Modeling the role of voyaging in the coastal spread of the Early Neolithic in the West Mediterranean

Neus Iserna a,1, João Zilhaob, c, Joaquim Forta, c, and Albert J. Ammerman d

*Complex Systems Laboratory and Physics Department, Universitat de Girona, 17003 Girona, Spain; bDepartament d’Història i Arqueologia (Grup de Recerca SGR2014-00108), University of Barcelona, 08001 Barcelona, Spain; cCatalan Institute for Research and Advanced Studies, 08010 Barcelona, Spain; and dDepartment of Classics, Colgate University, Hamilton, NY 13346

Edited by Melinda A. Zeder, National Museum of Natural History, Santa Fe, NM, and approved December 5, 2016 (received for review August 12, 2016)

The earliest dates for the West Mediterranean Neolithic indicate that it expanded across 2,500 km in about 300 y. Such a fast spread is held to be mainly due to a demic process driven by dispersal along coastal routes. Here, we model the Neolithic spread in the region by focusing on the role of voyaging to understand better the core elements that produced the observed pattern of dates. We also explore the effect of cultural interaction with Mesolithic populations living along the coast. The simulation study shows that (i) sea travel is required to obtain reasonable predictions, with a minimum sea-travel range of 300 km per generation; (ii) leapfrog coastal dispersals yield the best results (quantitatively and qualitatively); and (iii) interaction with Mesolithic people can assist the spread, but long-range voyaging is still needed to explain the archaeological pattern.

Accordingly, it is reasonable to think that the initial spread of the Neolithic in the west took place along coastal routes by first farmers ultimately coming from northwest Italy (proposed alternative routes, which are not supported by reliable evidence, are discussed in ref. 18).

The purpose of this article is to put forward a new and more quantitative approach to the question of the exceptionally fast spread of the Neolithic in the west, one that focuses on the role of voyaging. We perform computational simulations to identify the underlying mechanisms of voyaging and estimate the associated parameters that can account for the archaeological pattern. Previous computational approaches that included sea travel only analyzed its average effect on a continental scale (19). Thus, this modeling exercise is needed to understand better the processes at work in the coastal spread of first farmers in the West Mediterranean.

Although the models considered here are principally demic in character, such a fast coastal spread may well have entailed some interactions with local Mesolithic populations. Interaction can take place in a range of different forms, from exchanges of information and material culture to the adoption of farming as a way of life by a Mesolithic community. As a first step, we explore briefly two aspects of population interaction: cross-mating and acculturation [also called vertical and oblique/horizontal cultural transmission, respectively (20)].

Methods

Neolithic Chronology. The database used in this study (available as Dataset S1) contains high-quality dates from 42 Early Neolithic sites of the West Mediterranean, including northwest Italy, southern France, Spain, Portugal, and northern Morocco, most of which have a coastal location. For each site, the earliest date has been selected to estimate the time when the Neolithic first arrived in a given area. The sites are represented in Fig. 1A as circles, colored according to their calibrated dates. We analyze the dates by applying natural neighbor

Significance

The Neolithic expansion in Europe took place at an average rate of 1 km y−1. In the West Mediterranean, the archaeological record yields a much faster rate of spread, one that cannot be explained by classical overland models. Voyaging has been put forward as an alternative line of explanation. Here, we develop a computational model to identify the key elements and mechanisms and to estimate the values that yield outcomes that fit the observations. The results show that voyaging is indeed required to explain the pattern. We have also found that interactions with local hunter-gatherers played a lesser part in the fast rate of spread.

Author contributions: J.Z., J.F., and A.J.A. designed research; N.I. performed research; N.I. and J.F. contributed new reagents/analytic tools; J.Z. prepared the database; N.I. and J.F. analyzed data; and N.I., J.Z., J.F., and A.J.A. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

1To whom correspondence should be addressed. Email: neus.iserna@udg.edu.

This article contains supporting information available online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1613413114/-/DCSupplemental.

www.pnas.org/cgi/doi/10.1073/pnas.1613413114

PNAS | January 31, 2017 | vol. 114 | no. 5 | 897–902
Figs. 2–4. The dates shown as diamonds in Fig. 1 represent the calibrated dates before Christ (cal BC), and the error bars indicate their 2σ error ranges. The dates shown as diamonds in Fig. 1 and disregarding coastal sites with medians younger than 5000 BC to avoid interpolation (background color scale in Fig. 1) are database dates represented as diamonds in Fig. 1 (Table S1). Symbols represent the calibrated dates before Christ (cal BC), and El Zafrín (Islas Chafarinas; 4416 BC). (B) Database dates represented by areas and chronologically following, from bottom to top, the order in Dataset S1). Symbols represent the calibrated dates before Christ (cal BC), and the error bars indicate their 2σ error ranges. The dates shown as diamonds are the earliest for each area and are the dates used to test the validity of the different models; they are also shown as diamonds in Figs. 2–4.

The eight sites yielding such dates (listed in Table S1) are represented as diamonds in Fig. 1A, (Table S1) are represented as diamonds in Figs. 2–4. Horizontal/oblique cultural transmission relates to hunter-gatherers who teach or convert to farmers (29). Note that, in general, acculturation could also take place at different levels, with the local process.

Computational Simulations. We have written four computational models that aim to simulate the Neolithic spread in the West Mediterranean. A model is always a simplification of the complex processes taking place in human events; in this case, we use models that follow a dispersion-interaction-reproduction scheme, which allows us to study different possibilities for each of these steps. We apply a deterministic approach, which gives a good approximation to the averaged stochastic process, because we are not trying to reconstruct the exact processes that took place but, rather, to assess the range of scenarios that may have led to the observed pattern.

We run the simulations on a Cartesian grid of 50-km × 50-km cells, which are classified as inland, coastal, or sea cells. The map suggests an inland expansion from the interior toward the coast and at a later time; thus, this area is given for purposes of reference.

