Supporting Information

S1. Uncertainty analysis of the parameter values

In this section we estimate the generation time T and the initial growth rate a using ethnographic data.

S1a. Generation time T

For a given population, consider a sample of n parent-child pairs and let T_i stand for the generation time of pair $i = 1, 2, \ldots, n$. The mean generation time of the sample is

$$\bar{T} = \frac{\sum_{i=1}^{n} T_i}{n},\qquad(1)$$

and its standard deviation is [1]

$$\sigma = \frac{\sqrt{\sum_{i=1}^{n}(T_i - \bar{T})^2}}{\sqrt{n-1}} = \sqrt{Var},\qquad(2)$$

where Var is called the variance. Equation (2) refers to the distribution of the original data. In order to estimate the error of the mean (1), statistical theory also considers the distribution of the means, i.e., of values of $\bar{T}$ such that each value is the mean over a different sample of n pairs drawn from the original population. According to well-known theorems [1], (i) the mean of the distribution of means is the same as the mean of the distribution of the original data, Eq. (1); (ii) the standard deviation of the distribution of means is[1]

$$\sigma_m = \frac{\sigma}{\sqrt{n}},\qquad(3)$$

where σ is the standard deviation of the original data, Eq. (2). If the original distribution of T values has an approximately normal (i.e., Gaussian) distribution and $n \geq 30$, the error $\Delta\bar{T}$ in the average generation time $\bar{T}$ with 95% confidence level (CL) is [1]

$$\Delta\bar{T} = z_{0.975}\, \sigma_m,\qquad(4)$$

where $z_{0.975} = 1.96$ is the value of z for which the standard normal distribution $f(z)$ has probability 95% (i.e., the area under the curve is 0.95) [1]. Thus with 95% CL, the range of the mean is $(\bar{T} - \Delta\bar{T}, \bar{T} + \Delta\bar{T})$. It is important to keep in mind that, in contrast to Eq.

---

[1] It is easy to proof these theorems, i.e., to derive the mean and the standard deviation (3) by noting that T_1 takes different values for different samples, but its mean and variance will be the same as those of T_2, T_3, etc. Thus the expected value of the mean is $E(\bar{T}) = E\left(\frac{\sum_{i=1}^{n} T_i}{n}\right) = \frac{1}{n}\sum_{i=1}^{n} E(T_i) = \frac{1}{n}\sum_{i=1}^{n}\bar{T} = \frac{1}{n}n\bar{T} = \bar{T}$ and the standard deviation of the mean is $\sigma_m^2 = Var(\bar{T}) = Var\left(\frac{\sum_{i=1}^{n} T_i}{n}\right) = \frac{1}{n^2}\sum_{i=1}^{n} Var(T_i) = \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2 = \frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n}$, where we have applied that the variances of any random variables x, y satisfy that $Var(const \cdot x) = const^2\, Var(x)$ and, if x and y are independent, $Var(x + y) = Var(x) + Var(y)$.

(2), Eqs. (3)-(4) refer to the distribution of the mean $\bar{T}$ and not to that of the original variable T [1] [2].

Next we apply this approach to the pre-industrial populations for which we have found data for the generation time T. It is useful to keep in mind that the generation time T (age of a parent when a child is born) has to be considered for all children (not only for the first one) because all of them have an effect on the propagation of the population front [2].

Population 1. Yanomamö

In our excel supplementary file (table S5) we include the generation times of $n = 258$ parent-child pairs, which we have obtained from the same census that we have used to find out distances in our main paper [3]. Using Eqs. (1)-(4) we find that the mean is $\bar{T} = 28.1$yr and the standard deviation of the sample is $\sigma = 9.5$ yr, so the 95% error is $\Delta\bar{T} = 1.2$ yr and the 95% CL for the mean is

$$(\bar{T} - \Delta\bar{T}, \bar{T} + \Delta\bar{T}) = (26.9 - 29.3) \text{ yr.} \tag{5}$$

Population 2. Majangir

For this pre-industrial population of farmers in Ethiopia, an histogram with 4 bins and 96 data reported in Ref. [4], p. 58, table 5 has been used to estimate that 44 parent-child pairs had $T = 27$ yr, 49 pairs $T = 35.5$ yr, 2 pairs $T = 45.5$ yr and 1 pair $T = 55.5$ yr [2]. Then the mean generation time given by Eq. (1) is $\bar{T} = 32.0$ yr and Eq. (2) yields for the standard deviation $\sigma = 0.53$ yr. This is substantially smaller than that for population 1, due to the fact that two contiguous bins in the histogram concentrate 97% of the data. For this reason, Eqs. (3)-(4) also lead to a much smaller 95% CL error for the mean, namely $\Delta\bar{T} = 0.1$ yr and therefore

$$(\bar{T} - \Delta\bar{T}, \bar{T} + \Delta\bar{T}) = (31.9 - 32.1) \text{ yr.} \tag{6}$$

For population 3 below, we will also find a rather small value of $\Delta\bar{T}$ compared to that for population 1.

Population 3. Parma

In year 1967 Cavalli-Sforza and Zei [5] reported means and variances of generation times calculated for marriages in the high Parma valley (Italy). These data refer to rural populations and were obtained from parish books spanning the last 3 centuries. The raw data are not available and they published the mean and variance separately for mothers (m) and fathers (f). In order to find out the standard deviation σ for the complete sample, we write

---

[2] Equation (3) implies that $\sigma_m < \sigma$, i.e., the distribution of the mean $\bar{T}$ is always narrower than that of the original variable T. Thus the 95% CL interval given by Eq. (4), which contains 95% of the values of $\bar{T}$, does not contain 95% of the n original values of T. According to Eqs. (3)-(4), the mean $\bar{T}$ has a lower error $\Delta\bar{T}$ the higher the value of the number of data n over which it is calculated, as we should expect intuitively.

$$\sigma = \sqrt{Var}\,, \tag{7}$$

where, following Eq. (2), $Var = \frac{\sum_{i=1}^{n}(T_i - \bar{T})^2}{n-1}$ is the variance of the sample. We separate it into two contributions due to the n_m mother-child pairs and n_f father-child pairs,

$$Var = \frac{\sum_{i=1}^{n}(T_i - \bar{T})^2}{n-1} = \frac{\sum_{i=1}^{n_m}(T_i - \bar{T})^2 + \sum_{i=1}^{n_f}(T_i - \bar{T})^2}{n-1}. \tag{8}$$

Let $\bar{T}_m$ stand for the mean of the generation times of mother-child pairs, and $\bar{T}_f$ for that of father-child pairs. The first term in the numerator can be written as

$$\sum_{i=1}^{n_m}(T_i - \bar{T})^2 = \sum_{i=1}^{n_m}(T_i - \bar{T}_m + \bar{T}_m - \bar{T})^2 = \sum_{i=1}^{n_m}(T_i - \bar{T}_m)^2 + \sum_{i=1}^{n_m}(\bar{T}_m - \bar{T})^2 + 2(\bar{T}_m - \bar{T})\sum_{i=1}^{n_m}(T_i - \bar{T}_m). \tag{9}$$

The last term cancels out because $\sum_{i=1}^{n_m}(T_i - \bar{T}_m) = n_m\bar{T}_m - n_m\bar{T}_m = 0$. By repeating the same procedure for the second term in (8) and using Eqs. (2) and (7), we find

$$Var = \frac{(n_m-1)Var_m + n_m(\bar{T}_m - \bar{T})^2 + (n_f-1)Var_f + n_f(\bar{T}_f - \bar{T})^2}{n_m + n_f - 1}, \tag{10}$$

where $Var_m = \frac{\sum_{i=1}^{n_m}(T_i - \bar{T}_m)^2}{n_m - 1}$ is the variance of mother-child pairs, $Var_f = \frac{\sum_{i=1}^{n_f}(T_i - \bar{T}_f)^2}{n_f - 1}$ is the variance of father-child pairs, and the total number of pairs is $n = n_m + n_f$.

For the population studied by Cavalli-Sforza and Zei [5] $\bar{T}_m = 30.9$ yr, $Var_m = 42.6$ yr, $n_m = 1000$, $\bar{T}_f = 36.1$ yr, $Var_f = 53.3$ yr and $n_f = 1000$. Therefore $\bar{T} = \frac{30.9 + 36.1}{2} = 33.5$ yr and using these values into Eq. (10) yields $Var = 54.69$ yr and $\sigma = \sqrt{Var} = 7.4$ yr. Using Eqs. (3)-(4), we find that $\Delta\bar{T} = 0.3$ yr so the 95% CL range for the mean is

$$(\bar{T} - \Delta\bar{T}, \bar{T} + \Delta\bar{T}) = (33.2 - 33.8)\ \text{yr} \tag{11}$$
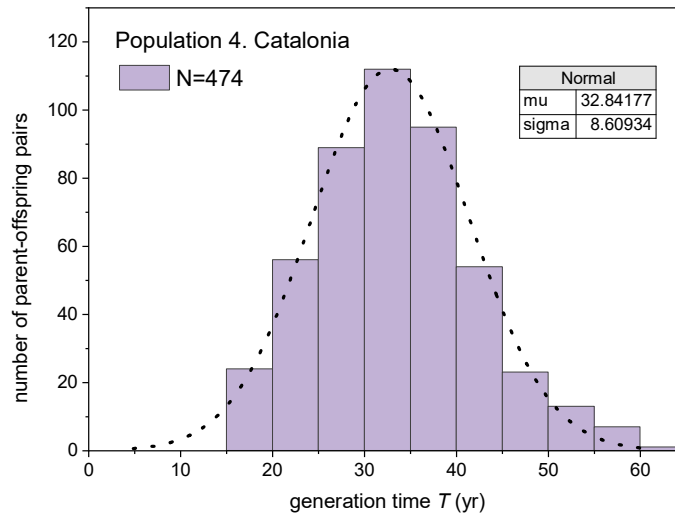
Population 4. Catalonia

We have used a book with genealogical trees of rural populations [6] and collected 474 parent-child pairs such that the parent was born before year 1,700 AD in Catalonia (present-day northeastern Spain and southeastern France). The oldest parent was born in year 1,295, so this database spans 4 centuries. We include the database and additional comments as table S6 in our excel supplementary file. Using Eqs. (1)-(4) we find that the mean is $\bar{T} = 32.8$ yr and the standard deviation of the sample is $\sigma = 8.6$ yr, so the 95% error is $\Delta\bar{T} = 0.8$ yr and the 95% CL for the mean is

$$(\bar{T} - \Delta\bar{T}, \bar{T} + \Delta\bar{T}) = (32.0 - 33.6)\ \text{yr} \tag{12}$$

Overall range

Next we will combine the results for the 4 populations above to derive an overall range for the generation time T. This is necessary to calculate front speeds (Figs. 2-4 in the main paper).

