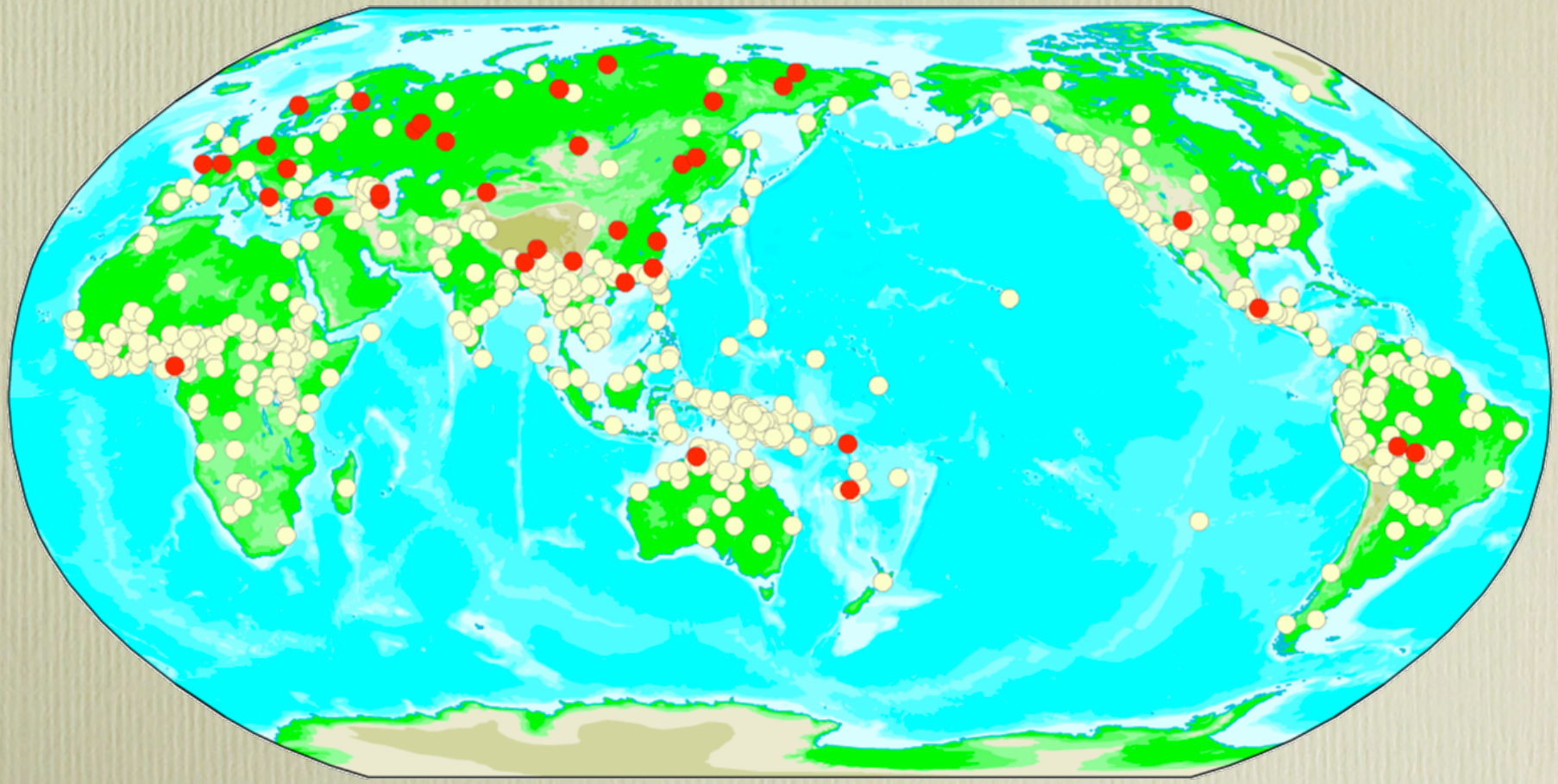
Identifying areas from typological surveys

Michael Cysouw
Max Planck Institute for Evolutionary Anthropology, Leipzig

World Atlas of Language Structures

- 140 worldmaps showing typological characteristics of languages
- Phonology, Morphology, Syntax, Lexicon
- Between 150 and 1,300 languages per map
- In total 58,000 datapoints
- Many maps appear to show areal patterns

