R.R. Sokal: Human Migrations

- 59 independent loci
- -but not uniformly or randomly distributed

-organized database into 5x5° land based quadrats because ethnic info. was organized this way → 85 of them in Europe

Gene Distances

① calculate mean gene frequency
② used to calculate genetic distances
③ ???
④ euclidian - to try and correct for sampling problems
-chose those quadrats that had each of 19 genetic loci samples
-these from Spain, France, Holland, Belgium, England Italy, Greece, Germany,

Ethnic Information

⑤ ~4000 years worth
-from 20 literature
-Dates, Sources, Languages, Location, Migrations Population Size (<20% of references)
~reduced to 420 most important records
⑥ restricted in 3 ways
① only info about language families
② uses 85 quadrats not detailed info
③ some action codes eliminated
⑥ 247 are mvmts (w/ >1100 quadrat changes)
~assume instantaneous assimilation

ARCDISTANCES ~use Cavalli-Sforza & Edwards distance estimates

COMPARISON – Mantel test of matrix correspondence

Comparison of Matrices

- Geographic distances affect <u>both</u> matrices
- ∴ Matrices are not independent
- So... to eliminate spatial autocorrelations

<u>Results</u>

Genetics & Geography correlated - low
Ethnohistory & Geography " - almost all
Genetics & Ethno " - low

Genetics & Ethno - bro = <u>some</u> correlated

What could nonlinearity relationship betw. genetics & ethno due to?

① genetic distance may not incr. past certain pt.

② randomized "history"
- significance of correlation decreases (& lost)

③ randomized geography (kept mvmts same)
- significance lower but still there

<u>Others</u>

① correlations signif. lost if use initial (~1000 BC) population and compare to genetics

② plot correlations over time ...
increase at ~1000 BC
even if leave out 4 big moves... still significant

③ <u>if</u> remove any group ... not much effect except w. Finnick.
Probably because of peculiar distribution of w. Finnicks.