Kandler, A. & Sherman, S. (2013). A non-equilibrium neutral model for analysing cultural change. Journal of Theoretical Biology, 330, pp. 18-25. doi: 10.1016/j.jtbi.2013.03.006

Original citation: Kandler, A. & Sherman, S. (2013). A non-equilibrium neutral model for analysing cultural change. Journal of Theoretical Biology, 330, pp. 18-25. doi: 10.1016/j.jtbi.2013.03.006

Permanent City Research Online URL: http://openaccess.city.ac.uk/7651/

A non-equilibrium neutral model for analysing cultural change

Anne Kandler*,[1,2,3] and Stephen Shennan[3]

[1]Department of Mathematics, City University London, Northampton Square, London EC1V 0HB, UK
*Corresponding author
Email: anne.kandler.1@city.ac.uk
Tel: +44 (0)20 7040 8464
Fax: +44 (0)20 7040 5060
[2]Santa Fe Institute, 1399 Hyde Park Road, Santa Fe (NM) 87501, USA
[3]Institute of Archaeology, University College London, 31–34, Gordon Square, London WC1H 0PY, UK

Abstract. Neutral evolution is a frequently used model to analyse changes in frequencies of cultural variants over time. Variants are chosen to be copied according to their relative frequency and new variants are introduced by a process of random mutation. Here we present a non-equilibrium neutral model which accounts for temporally varying population sizes and mutation rates and makes it possible to analyse the cultural system under consideration at any point in time. This framework gives an indication whether observed changes in the frequency distributions of a set of cultural variants between two time points are consistent with the random copying hypothesis. We find that the likelihood of the existence of the observed assemblage at the end of the considered time period (expressed by the probability of the observed number of cultural variants present in the population during the whole period under neutral evolution) is a powerful indicator of departures from neutrality. Further, we study the effects of frequency-dependent selection on the evolutionary trajectories and present a case study of change in the decoration of pottery in early Neolithic Central Europe. Based on the framework developed we show that neutral evolution is not an adequate description of the observed changes in frequency.

Keywords: Neutral model, Wright-Fisher model, Markov chains, conformity, Linear Pottery Culture

1. Introduction

Changes in frequencies of cultural variants over time reflect the dynamic of the underlying evolutionary processes. Those processes could be selection, random drift or other evolutionary forces ([7,8]) and the question remains whether inferences about the nature of the underlying evolutionary process can be made on the base of observed frequency patterns. In archaeological and anthropological applications those patterns are often the only information researchers have about past cultural traditions and the forces affecting them, so this question gains particular importance [24]. In archaeology especially there is an extensive literature going back many years concerned with distinguishing 'functional' from 'stylistic' aspects of artefact traditions and identifying the forces acting on them as they change through time (e.g. [4]). These discussions took a new direction when Dunnell [10] proposed that 'stylistic' aspects of artefacts could be defined as 'those not under selection'. Subsequently, Neiman [23] was able to put this insight to use by demonstrating how the mathematical foundations of the genetic theory of neutral evolution (e.g. [9,20]) could be used to form expectations about observable frequency patterns under the assumption that temporal frequency changes of cultural variants are solely the product of random drift and innovations. Neiman took a real archaeological case-study concerning changes through time in the frequencies of attributes of ceramics in pre-Columbian North America and showed that these could be accounted for by drift and innovation alone, a conclusion that had major substantive implications for the archaeology concerned, leading to a view of the prehistoric processes at work in this case that contradicted previous conclusions. Since then this model has been applied to diverse examples of cultural evolution, including the frequency distribution of first names [15], the popularity of purebred dogs [16] and stylistic motifs on prehistoric pottery [2,19,23,25,28][1].

Neutral evolution assumes that in finite populations cultural variants are chosen to be copied according to their relative frequency (and therefore changes in frequency occur only as a result of random drift), and new variants are introduced by a process of random mutation. Central to this approach is the notion of cultural homogeneity within a population at time t. Using Ewens sampling formula [12], which predicts properties of the underlying distribution for a neutral process, where an empirical sample of size N has been observed with k variants, one can infer how well the observed homogeneity pattern coincides with the pattern expected by neutral evolution. To do so two statistical test have been developed, the Ewens-Watterson test of homozygosity [30,31] and the Slatkin test [26,27]; the latter test is more general in using fewer assumptions about the aspects of distribution shape that might indicate selection [28].

In broad terms two different kinds of information can be gained from the neutral model. Firstly, if neutrality of the underlying evolutionary process is assumed one can draw inferences about variations in population characteristics such as population size and mutation rates over time, and secondly to test for departure from neutral evolution. In this latter context the neutral model has been proposed as a null hypothesis for studying cultural evolution, where rejection of the null hypothesis predicts the presence of collective selective forces [3,23,25][2]. However, as many scholars have noted, the application of the neutral model to cultural evolution must be done with caution as the underlying assumption that cultural variants whose frequencies are modelled are selectively neutral is hard to prove (for detailed discussions see [23,25,28]). Steele at al. [28] called for stronger tests of departure from neutrality than testing the goodness of fit between empirical data and theoretical frequency distributions. They showed, using the example of Hittite ceramic bowl types, that the neutral hypothesis was not rejected on the basis of the frequency distribution; however, examination of the characteristics of the bowl types revealed a correlation between those characteristics and their abundance ranking which indicated clear departure from the neutral hypothesis (The neutral model assumes functional equivalence and therefore predicts the absence of such a correlation). Further, Steele et al. [28] pointed out that a departure of the empirical frequency

---

[1] A different approach was taken by Bentley et al. [3] who based their consideration on Kimura and Crow's [20] result for the stationary distribution of variants possessing a certain frequency. In order to account for accumulation of cultural variants they developed a simulation approach and derived the expected magnitude of the power law slope between the variant frequencies and the number of variants occurring at that frequency under the assumptions of neutral evolution and different values for population size and mutation rates.

[2] In this context, Ewens ([12], p.105) noted that "any frequency configuration, and in particular that pertaining to a selectively neutral situation, can be explained by one or other selective system."

distribution from that expected under neutrality need not indicate departure from neutrality if the cultural system is not at equilibrium. Similarly, Blythe [6] argued that adding spatial structure to the neutral model can lead to behaviour different from that predicted by an unstructured model.

Here we develop a temporally explicit neutral model, which does not rest on the equilibrium assumption and therefore allows the analysis of the cultural system under consideration at any point in time. While equilibrium models evaluate the likelihood of the composition of an assemblage given the observed number of cultural variants under neutrality, our framework determines how likely it is that observed changes in the composition of the assemblage (especially changes in the number of present variants) between two consecutive time points can be produced by neutral evolution. Consequently our framework is only applicable to data sets which contain at least two data points describing the frequency distribution of the assemblage under consideration at different times. Our focus is on archaeological applications, which are often characterized by sparse data: frequency data of an assemblage of cultural artefacts at some consecutive time points are often the only available information. Utilizing this input data we base our consideration on the infinite alleles Wright-Fisher model but use tree-structured Markov processes to account for varying population sizes and mutation rates (We refer to “population” as the set of cultural artefacts present at time t.).

We note that the inherent assumption of the neutral model might be violated in the cultural context. Any act of copying, for example, of a ceramic motif, is unlikely to be random in terms of the model copied. However, if each act of copying one motif rather than another has a different motivation, the population-level result will be that there are no directional selective forces affecting what is copied [24]. The proposed framework aims to identify the existence of those directional selective forces by comparing the evolutionary dynamic obtained under the hypothesis of neutral evolution with the observed dynamic within the assemblage.

2. Model

Our framework tracks the frequency changes of different cultural variants over time. In particular we consider a situation where initially k different variants of a cultural trait are present at the frequencies n_i, $i=1,\dots,k$ and the composition of the set of cultural variants in the next time step is determined by a certain number of independent random copying events. The variants are regarded as selectively neutral and therefore their probability of being copied in the next time step is proportional to their relative frequency. Further, we assume that every copying event is only faithful with a probability of $1-\mu$. The mutation rate μ describes the probability of introducing a new (and different from the present) variant while copying one of the existing variants. By repeating this procedure we obtain the frequency distributions of the set of cultural variants at different times under the assumption of neutral evolution. The model takes the variant’s point of view and we refer to “population” as the set of cultural variants present at time t.

To describe the process of neutral evolution we model the frequencies $X_i(t)$ of the cultural variants i, $i=1,\ldots,k$ after t time steps as tree-structured Markov processes $X_i(t)$ with the distributions

$$P(X_i(t)) = \xi_i \prod_{s=1}^{t} M(s) \tag{1}$$

where $\xi_i=[0,\ldots,1,\ldots,0]$ is a unit row vector with 1 corresponding to the $X_i(0)$-th state (where $X_i(0)$ describes the initial frequency of variant i) and zero elsewhere (e.g. [14]). The transition matrices $M(s)$, $s=1,\ldots,t$ are defined by $M(s) = \begin{bmatrix} p_{00}(s) & p_{01}(s) & p_{0N(s)}(s) \\ p_{10}(s) & p_{11}(s) & p_{1N(s)}(s) \\ & & \\ p_{N(s-1)0}(s) & p_{N(s-1)1} & p_{N(s-1)N(s)} \end{bmatrix}$ and contain the probabilities $p_{mn}(s)$ for a variant possessing frequency m at time step $(s-1)$ to be copied n out of $N(s)$ times and therefore to reach frequency n at time step s. The variable $N(s) = N\rho(s)$ describes the total number of cultural variants which are present at time step s. In more detail the probabilities $p_{mn}(s)$ are given by

$$p_{mn}(s) = \binom{N(s)}{n} \pi_m(s)^n (1-\pi_m(s))^{N(s)-n} = \binom{N(s)}{n} \left(\frac{m(1-\mu(s))}{N(s-1)}\right)^n \left(1-\frac{m(1-\mu(s))}{N(s-1)}\right)^{N(s)-n} \tag{2}$$

[17] where $\pi_m(s) = \frac{m(1-\mu(s))}{N(s-1)}$ stands for the probability that a variant with the relative frequency $m/N(s-1)$ is copied and consequently $1-\pi_m(s) = 1-\frac{m(1-\mu(s))}{N(s-1)}$ for the probability that this variant is not copied. The mutation rate at time step s is given by $\mu(s)$. Summarising, the row vector $\xi_i \prod_{s=1}^{t} M(s)$ contains the probabilities that variant i with the initial frequency $X_i(0)$ reaches the frequencies $[0,1,\ldots,N(t)]$ after t time steps under neutral evolution.