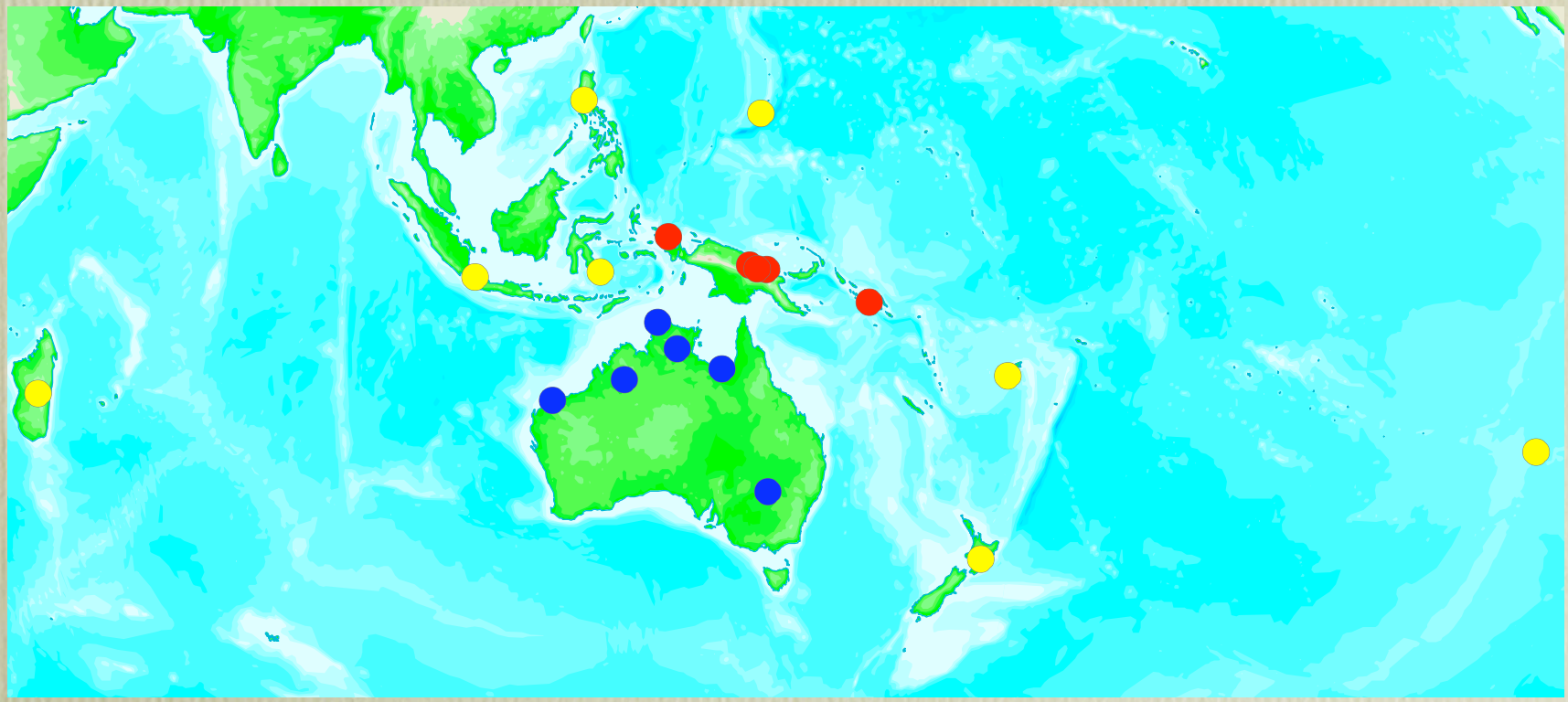
Front rounded vowels



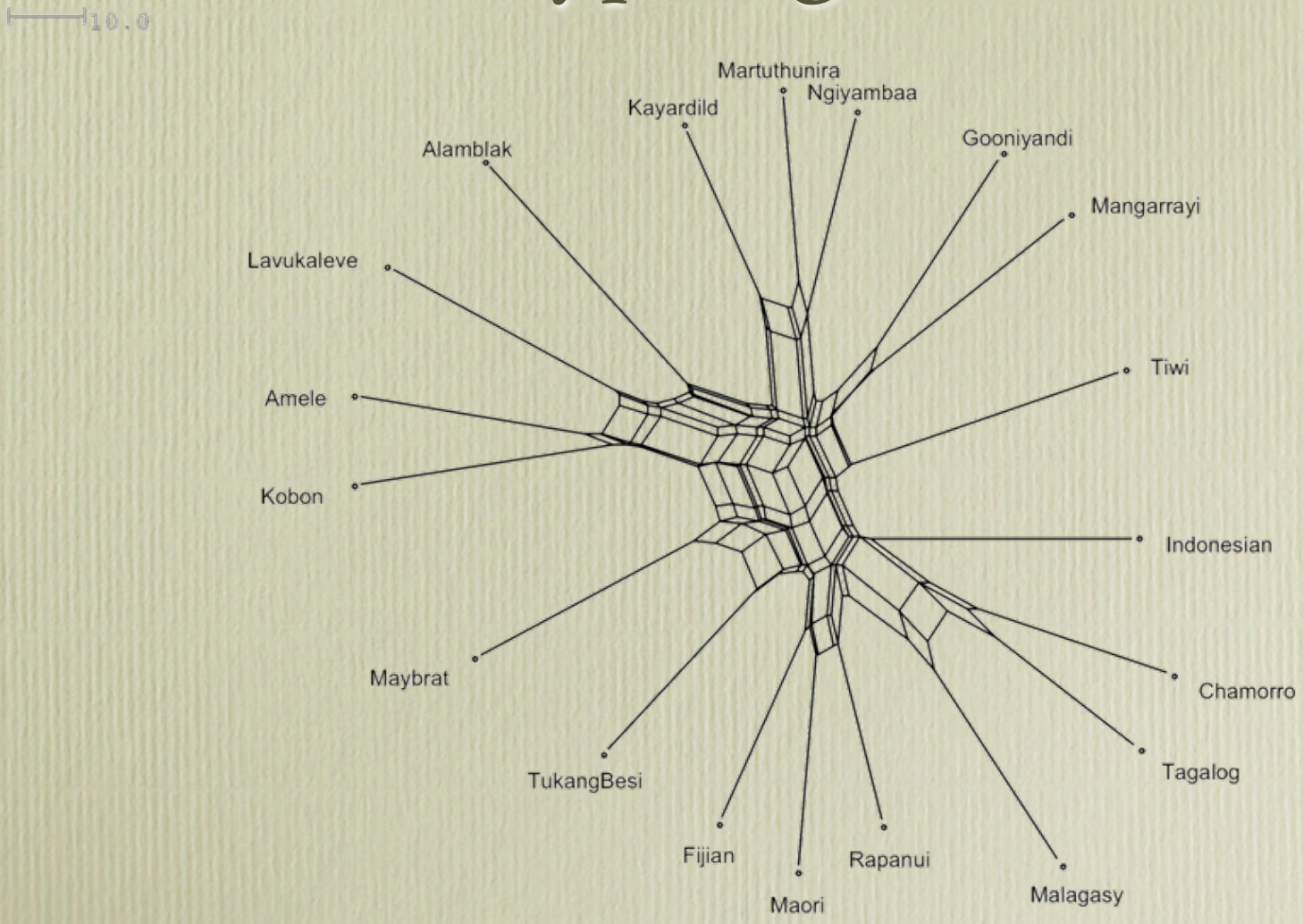
Maddieson, Ian (2005) Front rounded vowels.
In: *World Atlas of Language Structures,* Martin Haspelmath *et al.* Oxford: OUP.

What kind of (areal) patterns
can be found when
all maps are combined?

