Interbreeding and demic diffusion in the spread of the Neolithic in Europe

Joaquim Fort
Universitat de Girona (Catalonia, Spain)

World Neolithic Congress
Sanliurfa (Turkey), Nov. 5th 2024

ICREA Academia Grant 2021
MINISTERIO DE CIENCIA E INNOVACIÓN
AGENCIA ESTATAL DE INVESTIGACIÓN
Grant PID2023-150978NB-C22
Agència de Gestió d'Ajuts Universitaris i de Recerca
AGAUR
Grant 2021-SGR-00190

Models of Neolithic spread

- Demic diffusion = Farming populations spread = dispersal + net reproduction.

- Cultural diffusion = spread of ideas = transmission of plants, animals and knowledge from farmers to hunter-gatherers = = cultural transmission.

- Demic-cultural models

Demic models

1) Wave of Advance demic model

$$spread\ rate = s = \sqrt{\frac{r\,m}{T}}$$

Pre-industrial farmers:

$$\left.\begin{array}{r}\text{Reproduction:}\, r = 0.032\ \mathrm{yr}^{-1}\\ \text{Mobility:}\, m = 1544\ \mathrm{km}^2\\ \text{Generation time:}\, T = 25\ yr\end{array}\right\} \rightarrow s = 1.4\ \mathrm{km/yr}$$

Ammerman &
Cavalli-Sforza
(1973)

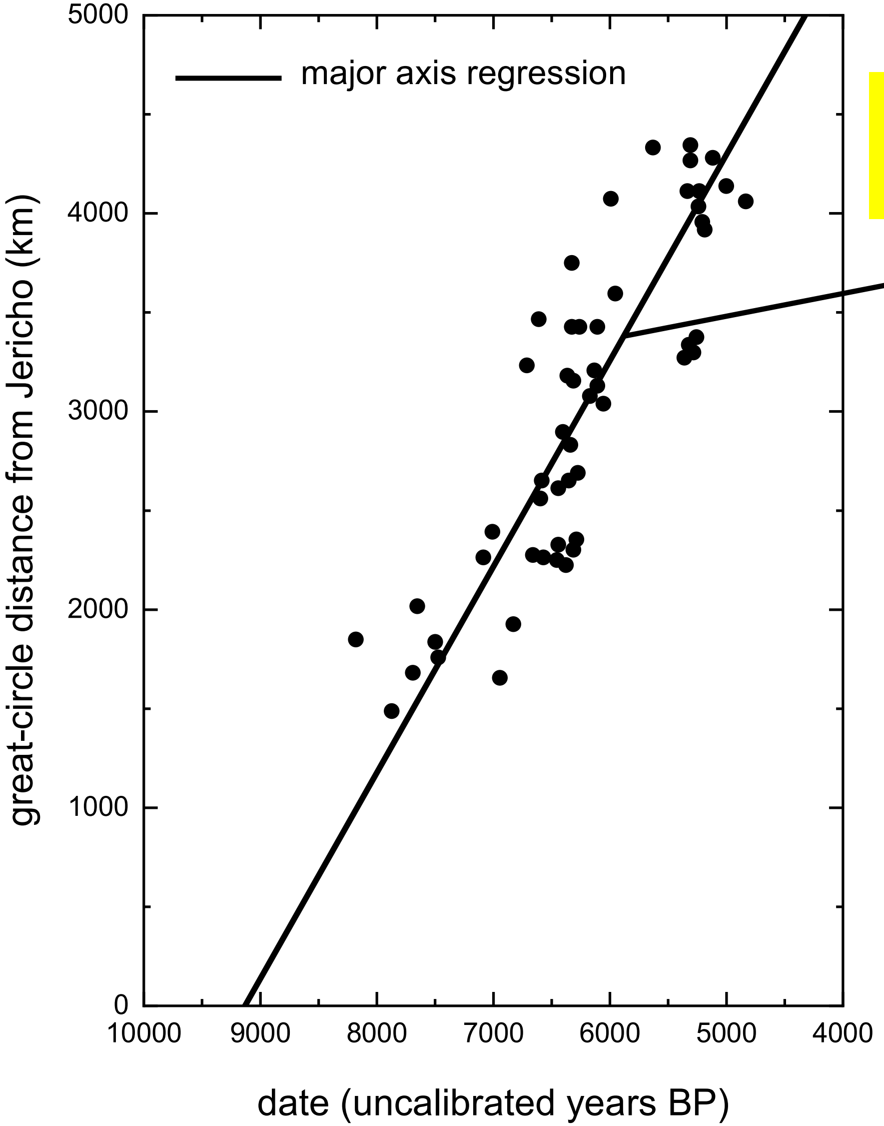
2) Time-delayed demic model

$$spread\ rate = \frac{s}{1+\frac{rT}{2}} = 1.0\ \text{km/yr}$$

Fort &
Méndez
(1999)

Archaeological data

Ammerman & Cavalli-Sforza (1971, 1984)



spread rate = 1 km/yr

53 sites in Europe

r = 0.89 (Jericho, highest-r origin)

Why did Ammerman & Cavalli-Sforza suggest a mainly demic process?

Because demic models agree
with the archaeological data:

Demic model ~ 1 km/yr

Archaeological data* ~ 1 km/yr

*Other databases also yield about 1 km/yr:

·Gkiasta, Russell, Shennan & Steele (2003): **510 sites**

·Pinhasi, Fort & Ammerman (2005): **735 sites**

·Baggaley, Sarson, Shukurov *et al.* (2012): **302 sites**

·Bocquet-Appel et al. (2012): **940 sites**

·Fort, Pujol & vander Linden (2012): **919 sites**

·etc.

Today there is no doubt that Ammerman & Cavalli-Sforza were right

Ancient genome-wide studies have shown that >90% of the genomic ancestry of early European farmers is due to a population source in Anatolia (Mathieson *et al*., Nature 2015).

This confirms that it was mainly demic.

But such studies have not answered some key questions:

1. Mean distance moved by early farmers?
2. Did it depend on the route (inland/coast)?
3. % of early farmers that interbred with HGs?
4. Did it depend on the route (inland/coast)?

In this talk we will deal with these questions.

Demic-cultural model

Initially:

· Farmers only in a region (e.g., corresponding to PPNB/C sites).

· The rest of the surface empty of farmers and has HGs at their saturation density.

