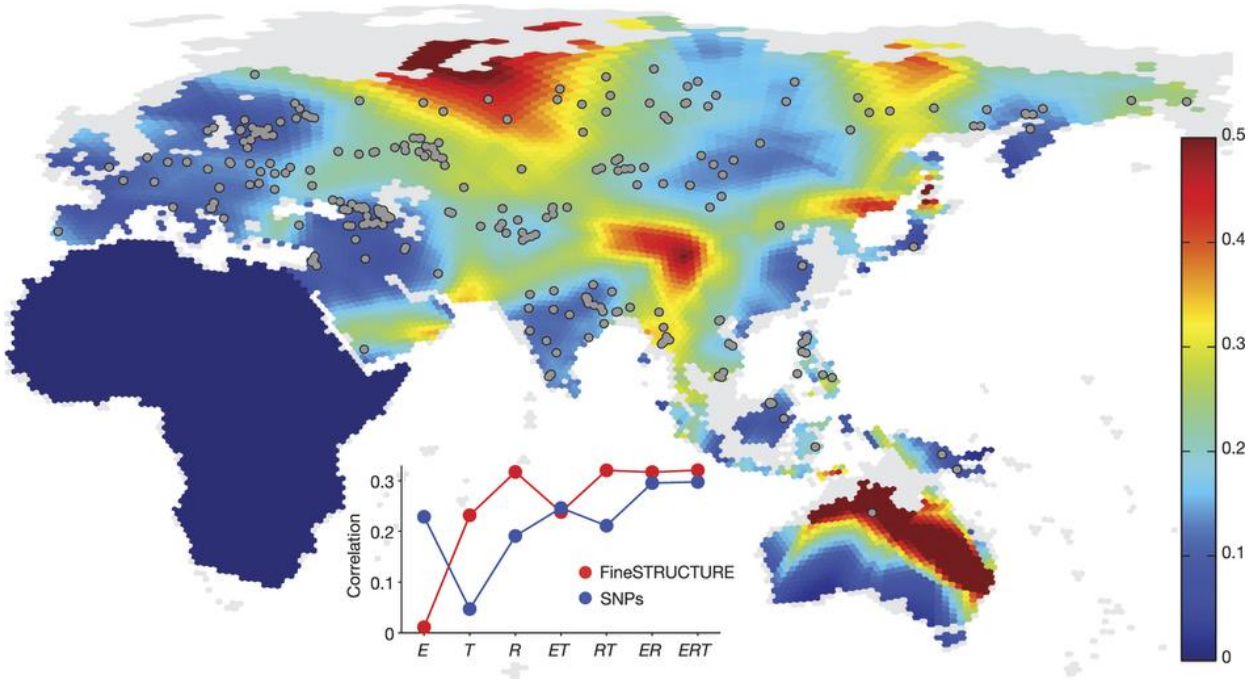
Quantifying relationships among populations: Many populations

