



CrossMark
click for updates

Research

Cite this article: Tamura K, Kobayashi Y, Ihara Y. 2015 Evolution of individual versus social learning on social networks. *J. R. Soc. Interface* **12**: 20141285.

<http://dx.doi.org/10.1098/rsif.2014.1285>

Received: 20 November 2014

Accepted: 2 January 2015

Subject Areas:

computational biology

Keywords:

cultural evolution, social structure, maladaptive culture, relatedness

Author for correspondence:

Kohei Tamura

e-mail: tmr.kohei@gmail.com

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rsif.2014.1285> or via <http://rsif.royalsocietypublishing.org>.

Evolution of individual versus social learning on social networks

Kohei Tamura^{1,2,3}, Yutaka Kobayashi⁴ and Yasuo Ihara¹

¹Department of Biological Sciences, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

²Department of Creative Informatics, Graduate School of Information Science and Technology, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan

³CREST, JST, 4-1-8, Honcho, Kawaguchi, Saitama 332-0012, Japan

⁴Department of Management, Kochi University of Technology, Tosayamada, Kami-city, Kochi 782-8502, Japan

A number of studies have investigated the roles played by individual and social learning in cultural phenomena and the relative advantages of the two learning strategies in variable environments. Because social learning involves the acquisition of behaviours from others, its utility depends on the availability of 'cultural models' exhibiting adaptive behaviours. This indicates that social networks play an essential role in the evolution of learning. However, possible effects of social structure on the evolution of learning have not been fully explored. Here, we develop a mathematical model to explore the evolutionary dynamics of learning strategies on social networks. We first derive the condition under which social learners (SLs) are selectively favoured over individual learners in a broad range of social network. We then obtain an analytical approximation of the long-term average frequency of SLs in homogeneous networks, from which we specify the condition, in terms of three relatedness measures, for social structure to facilitate the long-term evolution of social learning. Finally, we evaluate our approximation by Monte Carlo simulations in complete graphs, regular random graphs and scale-free networks. We formally show that whether social structure favours the evolution of social learning is determined by the relative magnitudes of two effects of social structure: localization in competition, by which competition between learning strategies is evaded, and localization in cultural transmission, which slows down the spread of adaptive traits. In addition, our estimates of the relatedness measures suggest that social structure disfavors the evolution of social learning when selection is weak.

1. Introduction

Learning is observed in a broad range of animal taxa and is of a pinnacle importance in shaping human behaviour. The unparalleled cultural diversity in modern humans (i.e. *Homo sapiens*), along with their global expansion into diverse environments, could not have been possible without their prominent learning capacity [1,2]. Learning is divided into two categories: individual learning and social learning (also known as exploration and exploitation, respectively) [2,3]. Individual (or asocial) learning, such as trial and error and insight, refers to learning that occurs independently of any social influences. Social learning, including imitation, involves the transmission of traits or behaviours through interaction with other individuals, in which a characteristic exhibited by a 'model' is acquired by an 'observer'. There is a wide theoretical agreement that individual and social learning are adaptations to temporally and/or spatially variable environments, as innate behaviour is sufficient to cope with a constant environment [4–7]. The relative advantages of individual and social learning depend on the environmental constancy: in short, unless the environment is extremely stable so that innate behaviour is most advantageous, social learners (SLs) are advantageous in a moderately variable environment because they can acquire adaptive behaviour without engaging in costly individual learning, but run a risk of carrying an inept behaviour in a highly variable environment.

Animal societies are structured in the sense that a given individual in a society is more likely to interact with certain individuals than with others in the same society. Social structure of this kind introduces a level of complexity to animal societies, which has been hypothesized to underlie the evolution of communicative complexity across species [8] as well as larger brains in primates [9]. Societies may be structured on the basis of genetic relatedness as well as other criteria. In humans, in particular, it has been emphasized that cultural variations within and between societies play an important role in creating social structure [10]. This adds a further complexity to human societies, which might be causally related to the evolutionary origins of the human cognitive capacity [11].

Social structure also governs the flow of information via social learning. Recent empirical studies have shown that cultural transmission in humans and other animals is affected by social network architecture [12–15]. Furthermore, the relative advantage of SLs depends on from whom they acquire information [16–20], which indicates that social structure should play an essential role in the evolution of learning.

However, possible effects of social structure on the evolution of learning have not been fully explored so far with some notable exceptions [21–24]. As summarized by a recent study, social or spatial structure may affect the evolution of learning in two ways [23]. First, information flow is localized in a structured population in the way that social learning occurs only between spatially or socially neighbouring individuals (i.e. localization in cultural transmission). Thus, it takes more time for any cultural trait to spread in a structured population than in a well-mixed population, which disfavours social learning by increasing the time they spend with outdated behaviours. Second, competition is also localized in a structured population, so that individuals tend to compete with those who are genetically similar to themselves (i.e. localization in competition). This benefits SLs with maladaptive behaviour by letting them circumvent competing with individual learners (ILs).

Then, a question arises: does social structure favour or dis-favour the evolution of social learning after all? Or, to put it in another way, which of the two effects of localization in a structured population plays the dominant role? The question remains open, as recent studies have produced mixed results [21,23]. In a landmark study, Rendell *et al.* [21] conducted a series of simulations to explore evolution of social and asocial learning in a two-dimensional, lattice-structured population [21]. They found that the long-term average frequency of SLs is typically higher when both dispersal and learning are limited within neighbouring sites than when both occur globally. This supports the notion that population structure favours the evolution of social learning. More recently, Kobayashi & Wakano [23] investigated the evolution of social and individual learning using an infinite island model, in which a population is divided into infinitely many, finite-sized islands connected by migration [23]. According to their analytical results, social learning is favoured when each island is large or the migration rate is high, suggesting that spatial subdivision disfavours the evolution of social learning, which is in stark contrast to Rendell *et al.* [21].

The discrepancy between the two studies implies that whether or not social structure favours the evolution of social learning depends on the precise nature of the structure. To understand the effect of structure, therefore, it is essential

Table 1. Pay-offs to three phenogenotypes.

phenogenotype	pay-off
IL	$1 - c$
SLC	1
SLW	$1 - s$

to formulate the evolution of learning in a social structure that is as general as possible. In this study, we model evolutionary dynamics of learning strategies in a general social network. In our analysis, we use the concept of relatedness, a statistical measure of association between interacting individuals applicable to both genetically and culturally inherited traits [25–28]. Relatedness has been applied to a variety of biological phenomena [25,29–31], particularly in spatially structured populations [32–34]. We first describe our evolutionary model and derive the condition for natural selection to favour the evolution of social learning. We then obtain an approximated expression of the long-term average frequency of SLs for homogeneous networks, where each individual has the same number of neighbours, and specify the condition under which social structure favours the evolution of social learning as a function of relatedness measures. We also examine the validity of our approximation by Monte Carlo simulations in three types of social structures: complete graphs, regular random graphs and scale-free networks.