3 steps or processes take place every generation (1 generation =32±3 yr, from ethnography):

(1) Reproduction: logistic, with growth rate of farmers r from ethnographic data: $r = (0.028\ \pm 0.005)\mathrm{yr}^{-1}$.

(2) Cultural transmission:

P_F = number of farmers (in the cell considered)

P_{HG} = number of hunter-gatherers (“ “ “ “)

Cultural transmission theory [1-3]:

$$P_F(t+1) = P_F(t) + f\frac{P_F(t)P_{HG}(t)}{P_F(t)+\gamma P_{HG}(t)} \approx P_F(t) + \textcircled{C}\, P_F(t)$$

if $P_F \ll P_{HG}$

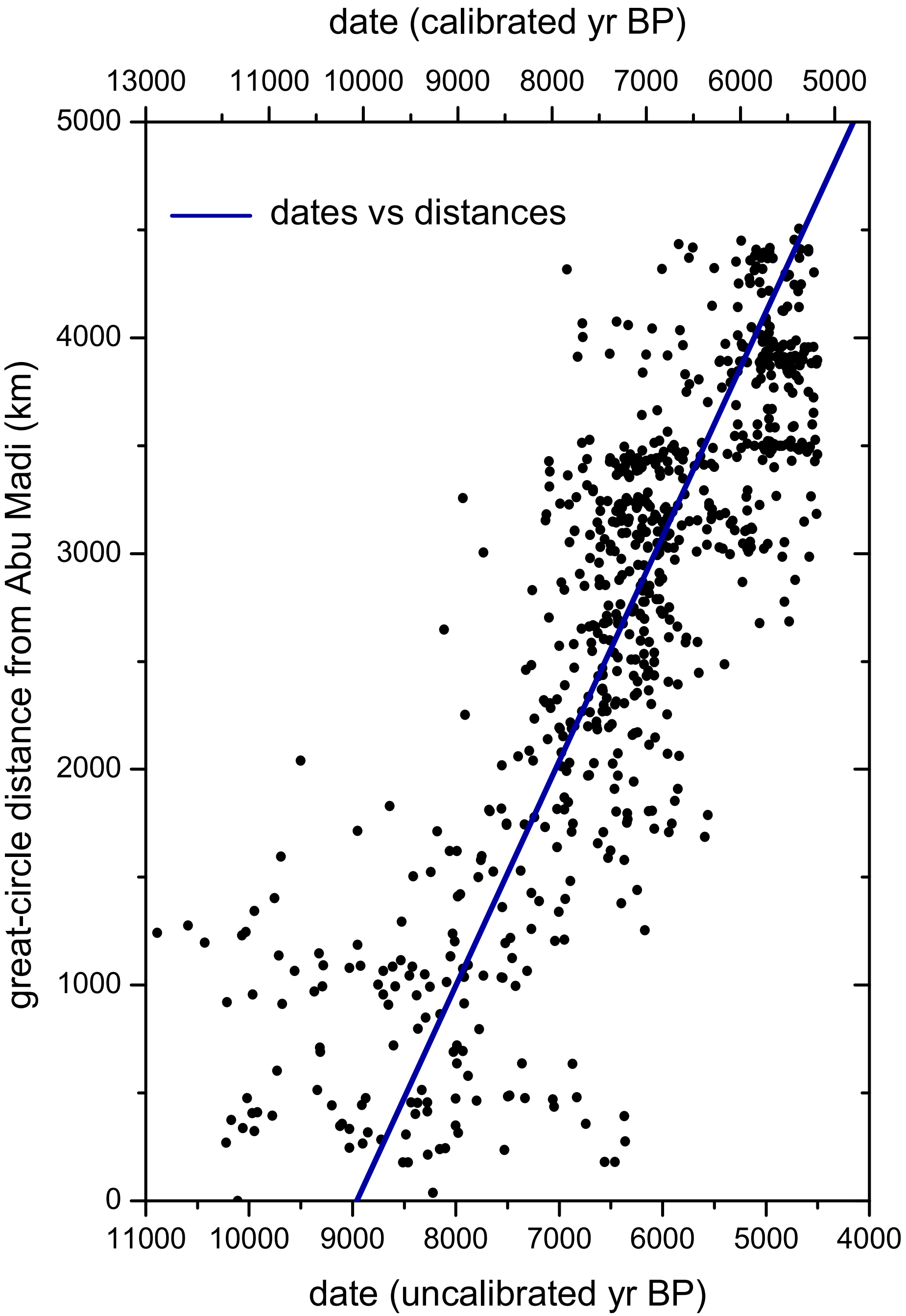
$$P_{HG}(t+1) = P_{HG}(t) - f\frac{P_F(t)P_{HG}(t)}{P_F(t)+\gamma P_{HG}(t)} \approx P_F(t) - \textcircled{C}\, P_F(t)$$

$$C = \frac{f}{\gamma} = \text{intensity of cultural transmission}$$

[1] Cavalli-Sforza & Feldman, *Cultural transmission & evolution*, Princeton 1981

[2] Fort, *Phys. Rev. E* 2011 [3] Fort, *PNAS* 2012

(3) Dispersal: distances & probabilities from ethnography, e.g.: {2.3,7.3,15,25,35,45,55,100}km, probabilities={42,23,16,8,7,2,1,1}%.



