

Toivatko N-miehet suomen kielen?

Jaakko Häkkinen 24.12.2024

Sisälllys	Sivu
Johdanto	1
1. Suomen kielen alkuperä	2
1.1. Kantauralista kantasuomeen	2
1.2. Kantasuomesta suomeen	4
1.3. Kantasuomi levisi monessa aallossa	6
1.4. Kielentasojen ajoitukset perustuvat lainasanakerrostumiin	8
2. Suomalaisten N-isälinjat	9
2.1. Suomalaisten isälinjat	9
2.2. N-Z1936	12
2.3. N-CTS2929	14
3. Lisää alueellista ja ajallista tarkkuutta	15
4. Yhteenveto	18
Kirjallisuus	19

Johdanto

On tunnettua, että suomalaismiehistä noin 60 % edustaa N-isälinjoja, joiden alkuperä on itäisessä Aasiassa. Internet onkin pullollaan väitteitä, joiden mukaan N-isälinjat liittyvät uralilaisiin kieliin. Tätä yhteyttä ei kuitenkaan ole vielä todistettu. Uralilainen kielikunta on kyllä osa suurta euraasialaista N-isälinjojen aluetta, mutta ihan yhtä lailla se on osa suurta euraasialaista R-isälinjojen aluetta. Osittainen alueellinen päällekkäisyys ei tietenkään yksinään riitä todisteeksi, vaan tarvitaan tarkka vastaavuus ajan, paikan ja leviämisen osalta kielten leviämisen jokaisessa vaiheessa.

Osoitan tässä kirjoituksessa, ettei tarkimmalla mahdollisella tarkastelutasolla edes suomen kielen esimuodon leviämislle Virossa Suomeen ole helppoa löytää N-isälinjoista vastinetta. Ainakaan suurella enemmistöllä suomalaismiesten N-isälinjoista ei nähtävästi ole mitään tekemistä kieleemme saapumisen kanssa.

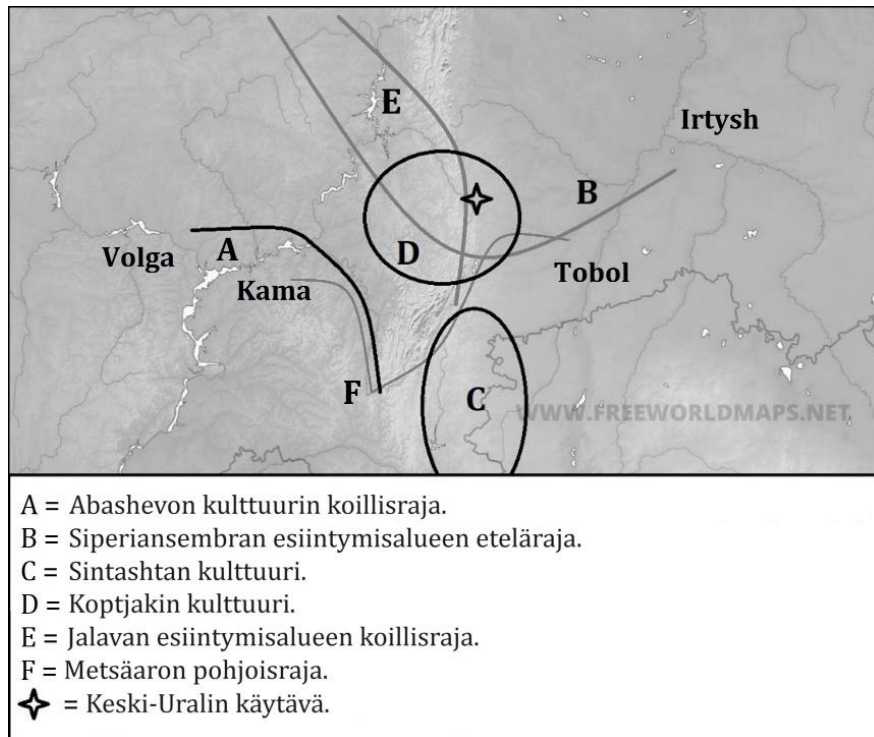
Käytän kirjoituksen kuvissa ajanlaskuun liittyen englanninkielisiä lyhenteitä, jotta vieraskielisetkin lukijat ymmärtäisivät ne oikein:

BCE = Before Common/Christian Era = ennen kristillistä ajanlaskua (eKr., eaa.)

CE = Common/Christian Era = kristillistä ajanlaskua (jKr., jaa.)

1. Suomen kielen alkuperä

1.1. Kantauralista kantasuomeen



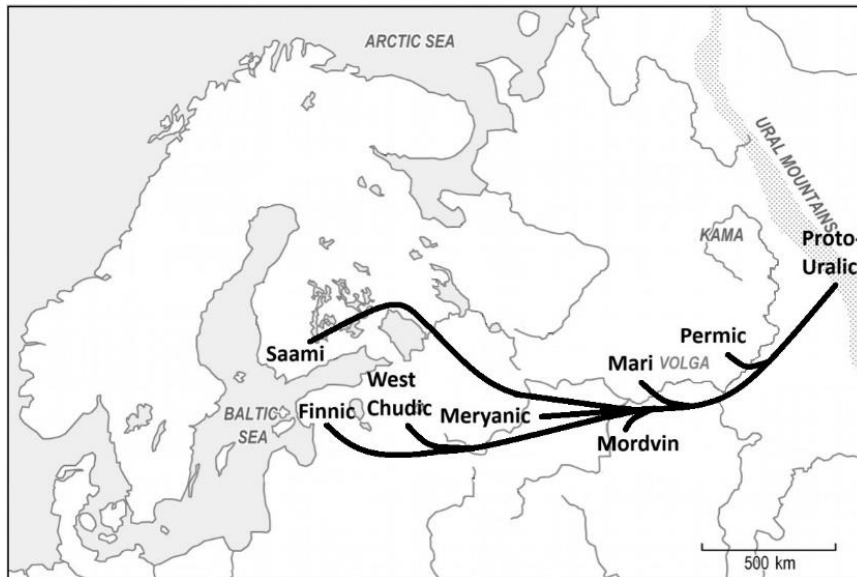
Kuva 1: Kantauralin puhuma-alue (D) ja siihen liittyvät paikannustodisteet.

Kantaurali on lähes aina paikannettu Ural-vuoriston ympäristöön, välillä Euroopan puolelle ja välillä Siperian puolelle. Kun huomioidaan kaikki pätevät todisteet, joilla on merkitystä myöhäiskantauralin paikannuksen kannalta, kielikuntamme alkukoti näyttää sijainneen Uralvuoriston keskivaiheilla, havumetsävyöhykkeen eteläreunalla kartan alueella D. Alueella vaikuttanut Koptjakin kulttuuri on hieman nuorempi kuin kantaurali, mutta sen aikana uralilaisten kielten puhuma-alue ei vielä ollut alkanut laajentua. Tämän alueen lounais- ja eteläpuolella asui varhaisia indoiranilaisia, joilta uralilaiset ovat lainanneet useampia sanakerrostumia, ja juuri täällä kohtasivat idästä levinneet havupuut ja lännestä levinneet jalot lehtipuut (Häkkinen 2023).

Kielten kehitys kantauralista tytärikieliin tunnetaan hyvin, ja kunkin kielentason puhuma-alue voidaan sijoittaa karkeasti tiettyyn aikaan ja paikkaan. Unkari, mansi, hanti ja samojedi näyttävät alkaneen kehittyä Uralvuoriston itärinteillä, kun taas Euroopan puolella uralilaiset kielihaarat ovat muodostuneet nauhamaisesti Kaman alajuoksun ja Volgan yläjuoksun varrelle sekä siitä vielä lännen suuntaan jatkuen. Idästä länteen Euroopan puoleiset uralilaiset kielihaarat ovat permi, mari, mordva, merja, länsitsüudi ja itämerensuomi.

Merjalaiskielistä ja länsitsüudista tiedetään jotain vain Keski- ja Länsi-Venäjän paikannimistöissä säilyneiden aineiden perusteella (Rahkonen 2013). Saame on poikennut tältä reitiltä pohjoisen suuntaan, ja on mahdollista, että muitakin uralilaisia kieliharoja on ehtinyt kehittyä pohjoisempaan ennen katoamistaan. Toisaalta Pohjois-Venäjän

uralilaisperäinen paikannimistö voidaan selittää sujuvasti jo tunnettujen kielihaarojen pohjalta, eikä muita siis välttämättä tarvita (Saarikivi 2006).



Kuva 2: Uralilaisen kielen leviäminen länteen ja kielihaarojen synty.

Kantauralin hajoamisen alkamisesta kului vielä lähes vuosituhat, ennen kuin uralilainen kielialue oli venynyt Keski-Uralilta Ylä-Volgan alueelle. Tämän kauden aikana omaksuttiin varhais-, keski- ja myöhäiskantaindoiranilaisia lainasanoja sekä vielä kantairanilaisia ja hieman myöhäisempiä joskin vielä samanlaisesta lähtökielestä omaksuttuja vanhakan-taisia (arkaaisia) iranilaislainoja (Holopainen 2019; Häkkinen 2023). Lainasanoja näyt-täisi omaksutun myös jostain vanhakantaisesta indoeurooppalaisesta murteesta, jota on perinteisesti nimitetty luoteisindoeuroopaksi ja jonka leviäminen yhdistetään yleisesti nuorakeraamiseen kulttuurialueeseen (Anthony & Ringe 2015). Näiden lainasanojen uu-delleenarviointi on kuitenkin vielä kesken.