2. Model

2.1. Description of the model

We investigate possible effects of social structure on the evolution of social and individual learning, following the general framework developed by earlier studies [4–6,23]. Consider a population consisting of N individuals in a fluctuating environment, where there is only one behaviour that is adaptive to the current state of the environment and all other behaviours are equally maladaptive. The environment changes every T generations, where T measures the environmental stability. We adopt the infinite environmental state model, which assumes that the environment never reverts when it changes [4–6,23].

Two learning strategies are considered: each individual is either an IL or an SL. An IL engages in costly trial-and-error learning and always discovers the correct (adaptive) behaviour. An SL imitates others' behaviour without paying the cost of individual learning, but may acquire a wrong (maladaptive) behaviour as a result. Hence, there are three phenogenotypes, IL, SL with the correct behaviour (SLC) and SL with a wrong behaviour (SLW). The relative advantages of the three phenogenotypes are measured in terms of pay-offs (table 1). The pay-off to IL is reduced to $1 - c$ relative to SLC, where c represents the cost owing to having a more complex nervous system or engaging in trial-and-error learning ($0 < c < 1$). The pay-off to SLW reduces to $1 - s$ relative to SLC, where s represents the cost owing to behaving maladaptively ($0 < s < 1$).

We regard each individual in a population as occupying a vertex of a network whose edges represent social links between individuals. We assume that network structure is

constant for the duration of the evolutionary changes considered. In what follows, individuals occupying linked vertices are referred to as neighbours.

We consider changes in the frequencies of IL, SLC and SLW over discrete time periods. Each time period is subdivided into two phases: the reproductive phase and the learning phase. Environmental changes may occur after the reproductive phase. In the reproductive phase, all individuals are synchronously replaced by newborns, where the individual currently occupying a given vertex will be replaced either by that individual's offspring or offspring of a randomly chosen neighbour. The offspring will inherit the learning strategy of the parent. Note that although we use the term 'reproductive', our model is also consistent with a process without biological reproduction. In the latter interpretation, we effectively assume pay-off-biased cultural transmission, in which each individual may copy a neighbour's learning strategy depending on the relative pay-offs.

Consider a pair of individuals competing for a focal vertex, that is, the individual currently occupying that vertex (self) and a randomly chosen neighbour (opponent). We assume that the probability, f , that the self will be replaced by the opponent's offspring or adopt the opponent's learning strategy follows the logit choice rule (also known as the Fermi rule) [35–37]:

$$f(\pi_{\text{opp}}, \pi_{\text{self}}) = \frac{1}{1 + \exp[-\alpha(\pi_{\text{opp}} - \pi_{\text{self}})]}, \quad (2.1)$$

where π_{self} and π_{opp} are the pay-offs to the self and the opponent, respectively, and α is the intensity of selection.

In the learning phase, all individuals acquire correct or wrong behaviour. IL always acquires the correct behaviour, whereas SL copies the behaviour of cultural models. The cultural model of SL occupying a given vertex is chosen at random from those who occupied that vertex or the neighbouring vertices before the reproductive phase, which approximates the process of vertical and oblique transmission [1] (see the electronic supplementary material for a model assuming vertical and horizontal transmission). When an environmental change occurred, none of the pre-existing behaviours is adaptive and, as a result, SL cannot acquire the correct behaviour after an environmental change.

2.2. Overview of the analysis

From equation (2.1), we can derive the probability that an individual is SL after the reproductive phase, depending on phenotypes of its neighbours and costs of IL and SLW. Under the assumption of large population size, averaging the probability over the population gives the frequency of SL after the reproductive phase or a recursion equation for the frequency of SL in the population. From the recursion equation, we obtain the condition under which SL are selectively favoured over IL. We also identify a periodically stable state of changes in the frequency of SL to derive the long-term average frequency of SL in terms of coefficients of relatedness, which measure the association between two neighbouring individuals. See the electronic supplementary material for more details.

2.3. Relative advantage of social learning

Besides, we consider only homogeneous networks, where each individual has the same number of neighbours. In a

homogeneous network, everybody is equally likely to be somebody's opponent. In a heterogeneous network, on the other hand, there may be hub individuals having much more neighbours than others and thus highly likely to be somebody's opponent. Suppose, for example, a network consisting of one (hub) individual and five neighbours, with each neighbour being connected only to the hub (i.e. a star graph). In this network, if the hub is SL and the others are IL, the expected frequency of SL among the opponents of a randomly chosen individual is 5/6, which is larger than the frequency of SL in the population, 1/6. Hence, in this specific case, the frequency of SL may increase simply because SL is more likely to be somebody's opponent than IL is, even if no pay-off-dependent selective process such as modelled by equation (2.1) is at work. Because our main focus here is on investigating the behaviour of the system dominated by the pay-off-dependent selective process, we limit our attention to homogeneous networks. See the electronic supplementary material for more general analysis allowing for such cases.

We derive the condition under which SL is selectively favoured over IL under weak selection. Let $\lambda_{X,Y}$ denote the probability that the self is X and the opponent is Y , where X and Y denote either IL, SL or SLW. As shown in the electronic supplementary material, SLs increase in frequency if

$$\frac{c}{s} > \frac{\lambda_{\text{IL,SLW}} + \lambda_{\text{SLW,IL}}}{\lambda_{\text{IL,SL}} + \lambda_{\text{SL,IL}}}. \quad (2.2)$$

Thus, social learning is more likely to become prevalent in such social structure that renders the ratio $(\lambda_{\text{IL,SLW}} + \lambda_{\text{SLW,IL}})/(\lambda_{\text{IL,SL}} + \lambda_{\text{SL,IL}})$ smaller. This ratio measures the extent to which IL and SLW are socially connected in comparison with the extent of connection between IL and SL, and is likely to be small when, for example, SLW forms a cluster that is surrounded by SLC, thereby separated from clusters of IL (e.g. the edge effect in two-dimensional lattice; see [21]).

2.4. Long-term average frequency of social learners

Here, we derive the long-term average frequency of SL assuming weak selection and stable coexistence of the three phenotypes. Let p and p' denote the frequency of SL before and after the reproductive phase, respectively. Because we assume periodic change of environment, the population will not reach an equilibrium state where IL, SLC and SLW coexist in a constant ratio. However, the long-term average of p will be stable if change in p is, on average, zero over T time periods between two consecutive environmental changes. When this condition is met, we regard the population as having reached a periodically stable state. We also denote by w the proportion of SLW among SL.