In order to apply this model to a data set containing frequency information of an assemblage of k cultural variants at different points in time we use the observed frequency distribution $\boldsymbol{y}(t_1)=[y_1(t_1),\ldots, y_k(t_1)]$ at time t_1 as initial condition and determine statistical properties of the assemblage of these k variants at time t_2. Thereby the assemblage at t_2 is generated under the assumption of neutral evolution and in this way we evaluate how likely it is that the observed frequency changes between t_1 and t_2 could have been produced by random copying alone (It is assumed that t time steps and therefore t opportunities for copying occur between those time points.).

2.1. Detecting departures from neutral evolution

We define the random variable $K(t)$ which describes the number of the initially k present variants which remain present in the assemblage at time t and determine the probability distribution $P(K(t) = l|\boldsymbol{X}_0, \mathbb{H})$ under the hypothesis $\mathbb{H}$: 'neutral evolution' and the initial condition $\boldsymbol{X}_0$. Naturally the distribution $P(K(t) = l|\boldsymbol{X}_0, \mathbb{H})$ is time-dependent and it holds

$\lim_{t\to\infty} P(K(t) = 1|\boldsymbol{X}_0, \mathbb{H}) = 1$ if $\mu=0$,

$\lim_{t\to\infty} P(K(t) = 0|\boldsymbol{X}_0, \mathbb{H}) = 1$ if $\mu>0$.

Without mutation ($\mu=0$) diversity cannot be maintained and only one variant will survive in the long run. In contrast for $\mu>0$, all initially present variants will go extinct after sufficiently many time steps. Further, it is known ([13,20]) for $\mu=0$ and constant population size ($N(t) = N, \forall t$)

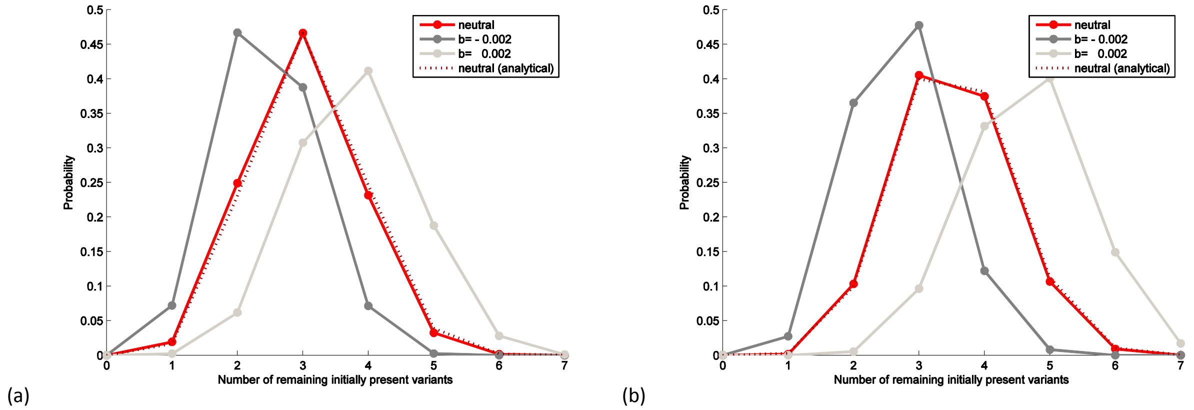
that the probability of at least $j \leq k$ variants remaining after t time steps is proportional to λ_j^t where $\lambda_j \leq 1$ is the jth eigenvalue of the transition matrix M (The multiplicity of λ_j, $j>1$ in the Wright-Fisher k-variant model is $(k+j-2)!/[(k-2)!j!]$ and k for $\lambda_1=1$.). Consequently a high number of remaining variants becomes increasingly unlikely with progressing time under neutral evolution.

More generally, the probability distribution $P(K(t) = l|\boldsymbol{X}_0, \mathbb{H})$ can be determined by

$$P(K(t) = l|\boldsymbol{X}_0, \mathbb{H}) = \sum_{s=1}^{l} (-1)^{l-s} \binom{k-s}{l-s} \sum P_0\left(t, p_{i_1} + \cdots + p_{i_{k-s}}\right), \quad l = 1, \dots, k-1, \qquad (3)$$

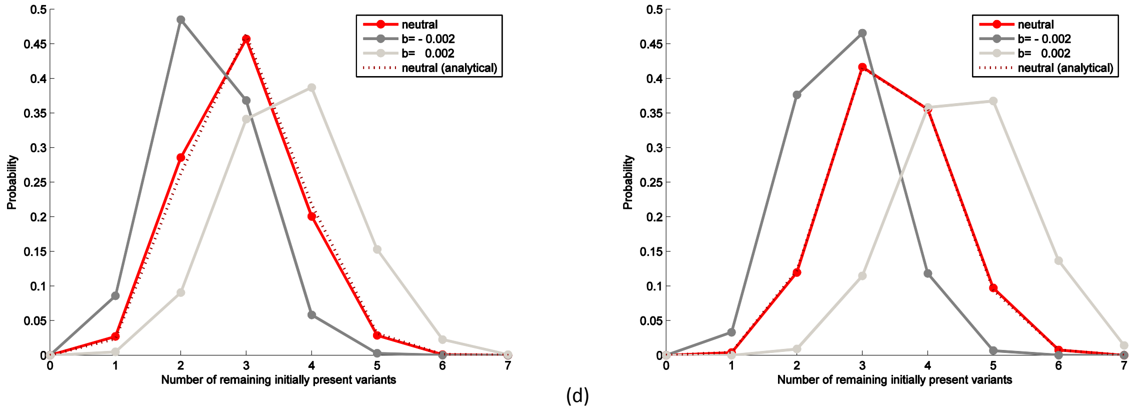
with $P(K(t) = 0|\boldsymbol{X}_0, \mathbb{H}) = P_0(p_1 + \cdots + p_k, t)$ and $P(K(t) = k|\boldsymbol{X}_0, \mathbb{H}) = 1 - \sum_{l=0}^{k-1} P(K(t) = l|\boldsymbol{X}_0, \mathbb{H})$. The second summation in equation (3) is taken over all possible combinations of (k-s) choices from the set {1,…,k} and P_0 describes the extinction probability of a variant with relative frequency $p_{i_1} + \cdots + p_{i_{k-s}}$ in the two-variant scenario [20]. The probability P_0 depends on mutation rate and population size and is obtained by diffusion approximations based on the Kolmogorov backward equation. The complete derivation of equation (3) and the calculation of the extinction probability P_0 can be found in Appendix A.

In order to illustrate equation (3) we assume an assemblage of seven variants with the initial frequencies $\boldsymbol{X}_0$=[1 2 5 10 20 40 60] and determine $P(K(t) = l|\boldsymbol{X}_0, \mathbb{H})$ after t=50 time steps. Figure 1 shows the distribution of the number of the seven initially present variants which remain in the assemblage assuming mutation rates μ=0; 0.0025 and constant and temporally varying population sizes. The changes in population size occur gradually over the considered time interval and we set $N(t) = N\rho(t)$ with $\rho(t) = \frac{r-1}{t_{max}} t + 1$ where r describes the relative change between the beginning and the end of the interval. It is obvious that changes in population size impact on the distribution and therefore have to be accounted for. We note that the timing of the changes will influence the shape of the distribution. On the other hand mutation rates of the order considered here ($\mu \sim O(1/N)$) do not influence the probability distribution greatly. The red dotted lines in Figure 1 show the distributions obtained by the diffusion approximation (3) and the red solid lines the distributions obtained by simulations. The results coincide very well and therefore the diffusion approximation (3) is applicable also in situations with relatively small population sizes.



(a)

(b)



(c) (d)

Figure 1. Probability distributions of the number of remaining initially present variants ($\boldsymbol{X}_0$=[1 2 5 10 20 40 60]) under the assumptions (a) $\mu = 0, \rho(t) = 1$ (b) $\mu = 0, \rho(t) = \frac{r-1}{t_{max}}t + 1, r = 2$ (c) $\mu = 0.0025, \rho(t) = 1$ (d) $\mu = 0.0025, \rho(t) = \frac{r-1}{t_{max}}t + 1, r = 2$. The red lines are obtained under the assumption of neutral evolution (solid line: simulations, dashed line: diffusion approximation (3)). The light grey lines assume weak anti-conformity (*b*=0.002) and the dark grey lines weak conformity (*b*=-0.002). All simulation results are obtained by averaging over 100,000 runs.

Now in order to detect departures from neutrality we examine the likelihood of the existence of the observed assemblage at time t_2 by determining the probability of the observed number of remaining initial *k* variants under the assumption of neutral evolution. The focus on remaining initial variants is partly motivated by the fact that the archaeological record might provide more accurate information on those variants than on innovations, which will be very rare at first appearance and may not appear in the archaeologist's sample, especially if they are short-lived. Also the calculation of the likelihood requires only information about presence and absence of variants and does not involve the whole frequency distribution at time t_2.

2.2 Sensitivity of $P(K(t) = l|X_0, \mathbb{H})$ to changes in the underlying process

Before we go on and apply the developed framework to an archaeological data set we analyse how the distribution $P(K(t) = l|\boldsymbol{X}_0, \mathbb{H})$ changes when the underlying process differs from random copying. In other words we ask whether different evolutionary processes can be distinguished on the base of the probability distribution of the number of remaining initially present variants. In the following we investigate the influence of frequency-dependent selection on $P(K(t) = l|\boldsymbol{X}_0, \mathbb{H})$.

