Biological Journal of the Linnean Society, 2019, **127**, 518–533. With 5 figures.

Downloaded from https://academic.oup.com/biolinnean/article-abstract/127/2/518/5475672 by Department of Plant Physiology, Faculty of Science, Charles University user on 24 May 2019

Non-particulate inheritance revisited: evolution in systems with Parental Variability-Dependent Inheritance

PETR TUREČEK[1,*], JAKUB SLAVÍK[2], MICHAL KOZÁK[3] and JAN HAVLÍČEK[4]

[1]*Department of Philosophy and History of Science, Faculty of Science, Charles University, Viničná 7, 128 44, Prague, Czech Republic*
[2]*Department of Stochastic Informatics, Institute of Information Theory and Automation, The Czech Academy of Sciences, Pod Vodárenskou věží 4, 182 08, Prague, Czech Republic*
[3]*Department of Mathematics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University in Prague, Trojanova 13, 120 00, Prague, Czech Republic*
[4]*Department of Zoology, Faculty of Science, Charles University, Viničná 7, 128 44, Prague, Czech Republic*

Received 3 January 2019; revised 7 March 2019; accepted for publication 7 March 2019

In non-genetic systems, such as culture, inheritance is often non-particulate. Owing to blending and consequent loss of variability, however, selection in such systems has been considered ineffective. The issue of loss of variability was solved by the Galton–Pearson model, which assumes a constant offspring variability and predicts gradual adaptation regardless of model parameters. The supposition of constant offspring variability is, however, arbitrary, and it is rather unrealistic in the context of social learning, because variability of inputs may affect the resulting trait acquisition. We present an alternative non-particulate inheritance model, 'Parental Variability-Dependent Inheritance', in which offspring variability is proportional to parental variability. Results of computer simulations show that despite its simplicity, this model can, even from the same initial conditions, result in one of two stable states: successful adaptation or loss of variability. Successful adaptation is more probable in larger populations with a larger relative offspring variability and an intermediate level of selection. A third possible outcome is an unstable, chaotic increase in variability, which takes place when relative offspring variability is too large to be trimmed by selection. Without any additional assumptions, this inheritance system results in punctuated evolution.

ADDITIONAL KEYWORDS: blending inheritance – computer simulations – cultural evolution – Galton–Pearson model – swamping argument.

INTRODUCTION

THE DOWNFALL OF NON-PARTICULATE INHERITANCE

Non-particulate inheritance has been considered outdated at least since the rediscovery of the principles of genetic inheritance at the beginning of the 20th century (Nurse, 2000). The paradigm of mutation, selection and admixture of discrete trait determinants solved some of the greatest difficulties of evolutionary theory of that time and dominated evolutionary biology for decades thereafter.

The main drawback of non-particulate inheritance lies in the assumption that offspring inherit an average trait value of homologous parental traits. This model, known as the 'paint-pot' theory of inheritance, leads to a loss of variability in each generation, which implies that after several generations, all individuals in a population ought to be indistinguishable and natural selection irrelevant. The first person to draw attention to this problem was Darwin's vocal opponent Henry Fleeming Jenkin (Fleeming Jenkin, 1867). He coined the term 'swamping' for the hypothetical cases where some exceptional traits of unusually fit individuals (so-called 'sports') are, together with the accompanying reproductive advantage, dissolved in a large population of average individuals. Although the thought experiment he used to illustrate the argument was racist (Bulmer, 2004) and mathematically unsound (Davis, 1871), it seems that Darwin was unsettled

---

*Corresponding author. E-mail: petr.turecek@natur.cuni.cz by this objection. In response, he started to stress the importance of a ubiquitous, 'common' variance over exceptional qualities provided by rare mutants (Darwin, 1857, 1866, 1867), and later introduced a particulate system parallel to Mendelism (Darwin, 1868) designed to sidestep non-particulate inheritance.

Other theorists defined numerous clever add-ons to the classical blending inheritance model, which served the same purpose; frequent introduction of individuals with advantageous mutations, i.e. sports (Davis, 1871), or error terms that secure a constant (or minimal) offspring variability (Galton, 1889; Pearson, 1896). Particulate inheritance prevailed owing to its broad applicability in life sciences, whereas non-particulate models were eventually abandoned, leaving behind only a vague idea of paint-pot inheritance as an example of a paradigm our predecessors were not yet able to discard.

A SIMULATION OF CULTURAL EVOLUTION

We are currently witnessing a growing interest in studying the evolution of human cultures and artefacts by methods used in population genetics (McElreath *et al.*, 2003), cladistics (Tehrani, 2013) and bioinformatics (Barbrook *et al.*, 1998). Cultural traits are, after all, variable, heritable, transmittable and subject to selection (Mesoudi, 2011; Henrich *et al.*, 2016). In contrast, cultural evolution differs from biological evolution, at whose core are the processes of mutation and selection, in some key aspects, such as different modes of transmission and, frequently, also a non-particulate nature of cultural variants (Henrich & McElreath, 2003; Mesoudi, 2017).

Mathematical models of cultural evolution that use non-particulate inheritance go back to the pioneering work of Cavalli-Sforza & Feldman (1973, 1978, 1981). Nevertheless, models of cultural change based on the idea of competition between discontinuous particulate traits and their subsequent abundance in the population received more attention (Boyd & Richerson, 1987, 1988a), and they are, in fact, still dominant in current theoretical work on cultural evolution (Buenstorf & Cordes, 2008; Strimling *et al.*, 2009; Li, 2017). They are used despite concerns that this approach might result in a 'phenotypic gambit', where complex interactive phenomena (cooperation, imitation and learning) are reduced to strategies determined by single 'alleles' (such as 'cooperate' vs. 'betray', 'imitate' vs. 'invent'), which might jeopardize the relevance of such models for real-world applications (Centola & Baronchelli, 2015; Cownden *et al.*, 2017). Moreover, inheritance in these models is frequently uniparental, which casts further doubt on their applicability (Enquist *et al.*, 2010).

Although some models of cultural inheritance are based on non-particulate systems (Boyd & Richerson, 1988b; Acerbi *et al.*, 2009), even these tend to be substantially influenced by particulate theories. They tend to assume, for example, a flawless replication (Mesoudi & O'Brien, 2008), competition between discontinuous phenotype variants (Doebeli *et al.*, 2017), identical parent–offspring pairs without any stochastic variation (Li, 2017), constant mutational rate (Chiou & Wang, 2017) or constant innovative change (Mesoudi & O'Brien, 2008; Acerbi *et al.*, 2009).

NON-PARTICULATE INHERITANCE REVISITED

If we assume that in systems with non-particulate inheritance the trait levels in the individual offspring cannot exceed the parental trait range, after several generations we indeed end up with a homogeneous population. That can, however, be avoided by introducing some minor changes to the classical non-particulate inheritance system while making certain that Fleeming Jenkin's assumptions remain almost unchanged, so that: (1) the expected average trait value of offspring is equal to parental average trait value; and (2) trait values of the individual offspring are normally distributed around the parental average value.

One way of complying with these requirements is to use the Galton–Pearson model (Galton, 1889; Pearson, 1896). This model was formulated in modern terms by Cavalli-Sforza & Feldman (1981) and has been used in a quantitative approach to cultural evolution before. In an early work by Boyd & Richerson (1988b), half of their models rely on this kind of non-particulate inheritance. In this model, the trait value of an individual offspring, t, is determined by the arithmetic mean of parental trait values $t_{p1}, t_{p2}, \ldots, t_{pM}$ for M number parents, and an 'error' term, ε, so that $t = \mu\left(t_{p1}, t_{p2}, \ldots, t_{pM}\right) + \varepsilon; \quad \varepsilon \sim N\left(0, \eta^2\right)$,

where $N\left(\tau, \varsigma^2\right)$ denotes normal distribution with a mean, τ, and standard deviation, ς, and $\mu\left(t_{p1}, t_{p2}, \ldots, t_{pM}\right)$ is the average of parental values $t_{p1}, t_{p2}, \ldots, t_{pM}$, that is,

$$\mu\left(t_{p1}, t_{p2}, \ldots, t_{pM}\right) = \frac{1}{M}\sum_{j=1}^{M} t_{pj}.$$

The error term is to be viewed as a random phenotypic mutation and η as a mutation standard deviation. In this model, the probability distribution of offspring trait values is characterized by a constant standard deviation. Cavalli-Sforza & Feldman (1981) specified all non-particulate models as biparental (i.e. $M = 2$). We follow this numerical approach in most of our computer simulations (see Material and Methods) but define the formulas for any M so as to allow for a possible generalization to multi-parental inheritance.

Downloaded from https://academic.oup.com/biolinnean/article-abstract/127/2/518/5475672 by Department of Plant Physiology, Faculty of Science, Charles University user on 24 May 2019

To the best of our knowledge, there are no experimental data available that would support this form of mutation standard deviation described above in cultural inheritance. It is an inheritance model based on Galton's observations of genetically determined traits, such as sweet pea seed size and human height (Stanton, 2001), which were later shown to be determined directly by Mendelian inheritance of continuous quantitative traits regulated by many genes with additive effects (Fisher, 1918). Thus, although useful in population genetics, the power of the Galton–Pearson model might be limited when it comes to models of cultural evolution, because quantitative cultural traits are not stored in distinct, mutually exclusive particles. With a constant error term, two identical parents are assumed to have offspring as variable as very diverse parents. This makes the model rather artificial and inaccurate when it comes to approximating social learning, which is heuristic and error prone. The outcome of homogeneous inputs ought to differ from the outcome of heterogeneous inputs, but the Galton–Pearson model implicitly assumes that it is equally easy to estimate the intermediate value for any set of cultural influences, irrespective of their variability. The distribution of offspring trait values might be approximated more accurately by an error term that scales according to the distribution of parental trait values (see Fig. 1).