For our purpose, we need to define three statistics, R , r and ρ , as follows. Note first that, from the symmetry between the self and opponent in homogeneous networks, we can write $\lambda_{\text{IL,SL}} = \lambda_{\text{SL,IL}} = (1 - P(\text{SL}|\text{SL}))p$, where $P(\text{SL}|\text{SL})$ is the probability that one of two neighbouring individuals is an SL under the condition that the other is an SL. This can be rewritten as $\lambda_{\text{IL,SL}} = \lambda_{\text{SL,IL}} = (1 - R)(1 - p)p$, where

$$R = \frac{P(\text{SL}|\text{SL}) - p}{1 - p}. \quad (2.3)$$

The coefficient R gives the standard genetic relatedness between two neighbouring individuals, which can be rewritten in terms of variances and covariances [38].

Second, we can write $\lambda_{IL,SLW} = \lambda_{SLW,IL} = (1-r)(1-p)pw$, where

$$r = \frac{P(SL|SLW) - p}{1-p}. \quad (2.4)$$

$P(SL|SLW)$ is the probability that one of two neighbouring individuals is an SL under the condition that the other is an SLW. Coefficient r , which measures the association between an individual's genotype and the neighbour's phenotype, has been formalized as the 'gene-to-culture relatedness' [39].

Finally, we introduce the following statistic

$$\rho = \frac{P(SL'|SLW) - p'}{1-p'}, \quad (2.5)$$

where $P(SL'|SLW)$ is the probability that an individual sampled after the reproduction phase is an SL under the condition that its cultural model is an SLW. Coefficient ρ measures the extent to which social learning occurs non-randomly, where positive ρ indicates that SL is more likely to imitate SLW than they would in well-mixed populations.

To proceed further, let us assume that once a periodically stable state is reached, changes in p as well as R , r and ρ are much smaller than those in w , so that they can be regarded as constant. This assumption may be justified when selection is weak. Under this assumption, we obtain (see the electronic supplementary material)

$$\hat{p} \approx 1 - \frac{s(1-\hat{r})}{cT(1-\hat{\rho})(1-\hat{R})}, \quad (2.6)$$

where the hat symbol represents long-term averages. The long-term average proportion of SLW among SL is also given by

$$\hat{w} = \frac{c(1-\hat{R})}{s(1-\hat{r})}. \quad (2.7)$$

Note that $\hat{R} = \hat{r} = \hat{\rho} = 0$ in a well-mixed population.

2.5. The effects of social structure

The frequency of SL at a periodically stable state in a homogeneous network can be greater than that would be in a well-mixed population in that case we can conclude that social structure promotes the evolution of social learning. From equation (2.6), this is the case if

$$\frac{1-\hat{R}}{1-\hat{r}} > \frac{1}{1-\hat{\rho}}. \quad (2.8)$$

Quantities $(1-\hat{R})/(1-\hat{r})$ and $1/(1-\hat{\rho})$ jointly specify characteristics of a homogeneous network, and both are one in large well-mixed populations. In structured populations, SL tends to form clusters, which render \hat{R} and \hat{r} larger. Because only IL can behave correctly after an environmental change, SL in the inside of a cluster of SL tends to acquire a maladaptive behaviour. On the contrary, SL on the edge of the cluster can acquire the correct behaviour from IL. As a result, SLW's opponent may be more likely to be SL than SLC's opponent is. This is true when $\hat{R} < \hat{r}$ in that case competition between SLW and IL tends to be evaded. As this effect is more prominent when $(1-\hat{R})/(1-\hat{r})$ is larger, this quantity can be regarded as a measure of localization in competition. Further, as mentioned above, $\hat{\rho}$ represents the extent to which social learning occurs non-randomly, and thus $1/(1-\hat{\rho})$ can be

considered as a measure of localization in cultural transmission. In sum, what inequality (2.8) proves is that social structure favours the evolution of social learning if the effect of localization in competition is stronger than that of localization in cultural transmission [23,24].

2.6. Monte Carlo simulations

To examine the validity of equations (2.6) and (2.7), we conduct a series of Monte Carlo simulations. We demonstrate that our approximations are valid for a wide range of parameter values. In fact, this turns out to be true even for heterogeneous networks and under relatively strong selection. The simulation is designed to model the same process as in our analytical model except that we now allow for mutation on the learning strategy in order to avoid the population becoming monomorphic by stochastic effect. Mutation, which switches an individual's learning strategy, is assumed to occur after the reproductive phase with probability $\mu = 10^{-5}$ per individual per time period. We obtain long-term average frequencies of phenogenotypes for $T \in \{5, 10, \dots, 100\}$ and $\alpha \in \{1, 10\}$, using $s = 0.1$ and $c = 0.02$. Results with other values for s and c are described in the electronic supplementary material. Throughout our simulations, we set $N = 10\,000$.

We track changes in the frequencies of the phenogenotypes in three types of networks: complete graphs, regular random graphs and scale-free networks. Complete graphs, in which each of N individuals is connected to $N-1$ others, are considered as a model of well-mixed populations. Regular random graphs, in which each individual has the same number, k , of neighbours chosen at random from the $N-1$ others ($1 < k < N-1$), represent homogeneous networks. Scale-free networks, in which the number of neighbours follows the power law with mean k , are included as an example of heterogeneous networks to evaluate whether or not our approximation derived under the assumption of homogeneous networks may be also applicable to heterogeneous networks. We generated scale-free networks based on the Barabási–Albert model [40]. In our simulations, we use $k \in \{8, 16\}$. Results presented are for $k = 8$ and they are not altered qualitatively by using $k = 16$ (see the electronic supplementary material).

In the beginning of each simulation run, either IL or SL is randomly assigned to each individual with equal probability. Each simulation run consists of 50 000 time periods. Long-term average frequencies of the three phenogenotypes are obtained by averaging over the last 10 000 time periods and then further averaged over 100 simulation runs. For complete graphs, the 100 runs are of the same network structure. For regular random graphs and scale-free networks, the 100 runs consist of five runs for each of 20 different network structures of the same type.

We assume that for a given network structure, R , r and ρ converge to certain values as long as selection is sufficiently weak. Accordingly, we conduct additional simulations for regular random graph and scale-free network in which selection is absent ($\alpha = 0$) to obtain estimates of the three coefficients. In these simulations, environmental change occurs only once at time period 40 000, and the three coefficients for each of the 100 time periods following the environmental change are calculated. We run 1000 simulations and then obtain estimates of \hat{R} , \hat{r} and $\hat{\rho}$ as the