Firstly it is interesting to consider, for each of the 4 populations, the Gaussian distribution with the same mean $\bar{T}$ and standard deviation σ as the corresponding sample (as estimated above). Figure S1 shows an example (population 4), and we see that the histogram of the original data is approximately Gaussian. This justifies the use of Eq. (4) for the 4 individual populations, as done above.

Fig. S1. Histogram of the generation time T for population 4 (Catalonia) and Gaussian distribution (dotted line) with the same mean ($\bar{T} = 32.8$ yr) and average ($\sigma = 8.6$ yr) as the histogram.
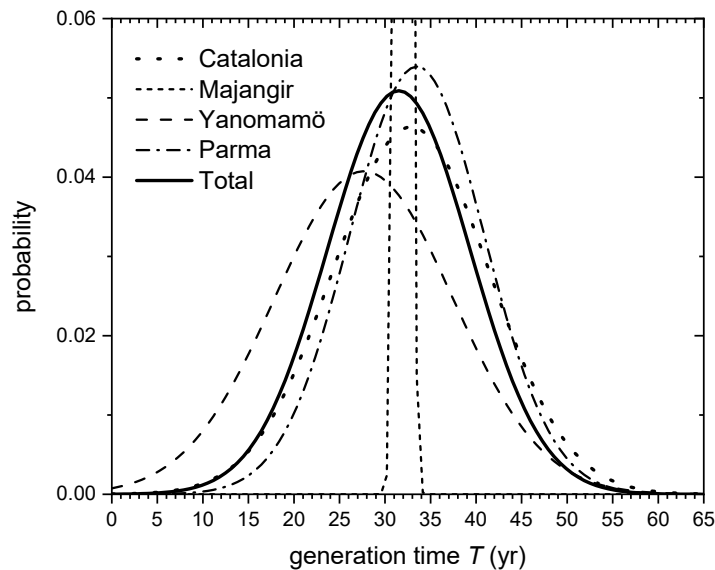
Fig. S2. Gaussian distributions of T for the 4 populations considered and total distribution (full line). The latter is the sum of the 4 distributions divided by 4 to ensure normality (i.e., that the area below the full curve is 1). The dotted line is the same as in Fig. S1.

In Fig. S2 we plot the sum of the four distributions[3] (divided by 4 to ensure normality) as a full line. In order to find out its mean and standard deviation, we consider the total distribution of generation times T (full line in Fig. S2)

$$f_{total}(T) = \frac{f_1(T)+f_2(T)+f_3(T)+f_4(T)}{4}, \tag{13}$$

where $f_i(T)$ is the distribution for population i (as also plotted in Fig. S2). The mean of $f_{total}(T)$ is

$$\bar{T}_{total} = \int_0^\infty T\, f_{total}(T)\, dT = \frac{1}{4}\sum_{i=1}^4 \int_0^\infty T\, f_i(T)\, dT = \frac{\bar{T}_1+\bar{T}_2+\bar{T}_3+\bar{T}_4}{4}. \tag{14}$$

Using the results above for the 4 populations, i.e., $\bar{T}_1 = 28.1$ yr, $\bar{T}_2 = 32.0$ yr, $\bar{T}_3 = 33.5$ yr and $\bar{T}_4 = 32.8$ yr, we find that

$$\bar{T}_{total} = 31.6 \text{ yr.} \tag{15}$$

The standard deviation of $f_{total}(T)$ is $\sigma_{total} = \sqrt{Var_{total}}$, and the variance of $f_{total}(T)$ is found from its definition [1] and Eq. (13),

$$Var_{total} = \int_0^\infty (T - \bar{T}_{total})^2\, f_{total}(T)\, dT = \frac{1}{4}\sum_{i=1}^4 \int_0^\infty (T - \bar{T}_{total})^2\, f_i(T)\, dT =$$

$$\frac{1}{4}\sum_{i=1}^4 \int_0^\infty (T - \bar{T}_i + \bar{T}_i - \bar{T}_{total})^2\, f_i(T)\, dT =$$

$$\frac{1}{4}\sum_{i=1}^4 \left[\int_0^\infty (T - \bar{T}_i)^2\, f_i(T)\, dT + \int_0^\infty (\bar{T}_i - \bar{T}_{total})^2\, f_i(T)\, dT + \int_0^\infty 2(T - \bar{T}_i)(\bar{T}_i - \bar{T}_{total})\, f_i(T)\, dT\right] =$$

$$\frac{1}{4}\sum_{i=1}^4 \left[Var_i + (\bar{T}_i - \bar{T}_{total})^2 + 2(\bar{T}_i - \bar{T}_{total})\int_0^\infty (T - \bar{T}_i) f_i(T)\, dT\right], \tag{16}$$

where we have applied the normalization of probability, i.e., $\int_0^\infty f_i(T)\, dT = 1$. The last term vanishes because $\int_0^\infty T f_i(T)\, dT = \bar{T}_i$, as already applied in Eq. (14). Therefore

$$Var_{total} = \frac{1}{4}\sum_{i=1}^4 [Var_i + (\bar{T}_i - \bar{T}_{total})^2] = \frac{1}{4}\sum_{i=1}^4 \left[Var_i + \bar{T}_i^2 + \bar{T}_{total}^2 - 2\bar{T}_i\bar{T}_{total}\right] =$$

---

[3] All of the 4 original distributions clearly overlap (Fig. S2). However not all of the 4 ranges (5), (6), (11) and (12) overlap, because these refer to the distribution of the mean $\bar{T}$ and its standard deviation σ_m. The reason is that, as mentioned above, Eqs. (3)-(4) correspond to the distribution of the mean $\bar{T}$ and not to the distribution of the generation time T [1].

$$\bar{T}_{total}^{\,2} - 2\bar{T}_{total}\frac{1}{4}\sum_{i=1}^{4}\bar{T}_i + \frac{1}{4}\sum_{i=1}^{4}[Var_i + \bar{T}_i^2\,]. \tag{17}$$

Using Eq. (14),

$$Var_{total} = \bar{T}_{total}^{\,2} - 2\,\bar{T}_{total}^{\,2} + \frac{1}{4}\sum_{i=1}^{4}[Var_i + \bar{T}_i^2\,], \tag{18}$$

and the standard deviation of the total distribution of T is finally $\sigma_{T\,total} = \sqrt{Var_{total}}$, i.e.

$$\sigma_{T\,total} = \sqrt{\frac{1}{4}\sum_{i=1}^{4}[\sigma_i^2 + \bar{T}_i^2\,] - \bar{T}_{total}^{\,2}}\;. \tag{19}$$

Using the results above for the 4 populations, i.e., $\bar{T}_1 = 28.1$yr, $\bar{T}_2 = 32.0$ yr, $\bar{T}_3 = 33.5$ yr, $\bar{T}_4 = 32.8$ yr, $\sigma_1 = 9.5$ yr, $\sigma_2 = 0.53$ yr, $\sigma_3 = 7.4$ yr and $\sigma_4 = 8.6$ yr, we find that the standard deviation of T is

$$\sigma_{T\,total} = 7.7 \text{ yr.} \tag{20}$$

This implies that 95% of the area below the total distribution (full line in Fig. S2) is comprised between $\bar{T}_{total} - 1.96\,\sigma_{total}$ and $\bar{T}_{total} - 1.96\,\sigma_{total}$ [1], i.e., between 17 yr and 47 yr, which agrees well with Fig. S2.

So far we have found the mean and the standard deviation of the total distribution of the generation time T. In order to find out the error of the mean, we need to consider the probability distribution function of the mean $\bar{T}$. According to Eq. (19), the standard deviation of the distribution of the mean $\bar{T}$ is

$$\tilde{\sigma}_{T\,total} = \sqrt{\frac{1}{4}\sum_{i=1}^{4}[\sigma_{mi}^2 + \bar{T}_i^2\,] - \bar{T}_{total}^{\,2}}\;, \tag{21}$$

where the standard deviation of each mean is given by Eq. (3) as

$$\sigma_{mi} = \frac{\sigma_i}{\sqrt{n_i}} \tag{22}$$

and using the number of observed values for each population above, i.e., $n_1 = 258$, $n_2 = 96$, $n_3 = 2{,}000$ and $n_4 = 474$, we find that

$$\tilde{\sigma}_{T\,total} = 2.09 \text{ yr}\,. \tag{23}$$

As explained at the beginning of this subsection, if $n \geq 30$ it would be justified to apply Eqs. (3)-(4), i.e.

$$\Delta\bar{T}_{total} = z_{0.975} \, \frac{\tilde{\sigma}_{T\,total}}{\sqrt{n}}. \tag{24}$$

However, in our case $n = 4$ because we have estimated the mean in Eq. (14) by using 4 values[4]. Since $n < 30$, Eq. (24) should be replaced by the more precise formula [1]

$$\Delta\bar{T}_{total} = t_{0.975}^{n-1} \, \frac{\tilde{\sigma}_{T\,total}}{\sqrt{n}}, \tag{25}$$

where $t_{0.975}^{n-1}$ is the value of t for which Student's t distribution with $n-1$ degrees of freedom has probability 95% (i.e., the area under the corresponding curve is 0.95)[5]. In our case $n = 4$, therefore $t_{0.975}^{n-1} = t_{0.975}^{3} = 3.18$ [1] and we find using also Eq. (23) that

$$\Delta\bar{T}_{total} = 3.3 \text{ yr.} \tag{26}$$

Thus finally the 95% CL for the mean of the total distribution is

$$(\bar{T}_{total} - \Delta\bar{T}_{total}, \bar{T}_{total} + \Delta\bar{T}_{total}) = (28 - 35) \text{ yr.} \tag{27}$$

This is the range that we have used to obtain Figs. 3-5 in the main paper. It seems reasonable intuitively because it is very similar to the interval implied by the minimum and maximum values of the ranges for the 4 means (5), (6), (11) and (12), namely 27-34 yr. However the latter interval is just an intuitive guess. For example, it does not take into account that three of the four distributions have larger overall values of T than the other one (Fig. S2). Therefore, we prefer to use the interval (27) that has been obtained according to statistical theory.

