

# Saamelaiset ja suomalaiset eurooppalaisina

## Saamis and Finns as Europeans

Jaakko Häkkinen, 27.5.2012

Tekeillä on lopultakin genomilaajuisia tutkimuksia, joissa verrataan saamelaisia ja suomalaisia toisiinsa. Nyt onkin viimeinen hetki kokeilla, mitä aiheesta saadaan irti epäsuorasti; onhan niin saamelaisia kuin suomalaisiakin verrattu erikseen muihin eurooppalaisiin väestöihin. Tässä kirjoituksessa esitetyt tulokset on sitten mielenkiintoista jälkikäteen verrata suoran vertailun tuottamiin tuloksiin.

Hyödynnän sekä yhteiset alleelit (ja siten yhteisen geenipohjan) näyttävää IBS-tasoa (*identity by state*) että väestöjen välillä tapahtuneen sekoittumisen tai sen puutteen eli isolaation (ja siten geneettisen etäisyyden) näyttävää FST-tasoa (*fixtation index: subpopulation to total*). Lisäksi tarkastelen lambda-tasoa (*inflation factor / overdispersion factor*).

### IBS-taso

Eri tutkimusten IBS-tulokset eivät valitettavasti ole yhteismitallisia. Ne voidaan kuitenkin kalibroida karkeasti toisiaan vastaaviksi vertaamalla väestöjen sisäisiä IBS-arvoja. Tutkimuksessa Salmela et al. 2011 pohjoiseurooppalaisten väestöjen sisäiset IBS-arvot vaihtelevat norrlantilaisten 0,7167:sta itäsuomalaisten 0,7195:een. Tutkimuksessa Huyghe et al. 2011 ranskalaisten IBS on 0,7461, venäläisten 0,7469, orkneyläisten 0,7475 ja saamelaisten 0,7506. Baskien ja sardinialaisten IBS on lähes tarkalleen saamelaisten tasoa (0,7505 ja 0,7503), ja koska ne ovat isolaatteja kuten saamelaisetkin, voidaan olettaa tämän suuruisen arvon olevan tyypillinen isolaateille. Koska itäsuomalaisetkin ovat samanlainen isolaatti, saadaan nämä eri tutkimusten korkeimmat arvot ankkuroitua toisiinsa.

There are finally appearing genome-wide studies which compare Saamis and Finns to each other. Now it is the last moment to find out what can be concluded about the relationship indirectly; they both have been compared to other European populations, after all. Afterwards it is then interesting to compare the results of this article to the results gained by direct comparison.

I utilize both the IBS-level (*identity by state*), which shows the shared alleles (and thus the common gene pool), and the FST-level (*fixtation index: subpopulation to total*), which shows the admixture versus isolation (and thus the genetic distance). Additionally I take a glimpse to the lambda-level (*inflation factor / overdispersion factor*).

### IBS-level

Unfortunately the IBS-results from different studies are not commensurate. Nevertheless, they can be roughly calibrated to each other by comparing the intra-population IBS-values. In the study Salmela et al. 2011 the internal IBS-values for Northern European populations vary from 0.7167 of Norrlanders (North Swedes) to 0.7195 of East Finns. In the study Huyghe et al. 2011 the internal IBS of the French is 0.7461, of Russians 0.7469, of Orcadians 0.7475, and of Saamis 0.7506. IBS-value of Basques and Sardinians is comparable to that of Saamis (0.7505 and 0.7503), and because all those populations are isolates, this kind of magnitude seems to be typical for isolates. Because East Finns are also a similar isolate, these highest values of different studies can be anchored to each other.

Kun siis Salmelan (2011) 0,7195 vastaa Huyghen (2011) 0,7506:a, on muuntokerroin Salmelasta Huygheen 1,0432245.

Näissä muuntamisissa on kuitenkin huomioitava se, että myös kokonaishajonta on erilainen: Huyghella (2011) pienin IBS-arvo on 0,7118 kun mukana ei ole afrikkalaisia väestöjä, ja Salmelalla (2011) noin 0,6610 kun mukana on nigerialaisia. Yhteismitalliseksi matalimmaksi arvoksi voidaan katsoa venäläisten etäisyys itäaasialaisiin (kiinalaisiin ja japanilaisiin): Salmelalla tämä IBS-arvo on noin 0,6835, Huyghella 0,7233–0,7236. Korkeimmat arvot ovat Salmelalla 0,7195 (itäsuomalaisten sisäinen) ja Huyghella 0,7506 (saamelaisien sisäinen). Vaihteluvälit ovat tutkimuksissa niin samankaltaiset, että pohjoiseurooppalaisten sisällä yksiköt ovat käytännössä yhteismitallisia.

Tutkimuksen Salmela et al. 2008 kohdalla korkein arvo on itäsuomalaisten sisäinen 0,8025 ja matalin saksalaisista ja englantilaisista itäaasialaisiin, noin 0,7685. Jälleen ero vaihteluvälissä on pohjoiseurooppalaisten osalta käytännössä mitätön, joten kaikkia tutkimuksia voidaan pitää mittaasteikoiltaan yhteensopivina. Muuntokerroin Salmelasta 2008 (itäsuomalaisten 0,8025) Huygheen (saamelaisten 0,7506) on 0,9353271.

Kerrointa on silti hyvä vielä testata. Kun itäsuomalaisten (jotka tältä osin ovat samankaltaisimpia venäläisten kanssa) etäisyys itäaasialaisiin tutkimuksesta Salmela et al. 2008 (0,7730) muunnetaan Huyghen asteikolle, siitä tulee oikeansuuntainen 0,7230 (Huyghella venäläisistä japanilaisiin 0,7233). Kun kuitenkin venäläisten IBS-arvo itäaasialaisiin tutkimuksesta Salmela et al. 2011 (0,6835) muunnetaan Huyghen mukaiseksi, siitä tulee vain 0,7130. En tiedä mistä ero johtuu, mutta jostain syystä Salmela et al. 2008 on yhteismitallisempi Huyghen et al. 2011 kanssa kuin Salmela et al. 2011.

Now that 0.7195 from Salmela (2011) corresponds to 0.7506 from Huyghe (2011), is the conversion multiplier from Salmela to Huyghe 1.0432245.