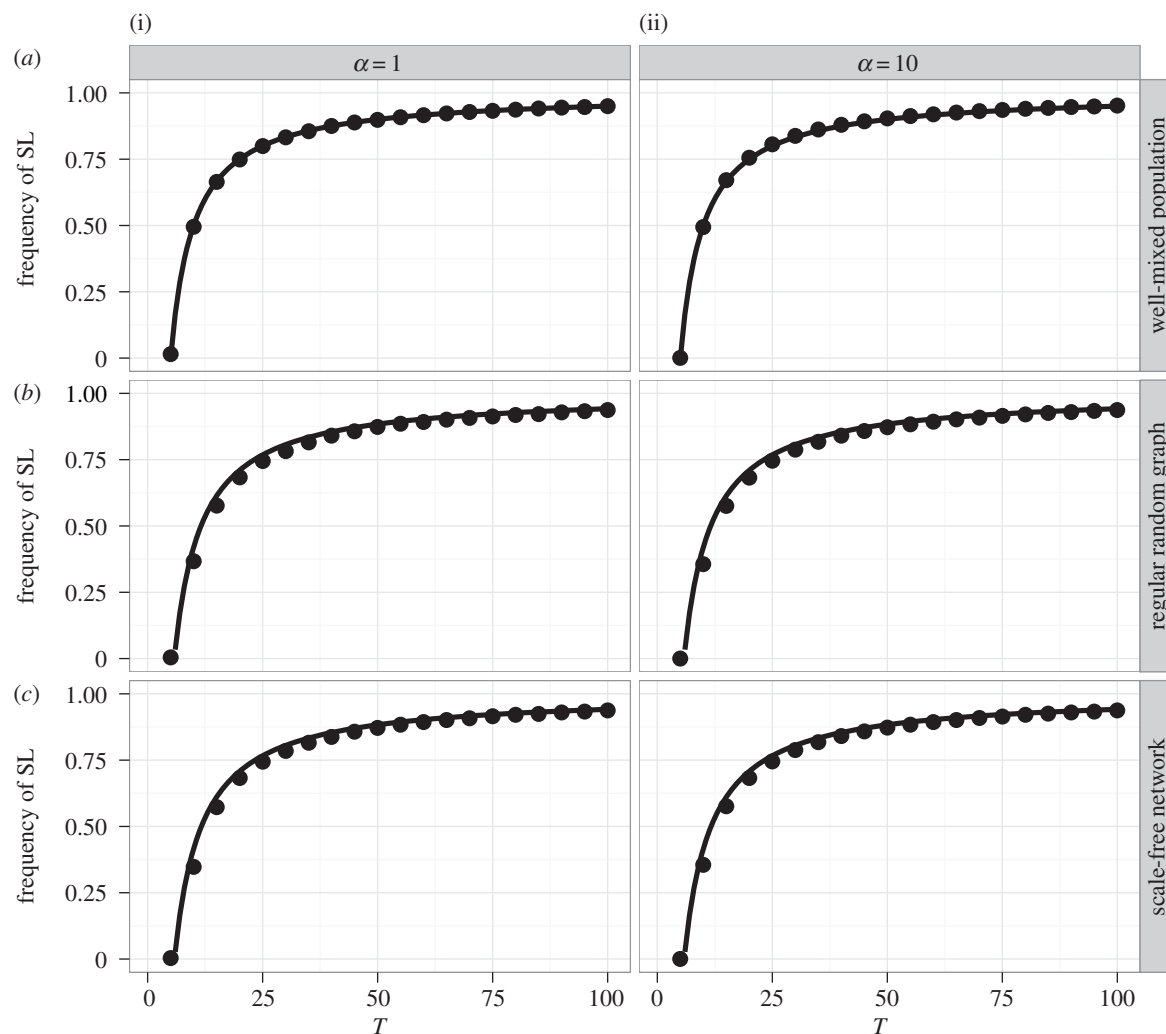


Figure 1. The long-term average frequency of SL, plotted against the environmental stability, T . The circles represent the results by Monte Carlo simulations and the curves represent the analytical approximations. The results for well-mixed populations (a), regular random graphs (b) and scale-free networks (c), respectively. For weak ($\alpha = 1$) (i) and strong ($\alpha = 10$) (ii) selection, respectively. Parameter values used are $s = 0.1$, $c = 0.02$.

Table 2. Estimates of the three relatedness measures for regular random graphs and scale-free networks.

	\hat{R}	\hat{r}	$\hat{\rho}$
regular random graph	0.100686	0.186913	0.220344
scale-free network	0.099856	0.180453	0.220302

median of the 100×1000 values of the three coefficients. Table 2 provides the estimates thus obtained.

The upper panels of figure 1 compare the long-term average frequency of SL for complete graphs obtained numerically by simulations and analytically by equation (2.6). For the parameter values examined, the equation predicts the numerical results very accurately. The middle panels of figure 1 show the long-term average frequency of SL for regular random graphs obtained numerically and analytically. For a broad range of T , equation (2.6) predicts the numerical results very accurately except when T is very small. The lower panels of figure 1 compare numerical results for scale-free networks with the analytical results. Numerical

results for scale-free networks are almost identical with those for regular random graphs, suggesting that equation (2.6) is also applicable to scale-free networks for the parameter values examined.

The upper panels of figure 2 compare the long-term average proportions of SLW among SLs for well-mixed populations obtained analytically by equation (2.7) and numerically by Monte Carlo simulations. Under weak selection ($\alpha = 1$), equation (2.7) predicts the numerical results very accurately except when T is small, whereas under strong selection ($\alpha = 10$), the prediction is slightly less accurate. The discrepancies between the analytical and numerical results are smaller for intermediate values of T , for which our assumption of stable coexistence of the three phenotypes is more likely to hold. The middle and lower panels of figure 2 show the results for regular random graphs and scale-free networks, respectively, obtained by equation (2.7) and Monte Carlo simulations. The general trends for the two types of networks are qualitatively the same as those for well-mixed populations, although equation (2.7) is less accurate for scale-free networks when T is large even when selection is weak.

Our estimates of the relatedness measures shown in table 2 do not satisfy inequality (2.8) for regular random graphs or scale-free networks. This suggests that social structure