Note that if we disregard the uncertainty within each population, Eq. (21) reads

$$\tilde{\sigma}_{T\,total} = \sqrt{\frac{1}{4}\sum_{i=1}^{4}[\sigma_{mi}^2 + \bar{T}_i^2] - \bar{T}_{total}^2} \approx \sqrt{\frac{1}{4}\sum_{i=1}^{4}\bar{T}_i^2 - \bar{T}_{total}^2} = 2.09 \text{ yr} \tag{28}$$

and this is (almost) the same as the more precise result (23)[6]. This means that the important contribution to the total uncertainty comes from the differences between the values of $\bar{T}_i$, not from the uncertainties in each of the 4 values of $\bar{T}_i$. This is not surprising if we note that the latter uncertainties ($\Delta\bar{T}_1 = 1.2$ yr, $\Delta\bar{T}_2 = 0.1$ yr, $\Delta\bar{T}_3 = 0.3$ yr and $\Delta\bar{T}_4 = 0.8$ yr) are much smaller than some differences between the values of $\bar{T}_i$ in our populations (e.g., $\bar{T}_3 - \bar{T}_1 \approx 5$ yr). This is why considering only one population (e.g., $\Delta\bar{T}_1 = 1.2$ yr) would seriously underestimate the total error (26). Thus our statistical

---

[4] Each mean is given by Eq. (14) for 4 populations drawn from a large set of different human populations. This is analogous to the fact that for Eq. (1) each value of T_i is drawn from a large set of values of T.

[5] Equation (24) can be applied if the distribution from which the original sample is drawn has an approximately Gaussian distribution. For $n < 30$ this is difficult to check and Eq. (25) is used instead of (4), which leads to larger errors due to this uncertainty in the shape of the original distribution. For large values of n $t_{0.975}^{n-1}$ tends to $z_{0.975} = 1.96$, so Eqs. (25) and (24) are the same.

[6] The result (28) can be also found from Eq. (2), i.e., $\sigma = \frac{\sqrt{\sum_{i=1}^{n}(\bar{T}_i - \bar{T}_{total})^2}}{\sqrt{n}} = \frac{\sqrt{\sum_{i=1}^{n}(\bar{T}_i - \bar{T}_{total})^2}}{\sqrt{4}} = 2.26$ yr. Here we have used $\sqrt{n}$ rather than $\sqrt{n-1}$ in the denominator because usually this correction is not applied for continuous distributions, and this is why the first equality in Eq. (17) neglecting the uncertainty within each population ($Var_i \approx 0$) is the same as Eq. (2) with n instead of $n-1$, i.e. 4 instead of 3. In fact using Eq. (2) without replacing $\sqrt{n-1}$ by $\sqrt{n}$ would lead to essentially the same result (28-35 yr).

analysis also leads us to the important conclusion that combining data from several populations is of utmost importance.

S1b. Initial growth rate a

Similarly to what we have done in the previous subsection for the generation time T, here we consider all populations for which we know data useful to perform statistically-sound estimations of the initial growth rate a, and then combine their results into an overall range. It is important to keep in mind that we can estimate the value of a only if we have data for small population numbers (compared to the saturation density), so that the population grows exponentially. The reason is that in general, biological populations reproduce according to the logistic equation,

$$\frac{\partial N}{\partial t} = a N \left[1 - \frac{N}{N_{max}}\right], \tag{29}$$

where N is the population number, t the time and N_{max} the carrying capacity. If $N \ll N_{max}$ then

$$\frac{\partial N}{\partial t} \approx a N \tag{30}$$

so the population grows exponentially,

$$N \approx N_0 e^{at}, \tag{31}$$

where N_0 is the initial population number. Thus if we have data of the population number (number of individuals) versus time and they are consistent with an exponential increase of the population, by fitting Eq. (31) to the data we can estimate the initial growth rate a. We have found useful data for the following 5 populations.

Population A. Pitcairn

Pitcairn is a Polynesian island situated about halfway between Australia and Chile. Birdsell [7] reported that the Pitcairn colony started in 1790 with 6 English men and 8 Polynesian women, and that cultivation was based on Polynesian tubers and fruits. The island remained essentially isolated, receiving very few immigrants until 1856. In Fig. S3 we reproduce the data in Ref. [7] and the exponential fit. It has exponent (initial growth rate) $a = 0.03744$ yr^{-1} and its standard deviation is $\sigma_a = 0.00149$ yr^{-1}.

The 95% CL interval for the initial growth rate is [1]

$$\Delta a = t_{0.975}^{n-2} \frac{\sigma_a}{\sqrt{n}} , \tag{32}$$

where $t_{0.975}^{n-2}$ is the value of t for which Student's t distribution with $n-2$ degrees of freedom has probability 95% (i.e., area below the curve equal to 0.95). Thus with 95% CL, the range of a is $(a - \Delta a, a + \Delta a)$.
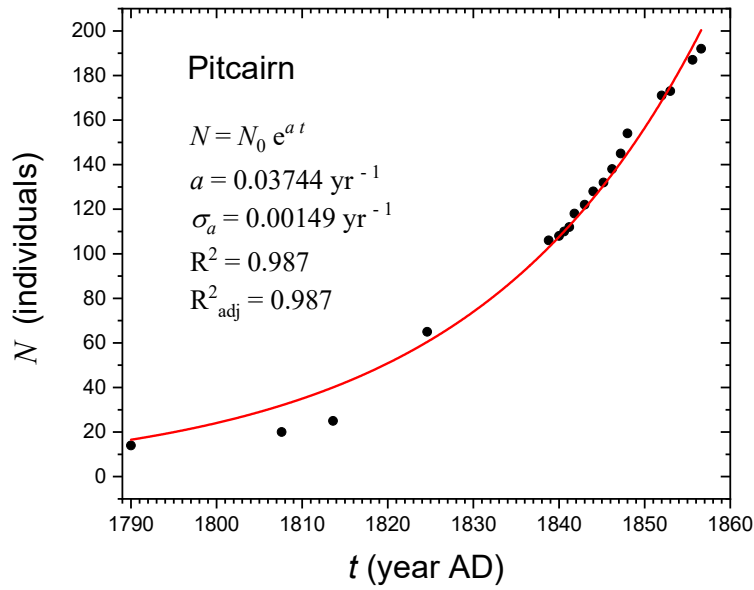
Fig. S3. Population number versus time and exponential fit for population A (Pitcairn island).

For the population of Pitcairn island the number of data points (Fig. S3) is $n = 19$, so $t_{0.975}^{n-2} = t_{0.975}^{17} = 2.11$ [1] and using the values of a and σ_a above into Eq. (32) we find that $a = (0.03744 \pm 0.00211)$ yr^{-1}, so the 95% CL interval for the initial growth rate is

$$(a - \Delta a, a + \Delta a) = (0.035 - 0.039) \text{ yr}^{-1}. \tag{33}$$

Population B. Bass Strait islands

These islands are located between Australia and Tasmania. Birdsell [7] reported population data beginning in year 1820 with 16 people, and mentioned that the economy was based on hunting, seafood, cultivation and domestic animals. In Fig. S4 we reproduce the data ($n = 19$) from Ref. [7] and the exponential fit. It has exponent (initial growth rate) $a = 0.02626$ yr^{-1} and its standard deviation is $\sigma_a = 0.00052$ yr^{-1}. Using these values in Eq. (32) yields $a = (0.02626 \pm 0.00110)$ yr^{-1}, so the 95% CL range for the initial growth rate is

$$(a - \Delta a, a + \Delta a) = (0.025 - 0.027) \text{ yr}^{-1}. \tag{34}$$

Fig. S4. Population number versus time and exponential fit for population B (Bass Strait islands).

Population C. United States of America

Lotka [8] published population data of the U.S.A. since 1790 when there were 3.29 millions of individuals until 1910 when the population had grown to 91.97 million. We include those data in Fig. S5 ($n = 13$), together with our exponential fit. It has exponent (initial growth rate) $a = 0.02326$ yr^{-1} and its standard deviation is $\sigma_a = 0.00071$ yr^{-1}. Using these values and $t_{0.975}^{n-2} = t_{0.975}^{11} = 2.20$ [1] into Eq. (32) yields $a = (0.02326 \pm 0.00156)$ yr^{-1}, so the 95% CL range for the initial growth rate is

$$(a - \Delta a, a + \Delta a) = (0.022 - 0.025) \text{ yr}^{-1}. \tag{35}$$

Fig. S5. Population number versus time and exponential fit for population B (United States of America).

Population D. Tristan da Cunha

This remote group of 5 small volcanic islands is located in the middle of the Atlantic Ocean (at less latitude than the tip of southern Africa) but only the largest island has been inhabited by humans, with a potato crop, cattle and fish as staple foods [7]. Roberts published population number data as a function of time [9]. Two brief periods of growth ended abruptly, and a third one begun in 1891 with 59 individuals. Its population numbers are reproduced in Fig. S6 ($n = 38$), together with the exponential fit. It has exponent $a = 0.02579$ yr^{-1} and its standard deviation is $\sigma_a = 0.00033$ yr^{-1}. Using Eq. (32) with these values and $t_{0.975}^{n-2} = t_{0.975}^{36} = 2.03$ [1] yields $a = (0.02579 \pm 0.00067)$ yr^{-1}, so the 95% CL range for the initial growth rate is

$$(a - \Delta a, a + \Delta a) = (0.025 - 0.026) \text{ yr}^{-1}. \tag{36}$$
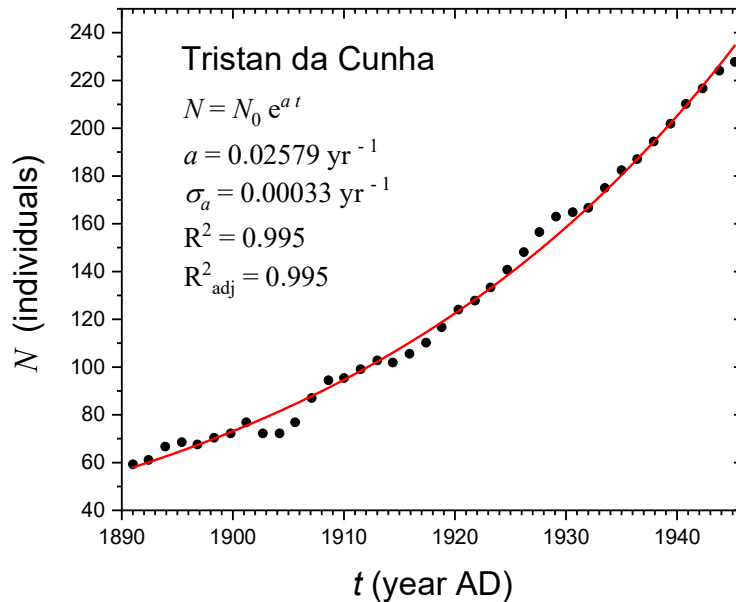
Fig. S6. Population number versus time and exponential fit for population D (Tristan da Cunha).