Step 1: Dispersion. Land dispersion takes place so that 38% of the population stays put in a cell. This value, which is called the persistence in demography, is taken from ethnographic data (21). The rest of its population relocates homogeneously to the four nearest neighbors. When one (or more) of the four neighbors is a sea cell, the population that should move there will be redistributed according to one of the four models that we used. Model 1 does not allow travel by sea, so the population that cannot "settle" in a sea cell does so in its other habitable neighbors (land/coastal cells). Model 2 allows voyaging along the coast up to a certain range, but with preferential attachment to the nearest cells (we use a Gaussian curve as the dispersal probability function). Model 3 allows for voyaging along the coast within a certain range, with all distances having an equal probability. Model 4 is based on a leapfrog process, with all voyagers relocating at a given distance measured along the coast (reached by means of cabotage), skipping over several cells between a starting place and a destination. Seagoing takes place in models 2–4 by means of cabotage [i.e., sea-travel distances are measured along the coast instead of along straight lines as in previous simulation models (19, 25)] and moves in an outward direction.

Step 2: Population interaction. When both farmer and hunter-gatherer populations are present in a given cell, they can interact. In broad terms, this interaction can involve rich and complex processes. Here, we consider two main forms of interaction: vertical and horizontal/oblique cultural transmission (20).

Vertical cultural transmission relates to cross-mating between late hunter-gatherers (M) and first farmers (N), whose children we take to become farmers in agreement with ethnographic evidence (26, 27). We compute the number of cross-matings per generation as

$$l_{ct} = \pi \frac{N - M}{N + M},$$

where \( \pi \in [0,1] \) indicates the degree of interbreeding (28).

Horizontal/oblique cultural transmission relates to hunter-gatherers who adopt farming; we compute the number of acculturated Mesolithic individuals per generation as

$$l_{ht} = f \frac{N - M}{N + \gamma M},$$

where the ratio \( f = c/\gamma \) gives the number of hunter-gatherers that a pioneer farmer (i.e., for \( N < c M \) teaches or converts to farmers (29). Note that, in general, acculturation could also take place at different levels, with the local population acquiring certain traits but not becoming full farmers in one generation; however, here we use Eq. 2 as a first approximation to the process.

Step 3: Population growth. Finally, we compute the new generation in each cell by applying population growth to the parent generation (couples of mixed, coupled, and couples of hunter-gatherers who have adopted farming). We compute the new size of the population as

$$N(t + \tau) = R_0 N(t) \quad \text{if} \quad N < N_{max},$$

$$N(t + \tau) = N_{max} \quad \text{if} \quad N \geq N_{max},$$

where the maximum population per cell is \( N_{max} = 3,200 \) individuals. This value is obtained from a maximum population density of 1.28 individuals/km² (30) and the size of the cells. The reproduction coefficient is \( R_0 = \exp(\alpha T) \) (21), where \( \alpha \) is the intrinsic growth rate and \( T \) is the generation time. Using a logistic approach instead of Eq. 3 will yield the same results; however, in more general cases, Eq. 3 avoids negative population numbers, whereas a logistic treatment does not (21, 31).
The position of the Neolithic front is identified as the cells that have reached, for the first time, a population of 300 Neolithic individuals during a given iteration. We have chosen this value (about 10% of the maximum cell population) as a practical threshold to compare the simulation outcomes with the archeological pattern. The simulated arrival times for the respective cells can be plotted on a map and then compared with the observed archeological dates to evaluate the goodness of fit between the two.

Results and Discussion

Our main interest is in reproducing the earliest arrival of the Neolithic in areas along the West Mediterranean coasts. For this reason, we compare the simulation results with a database consisting of the earliest high-quality Neolithic dates in the region (Dataset S1). As explained in Methods, we compare the simulation results with the earliest date for each area of interest (diamonds in Fig. 1B). We consider a model to predict the observed arrival time correctly if the simulated arrival time falls within the 2σ calibrated range (represented by a white diamond in Figs. 2 and 4). Otherwise, we consider the simulation to arrive either “too late” (the modeled expansion is too slow; represented by a black diamond) or “too early” (the modeled expansion is too fast; none of our maps contain such a case).

Below, we study the effects of each of the three steps in the simulation process (dispersal, interaction, and reproduction) to identify the necessary conditions for the spread to reach each of the respective areas at the right time.

Voyaging Mechanisms. We first study and try to characterize the importance and dynamics of coastal voyaging in the West Mediterranean expansion. The analysis of the database by interpolation (Fig. 1A) supports the inference of a process of inland expansion from nearly coeval coastal source points. This pattern fits with a fast coastal expansion characterized by long-distance jumps. However, for the sake of completeness, we explore four different dispersal mechanisms, taking into consideration voyages that range from 50 to 1,000 km. In all cases, we use a growth rate of $a = 2.8\%$. This value is taken from data on preindustrial farming populations (32), and it is consonant with other independent estimates of preindustrial farming populations, $a = 2.5\%$ (33), as well as with estimations obtained from Neolithic skeletal remains, $a = 2.4\%$ (34).

When no maritime dispersal is taken into account (model 1), the predicted Neolithic front is much slower than the observed arrival times, reaching Portugal more than 1,000 y too late. Indeed, comparing Fig. 2D with Fig. 1A it shows that this model does not fit the archeological data. This result is, of course, expected (supporting the position that overland dispersion cannot explain the data), and voyaging has to enter the picture.

Any model in which first farmers can travel by sea improves the previous simulation results. Considering a kernel where voyagers settle at successive locations along the coast within a specified range of distances (models 2 and 3) can yield “perfect” coastal fits (i.e., with all coastal areas reached within the 2σ range) when voyaging is of sufficient length. Model 2 requires voyages of up to 600 km to reach all of the coastal areas by observed arrival dates, whereas in the case of model 3, one can obtain a similar result with somewhat lower voyage ranges (over 450 km). For lower ranges, the simulations provide better results when a larger fraction of the population traveling by sea settles away from its point of departure. This behavior can be seen in Fig. 2, where for a given range of 350 km, model 3 (Fig. 2C) yields a slightly better fit (white diamonds) than model 2 (Fig. 2D).

Model 4, which takes a leapfrog approach, is the one that yields the best results with lower voyaging ranges. Indeed, assuming jumps of 350 km, all coastal areas are reached by the time of their observed arrival dates (Fig. 2D). For this model, good simulation results can also be obtained for coastal jumps of 450 km, and other jump distances between 350 km and 600 km are marginally consistent with the observed data, reaching one or two coastal areas either slightly too late (less than one generation) or else too early (less than two generations).