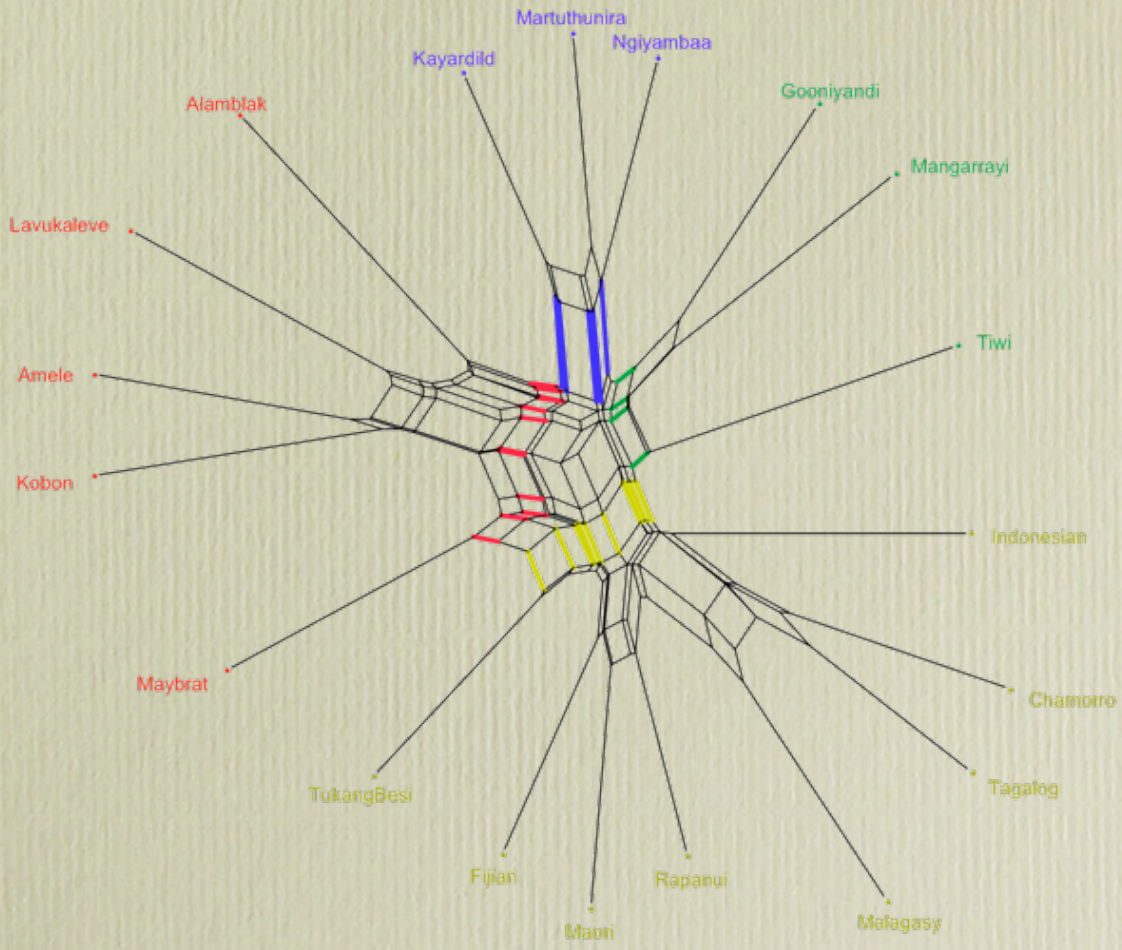
Oceania



NNet of typological distances

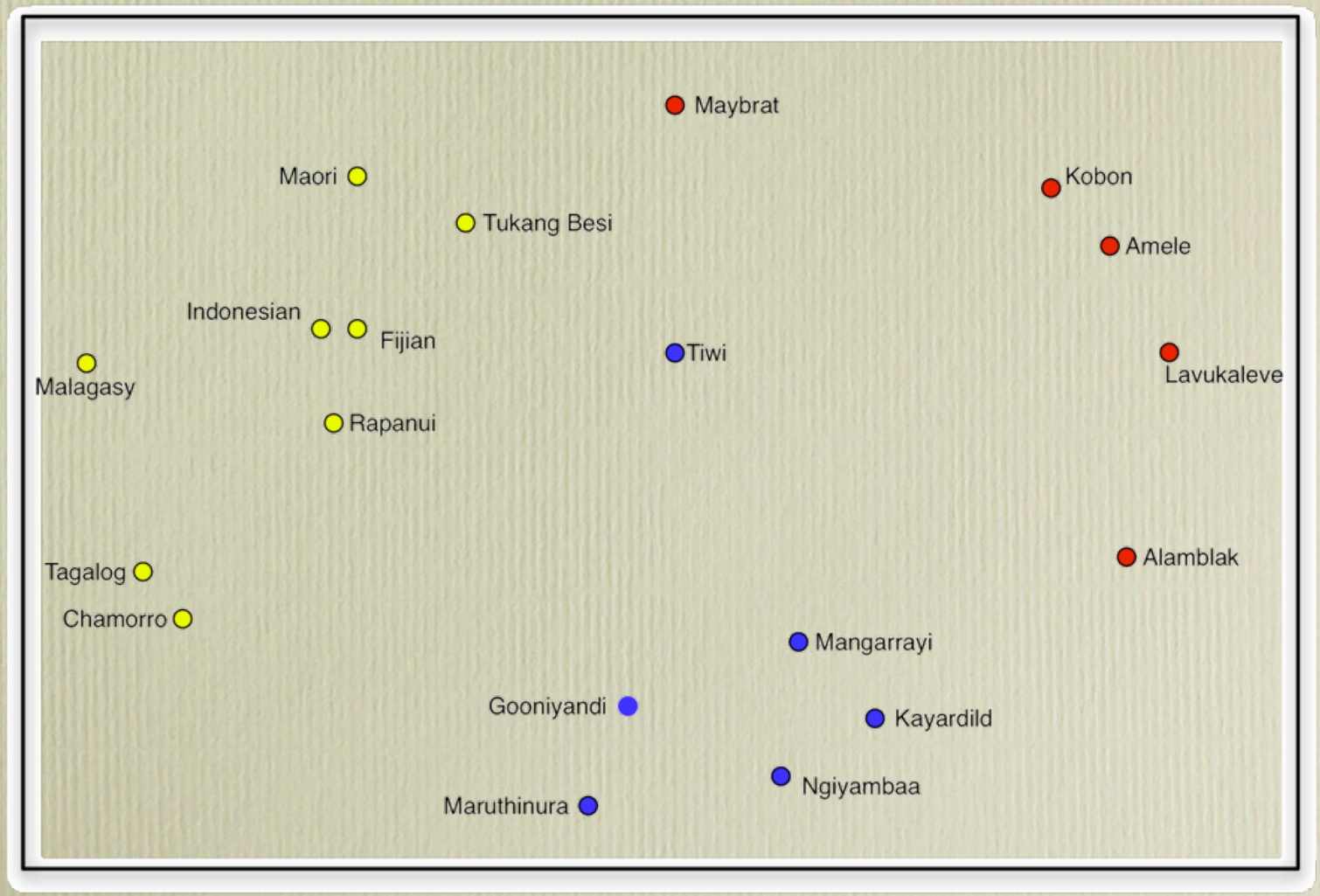


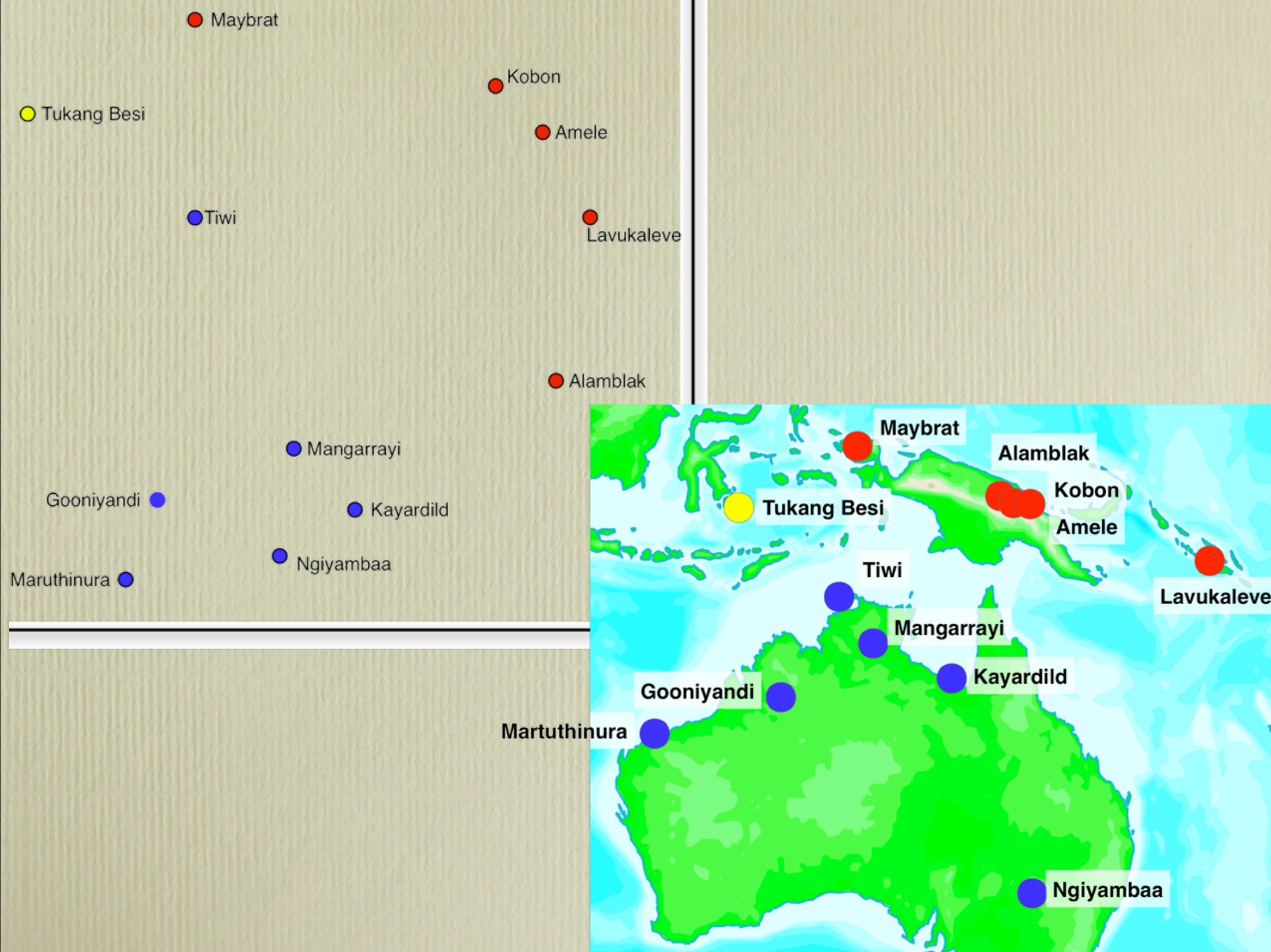
NNet of typological distances



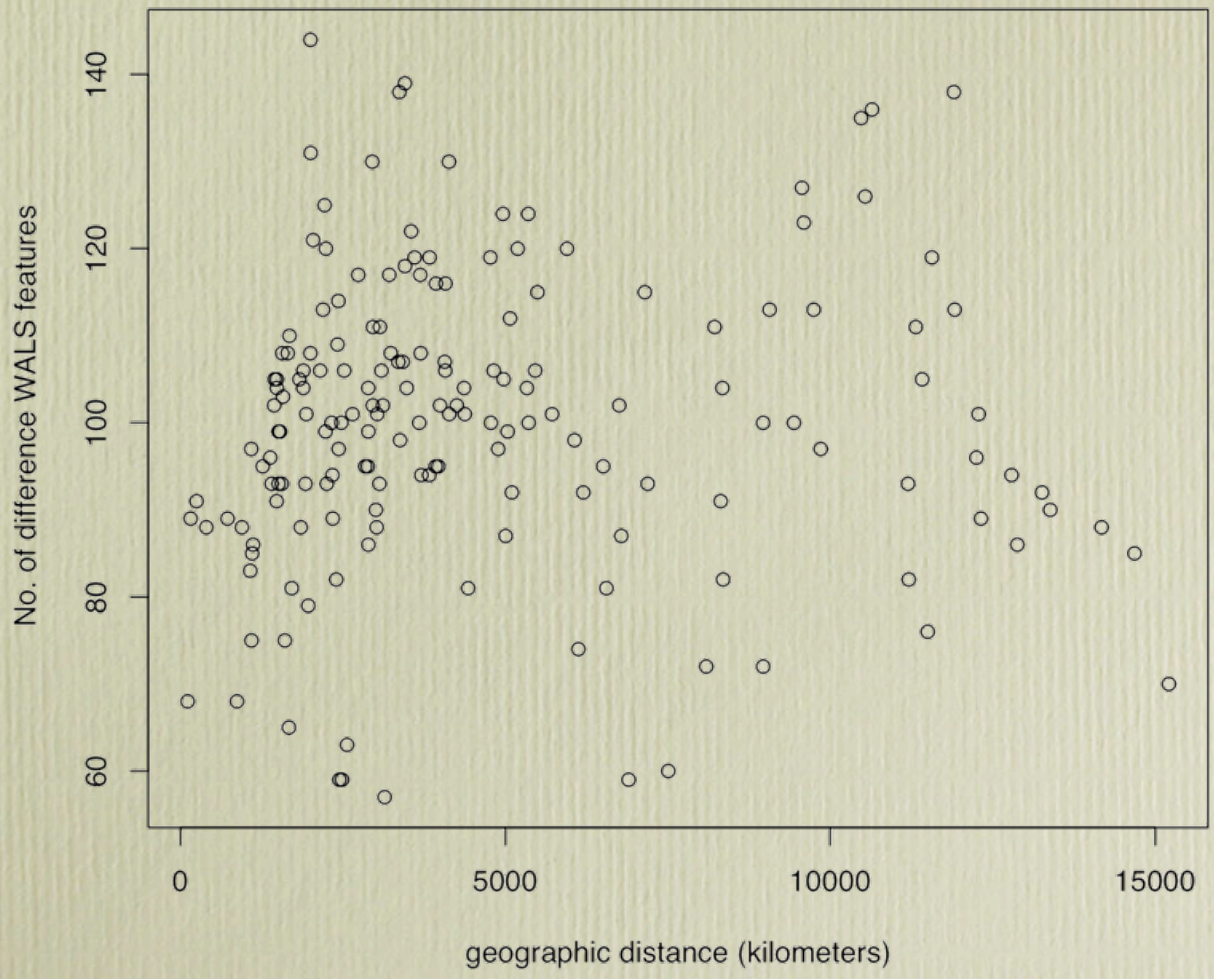
MDS of typological distances

MDS of typological distances




Maybrat
Kobon
Tukang Besi
Amele
Tiwi
Lavukaleve
Alamblak
Mangarrayi
Gooniyandi
Kayardild
Ngiyambaa
Maruthinura
Maybrat
Alamblak
Kobon
Tukang Besi
Amele
Lavukaleve
Tiwi
Mangarrayi
Kayardild
Gooniyandi
Martuthinura
Ngiyambaa