So far variants have been copied from the previous generation at random. As a result the variant's probability of being copied varies in response to how common the variant is. Now we assume that frequency information is used to either disproportionately prefer high-frequency to low-frequency variants or low-frequency to high-frequency variants. Boyd and Richerson [7] called such frequency-dependent biases conformity and anti-conformity, respectively. While conformity produces an exaggerated tendency to follow the majority [11], anti-conformity supports low-frequency variants disproportionately. To account for such biases in model (2) we redefine the probability of copying a variant with the frequency *m* by

$\pi_m(t) = \left(\frac{m(1-\mu)}{N(t-1)} + f\left(\frac{m}{N(t-1)}\right)\right)\bar{w}^{-1}$ with $\bar{w} = \sum_{i=1}^{\bar{k}} \pi_i\,(t)$. The function *f* is a greater than zero if *m/N(t-1)* exceeds the 'relative majority' $1/\bar{k}$ ($\bar{k}$ stands for the number of different variants at time *t*). And conversely, if the variant's frequency is below the threshold $1/\bar{k}$, the function *f* is smaller than zero[3]. Consequently for *b*<0, the probability that a variant is copied increases disproportionately if *m/N(t-1)* exceeds the 'relative majority' $1/\bar{k}$ while for *b*>0 the reverse is true.

Figure 1 illustrates the influence of frequency-dependent selection on the probability distribution $P(K(t) = l|\boldsymbol{X}_0, \mathbb{H})$ under $\mathbb{H}$: 'weak conformity, *b*=-0.002' (dark grey lines) and $\mathbb{H}$: 'weak anti-conformity, *b*=0.002' (light grey lines) (These distributions are obtained by using simulation techniques). It is obvious that the distributions are shifted to the left and right, respectively. Consequently observing a certain number of remaining initially present variants after 50 time steps is differently likely under the three assumed evolutionary models.

Of course the number of possible underlying evolutionary models is vast and Figure 1 only shows that neutral evolution and frequency-dependent selection can produce different probability distributions. However, the developed framework can be generalized to investigate the consequences of different cultural mechanisms by formalizing the appropriate copying probabilities π_m in equation (2) and simulation techniques are able to produce the probability distribution $P(K(t) = l|\boldsymbol{X}_0, \mathbb{H})$ easily[4].

Now if one wants to go on and analyse whether certain selective forces are consistent with observed changes in frequency, Figure 2a points to an interesting fact. It shows that almost every number of remaining variants can be obtained by varying the strength of the frequency-dependent bias. However, so far we only analysed the likelihood of the observed number of remaining initially present variants at time *t* and neglected information about the relative frequencies of these variants. While it might not be surprising that appropriate selection processes are able to produce the observed number of remaining initial variants it is less obvious that the frequency ranges for each variant at time *t* under the assumed selection process are consistent with the observed frequencies. In order to explore this fact we determine the frequency distributions $P(X_i(t)|\boldsymbol{X}_0, \mathbb{H})$ of the random variables $X_i(t)$, *i*=1,…,*k* under the assumed evolutionary hypothesis $\mathbb{H}$ at time *t* and determine the corresponding (1-α)100% confidence interval of the distribution. If the observed frequency y_i of variant *i* lies outside confidence interval we can conclude that the observed frequency change is not consistent with the assumed evolutionary process. Confidence intervals are used in the following to identify evolutionary processes which could not have produced the observed frequency changes. We note that the use of the (1-α)100% confidence intervals is not effective for detecting departures from neutrality. Neutral evolution leads to wide distributions $P(X_i(t)|\boldsymbol{X}_0, \mathbb{H})$ (as defined by equation (1)) and consequently wide confidence intervals,

---

[3] We assume a function of the form $f_b(p)=ap+b$ with the relation $a=-\bar{k}b$ where $\bar{k}$ is the number of different variants present at time *t*. The intersection with the x-axis is given by $1/\bar{k}$ and therefore the family of functions f_b models (for varying *b*) the effects of conformity (*b<0*) and anti-conformity (*b>0*) of different strengths.

[4] We note that the chosen nature of the framework allows us to model the effects of different properties of the cultural variants on the evolutionary dynamic but not of properties related to individuals possessing the cultural variants.

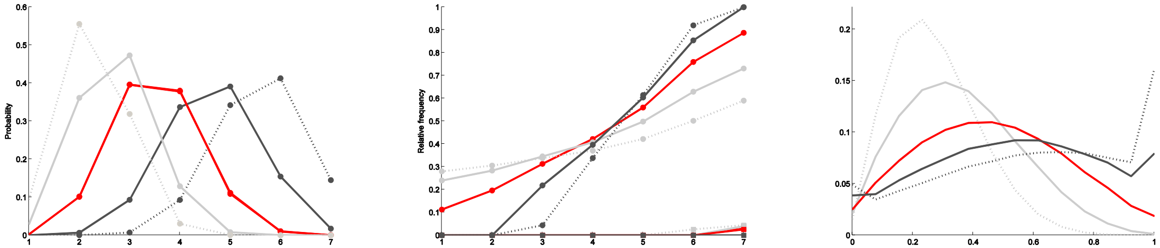
especially for medium and high frequency variants, which indicates that a broad range of frequency values are consistent with the neutral theory.

In order to compare the likelihoods of the observed frequencies $\boldsymbol{y}(t)$ of the assemblage of cultural variants under different evolutionary hypotheses $\mathbb{H}$ at time t we define

$$P(\boldsymbol{y}(t)|\boldsymbol{y}_0, \mathbb{H}) = \frac{1}{k}\sum_{i=1}^{k} \frac{p_i}{p_{max,i}} \tag{4}$$

with $p_i = P(X_i(t) = y_i(t)|\boldsymbol{y}_0, \mathbb{H})$, $p_{max,i} = \max P(X_i(t)|\boldsymbol{y}_0, \mathbb{H})$ and $\boldsymbol{y}_0$ describing the frequencies of the cultural variants at time t=0. The probability $P(\boldsymbol{y}(t)|\boldsymbol{y}_0, \mathbb{H})$ indicates how likely the observed frequencies $y_i(t)$ of the k initially present variants can occur under the assumed evolutionary process. It holds: the higher $P(\boldsymbol{y}(t)|\boldsymbol{y}_0, \mathbb{H})$ the more consistent are the observed frequency changes with hypothesis $\mathbb{H}$. Depending on the available data it might be appropriate to use a weighted average $P(\boldsymbol{y}(t)|\boldsymbol{y}_0, \mathbb{H}) = \sum_{i=1}^{k} \frac{p_i}{p_{max,i}} w_i$ with $\sum_{i=1}^{k} w_i = 1$ if more or less emphasis should be given to special variants.

Figure 2b shows that the presence of selective forces produces a more distinct pattern of the possible frequency ranges after 50 time steps than neutral evolution. Strong anti-conformity (light grey dashed lines) results in confidence intervals which are similar for all variants regardless of their initial condition. In contrast, strong conformity (dark grey dashed lines) results in large frequency ranges for initially high-frequency variants only. Figure 2c shows the approximated frequency distribution $P(X_7(t)|\boldsymbol{X}_0, \mathbb{H})$ exemplarily for variant 7 after t=50 time steps under the different hypotheses. It is obvious that frequency-dependent selection alters this distribution and therefore the observed frequency $y_7(t)$ will be differently likely under different assumptions.



(a) (b) (c)

Figure 2. (a) Probability distribution $P(K(50) = l|\boldsymbol{X}_0, \mathbb{H})$ with $\boldsymbol{X}_0$=[1 2 5 10 20 40 60] under the assumption of neutral evolution (red lines), conformity (b=-0.002 (dark grey solid line), b=-0.004 (dark grey dashed line)) and anti-conformity (b=0.002 (light grey solid line), b=0.004 (light grey dashed line)) (b) corresponding 95% confidence intervals of the possible frequency ranges for each variant (c) corresponding probability distribution $P(X_7(50)|\boldsymbol{X}_0, \mathbb{H})$ of variant 7 after t=50 time steps. The distributions $P(X_i(t)|\boldsymbol{y}_0, \mathbb{H})$ under the considered selection hypotheses $\mathbb{H}$ are obtained by using simulations and averaging over 100,000 runs.

Summarizing, we suggest a two-step analysis. If the probability $P(K(t) = n|\boldsymbol{y}_0, \mathbb{H})$ of the observed number (denoted by n) of remaining initially present variants under the assumed evolutionary hypothesis $\mathbb{H}$ has some positive value, we secondly examine the likelihood of the observed frequencies $\boldsymbol{y}(t)$ of the considered assemblage under $\mathbb{H}$ in the form of the (1-α)100% confidence intervals of the frequencies of the different variants and the probability $P(\boldsymbol{y}(t)|\boldsymbol{y}_0, \mathbb{H})$ given in equation (4).

3. Case study

To demonstrate the applicability of our approach we apply the developed model to an archaeological dataset. Like the earlier case-study by Shennan and Wilkinson [25], this one too focuses on the decorated pottery of the earliest farmers in Central Europe, the so-called Linear Pottery Culture (or LBK after its German name), c.7500-7000 years ago. Not only are the settlements of this culture well known but detailed quantitative seriation studies of the decorated pottery, supported by spatial analyses of the distributions of houses and other features, have established fine-grained chronological sequences for the 400-500 years that the culture lasted. Following the well-established pattern in LBK studies, Strien [29] carried out a correspondence analysis-based seriation of features from the LBK settlements of southwest Germany on the basis of the frequencies of the vessels decorated with different types of motif. He identified a chronological trend, which he divided into 8 phases. He was able to link this sequence to the sequence of building phases at the sites concerned and, following previous work on the chronology of the Linear Pottery Culture [22], proposed that some of the seriation phases lasted for more building phases than others. Each building phase, or house generation, has been shown to be last for c.25 years [22]. On this basis, the estimated lengths of the phases (in years) are as follows: Phase I[5] 62; phase II 62; phase III 50; phase IV 25; phase V 50; phase VI 25; phase VII. In temperate climates such as Central Europe pottery making is largely seasonal and generally takes place in one bout of activity per year ([1], chapter 3). Accordingly, the lengths of the phases are likely to correspond to the number of potential opportunities for copying within a given phase and thus the number of opportunities for the population of decorative motifs to change as a result of the operation of drift, innovation and other potential forces. Table 1 gives an overview of the characteristics of the different phases. The mutation rate μ quantifies the probability of introducing a new (and different from the present) variant while copying one of the existing variants and we use our modelling framework to determine the average mutation rate needed in order to obtain the observed number of innovations under neutral evolution.

