

Research



Cite this article: Allaby RG, Stevens C, Lucas L, Maeda O, Fuller DQ. 2017 Geographic mosaics and changing rates of cereal domestication. *Phil. Trans. R. Soc. B* **372**: 20160429.
<http://dx.doi.org/10.1098/rstb.2016.0429>

Accepted: 2 June 2017

One contribution of 16 to a theme issue 'Process and pattern in innovations from cells to societies'.

Subject Areas:

evolution, plant science

Keywords:

archaeology, archaeobotany, niche construction, origins of agriculture, Near East, China

Author for correspondence:

Dorian Q. Fuller
e-mail: d.fuller@ucl.ac.uk

Electronic supplementary material is available online at <https://dx.doi.org/10.6084/m9.figshare.c.3883027>.

Geographic mosaics and changing rates of cereal domestication

Robin G. Allaby¹, Chris Stevens², Leilani Lucas², Osamu Maeda³
and Dorian Q. Fuller²

¹School of Life Sciences, University of Warwick, Warwick, UK

²Institute of Archaeology, UCL, London, UK

³Institute for Comparative Research in Human and Social Sciences, University of Tsukuba, Tsukuba, Japan

DQF, 0000-0002-4859-080X

Domestication is the process by which plants or animals evolved to fit a human-managed environment, and it is marked by innovations in plant morphology and anatomy that are in turn correlated with new human behaviours and technologies for harvesting, storage and field preparation. Archaeobotanical evidence has revealed that domestication was a protracted process taking thousands of plant generations. Within this protracted process there were changes in the selection pressures for domestication traits as well as variation across a geographic mosaic of wild and cultivated populations. Quantitative data allow us to estimate the changing selection coefficients for the evolution of non-shattering (domestic-type seed dispersal) in Asian rice (*Oryza sativa* L.), barley (*Hordeum vulgare* L.), emmer wheat (*Triticum dicoccon* (Shrank) Schübl.) and einkorn wheat (*Triticum monococcum* L.). These data indicate that selection coefficients tended to be low, but also that there were inflection points at which selection increased considerably. For rice, selection coefficients of the order of 0.001 prior to 5500 BC shifted to greater than 0.003 between 5000 and 4500 BC, before falling again as the domestication process ended 4000–3500 BC. In barley and the two wheats selection was strongest between 8500 and 7500 BC. The slow start of domestication may indicate that initial selection began in the Pleistocene glacial era.

This article is part of the themed issue 'Process and pattern in innovations from cells to societies'.

1. Introduction

Domestication is the transformation of other species through interactions with humans and human-managed environments, such that the reproductive success of those species and their productivity for human-valued traits, such as food value or fibre-production, both increased [1]. This can be regarded as a form of co-evolution [2–5], in which the human side of the relationship takes place primarily through cultural evolution, i.e. development of techniques and technologies of ecological management transmitted culturally, while the plants or animals evolve through normal biological evolution (genetic transmission). Both cultural innovations and evolutionary developments in plant morphology and physiology occurred through the domestication process, making both human cultivators and cultivated crops better adapted, and able to exist in increased population sizes and across a greater range of geography. This is increasingly discussed in terms of niche construction theory [4,6–8]. Agriculture represents a major shift of biosphere systems in which human societies, and their growing populations, increasingly became a global driving force in environmental change [7–9]. Thus understanding evolutionary processes involved in domestication is fundamental to biology not just as an example of biological evolution but also as a key transition in biosphere systems. In terms of human history, agricultural production ecologies based on domesticated species provided the basis for all historical and urban civilizations.

Despite the centrality of domestication to evolutionary studies since the time of Darwin [10], much has changed in our understanding of these processes in recent years through the accumulation of empirical evidence, including the sub-fossil record provided by archaeology and discoveries through molecular genetics [3,11–13]. Genetic evidence makes clear that not all ‘domestication-related traits’ were evolving during initial domestication. As such it is useful to separate true domestication traits, those traits that underwent directional selection when a species first entered an intensive co-evolutionary relationship with humans, from those traits relating to diversification or crop improvement, which were selected subsequently to domestication, often creating geographically distinct varieties. Relating to archaeology, the growth of archaeobotanical evidence has provided a sub-fossil record in which phenotypic changes relating to domestication can be directly documented over time and space, allowing rates of change to be estimated for different traits in different taxa and regions based on empirical evidence. The current study provides an updated analysis of such data for cereal domestication, in particular in terms of the trait of non-shattering ears or panicles.

Non-shattering ears or panicles are often regarded as the *sine qua non* feature of cereal domestication, as this prevents natural seed dispersal, a process necessary in the wild, leaving cereals reliant on farmers to harvest and plant their seeds [14,15]. Recent accumulation of archaeobotanical evidence has made it possible to document empirically changes in the proportions of human-reliant (i.e. domesticated) cereals in past populations during the era of agricultural origins for some cereals, including West Asian wheat (*Triticum monoccocum* L. and *Triticum dicoccon* (Shrank) Schübl.) and barley (*Hordeum vulgare* L.) [16–18]. Data have also become available for at least part of the trajectory in Chinese rice (*Oryza sativa* L.) domestication [12,19,20]. Contrary to previous expectations that selection through sickle harvesting should be strong leading to rapid domestication, of the order of a century (20–200 years) [21–23], the archaeobotanical evidence of wheat and barley indicates protracted transitions with documented change in cereal morphology recorded over about 3000 years [12,16,24]. This indicates that the strength of selection for domestication traits is orders of magnitude weaker than previously supposed [12,25,26]. It is also plausible that selection for domestication traits was not always directional, and may have fluctuated around a meta-stable equilibrium of partially non-shattering [27], maintained by periods of fallow, variation in harvesting techniques and continued bolstering of human cereal stores from wild populations [27,28]. At the level of the meta-population, which included fallow fields, and feral crops, as well as those under active cultivation and sympatric wild populations, cereals undergoing domestication may be expected to have included processes of ‘fluctuating selection’ [29,30].

Given that domestication was a protracted process there is no reason to assume that human practices, ecological conditions and, therefore, selection pressures were constant and unchanging over several millennia. To assume that selection pressures were constant implies that environmental parameters, such as those relating to climate, and human behaviours were unchanging over a period of millennia or 100–200 human generations. More reasonable is to accept that changing environmental conditions, social conditions and technology could have all contributed to accelerating or slowing down, or even reversing, selection processes on early cultivated cereals. Previous

attempts to estimate rates of phenotypic evolution for domestication traits and associated selection coefficients, however, have assumed a single constant rate of domestication, based on linear regression of data (figure 1*a,b*). This has produced estimates of rates of phenotypic change, based on *Darwin* or *Haldane* units, that are in the same order of magnitude as estimates from modern microevolutionary studies unrelated to domestication, or even on the slow side [12,25,31]; this was somewhat surprising given that it was often assumed that the anthropogenic nature of domestication might lead to faster evolution [10,21,22]. A weakness of this approach, however, is that the single estimate of evolutionary rate estimate is really an approximate average of what are presumably varying rates, as selection strength varied over the course of the domestication process.

