

The Neolithic spread in Europe: modelling population dynamics and interactions

Joaquim Fort

Universitat de Girona (Catalonia, Spain)

***ENE 2025 conference,
Zadar (Croatia), May 23rd 2025***



Academia Grant 2022-2026



Grant PID2023-150978NB-C22



Grant 2021-SGR-00190

Ammerman and Cavalli-Sforza (1984)

$$\text{spread rate} = s = 2\sqrt{aD}$$

where

a = initial growth rate

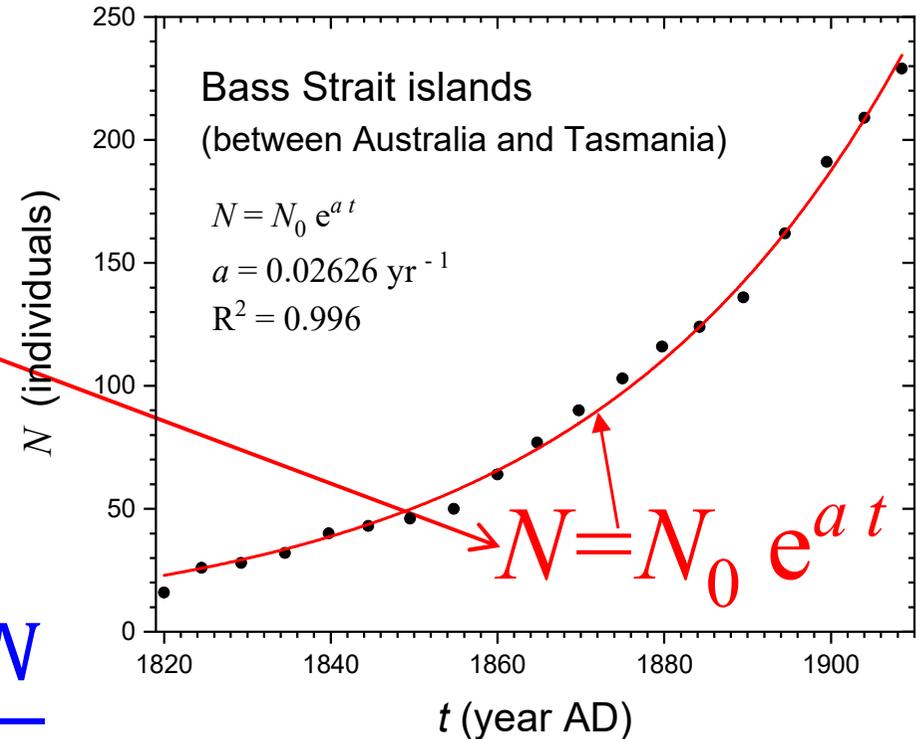
D = diffusion coefficient:

$$D = \frac{(r_1^2 + r_2^2 + \dots + r_N^2)/N}{4T}$$

with

r_1, r_2, \dots, r_N = intergenerational distance of individual 1, 2, ..., N .

T = generation time. What kind of distances?



Strictly, the correct distances are:

B distances = between birthplaces of parent and child

But very difficult to find. So 2 other distances have been used:

M distances = between birthplaces of spouses (M=mating)

R distances = between birthplace and place of residence

| Pre-industrial farming population | probabilities $\{p_1, p_2, p_3, \dots\}$ | distances $\{r_1, r_2, r_3, \dots\}$ |
|--|--|--|
| Majangir R ETHIOPIA. Amm. & C-S (1984) | {0.40, 0.17, 0.17, 0.26} | {2.4, 14.5, 36.2, 60.4} km |
| Issongos M CENTRAL AFRICAN REPUBLIC Ammerman & C-S (1984) | $\left\{ \begin{array}{l} 0.42, 0.23, 0.16, 0.08, 0.07, \\ 0.02, 0.01, 0.01 \end{array} \right\}$ | $\left\{ \begin{array}{l} 2.3, 7.3, 15, 25, 35, \\ 45, 55, 100 \end{array} \right\}$ km |
| Markazi B IRAN. Cobo et al. (2019) | $\left\{ \begin{array}{l} 0.799, 0.041, 0.022, 0.025, 0.064, \\ 0.004, 0.009, 0.021, 0.015 \end{array} \right\}$ | $\left\{ \begin{array}{l} 0.5, 5.5, 15, 25, 35, \\ 50.03, 57.20, 60.51, 97.65 \end{array} \right\}$ km |
| Bihar M INDIA. Cobo et al. (2019) | $\left\{ \begin{array}{l} 0.018, 0.081, 0.105, 0.129, 0.14, 0.125, \\ 0.107, 0.079, 0.068, 0.057, 0.036, 0.025, 0.03 \end{array} \right\}$ | $\left\{ \begin{array}{l} 2.5, 7.5, 12.5, 17.5, 22.5, 27.5, \\ 32.5, 37.5, 42.5, 47.5, 52.5, 57.5, 62.5 \end{array} \right\}$ km |
| Chandauli M INDIA. Cobo et al. (2019) | $\left\{ \begin{array}{l} 0.058, 0.122, 0.191, 0.256, 0.168, \\ 0.101, 0.069, 0.023, 0.012 \end{array} \right\}$ | $\left\{ \begin{array}{l} 4, 12, 20, 28, 36, \\ 44, 52, 60, 68 \end{array} \right\}$ km |