To apply this model we need a range for the spread rate from archaeological data

0.9-1.3 km/yr

great circles & shortest paths

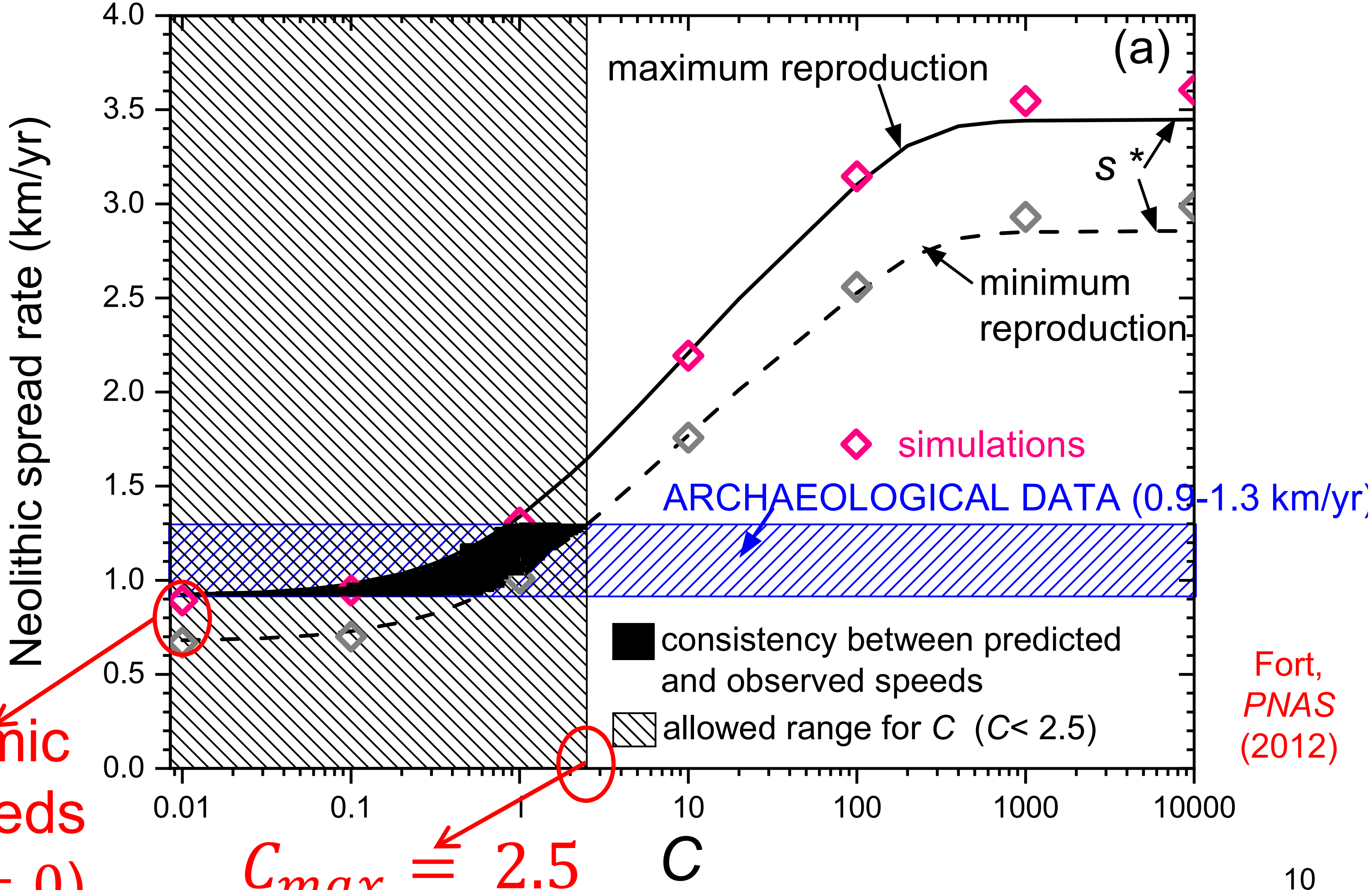
r = 0.83

(for both highest-*r* origins)

735 sites in Europe & Near East

Pinhasi, Fort & Ammerman, *PLoS Biol.* (2005)

Effect of the intensity of cultural transmission *C* on the spread rate



demic speeds $(C = 0)$

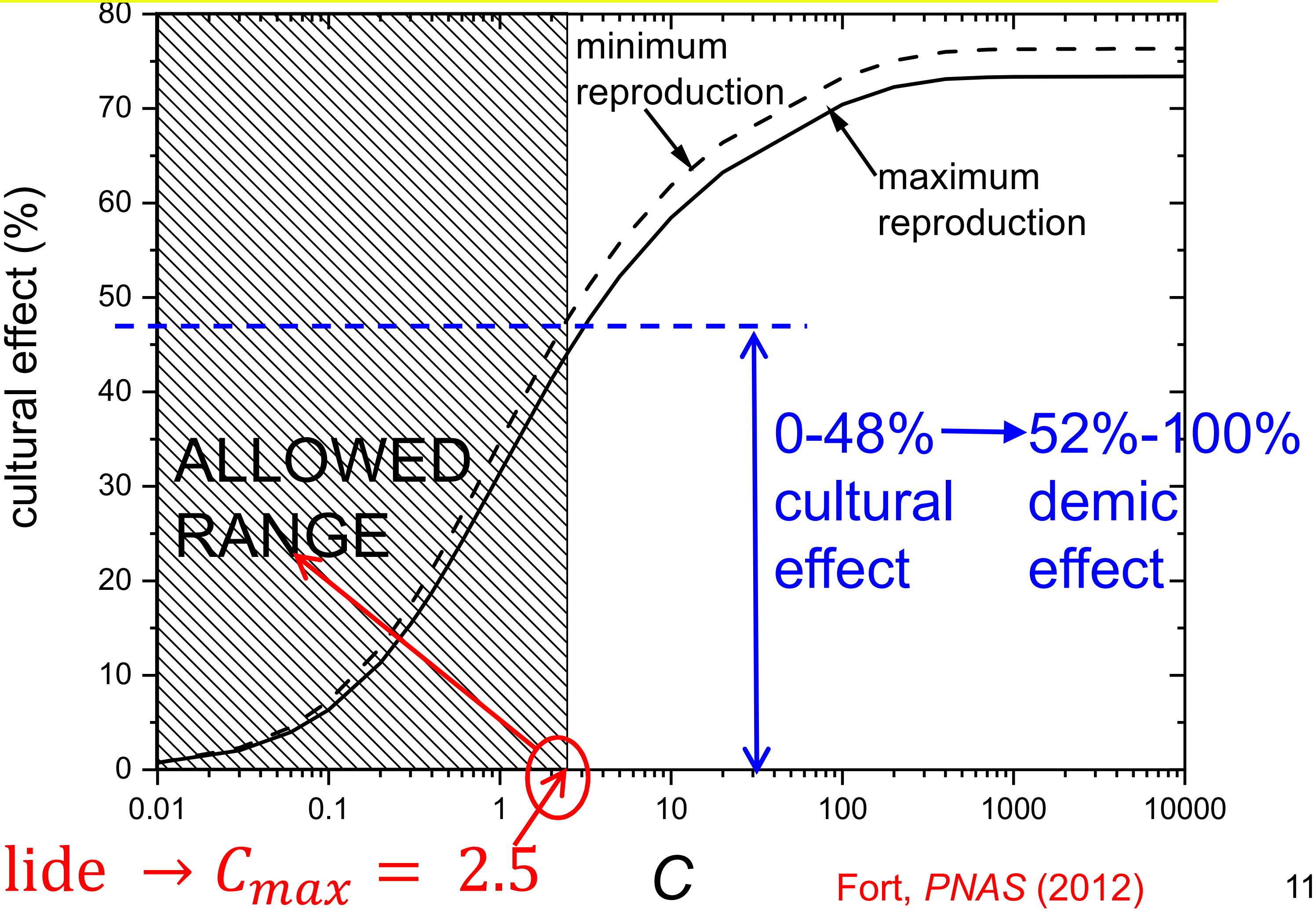
$C_{max} = 2.5$

Fort, *PNAS* (2012)