The present contribution takes the analysis of these data closer to a realistic estimate of the changing rates of evolution and selection pressures over the course of a given domestication episode. It is expected that any evolutionary process of directional selection will approximate a sigmoidal (logistic) pattern of change [2,21]. In the present contribution we explore the potential to use any pair of time series data points to represent the specific sections of the sigmoidal process, more localized in space and over a shorter time scale (figure 1*c*). As such we are modelling from archaeobotanical data—the fossil record of crop domestication—how the strength of selection changed over the course of the entire domestication process, in other words, whether different sigmoidal processes were in action over time. This in turn allows us to identify particular periods when human-driven selection was stronger, and it also provides a baseline for estimating backwards in time to when the earliest selection for domestication traits began. In other words we are able to predict when the beginnings of the effects of pre-domestication activity occurred, e.g. the gathering of wild grasses, as well as identifying when cereal domestication intensified.

2. Material and methods

We have collected all of the available sub-fossil data for cereal non-shattering from archaeological plant assemblages in the putative geographical regions of origin that indicate change towards domestication traits for four crops, including Asian rice (*O. sativa*), barley (*H. vulgare*), einkorn wheat (*T. monoccocum*) and emmer wheat (*T. dicoccon*) (electronic supplementary material). In the current study, we have excluded indeterminate remains of wheat and barley as well as immature (green-harvested) spikelet bases of rice. Previous studies have used the proportions of uncertain specimens to estimate error margins and standard deviations [12,25], whereas we have taken a different approach to this to account for sample size (see below). Assemblage data are placed on an absolute time scale by calculation of the median age estimate of the archaeological phases [12,24]. All dates are given in calibrated BC.

These data are drawn from across multiple sites, which are taken as representative of an evolutionary process of domestication taking place over a broader region and period. This assumes that archaeobotanical assemblages represent a regional meta-population undergoing domestication over an extended period of time. This is necessary as no one archaeological site that has yet been discovered was occupied through the entire domestication process (approx. 4000 years), and, therefore, there is no single stratigraphic archive that has preserved an entire domestication process *in situ*. Where single sites do demonstrate increasing domestication traits, such as the rise in domesticated type rice spikelet bases during the 300–400 years of occupation

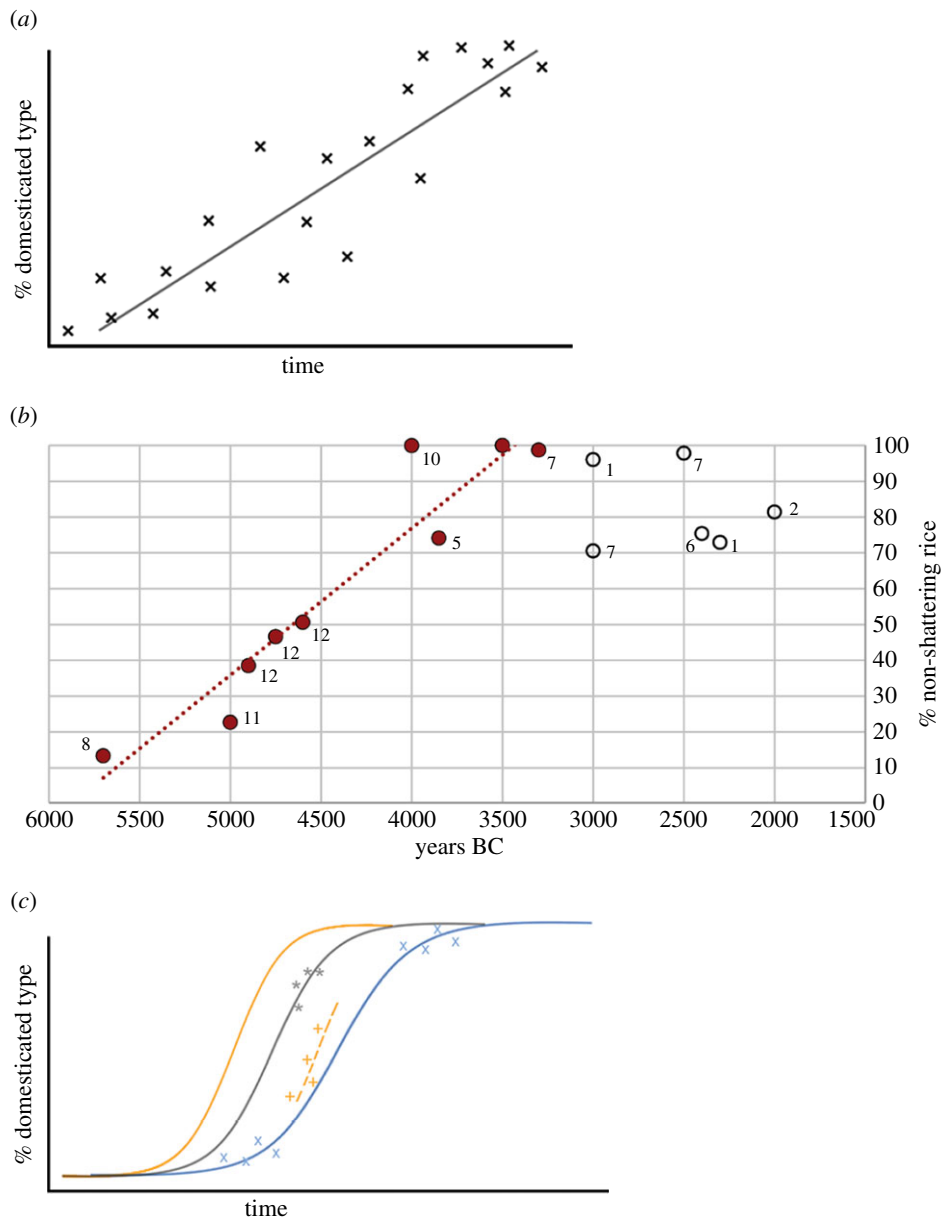


Figure 1. Models of domestication rate and rate estimates. (a) Hypothetical series of archaeobotanical data for non-shattering plotted against time with a least-squares linear regression as an estimate of rate of change [12,25,26]. (b) The example of rice spikelet bases from the Lower Yangtze and later northern China ($n = 30\,912$); data points numbered by sites in figure 2a (data: electronic supplementary material, table S1). (c) Hypothetical series of archaeobotanical data points forming selection chains fitted to closest logistic curves of differing rates (method developed in this paper).

at Tianluoshan, a gradual change is evident, similar to those from comparisons across sites [19]. As demonstrated by numerous case studies not related to domestication, geographical mosaics of co-evolution are recurrent features of evolution [32], and it can be suggested that both 'hotspots and 'cold-spots' of selection for domestication traits are likely to have occurred within the broader regions of agricultural origins [27,31]. Recently, genomic studies of some crops, including barley [33], emmer [34], einkorn [35] and Asian rice [36,37], indicate that modern domesticated populations derive in part from genetic diversity from numerous geographically disparate wild gene pools.