Two problems

- How can we test if the 3 kinds of distances give similar results?
- Histograms may not yield accurate spread rates, specially for long distances. Can we use lists of distances?

There is only one census with individual data (as far as we know), such that we can calculate B, M and R distances for the same population (Yanomamö, from Brazil and Venezuela):

Biella, P., Chagnon, N.A. & Seaman, G., *Yanomamö Interactive. The Ax Fight (CD-ROM)*, Orlando: Hartcourt Brace & Co. (1997).

We used this census to calculate lists (not histograms) of B, M and R distances (next slide)

Table S1 to Bancells & Fort, *Archaeol. Anthropol. Sci.* (2024)

B distances (i.e., between the birthplaces of parent and child)

| person | person | person | person | parent | parent | parent | parent | B distance |
|--------|---------|---------------|----------------|----------|---------------|---------------------|----------------------|-------------|
| ID | POB. ID | POB. latitude | POB. longitude | ID FA/MO | POB. ID FA/MO | POB. FA/MO latitude | POB. FA/MO longitude | DIST. FA/MO |
| | | ° N | ° W | | | ° N | ° W | km |
| 227 | 145 | 1.98 | 64.57 | 777 | 145 | 1.98 | 64.57 | 0.00 |
| 1022 | 126 | 1.60 | 65.27 | 651 | 126 | 1.60 | 65.27 | 0.00 |
| 1246 | 124 | 1.78 | 65.15 | 1929 | 124 | 1.78 | 65.15 | 0.00 |
| 1795 | 145 | 1.98 | 64.57 | 777 | 145 | 1.98 | 64.57 | 0.00 |
| 2384 | 128 | 1.68 | 65.27 | 1046 | 128 | 1.68 | 65.27 | 0.00 |
| 2531 | 126 | 1.60 | 65.27 | 522 | 126 | 1.60 | 65.27 | 0.00 |
| 2518 | 9 | 1.55 | 65.37 | 1834 | 113 | 1.55 | 65.38 | 1.85 |
| 159 | 126 | 1.60 | 65.27 | 2130 | 118 | 1.62 | 65.30 | 4.14 |
| 951 | 126 | 1.60 | 65.27 | 950 | 118 | 1.62 | 65.30 | 4.14 |
| 1125 | 126 | 1.60 | 65.27 | 1509 | 118 | 1.62 | 65.30 | 4.14 |
| 2380 | 126 | 1.60 | 65.27 | 1568 | 118 | 1.62 | 65.30 | 4.14 |
| 2398 | 126 | 1.60 | 65.27 | 950 | 118 | 1.62 | 65.30 | 4.14 |

In this way we found **257 B distances**.