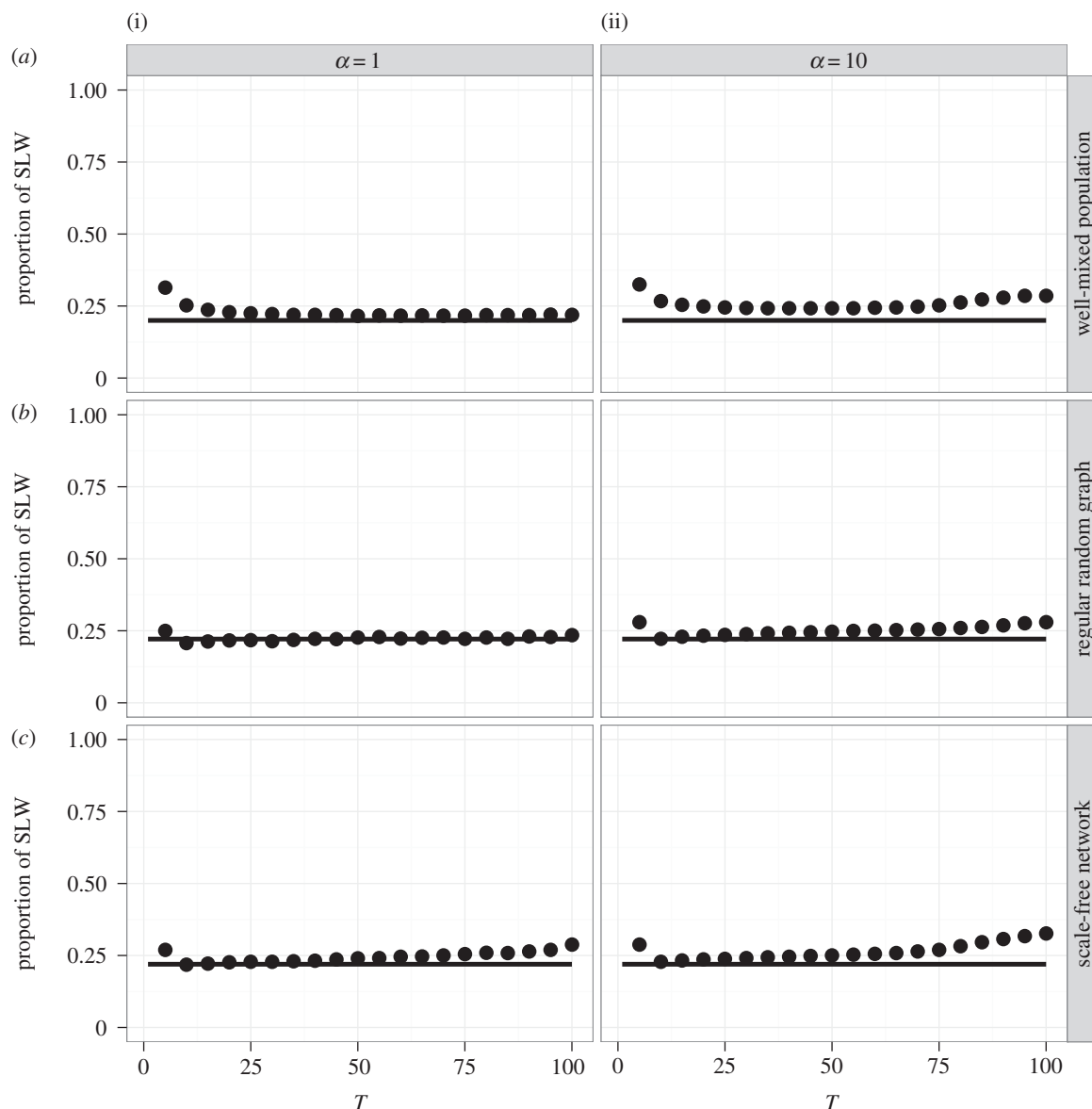


Figure 2. The long-term average proportion of SLW among SL, plotted against the environmental stability, T . The circles represent the results by Monte Carlo simulations and the curves represent the analytical approximations. The results for well-mixed populations (a), regular random graphs (b) and scale-free networks (c), shown respectively. For weak ($\alpha = 1$) (i) and strong ($\alpha = 10$) (ii) selection, shown respectively. Parameter values used are: $s = 0.1$, $c = 0.02$.

disfavours the evolution of social learning when selection is weak.

3. Discussion

We have explored the evolution of learning strategies on social networks. First, we have derived the condition under which natural selection favours the evolution of social learning, as opposed to individual learning, for a broad range of social networks (equation (2.2)). Second, we have obtained an analytical approximation of the long-term average frequency of SLs in homogeneous networks (equation (2.6)), from which we have specified the condition, as a function of three relatedness measures, for social structure to promote the long-term evolution of social learning (inequality (2.8)). Third, by Monte Carlo simulations, we have confirmed that our approximation is valid for wide range of parameter values in complete graphs, regular random graphs, and scale-free networks. Finally, we have found that our estimates of the relatedness measures (table 2) do not satisfy inequality

(2.8) for regular random graphs or scale-free networks, suggesting that social structure generally disfavours the evolution of social learning under weak selection.

Our mathematical analysis has shown that the long-term average frequency of SL depends on three coefficients, R , r and ρ , that represent non-random association between phenotypes. It has also demonstrated that the relative magnitudes of two quantities, $(1 - R)/(1 - r)$ and $1/(1 - \rho)$, which measure localization in competition and cultural transmission, respectively, determine whether the long-term average frequency of SL is larger or smaller in a homogeneously structured population than in an unstructured population. Therefore, the evolution of social learning is facilitated by the presence of such a social structure that localizes competition to a greater extent than it localizes cultural transmission.

In a structured population, SL acquires the correct behaviour less efficiently than in a well-mixed population because the structure localizes cultural transmission. When the degree of localization in cultural transmission is strong so that the inequality in (2.8) is reversed, SL tends to compete with neighbouring IL without acquiring the correct behaviour.

This results in lower long-term average frequency of SL in structured populations than in well-mixed populations. On the other hand, population structure also localizes competition in a way that competition between SL and IL tends to be evaded. More to the point, certain structures hinder competition between SLW and IL to a greater extent than they hinder competition between SLC and IL, in which case the left-hand side of inequality (2.8) exceeds 1. When the degree of localization in competition is relatively strong so that inequality (2.8) holds, SLW forms a cluster surrounded by SLC, whereby competition between SLW and IL is evaded. Some previous studies on the evolution of social learning on two-dimensional lattices have reported that the long-term average frequency of SL can be higher than those in unstructured populations, suggesting that lattice-structured populations can be more likely to favour the evolution of social learning, because the separation by SLC of SLW and IL can be readily achieved in this type of structure [21,24]. In such a structure, SL's disadvantage owing to the delay in information transmission can be more than compensated for by the advantage of readiness in surrounding clusters of ILs. The edge effect proposed by Rendell *et al.* is an example of this case [21].

While most of previous studies on the evolution of individual versus social learning heavily depended on numerical simulations, Kobayashi and co-workers have provided analytical formula for the approximated long-term average frequency of SLs [23,39]. In particular, Kobayashi & Ohtsuki's [39] inclusive-fitness analysis has pointed out the importance of a balance of two types of localization caused by population structure [39]. In one sense, our analysis is less restrictive than Kobayashi & Ohtsuki's [39]: a major difference between the two studies is that while Kobayashi and Ohtsuki used Wright's island model, which assumes a population consisting of infinitely many islands, our approximate formulas have been shown to be applicable to a fairly wide range of population structure. In another sense, our analysis is more restrictive: of the three mechanisms for stable coexistence of ILs and SLs investigated by Kobayashi & Ohtsuki [39], which are temporal fluctuation of environments, local population sparsity and spatial environmental heterogeneity, we focused on the first one. Our equation (2.6) is comparable with equation (2.1) in Kobayashi & Ohtsuki [39]. These equations are consistent in showing the importance of R and r , whereas ρ appears only in our equation. A possible reason why an additional measure is required in our model is as follows. In Wright's island model, each individual within an island is connected with all other individuals within the island, so that cultural transmission can be considered to occur within a small-sized well-mixed population, or possible cultural models for all individuals within an island is identical. In contrast, in a general homogeneous network, possible cultural models of two individuals can be different even if they are neighbours.