In the present analyses, time series data were modelled as selection chains, themselves inferred from a parsimonious model of selection and spread, making the populations represented by younger sites derivative of any older sites for which this is a plausible relationship. Selection chains were constructed by assuming that data points that are closer together in time, space and the proportion of non-shattering are more likely to be part of the same trajectory. For the most part we assumed that domestication is

directional and, therefore, that increasing percentages of non-shattering types are to be expected. We also considered the possibility that periods of negative selection could have taken place over shorter timescales within the overall domestication episode. This would be in keeping with the potential of fluctuating selection [29,30] or a period of semi-domestication [27]. This may have occurred during periods of fallow or changing practice, which could have led to some selection chains becoming disconnected in our analysis owing to falling frequencies of shattering over time. These episodes of negative selection are indicated by grey arrows in the figures below, and are discussed separately. Multiple potential pathways of evolution are possible and this allows us to increase the number of estimates of selection pressure over time and, therefore, look for general patterns in the strength of selection pressure.

We assume that it takes time for domesticates to spread between sites. In the case of early rice domestication in the Lower Yangtze, all sites are in fairly close proximity, lying within an elliptical area of approximately $65\,000\text{ km}^2$, accounting for evidence

dating between 6000 and 2000 BC. By contrast the greater 'Fertile Crescent' of southwestern Asia is much larger and requires an elliptical area of approximately 3.5 million km². In this region, therefore, some sites are quite distant from each other, and thus it is more likely that selection chains should be inferred for sites that are reasonably close to each other. In addition, material culture, such as lithic sickle typology, shows regional differences across Southwest Asia [24]. We therefore added an additional calculation to the construction of selection chains in wheat and barley by creating a time and space matrix that penalized longer distance and shorter time period inter-relationships.

This model considered two rate parameters, the rate with which frequencies change between two sites, and the velocity with which cultures would be required to move between sites. The rate of transition towards tough rachis dominance between sites was inferred through the tough rachis allele change between two sites represented by the selection coefficient (s). The minimum velocity (v_m) with which individuals would have to travel between sites in order for cultural continuity to be represented is estimated by utilizing the geographical distance divided by the maximum amount of time permitted by the archaeological dates. In the case of the Near East, the geographical distance was calculated with respect to viable routes that circumvented the central desert regions of the Near East by preferentially taking routes through the Fertile Crescent.

To calculate s we first had to infer the allele frequencies associated with phenotypic frequencies observed in the archaeological record. To do this, we calculated the inbreeding coefficient (F) associated with the mating system of the crops, which was assumed to be 98% inbreeding. The inbreeding coefficient was determined by simulation that began with Hardy–Weinberg proportions of genotypes P, R and Q (dominant homozygous, heterozygous and recessive homozygous genotypes, respectively) which were then mated with a 0.98 probability of selfing. The subsequent generation gave rise to new proportions of P, Q and R that deviated from Hardy–Weinberg, the reduction in heterozygosity being an estimate of F . Such that

$$F = \frac{R - q^2}{pq}, \quad (2.1)$$

where R is the proportion of heterozygotes, and p and q are dominant and recessive allele frequencies, respectively. Generations were iteratively produced, and F calculated each generation until it ceased to change by more than 0.00001, indicating the simulation had converged on the inbreeding equilibrium and a close approximation of F obtained. We assumed that tough rachis phenotypes represented homozygous recessive mutations, and, therefore, the Q value of genotypes, and associated with the q allele. Allele frequency was then determined by finding values of q that satisfy the equation

$$Q = q^2 + Fpq, \quad (2.2)$$

where Q is the observed proportion of tough rachis grains in the archaeobotanical record. To find s we simulated phenotypes of P, R and Q with proportions calculated from the determined allele frequencies and inbreeding coefficient beginning with the allele frequency of q associated with the oldest of a pair of archaeological sites. New generations were constructed as before, with P, R and Q genotypes mating with themselves with a 0.98 probability. Each generation the proportion Q was modified by s such that

$$Q' = Q(1 - s), \quad (2.3)$$

where Q' is the proportional value of Q after a round of selection. Note that s took negative values because the recessive homozygote is advantageous in this case. The simulations proceeded through a number of generations defined by the difference in age between the two archaeological sites under comparison. Values of s were then determined that coincided with the observed change in

tough rachis frequency between the two archaeological sites under comparison. Simulations to calculate s values from datasets of age and tough rachis proportions were carried out by a program written by R.G.A. (available on request). The simulations also recorded the date in generations that various frequencies would be reached ranging from 0.00001 to 1.

Selection chains were then determined from transition scores, which were calculated from the velocity of movement between sites and strength of selection. To emphasize the parsimony of slow movement and selection, the inverse of velocity and selection coefficient were used, such that a transition score (T_s) between two sites could be summarized as

$$T_s = \left(\frac{1}{v_m} \times \frac{1}{s} \right)^{1/2}, \quad (2.4)$$

which is the geometric mean of the inverses. The geometric mean is relatively robust to the different orders of magnitude associated with velocity and selection coefficient values; consequently in this model they are of equable influence on the transition score. Each archaeological locality had a series of transition scores relating to the other sites, and the highest scoring values were taken as the most likely connections of cultural continuity explained by a combination of the least selection and minimum travel speed required. However, under this model older sites will tend to become favoured over younger sites as likely points of cultural connection to later sites, even if associated with a stronger selection pressure, because of their tendency to be associated with lower velocities. To account for this we envisioned the possibility that sites could have existed over a longer time frame than the absolute dates suggest in the model and imputed a velocity that assumes all sites to be contemporaneous. In such a case the strength of selection is unaffected, because we are inferring what the frequency would have been at a particular time point, but the value s remains unchanged over time. The imputed minimum velocity (v_{mi}) was calculated as

$$v_{mi} = \frac{d}{c}, \quad (2.5)$$

where d and c are, respectively, the distance between sites and a constant, for which we used an arbitrary value of 2000. Consequently, the imputed transition score (T_{si}) was calculated as in equation (2.4), with v_i substituted for v_{mi} . Note that our choice of c has no effect on the relative magnitudes of T_{si} values because of the robustness of the geometric mean to differences in scale between terms.

Selection chains were then established by inferring most likely connections between sites based on the highest transition scores. We found that imputed transition scores down-weighted the influence of the oldest sites but made little difference to later connections between sites.

Selection chains defined which selection coefficients that had been calculated between archaeological sites should be retained. Ranges for selection coefficient values were generated by considering the 95% and 67% ranges of tough rachis proportions at sites, for their given sample sizes. These were obtained using the beta distribution, with parameters α and β defined as the counts of tough rachis and wild-type rachis plus one, respectively. The proportions for each pair of sites were then used to calculate values of s using the same program as above, highest and lowest values were taken as the range limits for 95% and 67%, respectively, and their associated age dates.

The selection chains were then used to determine which selection coefficients were associated with the domestication process, and particularly the earliest stages of the process. These had to be taken from tough rachis frequency values above zero found at the oldest archaeological sites, because in effect the true proportions at these sites is unknown. The earliest selection coefficients were then used to estimate an average early selection

coefficient. This value was then fed into the previous simulation as a fixed parameter to determine the ages at which various allele frequencies would be reached, given the allele frequency of a given site at a given time. We assumed a starting frequency for a new mutation to be 0.00001 (0.001%), which assumes an effective population size of 50,000, not unreasonable for cereals given genetic diversity estimates. However, ages associated with possible higher starting frequencies were also recorded.

