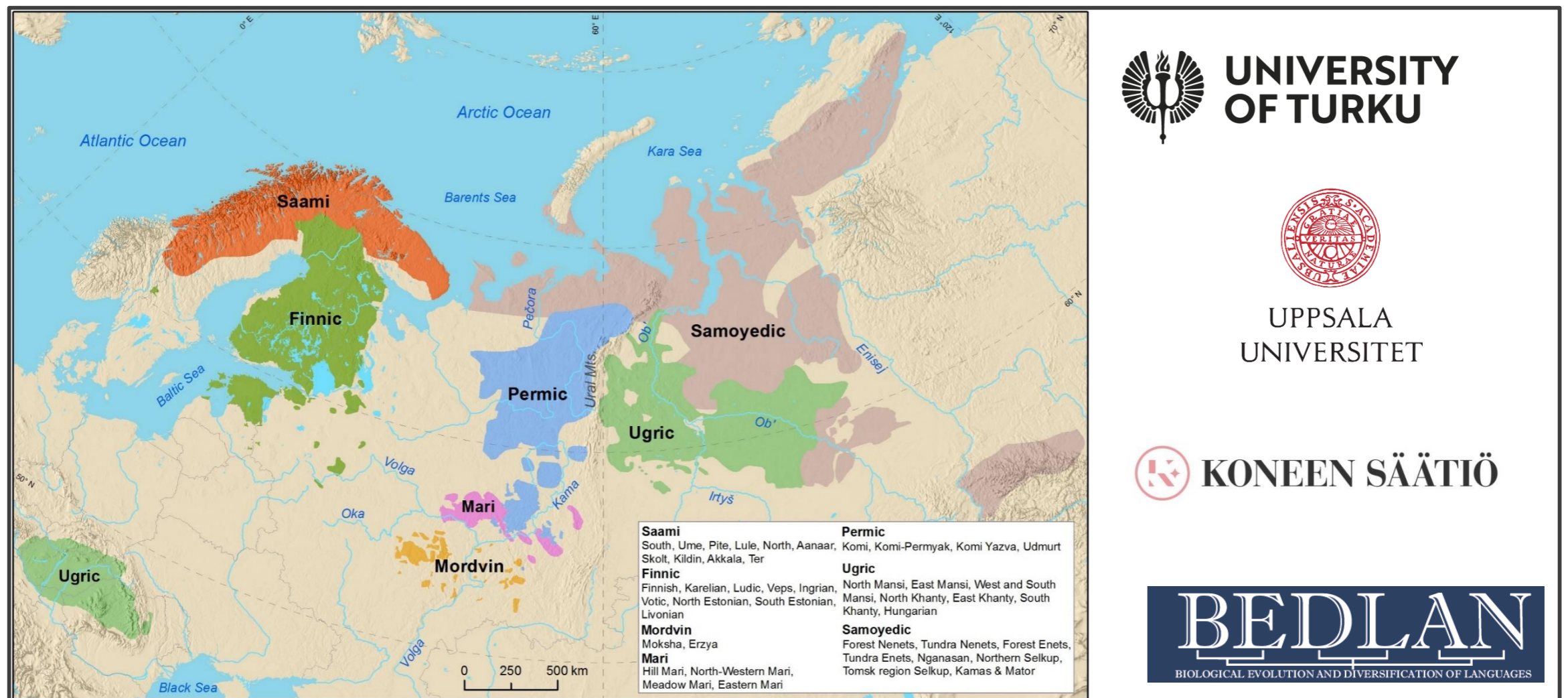
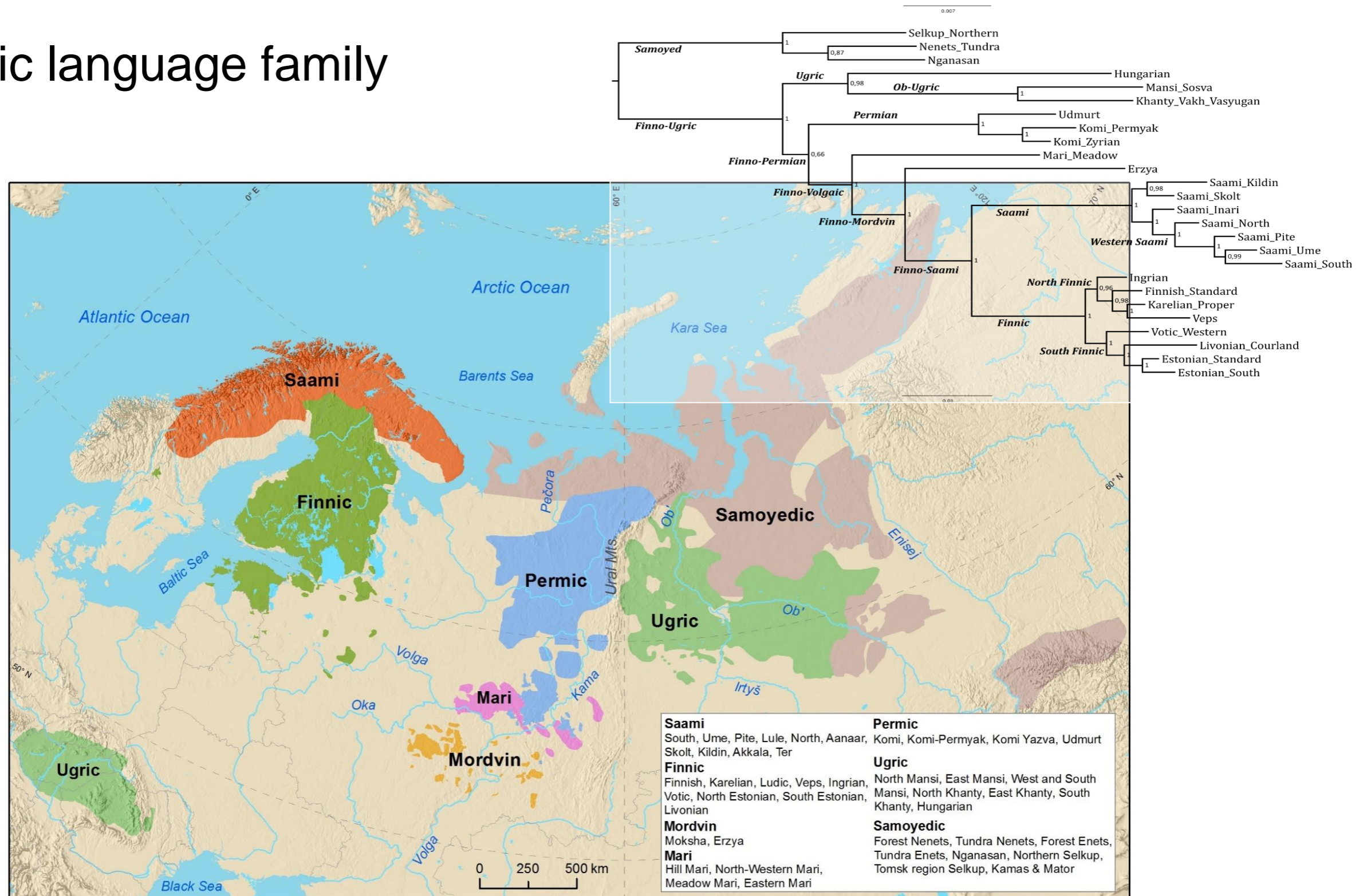