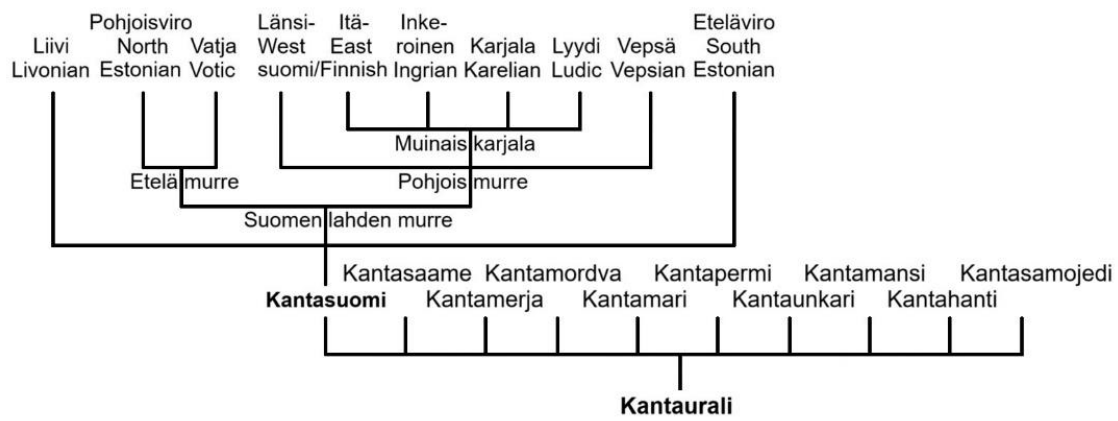
Voidaan olettaa yksi länteen suuntautunut muuttoliike, jonka reitin varrella alkoi sitten kehittyä erilaisia uralilaisia kielihaaroja kuvan 2 osoittamalla tavalla. Länsipään kielihaa-roja puhuttiin kuitenkin hyvin lähekkäin riittävän pitkään, niin että ne jakoivat keskenään yhteisiä muutoksia. Säilyneistä kielihaaroista saamea, itämerensuomea ja mordvaa voi-daankin nimittää länsiuralilaisiksi kieliksi, vaikka länsiurali ei näytäkään olleen varsinaisen välikantakieli vaan pikemminkin läheisten esimurteiden yhteinen kontaktivyöhyke. Tähän viittaa se, että enemmän löytyy sellaisia muutoksia, jotka vain jotkin kaksi kielihaa-raa jakavat, ja vähemmän sellaisia muutoksia, jotka kaikki kolme kielihaa-raa jakavat.

Toisen vuosituhannen eaa. loppua kohti länsiuralilaiset kielihaarat alkoivat levittäytyä, minkä seurauksena ne menettivät kosketuksen toisiinsa. Kaikissa Euroopan puoleisissa eli suomalais-permiläisissä kielihaa-roissa esiintyy paleoeurooppalaisia eli Euroopan tun-temattomista muinaiskielistä lainattuja sanoja. Saameen ja itämerensuomeen näitä sanoja on nähtävästi omaksuttu monessa vaiheessa levittäytymisen edetessä, viimeiset Lapissa noin 1500 vuotta sitten (Aikio 2004; 2012; 2015; 2021; Häkkinen 2009; Zhivlov 2015).

Varhaiskantasuomi-nimitystä käytettiin aikoinaan itämerensuomen ja saamen olete-tusta yhteisestä kantakielestä, mutta nykyään tällä nimityksellä viitataan vain

itämerensuomen varhaisimpien omien muutosten kauteen. Tällöin alkoivat kontaktit kantabaltin ja paleogermaanin kanssa, ja nähtävästi joitain iranilaisiakin lainasanoja vielä omaksuttiin (Kallio 2008; 2015; Junntila 2012; Holopainen 2020). Varhaiskantasuomi ajoitetaan vuoden 1000 BCE ympäristöön (ajoituksista ks. alaluku 1.4). Sitä seurannut keskikantasuomi on pitkä muutoskausi, kun taas myöhäiskantasuomi on viimeinen lyhyt vaihe, jonka aikana muutoksia on vielä levinnyt kaikkien itämerensuomalaisten kielten esimuotoihin. Pelkkä kantasuomi saattaa viitata koko kolmiosaiseen kehitysjatkuun, mutta se saattaa viitata myös pelkästään myöhäiskantasuomeen. Yleensä merkitys selviää lauseyhteydestä.

1.2. Kantasuomesta suomeen



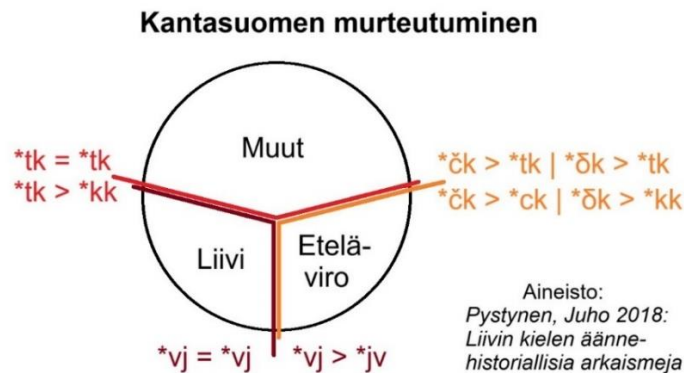
Kuva 3: Uralilainen ja itämerensuomalainen sukupu.

Nykytasolla suomen kieli on yksi itämerensuomalaisista kielistä. Itämerensuomalaisia kieliä lasketaan olevan vähintään yhdeksän: suomi, karjala, inkeroinen, lyydi, vepsä, vatja, pohjoisviro, eteläviro ja liivi. Myös Pohjois-Ruotsin meänkieli ja Pohjois-Norjan kveeni saatetaan laskea nykyään omiksi kielikseen, vaikka kielihistoriallisesti ne kuuluvatkin suomen kielen peräpohjalaismurteiden piiriin.

Kaikkien itämerensuomalaisten kielten yhteinen kantakieli oli kantasuomi. Nimitys on valitettavan hämäävä, koska suomi ei ole yhtään sen tärkeämpi kantasuomen jatkaja kuin muut itämerensuomalaiset kielet. Todellisuudessa myöskään mitään yhtenäistä suomen kieltä ei ole milloinkaan ollut olemassa, vaan tämän nimilapun alle on yhdistetty kaksi eri kehityslinjaa (muinaislänsisuomi ja osa muinaiskarjalan tytärkielistä), joita silloittaa toisiinsa keinotekoinen kirjakieli.

Toisin kuin terminologian perusteella voisi olettaa, kieli ei alkanut murteutua vasta myöhäiskantasuomen jälkeen vaan jo keskikantasuomen kaudella. Myöhäiskantasuomi nimittäin määritteli viimeisten kaikille itämerensuomalaisille kielille yhteisten muutosten perusteella. Viimeiset yhteiset muutokset ovat kuitenkin myöhäisempiä kuin ensimmäiset itämerensuomalaisia kieliä erottavat muutokset – muutoksen ikää ei siis suoraan näe sen levinneisyydestä, vaan iän selvittämiseen on historiallisessa kielitieteessä omat menetelmänsä (Häkkinen 2013). Kantakielen jakautuminen tytärkieliksi on

mutkikas prosessi, jossa tapahtuu sekä eriytymistä (divergenssiä) että yhdentymistä (konvergenssiä).

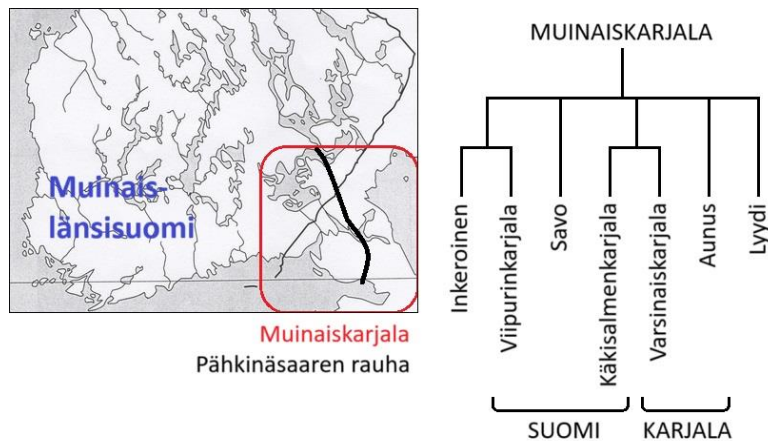


Kuva 4: Kantasuomen varhaisimmat murrerajat

Äännehistorian perusteella kantasuomen ensimmäiset murteutumiset tapahtuivat kielialueen eteläosassa, kun eteläviro ja liivi eriytyivät (Kallio 2007; 2014; Pystynen 2018). Eteläviron syvin sisäinen jakolinja puolestaan erottaa Koillis-Latviassa puhutun Leivun murteen muista (Kallio 2021). Yhtäältä on mahdollista, että kieli olisi murteutunut leviessään pohjoisesta eteläisiin suuntiin, mutta toisaalta on mahdollista, että etelässä tapahtunut murteutuminen heijastaa kielialueen sikäläistä vanhemmuutta suhteessa pohjoiseen kielialueeseen. Murteutumisyjärjestys ei siis itsessään anna suoria vastauksia kielen puhuma-alueesta.

Kantasuomen seuraava jakautuminen tapahtui Suomenlahden ympäristössä: Suomenlahden eteläpuolisesta murteesta kehittyivät pohjoisviro ja vatja, ja Suomenlahden pohjoispuolisesta murteesta kehittyivät muinaislänsisuomi, muinaiskarjala ja muinaisvepsä. Muinaislänsisuomen jatkajia ovat suomen kielen länsimurteet, kun taas muinaiskarjalan jatkajia ovat suomen itämurteet, karjala, inkeroinen ja lyydi. Äännehistoriallisen polveutumisen osalta kuva on kuitenkin monimutkaisempi. Niin karjalan kieli kuin suomen kielen itämurteetkin niputtavat yhteen muinaiskarjalaisen sukupuun eri oksia (Salminen 1998; Kallio 2018). Suomen kielen peräpohjalaismurteita puolestaan pidetään perinteisesti länsi- ja itämurteiden sekoituksena, vaikka Ruotsin ja Norjan alueille jääneissä peräpohjalaismurteissa ei itäistä vaikutusta juurikaan esiinny, vaan ne edustavat puhtaammin muinaislänsisuomen perintöä (Mantila 2000).