$$\text{cultural effect (\%)} = \frac{\text{spread rate} - \text{demic spread rate}}{\text{spread rate}} \cdot 100$$

$$\text{demic effect (\%)} = \frac{\text{demic spread rate}}{\text{spread rate}} \cdot 100$$

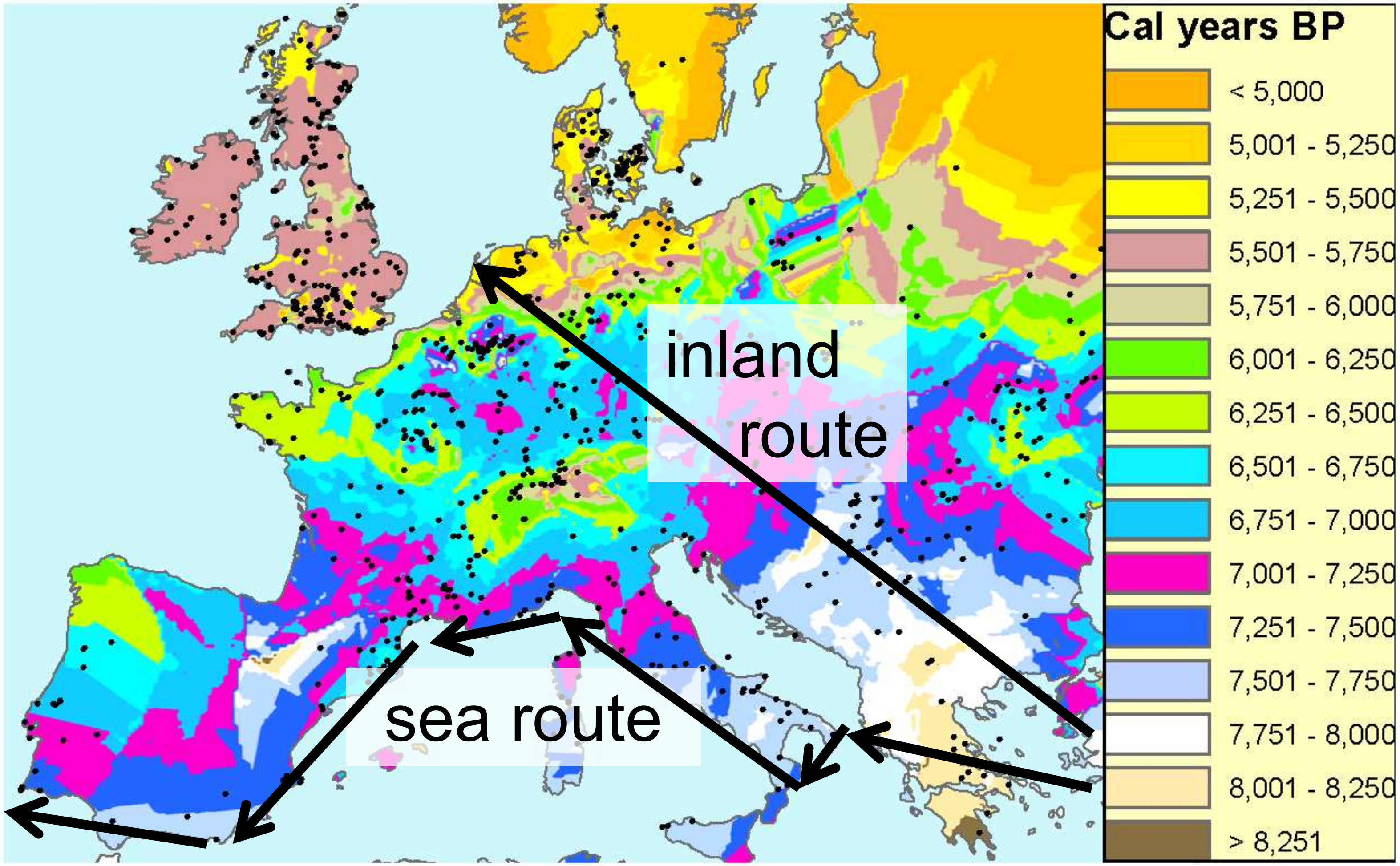
demic effect + cultural effect = 100 %



previuos slide → $C_{max} = 2.5$

Fort, *PNAS* (2012)