Genes and language in the prehistory of the Uralic speaking peoples

Outi Vesakoski, Terhi Honkola, Mervi de Heer, Michael Dunn

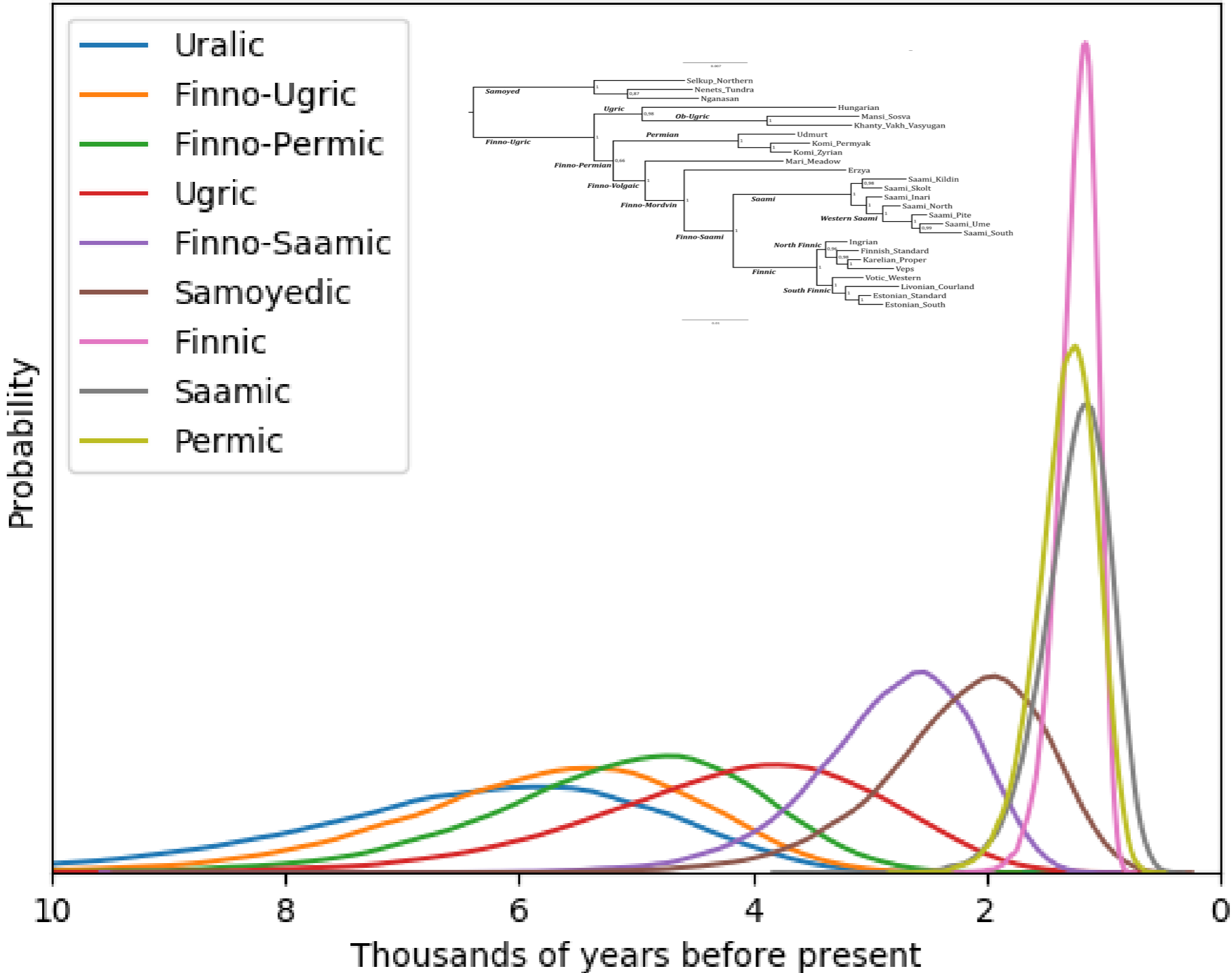


Uralic language family



Speaker area maps produced by BEDLAN “Cartographical database of Uralic languages”. Oxford Handbook of Uralic Languages (Rantanen et al. submitted).

Timing of divergence events of the Uralic language family: Ranges of probability distribution of the divergences based on Bayesian modelling