We propose an alternative, a Parental Variability-Dependent Inheritance model (PVDI), where the error term, ε, defined below is not constant but depends on the variability of parental values, σ, given by

$$\sigma^2\left(t_{p1},t_{p2},\ldots,t_{pM}\right)=\frac{1}{M}\sum_{j=1}^{M}\left[t_{pj}-\mu\left(t_{p1},t_{p2},\ldots,t_{pM}\right)\right]^2.$$

In this case, the probability distribution of offspring trait values is given by

$$t=\mu\left(t_{p1},t_{p2},\ldots,t_{pM}\right)+\ \varepsilon;$$
$$\varepsilon\ \sim N\left(0,\nu^2\sigma^2\left(t_{p1},t_{p2},\ldots,t_{pM}\right)\right),$$

where $\nu > 0$ is a constant representing relative offspring variability; we chose the expression $\nu^2\sigma^2(\ldots)$ mainly for simplicity. In this model, the greater the dissimilarity of parents with respect to a certain trait, the more variable in absolute values is the offspring they produce. In systems built around social learning (Hoppit & Laland, 2013), this is a reasonable assumption. When an individual learns from two different 'parents', it is reasonable to expect that individual to differ more (in absolute values) from the parental range. Unlike the Galton–Pearson model, the PVDI model predicts that the parental range is always exceeded by the same proportion of offspring. Both models are outlined briefly in flowcharts in Figure 1.

In the following, we use computer simulations to study the evolutionary dynamics in non-particulate PVDI systems. Each simulation starts with a homogeneous population far from the set optimum. This should help us to distinguish between directional and stabilizing selection (Phillips & Arnold, 1989). We observe that in PVDI models, the variability loss noted by Fleeming Jenkin and successful adaptation demonstrated by Cavalli-Sforza & Feldman (1981) are complemented by a third possible outcome, namely variability explosion. The probability of an outcome is influenced by the population size, N, proportion of surviving individuals, *s*, and relative offspring variability, ν. We also show that despite its simplicity, the PVDI model can generate punctuated evolutionary patterns. We complement these results with identical conclusions yielded by mathematical analysis of limiting cases in systems with Galton–Paerson inheritance and PVDI.

MATERIAL AND METHODS

We built a model of a population that is undergoing natural selection. To avoid population dynamics that could obscure the effect of the studied concept, the population size is held constant across generations in each simulation run. Population P_i is therefore represented by a vector of length N as

$$P_i=\left(t_1^i,t_2^i,\ldots,t_N^i\right)\ \in\ \mathbb{R}^N$$

where i denotes the number of the generation and real number t_j^i corresponds to the value of a hypothetical quantitative trait of j-th individual in i-th generation. To describe the process of generating the $(i + 1)$-th generation, we need the following set of parameters: (1) t_{opt} is the optimal trait value, i.e. any individual characterized by this trait value has the highest possible chance of survival; $0 < s \leq 1$ is the relative proportion of surviving individuals, low s therefore corresponds to strong selection; and $0 < \nu$ is the relative offspring variability.

Let us split the description of the model into two parts. First, we establish the model we used for the computer simulation; and second, we discuss possible ways of generalizing this model. The production of the $(i + 1)$-th generation from the i-th generation is composed of the following steps.

1. Handicaps of all individuals in the population are computed as square distances of their trait value from the optimum:

$$hand\left(t\right)=\left(t-t_{\text{opt}}\right)^2.$$

2. Truncation selection is used, i.e. only the fittest individuals survive and produce a new generation. The

Downloaded from https://academic.oup.com/biolinnean/article-abstract/127/2/518/5475672 by Department of Plant Physiology, Faculty of Science, Charles University user on 24 May 2019

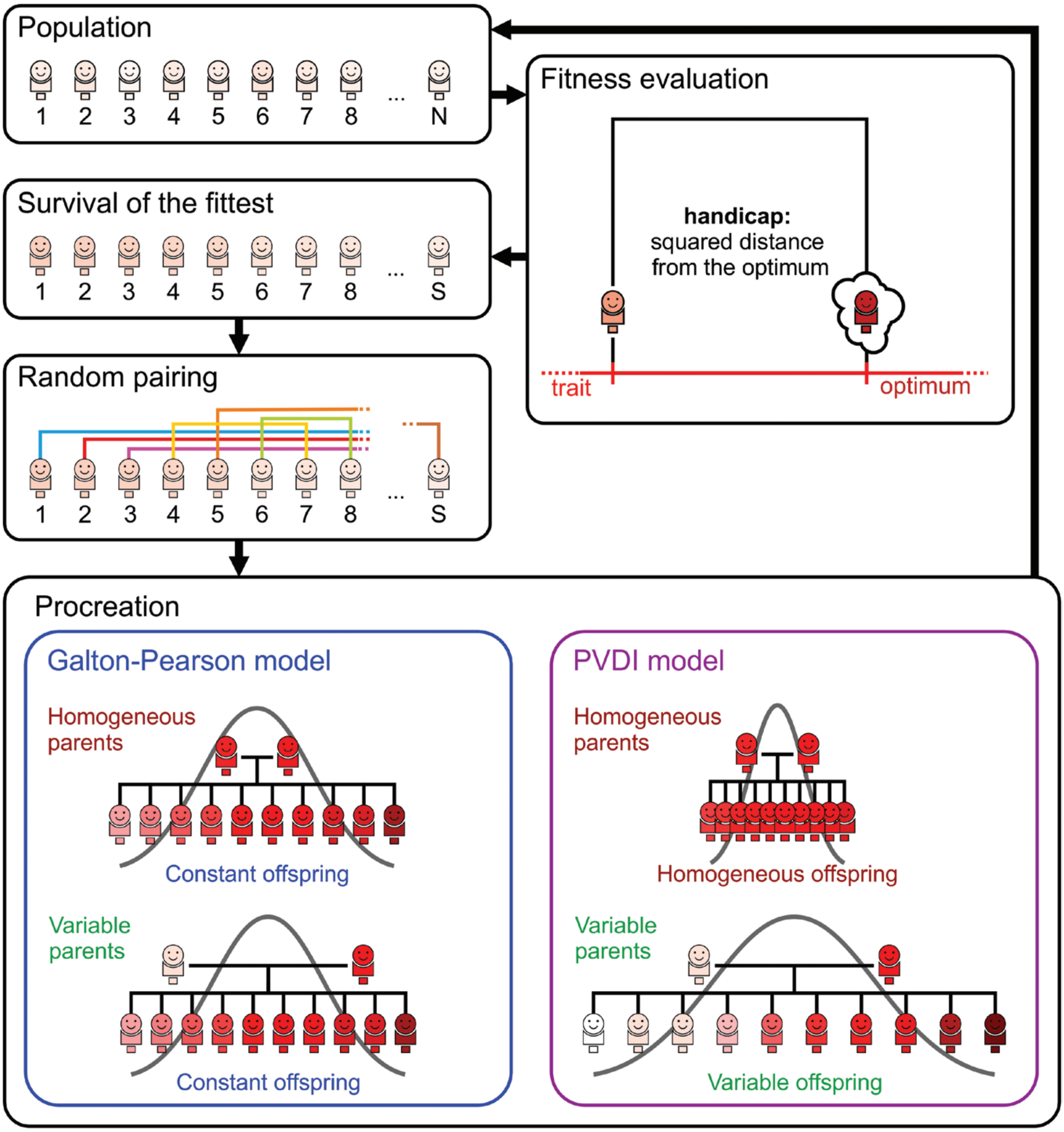

Figure 1. A flowchart of selection and reproduction in models of non-particulate inheritance. The colour shade of an individual represents a heritable quantitative trait. The Galton–Pearson and the Parental Variability-Dependent Inheritance model (PVDI) model differ with respect to the procreation step.

number of surviving individuals is $S = round(sN)$. For the sake of simplicity, let us assume that S is an even number and N is thus divisible by $S/2$. Then we find S individuals of P_i with the smallest handicap $hand\,(t_j)$ and pair them at random. This leaves us with $S/2$ random pairs of the S fittest individuals.

3. Each of the $S/2$ pairs generates an equal number of offspring, thus the size N of the original population is restored (each pair generates $2N/S$ individuals). We can now describe the generation of one of these offspring, the j-th member of a future $(i + 1)$-th generation. Let us denote the parents of the j-th individual of any generation by $p1\,(j)$ and $p2\,(j)$. The trait value t_j^{i+1} is then given by

$$t_j^{i+1} = \mu\left(t^i_{p1(j)}, t^i_{p2(j)}\right) + N\left(0, \nu^2\sigma^2\left(t^i_{p1(j)}, t^i_{p2(j)}\right)\right),$$

where μ and σ are the mean and standard deviation functions specified above. In this case, where an individual has exactly two parents, σ is equal to half of the parental span, i.e.

Downloaded from https://academic.oup.com/biolinnean/article-abstract/127/2/518/5475672 by Department of Plant Physiology, Faculty of Science, Charles University user on 24 May 2019

$$\sigma\left(t_{p1(j)},t_{p2(j)}\right)=\frac{1}{2}\left|t_{p1(j)}-t_{p2(j)}\right|.$$

4. The offspring generated in the previous step then forms a new $(i+1)$-th generation, again represented by a vector

$$P_{i+1}=\left(t_1^{i+1},t_2^{i+1},\ldots,t_N^{i+1}\right)\in\mathbb{R}^N.$$

Note that the result of this procedure is independent of the order in which individuals were originally sorted in vector P_i. A flowchart of this iterative algorithm can be found in Figure 1 (PVDI part). Editable script for an individual simulation run is available at https://github.com/costlysignalling/PVDI_python.

Where N is not divisible by $S/2$, we need to specify an additional process of offspring generation in order to keep the population size constant. In our numerical computations, we generate the remaining offspring using the parent pairing obtained in the previous part of offspring generation.