Johtuu pelkästään historiallisesta sattumasta, etteivät länsi- ja itäsuomi ole nykyään erillisiä kieliä. Muinaiskarjalan läntisimmät murteet jäivät pitkäksi ajaksi Ruotsin valtakunnan alueelle tai vaikutuspiiriin yhdessä muinaislänsisuomalaisen murteiden kanssa. Yhteydet rajan itäpuolelle jääneisiin muinaiskarjalan murteisiin heikkenivät, kun taas yhteydet länsipuolelle jääneiden erillään olevien murteiden välillä vahvistuivat. Valtiollisen rajan seurannaisvaikutuksina voidaan pitää kielikontaktien suuntaa ja kirjakielen syntyä. Seuranneiden vuosisatojen aikana ruotsin kielen vaikutus kehittyviin suomalaismurteisiin oli voimakasta kielen kaikilla tasoilla, ja yhtä voimakasta oli myös venäjän kielen vaikutus rajan itäpuolelle jääneisiin murteisiin, jotka kehittyivät karjalan kieleksi.



Kuva 5: Muinaiskarjalan tytärikielien ja Pähkinäsaaren rauhan raja (1323).

1500-luvulla kehitetty suomen kirjakieli perustui aluksi vahvasti lounaismurteisiin, mutta myöhemmin siihen otettiin myös itämurteiden piirteitä. Tämä näkyy edelleen kirjakielen sanaston kirjavuutena suhteessa murteiden äännehistorialliseen kehitykseen: osa sanoista päättyi kirjakieleen lounaismurteisessa asussa, osa itämurteisessa asussa. Esimerkiksi sana *vene* edustaa savolaista kehityslinjaa, kun taas sana *perhe* edustaa varsinaissuomalaisista kehityslinjaa.

Kantasuomi > *vars.sm* ~ *savo* ~ *karjala*

*veneh > venhe ~ **vene** ~ veneh

*pereh > **perhe** ~ pere ~ pereh

(*kirjasuomi*)

(*Literary Finnish*)

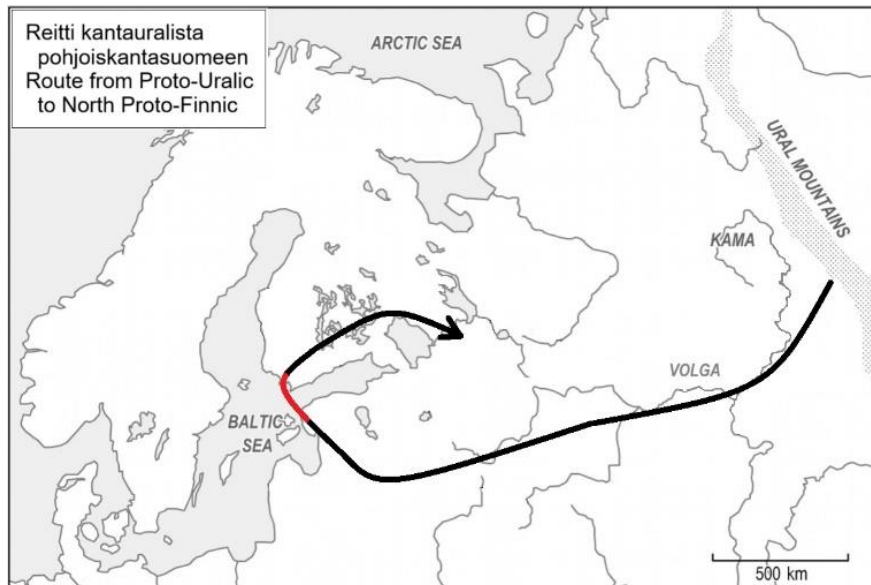
1.3. Kantasuomi levisi monessa aallossa

Paitsi alueellisesti, niin myös ajallisesti suomen kielen kehitystä voidaan pitää monijuurisena. Kantasuomalaiset näyttävät levittäytyneen monessa vaiheessa, ja esimerkiksi Lounais-Suomessa heitä näyttää olleen jo kauan ennen kuin suomen kielen esimuoto levisi Suomenlahden pohjoispuolelle. Esimerkiksi varhaiset germaanisiperäiset paikannimet lounaisessa Suomessa näyttävät polveutuvan jo keskikantasuomesta. Eurajoen alkuosa on lainattu vielä kantagermaanin tasoa edellyttävästä pitkän *ē:n sisältävästä sanasta *ēþrō- 'suoni, joki', ja Aurajoen alkuosa on lainattu sanan hieman myöhemmästä luoteisgermaanisesta asusta *āþrō- (Koivulehto 1987; Kroonen 2013).

Keskikantasuomeen nämä nimielementit omaksuttiin asuissa *ē(e)tra ja *a(a)tra, ja myöhemmin ne kehittyivät säännöllisesti muinaislänsisuomen asuihin *eura ja *aura (sama muinaislänsisuomalainen vokaaliutumisen nähdään esimerkiksi sanoissa *kakra > kaura ~ kakra | *kakla > kaula ~ kakra | *neekla > neula ~ niekla). Saamen kautta näitä nimielementtejä ei ole voitu omaksua, koska silloin lopputulos olisi aivan toisen näköinen. Keskikantasaameen lainattaessa olisi tapahtunut odotuksenmukaisesti sisäkonsonanttien metateesi: **erta ja **arta > myöhäiskantasaamen **eartē ja **vuortē → muinaislänsisuomen **äärtV ja **vuortV. (** = hypoteettinen rekonstruktio.)

Yhtäältä paikannimistön jatkuvuus Lounais-Suomessa jo kantagermaanin kaudesta lähtien (selvästi ennen ajanlaskun alkua) ja toisaalta suomen kielen esimuodon saapuminen Virossa vasta selvästi ajanlaskun alun jälkeen tarkoittaa, että Lounais-Suomen kantasuomalaisen siirtokunnan kieli on vielä pitkään päivittynyt Virossa. Tämä viittaa siihen, että siirtokunta on pysynyt suppea-alaisena ja vähäväkisenä. Leviämisaallot ovat olleet seuraavanlaiset:

- (1) Keskikantasuomi levisi Lounais-Suomeen;
- (2) keskikantasuomi kehittyi Virossa myöhäiskantasuomeksi (joka luultavasti levisi myös Lounais-Suomeen);
- (3) myöhäiskantasuomi kehittyi Virossa kantasuomen Suomenlahden murteeksi, joka levisi Lounais-Suomeen,
- (4) missä se kehittyi Suomenlahden pohjoismurteeksi eli pohjoiskantasuomeksi,
- (5) joka taas kehittyi edelleen muinaislänsisuomeksi, muinaiskarjalaksi ja muinaisvepsäksi levitessään idän suuntaan.



Kuva 6: Pohjoiskantasuomesta puhutaan Suomenlahden ylityksen jälkeen.

Käsittelen seuraavissa alaluvuissa mahdollisia vastineita pelkästään yhdelle etapille kielen leviämisessä, nimittäin Suomenlahden ylitystä (punainen linja kuvan 6 kartassa). Tämän tarkastelun tuloksilla ei ole seurannaisvaikutuksia kielen leviämisen aikaisempiin eikä myöhäisempiin vaiheisiin. Jokaista leviämisvaihetta on tarkasteltava itsenäisesti, koska kielilinjan jokainen leviämisvaihe on alkanut eri aikaan eri alueelta ja erilaisen geneettisen koostumuksen omaavasta väestöstä. On siis mahdollista, että jotkin muut N-isälinjat liittyvät esimerkiksi pohjoiskantasuomen leviämiseen Lounais-Suomesta Laatokalle ja että jotkin muut N-isälinjat liittyvät varhaiskantasuomen leviämiseen Ylä-Volgalta Baltiaan. Nämä etapit vaatisivat kuitenkin omat tarkastelunsa, eikä niistä voida sanoa tässä vaiheessa mitään varmaa.

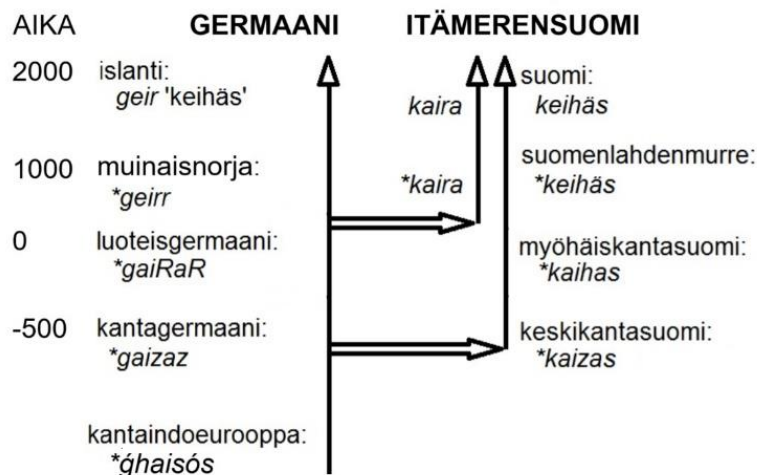
1.4. Kielentasojen ajoitukset perustuvat lainasanakerrostumiin

Lainasanakerrostumien ajoitus

Uralilainen kielikunta	Indoeurooppalainen kielikunta	
	Arkaainen indoeurooppa	3000 BCE
Kantaurali		2500 BCE
	Kantaindoirani	2000 BCE
Länsiurali	Arkaainen irani	1500 BCE
Varhais-kantasuomi	Kantabaltti Paleogermaani	1000 BCE
Keski-kantasuomi	Pohjoisbaltti Kantagermaani	500 BCE
	Luoteisgermaani	100 BCE
Myöhäis-kantasuomi	Kantaskandinaavi	200 CE

Kuva 7: Kielentasot ajoitetaan lainasanoissa nähtävien äänne muutosten avulla.