Population E. St. Lawrence valley

This region in present-day Canada had no sedentary population and was used for fishing and hunting until the seventeenth century, when European farmers arrived and gradually established new parishes along the banks of the St. Lawrence river. Charbonneau et al. reported population numbers since 1652 until 1803 [10]. Those data are shown Fig. S7 ($n = 31$), together with our exponential fit. It has exponent $a = 0.02433$ yr^{-1} and its standard deviation is $\sigma_a = 0.00022$ yr^{-1}. Using Eq. (32) with these values and $t_{0.975}^{n-2} =$

$t^{29}_{0.975} = 2.04$ [1] yields $a = (0.02433 \pm 0.00044)$ yr^{-1}, so the 95% CL range for the initial growth rate is

$$(a - \Delta a, a + \Delta a) = (0.024 - 0.025) \text{ yr}^{-1}. \tag{37}$$

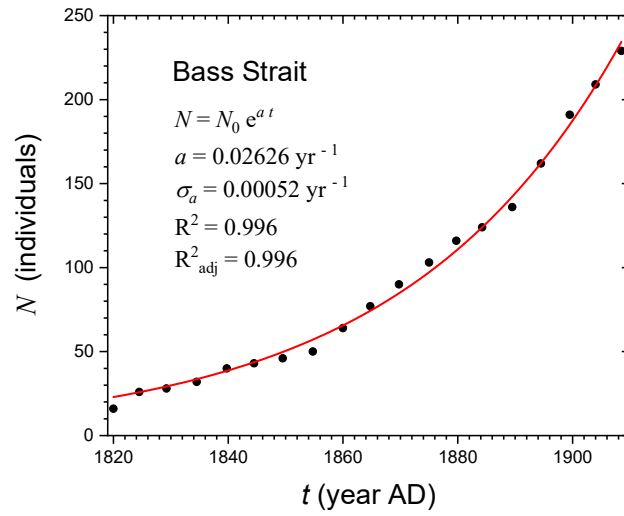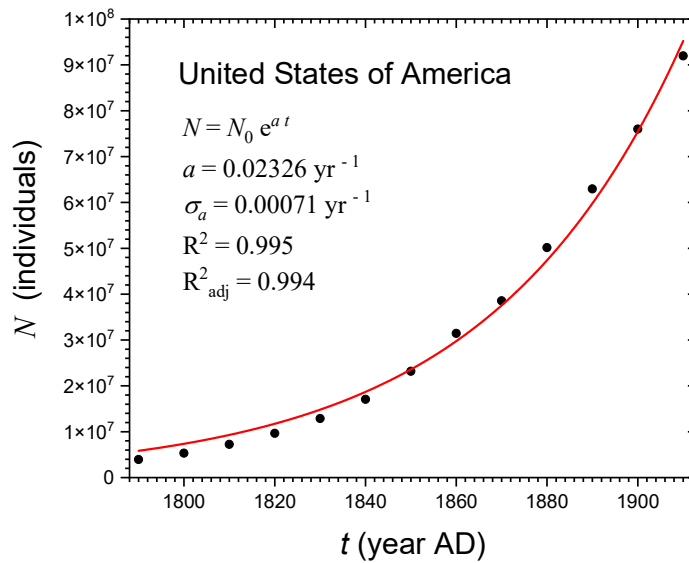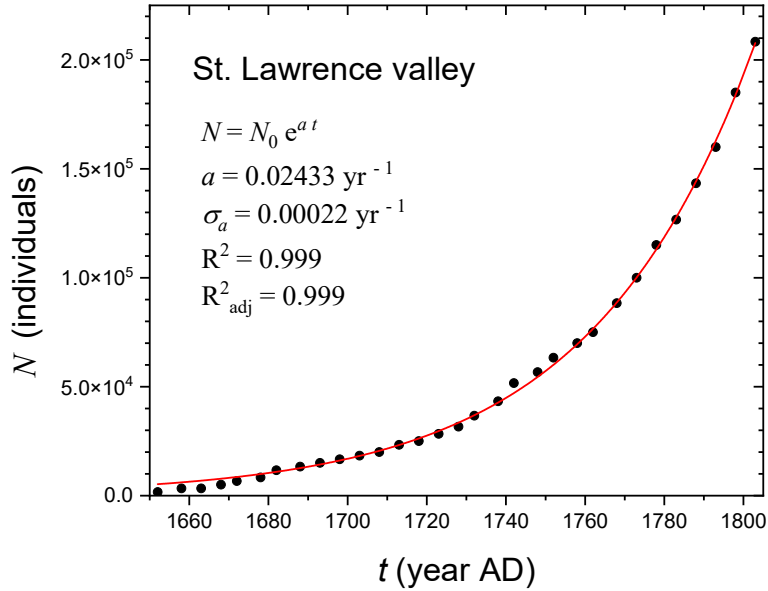**Fig. S7.** Population number versus time and exponential fit for population E (St. Lawrence valley).

Overall range

Let us combine the results for the 5 populations above to derive an overall range for the initial growth rate a. This is necessary to calculate front speeds (Figs. 2-4 in the main paper). Analogously to Eq. (13), the total probability distribution function of a is the sum of the 5 probability distribution functions considered above. Analogously to Eq. (14), the average of the total distribution is

$$\bar{a}_{total} = \frac{a_A + a_B + a_C + a_D + a_E}{5} = 0.02742 \text{ yr}^{-1}, \tag{38}$$

where we have used the results above for the 5 populations, namely $a_A = 0.03744$ yr^{-1}, $a_B = 0.02626$ yr^{-1}, $a_C = 0.02326$ yr^{-1}, $a_D = 0.02579$ yr^{-1} and $a_E = 0.02433$ yr^{-1}. The standard deviation is, analogously to Eq. (19),

$$\sigma_{a\ total} = \sqrt{\frac{\sigma_A^2 + a_A^2 + \sigma_B^2 + a_B^2 + \sigma_C^2 + a_C^2 + \sigma_D^2 + a_D^2 + \sigma_E^2 + a_E^2}{5} - \bar{a}^2} = 0.00519 \text{ yr}^{-1}, \tag{39}$$

where we have used the results for the standard deviations above, namely $\sigma_A = 0.00149$ yr^{-1}, $\sigma_B = 0.00052$ yr^{-1}, $\sigma_C = 0.00071$ yr^{-1}, $\sigma_D = 0.00033$ yr^{-1} and $\sigma_E = 0.00022$ yr^{-1}. Analogously to Eq. (25), the error in the growth rate is

$$\Delta \bar{a}_{total} = t_{0.975}^{n-1} \frac{\sigma_{a\,total}}{\sqrt{n}} = 0.00645 \text{ yr}^{-1}, \tag{40}$$

where we have used that $n = 5$ and $t_{0.975}^{n-1} = t_{0.975}^{4} = 2.78$. Thus finally the 95% CL for the mean of the total distribution is

$$(\bar{a}_{total} - \Delta \bar{a}_{total}, \bar{a}_{total} + \Delta \bar{a}_{total}) = (0.021 - 0.034) \text{ yr}^{-1}. \tag{41}$$

This is the range that we have used to obtain Figs. 3-5 in the main paper. It seems reasonable intuitively if we take into account that the minimum and maximum values of the ranges for the growth rates of the 5 populations (33)-(37) are 0.022 yr^{-1} and 0.039 yr^{-1}, respectively, but the latter corresponds to population A, which has a higher growth rate than the other four[7].

Finally, if we disregard the uncertainty within each population Eq. (39) reads

$$\sigma_{a\,total} \approx \sqrt{\frac{a_A^2 + a_B^2 + a_C^2 + a_D^2 + a_E^2}{5} - \bar{a}^2} = 0.00512 \text{ yr}^{-1}, \tag{42}$$

and Eq. (40) leads to $(0.021 - 0.034)$ yr^{-1}, i.e., the same as the more precise result (39). This means that, also for the growth rate, the important contribution to the total uncertainty comes from the differences between the values of a_i, not from the uncertainties in each of the 5 values of a_i. This is not surprising if we note that the latter uncertainties (e.g., $\Delta a_A \approx 0.002$ yr^{-1}) are much smaller than some differences between the values of a_i in our populations (e.g., $a_A - a_B \approx 0.011$ yr^{-1}). This is why considering only one population (e.g., $\Delta a_A = 0.00211$ yr^{-1}) would underestimate the total error (40). Thus, both for the growth rate (this subsection) and for the generation time (previous subsection), our statistical analysis leads us to the important conclusion that combining data from several populations is of utmost importance.

---

[7] For the overall range of the generation time T, Eq. (27), this effect is less pronounced because we have 3 populations with ranges rather different than another one (population 1). In contrast, for initial growth rate a we have 4 populations with ranges rather different than another one (population A) and this drives the overall range (41) further away from that of population A.

