

## Review



**Cite this article:** Kandler A, Powell A. 2018 Generative inference for cultural evolution. *Phil. Trans. R. Soc. B* **373**: 20170056. <http://dx.doi.org/10.1098/rstb.2017.0056>

Accepted: 21 December 2017

One contribution of 16 to a theme issue 'Bridging cultural gaps: interdisciplinary studies in human cultural evolution'.

**Subject Areas:**  
theoretical biology

**Keywords:**  
cultural evolution, social learning, generative inference, approximate Bayesian computation

**Author for correspondence:**  
Anne Kandler  
e-mail: [anne.kandler@eva.mpg.de](mailto:anne.kandler@eva.mpg.de)

## Generative inference for cultural evolution

Anne Kandler<sup>1</sup> and Adam Powell<sup>2,3</sup>

<sup>1</sup>Department of Human Behavior, Ecology and Culture, Max Planck Institute for Evolutionary Anthropology, 04103 Leipzig, Germany

<sup>2</sup>Department of Linguistic and Cultural Evolution, Max Planck Institute for the Science of Human History, 07745 Jena, Germany

<sup>3</sup>Department of Archaeogenetics, Max Planck Institute for the Science of Human History, 07745 Jena, Germany

AK, 0000-0003-3766-6597

One of the major challenges in cultural evolution is to understand why and how various forms of social learning are used in human populations, both now and in the past. To date, much of the theoretical work on social learning has been done in isolation of data, and consequently many insights focus on revealing the learning processes or the distributions of cultural variants that are expected to have evolved in human populations. In population genetics, recent methodological advances have allowed a greater understanding of the explicit demographic and/or selection mechanisms that underlie observed allele frequency distributions across the globe, and their change through time. In particular, generative frameworks—often using coalescent-based simulation coupled with approximate Bayesian computation (ABC)—have provided robust inferences on the human past, with no reliance on *a priori* assumptions of equilibrium. Here, we demonstrate the applicability and utility of generative inference approaches to the field of cultural evolution. The framework advocated here uses observed population-level frequency data directly to establish the likely presence or absence of particular hypothesized learning strategies. In this context, we discuss the problem of equifinality and argue that, in the light of sparse cultural data and the multiplicity of possible social learning processes, the exclusion of those processes inconsistent with the observed data might be the most instructive outcome. Finally, we summarize the findings of generative inference approaches applied to a number of case studies.

This article is part of the theme issue 'Bridging cultural gaps: interdisciplinary studies in human cultural evolution'.

## 1. Introduction

Understanding how human populations acquire, and use, social information is one of the central challenges of cultural evolution and the focus of a highly active, interdisciplinary debate [1]. Social learning, or cultural transmission, is defined as learning that is facilitated by observations of, or interactions with, another individual or their cultural products [2,3]. It supports the spread of adaptive information, accumulated over generations [4–6], yet also bears the risk of transmitting outdated, misleading or inappropriate information, especially in changing environmental conditions [7]. But there is no unique way in which social information can be acquired; in fact, a large number of social learning processes have been identified in the literature (e.g. [5,6,8]). Research aimed at identifying learning processes in human populations can be roughly divided into two groups: experimental laboratory-based and theoretical modelling-based approaches. Laboratory-based experiments, in particular 'microsocieties' (e.g. [9–12]) and diffusion chain experiments (e.g. [13–16]), have focused on uncovering the variety and subtlety of human social learning strategies, providing a powerful framework for studying cultural evolution empirically (see [1,17] for comprehensive review).

In the following, we focus on theoretical modelling-based approaches. These evolutionary models of learning have mainly focused on understanding which individual and social learning strategies are expected to have evolved in

spatially and temporally changing environments (see [18] for a comprehensive review of this literature). These models provided an elegant characterization of the long-term outcomes of evolution through natural selection, as well as their associated evolutionary trajectories, and therefore produce predictions of which learning processes are expected to be present in the population.

However, in order to verify those predictions, social learning processes would need to be observed directly so that fine-grained individual-level data detailing who learns from whom can be generated. But outside of controlled experimental conditions, large longitudinal datasets of this kind are difficult to obtain, especially in historical or anthropological contexts (e.g. [19], but see [20,21] for two cultural evolutionary case studies and [22] for a research program dedicated to addressing this issue). This is not to say that no such data exist, but in many case studies of interest the available data are in the form of frequencies of different variants of a cultural trait in the population at one or several points in time. While many modern datasets possess a rich temporal resolution (e.g. those describing the choice of first names in modern populations, which record the number of instances of a specific name each year), prehistoric or anthropological datasets—the focus of this paper—typically describe the frequencies of different cultural variants in sparse samples from the whole population.

So if we want to infer social learning processes from available data, we face a classical inverse problem: we can only observe aggregated, population-level frequency data but aim at identifying the underlying individual-level learning processes that gave rise to them. Recent approaches to address this inverse problem have focused on, among other things: the shape of adoption curves (e.g. [4,23,24]); the comparison between observed levels of cultural diversity or cultural accumulation and the ones expected under various processes of social learning (in particular unbiased transmission (or neutral evolution) (e.g. [25–28])); the shape of rank-abundance distributions (e.g. [29–31]); and the comparison between observed turnover rates and the ones expected under unbiased transmission (e.g. [32,33]) or phylogenetic analyses (e.g. [34]). This research has clearly shown that robustly inferring the underlying processes of social learning from population-level frequency data becomes a challenging task, especially in the light of equifinality, i.e. in situations where various learning processes can result in very similar population-level characteristics (e.g. [35,36]).

However, the inverse problem described above is of course not unique to cultural evolution. In fact, other scientific fields have successfully overcome similar challenges, in particular population genetics, which aims to understand the evolutionary mechanisms that produced the allele frequency distributions observed both now and in the past. Here, recent developments have provided elegant means for building complex evolutionary models, and allowing the application of efficient generative inference frameworks, which made possible the statistical testing of increasingly realistic demographic hypotheses. In general, the generative approach proceeds by building a fully specified probabilistic model, in which the hypothesized causal mechanisms are explicitly defined. This model is then used to repeatedly simulate pseudo-datasets under known parameter values, such that their expected distribution can be statistically compared with observed data, through techniques such as approximate Bayesian

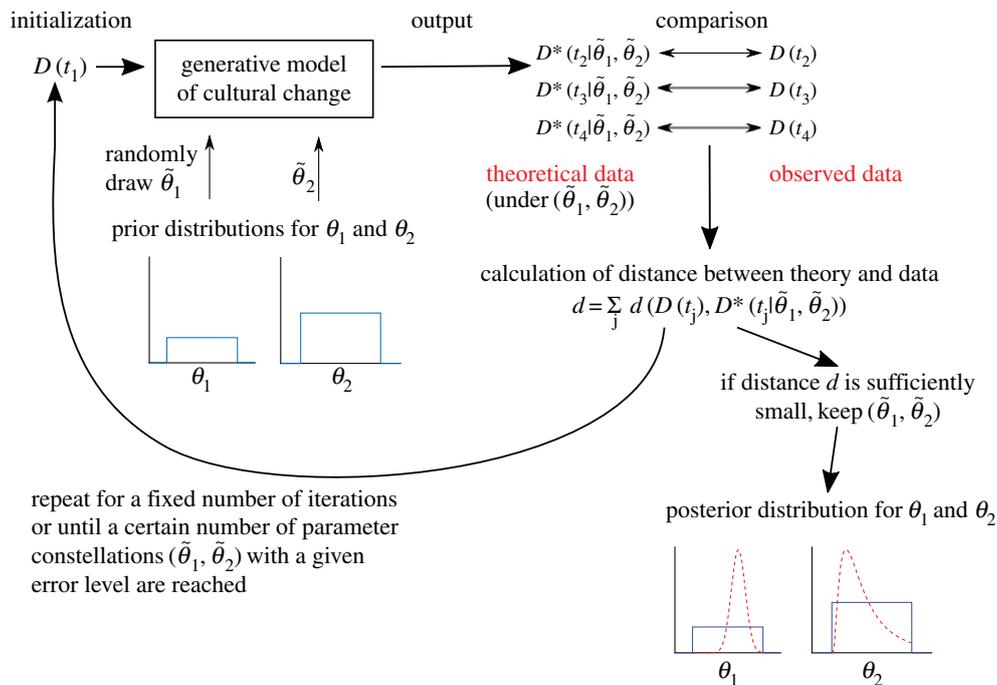
computation (ABC). This comparison allows certain hypothesized mechanisms to be rejected as inconsistent with the empirical data, and the estimation of model parameters that provide the best fit.