Toisin kuin pitkään uskottiin, myöhäiskantasuomen taso saavutettiin vasta alle 2000 vuotta sitten. Kielentasojen ajoitukset perustuvat siihen, että eri-ikäisten ja eri indoeurooppalaisista kielihaaroista lainattujen sanojen avulla uralilaisten kielten äänne muutokset saadaan ankkuroitua indoeurooppalaisten kielten aikakehykseen. Varhaisempia lainasanakerrostumia ovat vanhakantainen (luoteis)indoeurooppalainen sekä indoiranilaiset kerrostumat, ja myöhäisempiä balttilaiset ja germaaniset kerrostumat.



Kuva 8: Sama germaaninen sana on lainattu itämerensuomeen kahdessa eri vaiheessa.

Indoeurooppalaisten kielentasojen kronologia on tukevalla pohjalla, koska kirjallisista jäänteistä voidaan seurata kielissä tapahtuneita muutoksia. Germaanisista kieliä alettiin kirjoittaa jo lähes 2000 vuotta sitten, ja varhaisimmat indoeurooppalaisten kielten kirjoitukset ovat lähes 4000 vuoden takaa. Olivatpa sanakerrostumat omaperäisiä tai lainattuja, niin nyrkkisääntö on, että nuorempiin kerrostumiin kuuluu enemmän sanoja kuin vanhempiin, koska vanhoilla sanoilla on ollut pidempi aika kadota kielestä. Esimerkiksi

suomessa on paljon enemmän ruotsalaisia lainasanoja kuin kantaauralista perittyjä sanoja, mutta se ei tietenkään tee suomesta germaanista kieltä.

2. Suomalaisten N-isälinjat

Geneettinen isälinja selvitetään Y-kromosomin DNA:sta. Isälinjoissa tapahtuneita mutatioita voidaan seurata satojentuhansien vuosien taakse. Isälinjoihin viitattaessa käytetään haploryhmiä (esim. I, N, R) tai niiden alahaploryhmiä (esim. I1-L258, N3-Z1936, R1a-Z93). Valitettavasti N-haploryhmässä esiintyy edelleen useita kilpailevia nimityskäytäntöjä, joiden juuret ovat tutkimushistoriassa: N2 vs. N3 = N1b vs. N1c = N1a2 vs. N1a1. Eronteko on siirtynyt nimirimpussa eteenpäin sitä mukaa kun sukupuun tyvestä on löydetty uusia mutaatioita. Suomalaismiesten yleisin haploryhmä on 1990-luvulla Tat-mutaatiosta määritelty Hg16 => N3 => N1c => N1a1, mutta jatkossa käytän vain sellaisia neutraaleja alahaploryhmänimityksiä kuin N-Z1936.

2.1. Suomalaisten isälinjat

Suomalaiset:

Suomalaismiehiä 8357

- joista N-miehiä 5118 (61,2 %)

- joista I-miehiä 2251 (26,9 %)

- joista R-miehiä 846 (10,1 %)

- joista muita miehiä 142 (1,7 %)

Muut:

Latvialaismiehiä 420, joista N-miehiä 94 (22,4 %)

Virolaismiehiä 294, joista N-miehiä 126 (42,9 %)

Saamelaismiehiä 36, joista N-miehiä 22 (61,1 %)

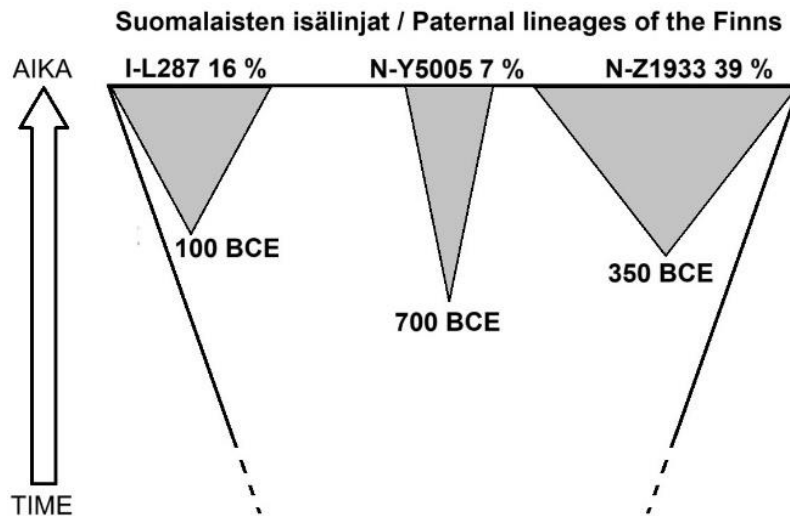
Karjalaismiehiä 29, joista N-miehiä 18 (62,1 %)

(FamilyTreeDNA:n Discover 7.11.2024)

Suomalaisten yleisin Y-kromosomaalinen haploryhmä on N, johon kuuluu 61 % suomalaismiehistä. Seuraavaksi yleisimmät haploryhmät ovat I ja R, joihin kuuluu yhteensä runsas kolmannes suomalaismiehistä. Muihin haploryhmiin kuuluu yhteensä vain alle 2 % suomalaismiehistä. Tämän tason päähaploryhmät ovat kuitenkin kymmeniä tuhansia vuosia vanhoja ja valtavan laajalle levinneitä, joten vertaileminen kieliin on mielekäästä vain paljon nuorempien alahaploryhmien osalta.

Seuraavissa alaluvuissa käsittelen N-haploryhmän alahaploryhmiä Z1936 (41 % suomalaismiehistä) ja CTS2929 (16 % suomalaismiehistä), jotka kattavat yhdessä noin 93 % suomalaismiesten N-isälinjoista ja jotka ovat alle 5000 vuotta vanhoja. Jäljempänä esitetyt mutaatiokuput sekä maiden ja kansojen pitoisuuspalkit perustuvat FamilyTreeDNA:n Discoverin näytteisiin. Vähäisen lukumääränsä vuoksi saamelaiset ja

karjalaiset jätetään seuraavassa tarkastelussa vähemmälle huomiolle; painotus on suomalaisten ja virolaisten isälinjoissa.

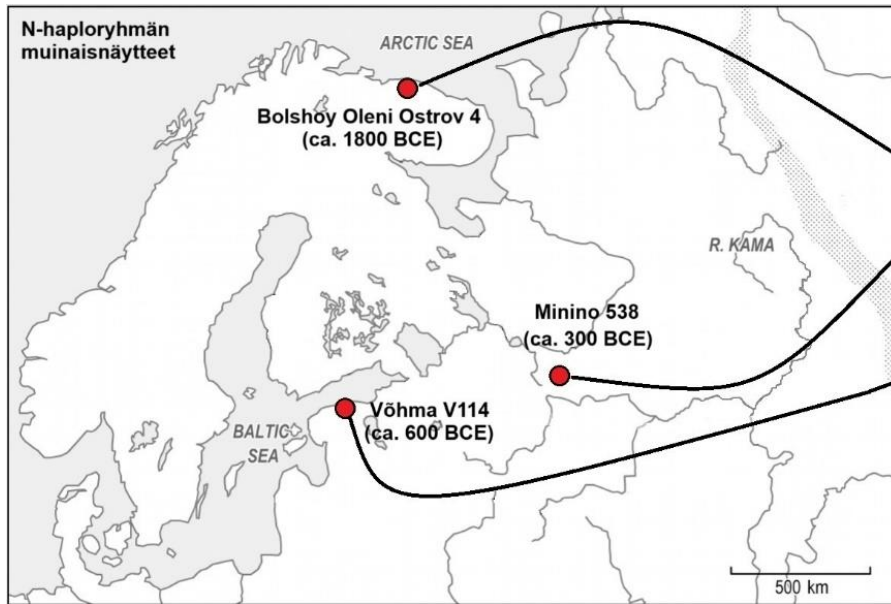


Kuva 9: Suomalaismiesten kolme yleisintä alahaploryhmää.

Kun mennään ajassa taaksepäin runsaat 2000 vuotta, 62 % eli lähes kaksi kolmasosaa nykysuomalaisten isälinjoista palautuu vain kolmeen mieheen. N-Z1933 on N-Z1936:n suurin jälkeläislinja Suomessa, ja N-Y5005 on N-CTS2929:n suurin jälkeläislinja Suomessa. Näiden isälinjojen kohdalla lisääntyminen ja levittäytyminen on ollut erittäin tehokasta. Koko väestö ei ole kasvanut yhtä nopeasti kuin näiden isälinjojen kantajien määrä, vaan näiden isälinjojen suhteellinen osuus on kasvanut syrjäyttämällä muita väestöissä esiintyneitä isälinjoja. Tämä heijastaa luultavasti eroa maatalouden harjoittajien ja metsästäjä-keräilijöiden välillä, koska maatalous mahdollistaa paljon suuremman asukastiheyden. Osasyinä voi olla myös erilaisia sosiaalisia selityksiä, mutta niistä emme voi enää jälkeinpäin saada varmaa tietoa.

Välittömästi kunkin mutaation syntymän jälkeen on kriittinen vaihe, jolloin isälinja saattaa kadota kokonaan (eli mies saa pelkästään tyttäriä tai ei saa lainkaan lapsia) tai jäädä hyvin harvinaiseksi (eli perheen mieslinjaiset jälkeläiset eivät muuta laajemmalle alueelle jatkamaan sukuaan). Edelleen löytyy sellaisiakin suomalaisia isälinjanäytteitä, joiden lähin osuma on eronnut yli 1000 vuotta sitten. Kuitenkin suurin osa harvinaisiksi jääneistä isälinjoista on luultavasti kaikkina aikoina päätenyt sammumaan, mikä tarkoittaa, että mutaatioita on aikoinaan tapahtunut paljon enemmän kuin niitä on säilynyt nykypäivään.