S2. Examples of application to other contexts

Section 3.2 in the main paper considers contexts with observed spread rates different from that of the Neolithic in Europe. Here we include some figures that are quoted in that section and necessary to understand how some results in table 3 in the main paper have been obtained.
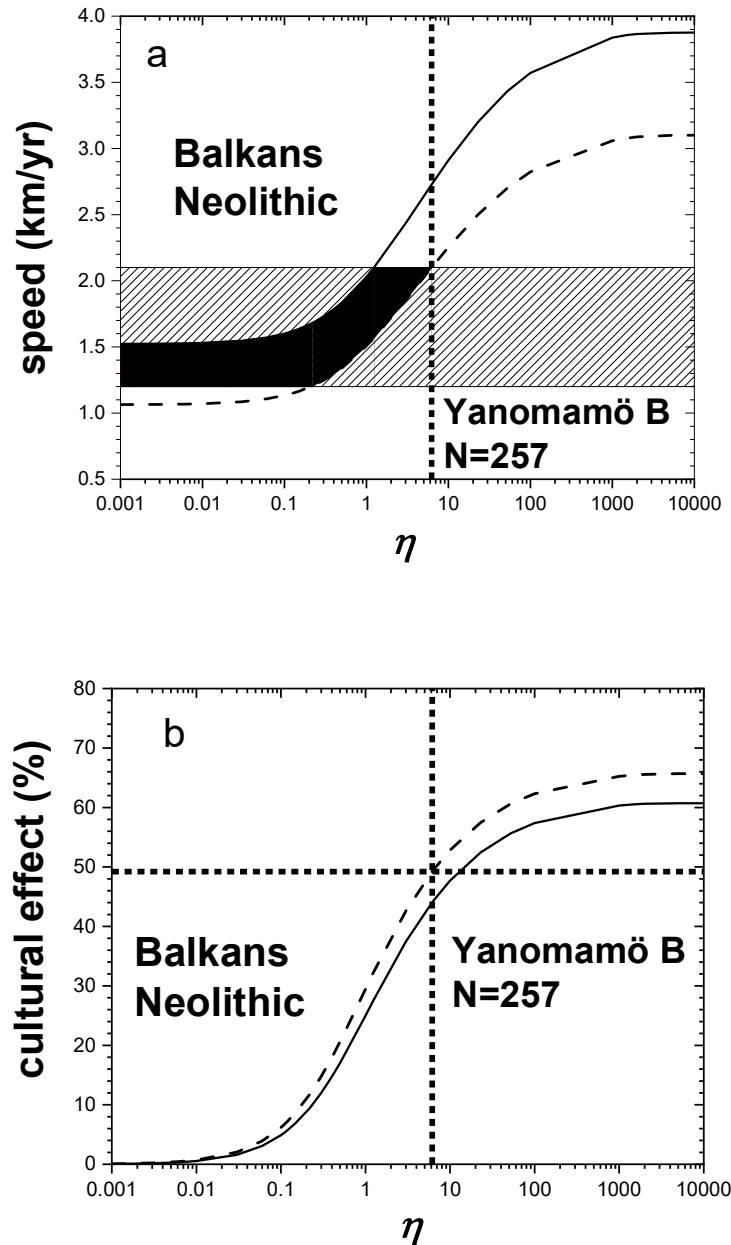
Fig. S8. The curves are the same as in Fig. 3a in the main paper because the same dispersal kernel (Yanomamö B) has been applied. However the context here is the Neolithic in the Balkans, so the hatched rectangle (observed spread rate) is different (1.2-2.1 km/yr here versus 0.9-1.3 in Fig. 3a). Thus the shape of the consistency region (black area) is different from that in Fig. 3a. But, again, consistency is attained for values of η between $\eta_{min} = 0$ and η_{max} given by the intersection of the upper side of the hatched rectangle and the dashed curve. This yields $\eta_{max} = 6.23$ from panel a and a maximum cultural effect of 49% from panel b, as reported in table 3 in the main paper, row 'Yanomamö B'.
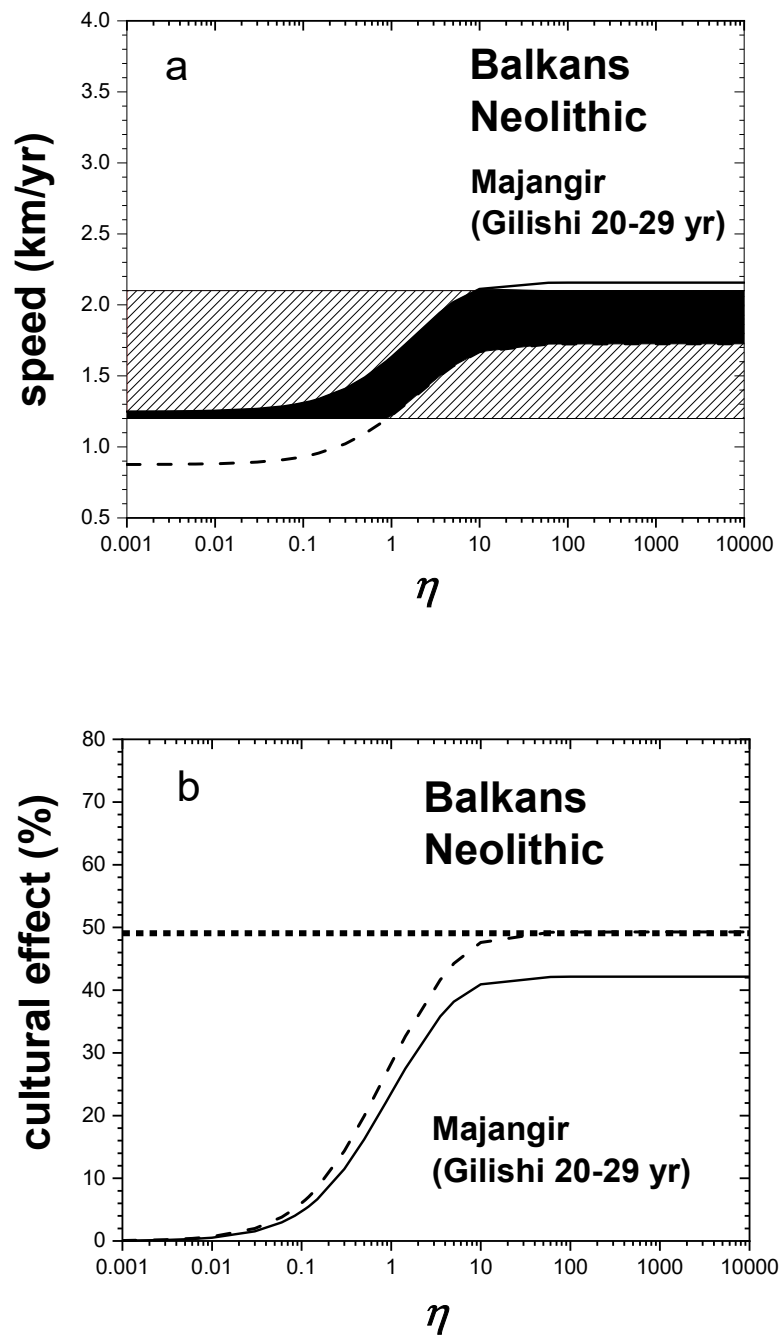
Fig. S9. The context (hatched rectangle) in panel a is the same as in Fig. S8 (the Neolithic in the Balkans, 1.2-2.1 km/yr) but the dispersal kernel used to calculate the curves here is that of the Majangir (Gilishi 20-29 yr). For this kernel the speeds are slower than in Fig. S8 and for this reason, the minimum calculated speed (dashed curve) does not cross the upper limit of the observed speed (2.1 km/yr) so we cannot find any upper bound for η, i.e., $\eta_{max} = \infty$ and panel b yields a maximum cultural effect of 49%, as reported in table 3 in the main paper, row 'Majangir R (Gilishi 20-29 yr).
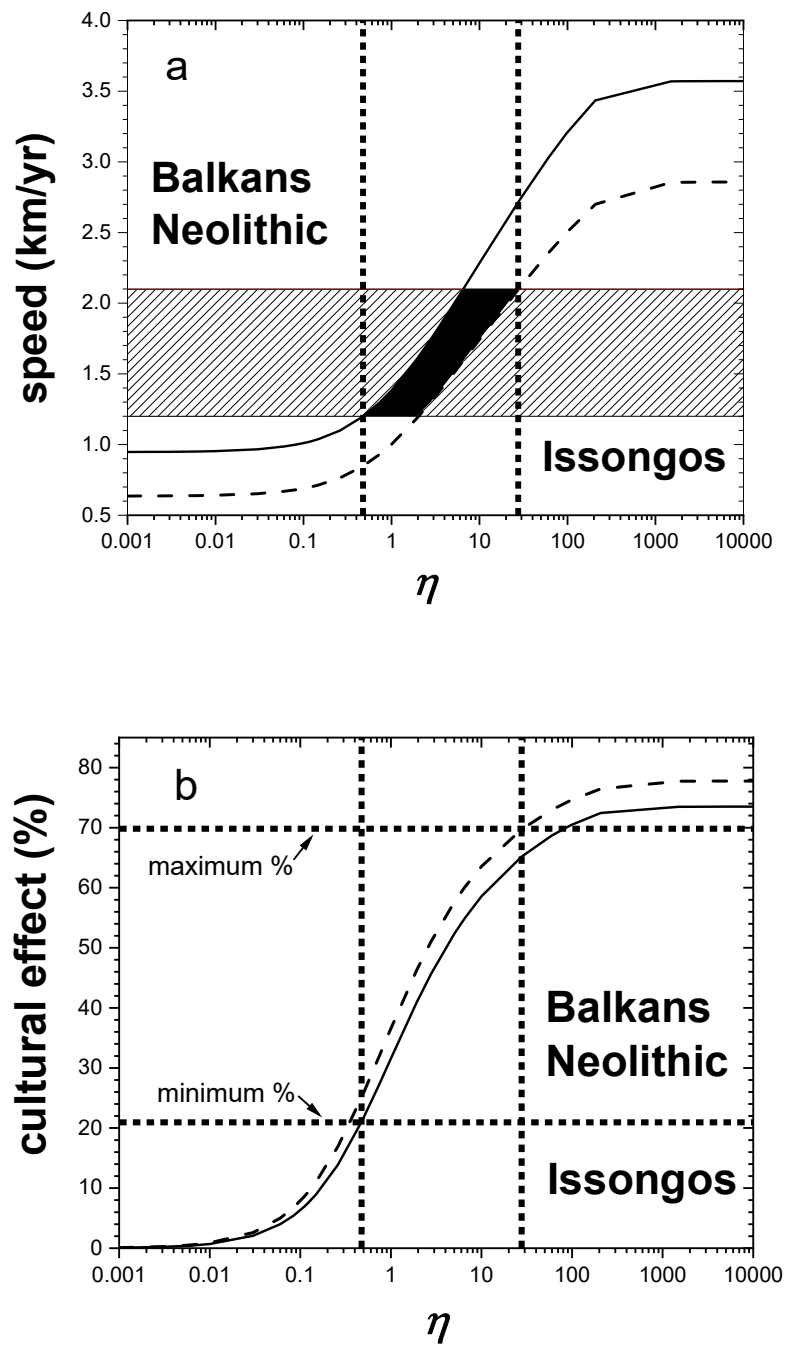
Fig. S10. The context (hatched rectangle) in panel a is the same as in Figs. S8-S9 (the Neolithic in the Balkans, 1.2-2.1 km/yr) but the dispersal kernel used to calculate the curves here is that of the Issongos. For this kernel the speeds are so slow for $\eta = 0$ that they do not cross the observed range (hatched rectangle). For this reason, the minimum value of η is different from zero, namely $\eta_{min} = 0.48$. However, in contrast with Fig. S9, the minimum speed (dashed curve in panel a here) crosses the upper limit for the observed speed (upper side of the hatched rectangle), so there is also a finite maximum value of η, namely $\eta_{max} = 27.79$. Panel b yields the minimum and maximum cultural effects, namely 21% and 70%. These four results are reported in table 3 in the main paper, row 'Issongos M'.