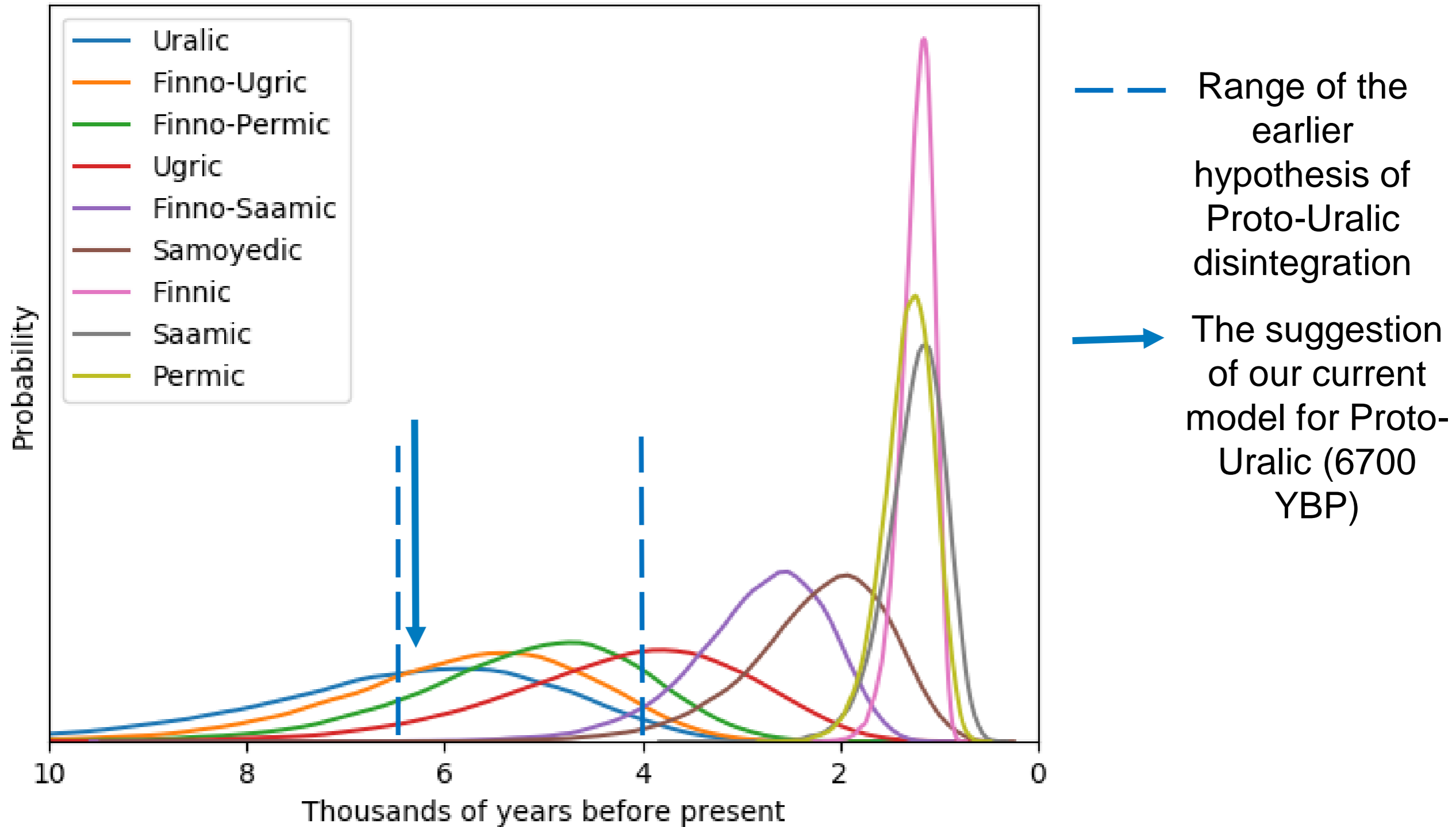


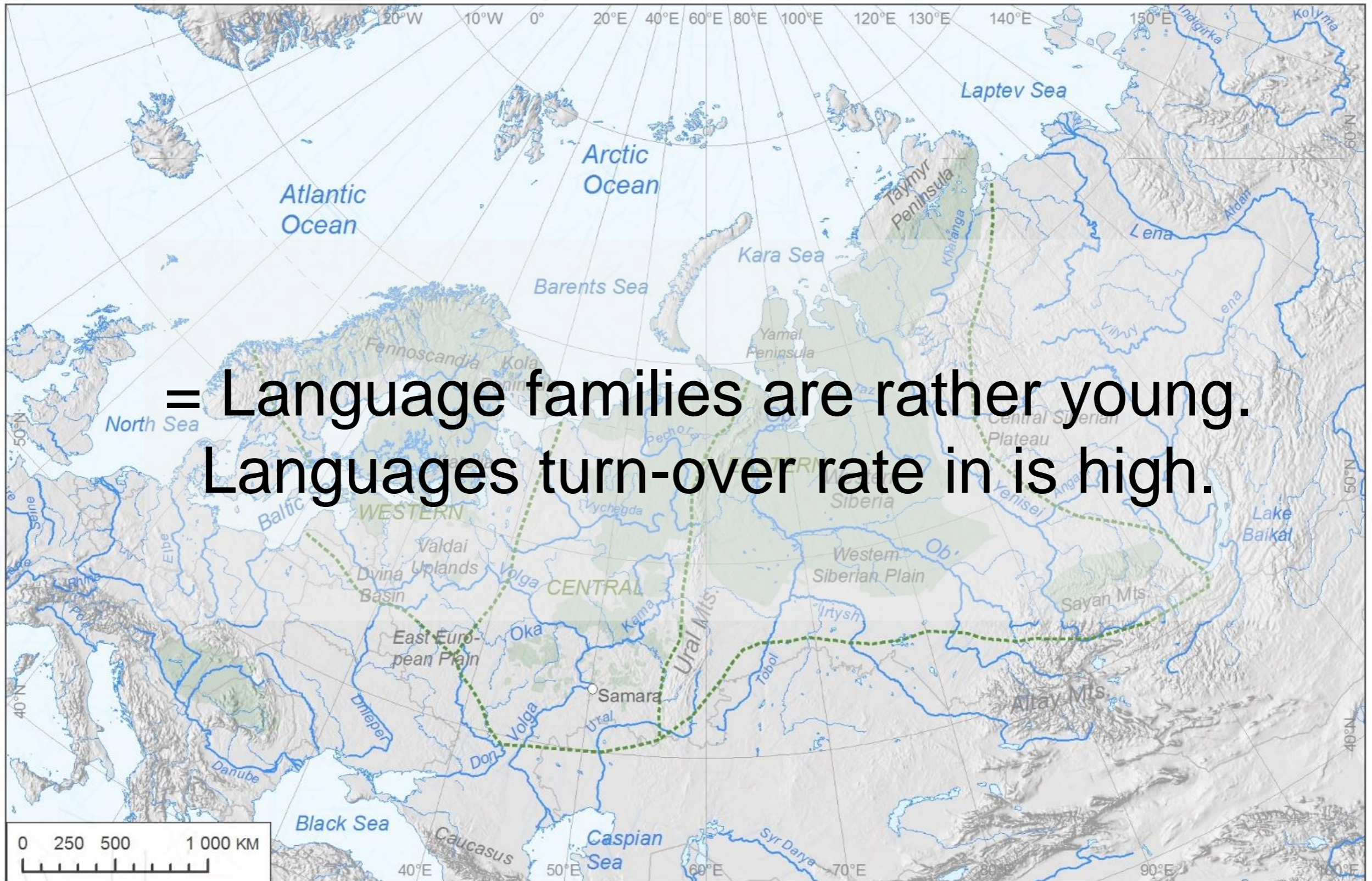
Uralex data : Basic vocabulary with cognate coding for 313 meanings in 26 language (Syrjänen et al. 2018)

BEAST analyses with binary Covarion model (Maurits et al. ms)

Calibrations points and their usage in Maurits, de Heer et al. (accepted ms)