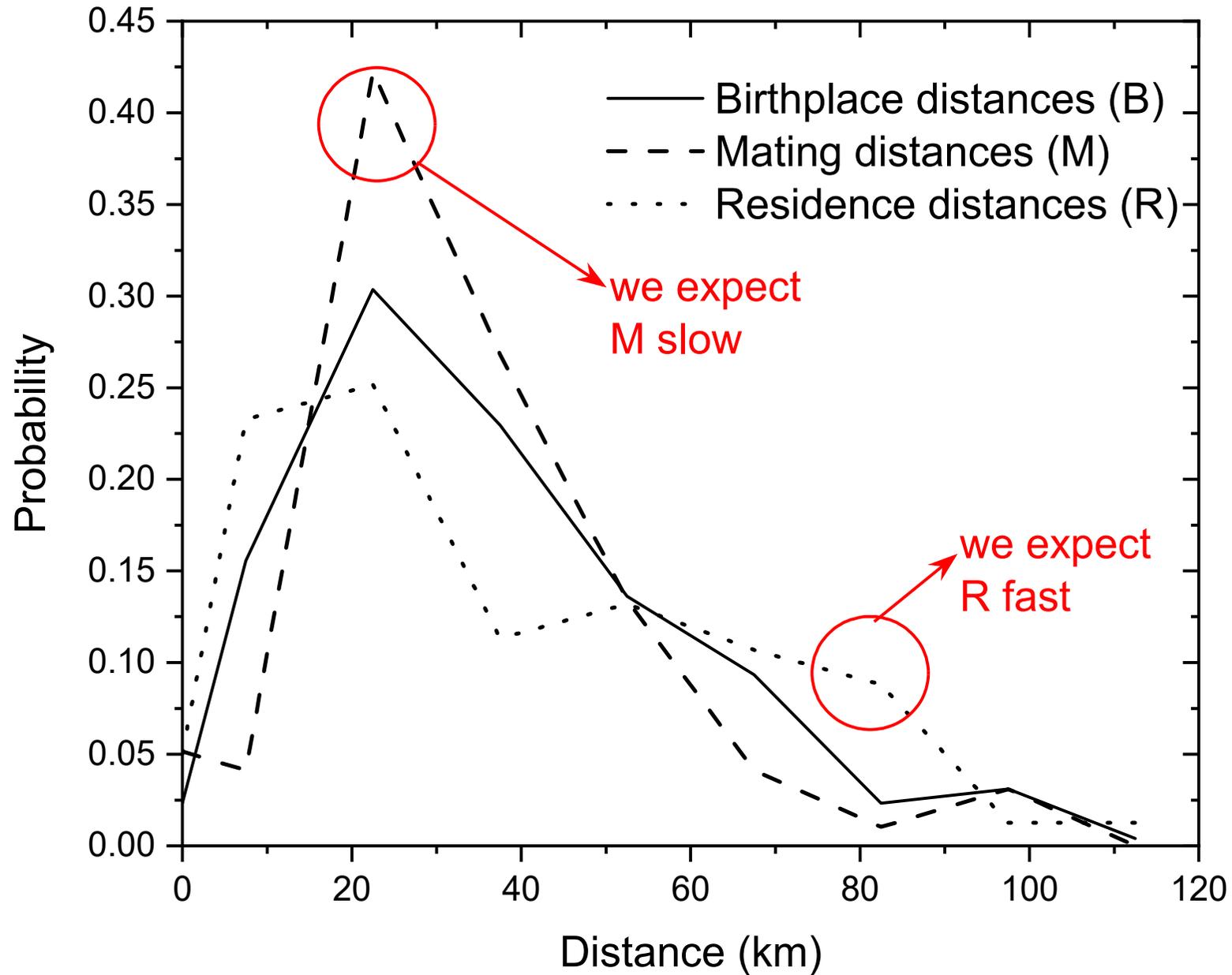
Similarly, we found

159 R distances and

97 M distances

We admit that the Yanomamö are a single population of horticulturalists living in the Amazonian forest. So we need more data, ideally from Neolithic Europe (using aDNA). Or at least censuses for other populations of pre-industrial farmers. If you know one, please let me know! This is only a first step.

Yanomamö



Simulations

Each generation ($t = 1, 2, 3, \dots$) :

- 1) Reproduction (logistic, with a and T from ethnography)
- 2) Dispersal, using the **list of B, R or M distances**
- 3) Interbreeding: $P_N = \text{farmers/km}^2$, $P_{HG} = \text{HGs/km}^2$