It should also be noted that the total dispersion is different, too: the lowest value in Huyghe (2011) is 0.7118 without Africans, and in Salmela (2011) about 0.6610 with Nigerians. The lowest commensurate value is from Russians to East Asians (Chinese and Japanese): in Salmela this IBS-value is about 0.6835, in Huyghe 0.7233–0.7236. The highest values are in Salmela 0.7195 (East Finnish internal) and in Huyghe 0.7506 (Saami internal). The ranges are so similar that the units are practically equal concerning North European.

In the study Salmela et al. 2008 the highest value is East Finnish internal 0.8025, and the lowest from Germans and Brits to East Asians, about 0.7685. Again the difference in the range is practically meaningless, so all the mentioned studies can be seen as compatible by their measure. The convert multiplier from Salmela 2008 (East Finnish 0.8025) to Huyghe (Saami 0.7506) is 0.9353271.

It is still good to test for the multiplier. When we change the distance from East Finns (who in this respect remind Russians) to East Asians from the study Salmela et al. 2008 (0.7730) to the Huyghe scale, the result is correct 0.7230 (in Huyghe from Russians to Japanese 0.7233). But when we change the IBS-value from Russians to East Asians from the study Salmela et al. 2011 (0.6835) to the Huyghe scale, the result is only 0.7130. I don't know the reason for this, but somehow Salmela et al. 2008 is more commensurate with Huyghe 2011 than Salmela et al. 2011.

	<b>Salmela 2011</b>	<b>Huyghe 2011</b>	<b>Salmela 2008</b>
Matalin/lowest:	0,6835	0,7118	0,7685
Korkein/highest:	0,7195	0,7506	0,8025
Vaihtelu/variation:	0,0360	0,0388	0,0340
Yksikköjen suhteet/ relation of units:	0,0001 => (1/360)	=> 0,0001078 0,000114 <=	<= 0,0001 (1/340)
	0,000093 <=	<= 0,0001 => (1/388)	=> 0,000088
Muuntokerroin/ conversion multiplier:	1,0432245 =>		<= 0,9353271

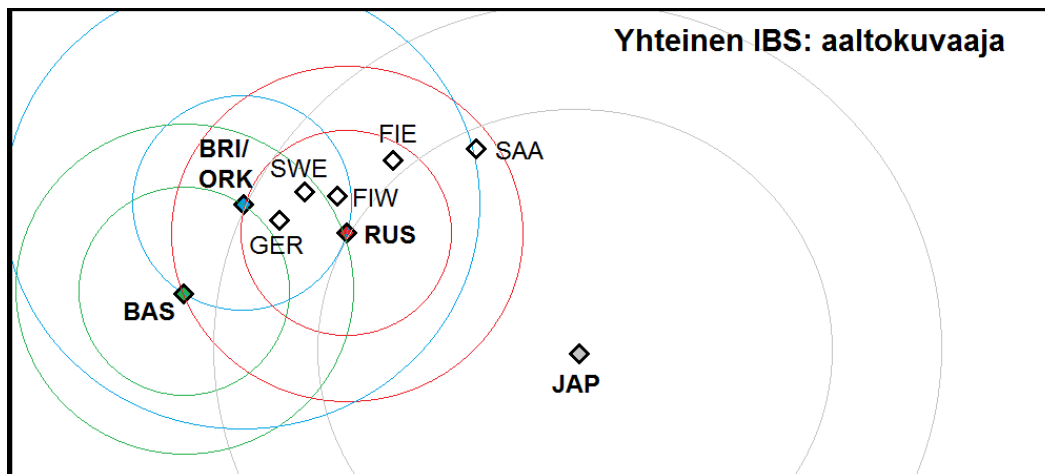
Seuraavassa parivertailutaulukossa IBS-arvot on muunnettu Huyghen muotoon, koska hänellä tuloksia on useammasta väestöstä. Länsi- ja itäsuomalaisten, ruotsalaisten, saksalaisten ja brittien väliset IBS-arvot on muunnettu tutkimuksesta Salmela et al. 2008, samoin näiden IBS-arvo itäaasialaisiin; näiden (paitsi brittien) IBS-arvo venäläisiin on tutkimuksesta Salmela et al. 2011.

In the next table the IBS-values have been converted to the Huyghe scale, because he has results from more populations. The IBS-values between West and East Finns, Swedes, Germans and Brits have been converted from the study Salmela et al. 2008, and likewise their IBS-value to East Asians; but their (except Brits') IBS-value to Russians is from the study Salmela et al. 2011.

	fiE	fiW	svea	ger	bri/orc	saa	rus	bas	jap
fiE	<b>0,7506</b>	0,7462	0,7448	0,7443	0,7441	-----	0,7455	-----	0,7230
fiW	0,7462	<b>0,7485</b>	0,7455	0,7452	0,7451	-----	0,7463	-----	0,7221
svea	0,7448	0,7455	<b>0,7479</b>	0,7455	0,7455	-----	0,7458	-----	0,7197
ger	0,7443	0,7452	0,7455	<b>0,7481</b>	0,7459	-----	0,7460	-----	0,7188
bri/orc	0,7441	0,7451	0,7455	0,7459	<b>0,7475</b>	0,7408	0,7446	0,7446	0,7188
saa	-----	-----	-----	-----	0,7408	<b>0,7506</b>	0,7434	0,7388	0,7272
rus	0,7455	0,7463	0,7458	0,7460	0,7446	0,7434	<b>0,7469</b>	0,7431	0,7233
bas	-----	-----	-----	-----	0,7446	0,7388	0,7431	<b>0,7505</b>	0,7159
jap	0,7235	0,7221	0,7197	0,7188	0,7188	0,7272	0,7233	0,7159	<b>0,7759</b>

Vaikka saamelaisia ja suomalaisia vertailevaa genomilajajuista tutkimusta ei vielä ole ilmestynyt, on mielenkiintoista katsoa tilannetta epäsuorasti, kolmansien osapuolten kautta. Seuraava nelinapainen aaltokuvaaja perustuu baskien, orkneyläisten/brittien, venäläisten ja japanilaisten keskinäisiin IBS-arvoihin, ja saamelaiset, suomalaiset, ruotsalaiset ja saksalaiset on lisätty kuvaan jälkikäteen näihin suhteuttaen. Vain saamelaisilta tunnetaan IBS-arvo kaikkiin neljään napaväestöön, muilta puuttuu vertailuarvo baskeihin. Kuvaa näyttää suhteellisen läheisyyden: väestöt ovat oikeassa järjestyksessä suhteessa kaikkiin muihin väestöihin.