	Phase I→II	Phase II→III	Phase III→IV	Phase IV→V	Phase V→VI	Phase VI→VII
Duration in years	62	50	25	50	25	25
ΔN	-62	166	47	660	-810	6
μ	$1.3\cdot10^{-3}$	$4.3\cdot10^{-3}$	$2.9\cdot10^{-3}$	$1.3\cdot10^{-3}$	$2.4\cdot10^{-4}$	$4.8\cdot10^{-3}$

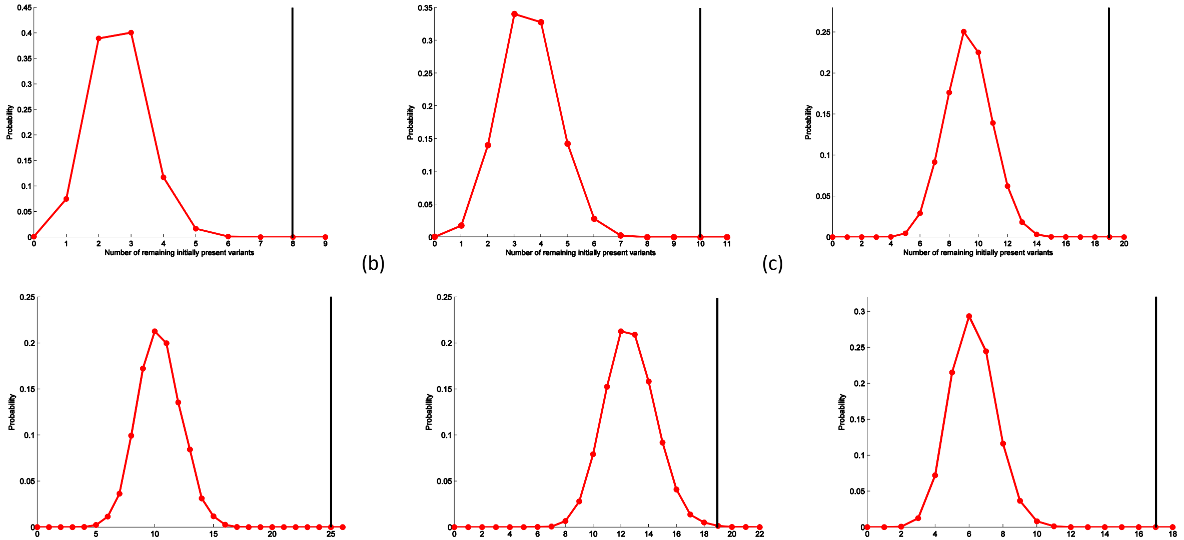
Table 1. Duration, difference in the total number of pots with different decorative motives at the end of two consecutive phases (ΔN) and the mutation rates for different phases.

Figure 3 shows the probability distribution of the number of remaining initially present variants at the end of each phase under the hypothesis of neutral evolution. The vertical lines indicate the observed number of variants. It is obvious that the observed number does not coincide with the hypothesis of neutral evolution; too many variants are maintained. We note

---

[5] The data for phase 1 in [28] has been removed for this analysis because its archaeological record is not comparable with those of the later phases; the phases in this analysis have been re-numbered accordingly.

that the assumed one copying opportunity per year is likely to be a minimum estimate. Increasing the number of copying opportunities results in probability distributions even more shifted to the left. Consequently we reject the hypothesis of neutral evolution and conclude the existence of collective selective forces in the shaping of the frequency distributions. Further, changes in population size and their timing have an influence on the probability distribution.



(a) (b) (c)

(d) (e) (f)

Figure 3. Probability distribution $P(K(t) = l|\boldsymbol{y}_0, \mathbb{H})$ obtained by the diffusion approximation (3) under the assumption of neutral evolution and the change in population size occurring gradually over the whole phase (a) transition from phase I to II; observed number: 8 (b) transition from phase II to III; observed number: 10 (c) transition from phase III to IV; observed number: 19 (d) transition from phase IV to V; observed number: 25 (e) transition from phase V to VI; observed number: 19 (f) transition from phase VI to VII; observed number: 17.

Now one may ask what kind of selective forces could have shaped the observed frequency distributions over time. As already mentioned the list of possible candidates is long and we will only show for the example of frequency-dependent selection, in particular of anti-conformity[6], how the developed framework can be used to test for evolutionary forces other than neutral evolution.

Figure 4a shows the probability distribution $P(K(t) = l|\boldsymbol{y}_0, \mathbb{H})$ under hypothesis $\mathbb{H}$: 'anti-conformity with strength *b*' (*b*=0.005;0.01;0.015;0.02) for the transition from phase II to phase III (The complete analysis of all phases can be found in Appendix B.). It is obvious that only a very strong preference for low-frequency variants (solid (*b*=0.015) and dashed (*b*=0.02) lines in Figure 4a) can maintain the high number of remaining initially present variants. However, these strong anti-conformist biases result in frequency ranges for different variants which are not consistent with the observed frequencies $\boldsymbol{y}(t)$. Figure 4b shows the 95% confidence intervals for each variant and for *b*=0.015;0.2 (solid and dashed lines) four out of the ten variants have frequencies outside the ranges possible under the assumed anti-conformity hypotheses. This is also reflected in the probability $P(\boldsymbol{y}(t)|\boldsymbol{y}_0, \mathbb{H})$ of observed frequencies $\boldsymbol{y}(t)$

---

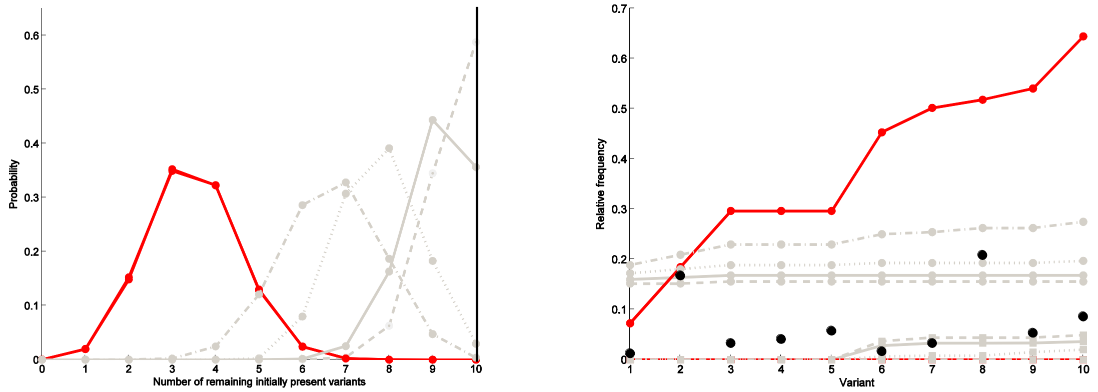
[6] As consistently in all phases the number of observed variants is larger than expected under neutral evolution conformist bias could not have produced the observed frequency changes.

after *t*=62 time steps (see Table 2). $P(\boldsymbol{y}(t)|\boldsymbol{y}_0,\mathbb{H})$ has the highest value for *b*=0.01 indicating that on average the observed frequencies are most consistent with the hypothesis $\mathbb{H}$: 'anti-conformity, *b*=0.01'. However in this case, the observed number of ten remaining initially present variants occurs only with a probability of 0.023 (cf. dotted line in Figure 4a).

	b=0	*b*=0.005	*b*=0.01	*b*=0.015	*b*=0.02
$P(\boldsymbol{y}(t)|\boldsymbol{y}_0,\mathbb{H})$	0.1508	0.4374	0.5211	0.4167	0.3468

Table 2. Probability $P(\boldsymbol{y}(t)|\boldsymbol{y}_0,\mathbb{H})$ under $\mathbb{H}$: 'anti-conformity with strength *b*' after *t*=62 time steps.

Therefore in order to explore which selective processes could and more importantly could not have produced the observed changes in frequency, both, the probability $P(K(t)=n|\boldsymbol{y}_0,\mathbb{H})$ of the observed number (denoted by *n*) of remaining initially present variants and the likelihood of the observed frequencies $\boldsymbol{y}(t)$ of the assemblage being considered under the assumed evolutionary hypothesis $\mathbb{H}$ in the form of (1-α)100% confidence intervals of the different variants and equation (4), should be investigated.



(a) (b)

Figure 4. (a) Probability distribution $P(K(62)=l|\boldsymbol{y}_0,\mathbb{H})$ for the transition from phase II to phase III under the assumption of neutral evolution (red lines) and anti-conformity with *b*=0.005 (dash-dotted line), *b*=0.01 (dotted line), *b*=0.015 (solid line), *b*=0.02 (dotted line) (b) corresponding 95% confidence intervals of the possible frequency ranges for each variant after 62 time steps.

We obtained a similar behaviour across all phases with the exception of the transition from phase V to phase VI. Here anti-conformity with *b*=0.001 could explain the observed frequency changes. Summarising, we found no consistent evidence that the observed frequency changes in decorative motifs on LBK pottery over the 400-500 year period were caused solely by neutral evolution or frequency-dependent selection in the form of conformity and anti-conformity.