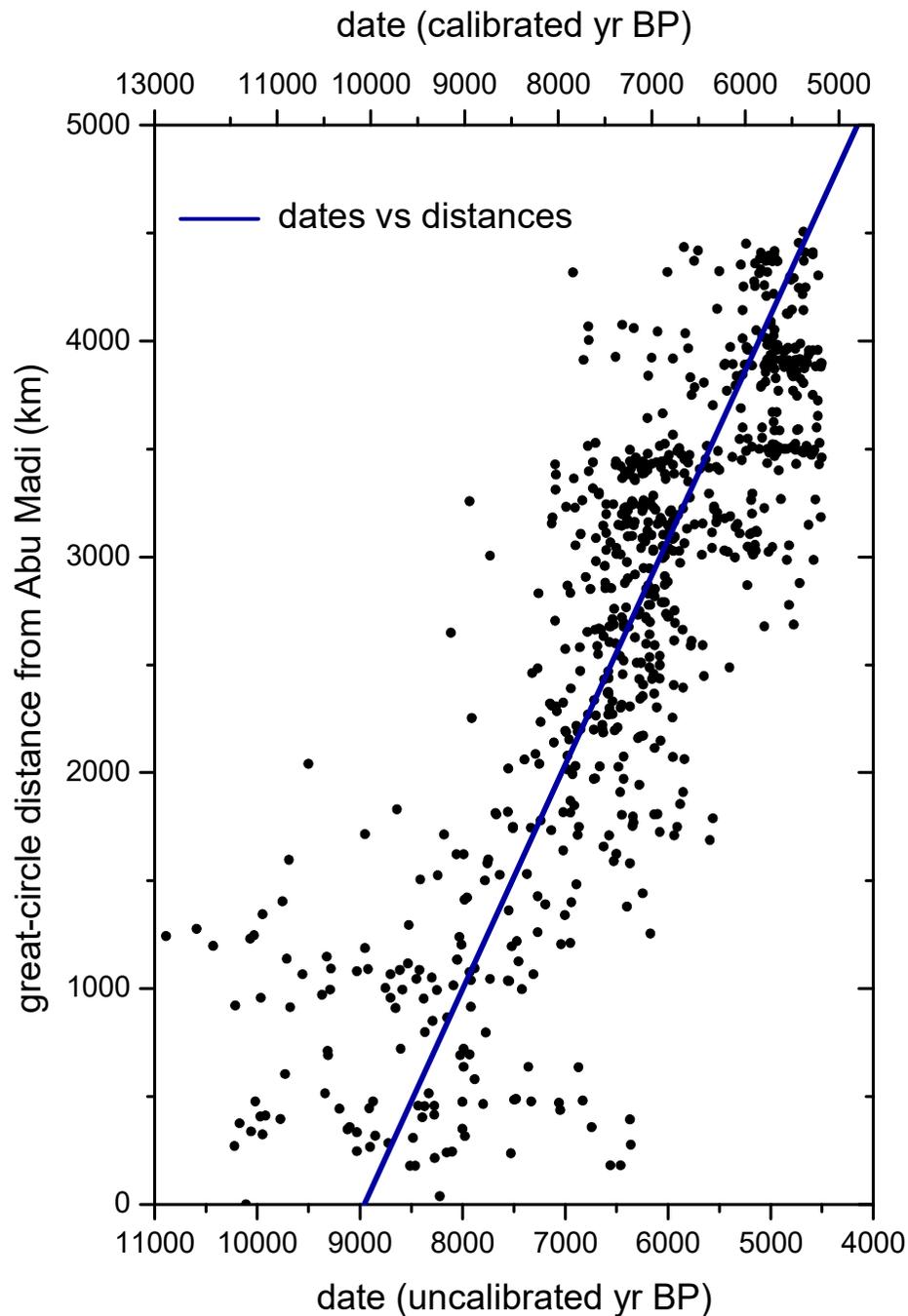
$$P_N(t + 1, x, y) - P_N(t, x, y) = C \frac{P_{HG} P_N}{P_{HG} + P_N} \approx C P_N \text{ if } P_N \ll P_{HG}$$

$$P_{HG}(t + 1, x, y) - P_{HG}(t, x, y) = C \frac{P_{HG} P_N}{P_{HG} + P_N} \approx -C P_N$$

$C =$ fraction of early farmers who interbreed with a HG.

For acculturation we can use the same Eqs. Then:

$C =$ number of HGs acculturated by an early farmer.