Our goal in this paper is to demonstrate how the generative inference approach can help answer the question of how human populations use social information, based on observable empirical evidence. We note that the general idea of generative modelling has already been applied to socio-cultural evolution. Significant early examples include Schelling's segregation model [37] and the influential agent-based economic modelling framework of *Sugarscape* [38]. These approaches and subsequent generalizations (see e.g. [39]) investigated the effects of explicitly defined individual-level causal mechanisms on population-level outcomes, which could then be compared with observed data. While one of the major advantages of this line of work is that the complex nature of the models considered allowed for more realistic expected outcomes, the principal limitation has been the lack of a robust statistical methodology capable of comparing these outcomes to empirical data. However, careful application of techniques like ABC, as mentioned above, is beginning to remove this limitation to inference (e.g. [40]).

We believe that the generative inference approach reviewed in this paper may link theoretical and empirical work in cultural evolution closer together by providing a framework that is able to evaluate the consistency between different individual-level processes and observed population-level patterns; in our case, between different processes of social learning and observed patterns of cultural change. Similarly to population genetics, such an inference framework consists of building a generative model that establishes a causal link between individual-level learning processes and observable population-level frequency data that then are evaluated for statistical consistency. The outcome of this approach is not only the identification of the most likely underlying learning process given the empirical data but a description of the breadth of processes that could have produced these data equally well, which in turn can be interpreted as an informal measure of the level of equifinality. Additionally, the inference framework may provide insight into the temporal and/or spatial resolution of the population-level frequency data that are needed to reliably distinguish between different processes of social learning (see [41] for a related discussion). In §1a, we briefly review some of the relevant key developments in population genetics, before exploring the applicability of the generative inference approach to cultural frequency data in §2.

### (a) Population genetics and generative inference

Classical population genetics—with its *prospective* approach [42]—provided many important theoretical insights into how the processes of mutation, drift, selection, migration and demographic change may shape the genetic variation expected in a population at equilibrium (e.g. [43–45]). But the development of coalescent theory in the early 1980s [46] (see also [47–49]) offered an alternative *retrospective* view of genetic evolution, providing a statistical model for the genealogical relationships between just a sample of individuals rather than the entire population. One major advantage of this coalescent framework is that, given an explicit model of



**Figure 1.** Schematic representation of the proposed generative inference framework. This non-equilibrium framework requires multiple observation (i.e. at least two) of cultural data  $D(t_j)$ , in our case population-level frequencies of different cultural variant types, at known times  $t_j$ .  $D^*(t_j | \tilde{\theta}_1, \tilde{\theta}_2)$  denotes the theoretical data produced under the social learning process described by  $(\tilde{\theta}_1, \tilde{\theta}_2)$ . The generative model is initialized with the data observed at the beginning of the time series  $t_1$ . (Online version in colour.)

demographic history and a mutation model, it allows for very efficient simulation of genetic—or genomic-scale—data for an observed sample with no *a priori* assumption of equilibrium. This has proved very useful in inferring population history, and while there is a wide array of other methodological approaches (e.g. [50–55]), the generative approach—in which simulated genetic data are statistically compared to the observed data—is growing in popularity, with the models of demographic history becoming increasingly complex and realistic (e.g. [56–58]).

However, generative inference crucially relies on the ability to make an evaluation of the quality of the model used. Rather than simply rejecting those demographic models or hypotheses that generate genetic variation inconsistent with what we observe (as in [59,60]), there exists a large and growing body of statistical techniques that allow for the explicit comparison of competing scenarios and the estimation of their underlying parameters. One such approach, ABC [61,62], was developed by statistical and population geneticists to circumvent the difficulty, or impossibility, of specifying the likelihood functions for complex models. ABC relies on repeatedly simulating pseudo-data under an explicitly specified model and, by retaining just those parameter values that generate data ‘close’ to the observed data, allows estimation of their posterior distributions (full details are given in §2b). A number of researchers have used this pairing of coalescent-based simulation and ABC to answer diverse questions about human demographic history, from early population differentiation in sub-Saharan Africa [63], to the global expansion of modern humans during the Late Pleistocene [58], to hunter–gatherer population replacement in Europe [64] and the initial colonization of the Americas [57] at the end of the last Ice Age.

## 2. Generative inference for cultural evolution

In the following, we demonstrate how generative inference procedures can be constructed and used to infer social learning processes from cultural data in the form of time-series detailing the usage or occurrence frequencies of different cultural variants. Similar to the population genetic applications, the inference procedure consists of two steps. First, we develop a non-equilibrium generative model capturing the main cultural and demographic dynamics of the considered system. This model describes the frequency evolution of different cultural variants present in a population at given time points under an assumed social learning hypothesis. Second, ABC techniques are used to derive conclusions about which (mixtures of) learning strategies are consistent with the observable frequency data and which are not. The aim of this framework is to allow researchers to ‘reverse engineer’, which learning strategies are likely to have been used in current or past populations, given knowledge of how frequencies have changed over time, independent of optimality or equilibrium assumptions. Figure 1 summarizes the steps of the generative inference framework described in this section.

We stress that this particular inference framework is designed to analyse the temporal dynamic of cultural change, defined as the change in frequency of different variants of cultural traits. If the observed data are of a different nature, e.g. describing the continuous variation of certain attributes of cultural artefacts, such as the dimensions of an arrowhead, then researchers have to first construct a hypothesis about the relationship between temporal variation of the attribute and the social learning processes considered in order to apply a similar inference procedure.

In §2a,b we describe the two steps of the inference framework and discuss in §2c the theoretical limits to inference; specifically, we ask how much information about underlying social learning processes we should expect to infer from population-level frequency data of a given temporal resolution. In §2d(ii) we show how the generative approach has been applied to cultural case studies. Lastly, in §2e we discuss some issues researchers should consider before applying the proposed, or a similar, inference framework.

### (a) Generative model

As mentioned above, the generative model aims at capturing the main cultural and demographic dynamics of the cultural system. Importantly, the generative model has to produce pseudo-data—in our case, population-level frequencies of different variants of a cultural trait at different points in time conditioned on the assumed social learning process—so that theoretical predictions can be compared to empirical observations. Thereby different learning processes are expressed by different model parameterizations; the model parameters are denoted by  $\theta = (\theta_1, \dots, \theta_k)$  in the following. In other words, the generative model establishes an explicit causal relationship between the assumed processes of social learning defined by  $\theta$  and observable population-level patterns of cultural change.

We note that there are no restrictions on the type of generative model used. Models ranging from systems of partial differential equations to agent-based simulations have also been used successfully; in fact, a number of the models mentioned in §1 could, with an appropriate choice of generative model, feasibly be adapted for use within this inference framework. As we want to generate frequency data at different time points, we advocate the use of non-equilibrium models, which can also account for temporal changes in demographic properties of the cultural system (e.g. variations in the total size of the population of cultural variants). This modelling choice aims at reducing the risk of misinterpreting non-equilibrium dynamics as evidence for the presence or absence of particular social learning processes (see [65] for a detailed discussion). For instance, the rejection of the hypothesis of neutral cultural evolution, based on empirical data, has usually been interpreted as evidence for the existence of selective biases in the population. But it has been pointed out that such a rejection can also be indicative of non-equilibrium dynamics or simply violations of the inherent assumptions of the neutral model (e.g. [28,66]). We note, however, that the relaxing of the equilibrium assumption requires accurate knowledge about e.g. the time points at which the observed frequencies are recorded. We return to this issue in §2e.