3. Results

(a) Rice

Oryza sativa has long been a major staple of civilizations in East, South and Southeast Asia. Recent genomic data strongly support domestication of subspecies *japonica* from perennial *Oryza rufipogon*, with the closest living wild populations found in southern China [38,39]. However, former wild populations in central China extirpated over the last 1000 years are likely to be closer to the ancestor (figure 2a). Domesticated rice arrived at Shixia in Guangdong at ca 2600 BC [41]. Other Asian rice subspecies, notably subspecies *indica* and the *aus* group, have distinct geographical origins [37,39,40]. Archaeobotanical data points to the earliest rice cultivation in the Yangtze Basin, with two candidate regions around the Middle and Lower Yangtze Basin having the greatest support from current archaeobotanical data [20,42]. While selection for non-shattering may have begun in more than one part of the Yangtze and its tributaries the most completely documented domestication is that of the Lower Yangtze.

Rice domestication in the Lower Yangtze is represented by a clear archaeological sequence of assemblages that show growing proportions of non-shattering spikelet bases (figures 1b and 2b), rising from less than 15% at ca 5700 BC to 46.7% a millennium later. After this period we have the potential for divergent selection paths into later periods (figure 2b). Given that some subsequent sites have lower percentages, such as Liangzhu and sites on the Yellow River, it is possible that these represent dispersal northwards starting from populations in which this domestication trait was not yet fixed and that began dispersing northwards prior to the occupation of Caoyieshan. Allowing fluctuations of negative selection, however, provides a better explanation for the reduction in non-shattering in the later phases of Maoshan and for deriving the Sushui (Yangshao) from Caoyieshan (electronic supplementary material, table S3). These periods of selection reversal make sense as increased irrigation [43] would have provided habitat for wild or weedy rice.

Taking the Tianluoshan–Caoyieshan selection chain, the maximum selection coefficient reaches nearly 0.003 in the early 5th millennium BC (figure 2c). Taking 500-year bin averages places a peak selection strength of approximately 0.0033 at ca 4750 BC, with selection strength declining thereafter. This corresponds to the inference that Tianluoshan records a key tipping point in the Lower Yangtze domestication process of rice, with weaker selection under pre-domestication cultivation before this period, and slowing selection as the domestication process finished shortly after 4000 BC [19,37]. Based on our new calculations we estimate that non-shattering would have been as low as 0.00001 between 12 714 and 13 361 years before present (electronic supplementary material, tables S2 and S13).

(b) Barley

Hordeum vulgare was domesticated in Southwest Asia (figure 3a), drawing from geographically widely dispersed wild populations [33]. It is the best documented of the Near Eastern cereals in terms of the number and size of archaeobotanical rachis assemblages, which indicate almost purely wild-type, fully shattering populations before 9500 BC, with the gradual appearance of non-shattering rachides, rising to as much as 30% in a few assemblages, between 9000 and 8500 BC [18,24]. After 7400 BC some sites show apparently fixed, 100% non-shattering barley. As barley is widespread, occurring in all parts of the Fertile Crescent, data likely represent multiple regional domestication chains, and in accordance with this we find several probable distinct selection chains (figure 3b). We also checked potential negative selection fluctuations, but none of these had transition scores as high as for positive selection, indicating that they are less likely than the positive selection chains (electronic supplementary material, table S6). Barley selection coefficients increase from <0.002 to >0.003 between the early and late 9th millennium BC, with stronger selection (greater than 0.005) estimated for the Northern Levant (figure 3c). Selection coefficients appear to decrease from the later 8th millennium BC. Based on our new calculations we estimate that non-shattering would have been as low as 0.000 01 between 10 405 and 10 705 BC in the Northern Levant (based on Jerf el Ahmar), but between 17 881 and 21 792 BC, average ca 19 996 BC, in the Southern Levant (based on similar estimates from several early sites: electronic supplementary material, tables S5 and S13). There is some support in the data for two discernible barley domestications occurring in the Southern Levant, which are picked out by the selection chains associated with the Netiv Hagdud and Zahrat Adh-Dhra lineages.

(c) Emmer

Triticum dicoccon derives from wild populations in the Fertile Crescent (figure 4a). There are fewer studied and quantified assemblages of spikelet fork material from this species, with a total identifiable sample size of only 672. The earlier assemblages, up to 8300 BC, are all from the Southern Levant, and demonstrate approximately 26% non-shattering, with later assemblages from the north and east. Nevertheless, our spatio-temporal heatmap (electronic supplementary material, figure S5) indicates the likelihood of at least two selection chains in emmer (figure 4b). We tested for potential negative selection events, but none are likely (electronic supplementary material, table S9). At 8000 BC selection is close to the approximately 0.001 level that is typical of many of our estimates across crops. Although limited, data from northern and eastern sites may indicate an increase in selection on emmer shattering to greater than 0.004 in the 7th millennium BC. Based on our calculations we estimate the origins of selection, i.e. non-shattering at 0.00001, occurred between 18 346 and 25 606, with older part of this range favoured by the Southern Levant evidence (electronic supplementary material, tables S8 and S13).

(d) Einkorn

Triticum monococcum derives from wild populations with a more northerly distribution than emmer, and most of our data comes from northern Fertile Crescent sites, although einkorn reached at least as far south as Tell Qarassa (figure 4a).