4. Summary and discussion

We developed a non-equilibrium neutral model to investigate whether observed changes in the frequency distribution of a set of cultural variants at two consecutive time points t_1 and t_2 are consistent with the hypothesis of neutral evolution. The model is based on the well-known

Wright-Fisher model but we used tree-structured Markov processes to account for temporally changing population sizes. We studied the likelihood of the existence of the observed assemblage under neutral evolution and calculated the theoretical probability distribution $P(K(t) = l|\boldsymbol{y}_0, \mathbb{H})$ of the number of remaining initially present variants at the end of the considered time period. The distribution $P(K(t) = l|\boldsymbol{y}_0, \mathbb{H})$ can be found analytically by using diffusion approximations of the form (3). Based on this theoretical result we can derive the probability of the observed number of remaining initially present variants under neutral evolution. Small probabilities indicate that neutral evolution is not able to maintain the observed number of variants throughout the phase. We found that variations in population size have an effect on the change dynamic and need to be accounted for. Further the probability distribution $P(K(t) = l|\boldsymbol{y}_0, \mathbb{H})$ is a meaningful indicator for detecting departures from neutrality as the distribution is altered by small changes in the underlying process.

We stress that the framework developed here assumes that the data set being analysed is a complete description of the assemblage of cultural variants being considered (or at least provides a representative sample). Future work is aimed at generalizing the model to investigate sampling issues.

On the base of the theoretical probability distribution $P(K(t) = l|\boldsymbol{y}_0, \mathbb{H})$ we were able to reject the neutral hypothesis for the decorated LBK pottery data set. We found that too many, especially initially low-frequency, variants are maintained in the population over a relatively long time period. In this context it is worth pointing out that if our estimate of the number of copying opportunities per phase as one per year is incorrect, it is likely to be biased towards the very low end of the range; if the actual number was much higher, with multiple copying opportunities per year, these results apply with all the more force. In the light of these findings we stress the importance of initially low-frequency variants in order to detect departures from the random drift hypothesis. While under the random drift assumption variants with a low initial frequency are likely to go extinct this might not be the case under selection. The presence of a positive selection pressure in favour of low-frequency variants may increase their frequencies more quickly and steadily than expected under pure random drift.

The proposed framework can be generalized to investigate the consequences of different selection process by formalizing the appropriate copying probabilities π_m in equation (2). While especially for larger population sizes it is not surprising that certain selection processes can produce the observed number of remaining cultural variants, it is less obvious that the observed frequencies of the remaining variants will coincide with the frequency ranges possible under the assumed selection process. Therefore we suggest concentrating on two characteristics when analysing whether observed frequency changes are consistent with a particular selection process: firstly, the probability distribution $P(K(t) = l|\boldsymbol{y}_0, \mathbb{H})$ of the number of remaining initially present variants under the assumed evolutionary hypothesis $\mathbb{H}$, and secondly, the likelihood of the observed frequencies $\boldsymbol{y}(t)$ of the considered assemblage under $\mathbb{H}$ in the form of the $(1-\alpha)100\%$ confidence intervals of the frequencies of the different variants and the probability $P(\boldsymbol{y}(t)|\boldsymbol{y}_0, \mathbb{H})$ given in equation (4). While the confidence intervals provide information about evolutionary processes which could not have produced the observed frequency changes $P(\boldsymbol{y}(t)|\boldsymbol{y}_0, \mathbb{H})$ quantifies the probability of the observed frequencies $\boldsymbol{y}(t)$ of the assemblage of cultural variants under different evolutionary hypotheses. The use of simulation techniques allows for a relatively easy calculation of these characteristics under different hypotheses $\mathbb{H}$.

In this paper we only explored the effects of frequency-dependent selection and in regards to the decorated LBK data we found that only the observed frequency changes from phase V to VI are consistent with an anti-conformity hypothesis. As both transmission hypotheses, random drift and frequency-dependent selection, are not, or only partly, supported by the data we conclude the existence of selective forces acting on decorated LBK pottery. These findings have now motivated research into developing and applying selection models to archaeological data, as started by Brantingham and Perreault [5].

Acknowledgements. AK was supported by a Leverhulme Trust Early Career Fellowship and by an Omidyar Fellowship at the Santa Fe Institute. SS would like to thank Tim Kerig for discussions on the chronology of LBK pottery and the European Research Council for its Advanced Grant for project 249390, EUROEVOL, Cultural Evolution of Neolithic Europe, which has made possible his work on this paper. We thank three anonymous reviewers for their helpful comments on the manuscript.

References

[1] D.E. Arnold, *Ceramic Theory and Cultural Process*, Cambridge University Press, 1985.

[2] R.A. Bentley and S.J. Shennan, Cultural transmission and stochastic network growth, *Am. Antiq.* **68**:3 (2003) 459-485.

[3] R.A. Bentley, W. Hahn and S.J. Shennan S.J., Random drift and culture change, *Proc. R. Soc. Lond. B* **271** (2004) 1443-1450.

[4] L.R. Binford, "Red Ocher" Caches from the Michigan Area: A Possible Case of Cultural Drift. *Southwestern Journal of Anthropology* 19 (1963) 89-108.

[5] P.J. Brantingham and C. Perreault, Detecting the effects of selection and stochastic forces in archaeological assemblages. *J. Archaeol. Sci.* **37** (2010) 3211-3225.

[6] R.A. Blythe, Random copying in space. *Adv. Complex Syst.* **15** (2012) 3, DOI: 10.1142/S0219525911003396.

[7] R. Boyd and P.J. Richerson, *Culture and the evolutionary process*, University of Chicago Press, 1985.

[8] L.L. Cavalli-Sforza and M.W. Feldman, *Cultural transmission and evolution*, Princeton University Press, 1981.

[9] J.F. Crow and M. Kimura, *An introduction to population genetics theory*, New York: Harper & Row, 1970.

[10] R.C. Dunnell, Style and function: a fundamental dichotomy. *Am. Antiq.* **43** (1978) 192–202.

[11] C. Efferson, R. Lalive, P.J. Richerson, R. McElreath and M. Lubell, Conformists and mavericks: the empirics of frequency-dependent cultural transmission. *Evol. Hum. Behav*. **29**:1 (2003) 56-64.

[12] W. Ewens, The sampling theory of selectively neutral alleles, *Theor. Popul. Biol.* **3** (1972) 87-112.

[13] W. Ewens, *Mathematical Population Genetics,* second ed., Springer, New York, 2004.

[14] J.C. Fu and W.Y.W. Lou, *Distribution theory of runs and patterns and its applications.* World Scientific Publishing Co. Pte. Ltd, 2003.

[15] M.W. Hahn and R.A. Bentley, Drift as a mechanism for cultural change: an example from baby names. *Proc. R. Soc. Lond. B* **270** (2003) 120-123.

[16] H.A. Herzog, R.A. Bentley and W. Hahn, Random drift and large shifts in popularity of dog breeds, *Proc. R. Soc. Lond. B* **271** (2004) 353-356.

[17] M. Iosifescu, *Finite Markov processes and their applications,* John Wiley & Sons, 1980.

[18] A. Kandler and K.N. Laland, An investigation of the relationship between innovation and cultural diversity. *Theor. Popul. Biol.* **76** (2009) 59-67.

[19] T.A. Kohler, S. VanBuskirk and S. Ruscavage-Barz, Vessels and villages: evidence for conformist transmission in early village aggregation on the Pajarito Plateau, New Mexico, *J. Anthropol. Archaeol.* **23** (2004) 100-118.

[20] M. Kimura and J.F. Crow, The number of alleles that can be maintained in a finite population, *Genetics* **49** (1964) 725–738.

[21] R.A. Littler, Loss of variability at one locus in a finite population, *Math. Bio.* **25** (1975) 152-163.

[22] J. Luening and P. Stehli, *Die Bandkeramik im Merzbachtal auf der Aldenhovener Platte.* Bonn: Habelt, 1994.

[23] F.D. Neiman, Stylistic variation in evolutionary perspective, *Am. Antiq.* **60** (1995) 7–36.

[24] S.J. Shennan, Decent with modification and the archaeological record, *Phil. Trans. R. Soc. B* **366 :**1567 (2011) 1070-1079.

[25] S.J. Shennan and J.R. Wilkinson, Ceramic Style Change and Neutral Evolution: A Case Study from Neolithic Europe, *Am. Antiq.* **66**:4 (2001) 577-593.

[26] M. Slatkin, An exact test for neutrality based on Ewens sampling distribution, *Genet. Res.* **64** (1994) 71-74.

[27] M. Slatkin, A correction to the exact test for neutrality based on Ewens sampling distribution, *Genet. Res.* **68** (1996) 259-260.

[28] J. Steele, C. Glatz and A. Kandler, Ceramic diversity, random copying, and tests for selectivity in ceramic production, *J. Archaeol. Sci.* **37** (2010) 1348-1358.

[29] H.-C. Strien, Untersuchungen zur Bandkeramik in Württemberg, *Universitätsforschungen zur Prähistorischen Archäologie* 69, Bonn: Habelt, 2000.

[30] G.A. Watterson, Heterosis or neutrality, *Genetics.* **88** (1977) 789-814.

[31] G.A. Watterson, The homozygosity test of neutrality, *Genetics.* **88** (1978) 405-417.

Appendix A

In the following we consider an assemblage of k cultural variants with relative frequencies $\mathbf{X}_0 = [x_{0,1}, x_{0,2}, \ldots, x_{0,k}]$ at $t = 0$ and derive the probability distribution of the number of these k initially present variants that remain in the assemblage at time t. We model the number of remaining initially present variants by the random variable $K(t)$ and assume that changes in variant frequencies occur through neutral evolution. The process of neutral evolution is described by the well-know Wright-Fisher model (e.g. (Ewens, 2004)) and the frequencies of the different variants over time are expressed by Markov processes $X_i(t),\ i = 1, \ldots, k$ which are characterized by the state space $S_t = \{0, \ldots, N(t)\}$ and transition matrix $P(t) = [p_{mn}(t)]_{m=0,\ldots,N(t-1),n=0,\ldots,N(t)}$ with the transition probabilities

$$p_{mn}(t) = \binom{N(t)}{n} \left(\frac{m}{N(t-1)}(1-\mu)\right)^n \left(1 - \frac{m}{N(t-1)}(1-\mu)\right)^{N(t)-n}$$

where $p_{0\cdot} = 0$. The variable $N(t)$ stands for the population size at time t and we set $N(t) = N\rho(t)$ (Evans et al., 2007). Further, μ describes the mutation rate. The distribution

$$\mathbf{P}(K(t) = l | \mathbf{X}_0, \mathbb{H}), \quad l = 0, 1, \ldots, k$$

with the hypothesis $\mathbb{H}$: 'neutral evolution' is naturally time-dependent and it holds for large t

$$\lim_{t\to\infty} P_0(K(t) = 1|\mathbf{X}_0, \mathbb{H}) = 1 \quad \text{if mutation rate } \mu = 0,$$
$$\lim_{t\to\infty} P_0(K(t) = 0|\mathbf{X}_0, \mathbb{H}) = 1 \quad \text{if mutation rate } \mu > 0.$$

The aim is to determine the distribution $\mathbf{P}(K(t) = l|\mathbf{X}_0, \mathbb{H})$ at any time t. We start our derivation by calculating the time-dependent extinction probability in a two-variant scenario and expand the obtained results to the situation of k variants.