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

0.9-1.3 km/yr

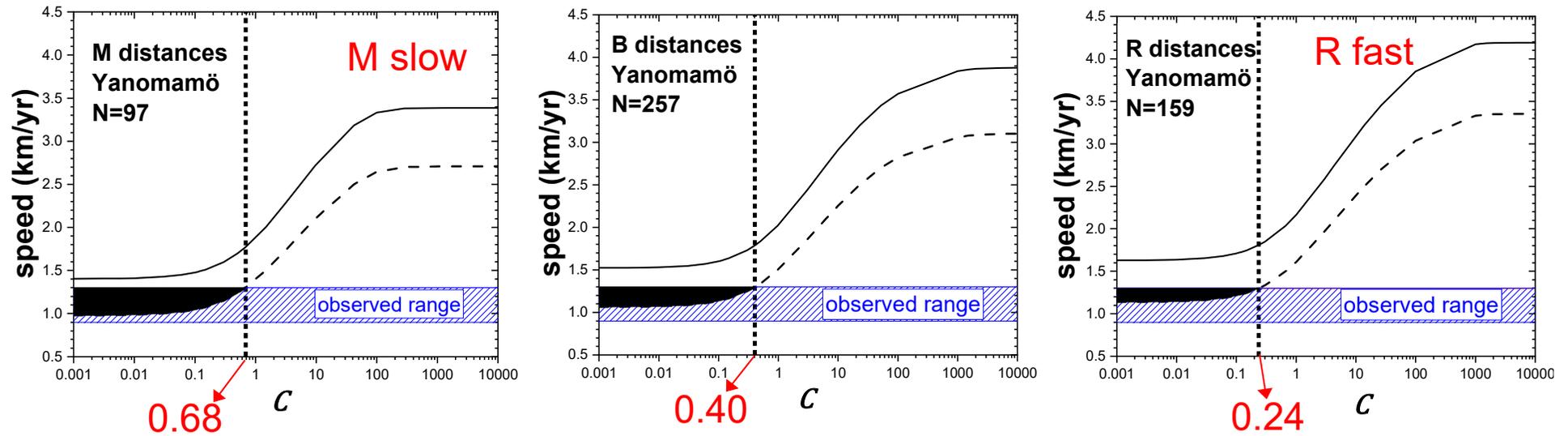
great circles & shortest paths

$$r = 0.83$$

(for both highest- r origins)

735 sites in Europe & Near East

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



The 3 distances lead to similar curves (and consistent with the observed range).

From these figs., we can find ranges for C :

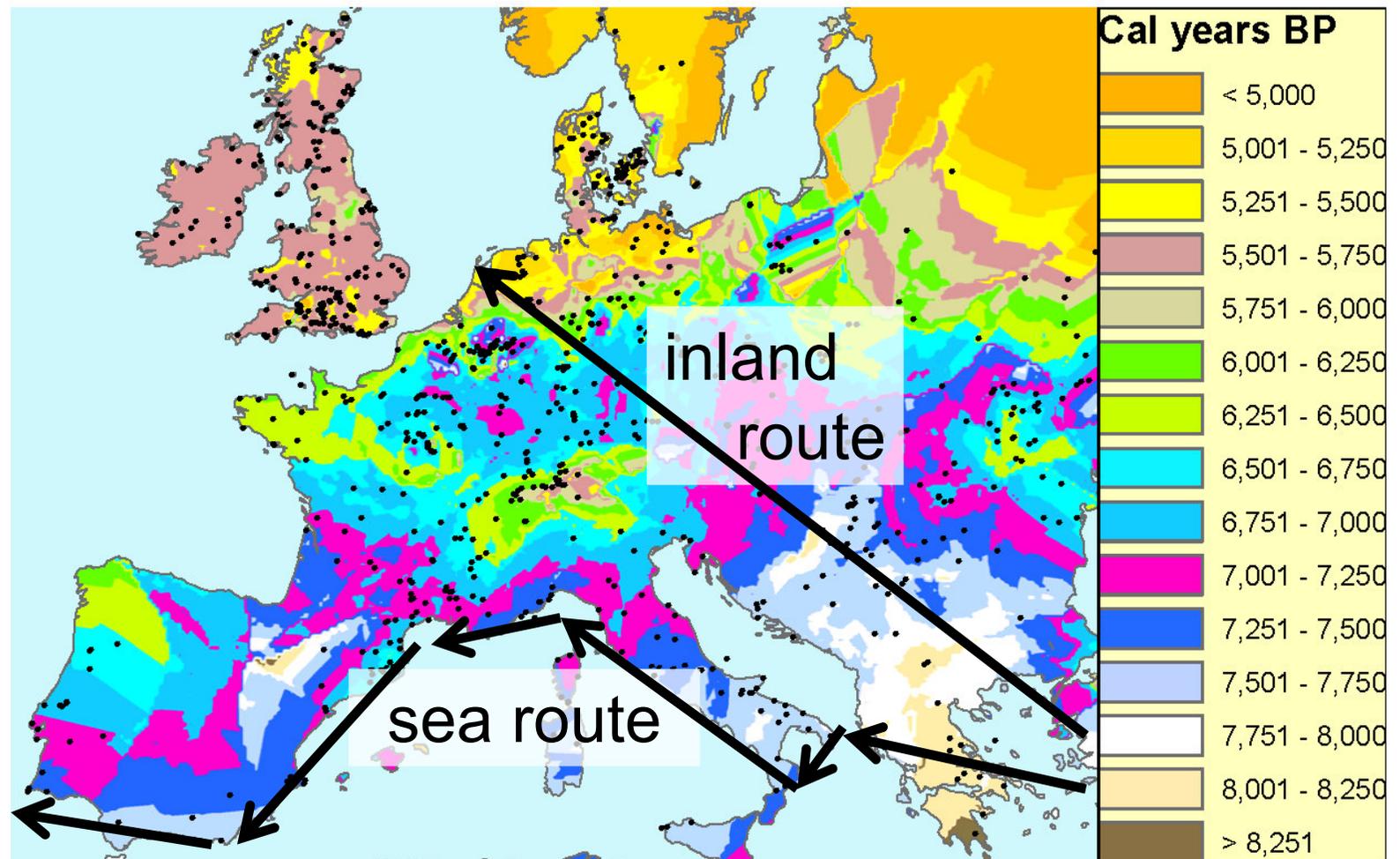
$0 < C < 0.68$ (M distances)

