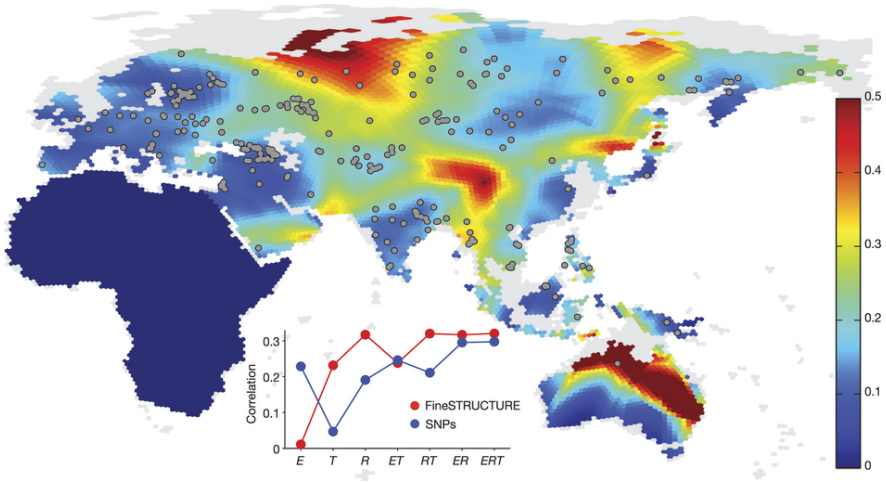
Quantifying relationships among populations: Many populations

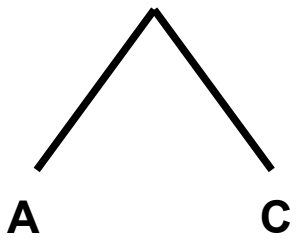


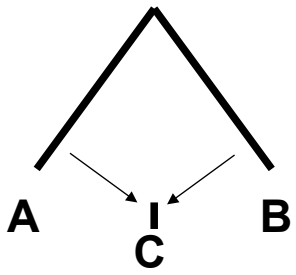
Andrea Manica

UNIVERSITY OF CAMBRIDGE

Outline

- Generalising f_4: # of waves
- Quantifying distances between populations
- Building trees
- Admixture Graphs
- Some spatial statistics





qpWave: generalising f_4

0 1 2 0 1 2

L: pop_a, pop_b, pop_c R: pop_d, pop_e, pop_f

$X(l_i,r_i) = f_4(l_0,l_i; r_0, r_i)$

rank of a X is # of linearly indep. rows (or cols)

rank + 1 ≤ n waves

f_4(a,b;d,e)	f_4(a,b;d,f)
f_4(a,c;d,e)	f_4(a,c;d,f)

qpAdm: estimating proportions from multiple sources

T: pop_t

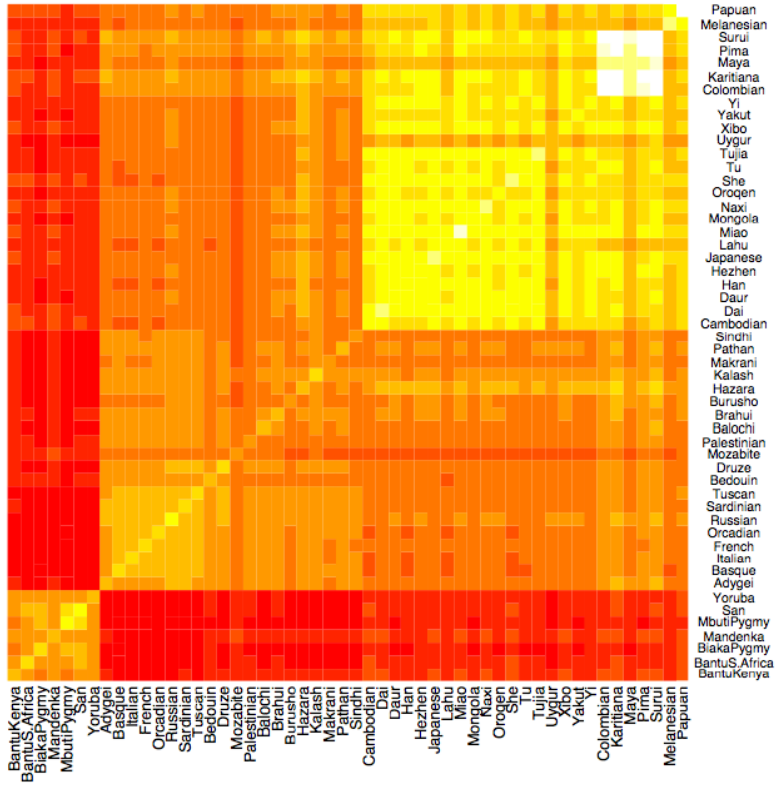
S: pop_a, pop_b, pop_c R: pop_d, pop_e, pop_f