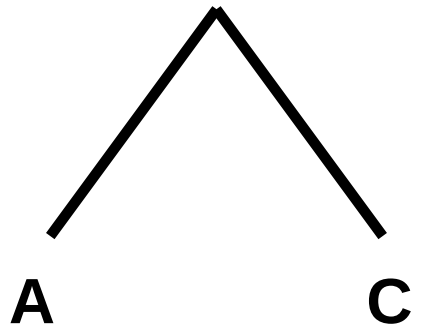


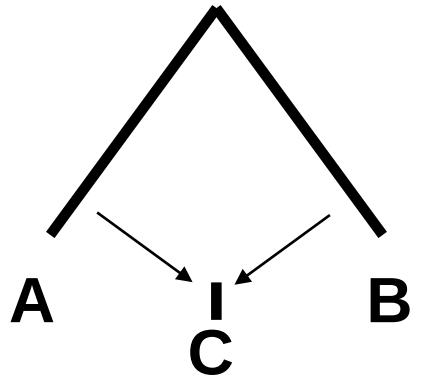
Andrea Manica

UNIVERSITY OF CAMBRIDGE

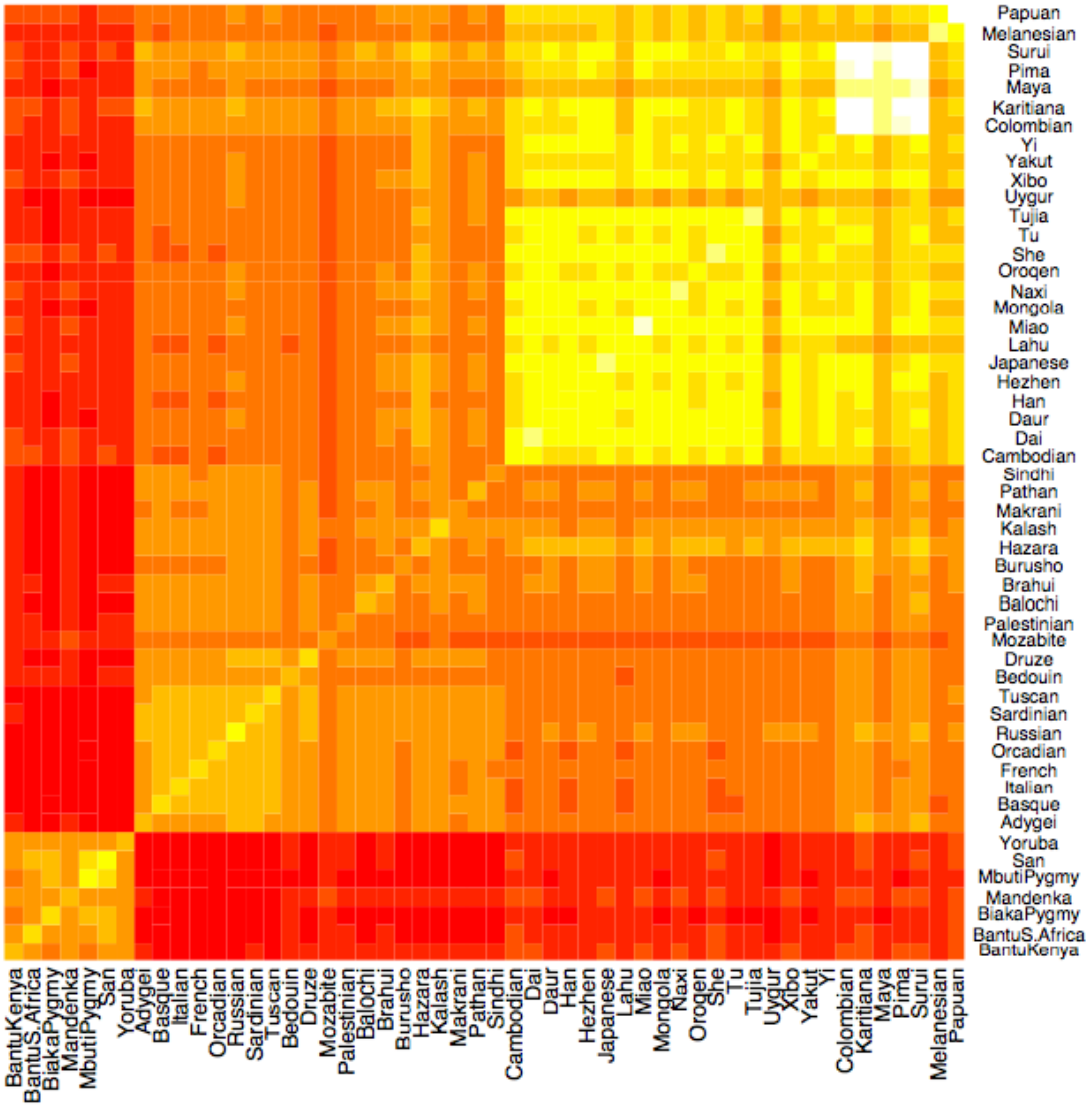
Outline

- Quantifying distances between populations
- Building trees
- Admixture Graphs
- Some spatial statistics