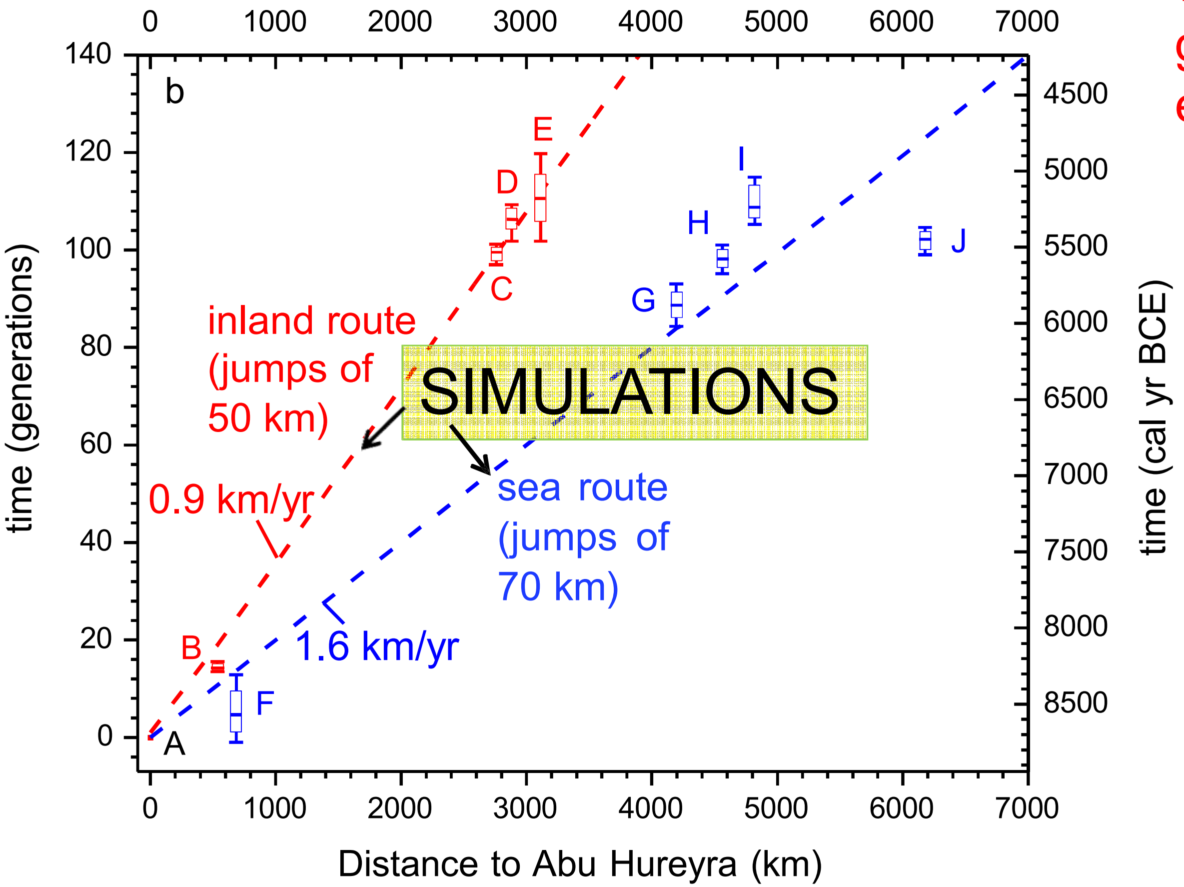
Two routes



Fort, *J. R. Soc. Interface* (2015)

Simulations + archaeological data

Initially there are farmers only at the cell with the **oldest PPNB site in Upper Mesopotamia (Abu Hureyra**, <9,038 cal BC) at a date (8,718 cal BC) such that the simulations agree with the data along the inland route.



Inland route: simulations with jumps of 50 km per generation (value from ethnography)

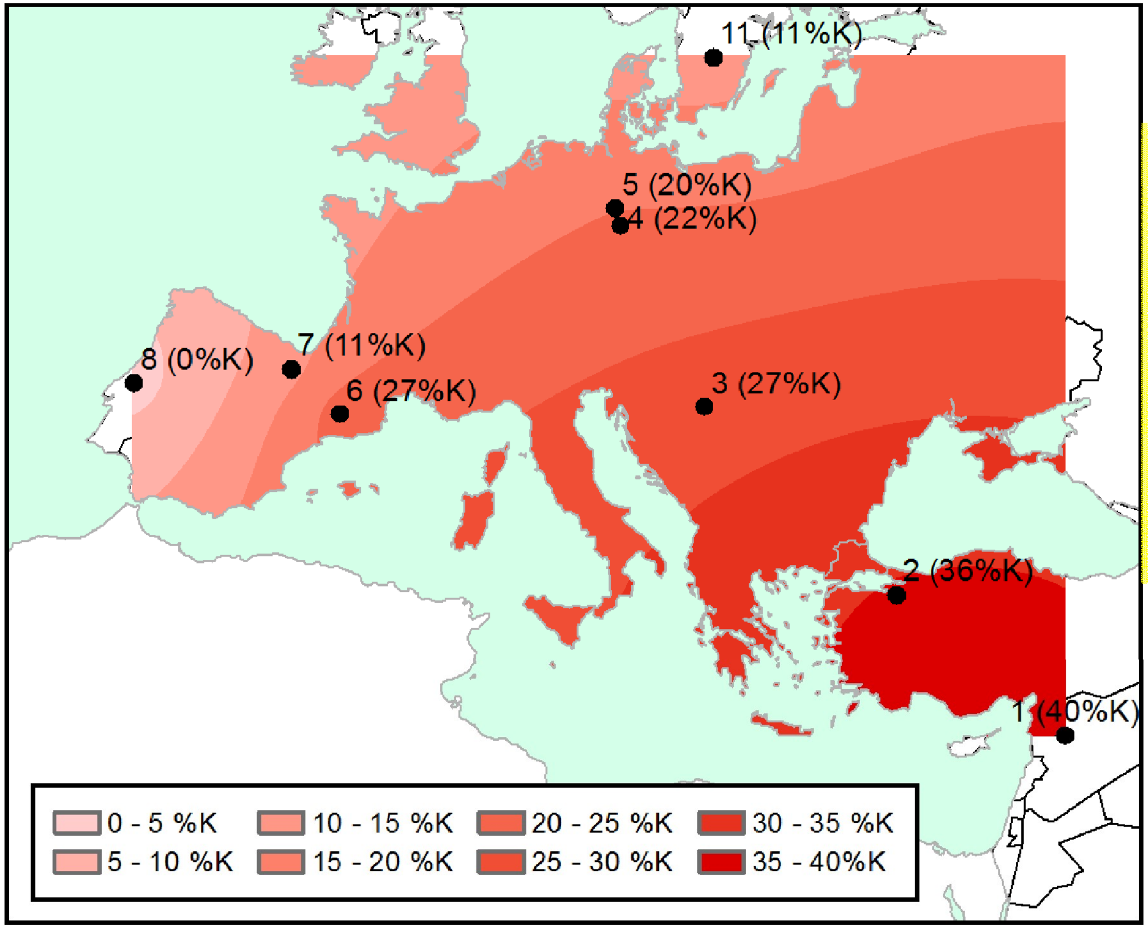
Sea route: best fit for simulations with jumps of 70 km per generation

A northern Mesopotamia
B Anatolia
C Germany
D northern France
E Belgium
F Cyprus
G Southern France
H Catalonia
I Navarre
J central Portugal

Fort & Pérez-Losada, *Nature Comm.* 2024

Ancient genetics

mtDNA haplogroup K is the most frequent one in early farmers. It is essentially absent in HGs



This pattern in early farmers suggests interbreeding with HGs

Isern,Fort & de Rioja, *Sci. Rep.* (2017)

Now we have data for more regions and can analyze both routes separately

Simulations

The genetic simulations are very similar to the archaeological ones explained above, but with 3 populations (not only 2):

P_N =farmers who *have* haplogroup K.