$$T = \sum_{i=1}^{n} w_i s_i$$

$$\sum_i w_i f_4(T, s_i, r_1, r_2) = f_4(T, T, r_1, r_2)$$

$$= 0$$

Quantifying distances among populations



Quantifying distances among populations

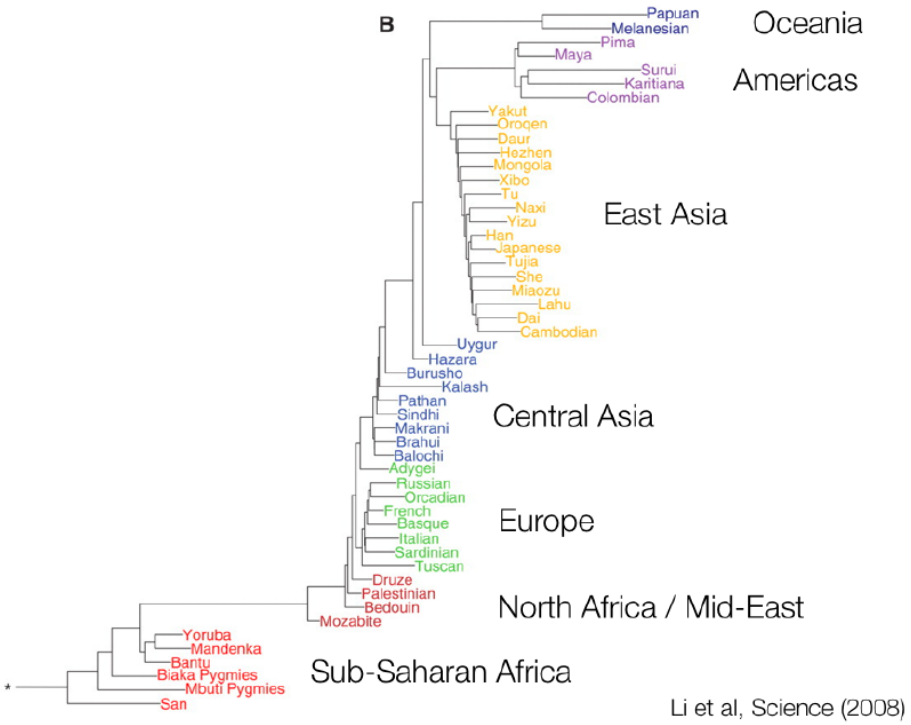
$$f_2(A,B) = E[(p_A - p_B)^2]$$

$$f_2(A,B) = \pi_{AB} - \tfrac{1}{2}(\pi_{AA} - \pi_{BB})$$

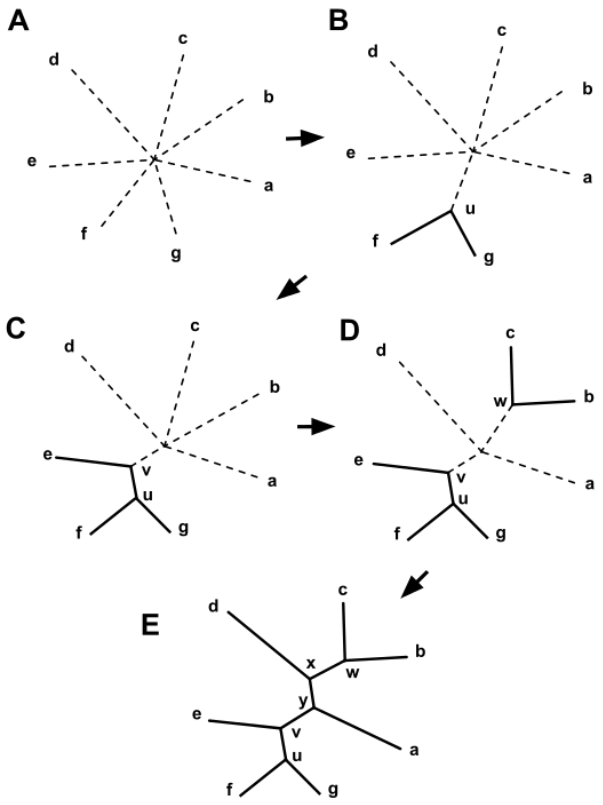
$$f_2(A,B) = \tfrac{1}{2} F_{ST} E H_{exp}$$