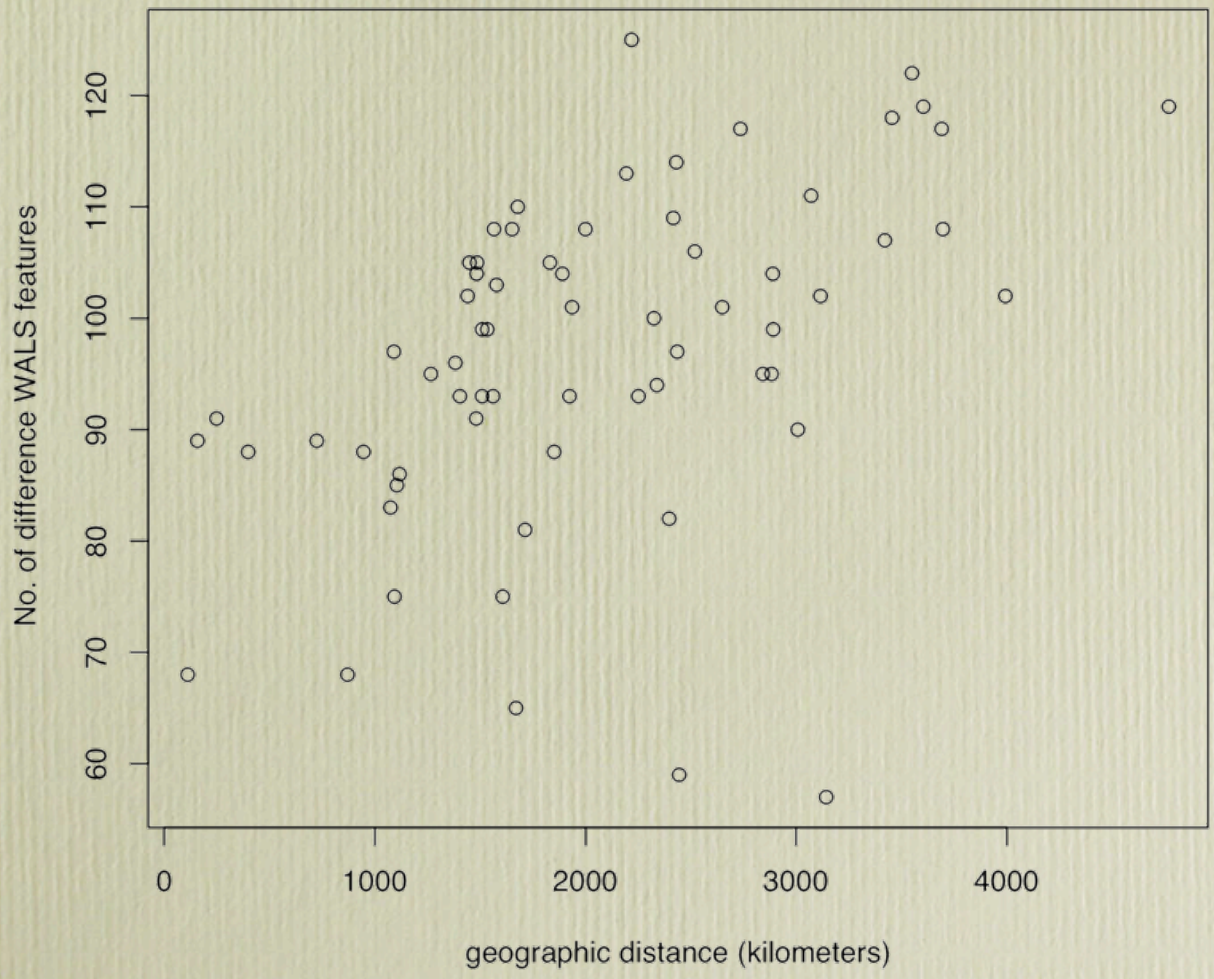
Typology/geography correlation



Mantel test

p = .349

Correlation for selection only



Mantel test

p = .001

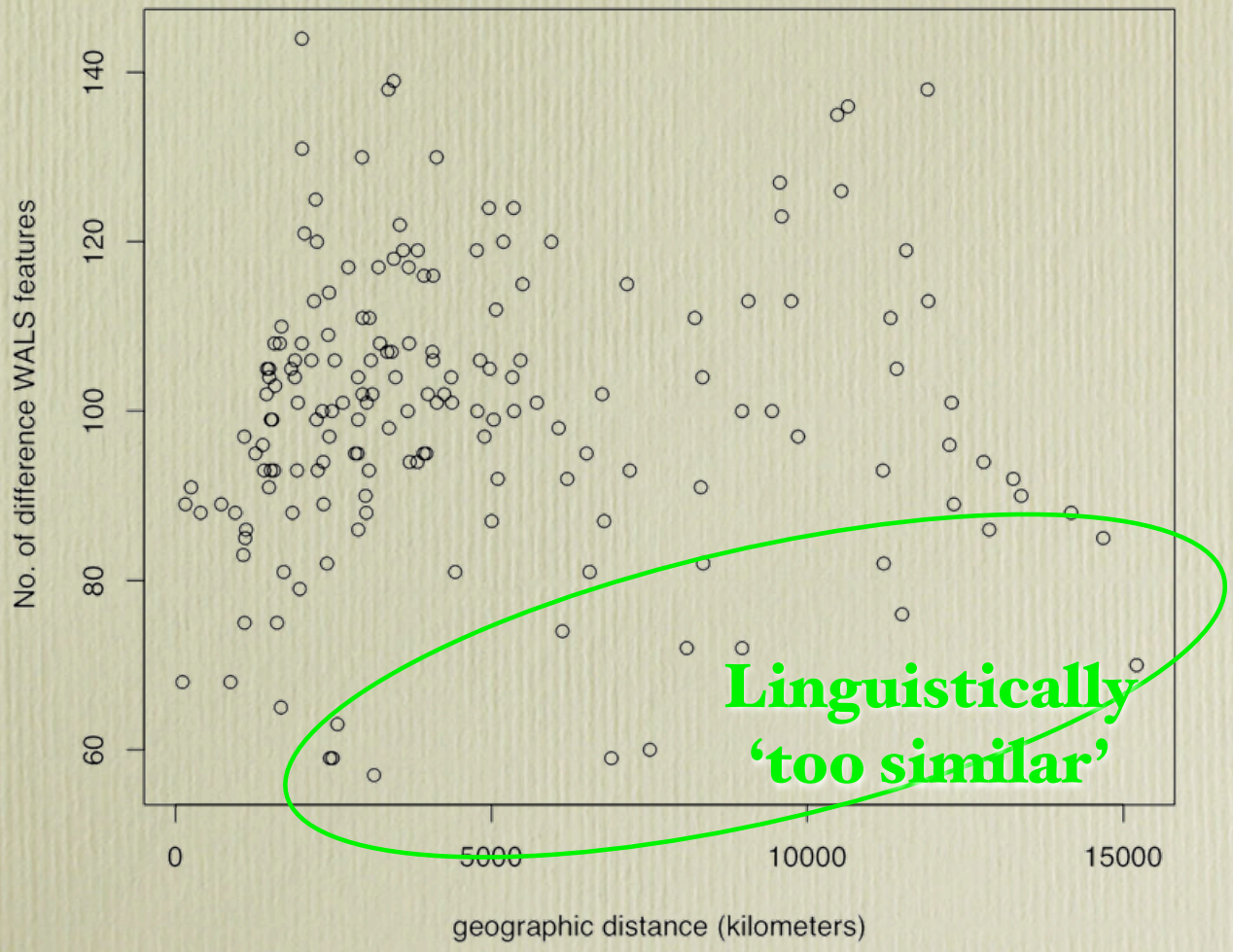
When does correlation improve?

	Pearson's *r*
Nothing removed	.035

When does correlation improve?