P_X = farmers who *do not have* haplogroup K.

P_{HG} = hunter-gatherers (all without haplogroup K).

$$\%K = \frac{P_N}{P_N+P_X} \cdot 100$$

Interbreeding [1-2]:

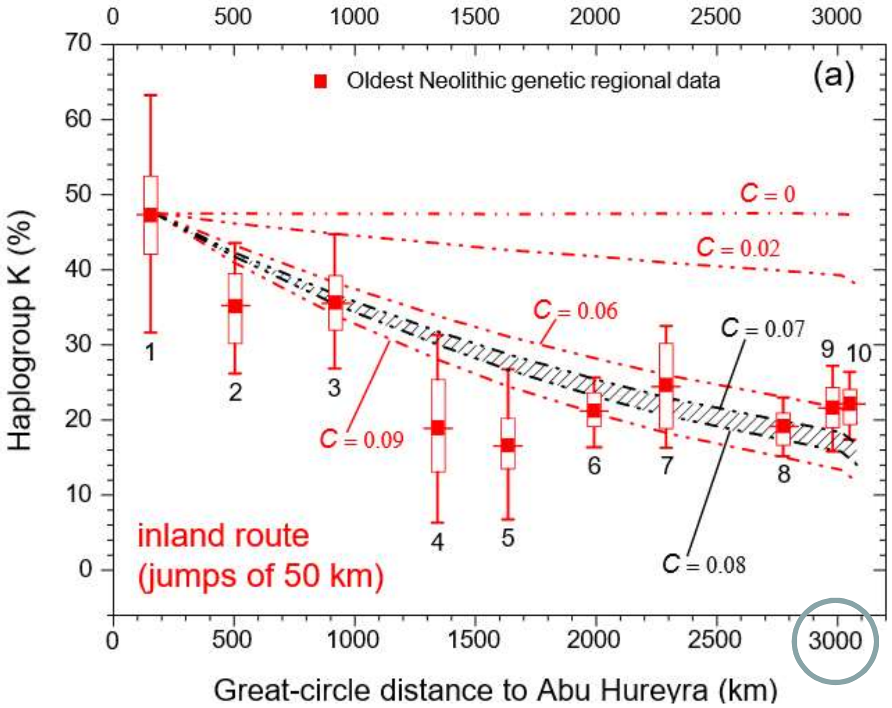
$$couples\ HN = C\frac{P_{HG}P_N}{P_{HG}+P_N+P_X}$$

$$couples\ HX = C\frac{P_{HG}P_X}{P_{HG}+P_N+P_X}$$

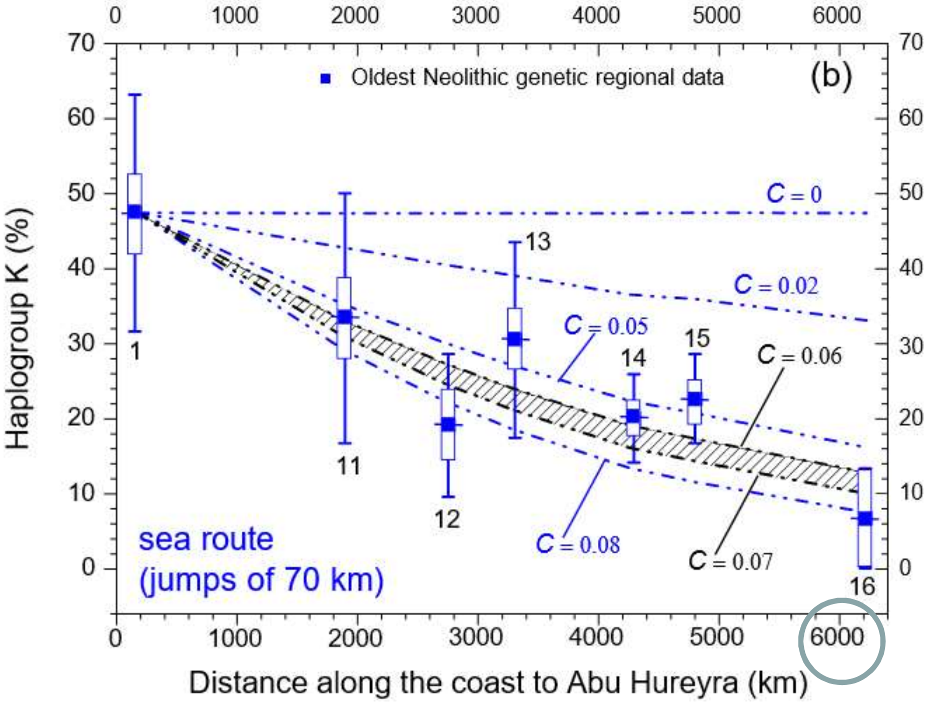
random mating for farmers→$couples\ NX = \frac{P_N P_X}{P_N+P_X}$

[1-2] Cavalli-Sforza & Feldman, book 1981; Fort, *Phys. Rev. E* 2011

Inland route



Sea route



Best fits: *C* = 0.07-0.08 **Best fits: *C* = 0.06-0.07**

Essentially the same!

Fort & Pérez-Losada, *Nature Comm.* 2024

·Previous slide: $C \approx 0.07$, so about
3.6% * of early farmers interbred with a HG.