Quantifying distances among populations



Quantifying distances among populations

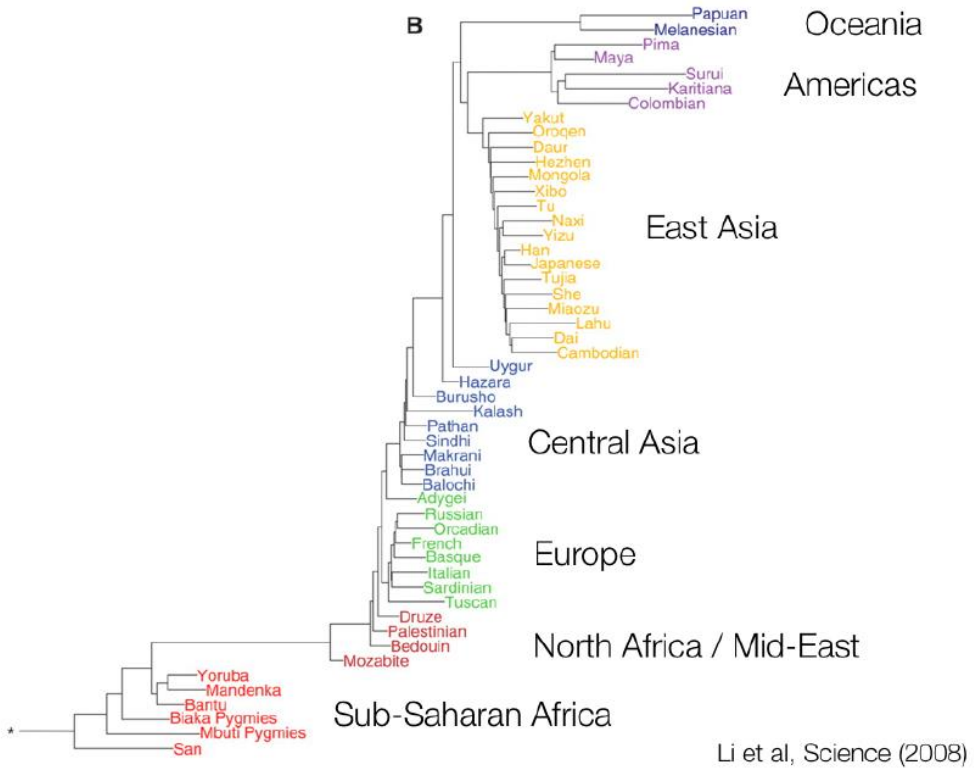
$F_2(A,B) = E[(p_A - p_B)^2]$

$$\hat{F}_2(P_1, P_2) = \pi_{12} - \frac{\pi_{11} + \pi_{22}}{2}.$$

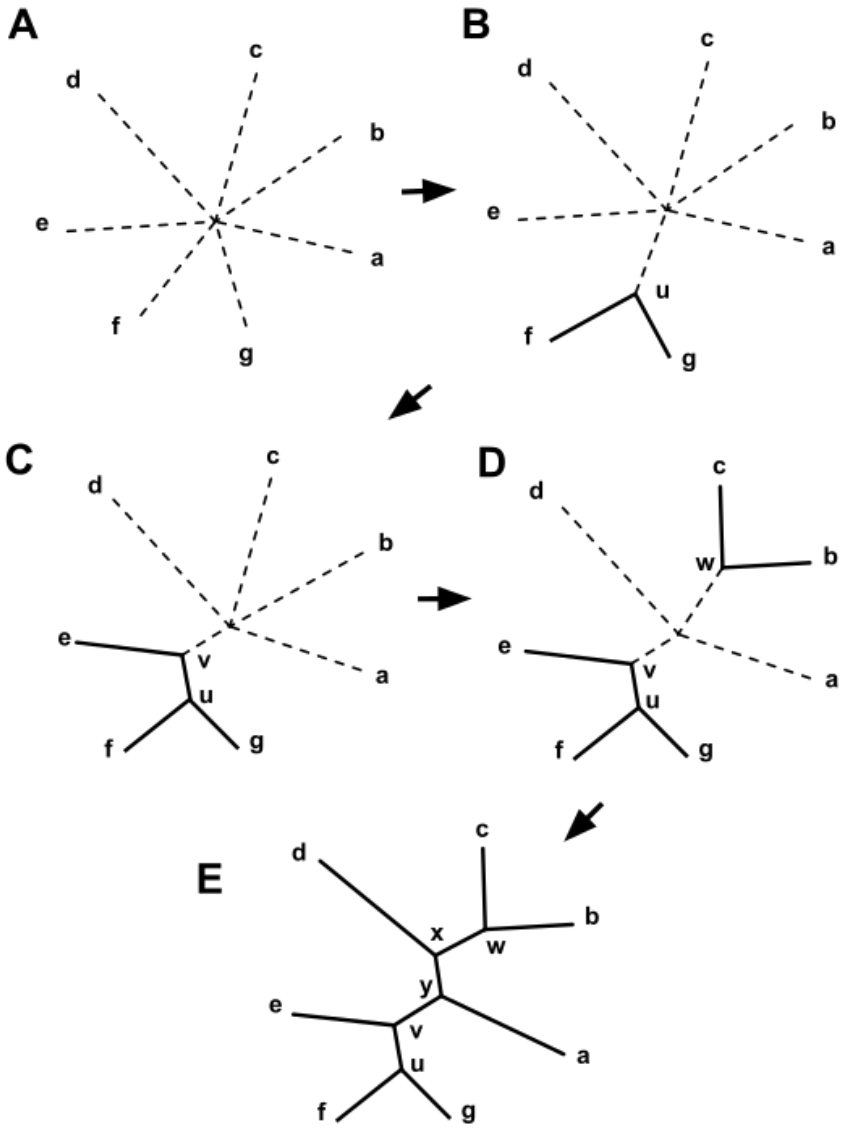
$$F_2(P_1, P_2) = \frac{1}{2} F_{\mathrm{ST}} \mathbb{E} H_{\mathrm{exp}}.$$