Extinction probability $P_0(t, \mathbf{X}_0)$

In the following we assume that only two variants, A and a, are present at time $t = 0$ and determine the extinction probability $P_0(t, \mathbf{X}_0)$ of variant A at time t assuming the initial condition $\mathbf{X}_0 = [p, 1-p]$ with $p = \frac{i}{N},\ i = 1, \ldots, N-1$ and mutations from variant A to a at rate μ. We do not allow for mutations from a to A and consequently it holds

$$\lim_{t\to\infty} P_0(t, \mathbf{X}_0) = 1.$$

The eventual loss of variant A is certain. As an explicit calculation of $P_0(t, \mathbf{X}_0)$ in the Wright-Fisher model proves to be difficult we use diffusion theory (see e.g. (Gardiner, 2010; Risken, 1989)) and approximate the discrete Markov process by a continuous time and continuous space diffusion process. If space is rescaled by a factor factor of $N\rho(t)$ and time by a factor of N, then the Markov process converges to a diffusion process on the interval $[0, 1]$. This diffusion process is a good approximation for large population sizes N and mutation rates of order $O(N^{-1})$. It can be shown that the transition probability density $f(x, p, t)$ satisfies the Kolmogorov backward equation of the form

$$\frac{\partial f(x, p, \tau)}{\partial \tau} = a(p)\frac{\partial f(x, p, \tau)}{\partial p} + \frac{1}{2}b(p)\frac{\partial^2 f(x, p, \tau)}{\partial p^2} \tag{1}$$

with appropriate initial and boundary conditions. For constant population size the coefficients a and b are determined by

$$\begin{aligned}&\mathbb{E}\{X(t+1)-X(t)|X(t)=x\}=a(x)N^{-1}+o\left(N^{-1}\right)\\&\mathbb{V}ar\{X(t+1)-X(t)|X(t)=x\}=b(x)N^{-1}+o\left(N^{-1}\right)\end{aligned}$$

with $a(p) = -N\mu p$ and $b(p) = p(1-p)$ (Ewens, 2004). To derive the extinction probabilities of variant A we define

$$P_0(\tau,p)=\int_0^{0^+} f(\tau,y,p)dy$$

and obtain

$$\frac{\partial P_0(\tau,p)}{\partial\tau}=a(p)\frac{P_0(\tau,p)}{\partial p}+\frac{1}{2}b(p)\frac{\partial^2 P_0(\tau,p)}{\partial p^2} \tag{2}$$

with the initial condition $P_0(0,p) = \begin{cases}1 & \text{if } p=0\\ 0 & \text{else}\end{cases}$ and appropriate boundary conditions (Ewens, 2004). In the case of no mutation ($\mu = 0$) the solution of (2) can be found by

$$P_0(\tau,p)=(1-p)+\sum_{i=1}^{\infty}(2i+1)p(1-p)(-1)^iF(1-i,i+2,2,1-p)e^{-0.5i(i+1)\tau} \tag{3}$$

where $F(1-i, i+2, 2, \cdot)$ represents the hypergeometric function (Kimura, 1955). For $\mu > 0$ the solution of (2) can be obtained using numerical solution techniques such as the finite element method (e.g. (Zienkiewicz and Taylor, 1991)). In order to account for variable (but deterministic) population sizes we set $N(t) = N\rho(t)$ and the coefficients a and b in equation (1) have the form

$$a(p)=-N\mu p \quad \text{and} \quad b(\tau,p)=\frac{p(1-p)}{\rho(\tau)}$$

(Evans et al., 2007).

Probability distribution $P(K(t) = l|\mathbf{X}_0, \mathbb{H})$

The distribution $P(K(t) = l|\mathbf{X}_0, \mathbb{H})$ describes how likely a certain number of the k initially present variants are still present at time t. We are only interested in changes in the assemblage of these k variants and therefore assume that each mutation event results in a variant of type $k+1$. Consequently we consider a $(k+1)$-variant model however, know that variant $k+1$ cannot go extinct. Using the extinction probability $P_0(\tau, \mathbf{X}_0)$ obtained above the probability distribution $P(K(t) = l|\mathbf{X}_0, \mathbb{H})$ is given by

$$P(K(t)=l|\mathbf{X}_0,\mathbb{H})=\sum_{s=s_b}^{l}(-1)^{l-s}\binom{k-s}{l-s}\sum P_0(\tau,x_{0,i_1}+\ldots+x_{0,i_{k-s}}),\quad l=s_b,\ldots,k-1 \tag{4}$$

with $s_b = 0$ for $\mu > 0$, $s_b = 1$ for $\mu = 0$ and

$$P(K(t) = k|\mathbf{X}_0, \mathbb{H}) = 1 - \sum_{i=s_b}^{k-1} P(K(t) = i|\mathbf{X}_0, \mathbb{H})$$

(Littler, 1975). The second summation in (4) is taken over all possible combinations of $(k - s)$ choices from the set $\{1, \ldots, k\}$.

Figure 1 shows the probability distributions $P(K(t) = l|\mathbf{X}_0, \mathbb{H})$ obtained by the the diffusion approximation (4) and simulations at different times $t = 25; 50; 75; 100$. We assume the existence of seven variants at $t = 0$ with $\mathbf{X}_0 = [1\ 2\ 5\ 10\ 20\ 40\ 60]$, mutation rates $\mu = 0; 0.0025$ and constant and temporally varying population sizes with $N(t) = N\rho(t)$ ($N = 138$ and $\rho(t) = \frac{r-1}{t_{max}}t+1$, $t_{max} = 100$, $r = 1; 2$). All simulation results are obtains by averaging over 100,000 runs.

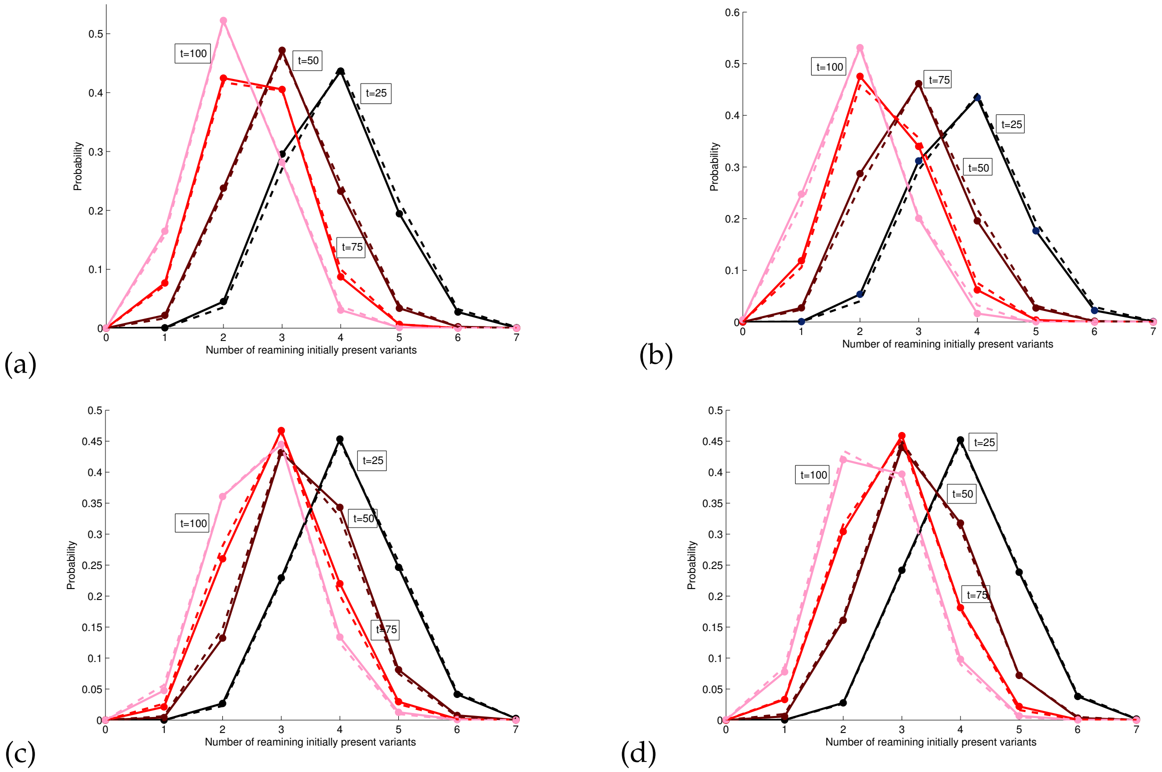


Figure 1: Probability distribution $P(K(t) = l|\mathbf{X}_0, \mathbb{H})$ with $\mathbf{X}_0 = [1\ 2\ 5\ 10\ 20\ 40\ 60]$ at times $t = 25; 50; 75; 100$ (black to pink lines). The dashed lines show the distribution obtained by the diffusion approximation (4) and the solid lines the distributions obtained by simulation for (a) $\mu = 0$, $r = 1$ (constant population size), (b) $\mu = 0.0025$, $r = 1$ (constant population size), (c) $\mu = 0$, $r = 2$ (gradually doubling of population size), (d) $\mu = 0.0025$, $r = 2$ (gradually doubling of population size).