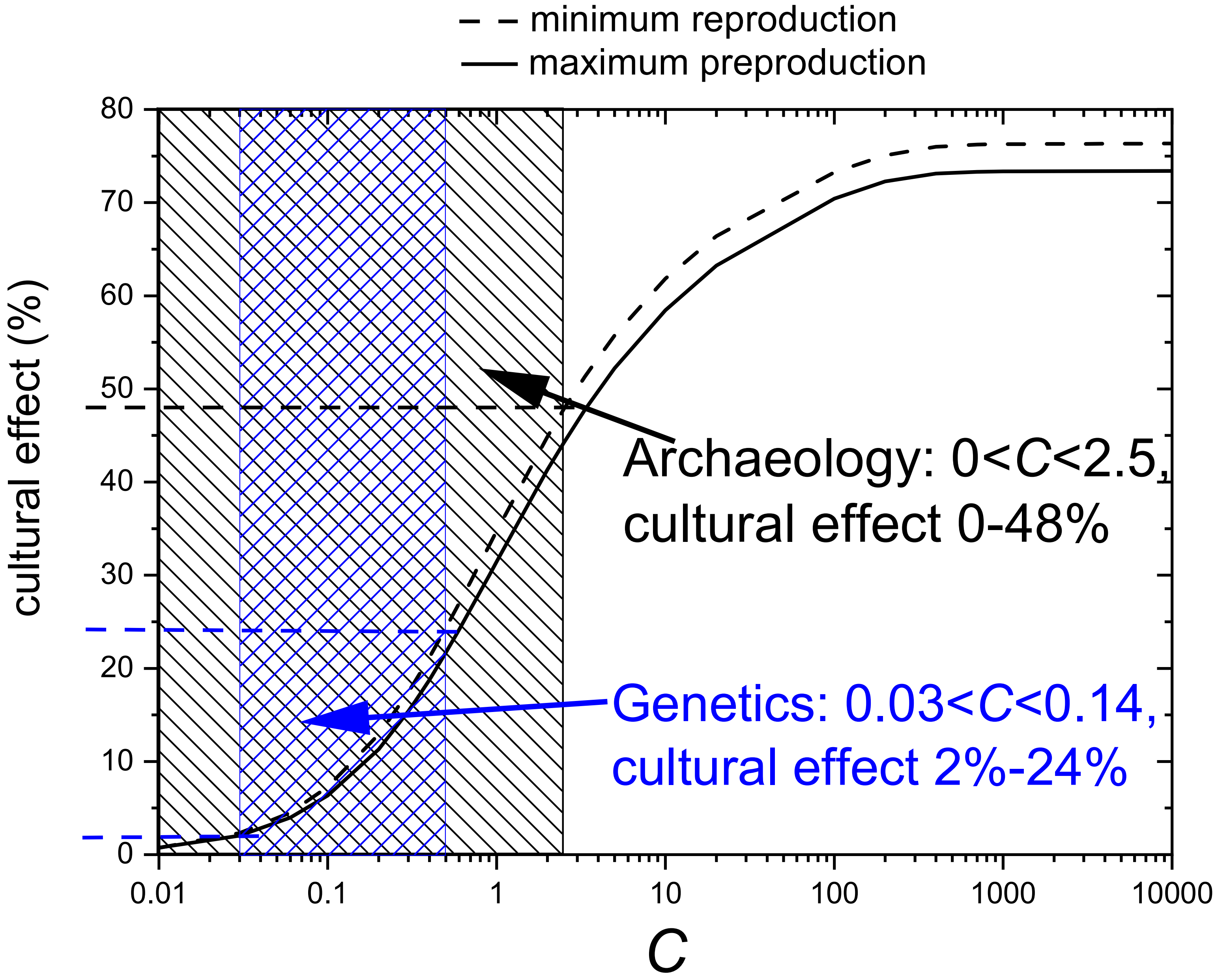
·Taking into account the uncertainties in the parameter values (r, $p_{F\,max}$, $p_{HG\,max}$, $p_{F\,min}$) and in the initial frequencies of haplogroup K: $0.03 < C < 0.14$, so
1% - 8% * of early farmers interbred with a HG.

Genetics → $0.03 < C < 0.14$

$$*\text{fraction of farmers} = \frac{P_F\,(t+1) - P_F(t)}{P_F(t)} = C\frac{P_{HG}(t)}{P_{HG}(t) + P_F(t)} = \frac{C}{1 + \frac{P_{F\,min}}{P_{HG\,max}}}$$