Lähialueiden muinaisnäytteistä on toistaiseksi tiedossa vain muutamia tarkasti analysoituja ja ajoitettuja N-isälinjaa edustavia miehiä. Kunkin löytöpaikan näytteet edustavat omanlaisiaan isälinjoja, ja niiden eriytyminen on tapahtunut jo Uralin takana Siperiassa. Näillä muinaisilla N-miehillä ei kuitenkaan ole juurikaan merkitystä nykyisten pohjois-eurooppalaisten kannalta: Kuolan näytteellä ei ole tiettävästi nykyään jälkeläisiä, ja Viron ja Vologdan näytteiden alahaploryhmiin kuuluu vain hyvin vähän nykymiehiä.



Kuva 10: Lähialueiden N-haploryhmän muinaisnäytteet edustavat eri isälinjoja.

Toisin kuin moni kuvittelee, Kuolaan lähes 4000 vuotta sitten Pohjois-Siperiasta muuttaneet N-miehet eivät ole voineet puhua uralilaista kieltä. Tämä tiedetään siitä, että kielitieteen tulosten perusteella uralilaista kieltä ei ole puhuttu sen enempää oletetun lähtöalueen kuin kohdealueenkaan lähellä vielä tuhansiin vuosiin ja että kun saamelaiskielet levisivät Jäämeren rannalle vasta alle 2000 vuotta sitten, ne syrjäyttivät siellä ei-uralilaisia kieliä. Tämä on varoittava esimerkki siitä, miten pahasti voi mennä metsään, jos epätieteellisesti kuvittelee voivansa nähdä kielen DNA:sta. Saman isälinjan kantajat voivat puhua eri kieliä, ja eri isälinjojen kantajat voivat puhua samaa kieltä. Kielihän ei tietenkään periydy DNA:ssa, vaan se opitaan ympäristöstä.



Kuva 11: Haploryhmä N:n levinneisyysalueella on monia kielikuntia.

Onkin tärkeää ymmärtää, että mikään vuosituhansia vanha haploryhmä ei liity vain yhteen kielikuntaan. Ei ole mielekästä olettaa, että N-miehet liittyisivät esimerkiksi vain uralilaiseen kielikuntaan, kun heitä kuuluu myös ainakin indoeurooppalaiseen, turkkilaiseen, mongolilaiseen, tunguusilaiseen, sino-tiibetiläiseen, korealaiseen, japanilaiseen, jukagiirilaiseen, kamtšatkalaiseen ja eskimo-aleutilaiseen kielikuntaan. N-isälinjojen

kokonaisuus ei siksi voi todistaa yhtään mitään uralilaisten kielten aiemmasta levinneisyydestä.

Me emme todellisuudessa vielä tiedä, mikä oli kantauralin puhujien geneettinen koostumus tai liittyvätkö N-isälinjat uralilaisten kielten leviämiseen kovinkaan varhain ja laajalti vaiko vasta myöhään ja paikallisesti (esimerkiksi isälinja N-BY22114 liittyy todistettavasti sukunimen Häkkinen leviämiseen Etelä-Savosta laajalle alueelle). Me emme tiedä vielä N-isälinjojen roolia edes suomen kielen esimuodon saapumisessa, mutta jäljempänä esitän siihen rooliin yhden ehdokkaan.

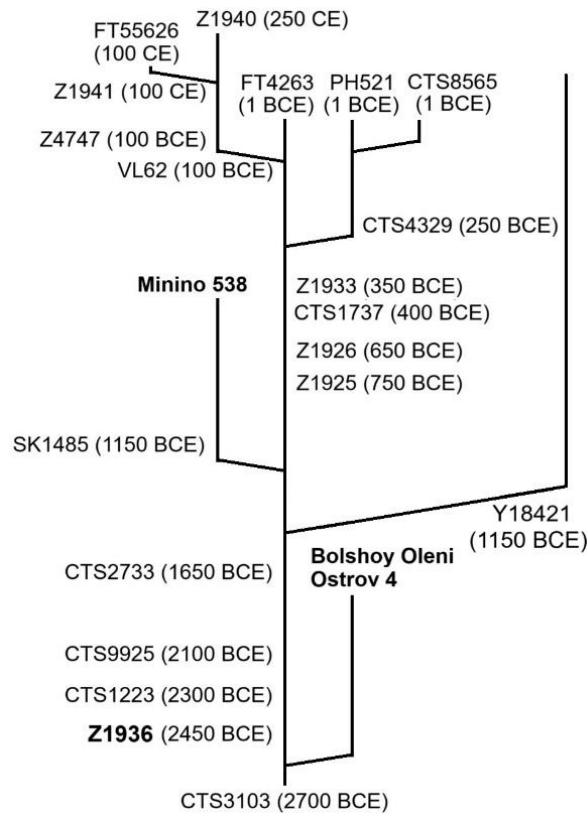
2.2. N-Z1936



Kuva 12: Alahaploryhmän N-Z1936 leviäminen länteen.

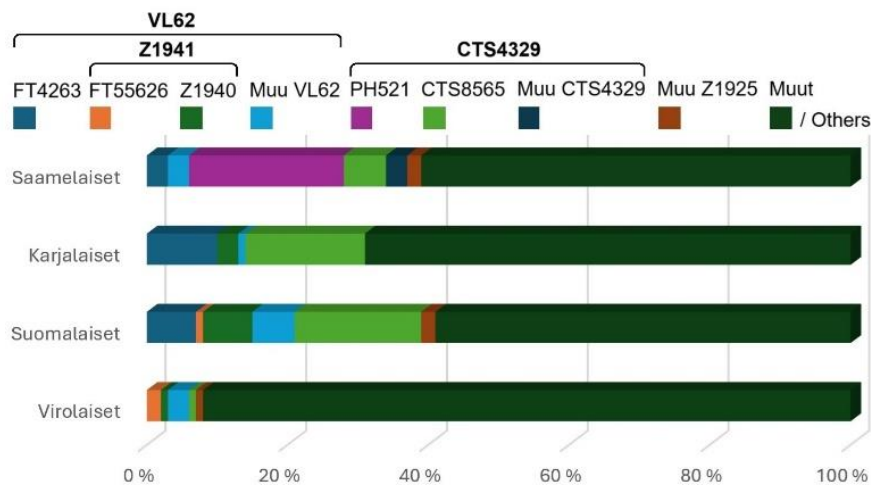
Mutaatio Z1936 on syntynyt jo Uralin itäpuolella Siperiassa; sen varhaisin edustaja löytyy Irtyšin keskijuoksulta Rostovkasta vuoden 2000 BCE tienoilta (Zeng et al. 2023). Tämä on sillä tavalla yksilätyvä alahaploryhmä, että valtava enemmistö sen miehistä kuuluu vasta Suomen lähistöllä syntyneeseen alahaploryhmään Z1925. Matkan varrelle on jäänyt joitain hyvin pieniä alahaploryhmiä kuten Y18421, jota tavataan vain karjalaisilta, ja SK1485, jota edustaa varhaisrautakauden vologdalainen Minino 538 ja kourallinen nykymiehiä, pääosin venäläisiä.

Z1925:n jälkeläislinjatkin keskittyvät selvästi Suomenlahden pohjoispuolelle, ja vasta mutaation Z1936 jälkeen tapahtuu merkittävä haarautuminen kahteen suureen pääryhmään, VL62:en ja CTS4329:än. Virolaisilta löytyy mainittavasti vain VL62:ta. Kuitenkin VL62:n kaikki kolme jälkeläislinjaa näyttävät vielä keskittyvän pohjoiseen, ja näistä vain Z4747:n jälkeläislinja Z1941 lienee syntynyt hieman etelämpänä, ehkä Karjalankannaksella. Sillä on kaksi jälkeläislinjaa, joista kumpaakin on vähän myös virolaisilla.



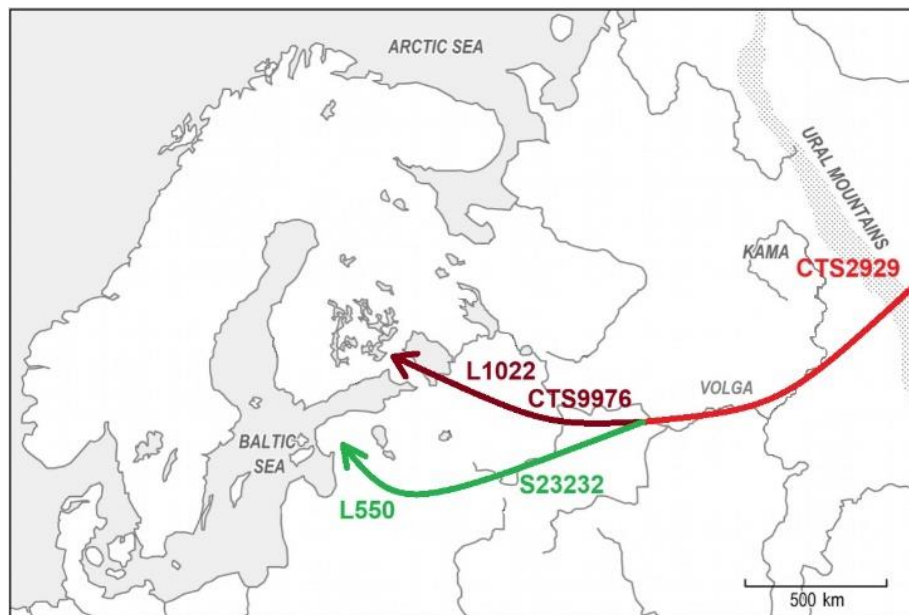
Kuva 13: Alahaploryhmän N-Z1936 SNP-mutaatiosukupuu.