Thus, we find that model predictions improve when more individuals relocate to the most distant places within a given range (model 4). This scenario, in turn, allows us to obtain good fits with lower voyaging ranges. We can also see in Fig. 2 that only in this scenario (model 4; Fig. 2D) can we recover a pattern with multiple points of coastal entry similar to the data (Fig. 1).
contrast, the other models yield a progressive expansion. In short, model 4 yields the best results, both from a qualitative point of view and from a quantitative one.

Note that although we are using a deterministic treatment where all sea jumps happen to have the same length (model 4), we do not mean to imply that this treatment reproduces exactly how the process took place. What we have found is that, to account for the coastal spread of the Early Neolithic across the whole of the West Mediterranean, first farmers must have advanced by means of cabotage jumps of at least 350 km, but the exact length of a given voyage may well have been determined by socioecological factors (e.g., presence of river mouths (22)) not included in model 4.

In addition, we can see that areas in the Iberian interior are often reached too late on the basis of the voyage ranges considered above (e.g., black diamond in Fig. 2D), implying that the land-based part of the expansion in these areas needs a more refined treatment. For the northeast, in particular, the role of the Ebro River in facilitating inland expansion along its banks, as suggested for the Danube-Rhine corridor in the LBK expansion (35), needs to be considered. A fuller study of this dynamic, however, is beyond the purpose of this article.

Coastal Speed. As a complementary visualization to the simulation results shown in Fig. 2 and to compute the coastal rate of spread, arrival times versus coastal distance to Arene Candide (the origin of the simulation) are plotted in Fig. 3 for each of the seven coastal areas (the much slower model 1 is excluded because it would fall out of range; note that distances are measured following the coastline and not from straight lines between locations). Fig. 3 shows that up to about 1,000 km from the origin, the three models that involve voyaging behave much the same and that it is only at further distances that a leapfrog scenario (model 4) provides a better fit.

Computing the speed from the archeological data, we obtain that the coastal expansion took place at about 8.66 km y⁻¹ (R = 0.830). For the parameters considered (a = 2.8% and voyage ranges of 350 km), the predicted rates of spread are: 1.70 km y⁻¹ (R = 0.925) for model 1, 4.22 km y⁻¹ (R = 0.999) for model 2, 5.02 km y⁻¹ (R = 0.998) for model 3, and 7.66 km y⁻¹ (R = 0.918) for model 4. Thus, model 4 agrees best with the observed spread rate. Note how a voyage range of 350 km does not imply that the coastal front advances at a speed of 350 km per generation (i.e., about 10.9 km y⁻¹). On the contrary, the front is substantially slower: Because only a fraction of the population living at a coastal site will travel by sea, the population at the front will sometimes be only a few individuals; thus, the Neolithic front may be stalled until the population number has grown. In fact, a general feature of reproduction-dispersal models is that the front speed is equal to the jump length divided by the time between successive jumps (350 km per generation in our case) only if the reproduction rate is sufficiently high; otherwise, the front travels at a slower speed (36).

Sensitivity to Simulation Origin. As argued earlier, the choice of a starting point located on the mainland (Arene Candide) is consistent with the distribution of obsidian artifacts in the region (obsidian from sources on the off-shore island of Sardinia are not found at Early Neolithic sites in Spain) and with voyaging in the form of cabotage as the means of the coastal spread (18). Because the choice of the cell of origin can affect the locations where the pioneer voyagers arrive in the models, it might modify, in principle, the ranges that yield the best fit. Therefore, we have repeated the simulations for model 4, shifting the origin first one or two cells eastward and then likewise to the west, while keeping the starting year at 7700 B.P. (5751 BC). Recall that the distance between the centers of two adjacent cells is 50 km. Studying the results from these other possible origins, the best fits are still obtained with leaps of 350–450 km (although “jumps” of up to 600 km may also yield agreement with the observed data). In other words, our results are not affected by the exact origin chosen for the simulation runs.

Growth Rate. We have also tested a wider range of growth rates in the case of model 4 to learn more about their effect. In short, increasing the growth rate improves the results when shorter jumps are involved, but not for values of voyaging below 300 km (even considering a = 4%, which is an unrealistically fast reproduction rate for human populations). Clearly, this finding implies that a minimum jump length is necessary.

In addition, we have evaluated the extent to which the growth rate could be lower than the maximum values reported in the ethnographic and archeological literature (a = 2.8%). In the case of model 4, we find that for jumps in the range of 350–450 km, we could lower the growth rate to a = 2.3%, close to the value of a = 2.4% (34) estimated on the basis of the archeological record, and still obtain much the same goodness of fit.

Population Interaction. Although the archeological record for the Late Mesolithic in many areas of the West Mediterranean is still quite modest, there are Late Mesolithic sites on the coast of Portugal that provide good evidence of the side-by-side coexistence of late hunter-gatherers and first farmers over a fair arc of time (8). On the other hand, there is still a shortage of good evidence for the nature and temporal length of Mesolithic and Neolithic coexistence along the Spanish coast (37). On the Mediterranean coasts of France and northern Italy, the evidence so far does not support the idea of coexistence of any real length of time (38, 39). In short, given the paucity of evidence on the Late Mesolithic currently available for most of the region, we explore a few basic scenarios in this section rather than test a specific hypothesis.

So far, all of the Late Mesolithic sites found in Iberia are located either near the coast or else on a river valley (e.g., the Ebro Valley), with no evidence for Mesolithic populations in the central Meseta (37). In this light, and with the aim of evaluating the effects of interaction, we consider that Mesolithic populations were restricted to coastal cells. Following the ethnographic evidence for populations of hunter-gatherers living on the coasts of southwest Tasmania (40), we assume a band of 50
hunter-gatherers per cell (50 km on each side; the results for a larger band size, 80 individuals, are given in SI Appendix).

Let us consider that Neolithic individuals relocate according to model 4 (with voyages of 350 km). Because the two processes of population interaction discussed briefly here (cross-mating and acculturation) result in the incorporation of Mesolithic people into Neolithic communities, the effective population growth is increased (i.e., not all of population growth is due simply to childbearing on the Neolithic side). Accordingly, in this section, for the first farmers’ side of the interaction, we consider the growth rate of $a = 1.8\%$; for example, Fig. 4A shows the results for a purely demic spread under such conditions.