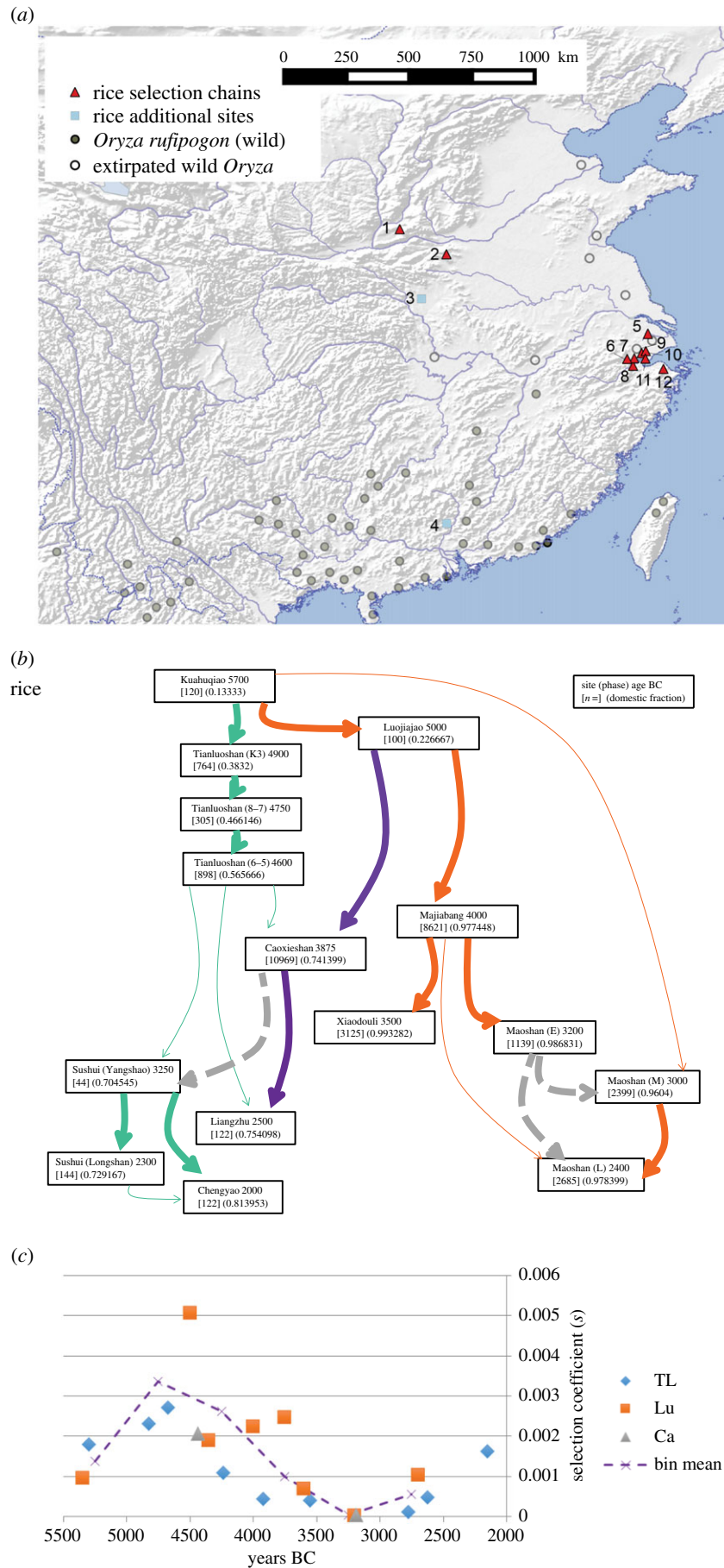


Figure 2. The evidence for the evolution of non-shattering Asian rice (*O. sativa*). (a) Map of evidence for rice domestication, indicating wild populations [40], and archaeological sites with spikelet base data, and additional sites: 1. Sushui valley, 2. Chengyao, 3. Baligang, 4. Shixia, 5. Caoxieshan, 6. Liangzhu, 7. Maoshan, 8. Kuahuqiao, 9. Luojiajiao, 10. Majiabang, 11. Xiaodouli, 12. Tianluoshan. (b) Selection chains for rice taking into account Lower Yangtze and central Chinese finds; negative selection events, dashed grey line. Boxes indicate site name, median age, sample size and proportion of non-shattering form. (c) Estimated selection coefficients (s) for different links in the selection chains in (b), and averages in 500-year bins; TL, Tianluoshan chain; Lu, Luojiajiao chain; Ca, Caoxieshan chain.

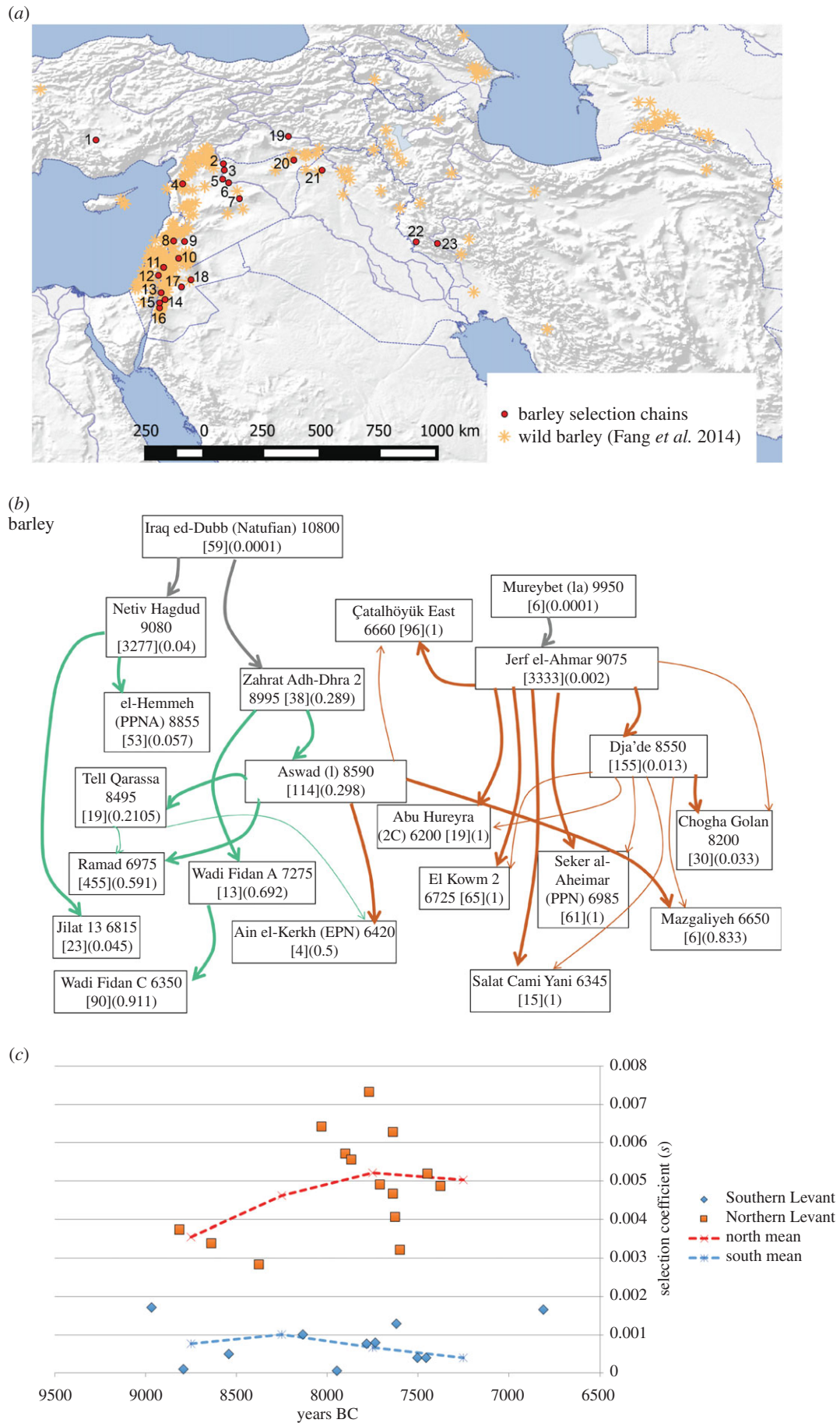


Figure 3. The evidence for the evolution of non-shattering in barley (*Hordeum vulgare*). (a) Map of evidence for barley domestication, indicating distribution of wild populations [44], and archaeological sites with rachis remains; sites numbered: 1. Çatalhöyük, 2. Dja'de, 3. Jerf el Ahmar, 4. Ain el-Kerkh, 5. Mureybet, 6. Abu Hureyra, 7. El Kowm 2, 8. Ramad, 9. Aswad, 10. Tell Qarassa, 11. Iraq ed-Dubb, 12. Netiv Hagdud, 13. Zahrat Adh-Dhra, 14. el-Hemmeh, 15. Wadi Fidan A, 16. Wadi Fidan C, 17. Jilat 7 and 13, 18. Azraq 31, 19. Salat Cami Yani, 20. Seker al-Aheimar, 21. Mazgaliyeh, 22. Chogha Golan, 23. East Chia Sabz. (b) Selection chains for barley. Boxes indicate site name, median age, sample size and proportion of non-shattering form. (c) Estimated selection coefficients (s) for different links in the selection chains in (b), indicating southern and northern averages for 500-year bins.