Timing of divergence events of the Uralic language family



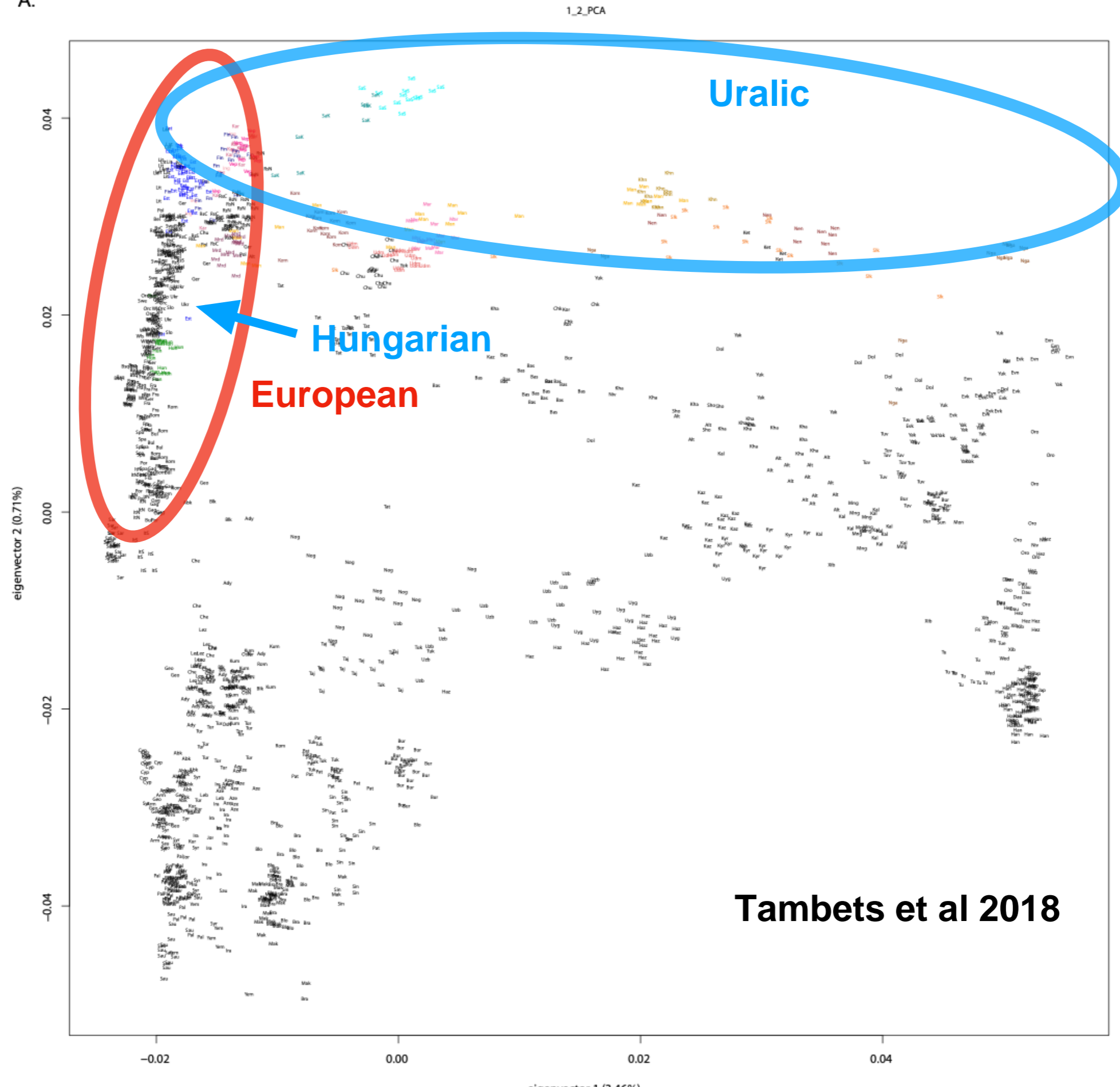




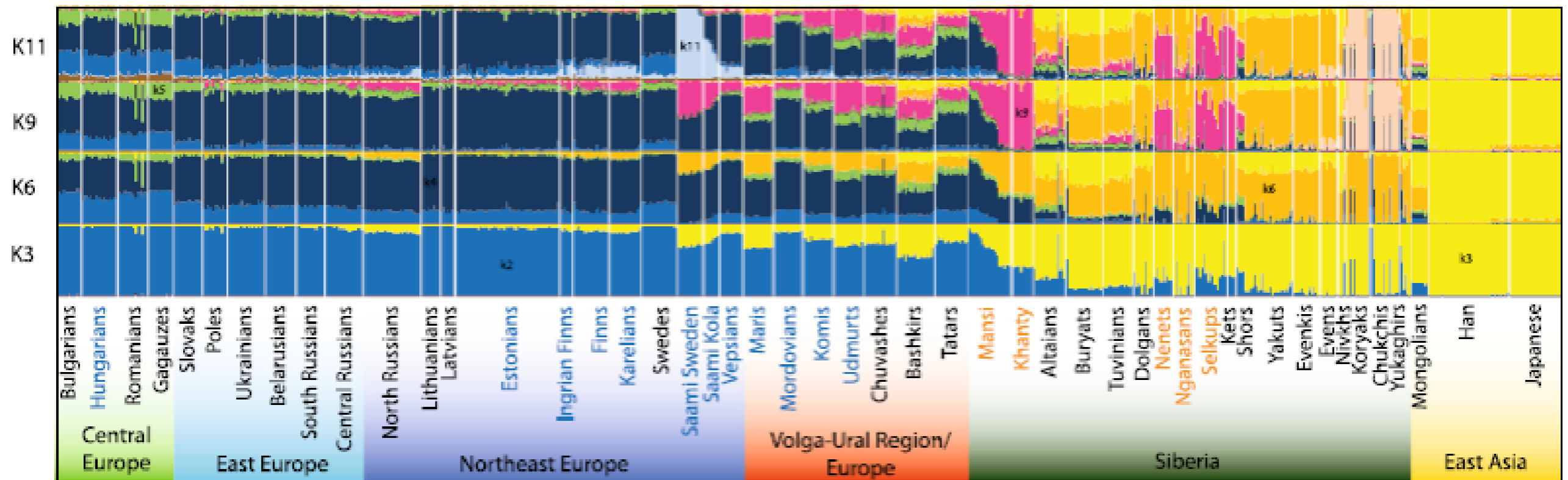
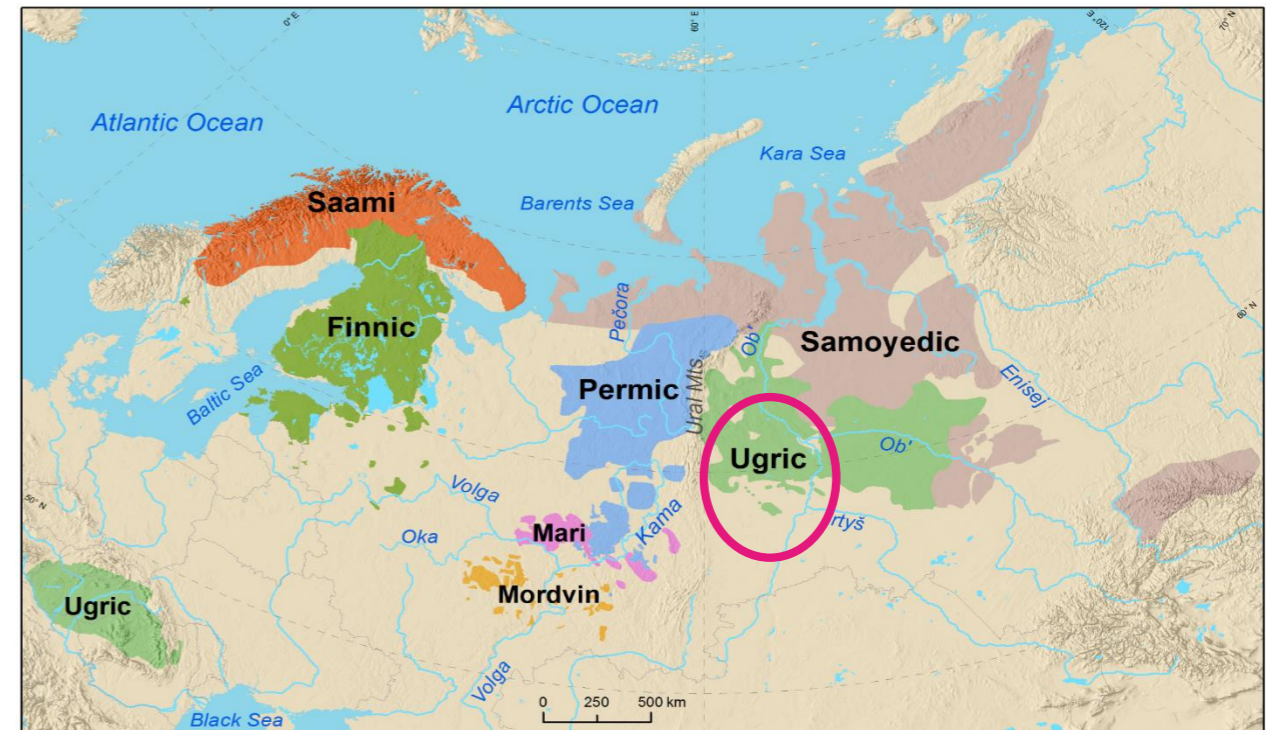
Do the Uralic speaker populations share any specific "Uralic genomic component"?

Autosomal variation of Uralic speaker populations in Eurasian context

A.

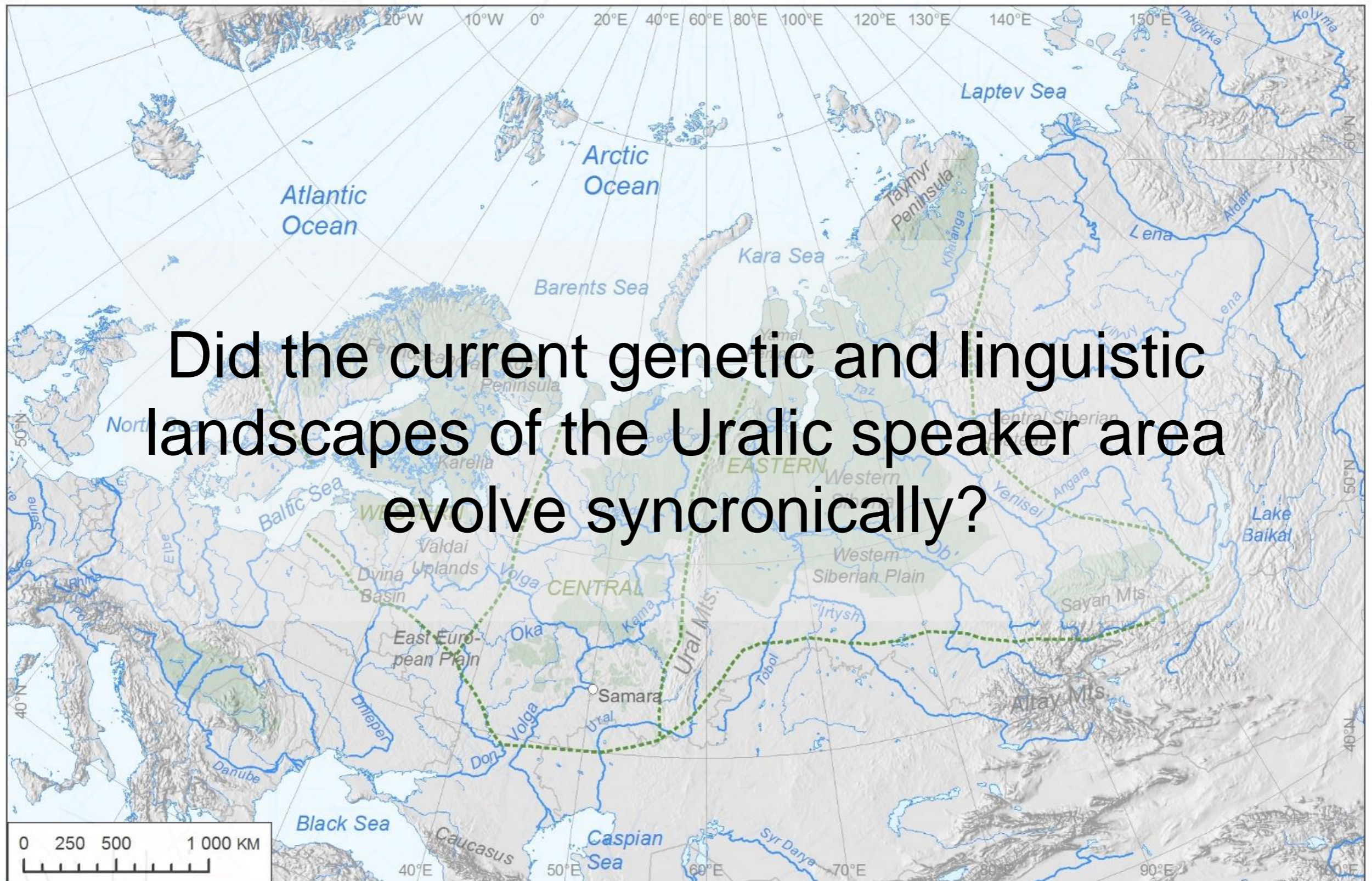


Central Siberian “origin” of the “Uralic genomic component” (autosomes)