Appendix B

In this appendix we present the complete analysis of the Linear Pottery (LBK) data set. The data set contains frequency information about pottery with specific types of decoration in 7 consecutive time periods of different length (see table 1 in the main text for details). In the following we describe the observed frequency distribution of the cultural variants at the beginning of a phase by $\mathbf{y}_0$ and at the end by $\mathbf{y}(t)$. In order to analyze whether the observed changes in frequency between the beginning and the end of the phase could have been produced by neutral evolution we calculate the probability $P(K(t) = n|\mathbf{y}_0, \mathbb{H})$ which quantifies the likelihood of the observed number (denoted by n) of remaining initially present cultural variants at the end of each phase under the assumption of neutral evolution (see Figure 3 in the main text). We found that the probability was negligible small in all phases and concluded the existence of directional selective forces. Now to explore whether anti-conformity (defined as the tendency to disproportionally support low-frequency variants (Boyd and Richerson, 1985)) could explain the observed change patterns we assume the hypotheses

$$\mathbb{H} : \text{'anti-conformity of strength } b', \ b > 0$$

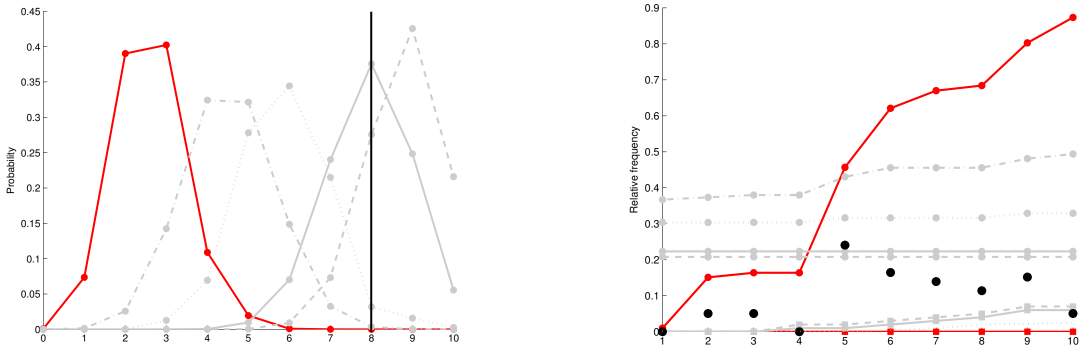
and determine firstly the probability $P(K(t) = n|\mathbf{y}_0, \mathbb{H})$ and secondly the likelihood of the frequency distribution $\mathbf{y}(t)$ at the end of the considered phase by analyzing the 95% confidence intervals of the frequency distribution $P(X_i(t)|\mathbf{y}_0, \mathbb{H})$ of variant i and the quantity

$$\beta(\mathbf{y}(t), \mathbf{y}_0, \mathbb{H}) = \frac{1}{k} \sum_{i=1}^{k} \hat{p}_i \tag{5}$$

where k stands for the observed number of variants at the beginning of the phase, $\hat{p}_i = \sum_{j \in V_i} P(X_i(t) = j|\mathbf{y}_0, \mathbb{H})$ with $V_i = \{j : P(X_i(t) = j|\mathbf{y}_0, \mathbb{H}) > P(X_i(t) = y_i(t)|\mathbf{y}_0, \mathbb{H})\}$. As the joint probability $P(X_1 = y_1(t), \ldots, X_k = y_k(t)|\mathbf{y}_0, \mathbb{H}))$ of the observed frequencies $\mathbf{y}(t)$ under the assumed hypothesis $\mathbb{H}$ becomes intractable for larger numbers of variants, population sizes and time steps β determines the average probability that the considered evolutionary process results in frequencies more likely than the observed frequencies. In more detail $\hat{p}_i$ describes the probability that under hypothesis $\mathbb{H}$ variant i possesses a frequency at time t which is more likely than the observed frequency $y_i(t)$. It holds $\hat{p}_i \in [0, 1]$ and values of $\hat{p}_i$ close to 1 indicates that only a small percentage of possible frequencies are less likely than the observed frequency $y_i(t)$. Contrary a small value of $\hat{p}_i$ indicates the observed frequencies are more likely than almost all other possible frequencies.

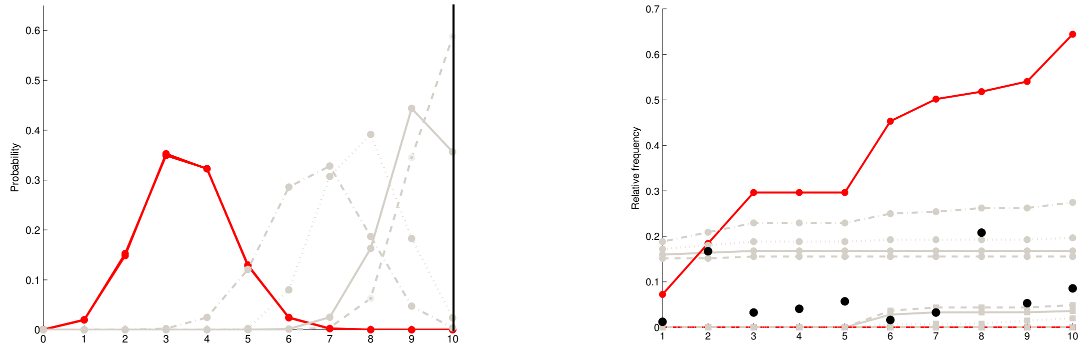
Figures 2-7a show the probability distributions $P(K(t) = l|\mathbf{y}_0, \mathbb{H})$ under the assumption of anti-conformity with different strengths b. It is obvious that in each phase the observed numbers of remaining initially present variants can be obtained with high probability under an appropriate strength $\bar{b}$. However, this strength $\bar{b}$ results in theoretical frequency ranges which do not coincide with the observed frequencies. Figures 2-7b show that apart for the transition from phase V to phase VI the frequencies of a number of variants lie outside the 95% confidence intervals of the distributions $P(X_i(t)|\mathbf{y}_0, \mathbb{H})$ under the assumption $\mathbb{H}$: 'anti-conformity with strength $\bar{b}$'. Further, Table 1 indicates that under the assumption of anti-conformity with strength $\bar{b}$ the obtained frequency distribution $\mathbf{y}(t)$ at the end of a phase has a considerable higher average probability $\beta(\mathbf{y}(t), \mathbf{y}_0, \mathbb{H})$ than under weaker anti-conformity.

Summarising, we found no consistent evidence that the observed frequency changes in decorative motifs on LBK pottery over the 400-500 year period were caused solely by neutral evolution or frequency-dependent selection in the form of anti-conformity.



(a) (b)

Figure 2: (a) Probability distributions $P(K(t) = l|\mathbf{y}_0, \mathbb{H})$ for the transition from phase I to phase II with red, solid line: neutral evolution, dash-dotted line: anti-conformity with $b = 0.005$, dotted line: anti-conformity with $b = 0.01$, solid line: anti-conformity with $b = 0.015$, dashed line: anti-conformity with $b = 0.02$. The black horizontal line indicates the observed number of variants. (b) corresponding 95% confidence intervals of the distribution $P(X_i(t)|\mathbf{y}_0, \mathbb{H})$ for each variant.



(a) (b)

Figure 3: (a) Probability distributions $P(K(t) = l|\mathbf{y}_0, \mathbb{H})$ for the transition from phase II to phase III with red, solid line: neutral evolution, dash-dotted line: anti-conformity with $b = 0.005$, dotted line: anti-conformity with $b = 0.01$, solid line: anti-conformity with $b = 0.015$, dashed line: anti-conformity with $b = 0.02$. The black horizontal line indicates the observed number of variants. (b) corresponding 95% confidence intervals of the distribution $P(X_i(t)|\mathbf{y}_0, \mathbb{H})$ for each variant.

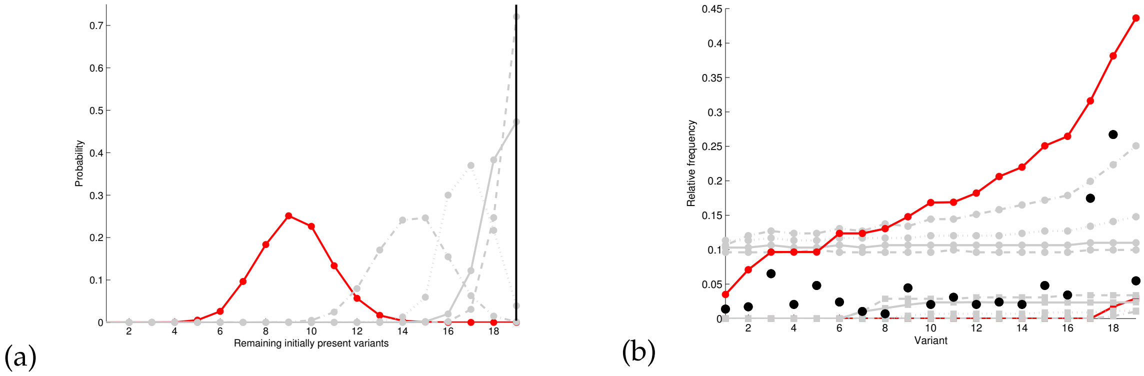


Figure 4: (a) Probability distributions $P(K(t) = l|\mathbf{y}_0, \mathbb{H})$ for the transition from phase III to phase IV with red, solid line: neutral evolution, dash-dotted line: anti-conformity with $b = 0.0025$, dotted line: anti-conformity with $b = 0.005$, solid line: anti-conformity with $b = 0.0075$, dashed line: anti-conformity with $b = 0.01$. The black horizontal line indicates the observed number of variants. (b) corresponding 95% confidence intervals of the distribution $P(X_i(t)|\mathbf{y}_0, \mathbb{H})$ for each variant.