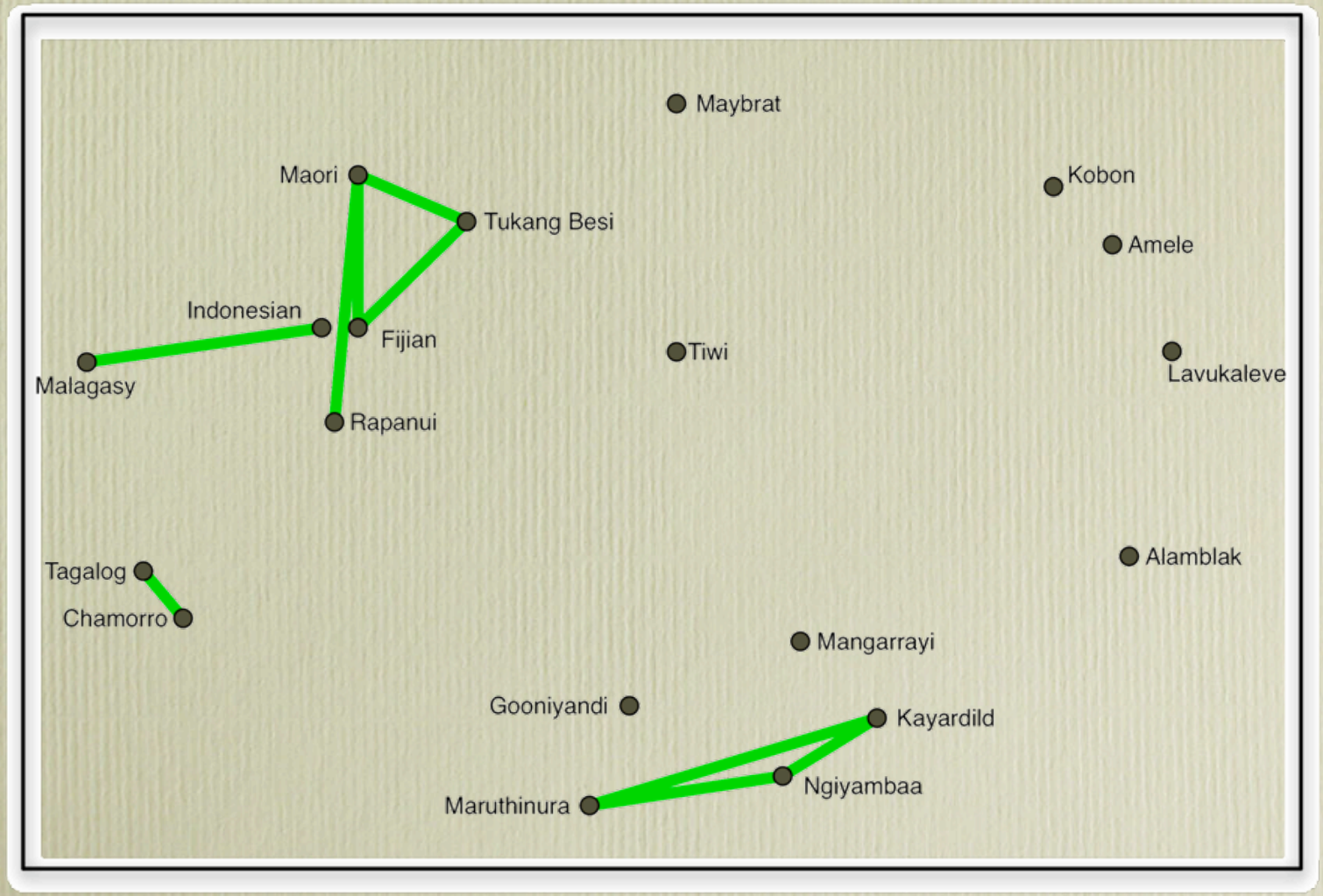
	Pearson's *r*
Nothing removed	.035
Rapanui	.186
Chamorro	.086
Indonesian	.076
Fijian	.073
Tagalog	.071
Maori	.062
Tukang Besi	.048

Investigation typology/geography relation



Linguistically ‘too similar’

Linguistically ‘too similar’



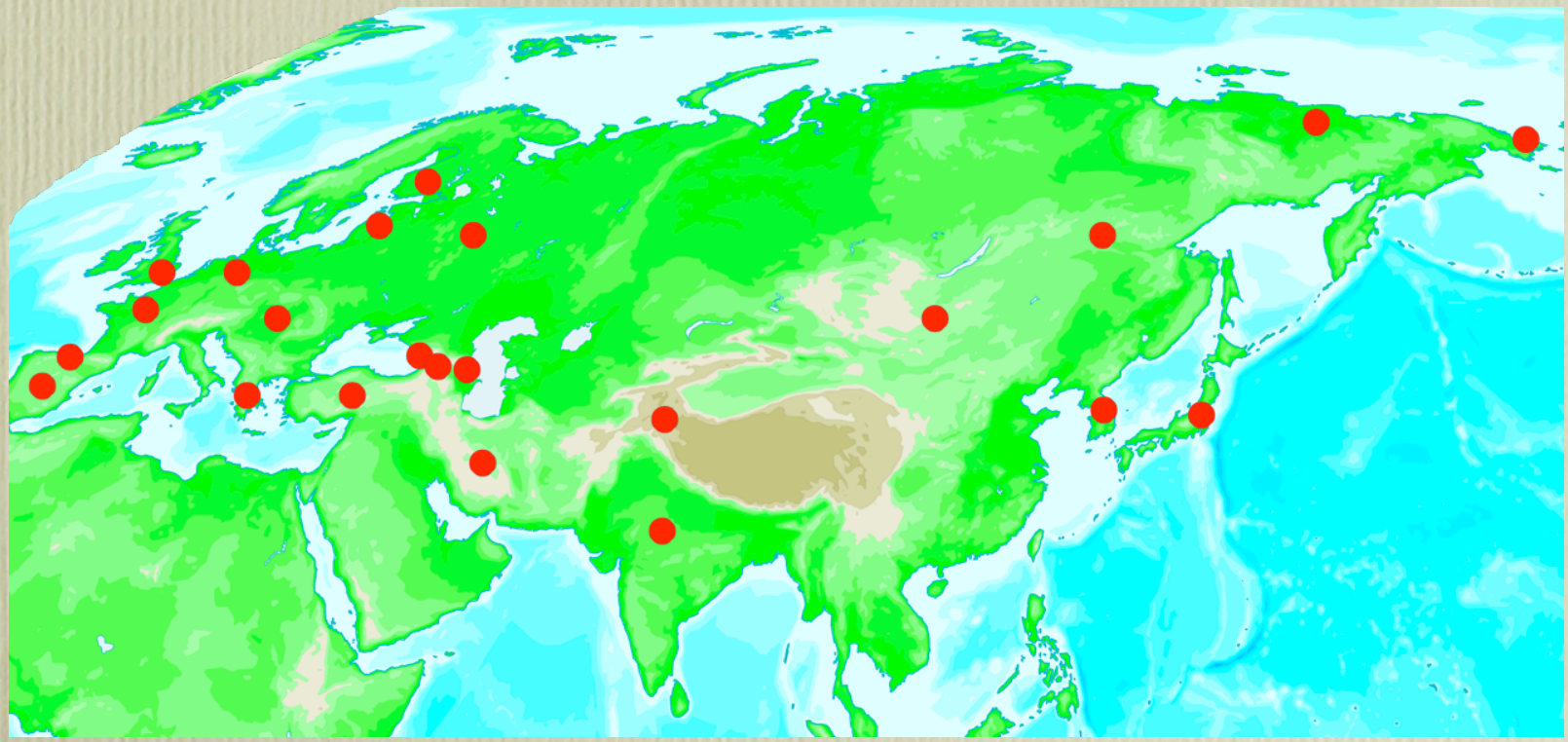
Summary

- Typology is correlated to genealogy
- but: typology is also correlated to geography
- When removing the (genealogically related) Austronesian languages, the typology/geography correlation improves
- The language-pairs that are typologically more similar than expected from geography are genealogically related