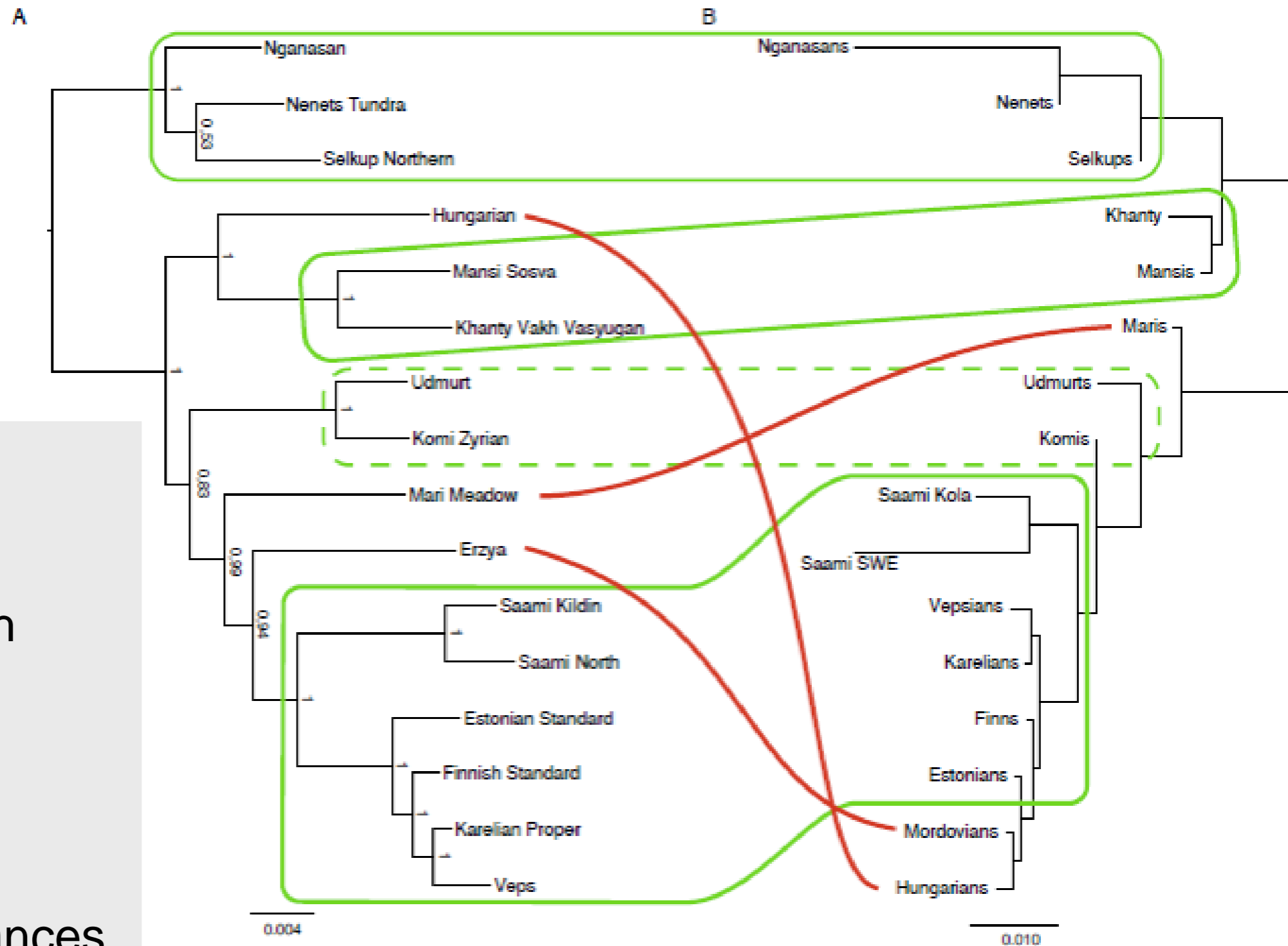


Tambets et al. 2018

Did the current genetic and linguistic landscapes of the Uralic speaker area evolve synchronically?



A match between language and genetic trees



Language tree:
UraLex data,
Bayesian tree with
MrBayes

Genetic tree:
Pairwise F_{st} -distances
Neighbour joining tree

Correlations between linguistic, genetic and geographical distances

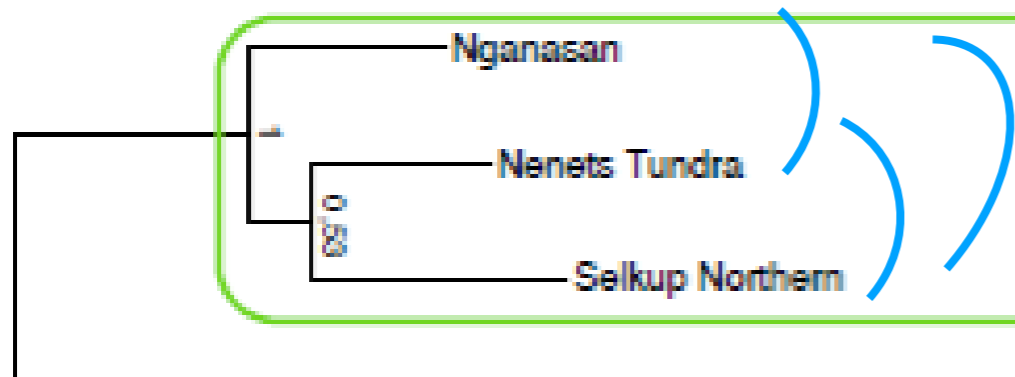
Partial mantel correlation with autosomes $p < 0.05$

But Y-chromosome and mtDNA *ns*

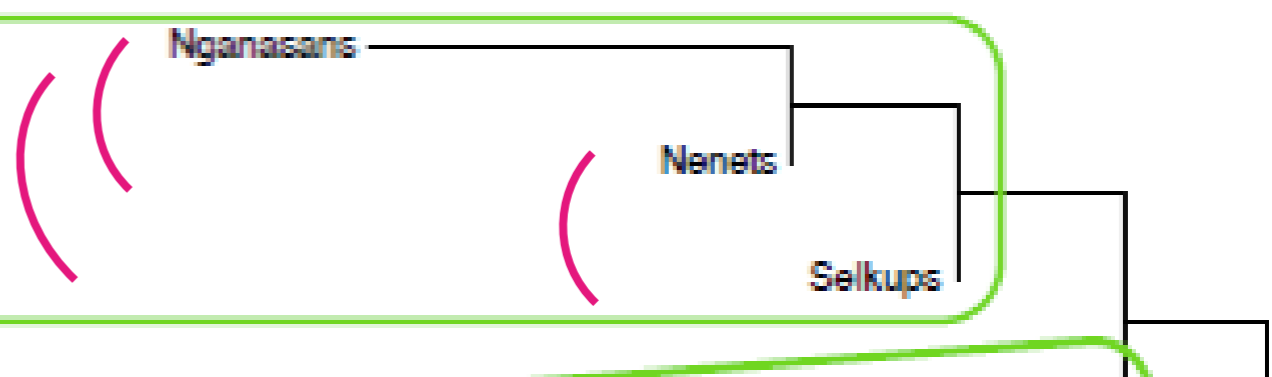
→ Synchronical development of genetic and linguistic landscapes?