The initial conditions in our numerical simulations were chosen to resemble Jenkin's original example with a uniform population and an individual with exceptional and advantageous mutation (the 'sport'). The simulation starts with a uniform population of size N, where $N-1$ individuals have a trait value 10, one individual (the 'sport') has trait value 11, and the optimum is set at $t_{\text{opt}}=120$. These values, although arbitrarily chosen, do not jeopardize the generality of our findings, because we are free to alter the scale of measurement of trait values (Boyd & Richerson, 2005). The range of constants N, s and v was chosen based on a preliminary analysis in order to demonstrate important transitions between systems with different frequencies of successful and unsuccessful adaptations. We worked with $400\le N\le 25\,600$, $0.1\le s\le 0.9$ and $0.6\le\nu\le 1.4$.

By successful adaptation at the i-th generation, we mean a situation where variability measured by the coefficient of variation (CV) is smaller than a threshold, ε_1, and the mean trait value of the population is at most ε_2 from the optimum, i.e.

$$\sigma(P_i)/\mu(P_i)<\varepsilon_1\quad\text{and}\quad|\mu(P_i)-t_{\text{opt}}|<\varepsilon_2.$$

By unsuccessful adaptation, we mean a situation where variability measured by the CV is smaller than the threshold, ε_1, and the mean trait value of the population is more than ε_2 from the optimum, i.e.

$$\sigma(P_i)/\mu(P_i)<\ \varepsilon_1\quad\text{and}\quad|\mu(P_i)-t_{\text{opt}}|>\ \varepsilon_2$$

In numerical simulations, the condition of vanished variability is realized by checking whether variability rounded to four decimal digits is zero. We chose $\varepsilon_2=0.1$.

We also marked the number of generations needed for the first individual to reach or rise above the optimum in each of the simulation runs. This measure is a better indicator of adaptation speed than the number of generations needed until stabilization around an optimum is reached, because variability loss is much slower in simulation runs where the ratio of surviving individuals, s, is large. This analysis was supplemented with an additional set of simulation runs, where $200\le N\le 51\,200$, and $0.001\le\nu\le 4.0$ varied beyond the original interval. Our aim here was to show all three possible outcomes (successful adaptation, variability loss and variability explosion) in a single three-dimensional parameter space.

In another part of the analysis, we investigated the punctuality of evolution in this non-particulate inheritance system. We counted how many transgenerational changes in the population mean are responsible for the majority (50 and 80%) of evolutionary transition from the populational mean in the first generation to the optimum. Using Pearson's product–moment correlation, we investigated how well the magnitude of transgenerational change between the i-th and the $(i+1)$-th generation predicts the magnitude of transgenerational change in the subsequent generation. The proportion of influential intergenerational changes is a measure of punctuality in general, and the correlation between the size of subsequent changes measures the clustering of influential intergenerational changes and periods of evolutionary stasis. In this part of the analysis, we used only those simulation runs that resulted in a successful adaptation. We omitted the end of each simulation run, i.e. the stage after optimum was achieved by the populational mean. We also investigated the correlation between populational variability, measured by the CV, and a subsequent change in population mean in the same subsample of intergenerational changes.

For the purpose of further analysis, we differentiated between directional and stabilizing selection. The part of the evolutionary process where the distribution of individuals does not reach above the optimum was viewed as a period of directional selection:

$$P_i=\left(t_1^i,t_2^i,\ldots,t_N^i\right);\text{for}\quad\text{all } j\in[1,N]:\ t_j^{i-1}<t_{\text{opt}}\quad\text{and}\quad t_j^i<t_{\text{opt}}.$$

We can be certain during directional evolution that the distribution is truncated only at one tail. If the distribution reaches above the optimum, we speak of a stabilizing selection:

$$P_i=\left(t_1^i,t_2^i,\ldots,t_N^i\right);\text{for}\quad\text{all } j\in[1,N]:\ t_j^{i-1}<t_{\text{opt}}\quad\text{and}\quad\text{for}\quad\text{some } k\in[1,N]:\ t_k^i\ge t_{\text{opt}}.$$

Downloaded from https://academic.oup.com/biolinnean/article-abstract/127/2/518/5475672 by Department of Plant Physiology, Faculty of Science, Charles University user on 24 May 2019

A total of 5000 independent runs were executed for every possible combination of the abovementioned parameters, N, s and ν. All agent-based models were created in the Python programming language (van Rossum, 1995). Python v.2.7 was used with the packages NumPy (van der Walt *et al.*, 2011) and SciPy (Jones *et al.*, 2001). Visualizations were created mostly with the help of Matplotlib (Hunter, 2007), and some parts of the analysis and data visualization were also conducted in R 3.3.1.

Analytical results were obtained for the limiting case where $M = S$. This left out the random pairing and $N \to \infty$, which by the law of large numbers allowed us to work with normal distribution functions instead of vectors of trait values. The initial distribution of the population was assumed to be normal.

RESULTS

Results of the computer simulations (for a summary, see Fig. 2) show that runs with a large population size, reasonably high offspring variability and a medium ratio of surviving individuals almost invariably resulted in successful adaptation (Fig. 3A). In smaller populations, where relative offspring variability was small and the ratio of surviving individuals either large or small, the population frequently stopped developing

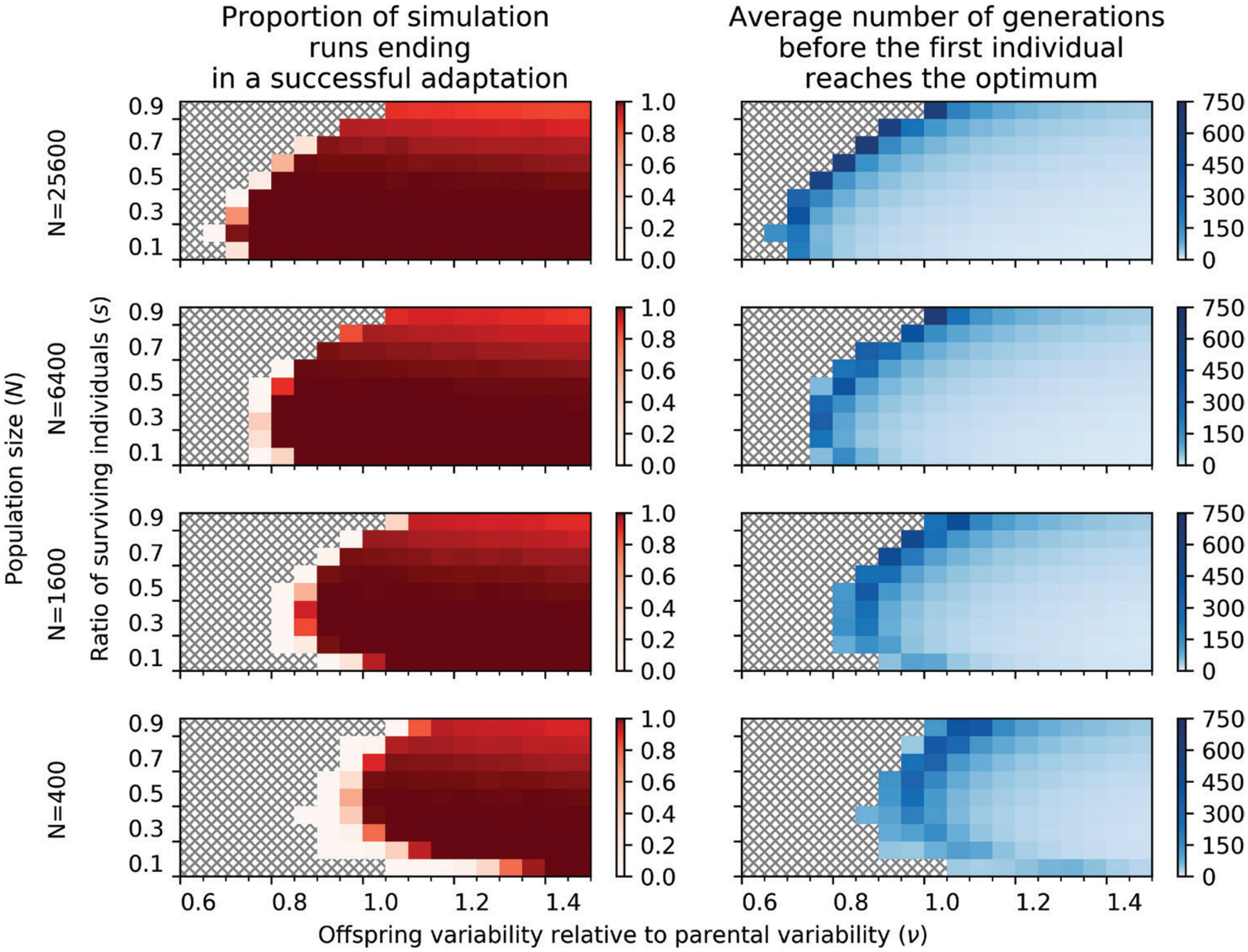

Figure 2. The probability of a successful adaptation (left) is determined by interactions between posterity variability, ν, the ratio of surviving individuals, s, and the population size, N (a logarithmic scale is used for more intelligible graphical representation of the population size effect). The mean number of generations needed until the first optimum is reached (right) has a slightly different relationship to the abovementioned variables. For each combination of parameters, 5000 simulations were run. Hatching distinguishes the area where zero simulation runs ended in a successful adaptation from the pale area where a small (e.g. 0.01%) but non-zero proportion was recorded. Logarithmic colouring is used. An equivalent summary for the Galton–Pearson model can be found in the Supporting Information (Fig. S1.1).