The first form of interaction, known as vertical cultural transmission (20), assumes that some newcomers (first farmers) mate with individuals in the local Mesolithic population. For example, in accord with ethnographic parallels, we expect voyaging male farmers to mate with local female hunter-gatherers, who will then become part of the farmer population (27). Such cross-mating, which is taken into account in Fig. 4B, will give rise to a slightly faster rate of spread (compare Fig. 4A and B, especially in the Portuguese case, where the effect is more pronounced due to the cumulative effect of continually adding population at the front). In this case, the arrival times for the respective coastal areas are correctly predicted only for the maximum possible value of the interbreeding parameter, $\eta = 1$ (28). It is, however, important to note that such a high interbreeding level is highly unlikely when we bear in mind that only a small proportion of female hunter-gatherers will be available for interbreeding (i.e., women between, say, 15 and 30 years old who do not already have a partner).

The other form of interaction, as mentioned above, is acculturation or recruitment [known also as horizontal/oblique cultural transmission (20)], that is, Mesolithic individuals or families who adopt first farming and become part of the Neolithic community. Because values for the acculturation parameter C (the number of hunter-gatherers incorporated in this way into the farming community per pioneer farmer per generation) are not well known today, we have chosen three values to illustrate three quite different scenarios: $C = 0.2$, $C = 1$, and $C = 10$. Again, as expected, the incorporation of acculturated farmers in Neolithic populations gives rise to a faster spread. When no vertical cultural transmission is taken into account, the model requires that at least one hunter-gatherer adopts farming per pioneer farmer ($C = 1$) in order for the expansion to reach all coastal areas in agreement with their observed arrival times. Interestingly, increasing the acculturation parameter to $C = 10$ yields almost no change in the simulations’ arrival times. The reason for this result is that the number of Mesolithic individuals per cell (e.g., up to 50 individuals) is relatively small, so the potential contribution of Mesolithic acculturation to increases in the rate of spread is limited.

Finally, we can combine these two forms of interaction by using somewhat lower intensities for each process in a simulation run. For example, as shown in Fig. 4C, we obtain a good fit with the observed arrival dates for an interbreeding level of $\eta = 0.5$ and an acculturation level of $C = 0.2$ (one Mesolithic individual brought into first farming for every five pioneer farmers per generation). Again, further increasing the value of $C$ yields little change in the simulation outcomes (maps are provided in SI Appendix).

From these results, we see that including such interactions gives rise to faster rates of spread with somewhat lower rates of population growth intrinsic to the spreading farmers. However, it does not eliminate the need for voyages that involve a distance of at least 300 km. In addition, it is worth adding that rates of interbreeding and incorporation that are too high will yield the complete acculturation of Mesolithic populations within one or two generations of first contact. However, this rate of acculturation is clearly at odds with the evidence for the coexistence of the two populations in Portugal over a notable span of time. Even for areas with shorter spans of coexistence, such a high pace of incorporation is probably unwarranted. In other words, it is more likely that cross-mating and acculturation took place at lower levels of intensity alongside other forms of interaction, such as a hunter-gatherer acting as a guide with local knowledge or else participating in a voyage led by a first farmer with a boat. The latter kinds of interaction are called mutualism (3), and they too would have made positive contributions to such things as choosing the new place for a first farmer to settle or having enough crew members on hand to set out on a voyage.

### Conclusions

The initial spread of the Neolithic in the West Mediterranean shows a clear pattern: a rapid expansion that involved multiple points of entry along the coast. This pattern led to the formulation of a model based on voyaging, which provides both a confirmation and an extension of the maritime pioneer model (8). In this article, we have explored several quantitative models of the spread (with and without voyaging) and found that (i) sea travel, as opposed to a land-based expansion, is necessary to explain the spread of the Early Neolithic in the West Mediterranean and (ii) the best fit with the observed pattern of arrival times is one in which voyaging, in the form of cabotage, takes

Fig. 4. Effect of Neolithic–Mesolithic interaction on the expansion process. The results derive from the use of model 4, with a voyage range of 350 km and an intrinsic growth rate of $a = 1.8\%$ for three cases: no interaction ($A$, $C = 0$), vertical cultural transmission only ($B$, $\eta = 0.5$, $C = 0$), and vertical and horizontal cultural transmission ($C$, $\eta = 0.5$, $C = 0.2$). We assume Mesolithic bands of 50 people, but the results for 80 people are similar (SI Appendix). White diamonds represent areas reached within the calibrated $2r$ range; black diamonds represent areas reached by the model later than the calibrated range.

Isern et al.  
PNAS | January 31, 2017 | vol. 114 | no. 5 | 901

ANTHROPOLOGY
place by means of long-distance relocations. These findings now make it possible to put forward distances of voyage dispersal in the range of 300–450 km. Although such distances might appear to be extreme (they came as a surprise to us), they are necessary to match the observed pattern of arrival dates. In fact, no value of voyaging below 300 km can fit with this pattern.

We have also found that to achieve a fast enough spread, an effective growth rate of at least 2.3% is called for. This growth rate can be achieved directly by means of reproduction and also through some level of interaction with local Mesolithic populations. Here, we have considered two cases: incorporation of Mesolithic individuals either by means of mating or by means of acculturation. The results show that population interaction can play a positive role in the rapid spread of the Early Neolithic, although its contribution is limited by the small size of Mesolithic populations. On the other hand, the level of incorporation called for to achieve a major effect on the rate of spread would be higher than what is consistent with the coexistence between the early Neolithic and the final Mesolithic observed in Portugal (8). Nonetheless, we think that some level of interaction was required to sustain such a high rate of spread over a span of some 300 y; thus, other mechanisms not requiring the adoption of early farming, such as mutualism, may have played a role as well.

In summary, our results point toward a spread driven by small groups of first farmers who relocated by means of voyaging over distances of 300–450 km at the expanding front. Interaction with local hunter-gatherers, in the form of cross-mating, acculturation, or mutualism (or some combination of the three), would have facilitated the growth of farming populations and, in turn, their rapid spread in the West Mediterranean.

Future Perspectives

The results presented here define some boundaries for the interpretation of population history; they also provide a basis for future inferences with respect to questions that have yet to be explored in greater depth. To develop a better knowledge of how the spread of first farming took place in the West Mediterranean, further research needs to be done along four main lines: (i) We need more and better dates for those areas where high-quality AMS dates are still not available; (ii) the best computational scenario obtained so far (model 4) can now serve as the starting point for a new generation of models with more complex dynamics (including a fuller treatment of the question of population interactions); (iii) in future work, it will be useful to incorporate more geographic detail (e.g., rivers, mountains) in modeling the inland spread; and (iv) the study of the DNA in Neolithic populations (to be done when more individuals come to light than the very limited sample available today) will provide an independent line of evidence for evaluating the conclusions reached in this article.