$0 < C < 0.40$ (B distances, the most reliable ones)

$0 < C < 0.24$ (R distances)

Therefore, Archaeology implies that **$0 < C < 0.40$ (B distances)**, i.e. the percentage of early farmers who interbred with a HG or acculturated one of them was **between 0% and 40%**.

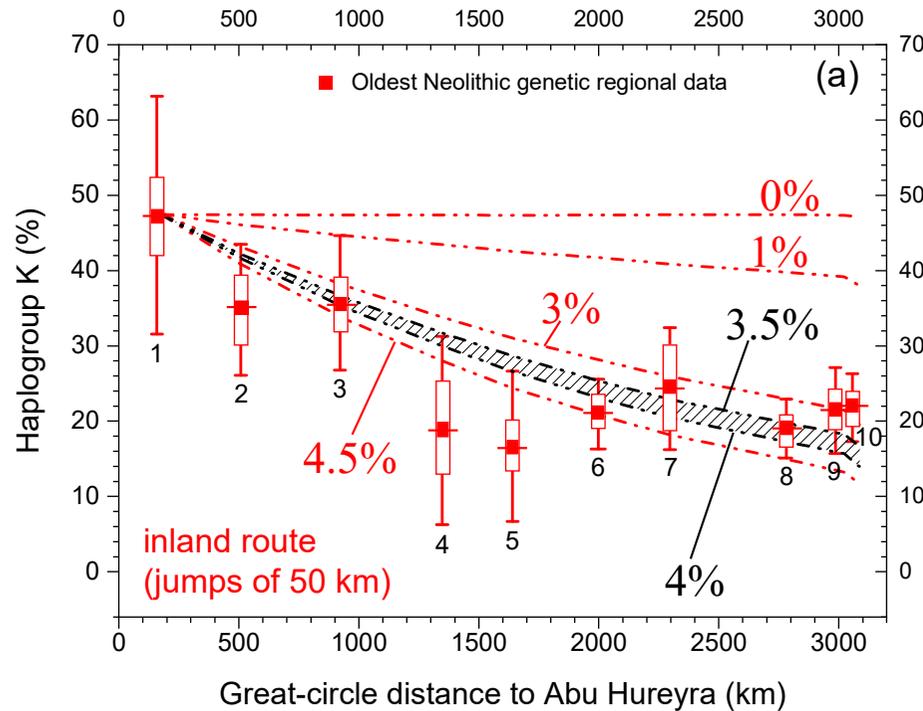
Ancient DNA along two routes



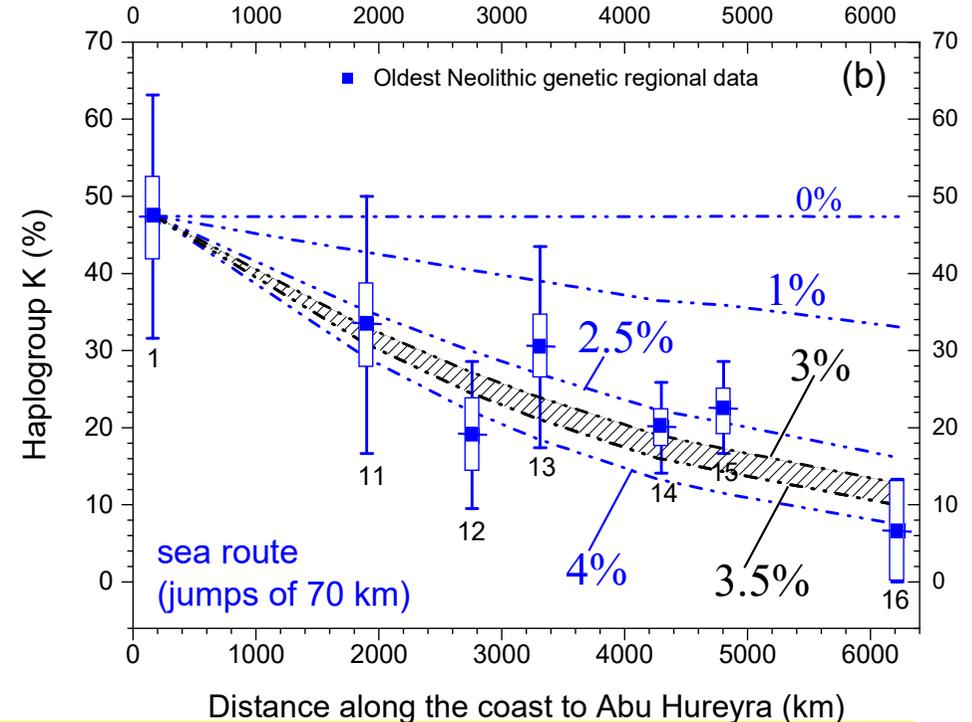
Interpolation from Fort, *J. R. Soc. Interface* (2015)