Although there still are no published genome-wide studies comparing Saamis and Finns, it is interesting to look at the situation through the third parties. The following four-pole wave graph is based on the IBS-values between the Basque, Orcadians/Brits, Russians and Japanese. Saamis, Finns, Swedes and Germans have been added in the figure related to these pole populations. Only Saamis have the IBS-value for all the four pole populations, all the other lack the value to Basques. The graph shows the relative closeness: the populations are in the right order compared to every other population.



- Japanilaisiin (JAP) verrattuna saamelaiset jakavat selvästi eniten geenipohjaa, seuraavina tulevat venäläiset (RUS) sekä itä- (FIE) ja länsisuomalaiset (FIW), sitten odotuksenmukaisesti ruotsalaiset (SWE), saksalaiset (GER) ja englantilaiset (BRI) sekä viimeisinä baskit (BAS).
- Venäläisiin verrattuna lähimpänä ovat länsisuomalaiset, sitten saksalaiset ja eteläruotsalaiset (svealaiset ja götalaiset), ja yllättäen sitten vasta itäsuomalaiset, joiden jälkeen pohjoisruotsalaiset (norrlantilaiset; ei merkitty); sitten tulevat britit, saamelaiset ja baskit. Kaikkein vähiten kaikkien eurooppalaisten kanssa jakavat tietysti japanilaiset.
- Britteihin verrattuna lähimpiä ovat saksalaiset, ruotsalaiset ja länsisuomalaiset, sitten baskit ja venäläiset; kauimpana ovat itäsuomalaiset ja saamelaiset.
- Baskeihin verrattuna lähimpänä ovat britit, saksalaiset, ruotsalaiset ja venäläiset, kauempana länsi- ja itäsuomalaiset sekä saamelaiset.

Itäsuomalaiset sijoittuvat selvästi länsisuomalaisista saamelaisten suuntaan. Silti niin länsi- kuin itäsuomalaisetkin ovat selvästi pohjoiseurooppalaisia, yhdessä saamelaisten kanssa (baskien ulomman aallon ulkopuolella). Länsisuomalaiset ja jopa venäläiset ovat lähellä länsieurooppalaisia, kun taas itäsuomalaiset ja saamelaiset ovat itäisempiä (brittien sisemmän aallon ulkopuolella). Väestöjen samankaltaisuus japanilaisiin on vastakkainen britteihin nähden.

Venäläiset edustavat itäeurooppalaisuutta, ja länsisuomalaiset ovat lähimpänä heitä. Yllättävää on, että itäsuomalaiset ovat lähes yhtä kaukana venäläisistä kuin baskit, mutta eri suunnalla, ja saamelaiset ovat vielä kauempana. Saamelaisten edustama pohjoinen akseli on siis aivan oma, itäeurooppalaisuudesta riippumaton tekijänsä.

- Compared to Japanese (JAP), Saamis share the most gene base, and the next are Russians (RUS), East (FIE) and West Finns (FIW), then as expected Swedes (SWE), Germans (GER), Brits (BRI) and the last Basques (BAS).
- Compared to Russians, the closest are West Finns, then Germans and South Swedes (Svealanders and Götalanders), and surprisingly only after them East Finns, and after them North Swedes (Norrlander; not in the graph); then come Brits, Saamis and Basques. Of course least with all European populations share the Japanese.
- Compared to Brits, the closest are Germans, Swedes and West Finns, then Basques and Russians; farthest are East Finns and Saamis.
- Compared to Basques, the closest are Brits, Germans, Swedes and Russians, farther are West and East Finns and Saamis.

East Finns are placed clearly from West Finns towards Saamis. Still, both West and East Finns are North Europeans, together with Saamis (outside the outer wave of Basques). West Finns and even Russians are close to West Europeans, while East Finns and Saamis are more eastern (outside the outer wave of Brits). The similarity to the Japanese is inverse compared to Brits.

Russians represent East Europeans, and West Finns are closest to them. It is surprising that East Finns are almost as far from Russians as Basques are, but in the different direction, and Saamis are still farther. The northern axis represented by Saamis is therefore an independent factor from East Europeanness.

## Lambda-taso ( $\lambda$ )

Lambda (*inflation factor / overdispersion factor*) osoittaa kuinka paljon väestöjen alleelikohtaiset frekvenssit poikkeavat odotetusta khiin neliötes-tissä. Toisin kuin IBS-tasolla, tässä eri tutkimus-ten tulokset ovat ilmeisesti laskentamenetelmästä johtuen yhteismitallisia: tutkimuksessa Salmela et al. 2011 ruotsalaisten (svealaisten) ja saksalaisten välinen arvo on 1,16, ja tutkimuksessa Nelis et al. 2009 se on 1,12 pohjoissaksalaisiin ja 1,21 etelä-saksalaisiin. Samoin länsi- ja itäsuomalaisten vä- linen arvo on Salmelalla 1,71, ja helsinkiläisten ja kuusamolaisten välinen arvo on Nelisillä 1,86. Salmelalla länsisuomalaisten ja itäsuomalaisten arvot ruotsalaisiin (svealaisiin) ovat 1,52 ja 2,62 ja Nelisillä helsinkiläisten ja kuusamolaisten arvot ruotsalaisiin ovat 1,89 ja 2,87. Nämä erot johtuvat siitä, että helsinkiläiset ovat geneettisesti länsi- ja itäsuomalaisten sekoitus, kun taas kuusamolaiset edustavat hyvin äärimmäistä ja ajautunutta itä-suomalaista osaväestöä.

Joissain tutkimuksissa lambda on kuitenkin laskettu FST-etäisyyksistä, ja sellaiset arvot eivät luonnollisestikaan vastaa suoraan alleeleista las- kettua lambdaa. Seuraavaan taulukkoon on kerätty lambda-arvot mainituista tutkimuksista: Salmelal- ta läntisiin väestöihin verrattuna ja Nelisiltä itäi- siin väestöihin verrattuna. Saamelaisista ei valitet- tavasti ole saatavilla lambda-arvoja.