### (b) Statistical inference

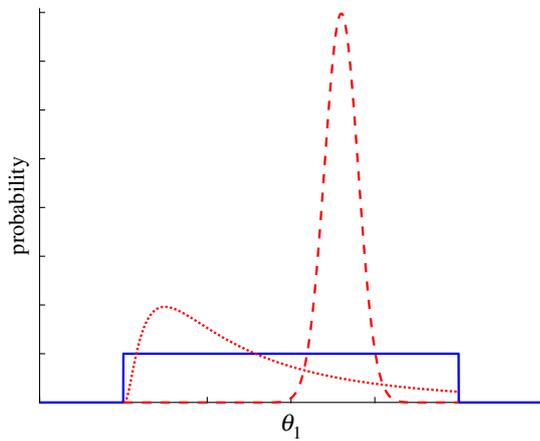
To infer which learning strategies are consistent with the observed data we would ideally determine the likelihood function of the generative model. However, in many cases (if not most in reality) the likelihood functions cannot be determined easily. As introduced in §1a, ABC [61,62] was developed to circumvent this difficulty. Given observed data  $D$ , this likelihood-free approach directly approximates the joint posterior density of the model parameters  $P(\theta|D)$ . It does this through repeatedly simulating data  $D^*$  under a generative model with parameter values drawn from their

prior distributions  $P(\theta)$ . These prior distributions describe the possible values that the parameter can assume or summarize all prior knowledge researchers may have. Retaining those parameter sets that generate data sufficiently ‘close’ to the observed data  $D$ , and rejecting the rest, results in a random sample from the distribution  $P(\theta|d(D, D^*) \leq \varepsilon)$ , where  $d(\cdot, \cdot)$  is a distance metric between the observed and simulated data, and  $\varepsilon$  is a tolerance level determining the approximation to the true posterior  $P(\theta|D)$ . Modal values and credible intervals for each model parameter can then be obtained from this approximate joint posterior.

Due to the high-dimensionality of most real-world datasets, the data  $D$  are often reduced to a summary statistic (or a set of summary statistics)  $S$ , so that we are really sampling from  $P(\theta|d(S, S^*) \leq \varepsilon)$  to approximate the posterior  $P(\theta|S)$ . The choice of appropriate summary statistics to maximize sufficiency (i.e. such that  $P(\theta|S) \rightarrow P(\theta|D)$ ) is not straightforward, and is an active area of statistical research (e.g. [67] and see also §2e). There have been many extensions to this initial basic—and inefficient—rejection algorithm, including weighting the retained parameter sets dependent on their exact distances  $d(\cdot, \cdot)$  through regression methods (e.g. [61,68]) or increasing the efficiency of sampling from the prior distributions (e.g. [69,70]).

The output of any ABC procedure is the joint posterior distribution of the model parameters  $\theta = (\theta_1, \dots, \theta_k)$  (and derived from that the marginal posterior distributions), indicating the range of the parameter space that is able to produce frequency data within a given tolerance level  $\varepsilon$  of the observed data, and consequently the learning strategies that are consistent with the data. We stress that the obtained posterior distribution is only a good approximation of the ‘true’, posterior distribution for small tolerance levels  $\varepsilon$ . Therefore if the obtained  $\varepsilon$  is large—and cannot be improved upon—the inferred parameter spaces are likely not meaningful. This situation may point to an inadequacy of the model, and therefore the assumed social learning processes, to explain the data. The explanatory value of the obtained posterior distribution can be investigated by posterior predictive checks [71]. These assess how well the parameter ranges specified by the posterior distribution explain the observed data (see [65] for further detail and §2d(i)). Additionally, cross validation tests or coverage plots have been developed to further investigate the accuracy of the results of the ABC analysis [40,72,73]. In practice, performing ABC analyses has been made relatively straightforward since the release of software such as DIY-ABC [74], ABCtoolbox [75], and R packages *abc* [72], *abctools* [76] and *EasyABC* [77].

Finally, we note that as well as estimating parameters, ABC has also been used to test between multiple competing models, by estimating Bayes factors from the relative proportions of simulations accepted from each model (e.g. [62]). While it has been shown that this approach is not theoretically justified [78] when reducing the data  $D$  to summary statistics  $S$ —as owing to the loss of information this approximation does not necessarily converge on the true Bayes factors—a number of authors have successfully applied various simulation-based power analyses to mitigate this problem (see for example [63,64,79]). And more recently another approach utilizing machine learning algorithms—and in particular random forests—has begun to prove successful for complex ABC model selection [80,81].



**Figure 2.** Blue solid line: uninformative, uniform prior distribution for parameter  $\theta_1$ ; red dashed and red dotted lines: potential marginal posterior distributions for  $\theta_1$ . (Online version in colour.)

### (c) Limits to inference

It is well-known that efforts to understand learning processes based on population-level data may be confounded by equifinality (e.g. [35,36] for a recent discussion). The inference framework introduced above generates posterior distributions of the model parameter describing the learning strategies that are consistent with the observed data. Therefore, the widths of these distributions, or their credible intervals, may provide an informal measure of the level of equifinality [82]. If the posterior distributions are narrow then only a small region of the parameter space is consistent with the data and therefore a large number of learning processes are *not* able to produce the observed frequency changes. In this case, the data carry a relatively strong signature of the underlying processes of social learning. By contrast, if the distributions are wide, a large region of the parameter space is consistent with the data and therefore many social learning processes are able to generate very similar population-level frequency patterns.

In this way the inference framework itself provides a way of exploring the inferential limits of population-level data of a given temporal resolution. For this, the generative model is used to simulate frequency data with a specific parameterization  $\theta$ , i.e. under a *known* process of social learning. Applying the inference procedure to this data produces posterior distributions, and while we know that the data have been generated with a specific parameter value, these distributions indicate all other values (and therefore learning processes) that could produce the ‘observed’ frequency changes equally well. Wide posterior distributions then mean that researchers should not expect cultural data—which is likely to be more noisy compared to pseudo-data produced by the generative model—with a similar temporal resolution to provide much information about underlying mechanisms.

But when do we consider a marginal posterior distribution narrow? One possibility is to compare the widths of prior and posterior distributions of the parameter in question. As mentioned above, the prior distribution describes the possible values that the parameter can assume or summarizes all prior knowledge researchers may have (see blue, solid line in figure 2 for an example of a uniform, uninformative prior distribution). If the parameter range covered by the posterior distribution is smaller compared to the range covered by the prior distribution (see the red, dashed line in figure 2

as an example) then the inference procedure led to the exclusion of some learning hypotheses: social learning processes described by parameter values not covered by the posterior distribution cannot generate theoretical data sufficiently close to the observed data and are consequently not considered to be consistent with the observations. Naturally, the smaller the credible interval the more the pool of potential hypotheses can be reduced, and the stronger the signature of underlying social learning processes in the observed population-level data. If, however, the parameter ranges covered by prior and posterior distributions are almost identical (see the red, dotted line in figure 2 as an example), then *a priori* knowledge of the researchers cannot be improved by analysing such data at the given resolution.

Additionally, cross validation analyses as suggested in [72] provide an alternative way of demonstrating how uninformative the data are about underlying social learning processes. In this context, we showed in [83] that we should not expect to be able to distinguish between unbiased transmission and moderately strong frequency-dependent selection based on frequency information of a population of cultural variants at two different points in time.

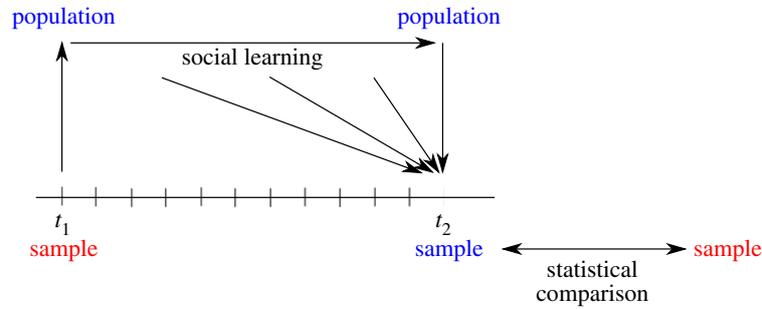
### (d) Application to studies of cultural evolution

#### (i) Cultural change in the linearbandkeramik (LBK) period

To demonstrate the applicability and utility of the generative inference framework described above, we summarize in the following the analysis of a cultural dataset from the earliest-known farming population in Central Europe, the so-called linearbandkeramik (LBK) from approximately 7500–7000 years ago (see [84] for the complete analysis). The dataset records the frequencies of different types of decorated vessels at seven different points in time, denoted by  $t_j$ ,  $j = 1, \dots, 7$  defining six phases of cultural change that vary in duration. The aim of this study was to explore whether observed frequency changes in different types of pottery between the beginning and the end of each of the six phases are consistent with a specific hypothesis about the underlying social learning processes, in particular unbiased transmission, frequency-dependent selection and pro-novelty selection. For the sake of brevity, we consider in the following unbiased and frequency-dependent selection only.