$$f_2(A,B) = 2E T_{AB} - E T_{AA} - E T_{BB}$$

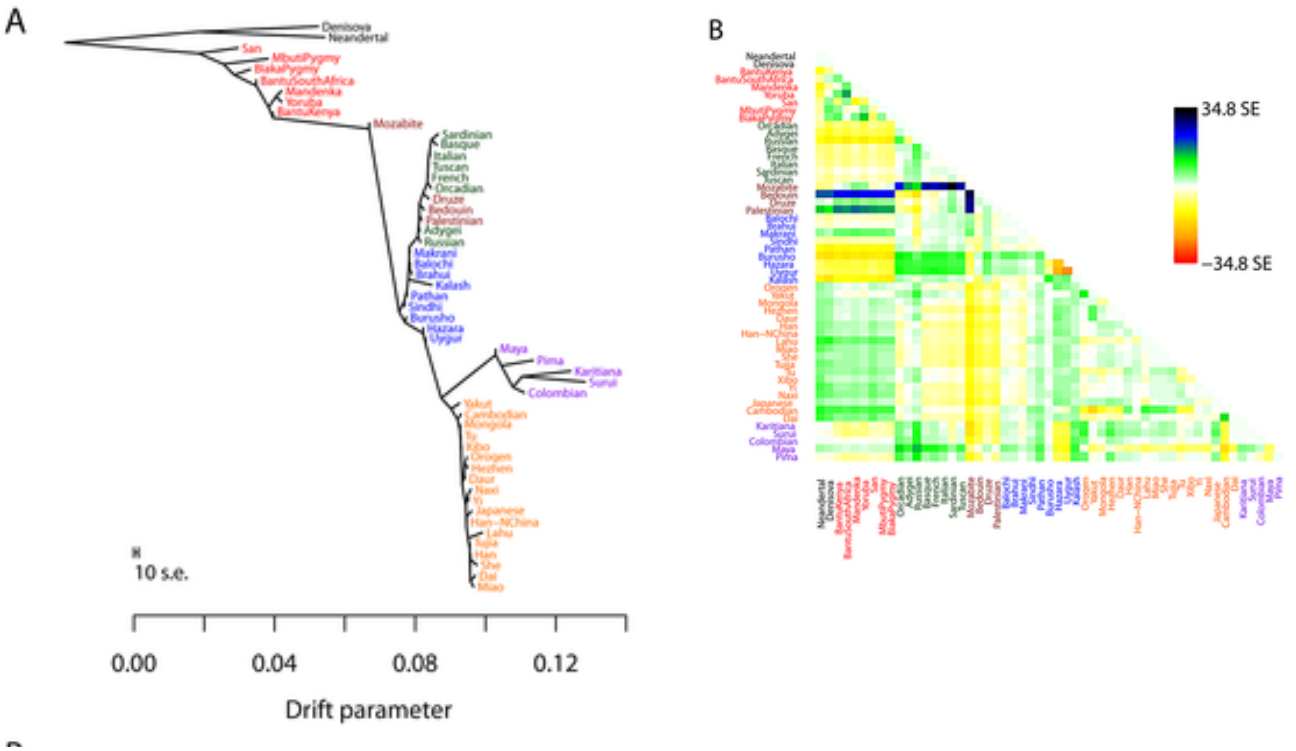
Quantifying distances among populations



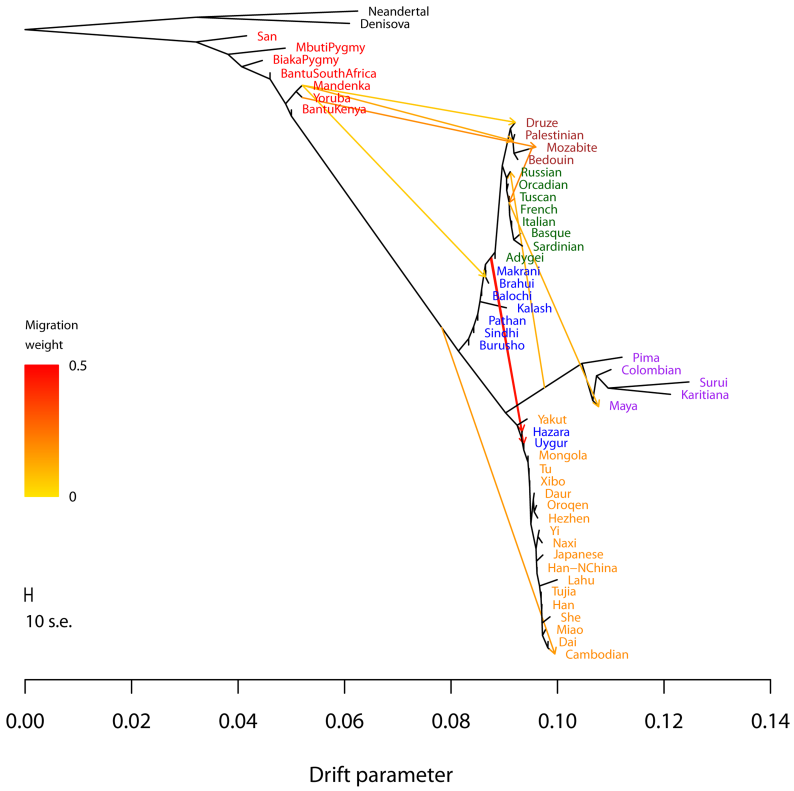
Neighbour Joining trees



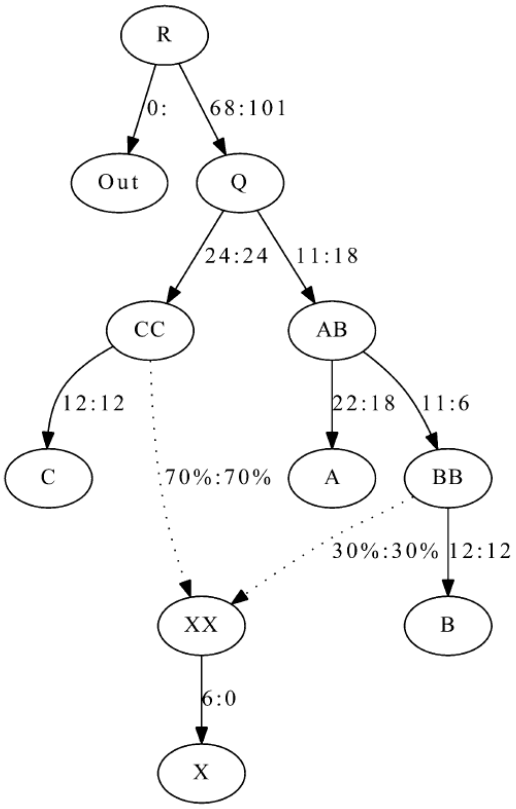
Breaking the tree



Treemix



qpGraph