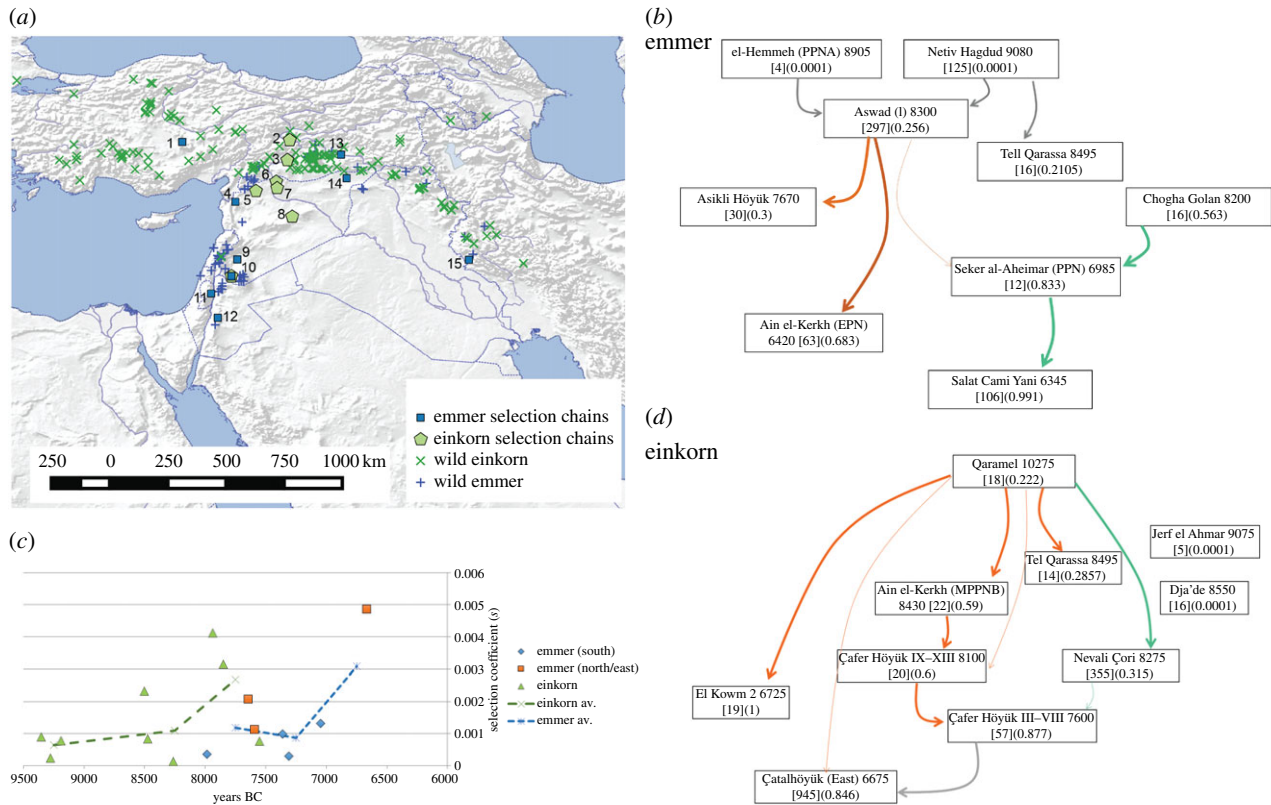


Figure 4. The evidence for the evolution of non-shattering in wheats. (a) Map of evidence for wheat domestication, indicating the distribution of wild *Triticum dicoccoides* [34], wild *Triticum boeoticum* [35], and archaeological sites with spikelet base remains; sites numbered: 1. Çatalhöyük, 2. Çafer Höyük, 3. Nevali Çori, 4. Ain el-Kherkh, 5. Qaramel, 6. Dja'de, 7. Jerf el Ahmar, 8. El Kowm 2, 9. Aswad, 10. Tell Qarassa, 11. Netiv Hagdud, 12. el-Hemmeh, 13. Salat Cami Yani, 14. Seker al-Aheimar, 15. Chogha Golan. (b) Selection chains for emmer and einkorn wheat, negative selection shown in grey in the case of einkorn. (c) Estimated selection coefficients (s) for emmer and einkorn wheats, showing averages in 500-year bins. (d) Selection chains for einkorn wheat. In (b) and (d), boxes indicate site name, median age, sample size and proportion of non-shattering form.

The total number of rachis remains is quite low ($n = 504$), and the earliest site, Qaramel, has surprisingly high levels of non-shattering, approximately 22%, prior to 10 000 BC, although dates have wide error margins and could be up to 1000 years later [24,26]. Later sites, including Jerf el Ahmar and Dja'de, appear to be entirely shattering, although cultivation is indicated at these sites by other lines of evidence [45]. This leads us to infer the main trajectory to domestication may not have involved this pair of sites (figure 4d). Setting aside those sites, the data indicate a coefficient s of approximately 0.001 up to 8250 BC which rises to approximately 0.003 after 8000 BC and a subsequent reduction in pressure (figure 4c). Çatalhöyük is best explained in our model as a result of a loss of selection after Çafer Höyük (electronic supplementary material, table S12). The levels of non-shattering reached at Qaramel lead to the inference of starting frequencies of 0.00001 pure wild (shattering) einkorn at approximately 30 000 BC, although estimates from later parts of the domestication chains provide start dates between *ca* 14 000 and *ca* 10 000 BC (electronic supplementary material, tables S11 and S13).

4. Discussion and conclusion

The evidence presented in this study supports the notion that the process of domestication was a slow evolutionary process. In general the selection coefficients recovered from this study corroborate the low values previously calculated for cereals using alternative methodologies [25], but contrast with later estimates [12].

An obvious question is why such low magnitudes for selection are seen within the archaeobotanical data when field experiments have demonstrated that higher rates would have been possible [21]? This study goes part way to addressing this question by separating out the strength of selection over time. A discernible change in the strength of selection occurs around 8000 BC in the Near Eastern cereals, which appears coincident with the rise of lithic technologies [24]. This shift in the strength of selection appears greatest in the Northern Levant, where a steep rise in selection coefficients is observed alongside the increase in group 3 and 4 sickle technologies [24]. Interestingly, the Southern Levant cultures do not appear to increase selection strength over time, implying a distinct difference in cultural practice between the two groups. Therefore, this study shows that different cultural practices led to different selection regimes that perhaps were not reflected in the one model system of field experiments. A second dimension to the strength of selection is in the amount of selection a population can endure. Selection comes at a cost, and while just a single trait that is under simple genetic control was considered in this study, in reality a number of genes would have been simultaneously under selection during the domestication process. An increasing number of genes under selection require that selection be proportionately weaker across loci, and the values of s found in this study tally closely with expectations of a selection cost model for the dozens of genes expected to have been affected by domestication [46].