The first step of the inference framework is the development of the *generative model*. To make use of all available archaeological information, we used a simulation approach that accounted for the fact that the observed frequencies describe a sample and not the population of pottery types. Starting from observed data, the absolute frequencies  $D(t_j) = [n_1, \dots, n_k]$  of  $k$  different variant types in the sample of size  $n(t_j)$  at the beginning of the phase,  $t_j$ , we generated a population of cultural variants  $\mathbf{P}(t_j) = [R_1, \dots, R_k, R_{k+1}]$  from which the sample could have been drawn at random using the Dirichlet distribution approach [71]. The variables  $R_i$  represent the absolute frequency of variant type  $i$  in the population. Importantly, the population consists of  $k + 1$  variant types, where the type  $k + 1$  contains all variants of types not observed in the sample at  $t_j$ .

Based on this population  $\mathbf{P}(t_j) = [R_1, \dots, R_k, R_{k+1}]$  and an estimate of the population size  $N(t_j)$  at time  $t_j$  (if no other information is available the population size  $N(t_j)$  at time  $t_j$  is inferred from the size of the sample at this time), we generated population-level frequencies of the  $k + 1$  variant



**Figure 3.** Schematic representation of the proposed inference framework. As processes of social learning act on the population-level (as opposed to the sample-level), we generate population distributions from which the observed sample of cultural variants at the beginning of the phase,  $t_1$ , could have been drawn at random using the Dirichlet distribution approach. Subsequently, we transform these populations over time until the end of the phase,  $t_2$ , conditioned on the assumed social learning process. To obtain theoretical samples at time  $t_2$  that can be compared to the observed sample we randomly select cultural variants from the generated populations between  $t_1$  and  $t_2$ . (Online version in colour.)

types conditioned on a specific process of social learning at each time step  $t = 1, \dots, t_{j+1} - t_j$ . For that, we assumed that in each time step a fraction  $r$  of the population of cultural variants is removed and new variants are subsequently added (in this way the framework can accommodate temporal changes in population size). While the removal process is random, the replacement process is defined by the assumed process of social learning. In detail, a variant type  $i$ ,  $i = 1, \dots, k$  is chosen to be added to the population according to the probability

$$p_i(t) = \begin{cases} 0, & \text{if } \frac{N_i(t)-u_i}{N(t)-u} + b_{\text{freq}} \left[ \hat{k} \frac{N_i(t)-u_i}{N(t)-u} - 1 \right] < 0 \\ \left[ \frac{N_i(t)-u_i}{N(t)-u} + b_{\text{freq}} \left[ \hat{k} \frac{N_i(t)-u_i}{N(t)-u} - 1 \right] \right] (1 - \mu) & \text{otherwise,} \end{cases} \quad (2.1)$$

where  $N(t)$  denotes the population size at time  $t$ ,  $N_i(t)$  is the number of variants of type  $i$ ,  $u$  is the total number of variants removed at this time step,  $u_i$  is the number of variants removed of type  $i$ ,  $b_{\text{freq}}$  controls the strength of frequency-dependent selection and  $\hat{k}$  is the number of variant types present at time  $t$ . Importantly, choosing  $b_{\text{freq}} = 0$  in equation (2.1) models unbiased transmission, whereas  $b_{\text{freq}} > 0$  describes the selective advantage for high-frequency variant types and  $b_{\text{freq}} < 0$  for low-frequency types. Further, the variable  $\mu$  defines the probability with which a novel variant type not previously seen in the population is introduced into the system. A similar probability as in equation (2.1) is defined for variant type  $k + 1$ , containing all variant types not observed in the sample at  $t_1$  and, per definition, all subsequent innovations.

Lastly, to generate theoretical samples at the end of the phase  $t_{j+1}$ , we randomly drew  $n(t_{j+1})$  cultural variants from the (theoretical) populations  $\mathbf{P}(t_j + t)$ ,  $t = 1, \dots, t_{j+1} - t_j$ .

In summary, the output of this framework is sample frequencies of the variant types that were present at the beginning of the phase,  $t_j$ , and an additional type containing all unobserved variants at the end of the phase,  $t_{j+1}$ , conditioned on the social learning process specified by the parameter  $b_{\text{freq}}$  in equation (2.1), which controls the strength of the frequency-dependent selection.

To infer the learning processes consistent with the observed changes in frequency between the beginning and the end of the phases, we applied an ABC procedure—specifically SMC ABC (e.g. [70])—and determined the joint posterior distributions of  $(b_{\text{freq}}, r)$ . The replacement fraction  $r$

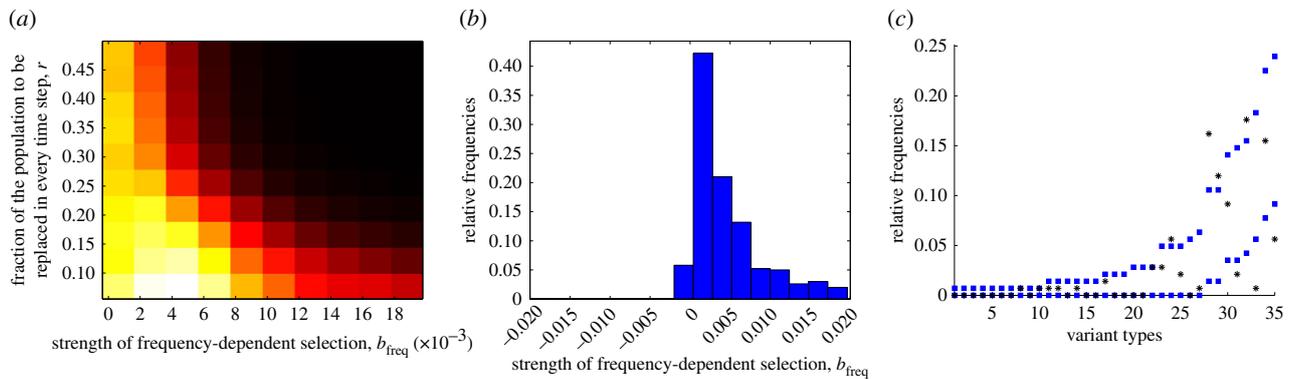
cannot be estimated from external sources and therefore has to be inferred from the data as well. Thereby the comparison between empirical and theoretical patterns was based on the absolute difference of the theoretical and observed frequencies of the  $k$  variant types present at the beginning of the simulation. Additionally, we required the same number of initially present variant types to have gone extinct at the end of the phase. The general scheme of the proposed generative inference framework is illustrated in figure 3.

Applying this analysis to all six phases, we concluded that

- (i) frequency-dependent selection does not describe the cultural dataset from the earliest farming population in Central Europe better than unbiased transmission. In fact, the credible intervals of all six marginal posterior distributions for  $b_{\text{freq}}$  contained the value 0, which means that unbiased transmission cannot be excluded as a potential explanation of the data by this analysis (see figure 4*a,b* for an example);
- (ii) frequency-dependent selection and unbiased transmission may not be the best model to explain the observed data as the achieved tolerance levels (i.e. the ‘distance’, between empirical and observed patterns) of the ABC analysis were relatively large.

Point (ii) suggests that the social learning hypotheses considered are not consistent with the data, which requires a re-evaluation of the generative model. Indeed, we showed in [84] that pro-novelty selection, which captures the preference for ‘young’, or recently introduced, cultural variant types, is able to replicate the observed frequency changes between the different phases and is therefore a *possible* explanation of the data.

