

R.R. Sokal: Human Migrations

- 59 independent loci
- but not uniformly or randomly distributed
- organized database into $5 \times 5^\circ$ land based \rightarrow 85 of them quadrats because ethnic info. was in Europe organized this way

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

① ~ 4000 years worth

- from 2^o literature

- Dates, Sources, Languages, Location, Migrations
Population Size (20% of references)

- reduced to 420 most important records

② restricted in 3 ways

③ only info about language families

④ uses 85 quadrats not detailed into

⑤ some action codes eliminated

⑥ 247 are mutants (w/ > 100 quadrat changes)

- assume instantaneous assimilation

ARC DISTANCES ~ use Cavalli-Sforza & Edwards distance estimates

COMPARISON - Mantel test of matrix correspondence

Comparison of Matrices

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

Results

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

Genetics & Ethno - Geo = some correlated

What could ^{non-linearity} relationship betw. genetics & ethno due to?

① genetic distance may not incr. past certain pt.

② randomized history

- significance of correlation decreases (a lot)

③ randomized geography (kept munts same)

- significance lower but still there

Others

① correlations signif. lost if use initial (4000 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 if remove any group... not much effect except w. Finnick.

Probably because of peculiar distribution of w. Finnick's.