CTS4329:llä on kaksi jälkeläislinjaa. Pohjoisempi PH521 keskittyy nykysaamelaisiin, ja siitä polveutuu myös historiallisen Kemin Lapin sukuja. Eteläisempää CTS8565:tä löytyy vähän myös saamelaisilta, vaikka se onkin nykyään yleisempi suomalaisilla ja karjalaisilla. Kaikilla näillä isälinjoilla on juuret Suomen alueella, eikä niiden setä- ja isosetälinjoja löydy lainkaan venäläisiltä tai virolaisilta. Siksi niistä ei myöskään löydy sopivaa vastinetta kantasuomalaisten Suomenlahden ylitykselle, vaikka teoriassa jotkin alahaploryhmät ovat saattaneet käydä Virossa kantasuomalaistumassa ja palata sen jälkeen takaisin.



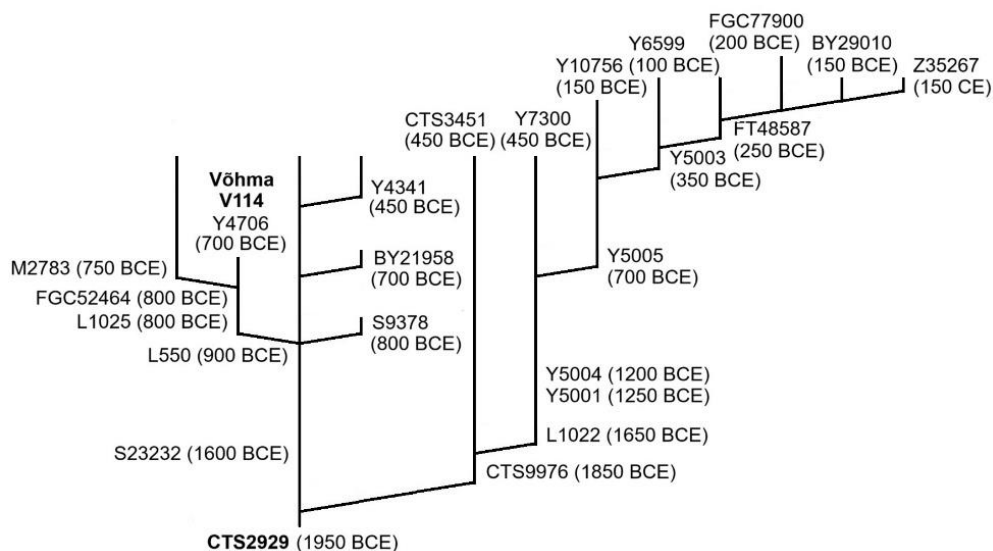
Kuva 14: Alahaploryhmien N-VL62 ja N-CTS4329 pitoisuudet väestöissä.

2.3. N-CTS2929



Kuva 15: Alahaploryhmän N-CTS2929 leviäminen länteen.

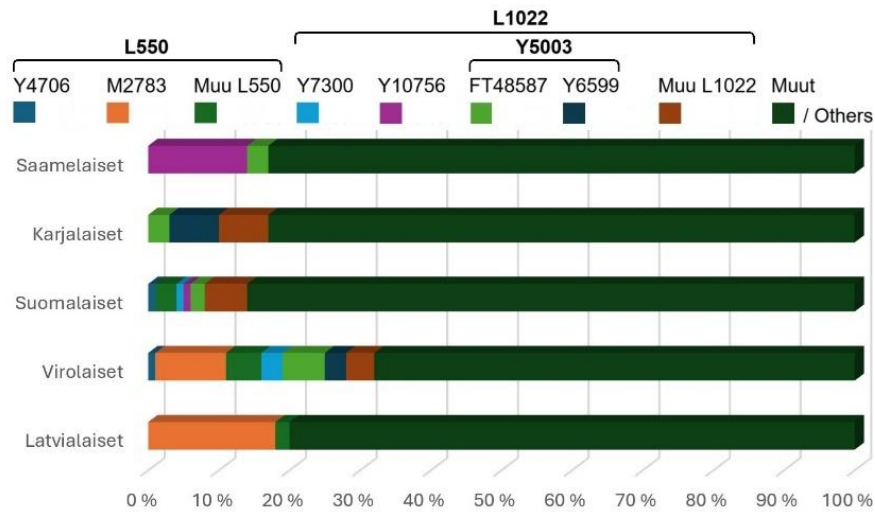
CTS2929 (myös VL29) jakautui jo varhain kahteen linjaan: pohjoisempaa reittiä levisi L1022, ja eteläisempää reittiä levisi L550. L1022:lla on kaksi nuorempaa veljeslinjaa, jotka osoittavat, että myös sen esimuotoa CTS9976 päätyi Itämeren rannoille saakka: CTS3451:tä löytyy saamelaisilta ja suomalaisilta, ja BY6010:a löytyy harvinaisena ruotsalaisilta.



Kuva 16: Alahaploryhmän N-CTS2929 SNP-mutaatiosukupuu.

L1022:n varhaiset jälkeläislinjat keskittyvät Suomenlahden pohjoispuolelle, kunnes mutaation Y5004 jälkeläislinjoissa nähdään sellaisiakin, joita on eniten virolaisissa: Y7300 ja erityisesti vielä nuorempi FT48587. Näiden isälinjojen muutto pohjoisesta etelään on

ollut vielä varhaisempi tapahtuma kuin edellä käsiteltyjen N-Z1936-linjojen muutto pohjoisesta etelään.



Kuva 17: Alahaploryhmien N-L550 ja N-L1022 pitoisuudet väestöissä.

Vaikka joitain erityisesti L550:n jälkeläislinjoja onkin levinnyt Virosta Suomeen, niiden osuus jää täällä hyvin vähäiseksi. On vaikea nähdä, että näin harvinaiset linjat olisivat riittäneet levittämään pohjoiskantasuomen Lounais-Suomeen ja edelleen Karjalaan.

3. Lisää alueellista ja ajallista tarkkuutta

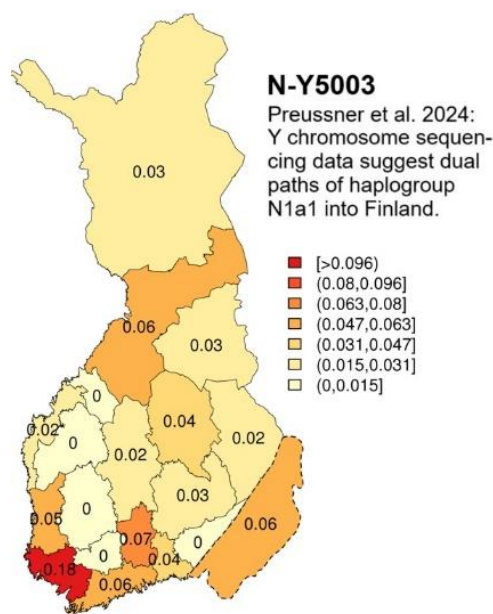
Preussner et al. 2024:

Y chromosome sequencing data suggest dual paths of haplogroup N1a1 into Finland.
– *European Journal of Human Genetics*.

Alueellista tarkkuutta saadaan äsken ilmestyneestä tutkimuksesta, jossa suomalaisten isälinjoja on tarkasteltu maakunnittaisten pitoisuuksien osalta. Ajallista tarkkuutta puolestaan löytyy FamilyTreeDNA:n Discover-työkälystä. Erityisen mielenkiintoinen on alahaploryhmä Y5003, jonka mutaatiopolku on seuraava:

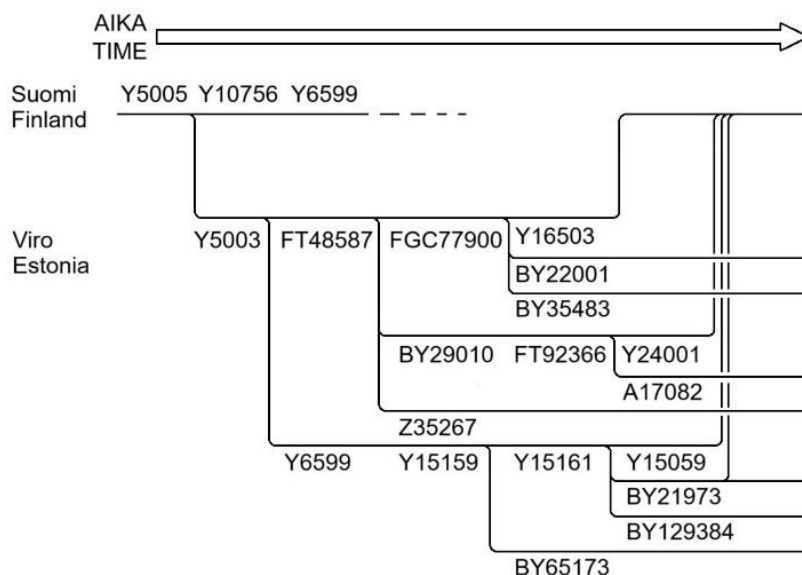
- CTS2929 (1900 BCE Keski-Venäjällä)
- > L1022 (1650 BCE Luoteis-Venäjällä)
- > Y5004 (1200 BCE Etelä-Suomessa)
- > Y5005 (700 BCE Etelä-Suomessa)
- > **Y5003** (350 BCE Pohjois-Virossa)

Suomessa tämän alahaploryhmän esiintymishuippu on Varsinais-Suomessa, peräti 18 % alueen isälinjoista. Valitettavasti edellä mainitussa tutkimuksessa ei ole pitoisuuskarttoja erikseen Y5003:n lukuisille jälkeläislinjoille.



Kuva 18: Alahaploryhmän N-Y5003 pitoisuus Suomen maakunnissa.