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

Inland route



Sea route



Best fits: 3.5% - 4%

Best fits: 3% - 3.5%

The percentage of early farmers who interbred with HGs was essentially the same (~3.5%)!

Previous slide:

~3.5% of early farmers interbred with
a HG

Taking into account the uncertainties in the
parameter values and in the initial
frequencies of haplogroup K:

1% - 8% of early farmers interbred with a HG

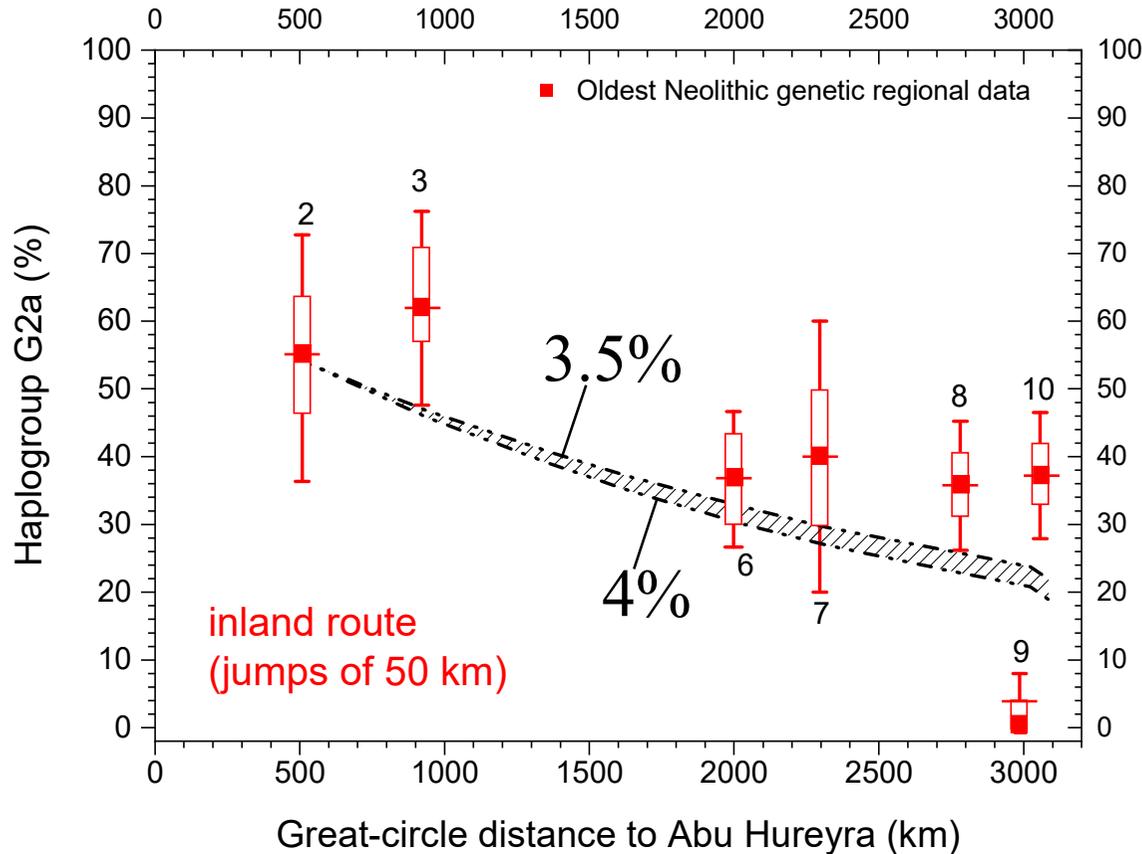
New results (unknown 1 year ago):

