan pitää minkään tietyn sukunimen kantajana.

Katsotaan sukunimikartastoista vielä oman ryhmäni **8a1a** muut sukunimet: Marttinen, Oino- nen, Lampinen, Kemppainen ja Porkka. Ainoa näitä sukunimiä yhdistävä tekijä on keskittymä Ete- lä-Savossa, joten yhteinen esi-isämme näyttäisi tosiaan syntyneen siellä.

## Jatkosuuntaa sukututkimukselle

Geenisukututkimus tarjoaa nyt uuden suunnan myös perinteiselle, historiallisiin dokumentteihin perustuvalla sukututkimukselle. Sukututkimuksen keinoin ei suomalaisten sukujen kohdalla päästä 1500-luvun taakse, mutta voisi olla palkitsevaa selvittää kaikkien samaan ryhmään kuuluvien testat- tujen varhaisin suoran isälinjan esi-isä: kuinka lähelle Etelä-Savoa päästään? Löytyykö jotain muuta yhteistä?

Tätä työtä ei tietenkään tarvitse tehdä itse, vaan Suomen sukututkimusseuran sukututkijaluet- telosta (<http://www.genealogia.fi/linktoold/sukututk/frames.php?type=1&kieli=fi>) voi etsiä jonkun näitä sukuja jo tutkineen.

Jos olet vielä tämän oppaan jälkeenkin eksyksissä, voit kirjoittaa tekijälle osoitteeseen: [jaakko.hakkinen@helsinki.fi](mailto:jaakko.hakkinen@helsinki.fi)

## **Viitattu kirjallisuus:**

Mikkonen, Pirjo & Paikkala, Sirkka: *Sukunimet*. Otava 2000

Pöyhönen, Juhani 1998: *Suomalainen sukunimikartasto 1*. Suomalaisen Kirjallisuuden Seura.

Pöyhönen, Juhani 2003: *Suomalainen sukunimikartasto 2 – Karjalaiset nimet*. Suomalaisen Kirjallisuuden Seura.

## **Viitatus verkkosijainnit:**

FamilyTreeDNA, geenitestejä tekevä yritys: <https://www.familytreedna.com/>

FamilyTreeDNA Projects: <https://www.familytreedna.com/projects.aspx>

FamilyTreeDNA:n SuomiDNA-projekti: <http://www.familytreedna.com/public/Finland/>

YSearch-tietokanta isälinjojen etsimiseen: <http://www.ysearch.org/>

MitoSearch-tietokanta äitilinjojen etsimiseen: <http://www.mitosearch.org/>

Artikkeli jossa on karttoja N-haploryhmistä: [http://volgagermanbrit.us/documents/Rootsi\\_et\\_al.pdf](http://volgagermanbrit.us/documents/Rootsi_et_al.pdf)

Suomalaisten geneettiset isälinjat: <http://www.mv.helsinki.fi/home/jphakkin/Isalinjat.html>

Google Maps: <http://maps.google.fi>

Suomen maakunnat, kuntalista: [http://fi.wikipedia.org/wiki/Suomen\\_maakunnat](http://fi.wikipedia.org/wiki/Suomen_maakunnat)

Luovutettu Karjala, kuntalista: [http://fi.wikipedia.org/wiki/Luovutettu\\_Karjala](http://fi.wikipedia.org/wiki/Luovutettu_Karjala)

TMRCAlaskuri, aika yhteiseen esi-isään: <http://dna-project.clan-donald-usa.org/tmrca.htm>

Sukututkijaluettelo: <http://www.genealogia.fi/linktoold/sukututk/frames.php?type=1&kieli=fi>

Karjalan kielen sanakirja: [http://kaino.kotus.fi/cgi-bin/kks/kks\\_etusivu.cgi](http://kaino.kotus.fi/cgi-bin/kks/kks_etusivu.cgi)