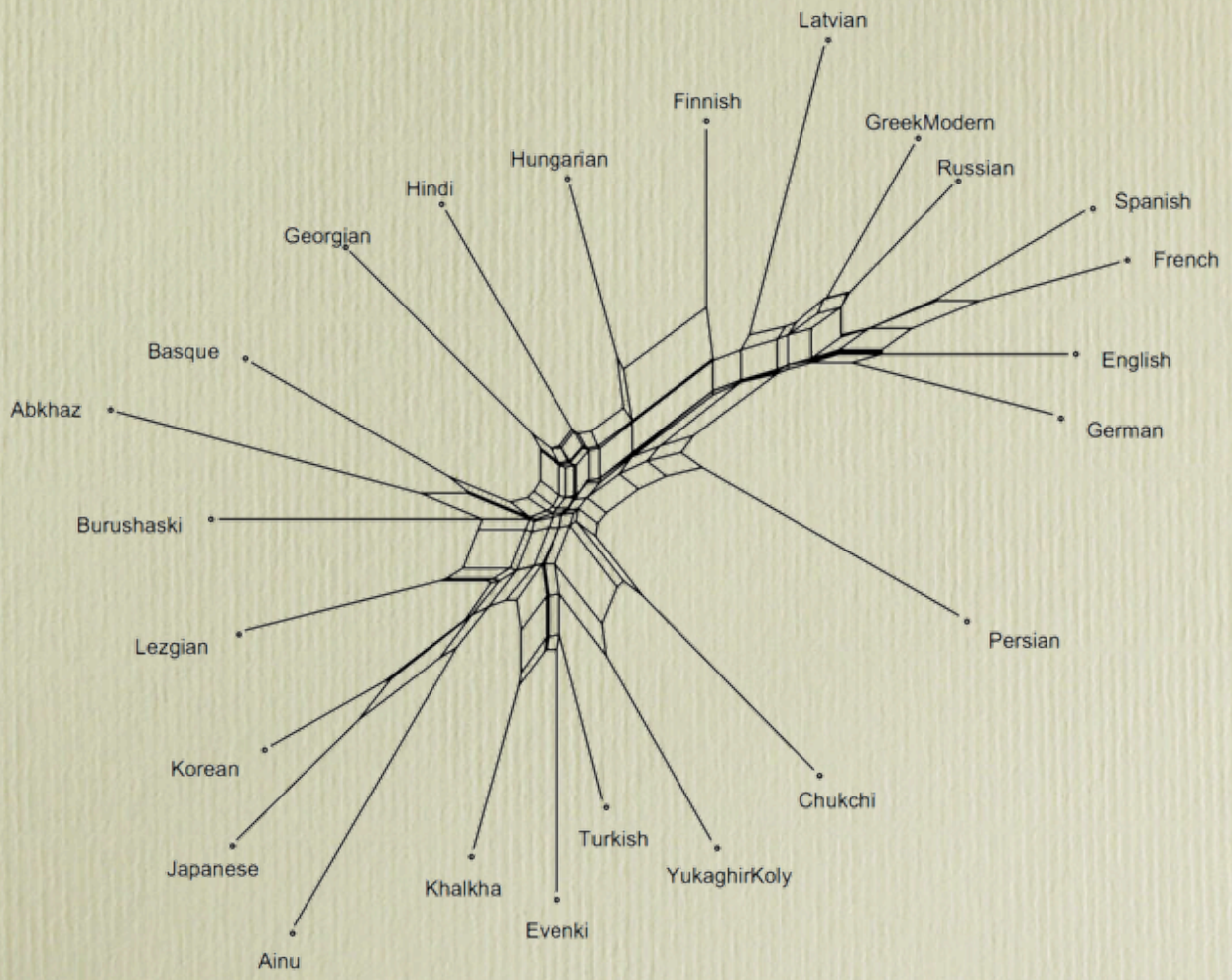
Towards an interpretation

- There are two different kinds of typological interaction between languages:
  - genealogically related languages are similar
  - areally close languages are similar
- In longterm static (areal) interaction typological features diffuse individually, leading to regular geographical clines
- In relatively recent (genealogical) spread bundles of features 'move' together, leading to stronger similarities as expected from geography