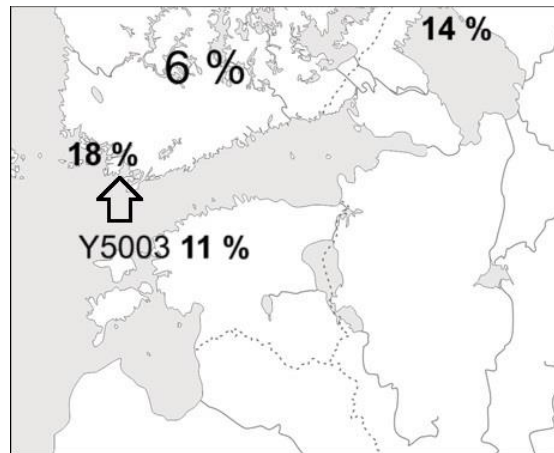
Näyttää kuitenkin siltä, että vaikka sen esimuodot vaikuttivatkin Suomessa, itse Y5003 vaikutti Suomenlahden eteläpuolella. Tähän viittaa se, että sen jälkeläislinjoista useampi esiintyy Virossa kuin Suomessa. Kuitenkin joukko sen jälkeläislinjoja on ensimmäisellä vuosituhanella ajanlaskun alun jälkeen palannut Suomen puolelle, mikä tekee niistä mahdollisia ehdokkaita suomen kielen esimuodon tuojiksi.



Kuva 19: Alahaploryhmän N-Y5003 jälkeläislinjojen levinneisyys ja muutot.

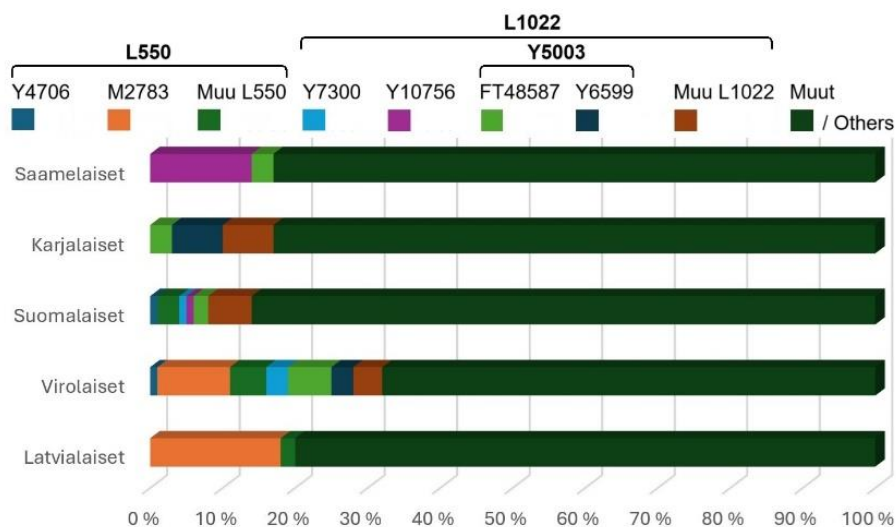
Kun kaikkiin L1022:n jälkeläislinjoihin kuuluu varsinaissuomalaisista miehistä 23 %, niin peräti 18 % kuuluu juuri Y5003:n alle. Koko Suomen osalta pitoisuus on tästä vain kolmasosa eli 6 %, virolaisilla puolestaan 11 %. Karjalaisilla pitoisuus on 14 %, mutta

tämä tarkoittaa Discoverin aineistossa kuitenkin vain viittä miestä, joten siitä ei voi päätellä kovinkaan paljon.



Kuva 20: Alahaploryhmän N-Y5003 pitoisuus Suomenlahden ympäristössä.

Kun lopputuloksesta tiedämme tulokkaiden kielen päässeen Lounais-Suomessa voitolle, niin silloin hyvin todennäköisesti myös heidän isälinjansa ovat olleet menestyksekkäitä ja syrjäyttäneet alueen aikaisempia isälinjoja. Linjan liittyminen pohjoiskantasuomen tuojiin siis selittäisi sen, miksi Y5003-isälinjaston osuus Varsinais-Suomessa on niin suuri. Tässä vielä uudelleen esiintymispalkit:

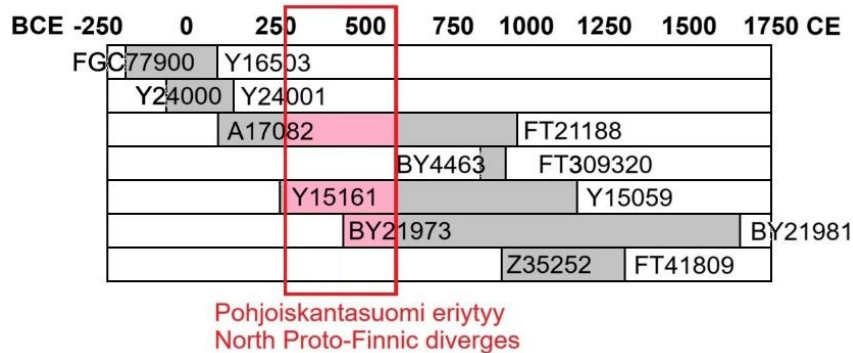


Kuva 21: Alahaploryhmien N-L550 ja N-L1022 pitoisuudet väestöissä.

Lisäksi karjalaisiin on päätynyt Y5003:n jälkeläislinjaa Y6599, ja vaikka tämän osuus suomalaisissa ei saavuta edes yhtä prosenttia, lukumääräisesti suomalaisia miehiä kuuluu tähänkin alahaploryhmään enemmän kuin karjalaisia. Sekin on siis saattanut levitä itään Varsinais-Suomen kautta.

Kuvassa 20 tummennettujen palkkien vasen reuna edustaa sellaisen mutaation syntymäaika, jota löytyy vielä virolaisiltakin, ja oikea reuna edustaa sellaisen jälkeläismutaation syntymäaika, jota ei enää löydy virolaisilta. Palkit osoittavat siis sitä aikahaarukkaa,

jonka sisällä kunkin isälinjan muutto Suomeen näyttää tapahtuneen. Punaisella on merkitty väljästi se ajanjakso, jolloin kielen leviäminen Suomenlahden pohjoispuolelle ja pohjoiskantasuomen kehitys on tapahtunut. Kuvasta nähdään, että näistä Y5003:n jälkeläislinjoista lopulta vain kolme näyttää ajallisesti mahdollisilta ehdokkailta pohjoiskantasuomen tuojien isälinjoiksi: A17082, Y15161 ja BY21973. Mahdollisuus ei kuitenkaan ole vielä todiste, koska kaikilla näillä isälinjoilla siirtymisen aikahaarukka on niin laaja. On hyvinkin mahdollista, etteivät ne todellisuudessa olekaan muuttaneet pohjoiseen juuri oikeaan aikaan.



Kuva 22: Ajallisesti sopivia N-Y5003:n jälkeläislinjoja.

Vaikka siis alahaploryhmä Y5003 kokonaisuutena täyttäisi pitoisuuskriteerin, niin sen jälkeläislinjoista vain kolme täyttäisi ajoituskriteerin: A17082 (12 miestä), Y15161 (23 miestä) ja BY21973 (6 miestä). Kuten nähdään, näihin linjoihin kuuluu yhteensä vain 41 suomalaismiestä, mikä on 0,5 % kaikista suomalaismiehistä. Vaikka käytettäisiin Y5003:n kohdalla nähtyä suhdelukua 3 : 1 Varsinais-Suomen ja koko Suomen välillä, saataisiin näiden isälinjojen yhteispitoisuudeksi Varsinais-Suomessakin vain 1,5 %. Ei vaikuta kovin uskottavalta, että näin vähäinen määrä muuttajia olisi riittänyt levittämään alueelle uuden kielen.

4. Yhteenveto

N-isälinjoja on tullut Suomeen eri aikoina idästä, etelästä ja jopa lännestä. Jotta voisimme yhdistää tietyn isälinjan suomen kielen esimuodon tuojiin, sen pitäisi olla saapunut Vi-rosta Varsinais-Suomeen suunnilleen ensimmäisen vuosituhannen toisen neljänneksen aikana. Sillä pitäisi myös olla riittävän suuri pitoisuus ainakin Varsinais-Suomessa, että sen tuojien kielellä olisi ollut realistinen mahdollisuus päästä voitolle. Kuitenkin mitä tarkem-malla tasolla asiaa katsomme, sitä vaikeampaa on löytää N-isälinjoista uskottavia vasti-neita suomen kielen tuojille. Tässä vaiheessa jäljellä on nähdäkseni enää kaksi mahdolli-suutta:

1. Joko pohjoiskantasuomen tuojat edustivat pääosin ihan muita kuin N-isälinjoja, eli joko I- tai R-isälinjoja tai joitakin äitilinjoi-
2. tai pohjoiskantasuomi onnistui leviämään hyvin vähäisen muuttajajoukon keskuu-desta paljon suuremman väestön äidinkieleksi.

R1a-isälinjoja on virolaismiehillä peräti 24 % ja suomalaismiehillä 5 %. I1-isälinjoja puolestaan on suomalaismiehillä 26 % ja virolaismiehilläkin 14 %. Näistä saattaisi hyvinkin löytyä sopivia vastineita suomen kielen esimuodon tuojille, mutta tähän mennessä niitä ei ole etsitty, koska on jostain syystä haluttu keskittyä N-isälinjoihin.

Äitilinjojen ongelmana on se, että mitokondriaalisen DNA:n hyvin pienen koon vuoksi mutaatioita tapahtuu paljon harvemmin, joten niiden kohdalla ei päästä läheskään yhtä suureen ajalliseen tarkkuuteen kuin isälinjojen kohdalla. Siksi äitilinjojen mahdollista osuutta kielen leviämässä ei ehkä koskaan pystytä todistamaan aukottomasti. Niitäkin kannattaisi kuitenkin tarkastella mahdollisimman suurella tarkkuudella eli laajalla aineistolla maakuntien tasolla ja jokaisen mutaation levinneisyys ja maakunnallinen pitoisuus selvittäen.