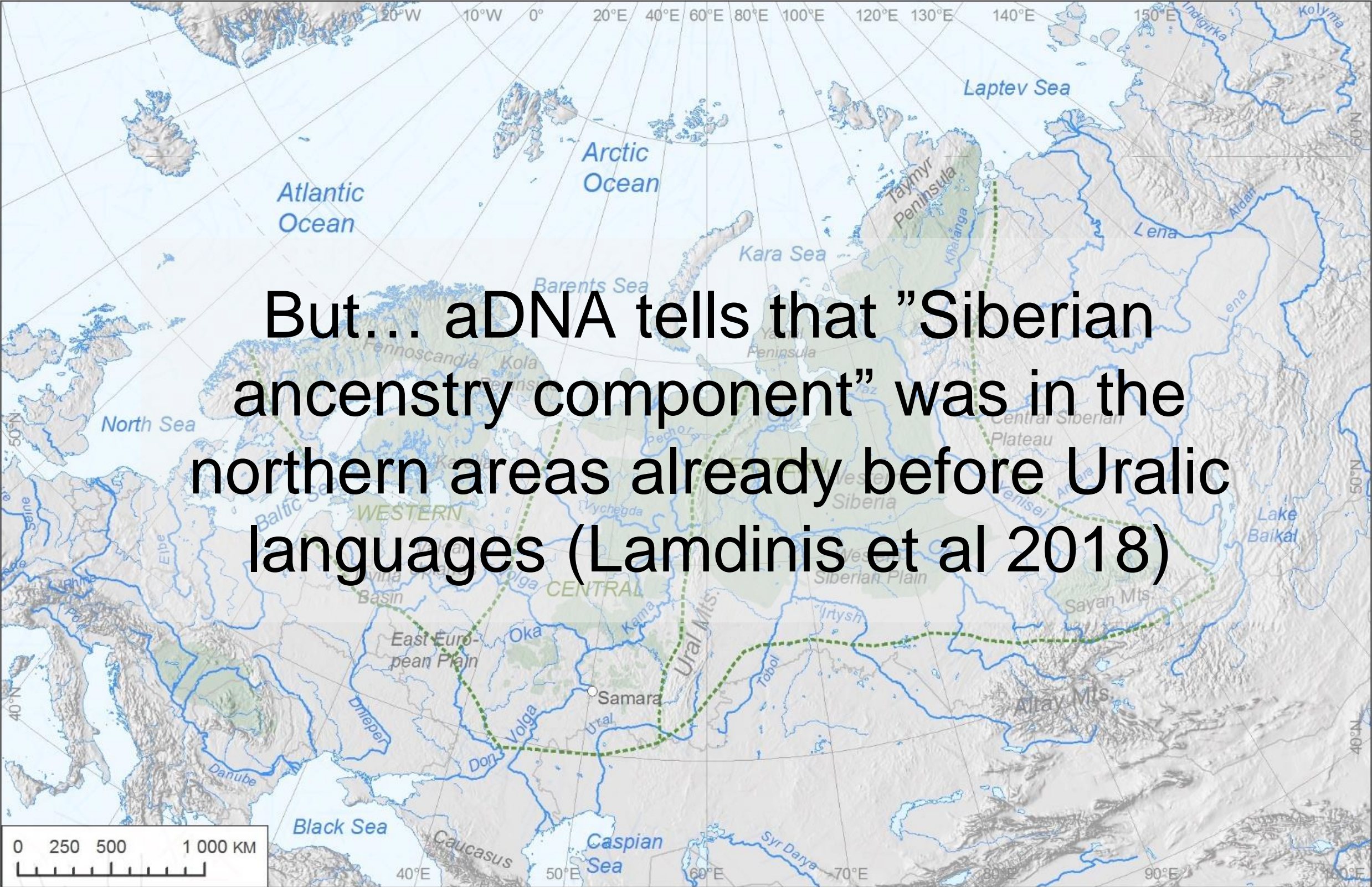


A



B





The map displays Eurasia with a grid of latitude and longitude lines. A dashed green line traces a path from the Taymyr Peninsula in the north, through the Ural Mountains, and across the Siberian Plain towards the Altay Mountains. Major geographical features like the Arctic Ocean, Barents Sea, Kara Sea, Laptev Sea, North Sea, Baltic Sea, Black Sea, Caspian Sea, and various rivers (Lena, Volga, Danube, Syr Darya) and mountains (Ural Mts, Sayan Mts, Altay Mts) are labeled. The text is overlaid on the central part of the map.

But... aDNA tells that "Siberian ancestry component" was in the northern areas already before Uralic languages (Lamdinis et al 2018)

North Siberian genomic component maximizes in Nganasan speakers

.. Nganasan **language speakers** immigrated the area 400 YBP and met speakers of unknown Palaeo-Siberian languages -> Linguistic contact, where "Palaeo-Siberian" speakers shifted their language to a Uralic language?

North Siberian genomic component could originate/is preserved **from/in the geographical area** of current Nganasan. Pure coincidence, that they speak a Uralic language now.