Downloaded from https://academic.oup.com/biolinnean/article-abstract/127/2/518/5475672 by Department of Plant Physiology, Faculty of Science, Charles University user on 24 May 2019

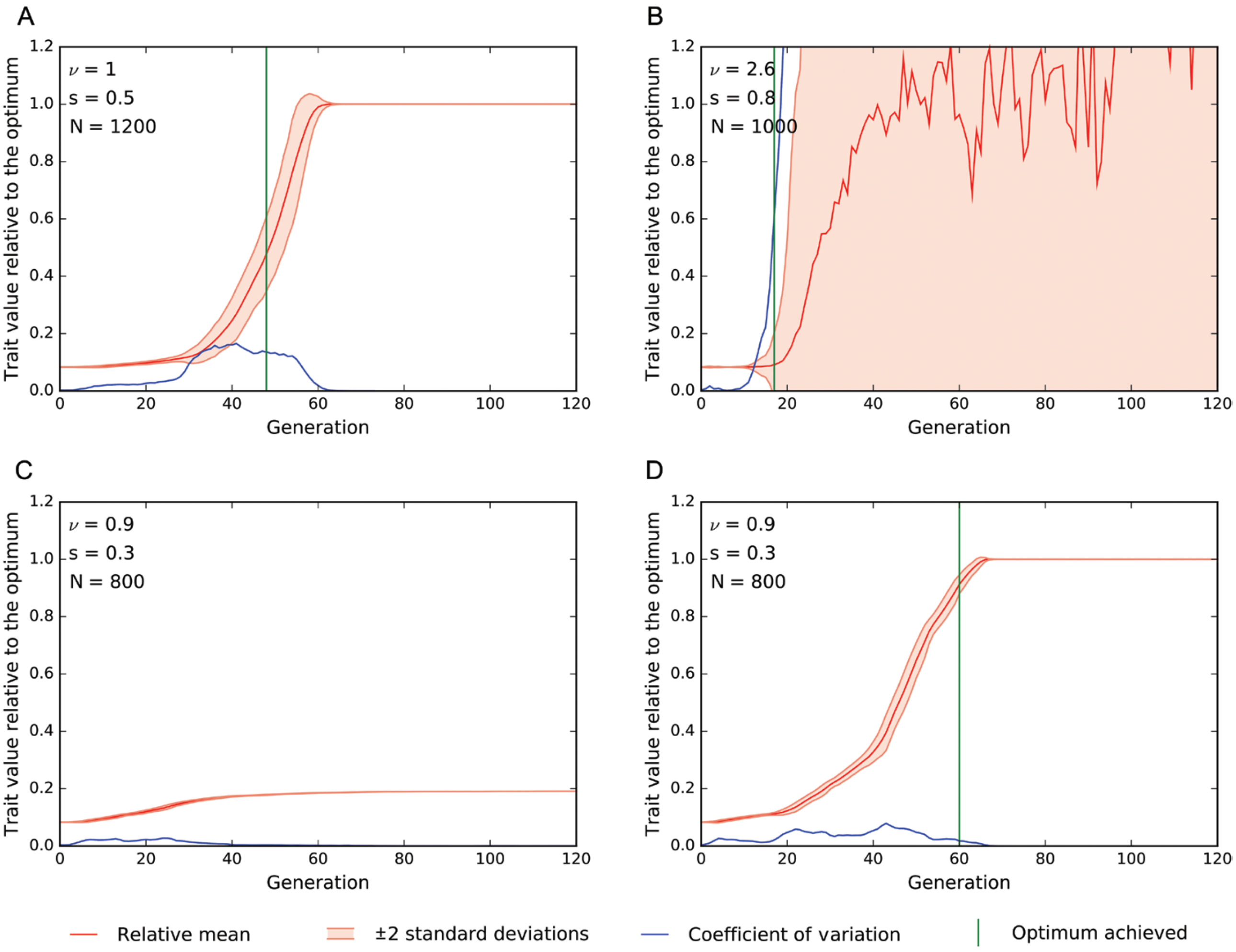

Figure 3. Exemplars of individual simulation runs. Every population is homogeneous at the beginning, and one sport is introduced in generation 1. Model parameters (stated in top left corner of each panel) in panels C and D are equal. Differences between exemplars in panels C and D are attributable to the stochastic nature of the model. The coefficient of variation (CV) is a relative variability measure: CV = σ/μ. Equivalent simulations for the Galton–Pearson model can be found in the Supporting Information (Fig. S1.2).

at a suboptimal level, because initial variability had vanished and selection with no variability could not transform the populational mean (Fig. 3C). In such cases, offspring variability was too small to resist the loss of variability attributable to averaging and selection.

In other configurations of parameters where adaptation was not successful, offspring variability was too large to be manageable either by averaging or by selection. The total population variability therefore increased throughout the whole simulation run, and the absolute value of fluctuation of the populational mean over time grew (Fig. 3B).

At the edge of the area where successful adaptation was universal, both successful and unsuccessful adaptations are possible, and stochastic mechanisms (random pairing and offspring generation) decide whether a simulation run ends with a successful adaptation or with variability loss (Fig. 3C, D).

Probabilities of the three possible outcomes specified above divided the three-diensional space of model parameters into three distinct areas: a zone of adaptation (Fig. 3A); a zone of variability explosion (Fig. 3B); and a zone of variability loss (Fig. 3C), which was the most likely outcome. These areas are outlined in a parameter space in Figure 4. Variability loss can also arise in the zone of variability explosion, because in the simulation runs described above there is at the outset only one exceptional individual. In principle, all her descendants could score lower than the homogeneous offspring of average population members. The likelihood of this happening at least

Downloaded from https://academic.oup.com/biolinnean/article-abstract/127/2/518/5475672 by Department of Plant Physiology, Faculty of Science, Charles University user on 24 May 2019

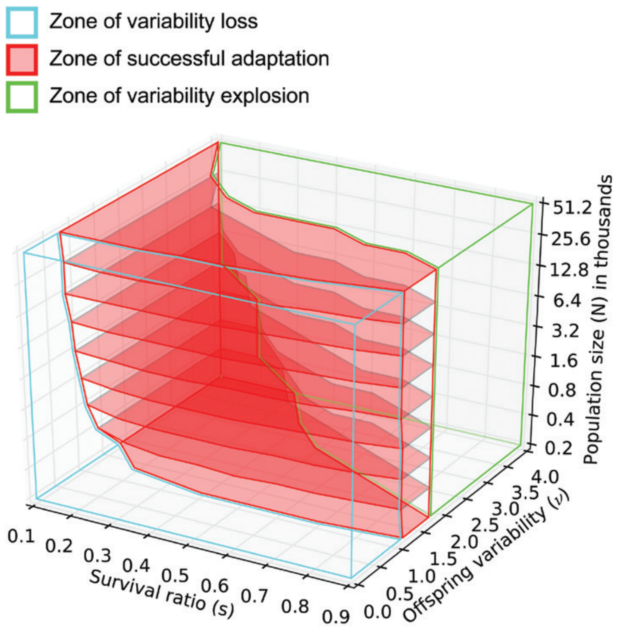

Figure 4. Zones of variability loss, adaptation and variability explosion in a parameter space outlined by ν, *s* and *N* (*N* is shown on a logarithmic scale). The colours outline areas where the outcomes in question are the most likely. Gradual overlaps (shown in Fig. 2) are neglected for the sake of clarity.

once in the first few generations after the introduction of this ‘sport’ increases in relationship to the growth of offspring variability (ν) and the ratio of surviving individuals (s), which also lowers the number of offspring an individual can have. When we substituted the initial conditions with a homogeneous population containing a sport for a population of variable individuals (such that $t_j^1 \sim N\left(\tau, \varsigma^2\right)$; $\varsigma \neq 0$), then in the zone of variability explosion a chaotic outcome became almost universal. It should be noted that sudden variability loss owing to a universal production of inferior offspring from promising individuals can also take place in the zone of adaptation (especially with growing ν).

The longest time until the first individual reaches the optimum is not associated with the combination of parameters that has the lowest probability of successful adaptation. The peak of this measure copies the border of an area where the optimum can be achieved but remains inside this border if the population size is small and evolution is therefore less predictable (Fig. 2). Where successful adaptation is unlikely, only individuals in runs where variability increases quickly tend to reach the optimum. If the parameter configuration is more favourable to a successful adaptation, even individuals in ‘less lucky’ runs, where a considerable amount of time is required to reach the populational optimum owing to stochastic processes acting against adaptation (averaging, weak selection), eventually reach the optimum.

Multi-parental inheritance, $M > 2$, conforms to all the abovementioned conclusions. Only with higher M, simulations become more deterministic and the zone of adaptation is more restricted by growing regions of adverse outcomes associated with variability loss or variability explosion. Adaptation either ends abruptly as a result of swamping of higher variability, attributable to averaging in larger parental sets, or the positive feedback of growing variability in already variable parental sets does not allow the populational mean to stabilize around the optimum. Graphical summaries for M equal to four, eight and 40, equivalent to Figures 2–4, can be found in Supporting Information, S2.

Analysis of the individual runs has revealed that in our model, evolution tends to follow a punctuated pattern, in which adaptation consists of periods of relative stasis and rapid change. On average, 27.9% (SD = 9.3%) of transgenerational changes in the populational mean were responsible for 80% of the total transition from the initial populational mean to the optimum. And although the punctuated character of the evolution was most apparent in simulation runs with large population size, weak selection and variable offspring (for complete results, see Supporting Information, S3), it was present in the whole domain of studied parameter combinations. In the vast majority (97.5%) of all simulation runs ending in a successful adaptation, the proportion of transgenerational changes responsible for 80% of total transitions was < 50%.

Transgenerational change between the populational mean in generation i, $\mu(P_i)$, and the population mean in generation $i+1$, $\mu(P_{i+1})$, predicted the magnitude of transgenerational change between $\mu(P_{i+1})$ and $\mu(P_{i+2})$ rather well. The average correlation between subsequent changes in the populational mean was 0.92 (SD = 0.08; 95% of all simulation runs scored between 0.70 and 0.99). In other words, large and small intergenerational changes in the populational mean tended to cluster. The examples in Figure 3A, D illustrate possible dynamics. At the outset, variability is small (note that the initial population is almost homogeneous) and selection cannot significantly affect the populational mean, but as variability grows, selection can lead to a substantial shift within a few generations. Populational variability (measured by CV) is well correlated with the subsequent change in the populational mean (average correlation was 0.75 with SD = 0.11; 95% of runs were between 0.49 and 0.91). For a depiction of how these values change in a parameter space, see Supporting Information, S3.