Posterior predictive checks further highlighted the problem raised in point (ii). To perform this, we sampled values of the model parameters from the joint posterior distribution, inserted these into the generative model and produced theoretical frequencies at the end of each phase. Repeating this procedure generated theoretical expectations of the frequency ranges for each individual variant type based on the joint posterior distribution. The comparison of the observed frequencies of each variant type with these frequency ranges allowed the explanatory power of the derived posterior distribution to be assessed. If observations are outside the theoretical expectations then the inferred social learning processes cannot replicate all aspects of the dynamic



**Figure 4.** (a) Joint posterior distribution for the selection strength  $b_{\text{freq}}$  and the replacement fraction  $r$  for phase V of the considered dataset. (b) Corresponding marginal distribution for the selection strength  $b_{\text{freq}}$ . It is obvious that  $b_{\text{freq}} = 0$  is contained in the credible interval. (c) Outcome of the posterior predictive check: 95% prediction intervals, of the marginal frequency distribution of the different variant types at the end of the phase. The observed frequencies are indicated by the '\*' symbol. Many observations lie outside the prediction intervals, pointing to an inadequate description of the data by frequency dependent and unbiased transmission. (Online version in colour.)

of cultural change, indicating a mismatch between theory and data. This analysis also has the potential to reveal single variant types whose temporal frequency patterns deviate from the general population trend (see [65] for more details). Applying the posterior predictive check to the case study showed that a number of observations were outside their (theoretical) frequency ranges as determined by the joint posterior distributions (see figure 4c for an example).

### (ii) Further applications

In the last section, we demonstrated the application of a generative inference framework to a specific archaeological dataset. Traditionally, Bayesian inference in archaeology has been largely limited to age estimation via  $^{14}\text{C}$  analyses (e.g. [85,86]), but recently the scope of inference techniques has been vastly broadened, with ABC approaches enjoying increasing popularity (e.g. [65,82,87–90]). In one of the first archaeological applications, Crema *et al.* [87] studied frequency changes of weaponry types in the Jura region of southeast France. The dataset comprises arrowheads of 20 types attributed to 9 chronologically distinct phases. The aim of this study was to analyse whether the temporal frequency change of the different arrowhead types contained evidence for, or against, unbiased transmission or frequency-dependent selection. Using an agent-based simulation as their generative model, the authors produced frequency change patterns under different hypotheses of social learning and under the assumption that the cultural system is at equilibrium. They compared these theoretical patterns to the observed data by measuring the dissimilarity between assemblages. Applying an ABC model selection framework, they concluded that both unbiased transmission and negative-frequency dependent selection could have generated the observed frequency differences within the phases and therefore excluded positive-frequency dependent selection as a possible mechanism of cultural evolution.

But ABC frameworks have not been exclusively used to infer underlying social learning strategies. Porčić & Nikolić [88] analysed the demographic properties of the Mesolithic–Neolithic transition in the Central Balkan region, in particular growth rates and population size estimates for the Lepenski Vir population. Their model generated the expected number of accumulated houses for a

large range of demographic scenarios which could then be compared to that observed in the archaeological record. The analysis revealed higher initial growth rates compared to other populations undergoing the Neolithic demographic transition and an increase in population size over time.

In order to highlight the breadth of questions that can be addressed within a generative inference framework we outline two further applications, one to historical studies (a field with no strong tradition of quantitative treatments) and to linguistics. Rubio-Campillo [91] investigated the evolution of combat. He explored the validity of different versions of Lanchester's law predicting the causalities of two enemy forces engaged in a land battle, with a dataset comprising the total number of combatants and causalities from 1080 land battles spanning from the middle of the seventeenth to the beginning of the twentieth century. The three most common formulations of Lanchester's law (linear, squared and logarithmic) can be operationalized using difference equations, and iterating these until one of the forces has suffered as many causalities as recorded in the historical record allowed for the comparison between theoretical and observed data. Besides confirming well-known results, the ABC framework pointed to a gradual decrease in the relevance of individual fighting abilities, suggesting that the plausibility of the models is not constant over the different periods.

Lastly, Thouzeau *et al.* [81] investigated the coevolution between genes and languages at a regional scale. They simulated population genetic and cognate data under various historical models encompassing divergences and multiple borrowings and admixture events between linguistic groups. They applied an ABC framework using linguistic and genetic data from across Central Asia, and were able to reconstruct the partly differing evolutionary scenarios underlying linguistic and genetic differentiation in the region.

### (e) Cautionary notes

Naturally, the application of the generative inference framework presented here has to proceed with caution. It is, after all, an analysis based on an underlying model of cultural change. If this model does not capture the main cultural and demographic processes contributing to the observed temporal frequency changes, the inferences obtained will likely be misleading. In the following, we outline some

issues researchers should consider before applying this, or a similar, inference framework.

In this paper, we advocate the use of non-equilibrium frameworks. While this modelling choice allows us to include knowledge about, for example, temporal changes in demographic properties and to initialize the model with observed variant frequencies, it also introduces a time-dependency. The inference framework evaluates whether frequency changes between different time points are consistent with the changes expected under a specific learning process (instead of evaluating whether statistics such as the level of cultural diversity at each point time are consistent with the equilibrium diversity prediction) and therefore misspecifications of time points and consequently the duration of the period over which the frequency changes are measured can produce erroneous theoretical expectations. Crema *et al.* [65] argue that the equilibrium assumption should serve as a hypothesis to be tested, rather than simply held *a priori*. They applied equilibrium and non-equilibrium versions of the generative model of cultural change to a dataset similar to the one described in §2d(i). They concluded that the cultural system was likely not at equilibrium and found hints for shifts between negative and positive frequency-dependent selection for different phases of the archaeological record.

In the archaeological case study described in §2d(i), the temporal change in population size between the beginning and the end of each archaeological phase has been inferred from the change in sample size, and any increase or decrease was assumed to occur in a linear fashion over the relevant time interval. While the assumption of linear change seems plausible, especially in the absence of other information, drastic, unobserved demographic events such as population bottlenecks may be an alternative scenario. Similar to the discussion about equilibrium versus non-equilibrium models, such hidden demographic events have the potential to influence the dynamic of cultural change (e.g. [92]). As they are not included in the generative model, their influences may be mistakenly attributed to social learning processes that are able to produce a similar effect at the population level. But this potential pitfall is also itself amenable to testing with the generative inference framework. Researchers can at least evaluate the extent to which posterior distributions change when assuming a population bottleneck between the beginning and the end of the phase.

The accuracy of ABC inference depends partly on how the difference between observed and simulated data is calculated and on the achieved tolerance level. Calculating the difference based on summary statistics  $S$  instead of the full data  $D$  results in discarding likely useful information [93]. If a summary statistic (or set of) is not sufficient—as is generally the case in practice—the resulting posterior distribution will not be equal to that computed with the full data [94] (see also [93] for a review of strategies dealing with this issue). While the impact of using insufficient statistics on inference results can be mitigated by careful application, we note that by using the actual frequencies for calculating the difference between observed and simulated data this problem is circumvented entirely. Further, any posterior distribution with large tolerance levels does not approximate the ‘true’ posterior distribution and should be treated with caution. In this case, the generative model may not produce data that are sufficiently close to the observed data. Additional procedures such as posterior predictive checks, cross

validation tests or coverage plots offer additional insights into the accuracy of the inference results.

Lastly, we point to the relationship between data quality or completeness and inferential accuracy. A recent study [31] revealed the importance of rare variants for inferring underlying processes. Using the progeny distribution (which records the frequencies of cultural variant types that produce  $k$  new variants over a fixed period of time) as a statistic, the authors showed that analyses based on only the most popular variants, as is often necessarily the case in cultural evolutionary studies, can provide misleading evidence for underlying transmission hypotheses. Especially in archaeological case studies, the observed frequencies describe the composition of often relatively small samples of cultural variants, and consequently rare variant types are likely *not* sampled and therefore absent from the data. Even though statistical techniques such as the Dirichlet distribution approach mentioned in §2d(i) are available, the number of rare variant types, i.e. types that are not contained in the observed sample, is likely to be misspecified (e.g. [95]) and future work is needed to understand the influence of missing data on the accuracy of the generative inference frameworks presented in this paper.

### 3. Conclusion

Relatively recent developments in population genetics—namely coalescent modelling and ABC—have made generative inference possible, and shown it to be a powerful inferential framework for understanding the human past (e.g. [57,58,63,64]). Cultural evolutionary theory has been greatly advanced by adopting concepts and modelling paradigms originating in population genetics. In this spirit, the aim of this paper was to demonstrate how analogous generative inference frameworks can be applied to cultural frequency data, potentially allowing us to close the gap between theoretical modelling work and empirical work in cultural evolution.