$$F_2(P_1, P_2) = 2\mathbb{E}T_{12} - \mathbb{E}T_{11} - \mathbb{E}T_{12}$$

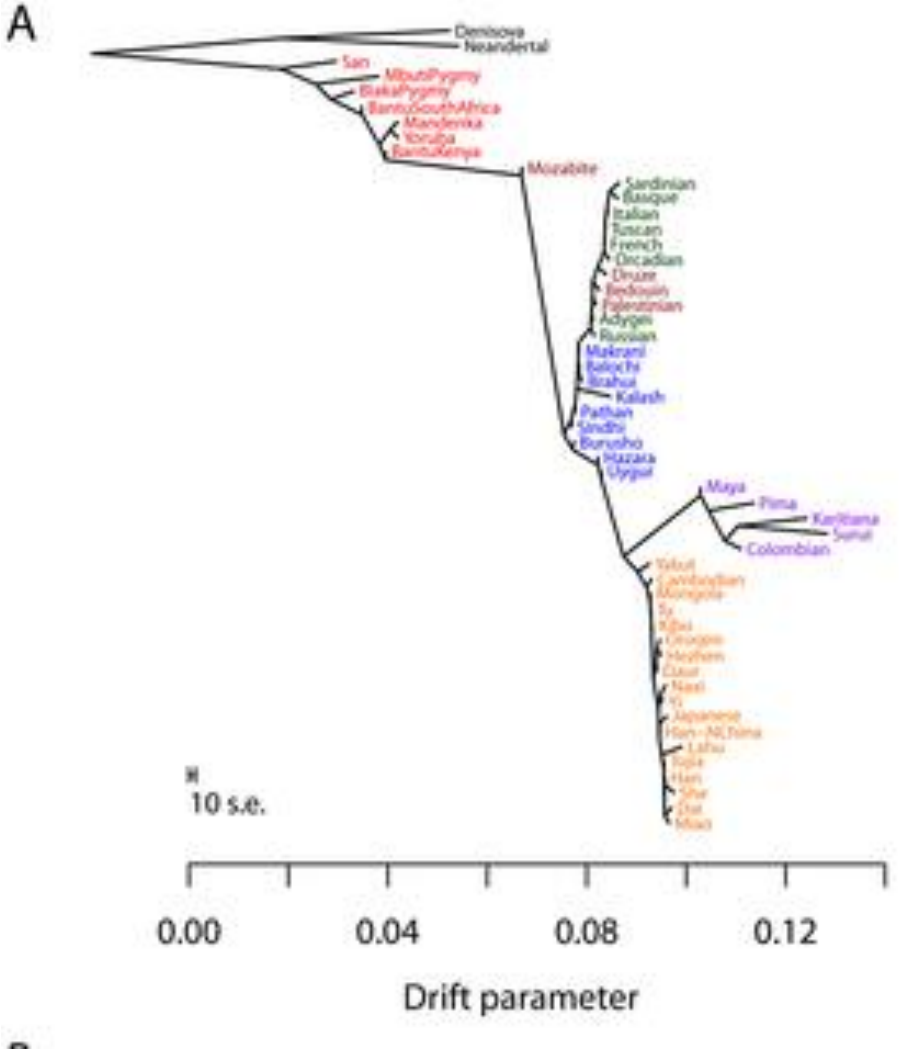
Quantifying distances among populations

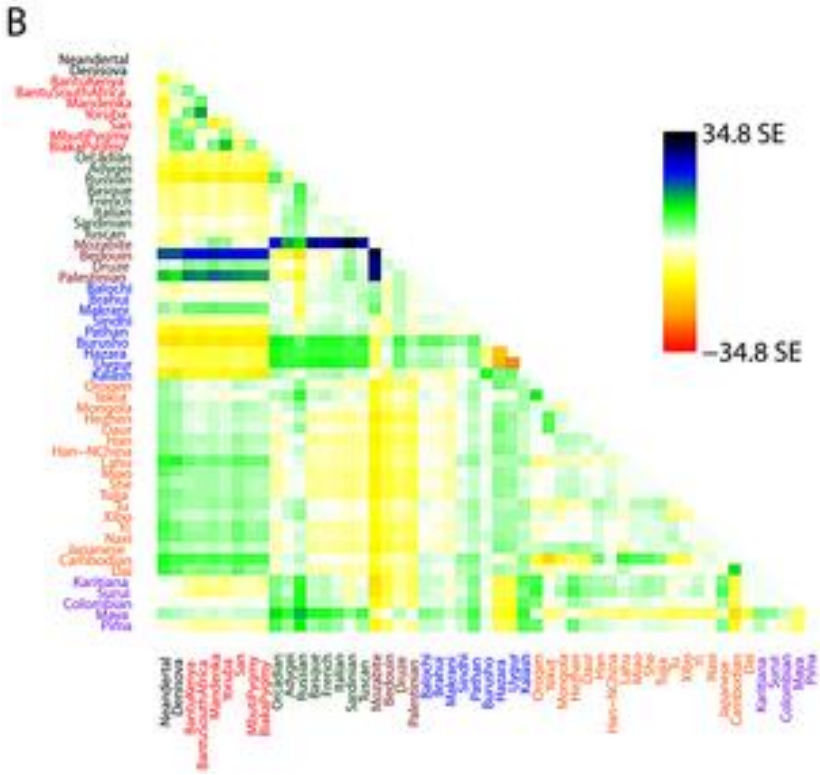


Neighbour Joining trees

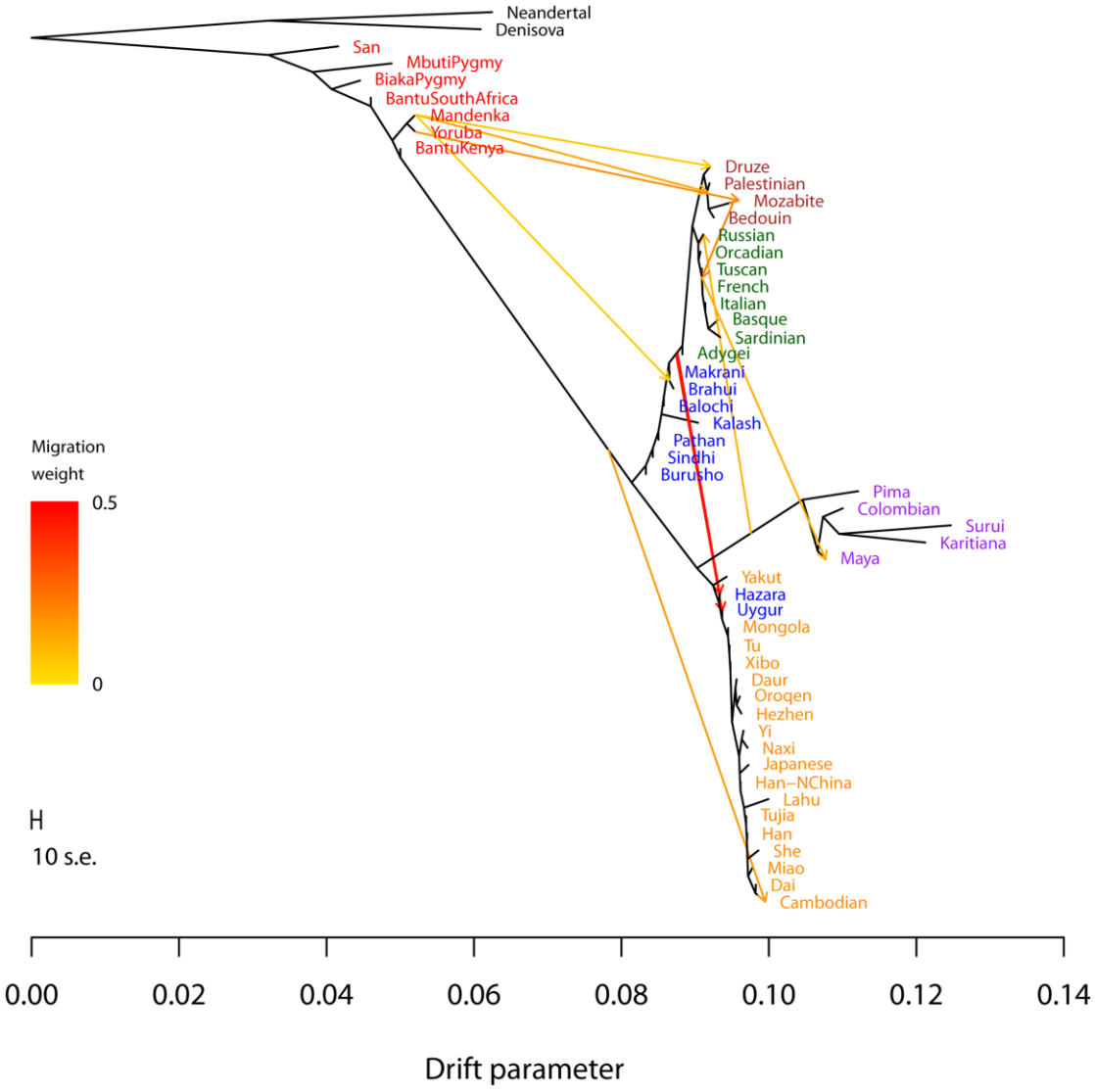


Breaking the tree

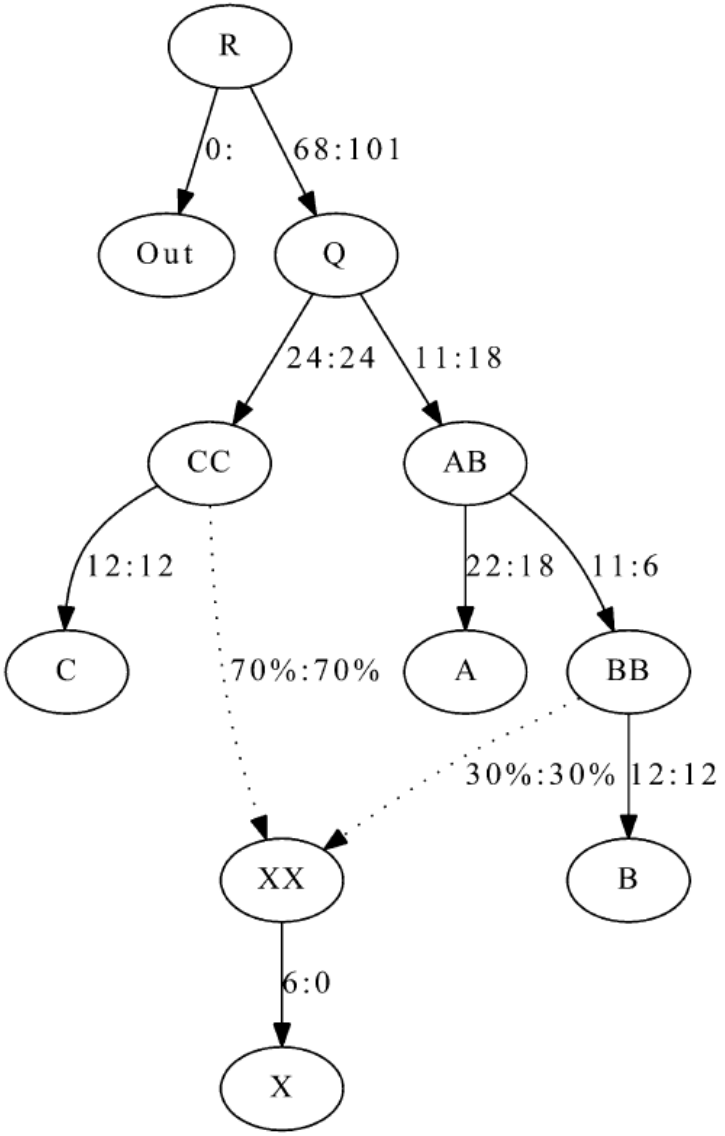




Treemix

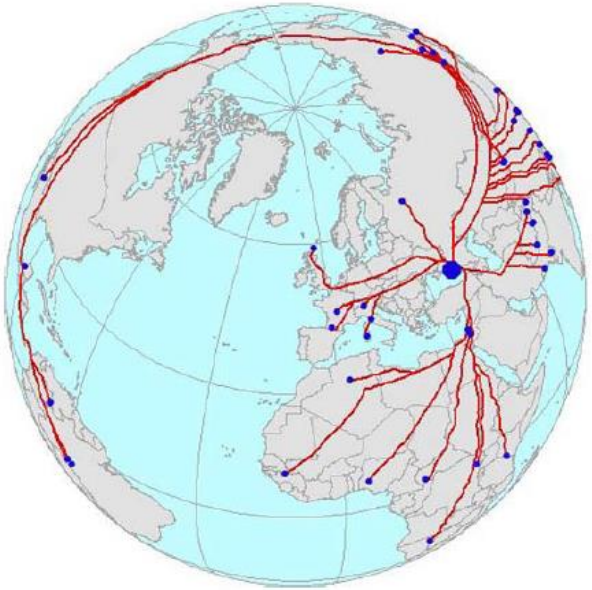

Neandertal
Denisova
San
MbutiPygmy
BiakaPygmy
BantuSouthAfrica
Mandenka
Yoruba
BantuKenya
Druze
Palestinian
Mozabite
Bedouin
Russian
Orcadian
Tuscan
French
Italian
Basque
Sardinian
Adygei
Makrani
Brahui
Balochi
Kalash
Pathan
Sindhi
Burusho
Pima
Colombian
Surui
Karitiana
Maya
Yakut
Hazara
Uygur
Mongola
Tu
Xibo
Daur
Oroqen
Hezhen
Yi
Naxi
Japanese
Han−NChina
Lahu
Tujia
Han
She
Miao
Dai
Cambodian
Migration
weight
0.5
0
H
10 s.e.
0.00
0.02
0.04
0.06
0.08
0.10
0.12
0.14
Drift parameter