Downloaded from https://academic.oup.com/biolinnean/article-abstract/127/2/518/5475672 by Department of Plant Physiology, Faculty of Science, Charles University user on 24 May 2019

An interesting difference was found between the phases of directional and stabilizing selection (described in the Material and Methods). During stabilizing selection, variability declined, which was expected, but during directional selection it grew. These tendencies were more pronounced when the relative offspring variability, ν, was larger. The overall relationship is visualized in Figure 5. An interesting pattern of points appears in the graphical depiction of directional selection in certain configurations of parameters (especially when selection is strong, e.g. with $s = 0.2$). It suggests the emergence of a systematic boost in variability that leads to adaptation within several generations. This discontinuity can serve as a hallmark of punctuated evolution in stochastic simulations of PVDI systems.

A mathematical analysis of limiting cases without any stochastic component leads to equivalent conclusions. In PVDI systems, the mean trait value and variance grow exponentially during a directional selection, or variance decreases until the populational mean halts at a certain suboptimal value. We can identify an exact analytical criterion, which is dependent on ν and *s* and distinguishes between these

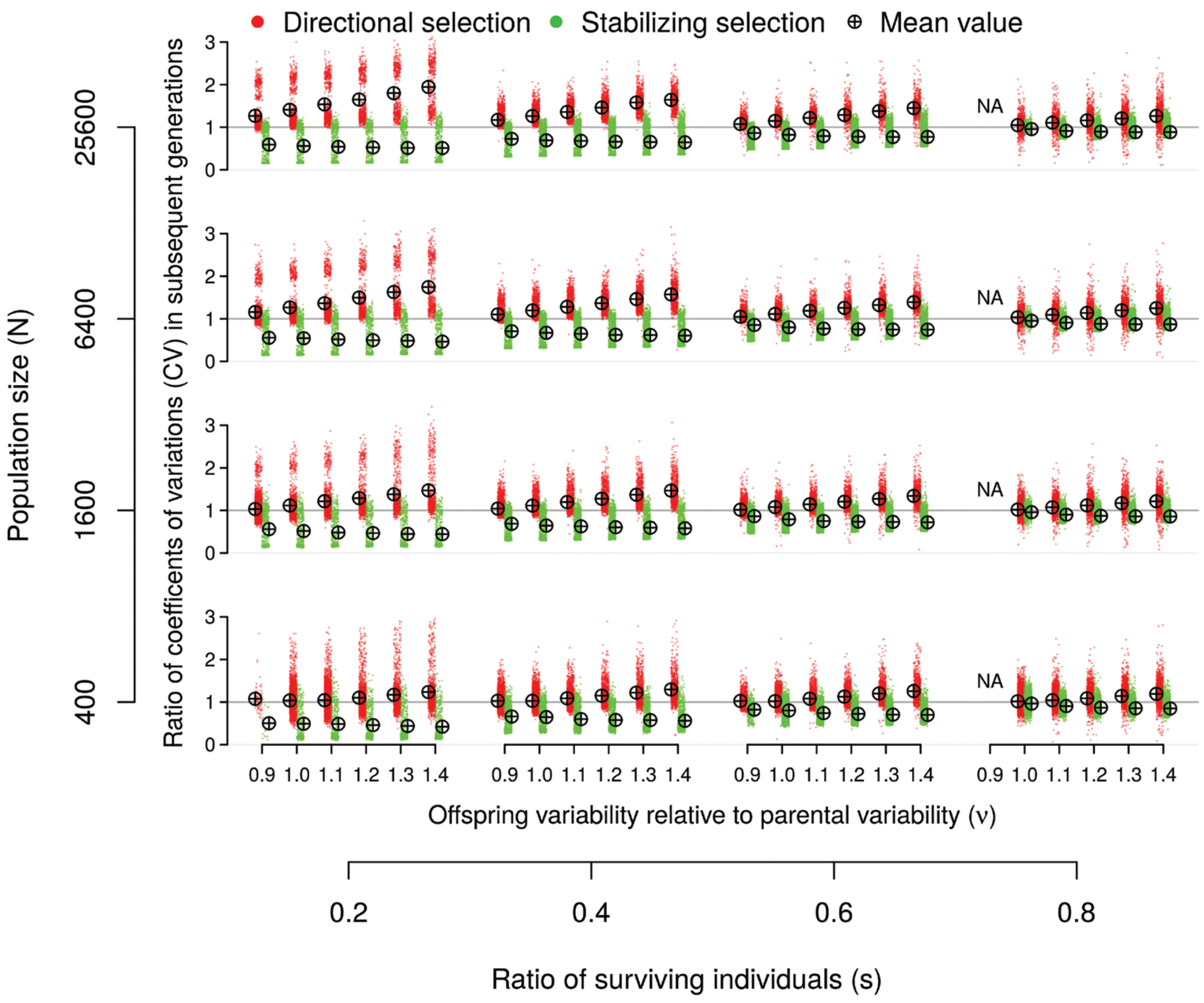

Figure 5. Stabilizing and directional selection have different impacts on the growth of variability. During directional selection the variability increases, whereas during stabilizing selection it declines. Moreover, directional selection seems to generate two rather distinct clusters of points at certain parameter configurations (e.g. $\nu = 1.1$, $s = 0.2$, $N = 6400$). This figure is based on 50 simulation runs ending in a successful adaptation for each parameter combination. Not available replace the subplot for parameter combinations where successful adaptation was not observed within all 5000 simulation runs. A dark grey line marks the ratio of CVs in subsequent generations equal to one. All points above this line indicate variability growth, whereas all points below indicate a decline of variability.

Downloaded from https://academic.oup.com/biolinnean/article-abstract/127/2/518/5475672 by Department of Plant Physiology, Faculty of Science, Charles University user on 24 May 2019

two cases. Selection in the Galton–Pearson model always results in a linear growth of mean trait value while variance remains constant. A complete analysis, with additional observations concerning the difference between an optimum in infinity and a real optimum, criterion of variability explosion, and the limiting case with absence of selection pressure, can be found in the Supporting Information, S4.

DISCUSSION

Previous studies of cultural evolution tended to rely on the Galton–Pearson model of inheritance, which assumes a constant offspring variability (Cavalli-Sforza & Feldman, 1981). This model, although originally meant to be non-particulate, was based on measurements influenced by the assumption of an underlying particulate structure of genes with additive effects (Fisher, 1918). For non-particulate cultural inheritance, as noted above, this assumption is rather unrealistic. We therefore proposed a model of non-particulate systems with PVDI. We managed to demonstrate that a minor modification in the mechanism of inheritance leads to a substantial boost in the complexity of evolutionary dynamics. Simulations based on this model show that once we introduce the stochasticity of random pairing and offspring generation, the same initial conditions can lead to either successful adaptation or loss of variability. These two previously contrasted outcomes are accompanied by a variability explosion, causing increasing fluctuations of the populational mean that dominates a substantial part of the investigated parameter space.

Earlier studies, with the aim of mathematical analysis of genetic variation and its phenotype realization across generations, modelled the variance of the offspring generation as a function of the variance of the parental generation (Cavalli-Sforza & Feldman, 1976, 1978; Lande, 1976; Feldman & Cavalli-Sforza, 1979). Inspired by Fisher (1918), these models postulate a large number of loci with additive effects, which after an independent selection of specific parental trait values results in a proportionality between the genotypic variance of progeny and one half of genotypic variance (plus a contribution of mutation to the within-family variance) in the parental population. This model can be expressed as

$$t_j^{i+1} = \mu\left(t_{p1(j)}^i, t_{p2(j)}^i\right) + N\left(m_g, \frac{\sigma^2\left(P_i^*\right)}{2} + M_g\right),$$

where m_g and M_g are constants that represent the influence of random mutation on changes in the mean trait value and variance of progeny respectively, and P_i^* is a population of parents after selection.

It has been shown that this model is an equivalent of the Galton–Pearson model for $m_g = 0$, whereas the constant mutation standard deviation, η, depends on other constant terms that drive selection and mutation (Fisher, 1918; Cavalli-Sforza & Feldman, 1976). Mathematical analyses of correlation between relatives use this model in order to be able to assume constant population variance over time (Lande, 1976; Feldman & Cavalli-Sforza, 1979). In the PVDI model, in contrast, the variance of sibling trait values depends on specific parental trait values. This makes the proposed model distinct from all other models of continuous variation, both genetic and non-genetic. Under the assumptions adopted in the PVDI, constant population variance, which is the key assumption of all of the abovementioned models, is unlikely.

Predictions of the PVDI model for systems with non-particulate inheritance can be summarized in the following three main claims.

1. Adaptive change is more probable and faster in larger populations with an intermediate ratio of surviving individuals and a higher, but not excessively high, relative offspring variability (Fig. 2).
2. The population mean can stabilize at a suboptimal level (Fig. 3C) or fluctuate chaotically (Fig. 3B) if relative offspring variability is either too small or too large with respect to other model parameters.
3. Population variability declines during stabilizing selection and tends to increase during directional selection (Figs 3A, D, 5).

The fact that successful adaptation is more frequent in large populations is intriguing and contradicts some findings from biological systems. It has been suggested that in large non-fragmented populations, substantial adaptive change is almost impossible owing to complex epistatic interactions between genes, which are all maintained in the population by frequency-dependent selection (Flegr, 2010). For an adaptive change to take hold, the population ought to be relatively small, but large enough to provide the requisite variability (Wright, 1930). It has also been proposed that adaptation can be achieved easily in large interbreeding populations, because substantial changes to allele frequencies have been observed in large natural populations on a year-to-year basis. In contrast, structured populations with semi-isolated units are still thought to be the most favourable setting for fast evolution (Fisher & Ford, 1950).