Fitted on f_2 or f_3

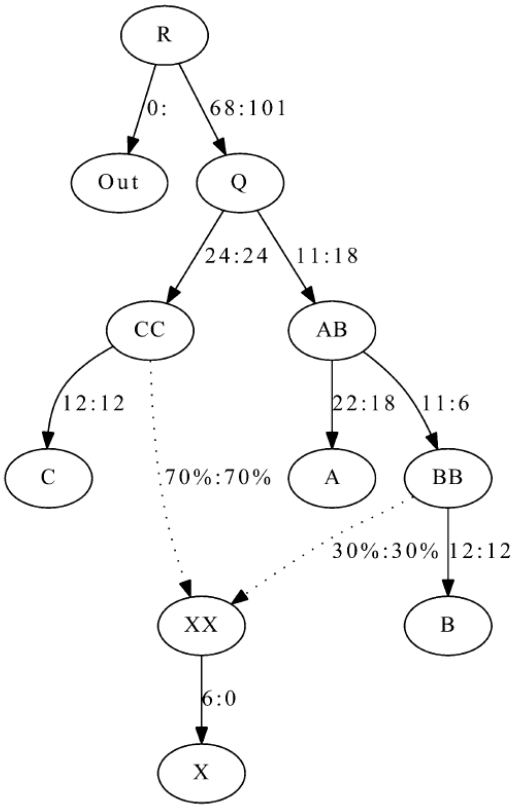
First fit unadmixed skeleton

Then test admixture scenarios

Check for mismatches in predicted vs observed f_3 and f_4

Not exhaustive, multiple graphs might fit data equally

qpGraph – changing philosophies