When reproduction and social learning are mediated by the same social network, as have been assumed in this study, the two types of localization tend to go hand in hand. In this case, therefore, social structure can favour or disfavour the evolution of social learning depending on details of the structure and parameter values, which may, in part, explain the discrepancy between the previous studies [21,23]. In some combinations of population structures and parameter values, localization in competition can overwhelm

localization in cultural transmission [21]. It may be interesting to consider two networks simultaneously, one through which individuals update learning strategies and the other through which SLs acquire cultural traits [41,42].

Social structure is also relevant to Rogers' paradox, which states that the mean fitness in a given population where ILs and SLs coexist at equilibrium is equal to that in the absence of SLs [43]. Rendell *et al.*'s [21] simulations suggested that the mean fitness at the stable coexistence of ILs and SLs tends to be lower in a lattice-structured population than in a well-mixed population, where the former harbours more SLs than the latter [21]. Kobayashi & Wakano [23] named this phenomenon strong Rogers' paradox and confirmed that the paradox also holds in their infinite island model [23]. Using equation (2.7), we derive the expectation of the mean pay-off, $\bar{\pi}$, in a periodically stable state

$$E_T[\bar{\pi}] = 1 - c - \hat{c}\hat{p} \left[\frac{1 - \hat{R}}{1 - \hat{r}} - 1 \right]. \quad (3.1)$$

Hence, the mean pay-off in a periodically stable state is on average smaller than the mean pay-off in the absence of social learning, $1 - c$, if $\hat{R} < \hat{r}$ is met. In a well-mixed population, for $R = r$ always holds, the mean pay-off in a periodically stable state equals the mean pay-off in the absence of social learning. In other words, equation (3.1) demonstrates that strong Rogers' paradox holds whenever social structure creates a positive effect of localization in competition.

We have not found any significant difference between regular random graphs and scale-free networks in terms of the long-term average frequency of SL. This may seem at odds with previous studies, in particular, those that addressed the evolution of cooperation. Those studies suggested that heterogeneous networks tend to facilitate the evolution of cooperation when compared with homogeneous networks ([44–46] but see [47,48]). This effect is likely to depend on the fact that in heterogeneous networks, there exist hub players who can enjoy large total pay-offs derived from a large number of interactions. Indeed, it has been pointed out that the advantage of scale-free networks in facilitating the evolution of cooperation vanishes if an individual's fitness is determined not by the total pay-off but by the averaged pay-off over interactions [49]. Because pay-offs do not accumulate over interactions in our model, our finding that homogeneous and heterogeneous networks produce almost the identical evolutionary outcomes does not necessarily contradict with the previous studies.

Our results have some anthropological and sociological implications. Kobayashi & Ohtsuki [39], as well as this study, showed that the evolution of social learning tends to be facilitated in unstructured populations, where there are more potential social partners than in structured populations. Based on their mathematical analysis and previous findings in anthropology [9,50,51], they proposed a possible evolutionary association between the heavy reliance on social learning and large group size in modern humans. In the electronic supplementary material, we confirmed that a larger number of neighbours, k , can provide the higher frequency of SLs even in the same type of network structure (electronic supplementary material, figure S1), suggesting that Kobayashi & Ohtsuki's [39] idea can be supported by models with explicit network structure. Because these results indicate the possible

existence of a feedback loop between the heavy reliance of humans on social learning and increasing group size, developing a mathematical model to consider the coevolution of sociality and learning strategies under various costs of having many social partners is a possible future direction of this study.

Homophily, or the degree to which individuals associate with similar others, may affect the spread of innovations. Although homophilic contacts are better at achieving mutual understanding, the importance of heterophilic contacts [52] and the 'strength of weak ties' [53] in the spread of innovations have been emphasized. In this study, equation (2.7) suggests that whenever social structure promotes localization of competition (i.e. $\hat{R} < \hat{r}$), the proportion of SLW among SL is heightened than expected in a well-mixed population. This may be in accord with the idea that homophilic contacts slow down the spread of innovations.

It has been argued that heterogeneity in a social network can affect cultural evolution in multiple ways. On the one hand, heterogeneity may promote the spread of adaptive cultural traits. Hub individuals in a heterogeneous network may have a greater influence on others than non-hub individuals, leading to one-to-many transmission, which has been considered as a key to rapid cultural evolution ([54,55] but see [56,57]). On the other hand, heterogeneous networks may also promote maladaptive cultural evolution. While adaptive aspects of culture have been emphasized, some maladaptive behaviours may be also culturally transmitted. For instance, the fertility transitions, prestige-seeking behaviours and contagious suicide have been of particular

interest to anthropologists as well as social scientists [1,2,51,58–66]. Learning biases and modes of cultural transmission have been deemed as factors contributing to maladaptive cultural evolution [1,2,61,63]. An interesting question is whether spread of some maladaptive behaviours can be also attributable to heterogeneous network architectures [66]. Using estimated values of \hat{R} , \hat{r} and $\hat{\rho}$ for regular random graphs and scale-free networks (table 2), we compared the long-term average proportion of SLW among SL between the two types of network architectures. It turns out that they are not significantly different, suggesting that heterogeneity plays only a minor role in the spread of cultural traits. Nevertheless, it is possible that heterogeneity makes a difference when it works in coordination with certain learning biases, including the pay-off-dependent bias. The pay-off-dependent bias may play an important role in shaping structure of cultural transmission networks: real cultural transmission networks are heterogeneous and the heterogeneity may have emerged from adaptive learning biases [67,68]. Theoretical investigation on a dynamic formation of cultural transmission networks through a process in which individuals preferentially choose models to acquire their behaviours may provide better understanding about human learning bias and the nature of cultural transmission.

Acknowledgements. We thank three anonymous reviewers for their helpful comments.

Funding statement. K.T. is supported by Grant-in-Aid for JSPS fellows and CREST, JST. Y.K. is supported by Monbukagakusho grant no. 22101004.