S3. Computer code used to generate the figures

This code runs in program *mathematica* and has been used to generate Fig. 3 (kernel B) in the main paper.

Figures 4 (kernel R) and 5 (kernel M) have been obtained with the same code but using:

· the corresponding number of distances, namely 159 for Fig. 4 (kernel R) and 97 for Fig. 5 (kernel M), instead of 257 for Fig. 3 (kernel B),

· the list of distances (in km) from table S2 (kernel R) for Fig. 4 and from table S3 (kernel M) for Fig. 5, instead of the list of distances from table S1 (kernel B) for Fig. 3.

The same procedure has been used to obtain Fig. 2 in the main paper.

This is the function to be minimized in Equation (3) in the main
    paper (here we use aN rather than *a* for the initial growth rate of farmers) :

In[ ]:= **f =**

aN / lambda + Log[(1 + eta) * (1 / 257) * (BesselI[0, lambda * r1] + BesselI[0, lambda * r2] +
    ⌊logaritmo   ⌊I de Bessel   ⌊I de Bessel

BesselI[0, lambda * r3] + BesselI[0, lambda * r4] + BesselI[0, lambda * r5] +
⌊I de Bessel   ⌊I de Bessel   ⌊I de Bessel

BesselI[0, lambda * r6] + BesselI[0, lambda * r7] + BesselI[0, lambda * r8] +
⌊I de Bessel   ⌊I de Bessel   ⌊I de Bessel

BesselI[0, lambda * r9] + BesselI[0, lambda * r10] + BesselI[0, lambda * r11] +
⌊I de Bessel   ⌊I de Bessel   ⌊I de Bessel

BesselI[0, lambda * r12] + BesselI[0, lambda * r13] + BesselI[0, lambda * r14] +
⌊I de Bessel   ⌊I de Bessel   ⌊I de Bessel

BesselI[0, lambda * r15] + BesselI[0, lambda * r16] + BesselI[0, lambda * r17] +
⌊I de Bessel   ⌊I de Bessel   ⌊I de Bessel

BesselI[0, lambda * r18] + BesselI[0, lambda * r19] + BesselI[0, lambda * r20] +
⌊I de Bessel   ⌊I de Bessel   ⌊I de Bessel

BesselI[0, lambda * r21] + BesselI[0, lambda * r22] + BesselI[0, lambda * r23] +
⌊I de Bessel   ⌊I de Bessel   ⌊I de Bessel

BesselI[0, lambda * r24] + BesselI[0, lambda * r25] + BesselI[0, lambda * r26] +
⌊I de Bessel   ⌊I de Bessel   ⌊I de Bessel

BesselI[0, lambda * r27] + BesselI[0, lambda * r28] + BesselI[0, lambda * r29] +
⌊I de Bessel   ⌊I de Bessel   ⌊I de Bessel

BesselI[0, lambda * r30] + BesselI[0, lambda * r31] + BesselI[0, lambda * r32] +
⌊I de Bessel   ⌊I de Bessel   ⌊I de Bessel

BesselI[0, lambda * r33] + BesselI[0, lambda * r34] + BesselI[0, lambda * r35] +
⌊I de Bessel   ⌊I de Bessel   ⌊I de Bessel

BesselI[0, lambda * r36] + BesselI[0, lambda * r37] + BesselI[0, lambda * r38] +
⌊I de Bessel   ⌊I de Bessel   ⌊I de Bessel

BesselI[0, lambda * r39] + BesselI[0, lambda * r40] + BesselI[0, lambda * r41] +
⌊I de Bessel   ⌊I de Bessel   ⌊I de Bessel

BesselI[0, lambda * r42] + BesselI[0, lambda * r43] + BesselI[0, lambda * r44] +
⌊I de Bessel   ⌊I de Bessel   ⌊I de Bessel

BesselI[0, lambda * r45] + BesselI[0, lambda * r46] + BesselI[0, lambda * r47] +
⌊I de Bessel   ⌊I de Bessel   ⌊I de Bessel

BesselI[0, lambda * r48] + BesselI[0, lambda * r49] + BesselI[0, lambda * r50] +
⌊I de Bessel   ⌊I de Bessel   ⌊I de Bessel

BesselI[0, lambda * r51] + BesselI[0, lambda * r52] + BesselI[0, lambda * r53] +
⌊I de Bessel   ⌊I de Bessel   ⌊I de Bessel

BesselI[0, lambda * r54] + BesselI[0, lambda * r55] + BesselI[0, lambda * r56] +
⌊I de Bessel   ⌊I de Bessel   ⌊I de Bessel

BesselI[0, lambda * r57] + BesselI[0, lambda * r58] + BesselI[0, lambda * r59] +
⌊I de Bessel   ⌊I de Bessel   ⌊I de Bessel

BesselI[0, lambda * r60] + BesselI[0, lambda * r61] + BesselI[0, lambda * r62] +
⌊I de Bessel   ⌊I de Bessel   ⌊I de Bessel

BesselI[0, lambda * r63] + BesselI[0, lambda * r64] + BesselI[0, lambda * r65] +
⌊I de Bessel   ⌊I de Bessel   ⌊I de Bessel

BesselI[0, lambda * r66] + BesselI[0, lambda * r67] + BesselI[0, lambda * r68] +
⌊I de Bessel   ⌊I de Bessel   ⌊I de Bessel

BesselI[0, lambda * r69] + BesselI[0, lambda * r70] + BesselI[0, lambda * r71] +
⌊I de Bessel   ⌊I de Bessel   ⌊I de Bessel

BesselI[0, lambda * r72] + BesselI[0, lambda * r73] + BesselI[0, lambda * r74] +
⌊I de Bessel   ⌊I de Bessel   ⌊I de Bessel

```
BesselI[0, lambda * r75] + BesselI[0, lambda * r76] + BesselI[0, lambda * r77] +
BesselI[0, lambda * r78] + BesselI[0, lambda * r79] + BesselI[0, lambda * r80] +
BesselI[0, lambda * r81] + BesselI[0, lambda * r82] + BesselI[0, lambda * r83] +
BesselI[0, lambda * r84] + BesselI[0, lambda * r85] + BesselI[0, lambda * r86] +
BesselI[0, lambda * r87] + BesselI[0, lambda * r88] + BesselI[0, lambda * r89] +
BesselI[0, lambda * r90] + BesselI[0, lambda * r91] + BesselI[0, lambda * r92] +
BesselI[0, lambda * r93] + BesselI[0, lambda * r94] + BesselI[0, lambda * r95] +
BesselI[0, lambda * r96] + BesselI[0, lambda * r97] + BesselI[0, lambda * r98] +
BesselI[0, lambda * r99] + BesselI[0, lambda * r100] + BesselI[0, lambda * r101] +
BesselI[0, lambda * r102] + BesselI[0, lambda * r103] + BesselI[0, lambda * r104] +
BesselI[0, lambda * r105] + BesselI[0, lambda * r106] + BesselI[0, lambda * r107] +
BesselI[0, lambda * r108] + BesselI[0, lambda * r109] + BesselI[0, lambda * r110] +
BesselI[0, lambda * r111] + BesselI[0, lambda * r112] + BesselI[0, lambda * r113] +
BesselI[0, lambda * r114] + BesselI[0, lambda * r115] + BesselI[0, lambda * r116] +
BesselI[0, lambda * r117] + BesselI[0, lambda * r118] + BesselI[0, lambda * r119] +
BesselI[0, lambda * r120] + BesselI[0, lambda * r121] + BesselI[0, lambda * r122] +
BesselI[0, lambda * r123] + BesselI[0, lambda * r124] + BesselI[0, lambda * r125] +
BesselI[0, lambda * r126] + BesselI[0, lambda * r127] + BesselI[0, lambda * r128] +
BesselI[0, lambda * r129] + BesselI[0, lambda * r130] + BesselI[0, lambda * r131] +
BesselI[0, lambda * r132] + BesselI[0, lambda * r133] + BesselI[0, lambda * r134] +
BesselI[0, lambda * r135] + BesselI[0, lambda * r136] + BesselI[0, lambda * r137] +
BesselI[0, lambda * r138] + BesselI[0, lambda * r139] + BesselI[0, lambda * r140] +
BesselI[0, lambda * r141] + BesselI[0, lambda * r142] + BesselI[0, lambda * r143] +
BesselI[0, lambda * r144] + BesselI[0, lambda * r145] + BesselI[0, lambda * r146] +
BesselI[0, lambda * r147] + BesselI[0, lambda * r148] + BesselI[0, lambda * r149] +
BesselI[0, lambda * r150] + BesselI[0, lambda * r151] + BesselI[0, lambda * r152] +
BesselI[0, lambda * r153] + BesselI[0, lambda * r154] + BesselI[0, lambda * r155] +
```

BesselI[0, lambda * r156] + BesselI[0, lambda * r157] + BesselI[0, lambda * r158] +

BesselI[0, lambda * r159] + BesselI[0, lambda * r160] + BesselI[0, lambda * r161] +

BesselI[0, lambda * r162] + BesselI[0, lambda * r163] + BesselI[0, lambda * r164] +

BesselI[0, lambda * r165] + BesselI[0, lambda * r166] + BesselI[0, lambda * r167] +

BesselI[0, lambda * r168] + BesselI[0, lambda * r169] + BesselI[0, lambda * r170] +

BesselI[0, lambda * r171] + BesselI[0, lambda * r172] + BesselI[0, lambda * r173] +

BesselI[0, lambda * r174] + BesselI[0, lambda * r175] + BesselI[0, lambda * r176] +

BesselI[0, lambda * r177] + BesselI[0, lambda * r178] + BesselI[0, lambda * r179] +

BesselI[0, lambda * r180] + BesselI[0, lambda * r181] + BesselI[0, lambda * r182] +

BesselI[0, lambda * r183] + BesselI[0, lambda * r184] + BesselI[0, lambda * r185] +

BesselI[0, lambda * r186] + BesselI[0, lambda * r187] + BesselI[0, lambda * r188] +