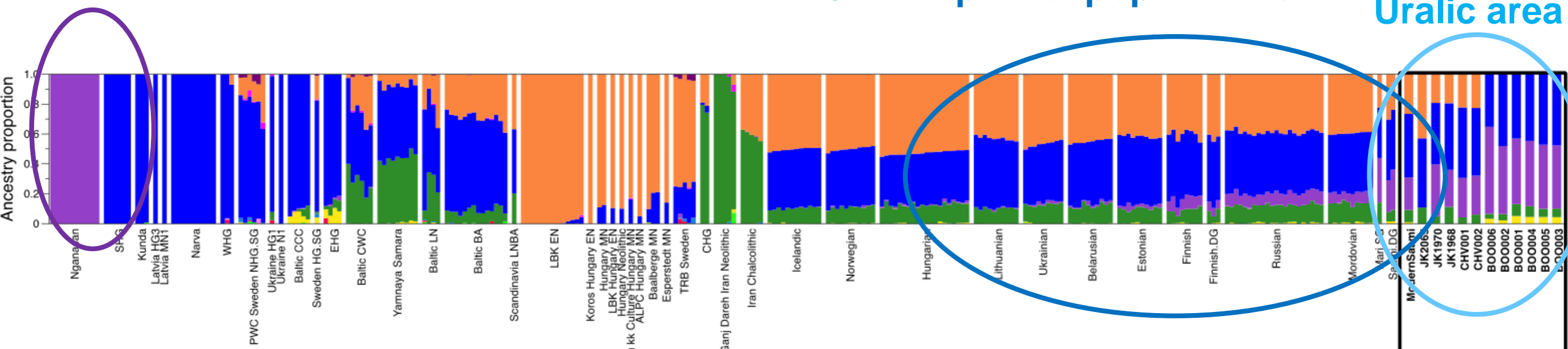


Lamnidis et al. 2018

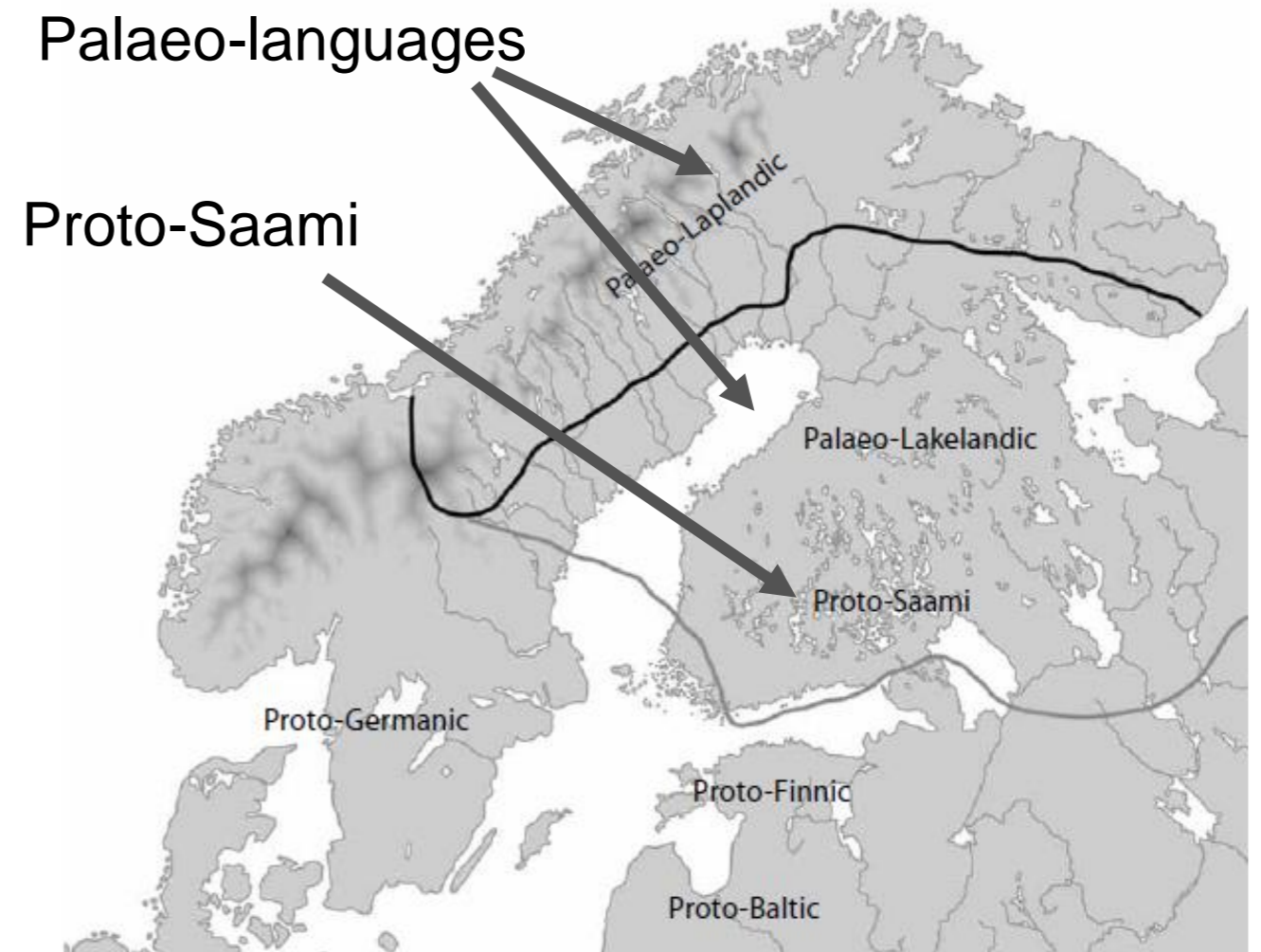
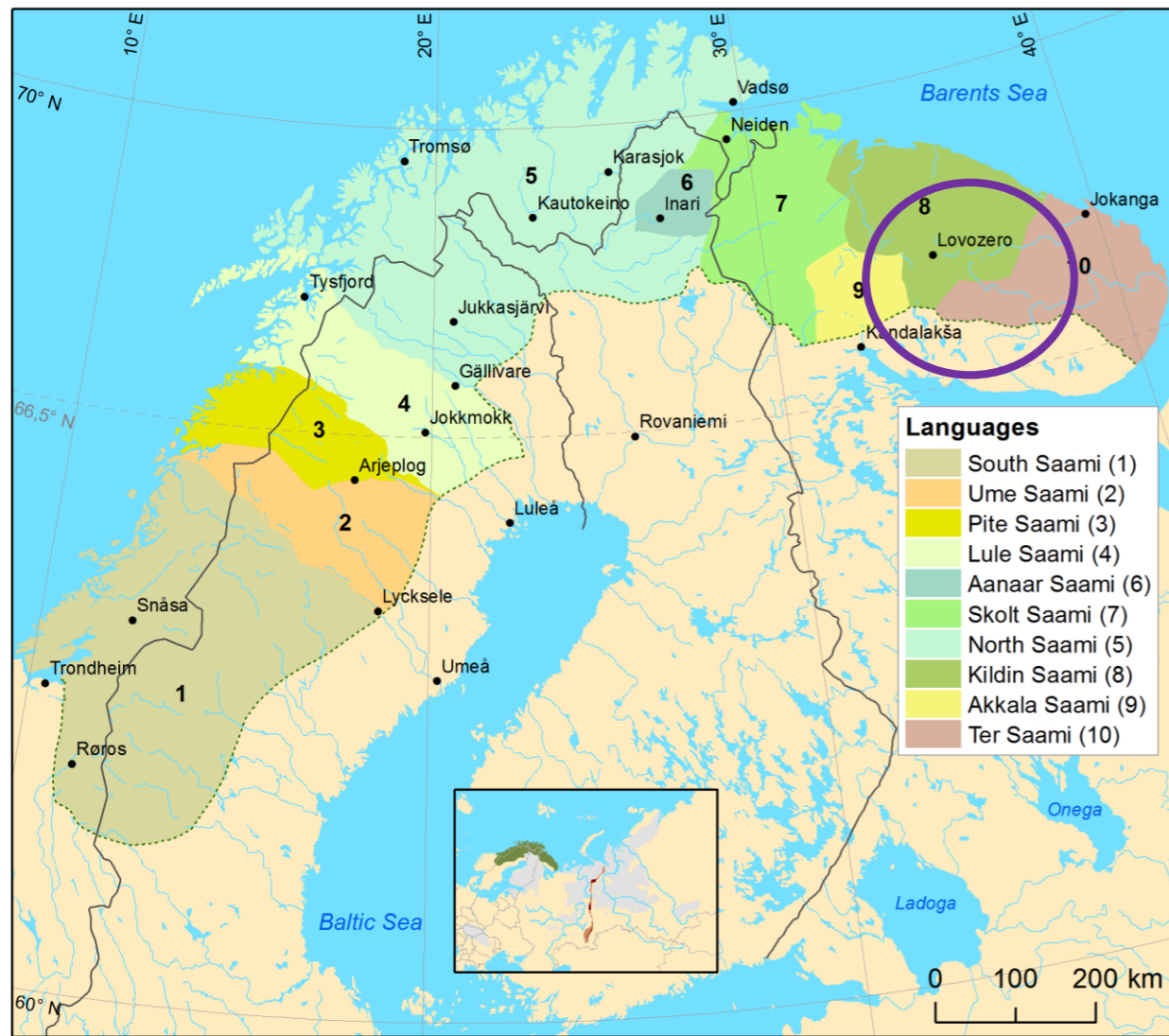
Nganasan

Uralic speaker populations

aDNA from Uralic area



Siberian genomic component to Lapland earlier than Uralic languages



Lamnidis et al. 2018: N haplotype and North Siberian genomic ancestry in Kola Peninsula 3500 YBP

Aikio 2012: Linguistic landscape 2500 YBP. Based on language substrate studies in Saami and Finnish languages and place name studies



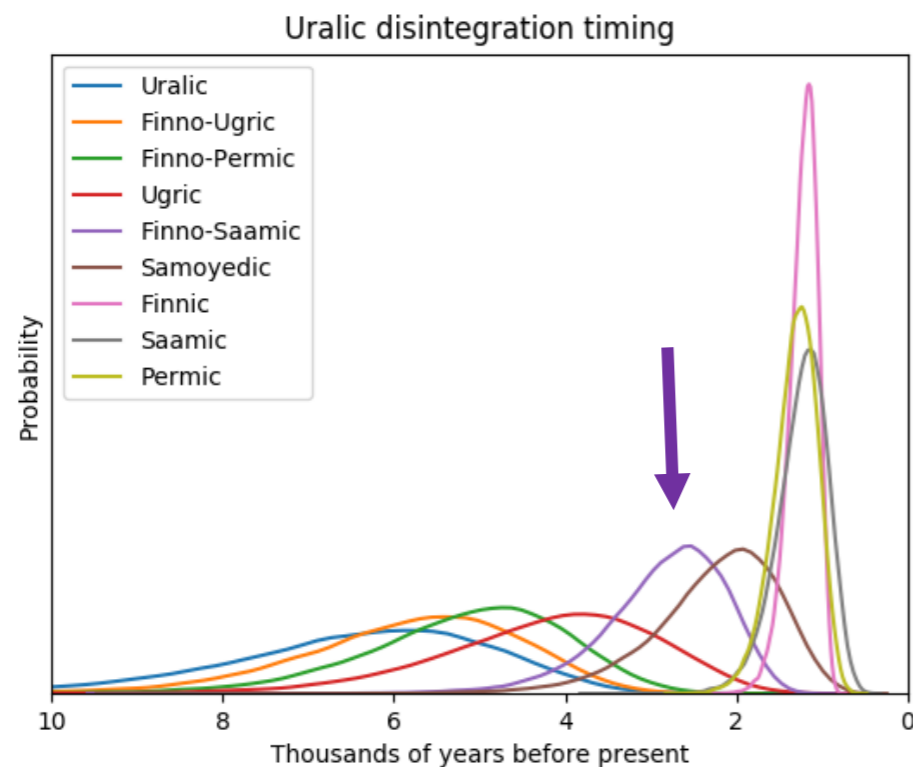
But in Estonia, Siberian influence may have arrived in synchrony to assumed arrival of Uralic languages.