qpGraph

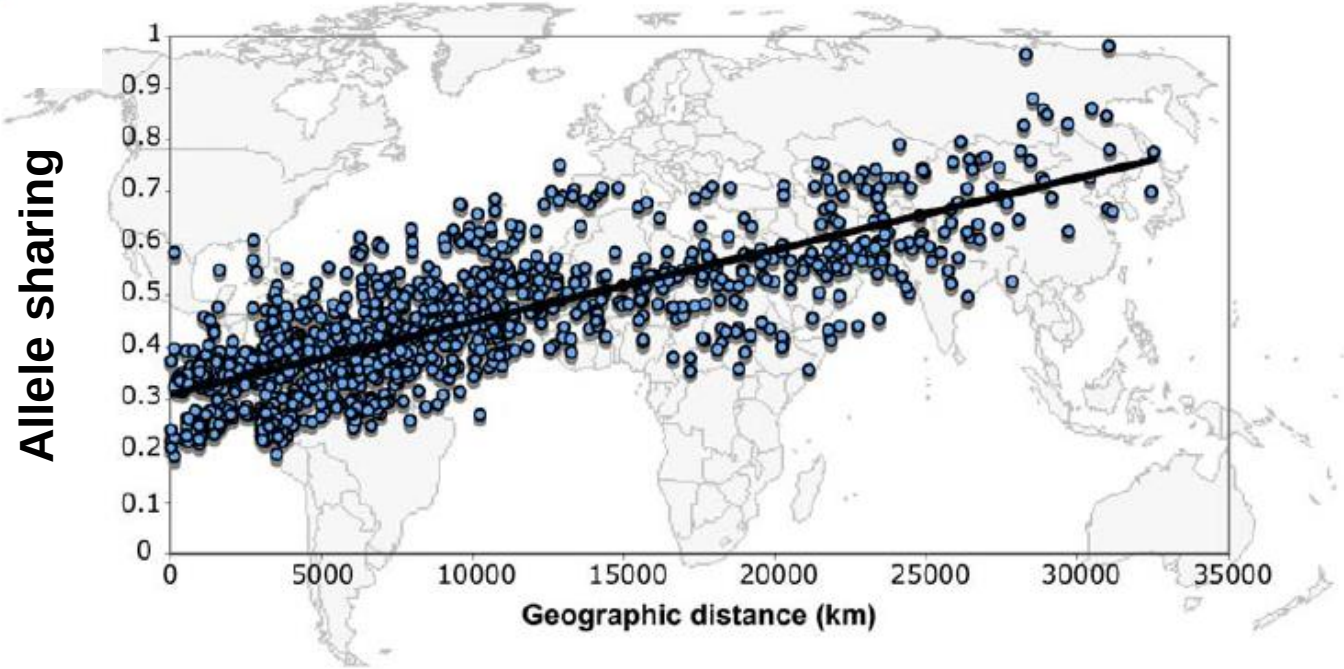


- Fitted on F_2
- First fit unadmixed skeleton
- Then test admixture scenarios
- Check for mismatches in predicted vs observed F_3 and F_4
- Not exhaustive, multiple graphs might fit data equally