In particular, we focused on the topic of inferring how human populations use social information based on the available empirical evidence. In many case studies of interest, the available data are in the form of frequencies of different variants of a cultural trait in the population at one or several points in time, which means that we face a classical inverse problem. Naturally, attempting to address this problem leads to the question of how much information about underlying processes of social learning can in fact be extracted from cultural frequency data of a given resolution. The framework outlined here allows us to address this equifinality problem. At the heart of this framework is a generative model, which captures the main cultural and demographic properties of the system considered. As noted, there are no restrictions on the type of model used, with the one described in §2d(i) simply an example tailored specifically to the observed population-level frequency data. Whatever their form, these models establish a causal link between model parameters controlling the strengths of underlying evolutionary processes and observable population-level patterns; in our case, between parameters controlling the strengths of social learning processes and population-level frequencies of cultural variant types. Bayesian inference techniques, such as ABC, can then evaluate whether this specific process of social

learning is able to produce frequency patterns consistent with the observed ones.

The outcome of this inference approach is posterior distributions of the model parameters describing the learning processes that are consistent with the observed data. As discussed in §2e, while there are a number of important factors potentially influencing the accuracy of the analysis to consider, the widths of the posterior distributions may be indicative of the amount of information about the underlying social learning processes contained in the data. Narrow posterior distributions indicate that the data carry a relatively strong signature of these processes, while wider distributions suggest that the data are largely uninformative or that the models considered do not provide an adequate description of the cultural system. Therefore, this approach does not only allow for the identification of the most likely underlying learning process given the empirical data, but also for a description of the breadth of processes that could have produced the these data equally well.

Revealing the presence of equifinality may appear to be a negative result, but we stress that one should not expect a unique mapping between (sparse) population-level frequency data and underlying processes of cultural evolution [4,36]. Nevertheless, the analysis of such data will help in excluding social learning processes that could not have produced the observed data. In this way inference frameworks will lead to a reduction in the pool of potential hypotheses (even though the level of reduction might vary from case study to case study) and to an understanding of which kinds of scientific questions can be answered by which kinds of

data. Additionally, we note that generative inference frameworks inform about the consistency of a limited set of possible underlying mechanisms with the data while not excluding the possibility that other mechanisms may be consistent as well. However, this should not necessarily be seen as a weakness, and as pointed out by Csilléry *et al.* [93, p. 413], ‘in reality scientific arguments often revolve around a limited number of hypotheses or scenarios without the need to consider an infinite set of alternative models. Models can always be improved and refined by other authors, allowing an open discussion that can greatly increase our understanding of the problem being studied.’

Undoubtedly, more research is needed to further develop and improve the statistical tools and to explore the influence of e.g. unobserved changes in the demographic properties of the system considered or of the quality of the observed data on the accuracy of generative inference frameworks, but we believe this is an exciting and promising new direction in cultural evolution that has already begun to produce interesting results.

**Data accessibility.** This article has no additional data.

**Competing interests.** We declare we have no competing interests.

**Funding.** No funding has been received for this article.

**Acknowledgments.** We thank Nicole Creanza, Oren Kolodny and Mark Feldman for inviting us to contribute to this special issue. Further, we thank two anonymous reviewers for their constructive comments and criticisms, which helped us improve this manuscript and members of the department of Human Behavior, Ecology and Culture at the Max Planck Institute for Evolutionary Anthropology for helpful comments on an earlier version of the manuscript.

## References

- Rendell L, Fogarty L, Hoppitt WJE, Morgan TJH, Webster MM, Laland KN. 2011 Cognitive culture: theoretical and empirical insights into social learning strategies. *Trends. Cogn. Sci. (Regul. Ed.)* **15**, 68–76. (doi:10.1016/j.tics.2010.12.002)
- Heyes CM. 1994 Social learning in animals: categories and mechanisms. *Biol. Rev.* **69**, 207–231. (doi:10.1111/j.1469-185X.1994.tb01506.x)
- Hoppitt W, Laland KN. 2013 *Social learning: an introduction to mechanisms, methods, and models*. Princeton, NJ: Princeton University Press.
- Cavalli-Sforza LL, Feldman MW. 1981 *Cultural transmission and evolution: a quantitative approach*. Princeton, NJ: Princeton University Press.
- Boyd R, Richerson PJ. 1985 *Culture and the evolutionary process*. Chicago, IL: University of Chicago Press.
- Rendell L *et al.* 2010 Why copy others? Insights from the social learning strategies tournament. *Science* **328**, 208–213. (doi:10.1126/science.1184719)
- Giraldeau L-A, Valone TJ, Templeton JJ. 2002 Potential disadvantages of using socially acquired information. *Phil. Trans. R. Soc. Lond. B* **357**, 1559–1566. (doi:10.1098/rstb.2002.1065)
- Laland KN. 2004 Social learning strategies. *Anim. Learn. Behav.* **32**, 4–14. (doi:10.3758/BF03196002)
- Coultras JC. 2004 When in Rome . . . an evolutionary perspective on conformity. *Group. Process. Intergroup. Relat.* **7**, 317–331. (doi:10.1177/1368430204046141)
- Baum WM, Richerson PJ, Efferson CM, Paciotti BM. 2004 Cultural evolution in laboratory microsocieties including traditions of rule giving and rule following. *Evol. Hum. Behav.* **25**, 305–326. (doi:10.1016/j.evolhumbehav.2004.05.003)
- McElreath R, Bell AV, Efferson C, Lubell M, Richerson PJ, Waring T. 2008 Beyond existence and aiming outside the laboratory: estimating frequency-dependent and pay-off-biased social learning strategies. *Phil. Trans. R. Soc. B* **363**, 3515–3528. (doi:10.1098/rstb.2008.0131)
- Morgan TJH, Rendell LE, Ehn M, Hoppitt W, Laland KN. 2012 The evolutionary basis of human social learning. *Proc. R. Soc. B* **279**, 653–662. (doi:10.1098/rspb.2011.1172)
- Mesoudi A, O'Brien MJ 2008 The cultural transmission of great basin projectile-point technology (i): an experimental simulation. *Am. Antiq.* **73**, 3–28. (doi:10.2307/25470521)
- Caldwell CA, Millen AE. 2009 Social learning mechanisms and cumulative cultural evolution: is imitation necessary? *Psychol. Sci.* **20**, 1478–1483. (doi:10.1111/j.1467-9280.2009.02469.x)
- Kirby S, Cornish H, Smith K. 2008 Cumulative cultural evolution in the laboratory: an experimental approach to the origins of structure in human language. *Proc. Natl Acad. Sci. USA* **105**, 10 681–10 686. (doi:10.1073/pnas.0707835105)
- Schillinger K, Mesoudi A, Lycett SJ. 2015 The impact of imitative versus emulative learning mechanisms on artifactual variation: implications for the evolution of material culture. *Evol. Hum. Behav.* **36**, 446–455. (doi:10.1016/j.evolhumbehav.2015.04.003)
- Whiten A, Caldwell CA, Mesoudi A. 2016 Cultural diffusion in humans and other animals. *Curr. Opin. Psychol.* **8**, 15–21. (doi:10.1016/j.copsyc.2015.09.002)
- Aoki K, Feldman MW. 2014 Evolution of learning strategies in temporally and spatially variable environments: a review of theory. *Theor. Popul. Biol.* **91**, 3–19. (doi:10.1016/j.tpb.2013.10.004)
- Guglielmino CR, Viganotti C, Hewlett B, Cavalli-Sforza LL. 1995 Cultural variation in Africa: role of mechanisms of transmission and adaptation. *Proc. Natl Acad. Sci. USA* **92**, 7585–7589. (doi:10.1073/pnas.92.16.7585)
- Beheim BA, Thigpen C, McElreath R. 2014 Strategic social learning and the population dynamics of human behavior: the game of Go. *Evol. Hum. Behav.* **35**, 351–357. (doi:10.1016/j.evolhumbehav.2014.04.001)
- Henrich J, Broesch J. 2011 On the nature of cultural transmission networks: evidence from Fijian villages