BesselI[0, lambda * r189] + BesselI[0, lambda * r190] + BesselI[0, lambda * r191] +

BesselI[0, lambda * r192] + BesselI[0, lambda * r193] + BesselI[0, lambda * r194] +

BesselI[0, lambda * r195] + BesselI[0, lambda * r196] + BesselI[0, lambda * r197] +

BesselI[0, lambda * r198] + BesselI[0, lambda * r199] + BesselI[0, lambda * r200] +

BesselI[0, lambda * r201] + BesselI[0, lambda * r202] + BesselI[0, lambda * r203] +

BesselI[0, lambda * r204] + BesselI[0, lambda * r205] + BesselI[0, lambda * r206] +

BesselI[0, lambda * r207] + BesselI[0, lambda * r208] + BesselI[0, lambda * r209] +

BesselI[0, lambda * r210] + BesselI[0, lambda * r211] + BesselI[0, lambda * r212] +

BesselI[0, lambda * r213] + BesselI[0, lambda * r214] + BesselI[0, lambda * r215] +

BesselI[0, lambda * r216] + BesselI[0, lambda * r217] + BesselI[0, lambda * r218] +

BesselI[0, lambda * r219] + BesselI[0, lambda * r220] + BesselI[0, lambda * r221] +

BesselI[0, lambda * r222] + BesselI[0, lambda * r223] + BesselI[0, lambda * r224] +

BesselI[0, lambda * r225] + BesselI[0, lambda * r226] + BesselI[0, lambda * r227] +

BesselI[0, lambda * r228] + BesselI[0, lambda * r229] + BesselI[0, lambda * r230] +

BesselI[0, lambda * r231] + BesselI[0, lambda * r232] +

BesselI[0, lambda * r233] + BesselI[0, lambda * r234] + BesselI[0, lambda * r235] +

```
BesselI[0, lambda * r236] + BesselI[0, lambda * r237] + BesselI[0, lambda * r238] +
   I de Bessel                    I de Bessel                    I de Bessel
BesselI[0, lambda * r239] + BesselI[0, lambda * r240] + BesselI[0, lambda * r241] +
   I de Bessel                    I de Bessel                    I de Bessel
BesselI[0, lambda * r242] + BesselI[0, lambda * r243] + BesselI[0, lambda * r244] +
   I de Bessel                    I de Bessel                    I de Bessel
BesselI[0, lambda * r245] + BesselI[0, lambda * r246] + BesselI[0, lambda * r247] +
   I de Bessel                    I de Bessel                    I de Bessel
BesselI[0, lambda * r248] + BesselI[0, lambda * r249] + BesselI[0, lambda * r250] +
   I de Bessel                    I de Bessel                    I de Bessel
BesselI[0, lambda * r251] + BesselI[0, lambda * r252] +
   I de Bessel                    I de Bessel
BesselI[0, lambda * r253] + BesselI[0, lambda * r254] + BesselI[0, lambda * r255] +
   I de Bessel                    I de Bessel                    I de Bessel
BesselI[0, lambda * r256] + BesselI[0, lambda * r257])] / (lambda * T)
   I de Bessel                    I de Bessel
```

Distances from table S2 = kernel B (km) :

In[◦]:=
```
r1 = 0;
r2 = 0;
r3 = 0;
r4 = 0;
r5 = 0;
r6 = 0;
r7 = 1.85;
r8 = 4.14;
r9 = 4.14;
r10 = 4.14;
r11 = 4.14;
r12 = 4.14;
r13 = 4.14;
r14 = 4.14;
r15 = 5.56;
r16 = 5.86;
r17 = 7.64;
r18 = 7.64;
r19 = 7.64;
r20 = 7.64;
r21 = 7.64;
r22 = 8.29;
r23 = 8.29;
r24 = 8.29;
r25 = 8.29;
r26 = 9.27;
r27 = 11.27;
r28 = 11.27;
r29 = 11.27;
r30 = 11.72;
r31 = 11.86;
```

```
r32 = 11.86;
r33 = 12.43;
r34 = 13.10;
r35 = 13.10;
r36 = 13.10;
r37 = 13.10;
r38 = 13.10;
r39 = 13.10;
r40 = 14.11;
r41 = 14.11;
r42 = 14.11;
r43 = 14.11;
r44 = 14.94;
r45 = 14.94;
r46 = 14.94;
r47 = 15.72;
r48 = 15.83;
r49 = 15.94;
r50 = 17.08;
r51 = 17.08;
r52 = 17.08;
r53 = 17.08;
r54 = 17.08;
r55 = 17.08;
r56 = 17.08;
r57 = 17.48;
r58 = 17.48;
r59 = 18.61;
r60 = 19.70;
r61 = 19.70;
r62 = 19.70;
r63 = 19.70;
r64 = 19.70;
r65 = 19.70;
r66 = 19.96;
r67 = 20.04;
r68 = 20.46;
r69 = 20.46;
r70 = 20.46;
r71 = 20.72;
r72 = 20.72;
r73 = 21.13;
r74 = 21.61;
r75 = 21.61;
r76 = 21.61;
r77 = 21.61;
r78 = 21.61;
r79 = 21.61;
r80 = 21.61;
r81 = 21.61;
```

```
r82 = 21.61;
r83 = 21.61;
r84 = 21.61;
r85 = 21.61;
r86 = 21.61;
r87 = 23.44;
r88 = 23.44;
r89 = 23.44;
r90 = 24.16;
r91 = 24.16;
r92 = 24.16;
r93 = 24.16;
r94 = 24.16;
r95 = 24.16;
r96 = 24.16;
r97 = 24.16;
r98 = 24.16;
r99 = 24.16;
r100 = 24.16;
r101 = 24.16;
r102 = 24.16;
r103 = 24.93;
r104 = 24.93;
r105 = 24.93;
r106 = 25.80;
r107 = 26.72;
r108 = 26.72;
r109 = 26.72;
r110 = 26.72;
r111 = 26.72;
r112 = 28.04;
r113 = 28.04;
r114 = 28.04;
r115 = 28.04;
r116 = 28.04;
r117 = 28.04;
r118 = 28.04;
r119 = 28.04;
r120 = 28.04;
r121 = 28.23;
r122 = 28.23;
r123 = 28.35;
r124 = 28.94;
r125 = 30.56;
r126 = 30.84;
r127 = 32.36;
r128 = 32.36;
r129 = 32.36;
r130 = 32.36;
r131 = 33.39;
```

```
r132 = 34.01;
r133 = 34.01;
r134 = 34.01;
r135 = 35.24;
r136 = 35.40;
r137 = 35.40;
r138 = 35.40;
r139 = 35.40;
r140 = 35.40;
r141 = 35.40;
r142 = 35.40;
r143 = 35.40;
r144 = 35.40;
r145 = 35.40;
r146 = 35.40;
r147 = 35.40;
r148 = 35.40;
r149 = 35.79;
r150 = 36.68;
r151 = 36.68;
r152 = 36.68;
r153 = 36.68;
r154 = 36.68;
r155 = 36.68;
r156 = 36.68;
r157 = 36.68;
r158 = 36.68;
r159 = 37.52;
r160 = 37.80;
r161 = 38.96;
r162 = 38.96;
r163 = 40.64;
r164 = 40.64;
r165 = 40.64;
r166 = 40.64;
r167 = 41.15;
r168 = 42.01;
r169 = 42.01;
r170 = 42.01;
r171 = 42.01;
r172 = 42.01;
r173 = 42.01;
r174 = 43.41;
r175 = 43.73;
r176 = 44.20;
r177 = 44.20;
r178 = 44.20;
r179 = 44.20;
r180 = 44.20;
r181 = 44.20;
```

```
r182 = 44.20;
r183 = 44.20;
r184 = 45.75;
r185 = 45.87;
r186 = 47.46;
r187 = 47.46;
r188 = 47.46;
r189 = 47.46;
r190 = 47.46;
r191 = 47.46;
r192 = 47.46;
r193 = 47.46;
r194 = 47.46;
r195 = 47.46;
r196 = 47.46;
r197 = 48.07;
r198 = 48.07;
r199 = 49.78;
r200 = 49.78;
r201 = 50.56;
r202 = 51.49;
r203 = 52.22;
r204 = 52.22;
r205 = 52.22;
r206 = 52.22;
r207 = 52.22;
r208 = 52.22;
r209 = 52.22;
r210 = 52.22;
r211 = 52.22;
r212 = 52.44;
r213 = 52.44;
r214 = 52.44;
r215 = 52.44;
r216 = 52.44;
r217 = 52.44;
r218 = 52.44;
r219 = 60.32;
r220 = 62.70;
r221 = 63.76;
r222 = 63.76;
r223 = 63.76;
r224 = 63.76;
r225 = 63.76;
r226 = 63.76;
r227 = 63.76;
r228 = 63.76;
r229 = 67.85;
r230 = 67.85;
r231 = 67.85;
```

```
r232 = 67.85;
r233 = 67.85;
r234 = 68.54;
r235 = 68.54;
r236 = 68.54;
r237 = 68.58;
r238 = 68.58;
r239 = 68.58;
r240 = 72.17;
r241 = 72.63;
r242 = 74.52;
r243 = 76.22;
r244 = 80.52;
r245 = 82.70;
r246 = 82.70;
r247 = 83.26;
r248 = 88.71;
r249 = 94.33;
r250 = 94.33;
r251 = 94.62;
r252 = 94.62;
r253 = 95.62;
r254 = 100.47;
r255 = 101.13;
r256 = 101.13;
r257 = 108.54
```
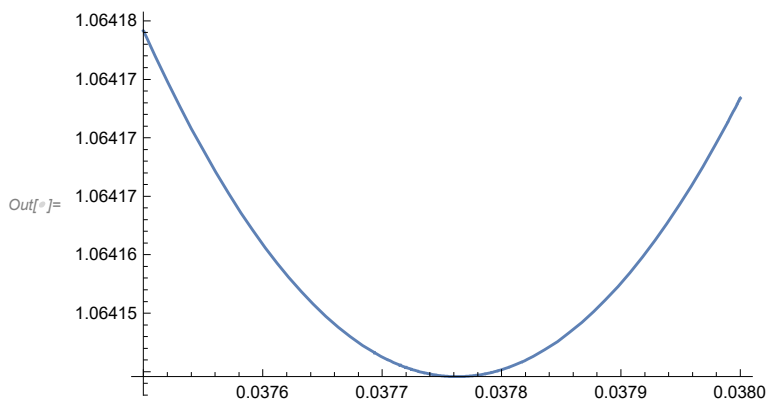