A natural corollary of the observation that selection was weak and slow is that it is likely that the selective processes

driving crops down the domestication evolution trajectory extended back in time beyond our currently accepted dates for the first appearance of domesticated phenotypes. The assumed starting frequency in this study for a tough rachis mutant is 10-fold higher than in previous studies [47], and so provides a conservative estimate with regard to time depth for the onset of selection. Our findings are in line with previous archaeobotanical evidence which has suggested an early onset of cultivation practices in the Pleistocene at Ohalo II in the Near East [48] and in the Yangtze valley by the end of the Pleistocene or earliest Holocene [49,50]. The variable nature of the climate during this time period is likely to have led to dead ends as settlements and ways of life were abandoned. This is likely the case for Ohalo II, which appears to be associated with a selection chain independent from those included here. Similarly, even within the Holocene epoch the selection chains established in our model suggest that the Netiv Hagdud associated chains ultimately died out in favour of the ZAD2 associated chain for barley.

The complexity of these domestications, with varying strengths of selection at different points of time and across distinct geographical regions, suggests a number of different processes in action. The selection chains identify this structure in the data. The clearest separation occurs between northern and southern cultures in the Levant, and the Luojiajiao and Tianluoshan associated cultural lineages in the Lower Yangtze. In the case of emmer wheat these seem to reflect independent domestication events, and even

within the Southern Levant there is evidence of two separate lineages that may reflect parallel rises of domesticated forms.

The nature of the selective agents driving crops down the domestication trajectory further back into the Late Pleistocene is currently a matter of speculation. It may be the case that direct cultivation was not at all involved, and that the signal we detect is associated with more ancient impacts on the ecology surrounding humans before the late glacial maximum. Such impacts could have been harvesting or gathering pressures in the wild which simply reduced the strength of wild-type dispersal adaptations. However, such weak selection pressures require substantially large population sizes to counter the effects of drift. Consequently, the effects we detect should have been systematic and widespread, requiring a significant number of agents to apply the pressures involved. If the selection pressures were derived from quite simple human disturbances of the ecology around them such as might be associated with a natural herbivore community, then it may be the case that the time depth of signal we observe may relate to a period in time in which sufficient population densities had been reached in the Levant and Yangtze regions, respectively.

Data accessibility. The datasets supporting this article have been uploaded as part of the electronic supplementary material.

Competing interests. We declare we have no competing interests.

Funding. Research for and writing of this paper by D.Q.F., L.L. and C.J.S. has been carried out as part of the Comparative Pathways to Agriculture project, supported by an Advanced Investigator Grant from the ERC (no. 323842).

References

- Larson G *et al.* 2014 Current perspectives and the future of domestication studies. *Proc. Natl Acad. Sci. USA* **111**, 6139–6146. (doi:10.1073/pnas.1323964111)
- Rindos D. 1984 *The origins of agriculture. An evolutionary perspective*. London, UK: Academic Press.
- Purugganan MD, Fuller DQ. 2009 The nature of selection during plant domestication. *Nature* **457**, 843–848. (doi:10.1038/nature07895)
- Zeder MA. 2015 Core questions in domestication research. *Proc. Natl Acad. Sci. USA* **112**, 3191–3198. (doi:10.1073/pnas.1501711112)
- Allaby RG, Kistler L, Gutaker RM, Ware R, Kitchen JL, Smith O, Clarke AC. 2015 Archaeogenomic insights into the adaptation of plants to the human environment: pushing plant–hominin co-evolution back to the Pliocene. *J. Hum. Evol.* **79**, 150–157. (doi:10.1016/j.jhevol.2014.10.014)
- Smith BD. 2011 General patterns of niche construction and the management of wild plant and animal resources by small-scale pre-industrial societies. *Phil. Trans. R. Soc. B* **366**, 836–848. (doi:10.1098/rstb.2010.0253)
- Ellis EC. 2015 Ecology in an anthropogenic biosphere. *Ecol. Monogr.* **85**, 287–331. (doi:10.1890/14-2274.1)
- Boivin NL, Zeder MA, Fuller DQ, Crowther A, Larson G, Erlandson JM, Denham T, Petraglia MD. 2016 Ecological consequences of human niche construction: examining long-term anthropogenic shaping of global species distributions. *Proc. Natl Acad. Sci. USA* **113**, 6388–6396. (doi:10.1073/pnas.1525200113)
- Ruddiman WF *et al.* 2016 Late Holocene climate: natural or anthropogenic? *Rev. Geophys.* **54**, 93–118. (doi:10.1002/2015RG000503)
- Darwin C. 1868 *The variation of animals and plants under domestication*. London, UK: John Murray.
- Meyer RS, Purugganan MD. 2013 Evolution of crop species: genetics of domestication and diversification. *Nat. Rev. Genet.* **14**, 840–852. (doi:10.1038/nrg3605)
- Fuller DQ, Denham T, Arroyo-Kalina M, Lucas L, Stevens CJ, Qin L, Allaby RG, Purugganan MD. 2014 Convergent evolution and parallelism in plant domestication revealed by an expanding archaeological record. *Proc. Natl Acad. Sci. USA* **111**, 6147–6152. (doi:10.1073/pnas.1308937110)
- Larson G, Fuller DQ. 2014 The evolution of animal domestication. *Annu. Rev. Ecol. Evol. Syst.* **45**, 115–136. (doi:10.1146/annurev-ecolsys-110512-135813)
- Zohary D. 1969 The progenitors of wheat and barley in relation to domestication and agricultural dispersal in the Old World. In *The domestication and exploitation of plants and animals* (eds PJ Ucko, GW Dimbleby), pp. 47–66. London, UK: Duckworth.
- Harlan JR, De Wet MJM, Price EG. 1973 Comparative evolution of cereals. *Evolution* **27**, 311–325. (doi:10.2307/2406971)
- Tanno K-I, Willcox G. 2012 Distinguishing wild and domestic wheat and barley spikelets from early Holocene sites in the Near East. *Veg. Hist. Archaeobot.* **21**, 107–115. (doi:10.1007/s00334-011-0316-0)
- Weide A, Riehl S, Zeidi M, Conard NJ. 2015 Using new morphological criteria to identify domesticated emmer wheat at the aceramic Neolithic site of Chogha Golan (Iran). *J. Archaeol. Sci.* **57**, 109–118. (doi:10.1016/j.jas.2015.01.013)
- Arranz-Otaegui A, Colledge S, Zapata L, Teira-Mayolini LC, Ibáñez JJ. 2016 Regional diversity on the timing for the initial appearance of cereal cultivation and domestication in southwest Asia. *Proc. Natl Acad. Sci. USA* **113**, 14 001–14 006. (doi:10.1073/pnas.1612797113)
- Fuller DQ, Qin L, Zheng Y, Zhao Z, Chen X, Hosoya LA, Sun GP. 2009 The domestication process and domestication rate in rice: Spikelet bases from the Lower Yangtze. *Science* **323**, 1607–1610. (doi:10.1126/science.1166605)
- Deng ZH, Qin L, Gao Y, Weisskopf A, Zhang C, Fuller DQ. 2015 From early domesticated rice of the Middle Yangtze Basin to millet, rice and wheat agriculture: archaeobotanical macro-remains from Baligang, Nanyang Basin, central China