ACKNOWLEDGMENTS. This work has been partially funded by Ministerio de Economía, Industria y Competitividad (Grants SimulPast-CSC-2010-00034, FIS-2012-37386, and FIS-2015-65625-P), Instituto de Astrofísica de Canarias, Foundation Workshop on Island Archaeology and the Origins of Seafaring in the Eastern Mediterranean, further research needs to be done along four main lines: (i) We need more and better dates for those areas where high-quality AMS dates are still not available; (ii) the best computational scenario obtained so far (model 4) can now serve as the starting point for a new generation of models with more complex dynamics (including a fuller treatment of the question of population interactions); (iii) in future work, it will be useful to incorporate more geographic detail (e.g., rivers, mountains) in modeling the inland spread; and (iv) the study of the DNA in Neolithic populations (to be done when more individuals come to light than the very limited sample available today) will provide an independent line of evidence for evaluating the conclusions reached in this article.

Moreover, our results are also suggestive of the need for a more detailed geographic approach to farming front propagation, where models of the initial phase can be further extended to cover more complex scenarios. In this regard, the study of the DNA in Neolithic populations (to be done when more individuals come to light than the very limited sample available today) will provide an independent line of evidence for evaluating the conclusions reached in this article.

Moreover, our results are also suggestive of the need for a more detailed geographic approach to farming front propagation, where models of the initial phase can be further extended to cover more complex scenarios. In this regard, the study of the DNA in Neolithic populations (to be done when more individuals come to light than the very limited sample available today) will provide an independent line of evidence for evaluating the conclusions reached in this article.
SI Appendix: Modeling the role of voyaging in the coastal spread of the Early Neolithic in the West Mediterranean

Neus Isern\textsuperscript{a,1}, João Zilhão\textsuperscript{b,c}, Joaquim Fort\textsuperscript{a,c}, and Albert J. Ammerman\textsuperscript{d}

\textsuperscript{a} Complex Systems Lab and Physics Department, Universitat de Girona, 17003 Girona, Spain.  
\textsuperscript{b} Departament d’Història i Arqueologia (SERP; Grup de Recerca SGR2014-00108), University of Barcelona, 08001 Barcelona, Spain.  
\textsuperscript{c} Catalan Institute for Research and Advanced Studies (ICREA), 08010 Barcelona, Spain.  
\textsuperscript{d} Department of Classics, Colgate University, Hamilton, NY 13346, United States of America.  
\textsuperscript{1} Corresponding author. Email: neus.isern@udg.edu

Database notes

Selection criteria

When preparing the database, only the earliest date for each site was retained, following the criteria in Ref. (1), i.e., excluding any and all rank 7 results, relying principally on results ranked 1-4, and resorting to results in ranks 5-6 only when nothing better was available. Despite the problems pointed out in Ref. (2), the database accepts the Nerja/Vestíbulo date (rank 5) on the basis of consistency with results for the adjacent geography (2). Two dates on shell, for the Portuguese sites of Cabranosa and Padrão, are included; their calibrated ages were calculated with the terrestrial curve after correction for the reservoir effect following Ref. (3).

The stringent criteria used in the selection of dating results find justification in the nature of the research question. The use of statistical techniques can be of help in sorting out signal from noise when dealing with large datasets, and such an approach makes sense when the questions being addressed concern such things as the duration of an archeological culture or the timing of its floruit. When dealing with “first arrival” issues, however, the “chronometric hygiene” approach used here is mandatory, and indeed the standard practice in comparable situations, namely the spread of the Lapita complex across the Pacific, or the time of settlement of New Zealand or the Hawaiian archipelago, not to mention, in the Mediterranean basin itself, the time of first human colonization of the Balearic Islands (4-6).

Two examples suffice to illustrate why, for the study of the spread of farming across the West Mediterranean, a “chronometric hygiene” approach is unavoidable. When comparing the two extremes of the distribution of the Impressa-Cardial complex, southern Italy and Portugal, a time lag of some 400 years is readily apparent; yet, as shown by the dating of the lakeside settlement of La Draga, that much can be the difference between the real age of the settlement as derived from short-lived, bone or cereal samples, and the apparent age obtained when dating the oak trees therein used as timber (7). The other example concerns the dating of Cueva de Nerja, and shows how the pattern of spread can be affected not only by old wood effects but also by issues of sample identification; here, a bone initially classified
zoo-archeologically as domesticated ovicaprid returned a result some 700 years too old, falling squarely in the time range of the regional Mesolithic and indicating that the sample was in fact of the extant, wild species *Capra pyrenaica* (2). Needless to say, uncritical use of such problematic dates would result in a “first arrival” scenario quite distinct from that in Fig. 1.

The downside of the “chronometric hygiene” approach consists in the risk of losing, or missing information concerning areas were quality dates are too few, resulting in underestimated “first arrival” dates for those particular areas. In our case, however, this is not a real problem, for several reasons, namely: (a) we are dealing with a process of spatio-temporal expansion in which the key factor in the calculation of such parameters as the rate of spread is a good control of the dating at the two ends—here, western Liguria and central Portugal—and the number of quality results in our dataset is, for these areas, sufficient for the chronological boundaries of the process to be considered robust; and (b) a rather large number of quality results is available for two key intermediate areas, Catalonia and Valencia, and they are consistent with the gradient to be expected if the process indeed was one of westward expansion with an origin in northern Italy.

The database was compiled in 2014 using the then available data and the models were developed on that basis; here (Dataset S1), we append to it results that could not be used in the models because they were obtained or published since (or only since came to our attention). For the sake of completeness, this new set includes one result for the open-air site of Cabecicos Negros (Almería) obtained on a sample of *Cerastoderma edule*, which we calibrated with the marine curve, as the local reservoir effect is unknown. Note that the site is located 2 km inland, meaning that the sample, if derived from a live mollusk collected for consumption, most probably comes from the immediate brackish water environment described by Camalich-Massieu et al. (8); therefore, the different between its true and apparent ages may well be even larger than if the sample had been collected in a marine environment. Except for a clearly anomalous 7280±40 BP date (Beta-347630) that most likely reflects the use of dead, beach-collected shells of *C. edule* for the manufacture of beads, an activity well documented at the site (9), the result we retained is the oldest date obtained for Cabecicos Negros (where all dates are on samples of that same taxon).