- Y chromosome
- whole genome

(next slides)

Y chromosome

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

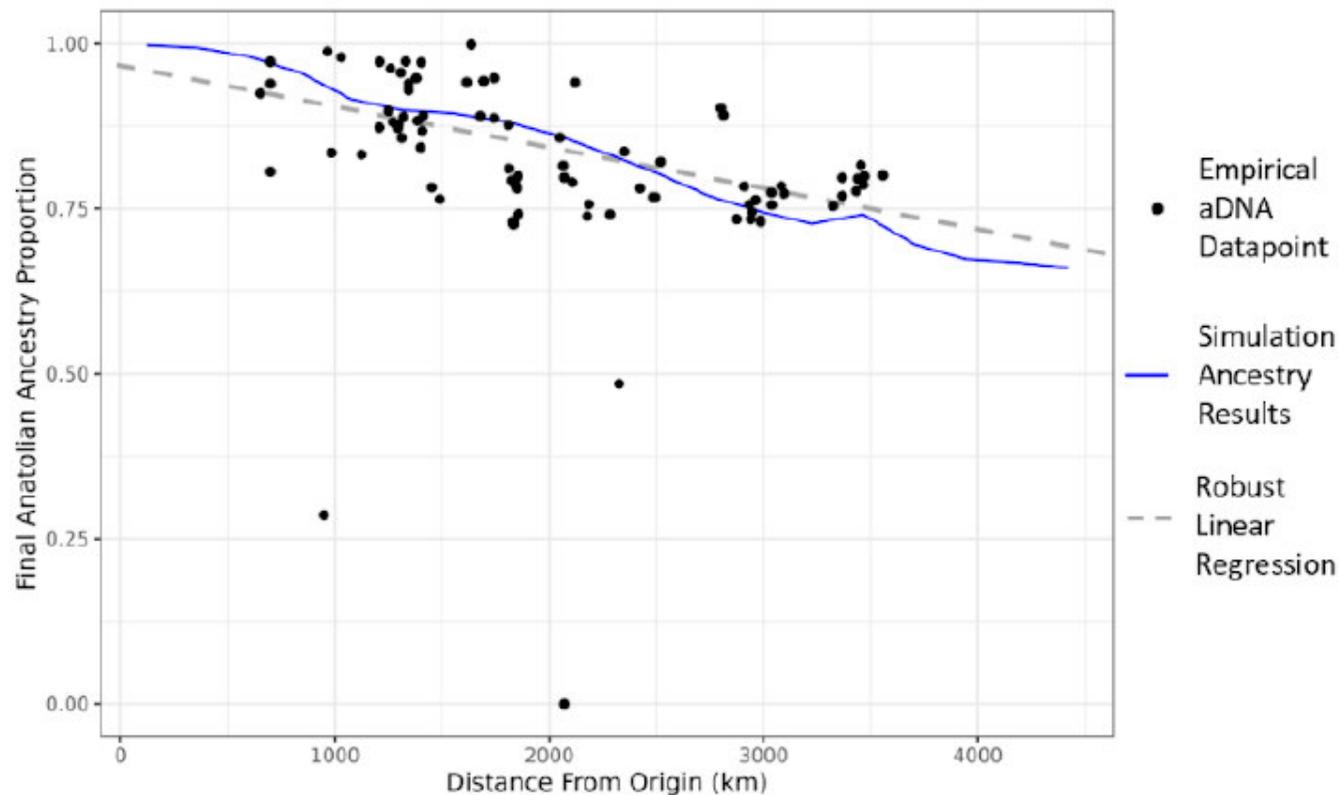


We obtain again ~3.5%, in agreement with the mt DNA results

For the sea route there are not enough data yet.

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

Whole genome



LaPolice,
Williams
& Huber,
bioRxiv
(2024)

Blue line assumes that 0.1% of early farmers acculturated a HG per year, i.e. about $0.1\% \cdot 32 \text{ yr} = 3.2\%$ per generation.

This is consistent with our estimation that 1% - 8% of early farmers interbred with a HG or acculturated him/her.

Conclusions

- B, R and M distances are reliable according to the data available.
- **Archaeology** implies that the percentage of early farmers who interbred with a HG or acculturated one of them was **between 0% and 40%**
- **Ancient DNA: 3 different kinds of data** (mtDNA, Y-chromosome and the whole genome) lead to the same result: **between 1% and 8%** of early farmers interbred with a HG or acculturated him/her.
- The range from **ancient DNA** is much more accurate than that from **Archaeology**.