## Lambda-level ( $\lambda$ )

Lambda (*inflation factor / overdispersion factor*) shows how much the allele-wise frequencies of populations differ from the expected values in Chi-square-test. Unlike in the IBS-level, here the results of different studies are commensurable, possibly due to the calculation formula. In the study Salmela et al. 2011 the value between Swedes (Svealanders) and Germans is 1.16, and in the study Nelis et al. 2009 it is 1.12 to North Germans and 1.21 to South Germans. Similarly the value between West and East Finns is 1.71 in Salmela, and the value between Helsinki Finns and Kuusamo Finns is 1.86 in Nelis. In Salmela the values from West and East Finns to Swedes (Svealanders) are 1.52 and 2.62, and in Nelis the values from Helsinki and Kuusamo Finns to Swedes are 1.89 and 2.87. These differences are due to the fact that Helsinki Finns are a mixture of West and East Finns, while Kuusamo Finns are an extreme and heavily drifted East Finnish subpopu- lation.

In some studies lambda is however calculated from the FST-distances, and such values naturally do not correspond to the directly from the alleles calculated lambda. In the following table there are relevant lambda-values from the mentioned stud- ies: from Salmela compared to western popula- tions, and from Nelis compared to eastern popula- tions. Unfortunately there are no lambda-values available from Saamis.

<i>Salmela 2011</i>	fiE	fiW	sweN	sweC	sweS	ger	bri		
fiE	<b>1,0</b>	1,71	2,59	2,62	2,91	3,08	3,30		
fiW	1,71	<b>1,0</b>	1,56	1,52	1,70	1,82	2,05		
sweN	2,59	1,56	<b>1,0</b>	1,12	1,20	1,36	1,46		
sweC	2,62	1,52	1,12	<b>1,0</b>	1,03	1,16	1,28		
sweS	2,91	1,70	1,20	1,03	<b>1,0</b>	1,13	1,21		
ger	3,08	1,82	1,36	1,16	1,13	<b>1,0</b>	1,11		
bri	3,30	2,05	1,46	1,28	1,21	1,11	<b>1,0</b>		
<i>Nelis 2009</i>	fiK	fiH	est	rus	lat	lit	pol	swe	gerN
fiK	<b>1,0</b>	1,86	2,80	3,16	3,33	3,37	2,49	2,87	3,27
fiH	1,86	<b>1,0</b>	1,71	2,10	2,31	2,33	1,75	1,89	2,17
est	2,80	1,71	<b>1,0</b>	1,21	1,24	1,28	1,17	1,49	1,53
rus	3,16	2,10	1,21	<b>1,0</b>	1,32	1,26	1,18	1,59	1,49
lat	3,33	2,31	1,24	1,32	<b>1,0</b>	1,20	1,26	1,89	1,84
lit	3,37	2,33	1,28	1,26	1,20	<b>1,0</b>	1,20	1,74	1,66
pol	2,49	1,75	1,17	1,18	1,26	1,20	<b>1,0</b>	1,30	1,18
swe	2,87	1,89	1,49	1,59	1,89	1,74	1,30	<b>1,0</b>	1,12
gerN	3,27	2,17	1,53	1,49	1,84	1,66	1,18	1,12	<b>1,0</b>

Salmelan tuloksissa on mielenkiintoista, että länsisuomalaiset ovat ruotsalaisista lähimpänä Keski-Ruotsin svealaisia (sweC) mutta itäsuomalaiset pohjoisen norrlandilaisia (sweN) – tosin erot eivät ole suuret. Molemmat ovat sen sijaan selvästi kauimpana etelän götalaisista (sweS). Götalaiset ovat jo yhtä lähellä brittejä ja jopa lähempänä saksalaisia kuin norrlandilaisia.

Nelisin tuloksissa huomiota kiinnittää, että helsinkiläisiin (fiH) verrattuna puolaiset (pol) ovat käytännössä yhtä lähellä kuin virolaiset (est), ja kuusamolaisiin (fiK) verrattuna he ovat jopa selvästi lähempänä kuin virolaiset. Sekä helsinkiläisiin että kuusamolaisiin verrattuna ruotsalaiset ovat puolalaisten ja virolaisten jälkeen kolmanneksi samankaltaisin väestö. Venäläiset ja saksalaisetkin ovat lähempänä suomalaisia kuin latvialaiset ja liettualaiset.

Ruotsalaiset ja saksalaiset ovat lähellä toisiaan ja myös puolalaisia; seuraavaksi lähimpänä heitä ovat virolaiset ja venäläiset, hieman kauempana baltit ja kaikkein kauimpana suomalaiset. Venäläiset ovat lähempänä virolaisia ja puolalaisia kuin baltteja, ja näiden jälkeen lähempänä saksalaisia ja ruotsalaisia kuin suomalaisia. Virolaiset ja baltit ovat lähellä toisiaan sekä venäläisiä ja puolalaisia, ja lähempänä myös ruotsalaisia ja saksalaisia kuin suomalaisia. Yllättäen virolaiset ovat lähempänä puolalaisia kuin latvialaisia ja liettualaisia tai edes nämä toisiaan.

Balteista katsoen suomalaiset ja virolaiset ovat lähempänä latvialaisia, kun taas venäläiset, puolalaiset, ruotsalaiset ja saksalaiset ovat lähempänä liettualaisia. Ruotsalaisiin nähden toiseksi lähimpänä heti saksalaisten jälkeen ovat puolalaiset, sitten virolaiset ja venäläiset, sitten baltit ja helsinkiläiset, ja kauimpana ovat kuusamolaiset.

It is interesting in the results of Salmela, that compared to Swedes, West Finns are closest to Svealanders (sweC) and East Finns are closest to Norrlanders (sweN); they both are farthest from Götalanders (sweS). Götalanders are as close to Brits as they are to Norrlanders, and even closer to Germans.

In the results of Nelis it draws attention, that compared to Helsinki Finns (fiH) Poles (pol) are practically as close as Estonians (est), and compared to Kuusamo Finns (fiK) they are even clearly closer than Estonians. Swedes are the third closest population to both Helsinki and Kuusamo Finns. Even Russians and Germans are closer to Finns than Latvians and Lithuanians are.

Swedes and Germans are close to each other and also to Poles; next to them are Estonians and Russians, a bit farther Balts and farthest Finns. Russians are closer to Estonians and Poles than to Balts, and after them closer to Germans and Swedes than to Finns. Estonians and Balts are close to each other and Russians and Poles, and also closer to Swedes and Germans than to Finns. Surprisingly Estonians are closer to Poles than to Latvians and Lithuanians, or the latter two to each other.