References

- Cavalli-Sforza LL, Feldman MW. 1981 *Cultural transmission and evolution: a quantitative approach*. Princeton, NJ: Princeton University Press.
- Boyd R, Richerson PJ. 1985 *Culture and the evolutionary process*. Chicago, IL: University of Chicago Press.
- Galef BGJ, Laland KN. 2005 Social learning in animals: empirical studies and theoretical models. *Bioscience* **55**, 489–499. (doi:10.1641/0006-3568(2005)055[0489:SLIAES]2.0.CO;2)
- Feldman MW, Aoki K, Kumm J. 1996 Individual versus social learning: evolutionary analysis in a fluctuating environment. *Anthropol. Sci.* **104**, 209–231. (doi:10.1537/ase.104.209)
- Wakano JY, Aoki K, Feldman MW. 2004 Evolution of social learning: a mathematical analysis. *Theor. Popul. Biol.* **66**, 249–258. (doi:10.1016/j.tpb.2004.06.005)
- Aoki K, Wakano JY, Feldman MW. 2005 The emergence of social learning in a temporally changing environment: a theoretical model. *Curr. Anthropol.* **46**, 334–340. (doi:10.1086/428791)
- Wakano JY, Aoki K. 2006 A mixed strategy model for the emergence and intensification of social learning in a periodically changing environment. *Theor. Popul. Biol.* **70**, 486–497. (doi:10.1016/j.tpb.2006.04.003)
- Freeberg TM, Dunbar RIM, Ord TJ. 2012 Social complexity as a proximate and ultimate factor in communicative complexity. *Phil. Trans. R. Soc. B* **367**, 1785–1801. (doi:10.1098/rstb.2011.0213)
- Dunbar RIM, Shultz S. 2007 Understanding primate brain evolution. *Phil. Trans. R. Soc. B* **362**, 649–658. (doi:10.1098/rstb.2006.2001)
- Ihara Y. 2011 Evolution of culture-dependent discriminate sociality: a gene-culture coevolutionary model. *Phil. Trans. R. Soc. B* **366**, 889–900. (doi:10.1098/rstb.2010.0247)
- Pearce E, Stringer C, Dunbar RIM. 2013 New insights into differences in brain organization between netherthals and anatomically modern humans. *Proc. R. Soc. B* **280**, 20130168. (doi:10.1098/rspb.2013.0168)
- Christakis NA, Fowler JH. 2007 The spread of obesity in a large social network over 32 years. *N. Engl. J. Med.* **357**, 370–379. (doi:10.1056/NEJMsa066082)
- Mann J, Stanton MA, Patterson EM, Bienenstock EJ, Singh LO. 2012 Social networks reveal cultural behavior in tool-using dolphins. *Nat. Commun.* **3**, 980. (doi:10.1038/ncomms1983)
- Claidière N, Messer EJE, Hoppitt WJE, Whiten A. 2013 Diffusion dynamics of socially learned foraging techniques in squirrel monkeys. *Curr. Biol.* **23**, 1251–1255. (doi:10.1016/j.cub.2013.05.036)
- Webster MM, Atton N, Hoppitt WJE, Laland KN. 2013 Environmental complexity influences association network structure and network-based diffusion of foraging information in fish shoals. *Am. Nat.* **181**, 235–244. (doi:10.1086/668825)
- Cavalli-Sforza LL, Feldman MW. 1983 Paradox of the evolution of communication and of social interactivity. *Proc. Natl Acad. Sci. USA* **80**, 2017–2021. (doi:10.1073/pnas.80.7.2017)
- Aoki K, Feldman MW. 1987 Toward a theory for the evolution of cultural communication: coevolution of signal transmission and reception. *Proc. Natl Acad. Sci. USA* **84**, 7164–7168. (doi:10.1073/pnas.84.20.7164)
- Aoki K, Feldman MW. 1989 Pleiotropy and preadaptation in the evolution of human language capacity. *Theor. Popul. Biol.* **35**, 181–194. (doi:10.1016/0040-5809(89)90016-6)
- Takahasi K, Aoki K. 1995 Two-locus haploid and diploid models for the coevolution of cultural transmission and paternal care. *Am. Nat.* **146**, 651–684. (doi:10.1086/285819)
- Tamura K, Ihara Y. 2011 Classes of communication and the conditions for their evolution. *Theor. Popul. Biol.* **79**, 174–183. (doi:10.1016/j.tpb.2011.03.001)
- Rendell L, Fogarty L, Laland KN. 2010 Rogers' paradox recast and resolved: population structure and the evolution of social learning strategies. *Evolution* **64**, 534–548. (doi:10.1111/j.1558-5646.2009.00817.x)