In our model, however, we explain cultural adaptation in large populations by hypothetically positing a positive feedback; a population can be large because cultural traits are well adapted, and cultural traits can be well adapted because the population is large (Powell *et al.*, 2009; Kempe & Mesoudi, 2014).

Downloaded from https://academic.oup.com/biolinnean/article-abstract/127/2/518/5475672 by Department of Plant Physiology, Faculty of Science, Charles University user on 24 May 2019

This might point to one of the fundamental differences between systems with particulate and non-particulate inheritance, because in particulate systems this positive feedback is absent.

Our findings are in line with assumed differences between changes in quickly and efficiently evolving cultural environments (such as contemporary complex societies), where many individuals participate in cultural transmission and innovation, and isolated traditional societies, which fit our description of simulated populations where variability loss is the most frequent outcome (small population size, small offspring variability). The case of the native Tasmanian population where massive loss of technological variability went hand in hand with absence of a constructive development also fits our findings (Henrich, 2004). However, the cultural decline of small and isolated societies can also be explainedwell by the notion of cultural drift (Richerson & Boyd, 2005), which relies on a particulate view of culture (Atkinson, 2011). The importance of larger group size for the construction of high-performance cultural artefacts is further supported by experimental evidence in laboratory settings (Derex *et al.*, 2013).

Predictions yielded by the PVDI model concur with a number of findings on cultural evolution that had no plausible explanation previously. Among other things, our simulations show that in cases of directional selection, variability begets variability. For instance, in the history of various technical inventions, it has been shown repeatedly that development starts with booming variability, which is followed by unification after the optimal design is reached (Klepper & Simons, 2000; Lake & Venti, 2009). Similar patterns were also identified in experimental settings that aimed to emulate the development of complex technologies (Derex & Boyd, 2016), social transmission of hunting skills (Acerbi *et al.*, 2016) and cooperation in groups (van den Berg *et al.*, 2015). The S-shaped curve (sigmoid curve with shallow slopes at the beginning and end, and a steep slope in the middle) illustrating the process of technological innovation (Foster, 1986; Lake & Venti, 2009) resembles the pattern symptomatic for adaptation in a PVDI system.

The punctuated character of evolution in PVDI systems is one of the most interesting outcomes of our study. This pattern has been identified previously in biological systems and described as punctuated anagenesis (MacLeod, 1991) or punctuated gradualism (Malmgren *et al.*, 1983). In such a scenario, evolution in a single non-fragmented population consists of periods of relative stasis and periods of rapid change, which are sometimes referred to as evolutionary 'pulses' or 'jerks' (Landis & Schraiber, 2017). Punctuated anagenesis should not be confused with the theory of punctuated equilibria, which requires lineage branching to explain apparent discontinuities in fossil records (Eldredge & Gould, 1972) and a variable rate of divergence (Levinthal, 1998; D'Huy, 2013).

Numerous studies report on the punctuated character of cultural evolution (Bar-Yosef, 2002; Grinin & Korotayev, 2009; Clark, 2011; d'Errico & Stringer, 2011; Kuhn, 2013; Lundgren *et al.*, 2018). The results of our simulations are also well compatible with prehistoric patterns of long periods of stasis, with patchy evidence of cognitive potential for a complex culture followed by major technological transition accompanied by booming cultural variability (Brumm & Moore, 2005; Kuhn, 2012). Such patterns tended to be explained by recourse to a positive feedback between population size and technological innovation (Baker, 2007; Grinin & Korotayev, 2009; Dow & Reed, 2011), but in the PVDI model they arise even if the population size is held constant.

In the PVDI system, stochasticity plays a more significant role then in systems with other previously suggested inheritance algorithms. The ongoing debate over the causes of and differences in the timing of 'symbolic revolutions' around the world might be beside the point (Bar-Yosef, 2002; Brumm & Moore, 2005; Powell *et al.*, 2009; Vaesen *et al.*, 2016), because if cultural transmission behaves in accordance with a PVDI model, major cultural transitions might happen at a given moment 'just because'. Although population size or density may influence the likelihood of occurrence of such an event, this model leaves a lot of space to contingency in the initiation of major cultural transitions. It may be impossible to identify any critical threshold in demographic parameters. It has also been hypothesized that punctuated anagenesis in cultural evolution is the result of abrupt 'macroinventions' (Mokyr, 1990) or the emergence of 'institutional friction' in large organizations (Lundgren *et al.*, 2018). Our simulated populations express this pattern even without introducing any of the abovementioned mechanisms or complicated spatiotemporal dynamics of punctuated equilibria.

We want to argue that the PVDI model is as simple and parsimonious as the Galton–Pearson model, which contradicts the widespread notion that models with heterogeneous evolutionary dynamics require more parameters (Hunt, 2008). Our results are compatible with simulation studies that used particulate models of cultural revolution to investigate its punctuated character while focusing on the combinatorial nature of innovation (Kolodny *et al.*, 2015). Both concepts might be deeply related. The number of possible new combinations increases with the number of existing 'parental' technologies. Even in a combinatorial model, the variability of potential outputs *s.l.* depends on the variability of inputs.

We tried to keep our model as simple as possible and make sure that is directly comparable to alternative

Downloaded from https://academic.oup.com/biolinnean/article-abstract/127/2/518/5475672 by Department of Plant Physiology, Faculty of Science, Charles University user on 24 May 2019

systems that work with non-particulate inheritance, such as the paint-pot theory of inheritance or Galton–Pearson model. This is why we kept survival deterministic and focused on biparental offspring production. Our simulations show that similar findings can be obtained when multi-parental inheritance is adopted (for details, see Supporting Information, S2). Moreover, it should be noted that both the Galton–Pearson model and the PVDI model can be viewed as special cases of general non-particulate inheritance models. Firstly, it would be easy to specify functions that would collapse the formulas describing these two models into a single, more general model, of which the Galton–Pearson and PVDI models are two special cases. Secondly, the parameters t_{opt}, s and ν can depend on time, in particular on, i, to capture possible time-dependent changes in the process. Thirdly, the requirement that N and M are constant could likewise be removed, in which case one would have to incorporate some functions that drive population dynamics. Fourthly, models that aspire to approximating the complexity of cultural transmission of a quantitative trait ought to contain certain randomness of survival (see Supporting Information, S5), inheritance with possible inequality of parental importance, and asynchronous generation of offspring.

There are other models of cultural evolution that abandoned the assumption of a normal distribution of offspring (Henrich, 2004; Powell *et al.*, 2009). Nonetheless, they still use distributions characterized by constant terms. The goal of this line of research was to link changes in cultural complexity with changes in the populational mean (Henrich, 2004). This step was, however, recently identified as problematic because cultural adaptation does not require changes in cultural variability (Vaesen *et al.*, 2016). In PVDI models, during directional selection populational variance tends to grow together with the populational mean. Hendrich's Gumbelian error term, characterized by constant α (bias away from optimum) and β (variability of offspring), might be generalized to a multi-parental inheritance, with β, and perhaps also α, depending on parental variance. This step would satisfy the abovementioned objections concerning constant populational variance.

Our simulations are based on a concept of cultural adaptation that includes a functional optimum value. Several recent studies on cultural evolution have used a related concept of adaptive landscape (i.e. local functional optima; Acerbi *et al.*, 2016; Derex & Boyd, 2016) or even continuous directional evolution (Henrich, 2004). The concept of optimal solution appears to be meaningful in the case of technologies, and particular technological inventions frequently evolve by optimizing their function (Foster, 1986; Petroski, 1994; Van Nierop *et al.*, 1997; Klepper & Simons, 2000). In non-technological domains of culture, such as symbols, languages and myths, however, the usefulness of the concept of cultural optimum is still disputed (Scott-Phillips & Kirby, 2010; Tehrani, 2011, 2013; D'Huy, 2013).

It has been suggested (Kleisner & Tureček, 2017) that biological and cultural evolution should be viewed as two aspects of one process. In particular, when a traditional genocentric model of biological evolution is enriched by other evolutionary processes, such as epigenetic inheritance, cognitively driven sexual selection or circular causations, non-particulate inheritance might provide more reliable predictions than models relying on between-allele competition (Muller & Wagner, 1991; Vaneechoutte, 1997; Lipkind & Tchernichovski, 2011; Jablonka *et al.*, 2014).

Future empirical studies on cultural macroevolution and transmission mechanisms will doubtless provide empirical support for either the traditional Galton–Pearson model or the alternative PVDI model. The PVDI model would be supported by findings of abrupt cultural changes preceded and accompanied by periods of increasing cultural variability. The Galton–Pearson model, on the contrary, would be supported by findings of gradual cultural changes and constant cultural variability. Likewise, if participants presented with a diverse set of cultural inputs (e.g. in a transmission chain) tend to deviate more from the mean input, then the PVDI model is supported, whereas a constant rate of innovation independent of input variability would provide support for the Galton–Pearson model of inheritance. We have conclusively demonstrated that if we want to describe and model cultural evolution faithfully, variation within culturally transmitted traits might be at least as important as their means. Computer models of cultural inheritance should consider the possibility of PVDI, because a failure to do so would lead to a neglect of a whole class of evolutionary dynamics stemming from the interaction between cultural variability, population size and innovation.

ACKNOWLEDGEMENTS

We would like to thank Anna Pilátová, PhD for her valuable comments and proofreading of the manuscript, Jan Toman for his comments on punctuated cultural evolution, and Oleg Sobchuk for his help with the issue of cultural adaptation and technological optima. We would also like to thank two anonymous reviewers for their extremely helpful comments.

Funding

This project was supported by the Charles University Grant Agency, project GA UK no. 52216.