Looking from Balts, Finns and Estonians are closer to Latvians, while Russians, Poles, Swedes and Germans are closer to Lithuanians. Looking from Swedes, the second closest after Germans are Poles, then Estonians and Russians, then Balts and Helsinki Finns, and the farthest are Kuusamo Finns.

## FST-taso

Tälläkin tasolla näyttää olevan yhteismitattomuutta Salmelan ja Huyghen välillä:

Venäläisten $F_{ST}$ (alkuperä/origin)	Salmela 2011 (Vologda)	Huyghe 2011 (Vologda)	Nelis 2009 (Tver)
RUS-FRA	50,4	18	50
RUS-ITN	74,8	17	90
RUS-GER	28,2	---	25
RUS-BAS	135,5	83	---
RUS-SWE	29,8	---	30
RUS-SAR	178,5	130	---
SAR-BAS	131,0	81	---
SWE-ITN	43,7	---	70

Nelis ja Salmela näyttävät olevan pääosin yhteensopivia, huolimatta venäläisten (Tver vs. Vologda) ja ruotsalaisten (Tukholma vs. koko maa) näytteiden erilaisesta alkuperästä. Ruotsalaisten pienempi etäisyys pohjoisitalialaisiin Salmelalla voi johtua osittain eteläruotsalaisten vedosta, nämä kun ovat ruotsalaisista lähimpänä läntisiä ja eteläisiä eurooppalaisia (ks. lambda-taso edellä).

Jotta saamelaisten FST-arvot saataisiin vertailukelpoisiksi, on tärkeää havaita että kaikissa suurimmissa eurooppalaisten välisissä etäisyyksissä Huyghen arvot ovat lähes tarkalleen 50 pienempiä kuin Salmelan arvot. Koska saamelaisten kaikki FST-arvot kuuluvat tähän suuruusluokkaan, pitää niihin lisätä +50 jotta ne saadaan edes summittaisesti vertailukelpoisiksi Salmelan arvojen kanssa. Nämä muunnetut arvot on merkitty taulukkoon merkillä \*.

Salmela	fiE	fiW	swe	ger	bri	orc	saa	rus	bas	jap	pat
fiE	----	30,3	70,3	85,0	96,0	126,2	----	69,0	202,5	----	----
fiW	30,3	----	24,2	34,5	44,0	71,0	----	34,4	141,6	----	----
swe	70,3	24,2	----	6,5	10,4	35,6	----	29,8	100,0	3.	----
ger	85,0	34,5	6,5	----	4,9	32,4	----	28,2	91,4	4.	----
bri	96,0	44,0	10,4	4,9	----	26,9	----	38,3	81,7	----	----
orc	126,2	71,0	35,6	32,4	26,9	----	194*	64,8	112,5	5.	2.
saa	----	----	----	----	----	194*	----	148*	266*	1.	4.
rus	69,0	34,4	29,8	28,2	38,3	64,8	148*	----	135,5	2.	1.
bas	202,5	141,6	100,0	91,4	81,7	112,5	266*	135,5	----	6.	3.
jap	----	----	3.	4.	----	5.	1.	2.	6.	----	----
pat	----	----	----	----	----	2.	4.	1.	3.	----	----

Japanilaisten kohdalla on se ongelma, että Salmelalla he eivät ole mukana vertailussa. Joka tapauksessa etäisyydet ovat niin suuria, että vain järjestyksellä on käytännön merkitystä: eurooppalaisista saamelaiset ovat lähimpänä japanilaisia, ja kohti Länsi- ja Lounais-Eurooppaa etäisyys kasvaa.

## FST-level

Also at this level there seem to be the lack of commensurate between Salmela and Huyghe:

Nelis and Salmela seem to be mainly commensurate, despite the different origins of Russians (Tver vs. Vologda) and Swedes (Stockholm vs. whole country). The smaller distance from Swedes to North Italians (ITN) in Salmela might be caused by the draw of South Swedes, who are of Swedes closest to Western and Southern Europeans (see lambda-level afore).

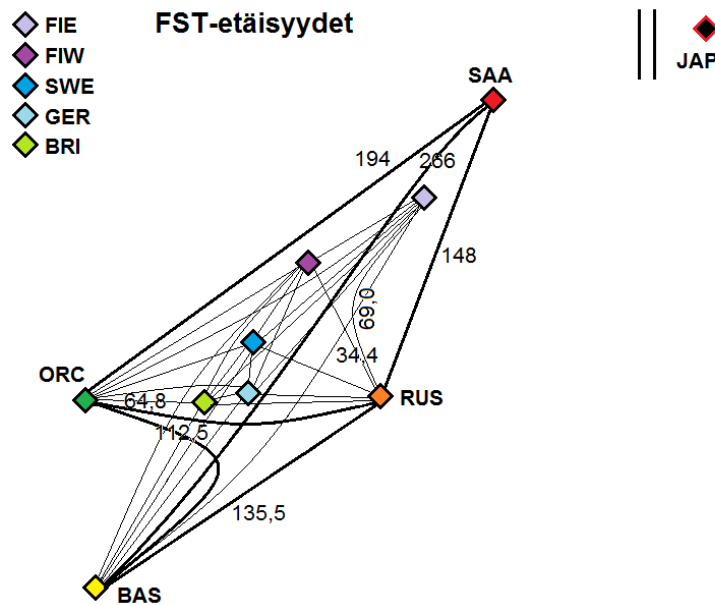
To make the FST-values of Saamis commensurate, it is important to note that in all the greatest distances between European populations the values from Huyghe are almost precisely 50 smaller than in Salmela. Because all the values from Saamis belong to this magnitude, +50 must be added to them to make them even remotely commensurate with the values from Salmela. These calibrated values are marked by \* in the following table.

Sama kuva pätee muihinkin itä- ja pohjoisaasialaisiin väestöihin, mutta sen sijaan keskiaasialaisiin väestöihin nähden venäläiset ovat lähempänä kuin saamelaiset, ja eteläaasialaisiin nähden jo kaikki muut eurooppalaiset paitsi sardinialaiset ovat lähempänä kuin saamelaiset; taulukossa esimerkkinä ovat pataanit.