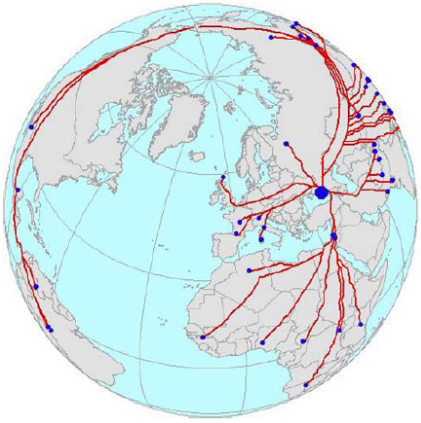


Even for simple scenarios, there are many graph that fit the data

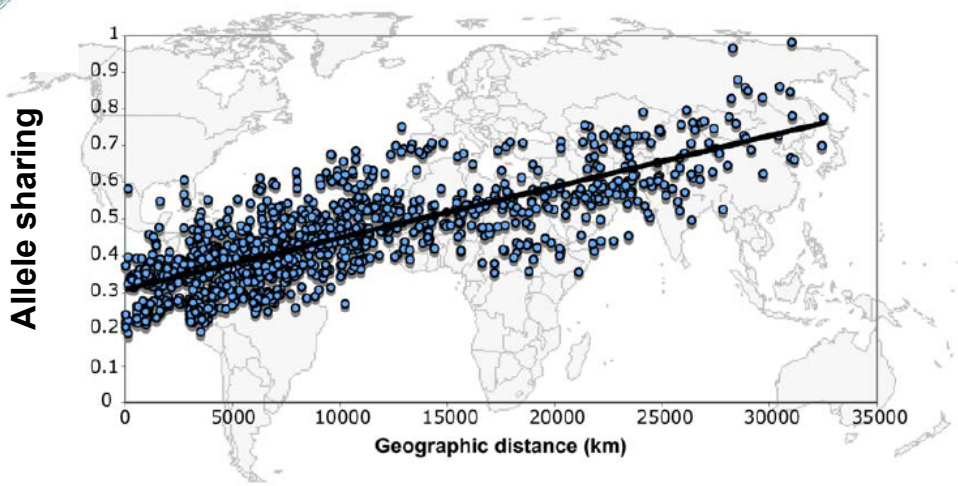
Manual searches are not enough

Extensive searches are needed, and evidence from admixture graphs needs to be complemented with other approaches