The appended results remain entirely consistent with the general picture in Fig. 1, supporting the notion that we can be confident that, within statistical error, we have indeed picked-up the time of first appearance of farming in the areas concerned. We have compared a few test runs against such new results (10-12), and indeed have found that at least in the case of model 4 (with no interaction and jumps of 450 km), the simulations remain consistent with them. Even though their inclusion might affect the goodness of the prediction for particular sets of parameters, we expect these additional results to modify neither the global picture nor our conclusions.

The only major geographical gap in our dataset concerning an area with a well-known and rather dense network of Early Neolithic sites is the French Languedoc, as is well apparent in Fig. 1. The lack of quality dates for these sites means that there is at present little that can be done to remedy the situation. On the other hand, this example serves to highlight another advantage of our “chronometric hygiene” approach, namely the possibility to use the general pattern in Fig. 3 to predict a time of first arrival of
the Neolithic in that area around 5650 cal BC. Indeed, such a time frame is suggested by some of the charcoal dates obtained for the sites of Peiro Signado and Pont de Roque-Haute, and would be consistent with the associated Ligurian ceramics (13).

**Comparison with model results**

As shown in Fig. 1, the archaeological data are clustered in nine areas, which we denote as: Italy/France, Catalonia, Valencia, Andalusia, Northern Africa, Algarve, Central Portugal, Cantabria and Interior. Even though only the first seven areas, those directly related to the coastal spread along the Mediterranean shores, are of interest here, we also compare the results with the Interior area as a reference for the inland expansion.

Since we are interested in the patterning of the spread, testing whether the models correctly match the archaeological evidence means assessing a given simulation run’s goodness of fit for each area. The eight dates selected to test the simulation outcomes (represented by diamonds in Fig. 1b) are listed in Table S1. We consider that the model predicts the observed arrival times in the respective areas correctly if the simulated arrival times fall within the 2σ calibrated range of these eight earliest dates. Otherwise, the simulations may be too slow (if some areas are reached after the 2σ range) or too fast (if some areas are reached before the 2σ range).

<table>
<thead>
<tr>
<th>Area</th>
<th>Site</th>
<th>2σ cal BC range</th>
<th>Median of the cal BC range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Italy/France</td>
<td>Arene Candide</td>
<td>5,490 – 5,791</td>
<td>5,711</td>
</tr>
<tr>
<td>Catalonia</td>
<td>Guixeres de Vilobí</td>
<td>5,490 – 5,643</td>
<td>5,582</td>
</tr>
<tr>
<td>Valencia</td>
<td>Mas d’Is</td>
<td>5,481 – 5,621</td>
<td>5,547</td>
</tr>
<tr>
<td>Andalusia</td>
<td>Cueva de Nerja</td>
<td>5,481 – 5,616</td>
<td>5,538</td>
</tr>
<tr>
<td>Northern Africa</td>
<td>Kef Taht el Ghar</td>
<td>5,076 – 5,483</td>
<td>5,333</td>
</tr>
<tr>
<td>Algarve</td>
<td>Cabranosa</td>
<td>5,374 – 5,623</td>
<td>5,513</td>
</tr>
<tr>
<td>Central Portugal</td>
<td>Galeria da Cisterna</td>
<td>5,326 – 5,481</td>
<td>5,414</td>
</tr>
<tr>
<td>Interior</td>
<td>Cueva de Chaves</td>
<td>5,478 – 5,615</td>
<td>5,528</td>
</tr>
</tbody>
</table>

**Computational model: extended description**

We have developed four computational models with the aim of simulating and reproducing the Neolithic coastal spread in the West Mediterranean. Here we present an extended description of our models, which have been outlined in the *Methods* section. We have written the models in Fortran 90/95, and the code is available at [http://copernic.udg.es/QuimFort/WestMedNeo_Program_Files.zip](http://copernic.udg.es/QuimFort/WestMedNeo_Program_Files.zip).

Our model follows a dispersion-interaction-reproduction scheme, which allows us to study different possibilities for each of the three steps. The two main effects we focus on are: the importance of coastal voyaging (which we explore using different travelling kernels) and the effect of cultural transmission (which allows us to explore a range of processes, from purely demic to mixed demic-cultural with varying degrees of interbreeding and acculturation).

We run our simulations on a grid of square cells of 50 km x 50 km, prepared using an Albers Equal Area Conic Projection for Europe, which we have adapted from Ref. (14). The 50 km value is the characteristic
distance moved by preindustrial farmers per generation, as estimated from observed mobility and persistence data (see (15), note [32]). In addition, the land-based model studied by Bernabeu Aubán et al. (16) for Iberia suggested that settlements were most likely spaced about 25–100 km.

Each cell in the grid is identified as being either a sea cell, a coastal cell or an inland cell. As shown in Fig. S1, coastal cells are defined so that no inland cell has a sea cell on any of its four nearest neighbors.

In agreement with archaeological evidence, we assume that Neolithic populations expanding along the West Mediterranean coast originate from northwest Italy. We take as the starting point of our simulations a cell located in northern Italy, at the northeastern extreme of our region of study, and with central coordinates (44.474N, 8.407E). The site of Arene Candide (site 1 in Table S1), which is the easternmost site in all figures and has the oldest calibrated date median, 7660 cal BP, falls within this cell. Using its date as representative, the simulations begin at 7700 BP (5751 BC). In the simulations we also explore the consequences of shifting this starting point one or two cells westward as well as eastward (keeping the same starting date).

We initially set the Neolithic population of this source cell at its maximum value, $N_{\text{max}} = 3,200$ individuals (initially, all other cells have no farming population). This value for the maximum population per cell is obtained from a realistic value of the carrying capacity of Neolithic populations, 1.28 individuals/km$^2$ (17)—knowing that each cell covers an area of $50 \text{ km} \times 50 \text{ km}$.

When we consider Neolithic-Mesolithic interactions, we assume that Mesolithic populations are present only in the coastal cells, following the currently available archaeological data (18). We estimate the Mesolithic population number and distribution from data on the aboriginal population of Southwest Tasmania (19), a hunter-gatherer population living in coastal and riverside contexts. From these data, we assume that the Mesolithic bands were distributed approximately every 50 km (one band per cell) with a population of 50–80 individuals per band.