Seuraavassa kuvaajassa saamelaisia ja baskeja on tuotu lähemmäs pääryhmää kuin lukuarvot edellyttävät; muuten väestöjen väliset suhteet ovat oikeassa mittakaavassa, paitsi että japanilaisten osalta näytetään vain suunta, ei etäisyyttä. Isolaation aiheuttama geneettinen ajautuminen saa eräät väestöt ”taipumaan” omiin suuntiinsa: esimerkiksi orkneyläisten ja baskien välinen etäisyys ei ole samassa avaruudessa muiden etäisyyksien kanssa vaan kiertää brittien kautta, ja itäsuomalaisten ja venäläisten välinen etäisyys kiertää länsisuomalaisten kautta.

The same is true also compared to other East and North Asian populations, but compared to Central Asians Russians are already closer than Saamis, and compared to South Asians all the other Europeans except Sardinians are closer than Saamis; an example population in the table are the Pathans.

In the following figure Saamis and Basques have been drawn closer to the main group than they would be by their values; otherwise the relations between the populations are in the right scale, except of the Japanese only the direction is shown, not the distance. Genetic drift caused by isolation makes some populations to “bend” to their own directions: for example the distance between Orcadians and Basques is not in the same space with the other distances, but it bends and rounds through Brits. The distance between East Finns and Russians rounds through West Finns.

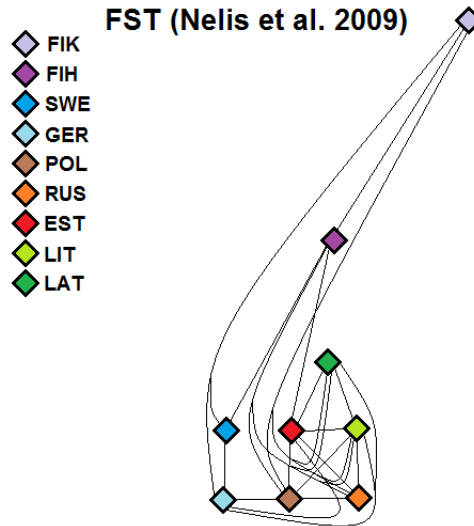


FST-tasolla näkyy sama ilmiö kuin IBS-tasollakin: ankkuroituna kolmeen eurooppalaiseen napaväestöön (venäläiset, orkneyläiset ja baskit) itäsuomalaiset lähestyvät saamelaisia.

At the FST-level the same phenomenon can be seen which was seen at the IBS-level: anchored to the three European pole populations (Russians, Orcadians, Basques) East Finns approach Saamis.



<i>Nelis</i>	fiK	fiH	swe	gerN	est	pol	rus	lat	lit
fiK	-----	50	110	120	90	120	120	130	130
fiH	50	-----	50	60	40	60	60	70	70
swe	110	50	-----	10	30	20	30	50	40
gerN	120	60	10	-----	30	10	20	40	30
est	90	40	30	30	-----	10	10	10	10
pol	120	60	20	10	10	-----	10	20	10
rus	120	60	30	20	10	10	-----	20	10
lat	130	70	50	40	10	20	20	-----	10
lit	130	70	40	30	10	10	10	10	-----



Kaikki ne etäisyydet, joita ei ole erikseen merkitty kuvaajaan, seuraavat lyhintä reittiä. Esimerkiksi ruotsalaisten etäisyys puolalaisiin: SWE 10 GER 10 POL = 20. Kuvaaja paljastaa monenlaisia kiertoreittejä. Puolalaisten (POL) etäisyys helsinkiläisiin (FIH) on 60, eli reitti ei kulje virolaisten kautta (POL 10 EST 40 FIH) vaan kiertää hieman. Sama kiertoreitti pätee venäläisiin ja baltteihin. Sen sijaan saksalaisten reitti helsinkiläisiin kulkee ruotsalaisten kautta: GER 10 SWE 50 FIH. kuusamolaisiin (FIK) nähden vain virolaisten reitti kulkee suoraan helsinkiläisten kautta, kun taas kaikilla muilla se kiertää hieman.

Koska FST-taso kertoo sekoittumisesta (arvo pienenee) tai sen puutteesta eli isoatiosta (arvo kasvaa), on tilanne mahdollista tulkita niin, että kiertoreitti kuvastaa alkuperäistä etäisyyttä ja oikein sekoittumista. Esimerkiksi ruotsinkieliset pohjalaiset ovat lähempänä sekä ruotsalaisia (FST ~20) että eteläpohjalaisia (FST 7) kuin näiden keskinäisen etäisyyden (FST ~40) perusteella voisi olettaa. Ruotsinkielisillä pohjalaisilla on siis genejä sekä ruotsalaisilta että eteläpohjalaisilta, mikä selittää tämän oikotien.

All the distances, which are not marked in the graph, follow the shortest route. For example the distance from Swedes to Poles: SWE 10 GER 10 POL = 20. The graph reveals many detours. The distance from Poles (POL) to Helsinki Finns (FIH) is 60, so it does not take the shortcut via Estonians (POL 10 EST 40 FIH). The same detour is true also in concern to Russians and Balts. Instead, the route from Germans to Helsinki Finns goes through Swedes: GER 10 SWE 50 FIH = 60. Compared to Kuusamo Finns, only Estonians follow the direct route through Helsinki Finns, all the other take detours.

Because the FST-level tells about admixture (value decreases) or isolation (value increases), the situation can be interpreted so that detour represents the original distance, and shortcut represents the result of mixing. For example, Swedish-speaking Ostrobothnians are closer to both Swedes (FST ~20) and Finnish-speaking Ostrobothnians (FST 7) than is expected on the direct distance of these populations (FST ~40). Swedish-speaking Ostrobothnians have genes from both Swedes and Ostrobothnians, which explains the shortcut.

Vastaavalla tavalla helsinkiläiset ovat kuusamo-laisten ja muiden pohjoiseurooppalaisten välissä (eli he ovat sekoittuneet molempien kanssa), mikä selittää muiden väestöjen kiertotiet. Virolaiset ovat sekoittuneet yhtäältä helsinkiläisten, toisaalta slaavien (POL, RUS) ja balttien (LIT, LAT) kanssa, mikä selittää heidän oikotiensä. Puolalaiset ovat sekoittuneet saksalaisten (GER) ja Baltian kansojen (EST, LIT, LAT) kanssa (saksalaisten etäisyys baltialaisiin kiertää venäläisten kautta). Sen sijaan ruotsalaiset eivät näytä oikaisevan saksalaisten ja länsisuomalaisten väliä, mikä merkitsee että geenivirta, joka on johtanut länsisuomalaisten sekoittumiseen, on tapahtunut ruotsalaisista länsisuomalaisiin eikä päinvastoin.