$P_F = P_N + P_X$

We can obtain independent estimations from archaeological and genetic data

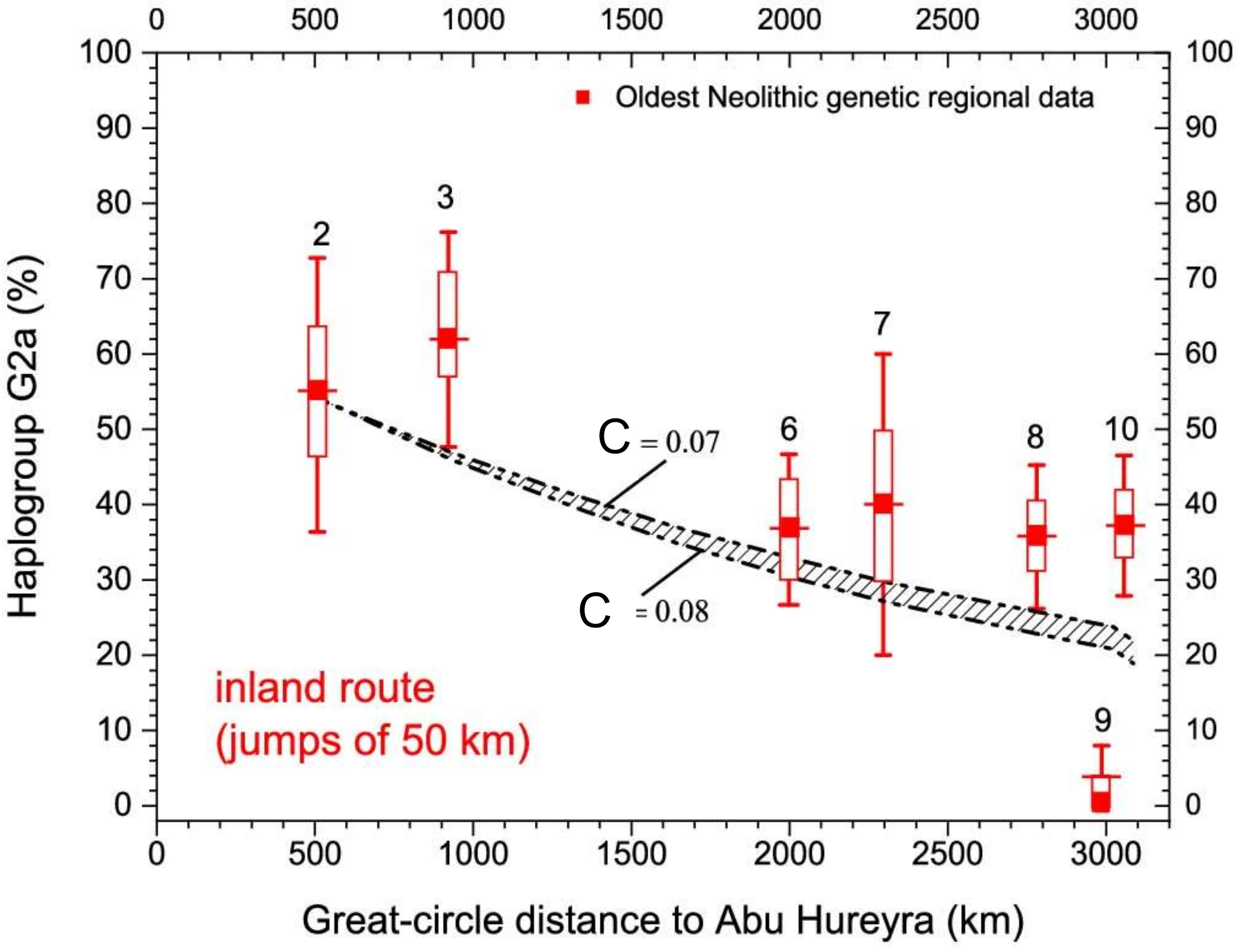


Archaeology: 0<*C*<2.5, cultural effect 0-48%

Genetics: 0.03<*C*<0.14, cultural effect 2%-24%

Y chromosome

Haplogroup G2 is the most frequent one in farmers.
It is essentially absent in hunter-gatherers

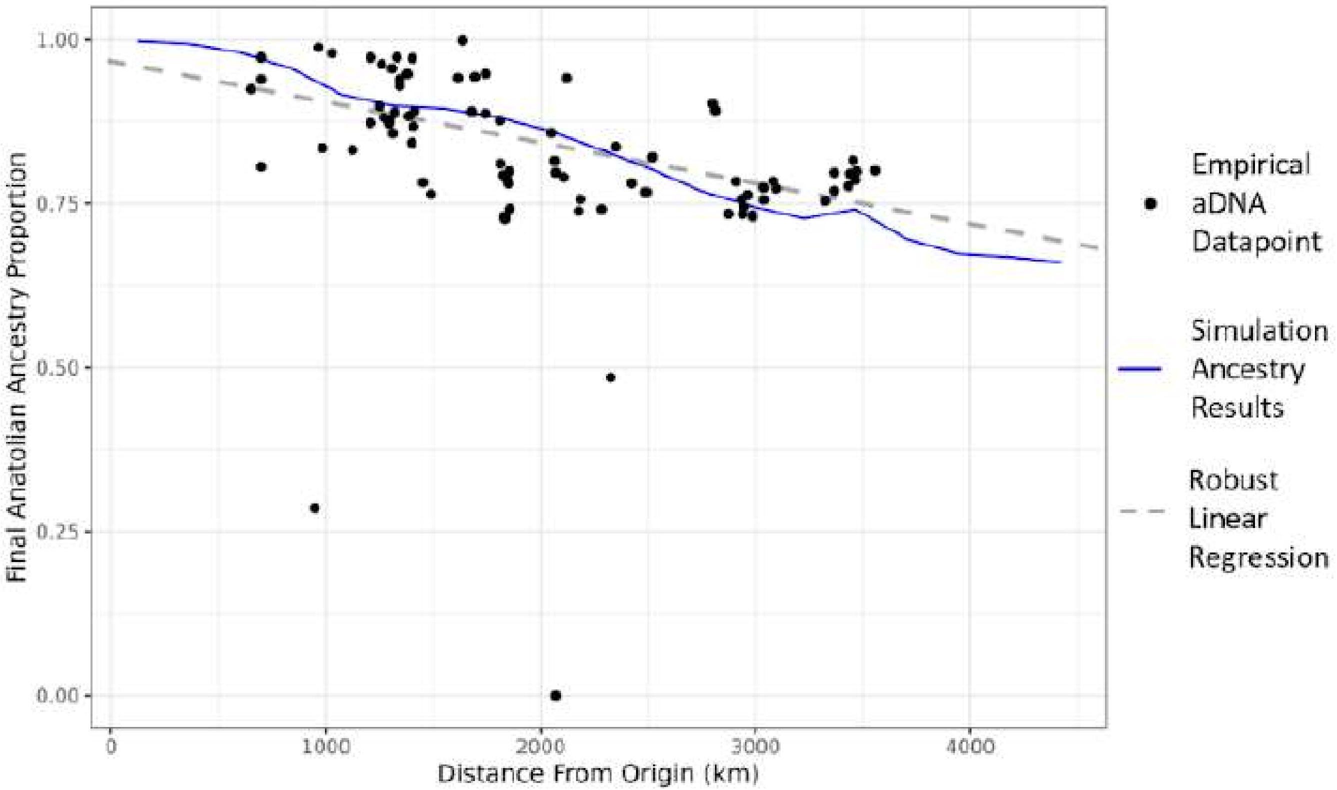


We obtain again *C ≈ 0.07,* in agreement with the mt DNA results

Fort & Pérez-Losada, *Nature Comm.* 2024

For the sea route there are not enough data

Whole genome



LaPolice, Williams & Huber, bioRxiv (2024)

They obtain that about 0.1% of early farmers interbred with a HG or acculturated him/her each year, i.e. about 0.1% ·32 yr = 3.2% per generation. This is consistent with our estimation (previous slides) that 1% - 8% * of early farmers interbred with a HG or acculturated him/her.

Conclusions

Archaeology tells us that:
Concerning dispersal, geography had a very important effect: early farmers moved longer distances per generation along the sea route than inland.

In sharp contrast to this, ancient genetics tells us that:
The interbreeding percentage of farmers was essentially the same along both routes. It did not depend on geography but only on the transition in the subsistence economy and its way of life.

Archaeology and genetics agree on the importance of cultural diffusion: 2%-24%