Downloaded from https://academic.oup.com/biolinnean/article-abstract/127/2/518/5475672 by Department of Plant Physiology, Faculty of Science, Charles University user on 24 May 2019

AUTHOR CONTRIBUTIONS

P.T. created the concept of PVDI. P.T. and J.H. developed conception and design of the study. J.S. and M.K. wrote the code of computer simulations. J.S., M.K. and P.T. performed data analysis and visualization. J.S. conducted the mathematical analysis of the limiting cases. All authors discussed the results and contributed to the final manuscript.

CONFLICTS OF INTERESTS

None declared.

REFERENCES

Acerbi A, **Enquist M**, **Ghirlanda S. 2009.** Cultural evolution and individual development of openness and conservatism. *Proceedings of the National Academy of Sciences of the United States of America* **106:** 18931–18935.

Acerbi A, **Tennie C**, **Mesoudi A. 2016.** Social learning solves the problem of narrow-peaked search landscapes: experimental evidence in humans. *Royal Society Open Science* **3:** 160215.

Atkinson QD. 2011. Phonemic diversity supports a serial founder effect model of language expansion from Africa. *Science* **332:** 346–349.

Baker MJ. 2007. A structural model of the neolithic revolution. *Journal of Economic Growth* **13:** 257–292.

Bar-Yosef O. 2002. The upper Paleolithic revolution. *Annual Review of Anthropology* **31:** 363–393.

Barbrook AC, **Howe CJ**, **Blake N**, **Robinson P**, **1998.** The phylogeny of *The Canterbury Tales*. *Nature* **394:** 839.

van den Berg P, **Molleman L**, **Junikka J**, **Mikael P**, **Weissing FJ. 2015.** Human cooperation in groups: variation begets variation. *Scientific Reports* **5:** 16144.

Boyd R, **Richerson PJ. 1987.** The evolution of ethnic markers. *Cultural Anthropology* **2:** 65–79.

Boyd R, **Richerson PJ. 1988a.** The evolution of reciprocity in sizeable groups. *Journal of Theoretical Biology* **132:** 337–356.

Boyd R, **Richerson PJ. 1988b.** *Culture and the evolutionary process*. Chicago: University of Chicago Press.

Boyd R, **Richerson PJ. 2005.** *The origin and evolution of cultures*. New York: Oxford University Press.

Brumm A, **Moore MW. 2005.** Symbolic revolutions and the Australian archaeological record. *Cambridge Archaeological Journal* **15:** 157–175.

Buenstorf G, **Cordes C. 2008.** Can sustainable consumption be learned? A model of cultural evolution. *Ecological Economics* **67:** 646–657.

Bulmer M. 2004. Did Jenkin's swamping argument invalidate Darwin's theory of natural selection? *The British Journal for the History of Science* **37:** 281–297.

Cavalli-Sforza LL, **Feldman MW. 1973.** Cultural versus biological inheritance: phenotypic transmission from parents to children (A theory of the effect of parental phenotypes on children's phenotypes). *American Journal of Human Genetics* **25:** 618–637.

Cavalli-Sforza LL, **Feldman MW. 1976.** Evolution of continuous variation: direct approach through joint distribution of genotypes and phenotypes. *Proceedings of the National Academy of Sciences of the United States of America* **73:** 1689–1692.

Cavalli-Sforza LL, **Feldman MW. 1978.** The evolution of continuous variation. III. Joint transmission of genotype, phenotype and environment. *Genetics* **90:** 391–425.

Cavalli-Sforza LL, **Feldman MW. 1981.** *Cultural transmission and evolution: a quantitative approach*. Princeton: Princeton University Press.

Centola D, **Baronchelli A. 2015.** The spontaneous emergence of conventions: an experimental study of cultural evolution. *Proceedings of the National Academy of Sciences of the United States of America* **112:** 1989–1994.

Chiou SC, **Wang YC. 2017.** The example application of genetic algorithm for the framework of cultural and creative brand design in Tamsui Historical Museum. *Soft Computing* **22:** 2527–2545.

Clark JL. 2011. The evolution of human culture during the later Pleistocene: using fauna to test models on the emergence and nature of 'modern' human behavior. *Journal of Anthropological Archaeology* **30:** 273–291.

Cownden D, **Eriksson K**, **Strimling P. 2017.** A popular misapplication of evolutionary modeling to the study of human cooperation. *Evolution and Human Behavior* **38:** 421–427.

d'Errico F, **Stringer CB. 2011.** Evolution, revolution or saltation scenario for the emergence of modern cultures? *Philosophical Transactions of the Royal Society B: Biological Sciences* **366:** 1060–1069.

D'Huy J. 2013. A cosmic hunt in the Berber sky: a phylogenetic reconstruction of a Palaeolithic mythology. *Les Cahiers de l'AARS* **16:** 93–106.

Darwin C. 1857. *Darwin correspondence project, 'Letter no. 2166'*. Available at: http://www.darwinproject.ac.uk/DCP-LETT-2166

Darwin C. 1866. *Darwin correspondence project, 'Letter no. 4989'*. Available at: http://www.darwinproject.ac.uk/DCP-LETT-4989

Darwin C. 1867. *Darwin correspondence project, 'Letter no. 5567'*. Available at: http://www.darwinproject.ac.uk/DCP-LETT-5567

Darwin C. 1868. *The variation of animals and plants under domestication, Vols 1 and 2*. London: John Murray.

Davis AS. 1871. The "North British Review" and the Origin of Species. *Nature* **5:** 161.

Derex M, **Beugin MP**, **Godelle B**, **Raymond M. 2013.** Experimental evidence for the influence of group size on cultural complexity. *Nature* **503:** 389–391.

Derex M, **Boyd R. 2016.** Partial connectivity increases cultural accumulation within groups. *Proceedings of the National Academy of Sciences of the United States of America* **113:** 2982–2987.

Doebeli M, **Ispolatov Y**, **Simon B. 2017.** Towards a mechanistic foundation of evolutionary theory. *eLife* **6:** e23804.

Downloaded from https://academic.oup.com/biolinnean/article-abstract/127/2/518/5475672 by Department of Plant Physiology, Faculty of Science, Charles University user on 24 May 2019

Dow GK, **Reed CG. 2011.** Stagnation and innovation before agriculture. *Journal of Economic Behavior and Organization* **77:** 339–350.

Eldredge N, **Gould SJ. 1972.** Punctuated equilibria: an alternative to phyletic gradualism. In: Schopf TJM, ed. *Models in Paleobiology*. San Francisco: Freeman, Cooper and Co., 82–115.

Enquist M, **Strimling P**, **Eriksson K**, **Laland K**, **Sjostranda J. 2010.** One cultural parent makes no culture. *Animal Behaviour* **79:** 1353–1362.

Feldman MW, **Cavalli-Sforza LL. 1979.** Aspects of variance and covariance analysis with cultural inheritance. *Theoretical Population Biology* **15:** 276–307.

Fisher RA. 1918. The correlation between relatives on the supposition of Mendelian inheritance. *Transactions of the Royal Society of Edinburgh* **52:** 399–433.

Fisher RA, **Ford B. 1950.** The "Sewall Wright Effect". *Heredity* **4:** 117–119.

Fleeming Jenkin HC. 1867. Review of the Origin of Species. *The North British Review* **46:** 277–318.

Flegr J. 2010. Elastic, not plastic species: frozen plasticity theory and the origin of adaptive evolution in sexually reproducing organisms. *Biology Direct* **5:** 2.

Foster RN. 1986. *Innovation: the attacker's advantage*. New York: Summit Books.

Galton F. 1889. *Natural inheritance*. London and New York: Macmillan.

Grinin L, **Korotayev A. 2009.** Social macroevolution: growth of the world system integrity and a system of phase transitions. *World Futures* **65:** 477–506.

Henrich J. 2004. Demography and cultural evolution: how adaptive cultural processes can produce maladaptive losses—the Tasmanian case. *American Antiquity* **69:** 197–214.

Henrich J, **Boyd R**, **Derex M**, **Kline MA**, **Mesoudi A**, **Muthukrishna M**, **Powell AT**, **Shennan SJ**, **Thomas MG. 2016.** Understanding cumulative cultural evolution. *Proceedings of the National Academy of Sciences of the United States of America* **113:** E6724–E6725.

Henrich J, **McElreath R. 2003.** The evolution of cultural evolution. *Evolutionary Anthropology* **12:** 123–135.

Hoppit W, **Laland KN. 2013.** *Social learning: an introduction to mechanisms, methods, and models*. Princeton and Oxford: Princeton University Press.

Hunt G. 2008. Gradual or pulsed evolution: when should punctuational explanations be preferred? *Paleobiology* **34:** 360–377.

Hunter JD. 2007. Matplotlib: a 2D graphics environment. *Computing in Science and Engineering* **9:** 99–104.

Jablonka E, **Lamb M**, **Zeligowski A. 2014.** *Evolution in four deminsions (revised edition): genetic, epigenetic, behavioral, and symbolic variation in the history of life*. Cambridge and London: MIT Press.

Jones E, **Oliphant E**, **Peterson P**. **2001.** *SciPy: open source scientific tools for Python*. Available at: http://www.scipy.org

Kempe M, **Mesoudi A. 2014.** An experimental demonstration of the effect of group size on cultural accumulation. *Evolution and Human Behavior* **35:** 285–290.

Kleisner K, **Tureček P. 2017.** Cultural and biological evolution: what is the difference? *Biosemiotics* **10:** 127–130.

Klepper S, **Simons KL. 2000.** The making of an oligopoly: firm survival and technological change in the evolution of the U.S. tire industry. *Journal of Political Economy* **108:** 728–760.

Kolodny O, **Creanza N**, **Feldman MW. 2015.** Evolution in leaps: the punctuated accumulation and loss of cultural innovations. *Proceedings of the National Academy of Sciences of the United States of America* **112:** E6762–E6769.

Kuhn SL. 2012. Emergent patterns of creativity and innovation in early technologies. In: Scott E, ed. *Origins of human innovation and creativity*. Amsterdam and Oxford: Elsevier, 69–87.