**Minimum speed (dashed curve in Fig. 3 in the main paper),
i.e., maximum value of T and minimum value
of aN from the ranges in Section S1, (27) and (41) :**

In[ ]:=   `T = 35; aN = 0.021`

In[ ]:=   `eta = 0`

In[◦]:= **Plot[f, {lambda, 0.0375, 0.038}]**

⌊representación gráfica

Out[◦]=
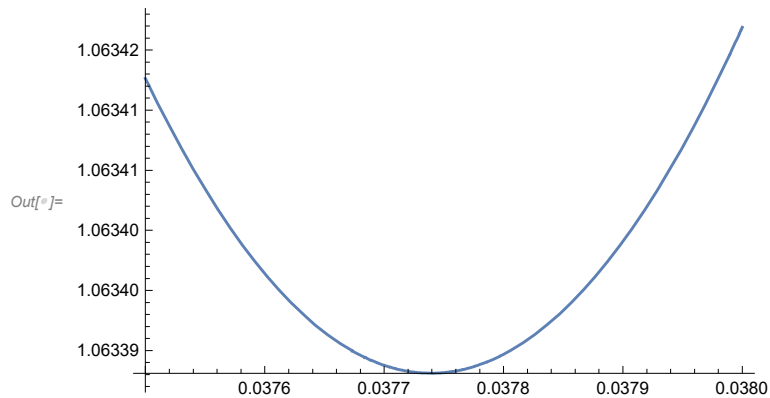
According to this plot, the minimum is :

<mark>**1.0634 km / yr**</mark>

This is the speed for eta = 0 (no cultural diffusion). It does not appear in Fig .3 a, but it is needed to compute the cultural
  effect (Fig. 3 b) for other values of eta, as done below.
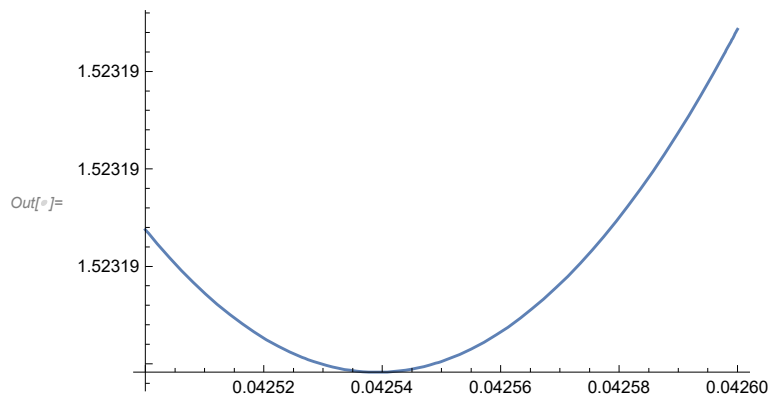

<mark>**Maximum speed (full curve in Fig. 3 in the main paper),**
 i.e., minimum value of T and maximum value
  of aN from the ranges in Section S1, (27) and (41) :</mark>

In[◦]:= <mark>**T = 28; aN = 0.034**</mark>

<mark>**eta = 0**</mark>

In[◦]:= **Plot[f, {lambda, 0.0424, 0.0426}]**

⌊representación gráfica

Out[◦]=
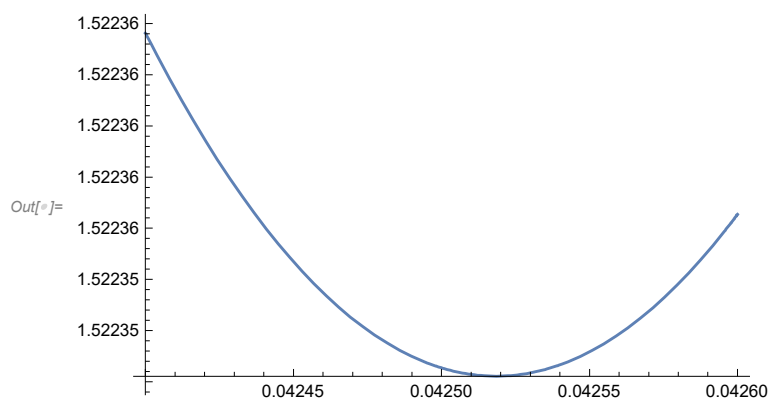
According to this plot, the minimum is :

<mark>**1.5224 km / yr**</mark>

This is the speed for eta = 0 (no cultural diffusion). It does not appear in Fig .3 a, but it is needed to compute the cultural
  effect (Fig. 3 b) for other values of eta, as done below.

In[◦]:= **T = 35; aN = 0.021**

In[◦]:= **eta = 0.001**

In[◦]:= **Plot[f, {lambda, 0.0375, 0.038}]**
representación gráfica

Out[◦]=

According to this plot, the minimum is :

1.0642 km / yr

This is the value in Fig. 3 a, dashed curve, for eta = 0.001.

In[◦]:= **T = 28; aN = 0.034**

In[◦]:= **eta = 0.001**

In[●]:= **Plot[f, {lambda, 0.0425, 0.0426}]**
⌊representación gráfica

Out[●]=

According to this plot, the minimum is :

1.5232 km / yr

This is the value in Fig. 3 a, full curve, for eta = 0.001.
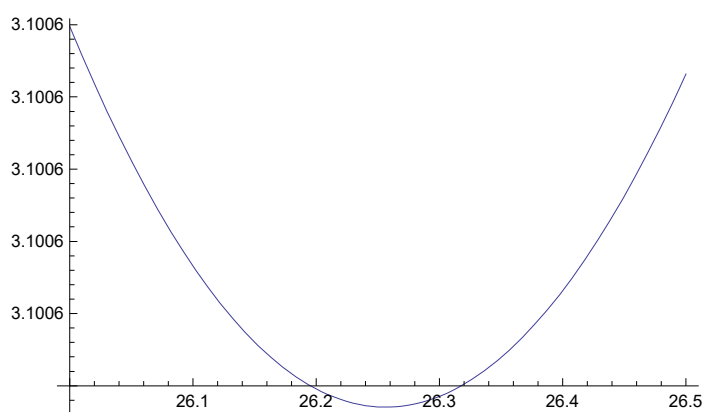

The speeds in Fig. 3 a for the remainining values of eta are computed in the same way, e.g. :

Minimum speed (dashed curve in Fig. 3 in the main paper),
i.e., maximum value of T and minimum value
  of aN from the ranges in Section S1, (27) and (41) :

In[●]:= **T = 35; aN = 0.021**

In[●]:= **eta = 10 000**

Plot[f, {lambda, 26, 26.5}]
⌊representación gráfica

According to this plot, the minimum is :

3.1006 km / yr

This is the value in Fig. 3 a , dashed curve, for eta = 10 000 .

The corresponding **cultural effect** (in %) is, according to Eq. (5) in the main paper :

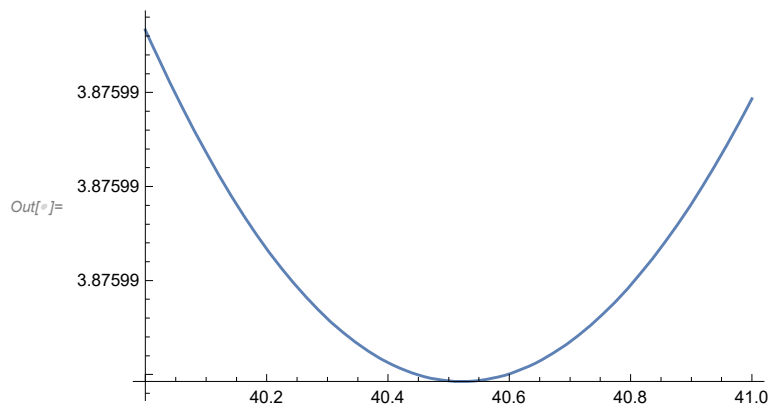*In[ ]:=* `100 * (3.1006 – 1.0634) / 3.1006`

Out[ ]:= `65.7034`

This is the value in Fig. 3 b , dashed curve, for eta = 10 000 .

Maximum speed (full curve in Fig. 3),
i.e., minimum value of T and maximum value
of aN from the ranges in Section S1, (27) and (41) :

In[ ]:= `T = 28; aN = 0.034`

In[ ]:= `eta = 10 000`

In[ ]:= `Plot[f, {lambda, 40, 41} ]`

representación gráfica

Out[ ]:=

According to this plot, the minimum is :

`3.8760 km / yr`

This is the value in Fig. 3 a , full curve, for eta = 10 000 .

CULTURAL EFFECT
(Fig. 3 b in the main paper) :

The corresponding **cultural effect** (in %) is, according to Eq. (5) in the main paper :

In[●]:= **100 * (3.8760 – 1.5224) / 3.8760**

Out[●]= 60.7224

This is the value in Fig. 3 b , full curve, for eta = 10 000 .

Supplementary references

[1] M. R. Spiegel, J. J. Schiller, R. A. Srinivasan and M. LeVan, Probability and statistics, New York: McGraw-Hill, 1975.

[2] Fort, J., Jana, D. & Humet, J., "Multidelayed random walks: theory and application to the neolithic transition in Europe," *Pys. Rev. E,* vol. 70, no. 031913, 2004.

[3] P. Biella, N. A. Chagnon and G. Seaman, Yanomamö interactive: the ax fight (CD-ROM), Orlando: Harcourt Brace & company, 1997.

[4] Stauder, J., The Majangir: ecology and society of a southwest Ethiopian people, Cambridge: Cambridge University Press, 1971.

[5] Cavalli-Sforza, L., Zei, G., "Experiments with an artificial population," in *Proceedings of the third congress of human genetics*, J. F. Crow and J. V. Neel, Eds., 1967, pp. 473-478.

[6] Viader, F., 127 genealogies, Girona: CCG Edicions, 2004.

[7] Birdsell, J. B., "Some population problems involving Pleistocene man," *Cold Spring Harbor Symp. Quant. Biol.,* vol. 22, pp. 47-69, 1957.

[8] Lotka, A. J., Elements of mathematical biology, New York: Dover, 1956.

[9] Roberts, D. F., "Genetic effects on population size reduction," *Nature,* vol. 220, pp. 1084-1088, 1968.

[10] Charbonneau, H., Desjardins, B., Légaré, J. & Denis, H., "The population of the St. Lawrence Valley, 1608-1760," in *A population history of North America*, Cambridge, Cambridge University Press, 2000, pp. 99-142.