Case Estonia and Finnic languages

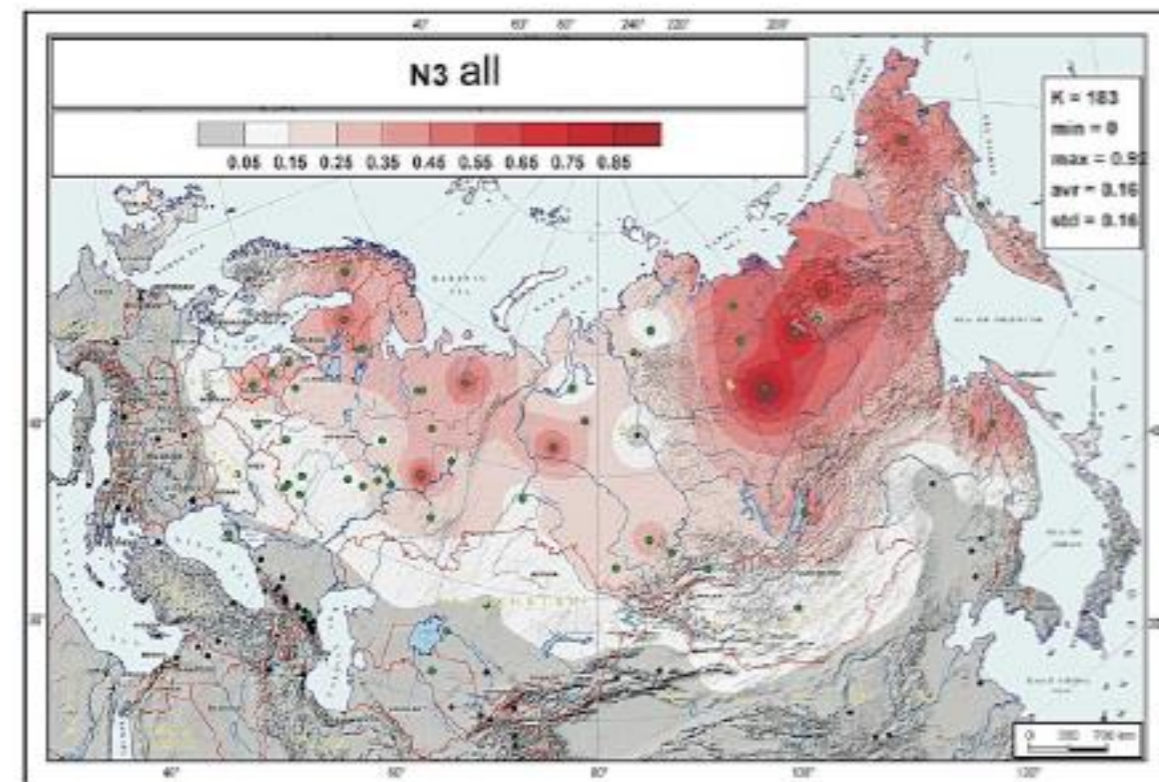
Saag et al. 2019:

N haplogroup arrived to Estonia 2500 YBP

Hist linguistics: Proto-Finnic languages to Baltic Sea coasts from east some 2500 YBP



Vesakoski et al. *ms* Finno-Saami disintegration 3000 YBP; (Proto-)Finnic evolved after that



Illumäe et al. 2016: N3 haplogroup distribution; Siberian origin, Uralic-specific?



Uralic sum up:

Joint spread of Siberian genomic compound and Uralic languages in the south...

...but in the north the "Uralic genetic landscape" predates the Uralic languages.

A map of Eurasia showing linguistic regions. The map includes labels for the Atlantic Ocean, Barents Sea, Laptev Sea, North Sea, Baltic Sea, Black Sea, Caspian Sea, and Syr Darya. It also shows major rivers like the Seine, Elbe, Danube, Dnieper, Volga, Ural, Ob, Yenisei, Lena, and Kolyma. Mountain ranges like the Sayan Mts and Altay Mts are marked. A scale bar at the bottom left indicates 0, 250, 500, and 1000 KM. A grid of latitude and longitude lines is overlaid on the map. The text is overlaid on the map in a semi-transparent white box.

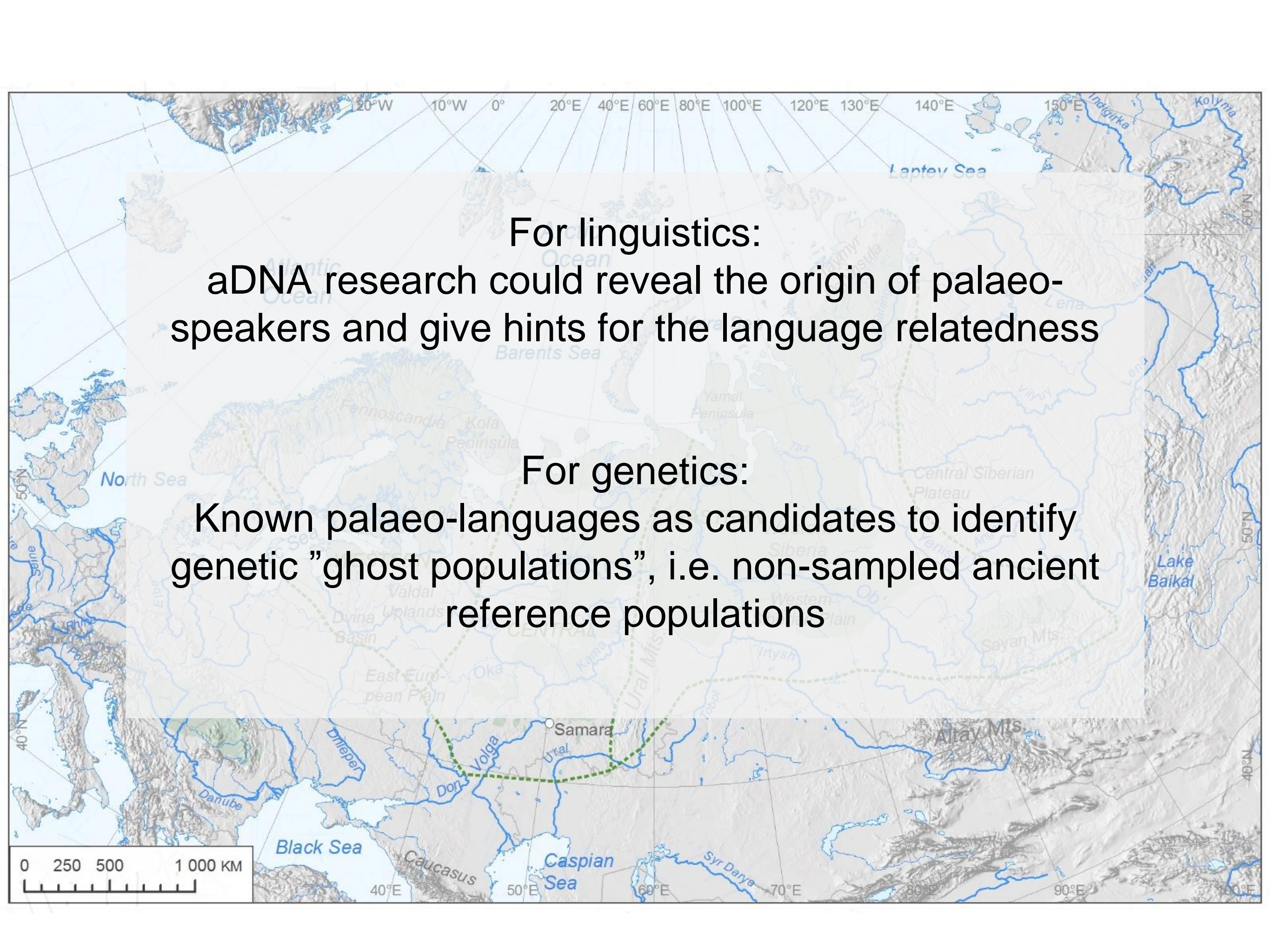
”Glottogenetic” sum up:

Current linguistic landscape is rather recent

Palaeo-language = ”linguistic ghost populations”

Extinct languages belonging to extinct language families

Traces left to current languages, revealed by contact linguistics, linguistic substrate studies and place name studies

A topographic map of Eurasia and surrounding regions, including parts of North America, the Atlantic Ocean, and the Arctic region. The map shows major rivers like the Danube, Volga, and Lena, and geographical features like the Ural Mountains and the Caspian Sea. A scale bar is located in the bottom left corner. Two text overlays are present: one in the upper center and one in the lower center, both in black font on a semi-transparent white background.

For linguistics:
aDNA research could reveal the origin of palaeo-speakers and give hints for the language relatedness

For genetics:
Known palaeo-languages as candidates to identify genetic "ghost populations", i.e. non-sampled ancient reference populations

Kiitos!



Lauri Kettunen

Unni-Päivä Leino



Kaj Syrjänen

Timo Rantanen

Mervi de Heer

Jenni Santaharju

Terhi Honkola