Kuhn SL. 2013. Cultural transmission, institutional continuity and the persistence of the Mousterian. In: Akazawa T, Nishiaki Y, Aoki K, eds. *Dynamics of learning in Neanderthals and modern humans*, *Vol. 1*. Tokyo: Springer Japan, 105–113.

Lake M, **Venti J. 2009.** Quantitative analysis of macroevolutionary patterning in technological evolution: bicycle design from 1800 to 2000. In: Shennan S, ed. *Pattern and process in cultural evolution*. Berkeley and Los Angeles: UCLA Press.

Lande R. 1976. Natural selection and random genetic drift in phenotypic evolution. *Evolution* **30:** 314–334.

Landis MJ, **Schraiber JG. 2017.** Pulsed evolution shaped modern vertebrate body sizes. *Proceedings of the National Academy of Sciences of the United States of America* **114:** 13224–13229.

Levinthal DA. 1998. The slow pace of rapid technological change: gradualism and punctuation in technological change. *Industrial and Corporate Change* **7:** 217–248.

Li Y. 2017. Evolution of cooperative imitators in social networks. *Physical Review E* **95:** 022303.

Lipkind D, **Tchernichovski O. 2011.** Quantification of developmental birdsong learning from the subsyllabic scale to cultural evolution. *Proceedings of the National Academy of Sciences of the United States of America* **108:** 15572–15579.

Lundgren M, **Squatrito T**, **Tallberg J. 2018.** Stability and change in international policy-making: a punctuated equilibrium approach. *Review of International Organizations* **13:** 547–572.

MacLeod N. 1991. Punctuated anagenesis and the importance of stratigraphy to paleobiology. *Paleobiology* **17:** 167–188.

Malmgren BA, **Berggren WA**, **Lohmann GP. 1983.** Evidence for punctuated gradualism in the late Neogene *Globorotalia tumida* lineage of planktonic foraminifera. *Paleobiology* **9:** 377–389.

McElreath R, **Boyd R**, **Richerson PJ. 2003.** Shared norms and the evolution of ethnic markers. *Current Anthropology* **44:** 122–130.

Mesoudi A. 2011. *Cultural evolution: how Darwinian thoery can explain human culture and synthesize the social sciences*. Chicago: University of Chicago Press.

Mesoudi A. 2017. Pursuing Darwin's curious parallel: prospects for a science of cultural evolution. *Proceedings*

Downloaded from https://academic.oup.com/biolinnean/article-abstract/127/2/518/5475672 by Department of Plant Physiology, Faculty of Science, Charles University user on 24 May 2019

of the National Academy of Sciences of the United States of America **114:** 7853–7860.

Mesoudi A, **O'Brien MJ. 2008.** The cultural transmission of great basin projectile-point technology II: an agent-based computer simulation. *American Antiquity* **73:** 627–644.

Mokyr J. 1990. Punctuated equilibria and technological progress. *The American Economic Review* **80:** 350–354.

Muller GB, **Wagner GP. 1991.** Novelty in evolution: restructuring the concept. *Annual Review of Ecology and Systematics* **22:** 229–256.

van Nierop OA, **Blankendaal ACM**, **Overbeeke CJ. 1997.** The evolution of the bicycle: a dynamic systems approach. *Journal of Design History* **10:** 253–267.

Nurse P. 2000. A long twentieth century of the cell cycle and beyond. *Cell* **100:** 71–78.

Pearson K. 1896. Mathematical contributions to the theory of evolution. III. Regression, heredity and panmixia. *Philosophical Transactions of the Royal Society of London A* **187:** 253–318.

Petroski H. 1994. *The evolution of useful things*. New York: Vintage Books.

Phillips PC, **Arnold SJ. 1989.** Visualizing Multivariate Selection. *Evolution* **43:** 1209–1222.

Powell A, **Shennan S**, **Thomas MG. 2009.** Late Pleistocene demography and the appearance of modern human behavior. *Science* **324:** 1298–1301.

Richerson PJ, **Boyd R. 2005.** *Not by genes alone*. Chicago: The University of Chicago Press.

van Rossum G. 1995. *Python tutorial*, *Technical Report CS-R9526*. Amsterdam: Centrum voor Wiskunde en Informatica (CWI).

Scott-Phillips TC, **Kirby S. 2010.** Language evolution in the laboratory. *Trends in Cognitive Sciences* **14:** 411–417.

Stanton JM. 2001. Galton, pearson, and the peas: a brief history of linear regression for statistics instructors. *Journal of Statistics Education* **9:** 1–13.

Strimling P, **Enquist M**, **Eriksson K. 2009.** Repeated learning makes cultural evolution unique. *Proceedings of the National Academy of Sciences of the United States of America* **106:** 13870–13874.

Tehrani JJ. 2011. Patterns of evolution in Iranian tribal textiles. *Evolution* **4:** 390–396.

Tehrani JJ. 2013. The phylogeny of little Red Riding Hood. *PLoS ONE* **8:** e78871.

Vaesen K, **Collard M**, **Cosgrove R**, **Roebroeks W. 2016.** Population size does not explain past changes in cultural complexity. *Proceedings of the National Academy of Sciences of the United States of America* **113:** E2241–E2247.

Vaneechoutte M. 1997. Bird song as a possible cultural mechanism for speciation. *Journal of Memetics - Evolutionary Models of Information Transmission* **1:** 1–10.

van der Walt S, **Colbert SC**, **Varoquaux G. 2011.** The NumPy array: a structure for efficient numerical computation. *Computing in Science and Engineering* **13:** 22–30.

Wright S. 1930. *Evolution in Mendelian populations*. Chicago: University of Chicago Press.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

S1. Simulation runs of a system with a Galton–Pearson model of inheritance equivalent to Parental Variability-Dependent Inheritance model (PVDI) runs shown in Figures 2 and 3.

Figure S1.1. Under Galton–Pearson inheritance, the probability of successful adaptation (left) is 100%. The mean number of generations needed to reach the first optimum (right) depends on a constant posterity variability, η, and the ratio of surviving individuals, s, but not on the population size, N. For each combination of parameters, 5000 simulations were run. The definition of adaptation had to be modified because in the Galton–Pearson inheritance framework, population variability never disappears completely. Sufficient condition of adaptation was defined as the achievement of the optimum by the populational mean and maintenance of a position sufficiently close to this optimum for the rest of the simulation run (2000 generations). Logarithmic colouring is used.

Figure S1.2. Examples of individual simulation runs in a system with a Galton–Pearson model of inheritance. All populations are homogeneous at the beginning, and one sport is introduced in generation 1. Model parameters (stated in the top left corner of each example) are equal in exemplars c and d. There is almost no difference between these two exemplars, because in the Galton–Pearson model, stochastic processes do not play a major role. The coefficient of variation (CV) is a measure of relative variability, $\mathrm{CV} = \sigma/\mu$.

S2. Parental variability-dependent inheritance model (PVDI): results of multi-parental inheritance.

Figure S2.1. The probability of a successful adaptation (left) and the mean number of generations required until the first optimum is reached (right) for $M = 4$.

Figure S2.2. The probability of a successful adaptation (left) and the mean number of generations required until the first optimum is reached (right) for $M = 8$.

Figure S2.3. The probability of a successful adaptation (left) and the mean number of generations required until the first optimum is reached (right) for $M = 40$.

Downloaded from https://academic.oup.com/biolinnean/article-abstract/127/2/518/5475672 by Department of Plant Physiology, Faculty of Science, Charles University user on 24 May 2019

Figures S2.1–2.3. For each combination of parameters, 1000 simulations were run. Hatching distinguishes the area where no simulation runs ended in a successful adaptation from the pale area, where a small (e.g. 0.01%) but non-zero proportion was recorded. Logarithmic colouring is used.
Figure S2.4. Examples of individual simulation runs for $M = 4$.
Figure S2.5. Examples of individual simulation runs for $M = 8$.
Figure S2.6. Examples of individual simulation runs for $M = 40$.
Figures S2.4–2.6. Simulation parameters were modified for each value of M to match the patterns demonstrated for biparental inheritance in the main text. Exemplar a represents a successful adaptation; exemplar b represents variability explosion; and exemplars c and d demonstrate a parameter setting where the outcome alters between variability loss and successful adaptation owing to the stochastic nature of the model. The coefficient of variation (CV) is a measure of relative variability, $\mathrm{CV} = \sigma/\mu$.
Figure S2.7. Zones of variability loss, adaptation and variability explosion in a parameter space outlined by ν, s and N (shown on a logarithmic scale) for $M = 4$, 8 and 40. The colours outline areas where the outcomes in question are the most likely ones. For each combination of parameters, 1000 simulations were run.
S3. Variation in average punctuality and variability dependence of evolution.
Figure S3.1. Population size = 400.
Figure S3.2. Population size = 1600.
Figure S3.4. Population size = 6400.
Figure S3.5. Population size = 25 600.
S4. Analytical comparison between a Parental Variability-Dependent Inheritance model (PVDI) and a Galton–Pearson model of inheritance.
S4.1. Truncated normal distribution.
S4.2. Optimum in infinity.
S4.3. Real optimum.
S4.4. Variability explosion after achieving optimum.
S4.5. The case of no selection.
S5. Random factor in survival.

SHARED DATA

This article has no additional data. Computer script that generates one simulation run can be accessed at https://github.com/costlysignalling/PVDI_python. Interested readers can use this script to generate the bulk of data statistically equivalent to those that was used in our analysis.

Downloaded from https://academic.oup.com/biolinnean/article-abstract/127/2/518/5475672 by Department of Plant Physiology, Faculty of Science, Charles University user on 24 May 2019

Downloaded from https://academic.oup.com/biolinnean/article-abstract/127/2/518/5475672 by Department of Plant Physiology, Faculty of Science, Charles University user on 24 May 2019