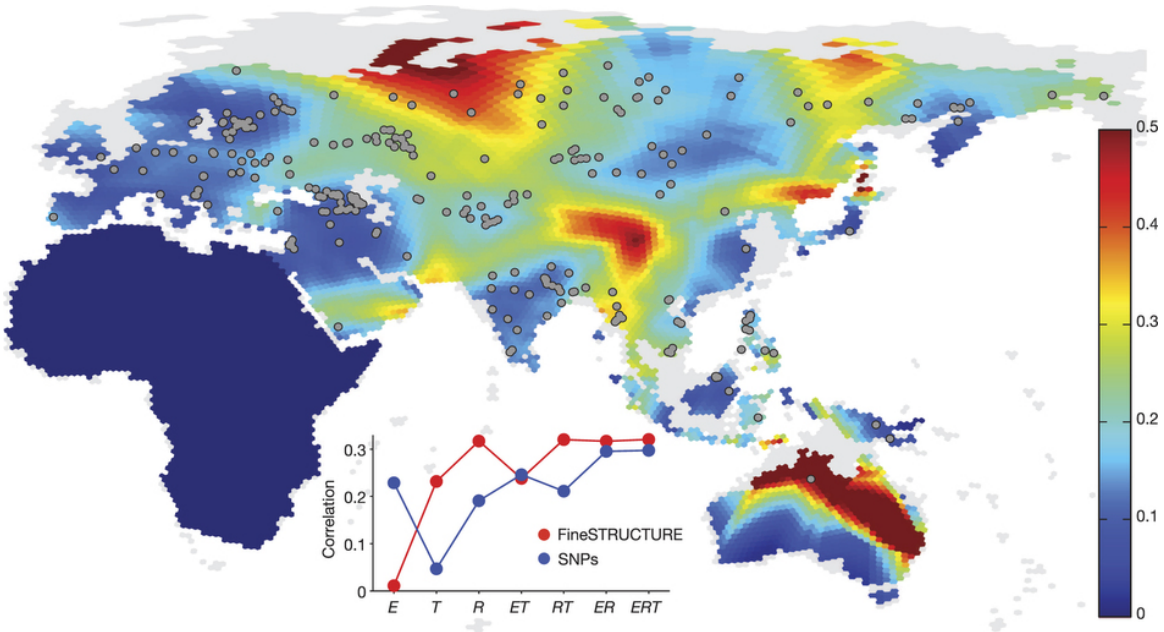
Using space to model many populations



Isolation by distance can explain a lot of differences

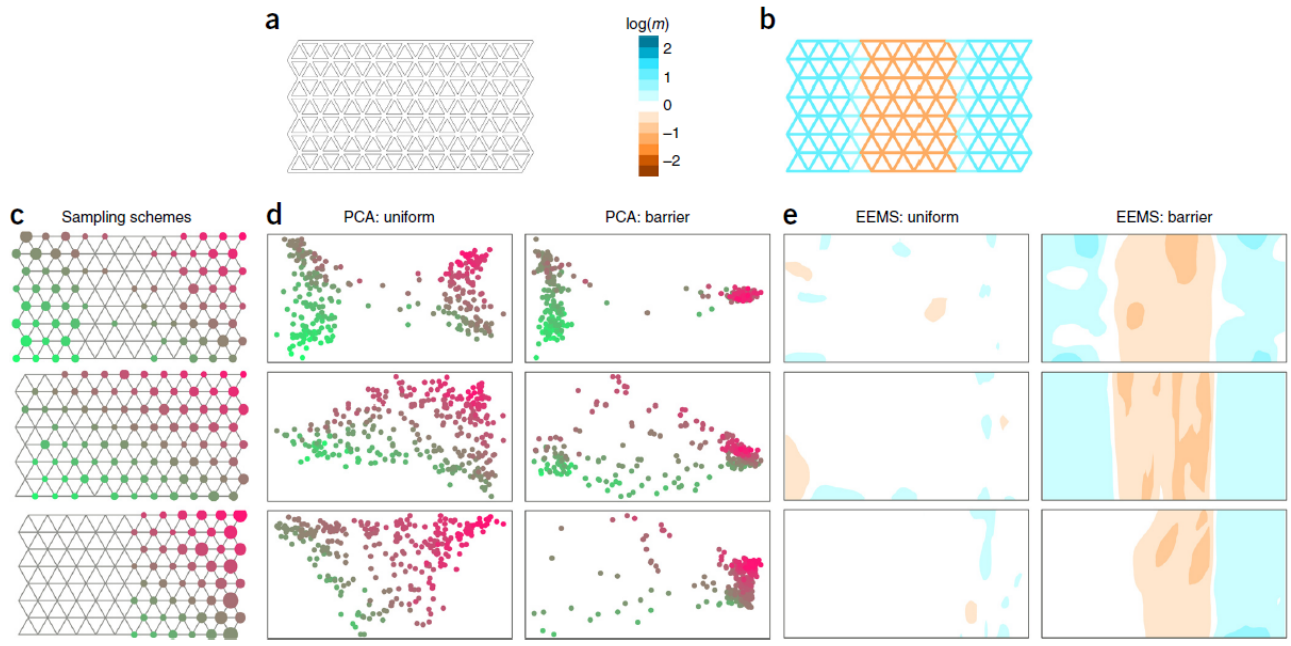


Barriers for human movement



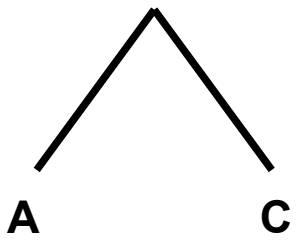
Pagani et al. 2016 Nature

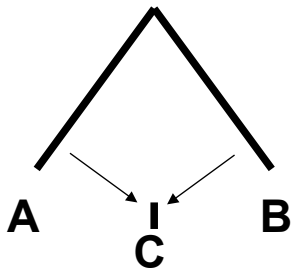
EMMS



Outline

- Generalising f_4
- Quantifying distances between populations
- Building trees
- Admixture Graphs
- Some spatial statistics





Practical

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