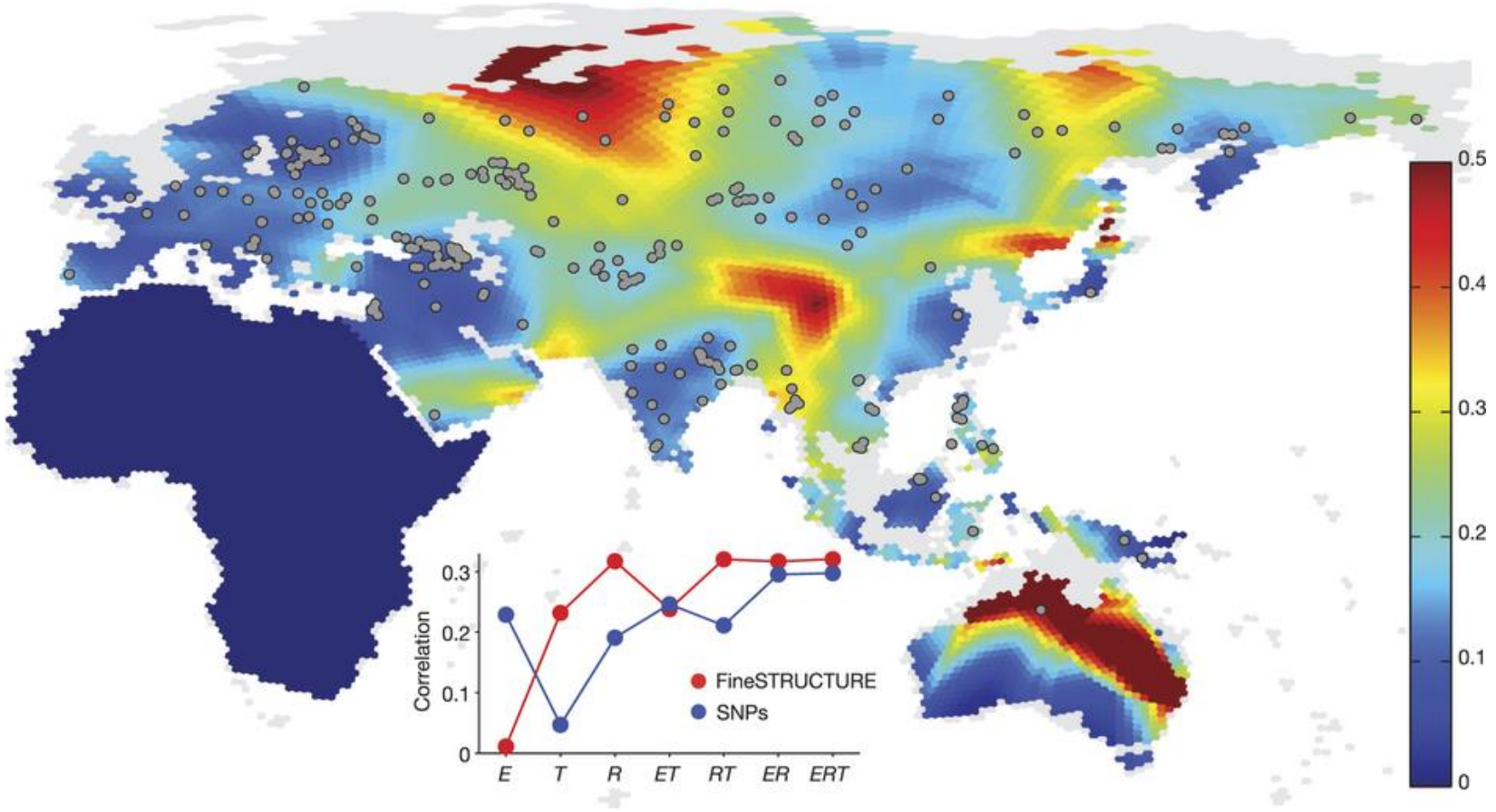
Using space to model many populations



Isolation by distance can explain a lot of differences

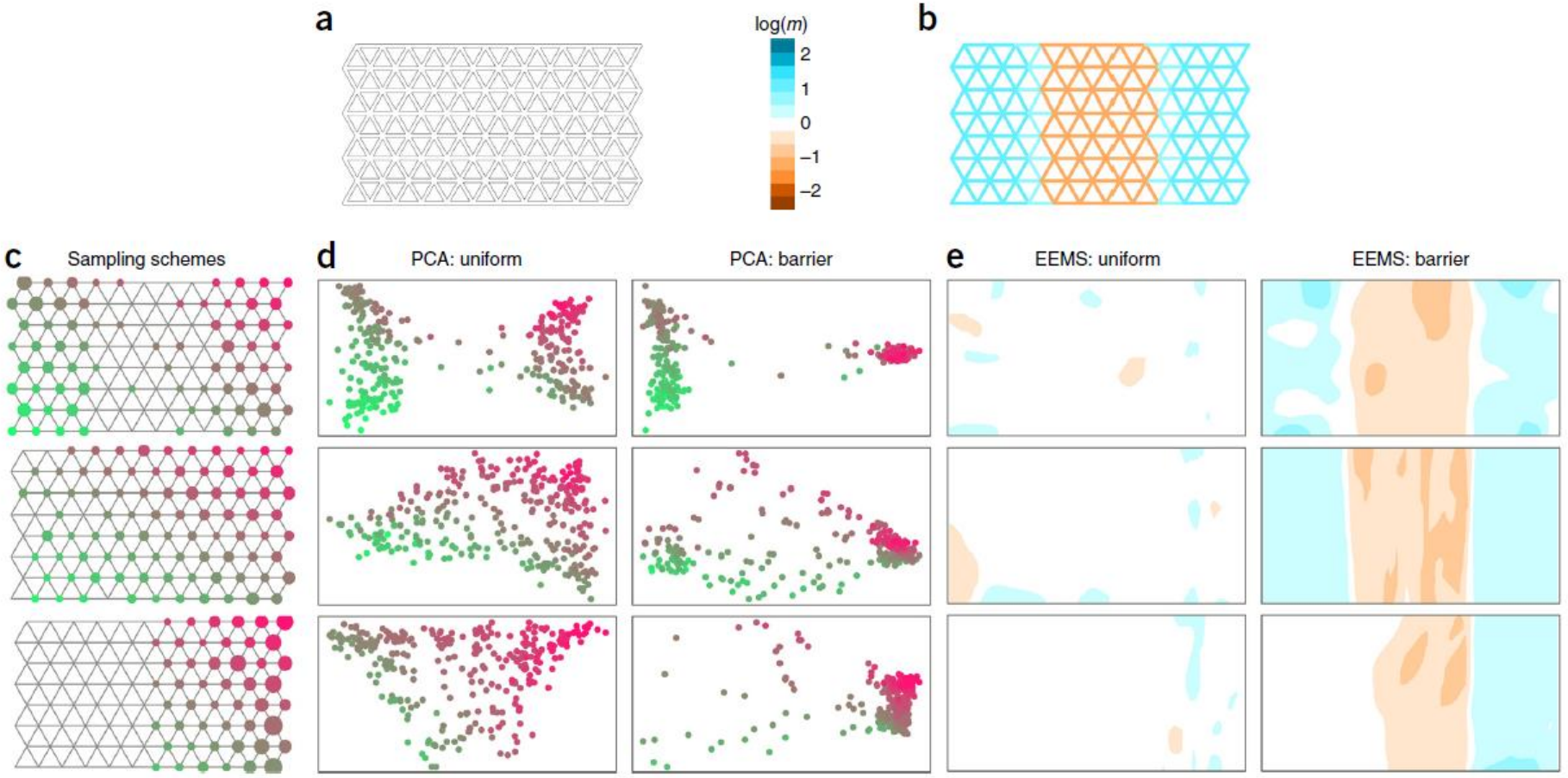


Barriers for human movement

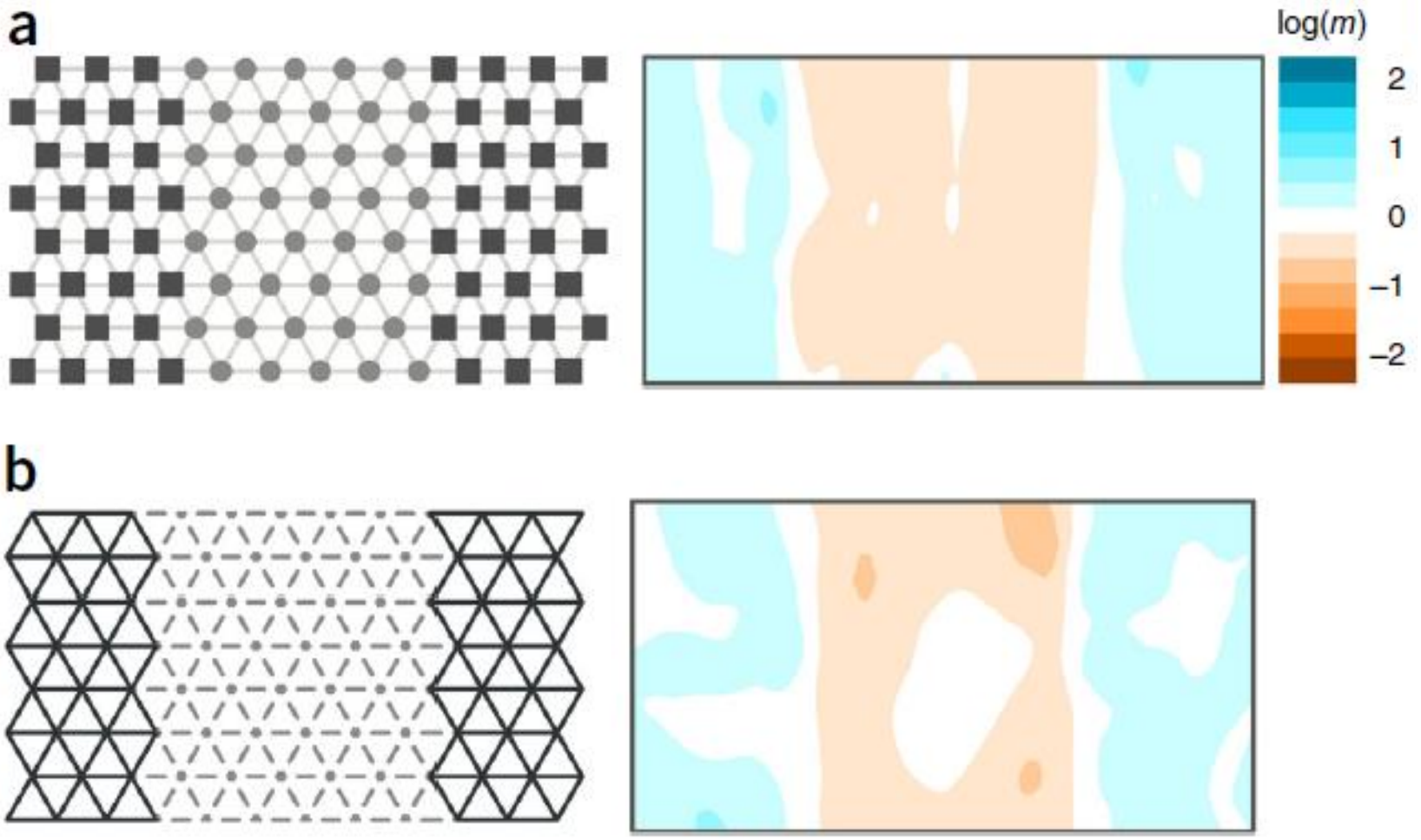


Pagani et al. 2016 Nature

EMMS

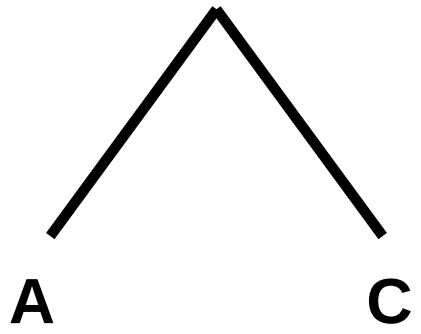

a
log(m)
2
1
0
−1
−2
b
c
Sampling schemes
d
PCA: uniform
PCA: barrier
e
EEMS: uniform
EEMS: barrier

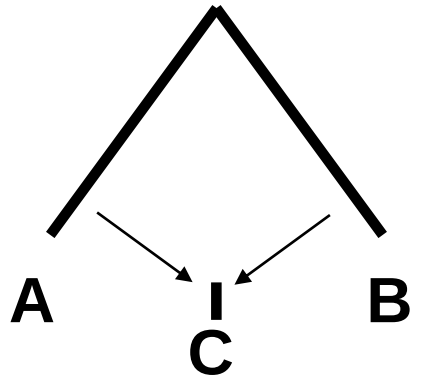
EMMS – migration vs population size



Summary

- Quantifying distances between populations
- Building trees
- Admixture Graphs
- Spatial statistics





Practical

- Use Admixtools to fit admixture graph
- Human dataset with modern and ancient

qpGraph input file

```
root    R

label   Name   NameInDataset
label   Mbuti  Mbuti

edge    Name          Source Destination
edge    R_Mbuti       R      Mbuti

Admix  Target  Source1       Source2       Prop1  Prop2
admix  pLBK    X1      BE    80      20
```