Kiertoreitti ei kuitenkaan aina edusta välttämättä alkuperäistä tilannetta, vaan kasvanut  $F_{ST}$ -etäisyys voi johtua isolaatiosta ja geneettisestä ajautumisesta. Näin näyttää tapahtuneen balttien ja erityisesti latvialaisten kohdalla: he ovat lähellä naapuriväestöjään kuten virolaisia, mutta selvästi kauempana helsinkiläisistä: he eivät siis seuraa reittiä LAT 10 EST 40 FIH = 50 vaan LAT 70 FIH. Latvialaisten ajautuminen ei kuitenkaan ole ollut yhtä suurta kuin itäsuomalaisilla, saamelaisilla, baskeilla tai sardinialaisilla, joilla ei ole lainkaan läheisiä väestöjä. Latvialaisten ajautumisen vahvistaa se, että lambda-tasolla he ovat vielä lähempänä virolaisia ja suomalaisia, mutta  $F_{ST}$ -tasolla liettualaiset ovatkin yhtä lähellä.

## Yhteenveto

Lambda-taso viittaa vahvasti siihen suuntaan, että alleelitasolla puolalaiset ja virolaiset olisivat lähimpänä länsisuomalaisia, mutta koska näistä väestöistä ei ole saatavana IBS-arvoja, ovat venäläiset IBS-tasolla lähimpänä länsisuomalaisia. IBS-tasolla länsisuomalaiset ovat jopa lähempänä venäläisiä kuin itäsuomalaisia, kun taas lambda-tasolla tilanne on päinvastainen. Nämä eroavaisuudet saattavat kuitenkin johtua hieman erilaisista näytteistä (IBS-tasolla  $f_{iW}$  ja  $f_{iE}$  vs. lambda-tasolla  $f_{iH}$  ja  $f_{iK}$ ).

Similarly, Helsinki Finns are between Kuusamo Finns and other North Europeans (they have mixed with them both), which explains the detours of the other populations. Estonians have mixed with Helsinki Finns, Slavs (POL, RUS), and Balts (LAT, LIT), which explains their shortcut. Poles have mixed with Germans (GER) and Baltian populations (EST, LIT, LAT); the route from Germans to Baltians detours through Russians. Instead, Swedes do not seem to take any shortcut to Helsinki Finns compared to Germans, which means that the gene flow has been from Swedes to Helsinki Finns and not to the other direction.

Detour does not necessarily always represent the original situation, but increased  $F_{ST}$ -value may be caused by isolation and genetic drift. This seems to be the case with Balts and especially Latvians: they are close to their neighbours like Estonians, but clearly more distant from Helsinki Finns: they don't follow the route LAT 10 EST 40 FIH but LAT 70 FIH. The drift in Latvians is however not as great as the drift in East Finns, Saamis, Basques or Sardinians, who have no similarly close populations at all. The drift in Latvians is confirmed by the lambda-level, at which they are still closer to Estonians and Finns than Lithuanians are, but at the  $F_{ST}$ -level Lithuanians are equally close.

## Summary

The lambda-level strongly hints that at the allele level Poles and Estonians are closest populations compared to West Finns, but because we don't have IBS-values between these populations, Russians are now closest to West Finns at the IBS-level. At the IBS-level West Finns are even closer to Russians than they are to East Finns, while at the lambda-level the situation is just the opposite. These differences may however be caused by different samples (at the IBS-level  $f_{iW}$  and  $f_{iE}$  vs. at the lambda-level  $f_{iH}$  and  $f_{iK}$ ).

Joka tapauksessa geenipohjansa osalta länsisuomalaiset kuuluvat selvästi itäeurooppalaisten joukkoon (rajan seuratessa Itämerta). Länsisuomalaisilla on vähemmän yhteistä länsieurooppalaisten kanssa kuin muilla itäeurooppalaisilla (virolaiset, venäläiset puolalaiset).

Lambda-tason tuloksilla täydennetyt IBS-tason perusteella hahmottuu kuitenkin kaksi muuta geenipohjaa länsi- ja itäeurooppalaisen geenipohjan lisäksi: saamelainen ja balttilainen. Saamelaisista geenipohjaa edustavat saamelaiset ja osittain itäsuomalaiset, ja balttilaista geenipohjaa edustavat latvialaiset ja liettualaiset. Nämä kaksi väestöparia ovat kaukana niin länsi- ja itäeurooppalaisista kuin toisistaankin. On monia mahdollisia syitä sille saamelaisten ja balttien erilaisuudelle:

- (1) He edustavat pohjoisen Euroopan jääkauden jälkeisiä alkuasukkaita, kun taas itäeurooppalaiset edustavat nuorempaa tulo-kaasväestöä.
- (2) He edustavat nuorempaa tulo-kaasväestöä tai vähintään vierasta sekoitusta, ja itäeurooppalaiset edustavat alkuasukkaita.
- (3) He edustavat alun perin ihan samaa itäeurooppalaista geenipohjaa, mutta isolaation ja vähäisen väkimäärän aiheuttama geneettinen ajautuminen on vienyt heitä omiin suuntiinsa.

Ajautumista näyttää F<sub>ST</sub>-tason perusteella tapahtuneen sekä saamelaisissa että balteissa, mutta vähemmän jälkimmäisissä. Riittävän suuri ajautuminen voi tietysti näkyä myös muilla tasoilla. Vierasta sekoitusta olisi todennäköisemmin saamelaisissa, ottaen huomioon Fennoskandian runsaat vaikutteet idän suunnalta esihistorian saatossa. Lisäksi aasialaisperäinen vaikutus näkyy saamelaisten äitilinjoissa (Ingman & Gyllensten 2007) ja genomilaajuusastekin IBS-tasolla (ks. edellä). Kuitenkin saamelaisissa on voinut säilyä myös muinaista eurooppalaista perimää, asuvathan he aivan maanosan pohjoisella reunalla.

Balttien kohdalla vieras vaikutus on epätodennäköisempi. Se ei voi olla samaa itäistä vaikutusta kuin saamelaisilla, koska lambda- ja F<sub>ST</sub>-tason perusteella baltit ovat erittäin kaukana itäsuomalaisista, jotka taas ovat lähempänä saamelaisia.

In any case, according to their gene base West Finns clearly belong among the East European populations (the border on Baltic Sea). West Finns have less common with the West European populations than the other East European populations do (Estonians, Russians, Poles).