Eurasia

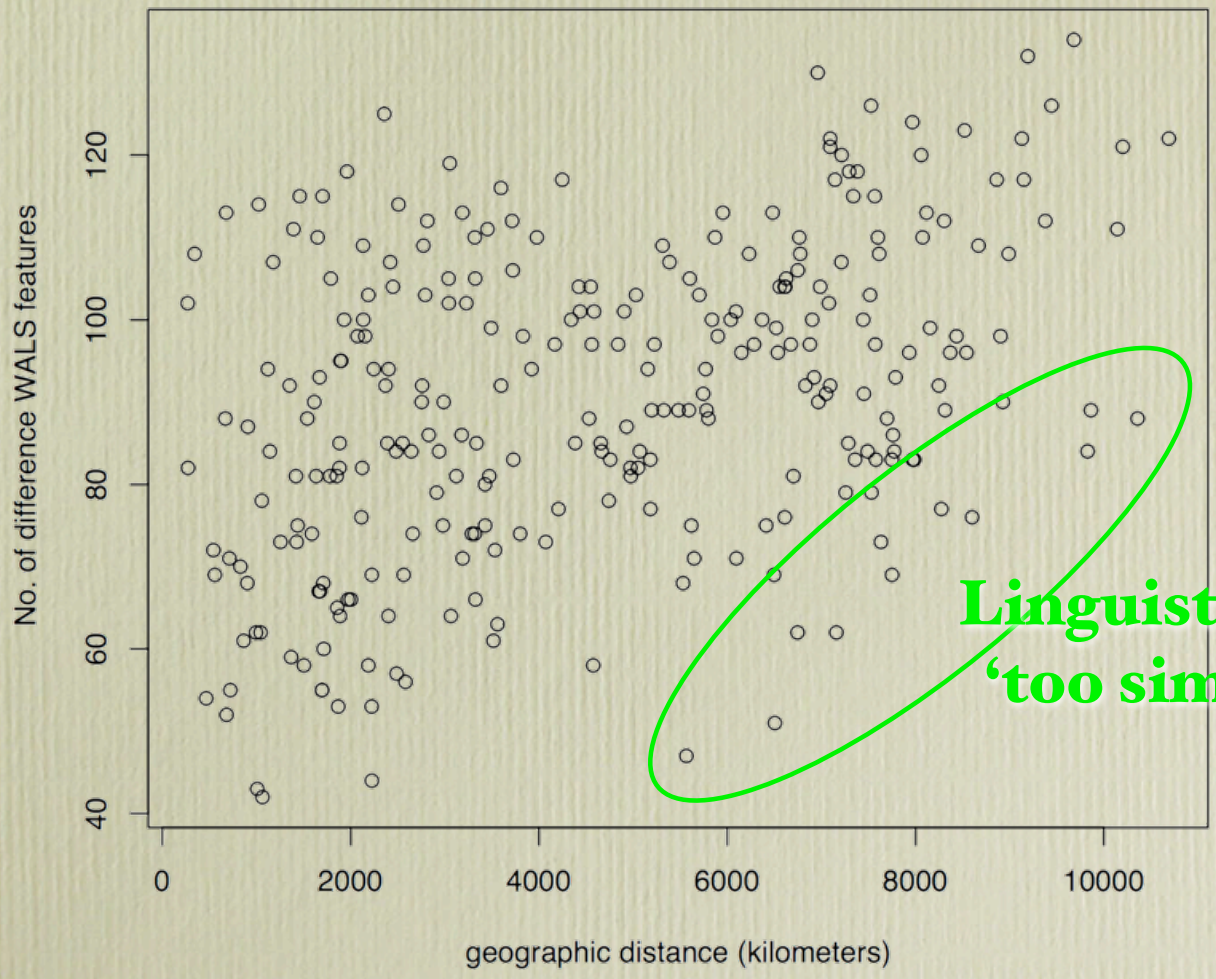


NNet of typological distances



MDS of typological distances

Typology/geography correlation

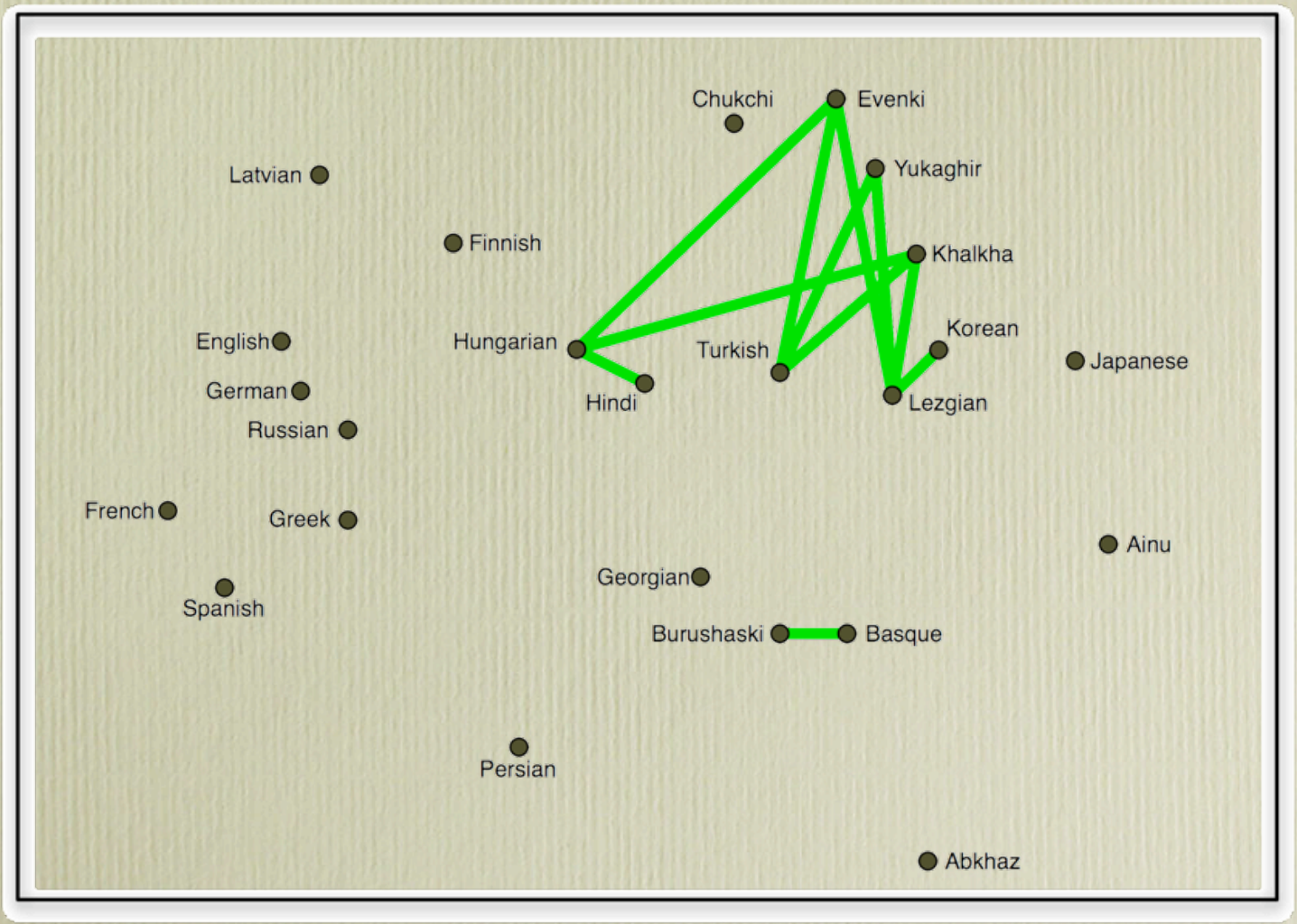


Mantel test

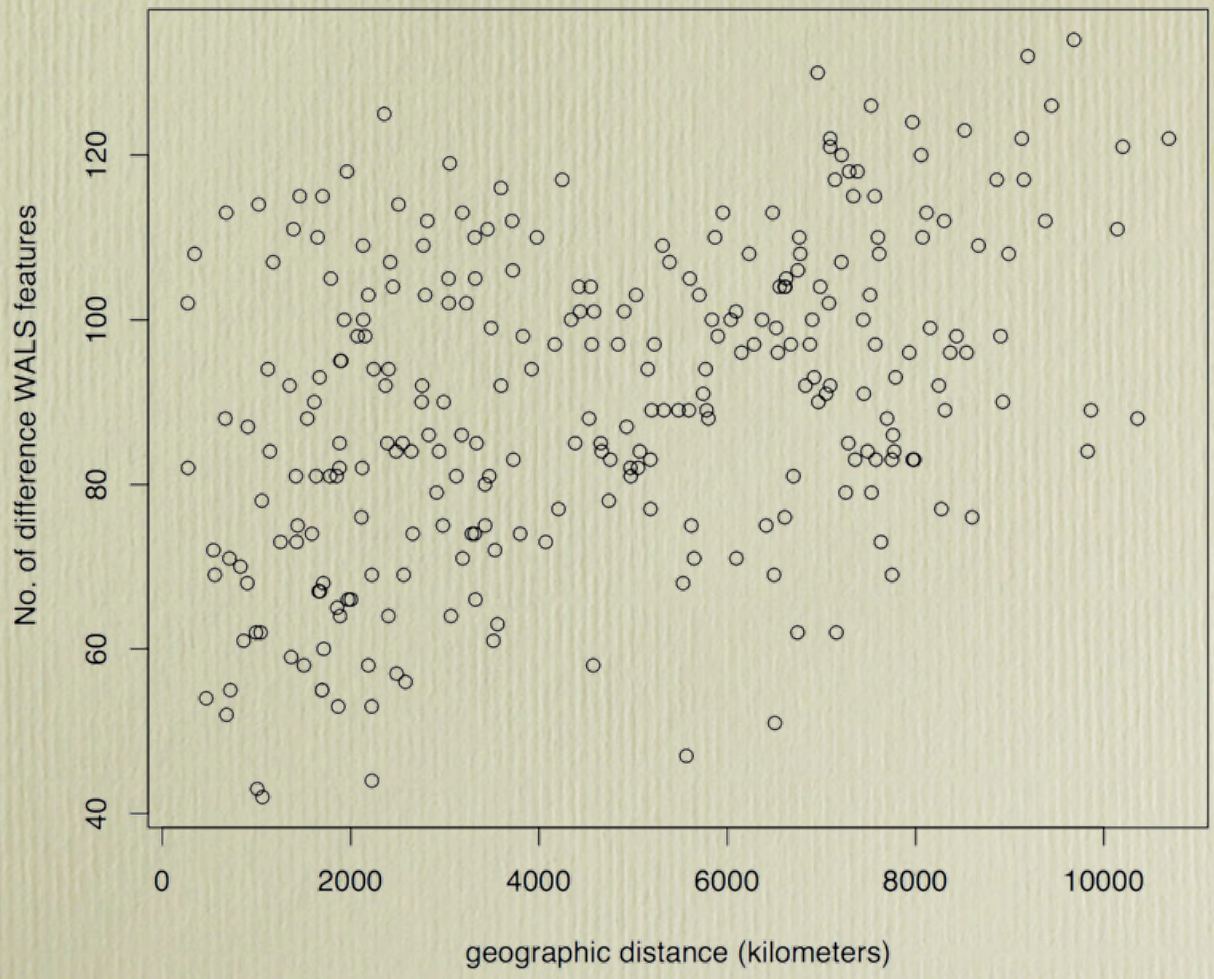
p = .001

MDS of typological distances

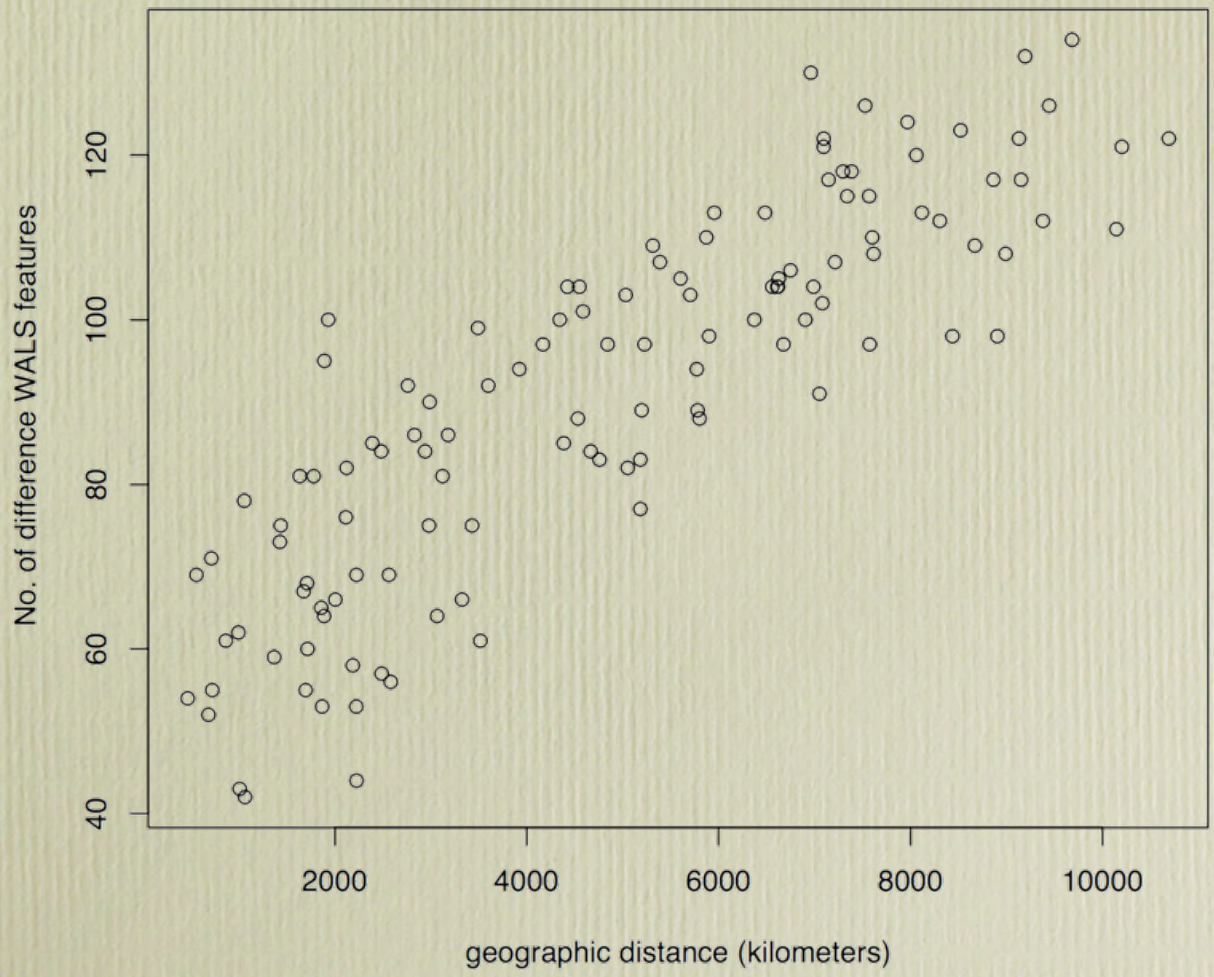
MDS of typological distances



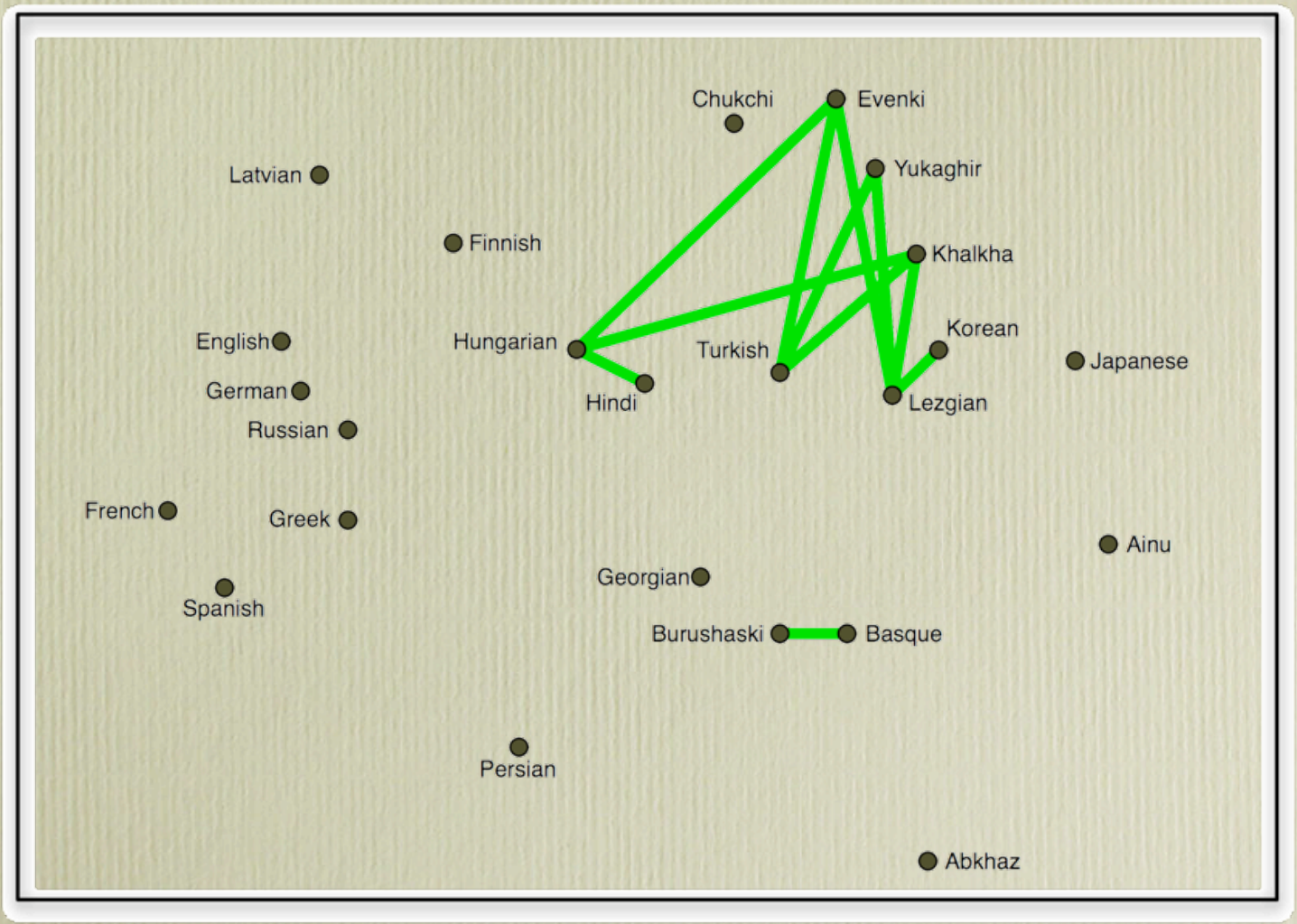
Remove 'worst-fitting' languages



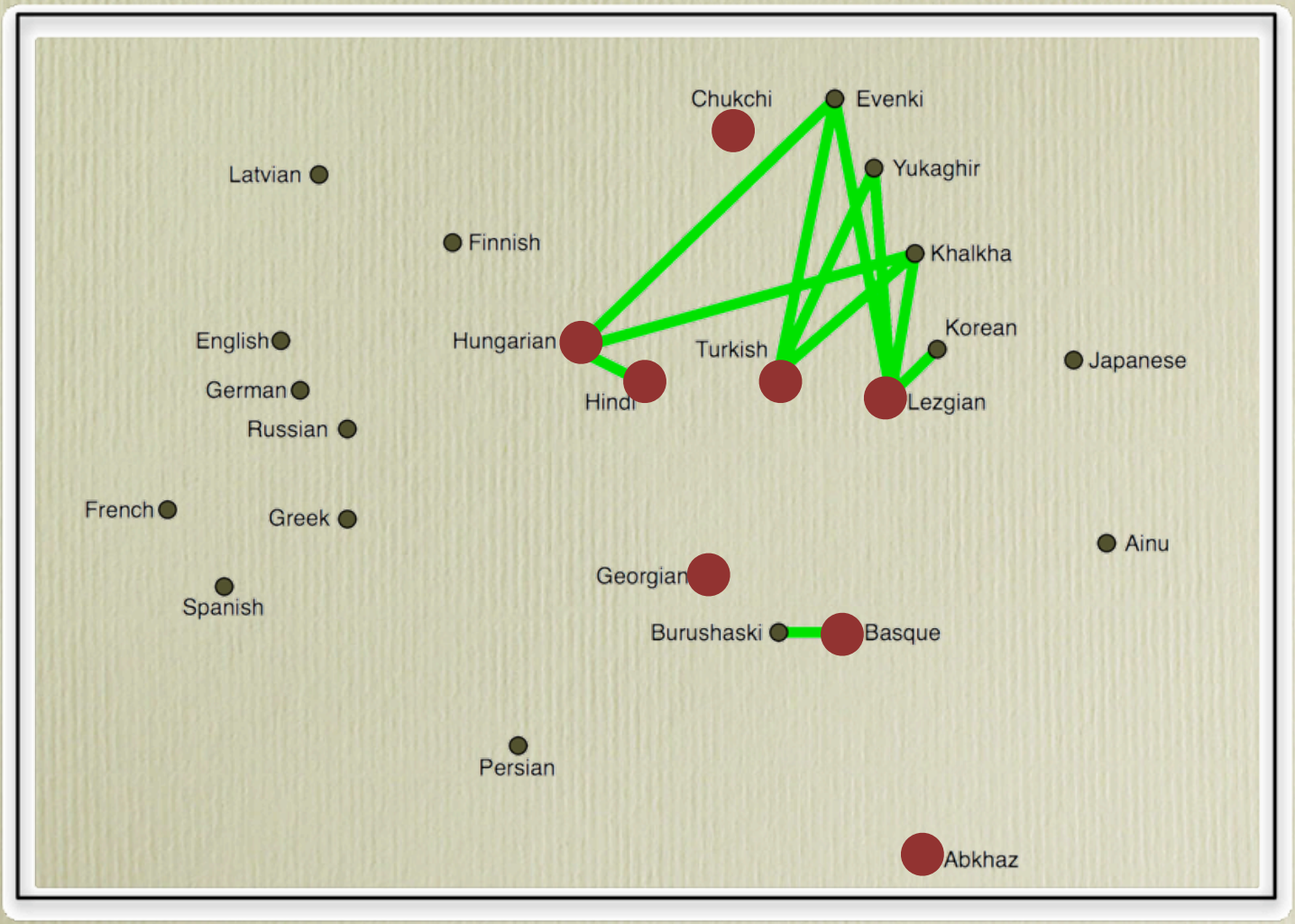
Remove ‘worst-fitting’ languages



MDS of typological distances



MDS of typological distances



Some interpretation

- Turkish and Hungarian are cases of relatively recent movement of whole languages
- But Lezgian (probably) not
- Link Hindi-Hungarian is unclear, and Burushaski-Basque is too cranky a speculation
- Chukchi, Georgian, Abkhaz simply unrelated, both genealogical and areal

Summary

- Investigating language similarity by combining many different characteristics
- The interesting effects are found by looking at the relationship between
  - typological similarity
  - geographical similarity
  - genealogical similarity
- Mismatches between these are most informative

MAX-PLANCK-GESELLSCHAFT