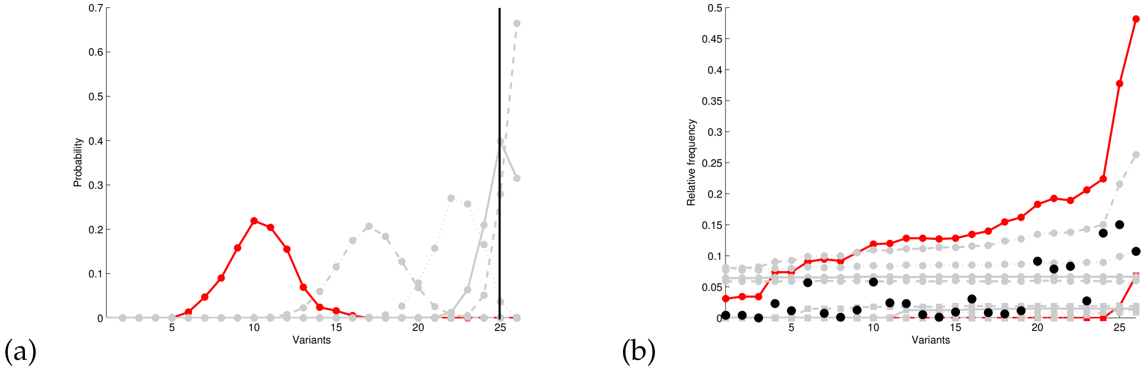


Figure 5: (a) Probability distributions $P(K(t) = l|\mathbf{y}_0, \mathbb{H})$ for the transition from phase IV to phase V with red, solid line: neutral evolution, dash-dotted line: anti-conformity with $b = 0.001$, dotted line: anti-conformity with $b = 0.0025$, solid line: anti-conformity with $b = 0.005$, dashed line: anti-conformity with $b = 0.0075$. The black horizontal line indicates the observed number of variants. (b) corresponding 95% confidence intervals of the distribution $P(X_i(t)|\mathbf{y}_0, \mathbb{H})$ for each variant.

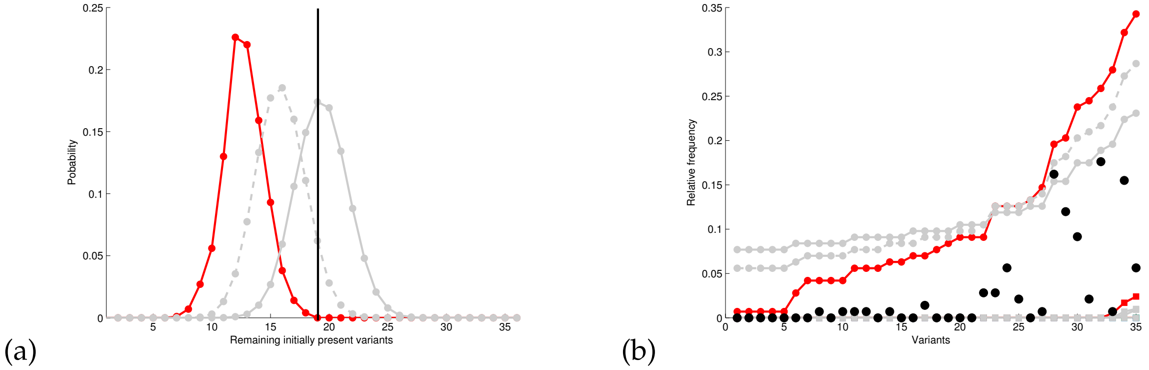


Figure 6: (a) Probability distributions $P(K(t) = l|\mathbf{y}_0, \mathbb{H})$ for the transition from phase V to phase VI with red, solid line: neutral evolution, dashed line: anti-conformity with $b = 0.0005$, solid line: anti-conformity with $b = 0.001$. The black horizontal line indicates the observed number of variants. (b) corresponding 95% confidence intervals of the distribution $P(X_i(t)|\mathbf{y}_0, \mathbb{H})$ for each variant.

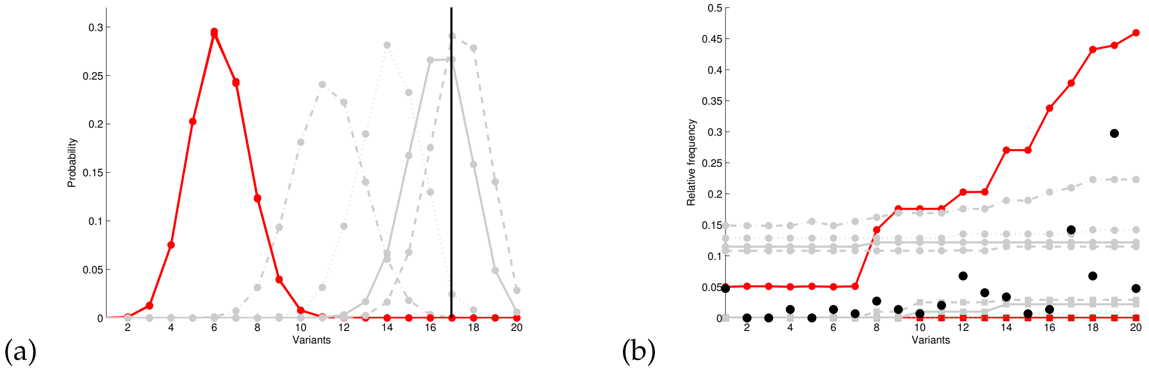


Figure 7: (a) Probability distributions $P(K(t) = l|\mathbf{y}_0, \mathbb{H})$ for the transition from phase VI to phase VII with red, solid line: neutral evolution, dash-dotted line: anti-conformity with $b = 0.005$, dotted line: anti-conformity with $b = 0.01$, solid line: anti-conformity with $b = 0.0125$, dashed line: anti-conformity with $b = 0.015$. The black horizontal line indicates the observed number of variants. (b) corresponding 95% confidence intervals of the distribution $P(X_i(t)|\mathbf{y}_0, \mathbb{H})$ for each variant.

	Phase I-II			Phase II-III	
	$\beta(\mathbf{y}(t), \mathbf{y}_0, \mathbb{H})$	$P(K(t) = n \mid \mathbf{y}_0, \mathbb{H})$		$\beta(\mathbf{y}(t), \mathbf{y}_0, \mathbb{H})$	$P(K(t) = n \mid \mathbf{y}_0, \mathbb{H})$
$b = 0$:	0.4420	0	$b = 0$:	0.6789	0
$b = 0.005$:	0.4007	0.0003	$b = 0.005$:	0.6642	0.0004
$b = 0.01$:	0.4594	0.0319	$b = 0.01$:	0.7625	0.0232
$\boxed{b = 0.015:}$	0.5906	0.3750	$b = 0.015$:	0.8054	0.3565
$b = 0.02$:	0.6149	0.2756	$\boxed{b = 0.02:}$	0.8214	0.5878

	Phase III-IV			Phase IV-V	
	$\beta(\mathbf{y}(t), \mathbf{y}_0, \mathbb{H})$	$P(K(t) = n \mid \mathbf{y}_0, \mathbb{H})$		$\beta(\mathbf{y}(t), \mathbf{y}_0, \mathbb{H})$	$P(K(t) = n \mid \mathbf{y}_0, \mathbb{H})$
$b = 0$:	0.6956	0	$b = 0$:	0.7154	0
$b = 0.0025$:	0.4871	0.0011	$b = 0.001$:	0.6524	0
$b = 0.005$:	0.5405	0.0391	$b = 0.0025$:	0.7991	0.0363
$b = 0.0075$:	0.6937	0.4731	$\boxed{b = 0.005:}$	0.8588	0.3981
$\boxed{b = 0.01:}$	0.7763	0.7208	$b = 0.0075$:	0.8800	0.2795

	Phase V-VI			Phase VI-VII	
	$\beta(\mathbf{y}(t), \mathbf{y}_0, \mathbb{H})$	$P(K(t) = n \mid \mathbf{y}_0, \mathbb{H})$		$\beta(\mathbf{y}(t), \mathbf{y}_0, \mathbb{H})$	$P(K(t) = n \mid \mathbf{y}_0, \mathbb{H})$
$b = 0$:	0.3982	0.0001	$b = 0$:	0.6237	0
$b = 0.0005$:	0.3689	0.0620	$b = 0.005$:	0.5608	0.0005
$\boxed{b = 0.001:}$	0.2813	0.1739	$b = 0.01$:	0.6240	0.0244
			$b = 0.0125$:	0.7275	0.2660
			$\boxed{b = 0.015:}$	0.7912	0.2908

Table 1: Average probability $\beta(\mathbf{y}(t), \mathbf{y}_0, \mathbb{H})$ that the evolutionary process $\mathbb{H}$ results in a frequency more likely than the observed frequency $\mathbf{y}(t)$ of the assemblage and probability $P(K(t) = n \mid \mathbf{y}_0, \mathbb{H})$ of n observed number of remaining initially present variants at the end of each phase under the hypotheses $\mathbb{H}$: 'anti-conformity with strength b'. The red boxes indicate the strength of anti-conformity which generates the observed number of variants with the highest probability.

References.

Boyd R. and Richerson P.J., 1985. *Culture and the evolutionary process*, University of Chicago Press, Chicago .

Ewens W., 2004. *Mathematical Population Genetics*, second ed., Springer, New York.

Evans S.N., Shets Y. and Slatkin M., 2007. Non-equilibrium of the allele frequency spectrum, *Theor. Popul. Biol.* 71, 109-119.

Gardiner C., 2010. *Stochastic Methods*, fourth ed., Springer, Berlin.

Kimura M., 1955. Genetic Drift in Multi-Allelic Locus, *Evolution*, 9 (4), 419-435.

Littler R.A., 1975. Loss of variability at one locus in a finite population, *Math. Bio.* 25, 152-163.

Risken H., 1989. *The Fokker-Planck Equation: Methods of Solution and Applications*, second ed., Springer-Verlag, Berlin Heidelberg.

Zienkiewicz O.C. and Taylor R.L., 1991. *The Finite Element Method: Its Basis And Fundamentals*, fourth ed., MacGraw-Hill, London.