When the IBS-level is complemented with the results of the lambda-level, two other gene bases can be sketched: the Baltic gene base and the Saamic genebase. The Saamic genebase is represented by Saamis and partially East Finns, while the Baltic gene base is represented by Latvians and Lithuanians. These two population pairs are distant to West and East Europeans and to each other. There are many possible reasons for the distinctiveness of Saamis and Balts:

- (1) They represent the original post-Ice Age settlers, while East Europeans represent the younger newcomers.
- (2) They represent the younger newcomers or at least some foreign admixture and East Europeans represent the original settlers.
- (3) They represent originally the same East European gene base, but the genetic drift caused by isolation and low population size has taken them to their own directions.

The drift can be confirmed both in Saamis and Balts according to the F<sub>ST</sub>-level, but less in the latter. Great enough drift can of course be seen also at other levels. Foreign admixture would be more probable in Saamis, considering the many influences from the east in prehistoric Fennoscandia. Additionally, the eastern influence in Saamis can be seen both in the maternal lineages (Ingman & Gyllensten 2007) and genome-wide at the IBS-level (see afore). Still, there may have been preserved also some ancient inheritance in Saamis, who live after all in the northernmost periphery of Europe.

In Balts a foreign influence is less probable. It could not be the same eastern influence than in Saamis, because according to both the lambda- and F<sub>ST</sub>-level Balts are very distant to East Finns, who are closer to Saamis.

Mikään muukaan vieras vaikutus ei tule kyseen, koska baltit ovat maantieteellisesti keskellä Eurooppaa. Todennäköisesti baltit edustavatkin muinaisen pohjoiseurooppalaisen geeniperimän saarekettä etelän suunnalta levinneiden länsi- ja itäeurooppalaisten välissä.

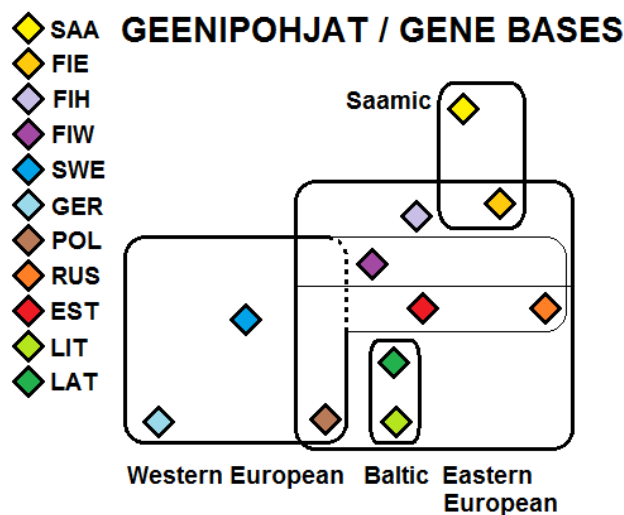
On oikeastaan yllättävää, että länsi- ja itäsuomalaiset sekä saamelaiset ovat suhteessa muihin eurooppalaisiin niin samanlaisia eri tasoilla. FST-tason lisäksi jo IBS-taso osoittaa itäsuomalaisten geenipohjan olevan länsisuomalaisten ja saamelaisten välissä, ja länsisuomalaiset toimivat siltana sekä länsi- että itäeurooppalaisiin. Ainoa selvä ero näillä tasoilla koskeekin länsisuomalaisia: geenipohjansa osalta he ovat selvästi lähempänä itäeurooppalaisia (IBS-tasolla venäläisiä ja lambda-tasolla lisäksi virolaisia ja puolalaisia), mutta FST-taso osoittaa heidän sekoittuneen vahvasti myös länsieurooppalaisiin: he ovat lähempänä ruotsalaisia kuin venäläisiä.

Länsisuomalaiset ovat siis alkuaan itäeurooppalainen, vahvasti länsieurooppalaista vaikutusta saanut väestö. (Oikeastaan kaikissa muissakin itäeurooppalaisissa näyttää olevan länsieurooppalaista vaikutusta.) Itäsuomalaiset sen sijaan näyttävät olevan itäeurooppalais-saamelainen sekäväestö.

No other foreign influence is probable, either, because geographically Balts are in the centre of Europe. Probably Balts represent the island of ancient North European inheritance between the south-originating West and East Europeans.

It is actually surprising that relationships between West Finns, East Finns and Saamis are so similar in different levels. In addition to the FST-level, also the IBS-level shows that the gene base of East Finns is between West Finns and Saamis, and West Finns bridge the gap to both West and East Europeans. The only clear difference in these levels concerns West Finns: by their gene base they are clearly closer to East Europeans (at the IBS-level Russians and at the lambda-level also Estonians and Poles), but the FST-level shows considerable West European admixture: they are closer to Swedes than and as close to Germans as they are to Russians.

Thus West Finns are originally East European but strong West European admixture received population. (Actually there seems to be West European admixture also in the other East Europeans except in Balts.) On the other hand, East Finns seem to be East European – Saamic admixture population.



## Lähteet / Sources

Huyghe JR, Fransen E, Hannula S, Van Laer L, Van Eyken E, Mäki-Torkko E, Aikio P, Sorri M, Huentelman MJ, Van Camp G 2011: A genome-wide analysis of population structure in the Finnish Saami with implications for genetic association studies. – *European Journal of Human Genetics* 19(3). <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3062008/?tool=pubmed>

Nelis et al. 2009: Genetic Structure of Europeans: A View from the North–East. – *PloS ONE* volume 4, issue 5. <http://www.plosone.org/article/info:doi%2F10.1371%2Fjournal.pone.0005472>

Salmela E, Lappalainen T, Fransson I, Andersen P. M, Dahlman-Wright K, Fiebig A, Sistonen P, Savontaus M-L, Schreiber S, Kere J, Lahermo P 2008: Genome-Wide Analysis of Single Nucleotide Polymorphisms Uncovers Population Structure in Northern Europe. – *PLoS ONE* volume 3, issue 10. <http://www.plosone.org/article/info:doi%2F10.1371%2Fjournal.pone.0003519>

Salmela et al. 2011: Swedish Population Substructure Revealed by Genome-Wide Single Nucleotide Polymorphism Data. – *PLoS ONE* volume 6, issue 2. <http://www.plosone.org/article/info:doi%2F10.1371%2Fjournal.pone.0016747>