**Fig. S1:** Model grid. The diagram shows the three kinds of cells used in the model: inland cells in brown, coastal cells in yellow and sea cells in light blue (coastal cells are defined so that no inland cell has a sea cell as any of its four nearest neighbors). The red lines joining coastal cells indicate how the coastal distance between cells is computed.
The simulations run for several iterations, each one corresponding to a generation, until the Neolithic has spread through all of the West Mediterranean. We use a realistic value of the generation time $T = 32\gamma$ (20); this corresponds to the average age difference between the parents and one of their children measured for pre-industrial populations. At the end of each iteration, we identify the position of the Neolithic front as the cells that have reached for the first time a population of 300 Neolithic individuals during that iteration. We use this value, about 10% of the maximum population per cell in the simulations (see above), as a practical threshold slightly above the minimum size required for a human reproductive network to be viable (21) and bearing in mind that the earliest archeologically detected Neolithic in a given cell may have had as-yet undetected local settlement antecedents. Using a lower or higher threshold would modify neither the minimum sea-travel range nor the ranges that yield the best results. The arrival times obtained for each cell can then be plotted on a map and/or be compared with the archeological data in order to quantify the goodness of fit with the simulated results.

The fact that our models use a dispersion-interaction-reproduction scheme means that at every iteration—namely, a generation—the model performs three sequential steps, which involve population dispersal, population interaction and population growth. We describe the three steps below. Changing the order of the three steps does not affect the results.

**Step 1: Population dispersal**
Regarding the dispersal process, we have designed four different approaches to describe how to include voyaging when modelling the West Mediterranean expansion. In all cases, voyaging takes place as a cabotage process (i.e. traveling along the coast); the differences between the four approaches depend then on how probable it is for the voyaging first farmers to settle at the possible destinations.

In all of our models, at each iteration and for each inhabited cell, a fraction $p_e = 0.38$ (persistence) of the Neolithic individuals remains in the same cell. This value for $p_e$ has been estimated previously from ethnographic data (15). The rest of the population moves to other cells according to the following travel rules, which depend on the type of the cell of origin: inland or coastal.

If the origin cell is an inland cell, the remaining individuals are equally distributed among the four nearest neighboring cells—each receiving a fraction $\left(1 - p_e\right)/4$ of the initial population (Fig. S2a). When a fractional value is obtained, the number of individuals who move is rounded to the nearest integer (with 0.5 rounded up to 1). The program checks to see that the final number of individuals—those who jump plus those who stay—is equal to the initial number; if necessary, the possible offset produced by rounding is applied to the number of individuals who stay at the cell of origin (the same rounding approach is applied when dealing with coastal cells).

When the cell of origin is a coastal cell, some of its nearest neighbors are sea cells, and thus ineligible destinations for settlement (Figs S2b, c). Below we describe how this situation is dealt with according to four different voyaging models.
Model 1. No voyaging. In this first approximation, we do not allow sea travel. Thus, here, the individuals that would move to sea cells are redistributed in equal parts among the other neighboring cells that are eligible destinations (i.e., those classified as inland or coastal cells; black arrows in Figs. 2b, c). In Fig. S2b, for example, the three land neighbors would each receive \((1 - p_c)/3\) of the original population and no first farmer would travel by sea.

Model 2. Voyaging with decreasing probability to settle with distance. In this model we assume that Neolithic individuals can travel by sea in the form of cabotage, establishing new settlements in each coastal cell within a certain range (e.g., 300 km), though with diminishing probability to settle with increasing distance (i.e., there is a preference to relocate close to home). Therefore, in this case, the individuals who would move from the cell of origin to sea cells are redistributed among the coastal cells within the travel range and according to a given probability distribution. We consider a Gaussian distribution

\[
p(x) = K_1 \exp\left(-\frac{1}{2} \frac{x^2}{(x_{\text{max}}/2)^2}\right),
\]

where \(x_{\text{max}}\) is the sea travel range, \(x\) is the distance to each coastal node between 50 km and \(x_{\text{max}}\), and \(K_1\) is a normalization constant. Fig. S3a shows how the voyaging population is distributed according to this model. Since we assume that voyaging takes place in the form of cabotage, the sea-travel distances are computed along the coast as the cumulative distance between coastal cells (as indicated by the red line in Fig. S1).

Model 3. Voyaging with uniform probability to settle. In this model we assume that Neolithic individuals can travel by sea and relocate along the coast up to a certain range (e.g., 300 km), with the probability of settling not depending on distance (i.e., all cells within the prescribed range receive the same fraction of voyagers; see Fig. S3b). Again, the distance between origin and destinations is computed along the coast, as shown in Fig. S1.
For models allowing sea travel (models 2–4), voyaging takes place in an onward direction. This means that voyaging at the front takes place toward unknown territories, rather than toward areas already occupied by first farmers. Increasing the probability that the individuals who relocate travel by sea rather than by land does not decrease the minimum sea-travel range necessary to fit the archaeological pattern, although slightly lower growth rates might be necessary.

We assume that the Mesolithic populations are well established and stationary in terms of population numbers, so any population movement between hunter-gatherer bands would be balanced.

**Step 2: Population interaction**

When assuming a purely demic process, the Neolithic population advances following a dispersal-reproduction scheme, so we would not apply this interaction step. However, if we consider that the expansion was a mixed cultural-demic process, we have to include cultural transmission in the model. Here we consider two processes: cross-mating (vertical cultural transmission) and acculturation (horizontal/oblique cultural transmission). At each iteration, we compute the new population incorporated into the farmer community per cell as a result of the interaction applying the following equations.

**Vertical cultural transmission.** This relates to cross-mating between populations (23). It is characterized by a parameter $\eta$ (with values between 0 and 1) that indicates the degree of interbreeding. If $N$ is the number of farmers and $M$ the number of hunter-gatherers, then the number of hunter-gatherers that mate with farmers is (24)

$$I_{VT} = \eta \frac{N \cdot M}{N + M},$$

(S2)

Note that this equation is not exactly the same as in Ref. (24). Here we are computing the number of mixed couples (or individuals of each population that will mate with individuals of the other population), while the interaction term in Ref. (24) computed the new population contributed by the mixed couples after reproduction; here we compute the population growth in the next step, which leads to the same final output.
In principle the case $\eta > 1$ is also possible, but it would lead to more cross-matings than under random mating (25), which is not realistic for the case of farmers and hunter-gatherers according to ethnographic parallels (25, 26).