On periaatteessa mahdollista, että kielentuojiensa isä- ja äitilinjat olisivat ehtineet sammua Suomessa viimeisten puolentoista tuhannen vuoden aikana. Isälinjahan sammuu, jos mies ei saa poikia, ja äitilinja sammuu, jos nainen ei saa tyttäriä. Heti saapumisen jälkeen todennäköisyys sammumiseen on suurin, kun linjat eivät ole vielä ehtineet levitä laajemmalle.

Sekin on mahdollista, että geneettinen jälki suomen kielen tuojista on löydettävissä enää autosomaaliselta tasolta, esimerkiksi suomalaisten ja virolaisten jakamista IBD-jaksoista (tästä onkin tulossa lähiaikoina uusia tutkimuksia). Muinais-DNA puolestaan edustaa aina vain pientä valikoimaa koko muinaisesta väestöstä, eikä se Suomen alueella ole yleensä säilynyt riittävän hyvin. Siltä suunnalta merkittävää apua ei ole välttämättä tuossa.

Myös Viron maakunnista tarvittaisiin suuret otokset, ja tarkasteluun pitäisi ottaa erikseen aivan jokaisen mutaation pitoisuudet Suomen ja Viron kaikissa maakunnissa. Alueellinen tarkkuus yhdistettynä koko Y-DNA:n kartoittavien testien mahdollistamaan ajalliseen tarkkuuteen voisi auttaa löytämään sopivia vastineita suomen kielen esimuodon tuojille. Valitettavasti on kuitenkin aivan mahdollista, että näin tarkkaa tutkimusta ei kukaan tule milloinkaan tekemään.

Kirjallisuus

Aikio, Ante 2004: An essay on substrate studies and the origin of Saami. – Irma Hyvärinen, Petri Kallio & Jarmo Korhonen (toim.): *Etymologie, Entlehnungen und Entwicklungen. Festschrift für Jorma Koivulehto zum 70. Geburtstag*, s. 5–34. Mémoires de la Société Néophilologique de Helsinki, LXIII. Helsinki: Uusfilologinen yhdistys ry. https://www.academia.edu/4811770/An_Essay_on_Substrate_Studies_and_the_Origin_of_Saami

- Aikio, Ante 2012: An essay on Saami ethnolinguistic prehistory. *A Linguistic Map of Prehistoric Northern Europe*. Suomalais-Ugrilaisen Seuran Toimituksia 266, 63–117. Helsinki: Suomalais-Ugrilainen Seura. https://www.sgr.fi/sust/sust266/sust266_aikio.pdf
- Aikio, Ante 2015: The Finnic ‘secondary e-stems’ and Proto-Uralic vocalism. *Suomalais-ugrilaisen seuran aikakauskirja* 95, 25–66. <https://journal.fi/susa/article/download/82642/41883/>
- Aikio, Ante 2021: The layers of substrate vocabulary in West Uralic. Presentation at “Sub-Indo-European Europe: Problems, Methods and Evidence”, Leiden, Aug 30, 2021. https://www.academia.edu/51152444/Layers_of_Substrate_Vocabulary_in_Western_Uralic
- Anthony, David W. & Ringe, Don 2015: The Indo-European Homeland from Linguistic and Archaeological Perspectives. *Annual Review of Linguistics* 2015. 1: 199–219. <https://www.annualreviews.org/content/journals/10.1146/annurev-linguist-030514-124812>
- Holopainen, Sampsa 2019: *Indo-Iranian Borrowings in Uralic. Critical Overview of the Sound Substitutions and Distribution Criterion*. Doctoral dissertation, University of Helsinki. <https://helda.helsinki.fi/handle/10138/307582>
- Holopainen, Sampsa 2020: Indo-Iranian loanwords in Finnic – a critical overview. *Acta Linguistica Petropolitana* 2020. Vol 16.3, 613–668. https://alp.il-ling.spb.ru/static/alp_XVI_3/19.pdf
- Häkkinen, Jaakko 2009: Kantauralin ajoitus ja paikannus: perustelut puntarissa. *Suomalais-Ugrilaisen Seuran Aikakauskirja* 92, 9–56. <http://www.sgr.fi/susa/92/hakkinen.pdf>
- Häkkinen Jaakko 2013: Kieltenvälinen vertailu historiallisessa kielitieteessä. Kolehmainen, Leena & Miestamo, Matti & Nordlund, Taru (toim.): *Kielten vertailun metodiikka*. Suomalaisen Kirjallisuuden Seura. <https://www.alkuperasivusto.fi/Vertailu.pdf>
- Häkkinen, Jaakko 2023: On locating Proto-Uralic. *Finnisch-Ugrische Forschungen* 68. <https://journal.fi/fuf/article/view/120910>
- Junttila, Santeri 2012: The prehistoric context of the oldest contacts between Baltic and Finnic languages. Grünthal, Riho & Kallio, Petri (toim.): *A Linguistic Map of Prehistoric Northern Europe*. Suomalais-Ugrilaisen Seuran Toimituksia 266, 261–296. https://www.sgr.fi/sust/sust266/sust266_junttila.pdf
- Kallio, Petri 2007: Kantasuomen konsonanttihistoriaa. Jussi Ylikoski & Ante Aikio (editors): *Sámit, sánit, sátnehámit. Riepmočála Pekka Sammallahtii miessemánu 21. beaivve 2007*. Suomalais-Ugrilaisen Seuran Toimituksia = Mémoires de la Société Finno-Ougrienne 253, 229–249. Helsinki: Suomalais-Ugrilainen Seura. http://www.sgr.fi/sust/sust253/sust253_kallio.pdf
- Kallio, Petri 2008: On the ‘Early Baltic’ Loanwords in Common Finnic. Alexander Lubotsky, Alexander & Schaeken, Jos & Wiedenhof, Jeroen (toim.): *Evidence and Counter-Evidence: Essays in Honour of Frederik Kortlandt 1. Balto-Slavic and Indo-European Linguistics*, 265–277. *Studies in Slavic and General Linguistics* 32. Amsterdam

- New York. https://www.academia.edu/1103450/On_the_Early_Baltic_Loanwords_in_Common_Finnic
- Kallio, Petri 2014: The Diversification of Proto-Finnic. In: Joonas Ahola and Frog (eds.), *Fibula, Fabula, Fact: The Viking Age in Finland*, pp. 155-168. *Studia Fennica Historica* 18. Helsinki 2014. https://www.academia.edu/10172149/The_Diversification_of_Proto_Finnic
- Kallio, Petri 2015: The Stratigraphy of the Germanic Loanwords in Finnic. John Ole Askedal, John Ole & Nielsen, Hans Frede (toim.): *Early Germanic Languages in Contact*, 23–38. John Benjamin Publishing Company. https://www.academia.edu/13615139/The_Stratigraphy_of_the_Germanic_Loanwords_in_Finnic
- Kallio, Petri 2018: Muinaiskarjalan uralilainen tausta. https://www.academia.edu/37709596/Muinaiskarjalan_uralilainen_tausta
- Kallio, Petri 2021: The position of Leivu. *ESUKA – JEFUL* 2021, 12–2: 123–143. <https://ojs.utlib.ee/index.php/jeful/article/view/jeful.2021.12.2.05>
- Koivulehto, Jorma 1987: Namn som kan tolkas urgermanskt. *Klassiska problem inom finlandssvensk ortnamnsforskning. Svenska litteratursällskapets jubileumssymposium på Hanaholmen 4–6 oktober 1985*. *Studier i nordisk filologi* 67, 27–42.
- Kroonen, Guus 2013: *Etymological Dictionary of Proto-Germanic*. Leiden Indo-European Etymological Dictionary Series 11. Leiden – Boston: Brill.
- Mantila, Harri 2000: Lapin rajat ja murrerajat. *Virittäjä* 2/2000, 188–202. <https://journal.fi/virittaja/article/view/40001>
- Preussner, Annina et al. 2024: Y chromosome sequencing data suggest dual paths of haplogroup N1a1 into Finland. *European Journal of Human Genetics* 2024. Published online 28 October 2024. <https://doi.org/10.1038/s41431-024-01707-7>
- Pystynen, Juho 2018: Liivin kielen äännehistoriallisia arkaismeja. Esitelmä: XLV Kielitieteen päivät 8.5.2018. https://www.academia.edu/36612432/Liivin_kielen_%C3%A4%C3%A4nnehistoriallisia_arkaismeja
- Rahkonen, Pauli 2013: *Southeastern Contact Area of Finnic Languages in the Light of Onomastics*. Doctoral dissertation, University of Helsinki. <https://helda.helsinki.fi/handle/10138/38908>
- Saarikivi, Janne 2006: *Substrata Uralica. Studies on Finno-Ugrian substrate in northern Russian dialects*. Väitöskirja. Tartu. <http://ethesis.helsinki.fi/julkaisut/hum/suoma/vk/saarikivi/substrat.pdf>
- Salminen, Tapani 1998: Pohjoisten itämerensuomalaisten kielten luokittelun ongelmia. Grünthal, Riho & Johanna Laakso, Johanna (toim.): *Oekeeta asijoo. Commentationes Fenno-Ugricae in honorem Seppo Suhonen sexagenarii*, 390–406. Suomalais-Ugrilaisen Seuran Toimituksia 228.
- Zeng, Tian Chen et al. 2023: Postglacial genomes from foragers across Northern Eurasia reveal prehistoric mobility associated with the spread of the Uralic and Yeniseian languages. Preprint in BioRxiv. doi: <https://doi.org/10.1101/2023.10.01.560332>
- Zhivlov, Mikhail. 2015. Неиндоевропейский субстрат в финно-волжских языках. Käsikirjoitus. <https://www.academia.edu/11689161/>