- for adaptive learning biases. *Phil. Trans. R. Soc. B* **366**, 1139–1148. (doi:10.1098/rstb.2010.0323)
22. McElreath R. 2018 A long-form research program in human behavior, ecology, and culture. White paper available under [http://www.eva.mpg.de/fileadmin/content\\_files/staff/richard\\_mcelreath/pdf/HBEC\\_whitepaper.pdf](http://www.eva.mpg.de/fileadmin/content_files/staff/richard_mcelreath/pdf/HBEC_whitepaper.pdf) (Last accessed 9 January 2018.).
  23. Henrich J. 2004 Cultural group selection, coevolutionary processes and large-scale cooperation. *J. Econ. Behav. Organ.* **53**, 3–35. (doi:10.1016/S0167-2681(03)00094-5)
  24. Rogers EM. 1995 *Diffusion of innovations*. New York, NY: Free Press.
  25. Neiman FD. 1995 Stylistic variation in evolutionary perspective: inferences from decorative diversity and interassemblage distance in Illinois woodland ceramic assemblages. *Am. Antiq.* **60**, 7–36. (doi:10.2307/282074)
  26. Shennan SJ, Wilkinson JR. 2001 Ceramic style change and neutral evolution: a case study from neolithic Europe. *Am. Antiq.* **66**, 577–593. (doi:10.2307/2694174)
  27. Kohler TA, VanBuskirk S, Ruscavage-Barz S. 2004 Vessels and villages: evidence for conformist transmission in early village aggregations on the Pajarito Plateau, New Mexico. *J. Anthropol. Archaeol.* **23**, 100–118. (doi:10.1016/j.jaa.2003.12.003)
  28. Premo LS. 2014 Cultural transmission and diversity in time-averaged assemblages. *Curr. Anthropol.* **55**, 105–114. (doi:10.1086/674873)
  29. Hahn MW, Bentley RA. 2003 Drift as a mechanism for cultural change: an example from baby names. *Proc. R. Soc. Lond. B* **270**(Suppl. 1), S120–S123. (doi:10.1098/rsbl.2003.0045)
  30. Bentley RA, Hahn MW, Shennan SJ. 2004 Random drift and culture change. *Proc. R. Soc. Lond. B* **271**, 1443–1450. (doi:10.1098/rspb.2004.2746)
  31. O'Dwyer JP, Kandler A. 2017 Inferring processes of cultural transmission: the role of rare variants for distinguishing neutrality from novelty biases. *Phil. Trans. R. Soc. B* **372**, 20160426. (doi:10.1098/rstb.2016.0426)
  32. Bentley RA, Lipo CP, Herzog HA, Hahn MW. 2007 Regular rates of popular culture change reflect random copying. *Evol. Hum. Behav.* **28**, 151–158. (doi:10.1016/j.evolhumbehav.2006.10.002)
  33. Acerbi A, Bentley RA. 2014 Biases in cultural transmission shape the turnover of popular traits. *Evol. Hum. Behav.* **35**, 228–236. (doi:10.1016/j.evolhumbehav.2014.02.003)
  34. Eerkens JW, Bettinger RL, McElreath R. 2006 Cultural transmission, phylogenetics, and the archaeological record. In *Mapping our ancestors: Phylogenetic methods in anthropology and prehistory* (eds CP Lipo, MJ O'Brien, M Collard, SJ Shennan), pp. 169–83. New York, NY: Aldine.
  35. Von Bertalanffy L. 1969 *General system theory: foundations, development, applications* (revised edition). New York, NY: George Braziller Inc.
  36. Premo LS. 2010 Equifinality and explanation: the role of agent-based modeling in postpositivist archaeology. In *Simulating change: Archaeology into the twenty-first century* (eds A Costopoulos, MW Lake), pp. 28–37. Salt Lake City, UT: University of Utah Press.
  37. Schelling TC. 1971 Dynamic models of segregation. *J. Math. Sociol.* **1**, 143–186. (doi:10.1080/0022250X.1971.9989794)
  38. Epstein JM, Axtell R. 1996 *Growing artificial societies: social science from the bottom up*. Washington, DC: Brookings Institution Press.
  39. Epstein JM. 2006 *Generative social science: Studies in agent-based computational modeling*. Princeton, NJ: Princeton University Press.
  40. van der Vaart E, Beaumont MA, Johnston ASA, Sibly RM. 2015 Calibration and evaluation of individual-based models using approximate Bayesian computation. *Ecol. Modell.* **312**, 182–190. (doi:10.1016/j.ecolmodel.2015.05.020)
  41. Kandler A, Wilder B, Fortunato L. 2017 Inferring individual-level processes from population-level patterns in cultural evolution. *R. Soc. open. sci.* **4**, 170949. (doi:10.1098/rsos.170949)
  42. Ewens WJ. 2004 *Mathematical population genetics 1: Theoretical introduction*. Berlin, Germany: Springer Science & Business Media.
  43. Fisher RA. 1930 *The genetical theory of natural selection: a complete variorum edition*. Oxford, UK: Oxford University Press.
  44. Wright S. 1931 Evolution in Mendelian populations. *Genetics* **16**, 97–159.
  45. Moran PAP. 1958 Random processes in genetics. *Math. Proc. Cambridge Philos. Soc.* **54**, 60–71. (doi:10.1017/S0305004100033193)
  46. Kingman JFC. 1982 The coalescent. *Stoch. Process. Their Appl.* **13**, 235–248. (doi:10.1016/0304-4149(82)90011-4)
  47. Hudson RR. 1983 Properties of a neutral allele model with intragenic recombination. *Theor. Popul. Biol.* **23**, 183–201. (doi:10.1016/0040-5809(83)90013-8)
  48. Tajima F. 1983 Evolutionary relationship of DNA sequences in finite populations. *Genetics* **105**, 437–460.
  49. Tavaré S. 1984 Line-of-descent and genealogical processes, and their applications in population genetics models. *Theor. Popul. Biol.* **26**, 119–164. (doi:10.1016/0040-5809(84)90027-3)
  50. Li H, Durbin R. 2011 Inference of human population history from individual whole-genome sequences. *Nature* **475**, 493–496. (doi:10.1038/nature10231)
  51. Schiffels S, Durbin R. 2014 Inferring human population size and separation history from multiple genome sequences. *Nat. Genet.* **46**, 919–925. (doi:10.1038/ng.3015)
  52. Pickrell JK, Pritchard JK. 2012 Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS. Genet.* **8**, e1002967. (doi:10.1371/journal.pgen.1002967)
  53. Patterson N, Moorjani P, Luo Y, Mallick S, Rohland N, Zhan Y, Genschoreck T, Webster T, Reich D. 2012 Ancient admixture in human history. *Genetics* **192**, 1065–1093. (doi:10.1534/genetics.112.145037)
  54. Hellenthal G, Busby GBJ, Band G, Wilson JF, Capelli C, Falush D, Myers S. 2014 A genetic atlas of human admixture history. *Science* **343**, 747–751. (doi:10.1126/science.1243518)
  55. Gronau I, Hubisz MJ, Gulko B, Danko CG, Siepel A. 2011 Bayesian inference of ancient human demography from individual genome sequences. *Nat. Genet.* **43**, 1031–1034. (doi:10.1038/ng.937)
  56. Currat M, Ray N, Excoffier L. 2004 SplatcHe: a program to simulate genetic diversity taking into account environmental heterogeneity. *Mol. Ecol. Resour.* **4**, 139–142. (doi:10.1046/j.1471-8286.2003.00582.x)
  57. Ray N, Wegmann D, Fagundes NJR, Wang S, Ruiz-Linares A, Excoffier L. 2009 A statistical evaluation of models for the initial settlement of the American continent emphasizes the importance of gene flow with Asia. *Mol. Biol. Evol.* **27**, 337–345. (doi:10.1093/molbev/msp238)
  58. Eriksson A, Betti L, Friend AD, Lycett SJ, Singarayer JS, von Cramon-Taubadel N, Valdes PJ, Balloux F, Manica A. 2012 Late Pleistocene climate change and the global expansion of anatomically modern humans. *Proc. Natl Acad. Sci. USA* **109**, 16 089–16 094. (doi:10.1073/pnas.1209494109)
  59. Bramanti B et al. 2009 Genetic discontinuity between local hunter-gatherers and central Europe's first farmers. *Science* **326**, 137–140. (doi:10.1126/science.1176869)
  60. Bollongino R, Nehlich O, Richards MP, Orschiedt J, Thomas MG, Sell C, Fajkošová Z, Powell AT, Burger J. 2013 2000 years of parallel societies in stone age central Europe. *Science* **342**, 479–481. (doi:10.1126/science.1245049)
  61. Beaumont MA, Zhang W, Balding DJ. 2002 Approximate Bayesian computation in population genetics. *Genetics* **162**, 2025–2035.
  62. Pritchard JK, Seielstad MT, Perez-Lezaun A, Feldman MW. 1999 Population growth of human Y chromosomes: a study of Y chromosome microsatellites. *Mol. Biol. Evol.* **16**, 1791–1798. (doi:10.1093/oxfordjournals.molbev.a026091)
  63. Veeramah KR, Wegmann D, Woerner A, Mendez FL, Watkins JC, Destro-Bisol G, Soodyall H, Louie L, Hammer MF. 2011 An early divergence of KhoeSan ancestors from those of other modern humans is supported by an ABC-based analysis of autosomal resequencing data. *Mol. Biol. Evol.* **29**, 617–630. (doi:10.1093/molbev/msr212)
  64. Posth C et al. 2016 Pleistocene mitochondrial genomes suggest a single major dispersal of non-Africans and a late glacial population turnover in Europe. *Curr. Biol.* **26**, 827–833. (doi:10.1016/j.cub.2016.01.037)
  65. Crema ER, Kandler A, Shennan SJ. 2016 Revealing patterns of cultural transmission from frequency data: equilibrium and non-equilibrium assumptions. *Sci. Rep.* **6**, 39122 (doi:10.1038/srep39122)
  66. Steele J, Glatz C, Kandler A. 2010 Ceramic diversity, random copying, and tests for selectivity in ceramic production. *J. Archaeol. Sci.* **37**, 1348–1358.
  67. Harrison J, Baker R. 2017 An automatic adaptive method to combine summary statistics in approximate Bayesian computation. (<http://arxiv.org/abs/1703.02341v1>)