**Horizontal/oblique cultural transmission.** This relates to acculturation of hunter-gatherers (23). Accordingly, it accounts for the number of hunter-gatherers that are incorporated in the farming community, though not for mating reasons (in other words, they can be couples). The number of Mesolithic people incorporated in the farming population due to this process is (27)

$$I_{HT} = f \frac{N \cdot M}{N + \gamma M}, \quad (S3)$$

where $f$ indicates the intensity of acculturation and $\gamma$ the preference by hunter-gatherers to copy either farmers ($\gamma < 1$) or other hunter-gatherers ($\gamma > 1$). However, as shown analytically in (27, 28), the speed of the Neolithic wave of advance depends only on a single acculturation parameter, namely $C = f / \gamma$. This parameter involves the number of hunter-gatherers incorporated in the farming community per pioneer farmer (i.e., when $N \approx 0$) and per generation (or it could also indicate the number of hunter-gatherer couples incorporated per pioneer farming couple and per generation, since they are not incorporated for mating reasons).

Since the value of $C$ is highly uncertain, we compute the results of applying horizontal/oblique cultural transmission for three illustrative values of the acculturation rate: $C = 0.2$ (five farmers are needed for a hunter-gatherer to adopt farming per generation), $C = 1$ (a hunter-gatherer is converted per every farmer and generation) and $C = 10$ (every farmer converts 10 hunter-gatherers into farming per generation). In the model, however, we have to input the values of the parameters $f$ and $\gamma$ when applying Eq. (S3). We have checked that, as expected, for a single value of $C = f / \gamma$, any other pair of values for the parameters $f$ and $\gamma$ yields roughly the same results (in agreement with Refs. (27, 28)), and thus we fix $\gamma = 10$ for the purpose of modelling the effect of acculturation.

**Step 3: Population growth**

The final population in each cell after a generation (iteration) is computed by applying a reproduction step. The population growth process it computed for each cell of the grid using a reproduction coefficient $R_0 = \exp(aT)$, where $a$ is the intrinsic growth rate and $T$ the generation time, and limiting the growth to the maximum value of individuals per cell $N_{\text{max}}$ (recall that $N_{\text{max}} = 3200$ individuals). Therefore, the number of farmers in each coastal or inland cell after the population growth process, $N(t + T)$, is computed as follows

$$\begin{align*}
N(t + T) &= R_0 N(t) \quad \text{if } N < N_{\text{max}} \\
N(t + T) &= N_{\text{max}} \quad \text{if } N \geq N_{\text{max}},
\end{align*} \quad (S4)$$
where \( N \) is the number of individuals per cell after the dispersion and interaction processes. Since the final number of farmers in a cell must be an integer value, the results from Eq. (S4) are rounded to the nearest integer.

Other approaches are also valid, e.g., a logistic (i.e., quadratic rather than linear) function in the first line in Eq. (S4); however, Eq. (S4) does not yield negative population numbers when using finite-difference equations (as in the present study), in contrast with the logistic function \((15, 29)\). Alternatively, we could use the complete solution of the logistic equation; however, including a higher-order approximation does not affect significantly the speed of the front.

Since we assume, as a first approximation, that the Mesolithic populations are well established and in a stationary state in terms of population numbers, we do not consider net population growth for Mesolithic individuals.

**Interaction parameters**

**Acculturation parameter**

In the main paper (Fig. 4) we show the results obtained when vertical and horizontal cultural transmission are applied with \( \eta = 0.5 \) and \( C = 0.2 \). Increasing the values of \( C \) does not modify, however, the results significantly, as shown in Fig. S4 for \( C = 1 \) (Fig. S4a) and \( C = 10 \) (Fig. S4b). Therefore, the exact value of \( C \) does not seem to be very important.

**Size of Mesolithic populations**

All of the results that include interaction discussed so far (both here and in the main paper) have been computed for bands of 50 hunter-gatherers. Increasing the number of people per band to 80 (also consistent with the evidence for Tasmanian populations \((19)\)) allows for a faster spread, but very slightly, so that it does not change the conclusions reached. This is shown by the following figures \((S5–S6)\), where the diamonds indicating the goodness of fit with the prediction remain unchanged when compared with the results for bands of 50 people (Figs. 4 and S4).
Fig. S4: Effect of combined vertical (cross-mating) and horizontal/oblique (acculturation) cultural transmission on the expansion process. Results for model 4, with $\sigma=1.8\%$, $\eta=0.5$ with bands of 50 people, and for different acculturation levels: (a) $C=1$, and (b) $C=10$. We used a sea-travel range of 350 km. White diamonds represent areas reached within the calibrated range; black diamonds represent areas reached by the model later than the calibrated range.
**Fig. S5:** This figure is the same as Figs. 4b and 4c in the main paper, but it is based on bands of 80 individuals (instead of 50 people). It shows the effect of Neolithic-Mesolithic interaction on the rate of spread. The results are those of model 4, with a sea-travel range of 350 km and intrinsic growth rate $a=1.8\%$ for two cases: (a) vertical cultural transmission only ($\eta=0.5$, $C=0$, as in the case of Fig. 4b), and (b) combined vertical and horizontal cultural transmission ($\eta=0.5$, $C=0.2$, as in the case of Fig. 4c). White diamonds represent areas reached within the calibrated range and black diamonds those areas reached later than the calibrated range (the pattern of the colored diamonds is the same as in Figs. 4b,c, so the conclusions do not change).
**Fig. S6:** This figure is the same as Fig. S4 above, but taking a band to have 80 individuals (instead of 50). It shows the combined effect of vertical and horizontal/oblique cultural transmission on the rate of spread. The results are for model 4, with $\sigma=1.8\%$, $\eta=0.5$ with bands of 80 people, and for different acculturation levels: (a) $C=1$, and (b) $C=10$. In all cases, the sea travel range is 350 km. White diamonds represent areas reached within the calibrated range and black diamonds areas reached later than the calibrated range (the pattern of diamonds is again the same as in Fig. S4, so the conclusions do not change).
Appendix. Model

ing the West Mediterranean Neolithic spread

References