- (6700–500 BC). *PLoS ONE* **10**, e0139885. (doi:10.1371/journal.pone.0139885)
21. Hillman GC, Davies MS. 1990 Domestication rates in wild-type wheats and barley under primitive cultivation. *Biol. J. Linn. Soc.* **39**, 39–78. (doi:10.1111/j.1095-8312.1990.tb01611.x)
 22. Harlan, JR. 1992 *Crops and Man*, 2nd edn. Madison, WI: American Society for Agronomy.
 23. Zohary D. 2004 Unconscious selection and the evolution of domesticated plants. *Econ. Bot.* **58**, 5–10. (doi:10.1663/0013-0001(2004)058[0005:USATEO]2.0.CO;2)
 24. Maeda O, Lucas L, Silva F, Tanno K-I, Fuller DQ. 2016 Narrowing the harvest: increasing sickle investment and the rise of domesticated cereal agriculture in the Fertile Crescent. *Quat. Sci. Rev.* **145**, 226–237. (doi:10.1016/j.quascirev.2016.05.032)
 25. Purugganan MD, Fuller DQ. 2011 Archaeological data reveal slow rates of evolution during plant domestication. *Evolution* **65**, 171–183. (doi:10.1111/j.1558-5646.2010.01093.x)
 26. Fuller DQ, Willcox G, Allaby RG. 2012 Early agricultural pathways: moving outside the ‘core area’ hypothesis in Southwest Asia. *J. Exp. Bot.* **63**, 617–633. (doi:10.1093/jxb/err307)
 27. Fuller DQ, Allaby RG, Stevens C. 2010 Domestication as innovation: the entanglement of techniques, technology and chance in the domestication of cereal crops. *World Archaeol.* **42**, 13–28. (doi:10.1080/004382409034296800)
 28. Willcox G. 1999 Agrarian change and the beginnings of cultivation in the Near East: evidence from wild progenitors, experimental cultivation and archaeobotanical data. In *The prehistory of food. Appetites for change* (eds C Gosden, J Hather), pp. 478–500. London, UK: Routledge.
 29. Bell G. 2010 Fluctuating selection: the perpetual renewal of adaptation in variable environments. *Phil. Trans. R. Soc. B* **365**, 87–97. (doi:10.1098/rstb.2009.0150)
 30. Thompson JN. 2013 *Relentless evolution*. Chicago, IL: University of Chicago Press.
 31. Fuller DQ, Asouti E, Purugganan MD. 2012 Cultivation as slow evolutionary entanglement: comparative data on rate and sequence of domestication. *Veg. Hist. Archaeobot.* **21**, 131–145. (doi:10.1007/s00334-011-0329-8)
 32. Thompson JN. 2005 *The geographic mosaic of coevolution*. Chicago, IL: University of Chicago Press.
 33. Poets AM, Fang Z, Clegg MT, Morrell P. 2015 Barley landraces are characterized by geographically heterogeneous genomic origins. *Genome Biol.* **16**, 173. (doi:10.1186/s13059-015-0712-3)
 34. Civaň P, Ivaničová Z, Brown TA. 2013 Reticulated origin of domesticated emmer wheat supports a dynamic model for the emergence of agriculture in the fertile crescent. *PLoS ONE* **8**, e81955. (doi:10.1371/journal.pone.0081955)
 35. Kilian B, Özkan, H, Walther A, Kohl J, Dagan T, Salamini F, Martin W. 2007 Molecular diversity at 18 loci in 321 wild and 92 domesticate lines reveal no reduction of nucleotide diversity during *Triticum monococcum* (einkorn) domestication: implications for the origin of agriculture. *Mol. Biol. Evol.* **24**, 2657–2668. (doi:10.1093/molbev/msm192)
 36. Londo JP, Chiang YC, Hung KH, Chiang TY, Schaal BA. 2006 Phylogeography of Asian wild rice, *Oryza rufipogon*, reveals multiple independent domestications of cultivated rice, *Oryza sativa*. *Proc. Natl Acad. Sci. USA* **113**, 9578–9583. (doi:10.1073/pnas.0603152103)
 37. Civaň P, Craig H, Cyman J, Brown TA. 2015 Three geographically separate domestications of Asian rice. *Nat. Plants* **1**, 15164. (doi:10.5281/zenodo.20740)
 38. Huang XH *et al.* 2012 A map of rice genome variation reveals the origin of cultivated rice. *Nature* **490**, 497–501. (doi:10.1038/nature11532)
 39. Choi JY, Platts AE, Fuller DQ, Hsing Y-L, Wing RA, Purugganan MD. 2017 The rice paradox: multiple origins but single domestication in Asian rice. *Mol. Biol. Evol.* **2017**, msx049. (doi:10.1093/molbev/msx049)
 40. Fuller DQ, Sato YI, Castillo C, Qin L, Weisskopf AR, Kingwell-Banham EJ, Song J, Ahn S-M, Etten J. 2010 Consilience of genetics and archaeobotany in the entangled history of rice. *Archaeol. Anthropol. Sci.* **2**, 115–131. (doi:10.1007/s12520-010-0035-y)
 41. Yang XY, Wang WW, Zhuang YJ, Zhao L, Ma ZK, Ma YC, Cui Y, Wei J, Fuller DQ. 2016 New radiocarbon evidence on early rice consumption and farming in South China. *Holocene* **2016**, 1–7. (doi:10.1177/0959683616678465)
 42. Silva F, Stevens CJ, Weisskopf A, Castillo C, Qin L, Bevan A, Fuller DQ. 2015 Modelling the geographical origin of rice cultivation in Asia using the rice archaeological database. *PLoS ONE* **10**, e0137024. (doi:10.1371/journal.pone.0137024)
 43. Weisskopf A, Qin L, Ding J, Ding P, Sun G, Fuller DQ. 2015 Phytoliths and rice: from wet to dry and back again in the Neolithic Lower Yangtze. *Antiquity* **89**, 1051–1063. (doi:10.15184/aqy.2015.94)
 44. Fang Z, Gonzales AM, Clegg MT, Smith KP, Muehlbauer GJ, Steffenson BJ, Morelle PL. 2014 Two genomic regions contribute disproportionately to geographic differentiation in wild barley. *Genes Genomes Genet.* **4**, 1193–1203. (doi:10.1534/g3.114.01056)
 45. Willcox G, Fornite S, Herveux L. 2008 Early Holocene cultivation before domestication in northern Syria. *Veget. Hist. Archaeobot* **17**, 313–325. (doi:10.1007/s00334-007-0121-y)
 46. Allaby RG, Kitchen JL, Fuller DQ. 2016 Surprisingly low limits of selection in plant domestication. *Evol. Bioinform.* **11**, 41–51. (doi:10.4137/EB0.S33495)
 47. Ladizinsky G. 1998 *Plant evolution under domestication*. New York, NY: Kluwer Academic Publishers.
 48. Snir A, Nadel D, Groman-Yaroslavski I, Melamed Y, Sternberg M, Bar-Yosef O, Weiss E. 2015 The origin of cultivation and proto-weeds, long before Neolithic farming. *PLoS ONE* **10**, e0131422. (doi:10.1371/journal.pone.0131422)
 49. Zhijun Z. 1998 The Middle Yangtze region in China is one place where rice was domesticated: phytolith evidence from the Diaotonghuan Cave, Northern Jiangxi. *Antiquity* **72**, 885–897. (doi:10.1017/S0003598X00087524)
 50. Ma Y *et al.* 2016 Rice bulliform phytoliths reveal the process of rice domestication in the Neolithic Lower Yangtze River region. *Quat. Int.* **426**, 126–132. (doi:10.1016/j.quaint.2016.02.030)