68. Blum MGB, François O. 2010 Non-linear regression models for approximate Bayesian computation. *Stat. Comput.* **20**, 63–73. (doi:10.1007/s11222-009-9116-0)
69. Marjoram P, Molitor J, Plagnol V, Tavaré S. 2003 Markov chain Monte Carlo without likelihoods. *Proc. Natl Acad. Sci. USA* **100**, 15 324–15 328. (doi:10.1073/pnas.0306899100)
70. Toni T, Welch D, Strelkowa N, Ipsen A, Stumpf MPH. 2009 Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *J. R. Soc. Interface* **6**, 187–202. (doi:10.1098/rsif.2008.0172)
71. Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. 2013 *Bayesian data analysis*. 3rd edn. Boca Raton, FL: CRC Press.
72. Csilléry K, François O, Blum MGB. 2012 abc: an R package for approximate bayesian computation (abc). *Methods Ecol. Evol.* **3**, 475–479. (doi:10.1111/j.2041-210X.2011.00179.x)
73. Prangle D, Blum MGB, Popovic G, Sisson SA. 2014 Diagnostic tools for approximate bayesian computation using the coverage property. *Aust. N. Z. J. Stat.* **56**, 309–329. (doi:10.1111/anzs.12087)
74. Cornuet J-M, Santos F, Beaumont MA, Robert CP, Marin J-M, Balding DJ, Guillemaud T, Estoup A. 2008 Inferring population history with DIY ABC: a user-friendly approach to approximate bayesian computation. *Bioinformatics* **24**, 2713. (doi:10.1093/bioinformatics/btn514)
75. Wegmann D, Leuenberger C, Neuenschwander S, Excoffier L. 2010 ABCtoolbox: a versatile toolkit for approximate Bayesian computations. *BMC Bioinformatics* **11**, 116. (doi:10.1186/1471-2105-11-116)
76. Nunes MA, Prangle D. 2015 abctools: an R package for tuning approximate Bayesian computation analyses. *R. J.* **7**, 189–205.
77. Jabot F, Faure T, Dumoulin N. 2013 EasyABC: performing efficient approximate Bayesian computation sampling schemes using R. *Methods Ecol. Evol.* **4**, 684–687. (doi:10.1111/2041-210X.12050)
78. Robert CP, Cornuet J-M, Marin J-M, Pillai NS. 2011 Lack of confidence in approximate Bayesian computation model choice. *Proc. Natl Acad. Sci. USA* **108**, 15 112–15 117. (doi:10.1073/pnas.1102900108)
79. Beaumont M. 2008 Joint determination of topology, divergence time and immigration in population trees. In *Simulations, genetics and human prehistory* (eds S Matsumura, P Forster, C Renfrew), pp. 135–154. Cambridge, UK: McDonald Institute for Archaeological Research.
80. Pudlo P, Marin J-M, Estoup A, Cornuet J-M, Gautier M, Robert CP. 2015 Reliable ABC model choice via random forests. *Bioinformatics* **32**, 859–866. (doi:10.1093/bioinformatics/btv684)
81. Thouzeau V, Menecier P, Verdu P, Austerlitz F. 2017 Genetic and linguistic histories in central Asia inferred using approximate Bayesian computations. *Proc. R. Soc. B* **284**, 20170706. (doi:10.1098/rspb.2017.0706)
82. Kandler A, Laland KN. 2013 Tradeoffs between the strength of conformity and number of conformists in variable environments. *J. Theor. Biol.* **332**, 191–202. (doi:10.1016/j.jtbi.2013.04.023)
83. Kandler A, Powell A. 2015 Inferring learning strategies from cultural frequency data. In *Learning strategies and cultural evolution during the Palaeolithic* (eds A Mesoudi, K Aoki), pp. 85–101. Berlin, Germany: Springer.
84. Kandler A, Shennan SJ. 2015 A generative inference framework for analysing patterns of cultural change in sparse population data with evidence for fashion trends in LBK culture. *J. R. Soc. Interface* **12**, 20150905. (doi:10.1098/rsif.2015.0905)
85. Ramsey CB. 2009 Bayesian analysis of radiocarbon dates. *Radiocarbon* **51**, 337–360. (doi:10.1017/S0033822200033865)
86. Buck CE, Kenworthy JB, Litton CD, Smith AFM. 1991 Combining archaeological and radiocarbon information: a Bayesian approach to calibration. *Antiquity* **65**, 808–821. (doi:10.1017/S0003598X00080534)
87. Crema ER, Edinborough K, Kerig T, Shennan SJ. 2014 An approximate Bayesian computation approach for inferring patterns of cultural evolutionary change. *J. Archaeol. Sci.* **50**, 160–170. (doi:10.1016/j.jas.2014.07.014)
88. Porčić M, Nikolić M. 2016 The approximate Bayesian computation approach to reconstructing population dynamics and size from settlement data: demography of the mesolithic–neolithic transition at Lepenski Vir. *Archaeol. Anthropol. Sci.* **8**, 169–186. (doi:10.1007/s12520-014-0223-2)
89. Edinborough K, Shennan SJ, Crema ER, Kerig T. 2015 An ABC of lithic arrowheads: a case study from south-eastern France. In *Neolithic Diversities. Acta Archaeologica Lundsensia*, Series 80, vol. 65 (eds K Brink, S Hydén, K Jenn bert, L Larsson, D Olausson), pp. 213–223. Department of Archaeology and Ancient History, Lund University: Sweden.
90. Kovacevic M, Shennan SJ, Vanhaeren M, d’Errico F, Thomas MG. 2015 Simulating geographical variation in material culture: were early modern humans in Europe ethnically structured? In *Learning strategies and cultural evolution during the Palaeolithic* (eds A Mesoudi, K Aoki), pp. 103–120. Berlin, Germany: Springer.
91. Rubio-Campillo X. 2016 Model selection in historical research using approximate Bayesian computation. *PLoS ONE* **11**, e0146491. (doi:10.1371/journal.pone.0146491)
92. Rorabaugh AN. 2014 Impacts of drift and population bottlenecks on the cultural transmission of a neutral continuous trait: an agent based model. *J. Archaeol. Sci.* **49**, 255–264. (doi:10.1016/j.jas.2014.05.016)
93. Csilléry K, Blum MGB, Gaggiotti OE, François O. 2010 Approximate Bayesian computation (ABC) in practice. *Trends. Ecol. Evol. (Amst.)* **25**, 410–418. (doi:10.1016/j.tree.2010.04.001)
94. Marjoram P, Tavaré S. 2006 Modern computational approaches for analysing molecular genetic variation data. *Nat. Rev. Genet.* **7**, 759–770. (doi:10.1038/nrg1961)
95. Mao CX, Colwell RK. 2005 Estimation of species richness: mixture models, the role of rare species, and inferential challenges. *Ecology* **86**, 1143–1153. (doi:10.1890/04-1078)