22. Fogarty L, Rendell L, Laland KN. 2012 The importance of space in models of social learning, cultural evolution and niche construction. *Adv. Complex Syst.* **15**, 1150001. (doi:10.1142/S0219525911003311)
23. Kobayashi Y, Wakano JY. 2012 Evolution of social versus individual learning in an infinite island model. *Evolution* **66**, 1624–1635. (doi:10.1111/j.1558-5646.2011.01541.x)
24. Tamura K, Ihara Y. 2012 Evolution of social learning in lattice-structured populations. *Lett. Behav. Evol. Sci.* **3**, 25–29. (doi:10.5178/lebs.2012.23)
25. Hamilton WD. 1964 The genetical theory of social behaviour I & II. *J. Theor. Biol.* **7**, 1–52. (doi:10.1016/0022-5193(64)90038-4)
26. Allison PD. 1992 Cultural relatedness under oblique and horizontal transmission rules. *Ethol. Sociobiol.* **13**, 153–169. (doi:10.1016/0162-3095(92)90031-X)
27. Lehmann L, Feldman MW. 2008 The co-evolution of culturally inherited altruistic helping and cultural transmission under random group formation. *Theor. Popul. Biol.* **73**, 506–516. (doi:10.1016/j.tpb.2008.02.004)
28. Lehmann L, Feldman MW, Foster KR. 2008 Cultural transmission can inhibit the evolution of altruistic helping. *Am. Nat.* **172**, 12–24. (doi:10.1086/587851)
29. Frank SA. 1998 *Foundation of social evolution*. Princeton, NJ: Princeton University Press.
30. Roussel F. 2004 *Genetic structure and selection in subdivided population*. Princeton, NJ: Princeton University Press.
31. West SA, El Mouden C, Gardner A. 2011 Sixteen common misconception about the evolution of cooperation in humans. *Evol. Hum. Behav.* **32**, 231–262. (doi:10.1016/j.evolhumbehav.2010.08.001)
32. Grafen A. 2007 An inclusive fitness analysis of altruism on a cyclical network. *J. Evol. Biol.* **20**, 2278–2283. (doi:10.1111/j.1420-9101.2007.01413.x)
33. Ohtsuki H. 2010 Evolutionary games in wright's island model: kin selection meets evolutionary game theory. *Evolution* **64**, 3344–3353. (doi:10.1111/j.1558-5646.2010.01117.x)
34. Lion S, Jansen VAA, Day T. 2011 Evolution in structured population: beyond the kin versus group debate. *Trends Ecol. Evol.* **26**, 193–201. (doi:10.1016/j.tree.2011.01.006)
35. Szabó G, Tóke C. 1998 Evolutionary prisoner's dilemma game on a square lattice. *Phys. Rev. E* **58**, 69–73. (doi:10.1103/PhysRevE.58.69)
36. Goeree JK, Holt CA. 1999 Stochastic game theory: for playing games, not just for doing theory. *Proc. Natl Acad. Sci. USA* **96**, 10 564–10 567. (doi:10.1073/pnas.96.19.10564)
37. Traulsen A, Nowak M, Pacheco J. 2006 Stochastic dynamics of invasion and fixation. *Phys. Rev. E* **74**, 1–5. (doi:10.1103/PhysRevE.74.011909)
38. McElreath R, Boyd R. 2007 *Mathematical models of social evolution: a guide for the perplexed*. Chicago, IL: University of Chicago Press.
39. Kobayashi Y, Ohtsuki H. 2014 Evolution of social versus individual learning in a subdivided population revisited: comparative analysis of three coexistence mechanisms using the inclusive-fitness method. *Theor. Popul. Biol.* **92**, 78–87. (doi:10.1016/j.tpb.2013.12.003)
40. Barabási AL, Albert R. 1999 Emergence of scaling in random networks. *Science* **286**, 509–512. (doi:10.1126/science.286.5439.509)
41. Suzuki R, Arita T. 2009 Evolution of cooperation on different pairs of interaction and replacement networks with various intensity of selection. In *2009 IEEE Congress on Evolutionary Computation, 19 May, Trondheim, Norway*, pp. 494–499. Piscataway, NJ: IEEE Press.
42. Ohtsuki H, Nowak MA, Pacheco J. 2007 Breaking the symmetry between interaction and replacement in evolutionary dynamics on graphs. *Phys. Rev. Lett.* **98**, 108106. (doi:10.1103/PhysRevLett.98.108106)
43. Rogers A. 1988 Does biology constrain culture? *Am. Anthropol.* **90**, 819–831. (doi:10.1525/aa.1988.90.4.02a00030)
44. Santos FC, Pacheco JM. 2005 Scale-free networks provide a unifying framework for the emergence of cooperation. *Phys. Rev. Lett.* **95**, 98104. (doi:10.1103/PhysRevLett.95.098104)
45. Santos FC, Pacheco JM, Lenaerts T. 2006 Evolutionary dynamics of social dilemmas in structured heterogeneous populations. *Proc. Natl Acad. Sci. USA* **103**, 3490–3494. (doi:10.1073/pnas.0508201103)
46. Santos FC, Rodrigues JF, Pacheco JM. 2006 Graph topology plays a determinant role in the evolution of cooperation. *Proc. R. Soc. B* **273**, 51–55. (doi:10.1098/rspb.2005.3272)
47. Masuda N. 2007 Participation costs dismiss the advantage of heterogeneous networks in evolution of cooperation. *Proc. R. Soc. B* **274**, 1815–1821. (doi:10.1098/rspb.2007.0294)
48. Konno T. 2011 A condition for cooperation in a game on complex networks. *J. Theor. Biol.* **269**, 224–233. (doi:10.1016/j.jtbi.2010.10.033)
49. Tomassini M, Pesterlacci E, Luthi L. 2007 Social dilemmas and cooperation in complex networks. *Int. J. Mod. Phys. C* **18**, 1173–1185. (doi:10.1142/S0129183107011212)
50. Aiello LC, Dunbar RI. 1993 Neocortex size, group size, and the evolution of language. *Curr. Anthropol.* **34**, 184–193. (doi:10.1086/204160)
51. Richerson PJ, Boyd R. 2004 *Not by genes alone: how culture transformed human evolution*. Chicago, IL: University of Chicago Press.
52. Rogers ME. 1995 *Diffusion of innovations*. New York, NY: Free Press.
53. Granovetter MS. 1973 The strength of weak ties. *Am. J. Sociol.* **78**, 1360–1380. (doi:10.1086/225469)
54. Cavalli-Sforza LL, Feldman MW, Chen KH, Dornbusch SM. 1982 Theory and observation in cultural transmission. *Science* **218**, 19–27. (doi:10.1126/science.7123211)
55. Hewlett BS, Cavalli-Sforza LL. 1986 Cultural transmission among Aka pygmies. *Am. Anthropol.* **88**, 922–934. (doi:10.1525/aa.1986.88.4.02a00100)
56. Aoki K, Lehmann L, Feldman MW. 2011 Rates of cultural change and patterns of cultural accumulation in stochastic models of social transmission. *Theor. Popul. Biol.* **79**, 192–202. (doi:10.1016/j.tpb.2011.02.001)
57. Seki M, Ihara Y. 2012 The rate of cultural change in one-to-many social transmission when cultural variants are not selectively neutral. *Lett. Evol. Behav. Sci.* **3**, 12–16. (doi:10.5178/lebs.2012.20)
58. Durham WH. 1991 *Coevolution: genes, culture, and human diversity/William H. Durham*. Stanford, CA: Stanford University Press.
59. Borgerhoff Mulder M. 1998 The demographic transition: are we any closer to an evolutionary explanation? *Trends Ecol. Evol.* **13**, 266–270. (doi:10.1016/S0169-5347(98)01357-3)
60. Laland K. 1998 Social transmission of maladaptive information in the guppy. *Behav. Ecol.* **9**, 493–499. (doi:10.1093/beheco/9.5.493)
61. Ihara Y, Feldman MW. 2004 Cultural niche construction and the evolution of small family size. *Theor. Popul. Biol.* **65**, 105–111. (doi:10.1016/j.tpb.2003.07.003)
62. Bearman PS, Moody J. 2004 Suicide and friendships among American adolescents. *Am. J. Public Health* **94**, 89–95. (doi:10.2105/AJPH.94.1.89)
63. Ihara Y. 2008 Spread of costly prestige-seeking behavior by social learning. *Theor. Popul. Biol.* **73**, 148–157. (doi:10.1016/j.tpb.2007.10.002)
64. Alvergne A, Gurmu E, Gibson MA, Mace R. 2011 Social transmission and the spread of modern contraception in rural Ethiopia. *PLoS ONE* **6**, e22515. (doi:10.1371/journal.pone.0022515)
65. Tanaka CM, Iwasa Y. 2012 Cultural evolution of a belief controlling human mate choice: dynamic modeling of the *hinoeuma* superstition in Japan. *J. Theor. Biol.* **309**, 20–28. (doi:10.1016/j.jtbi.2012.06.001)
66. Yeaman S, Schick A, Lehmann L. 2012 Social network architecture and the maintenance of deleterious cultural traits. *J. R. Soc. Interface* **9**, 848–858. (doi:10.1098/rsif.2011.0555)
67. Henrich J, Henrich N. 2010 The evolution of cultural adaptations: Fijian food taboos protect against dangerous marine toxins. *Proc. R. Soc. B* **277**, 3715–3724. (doi:10.1098/rspb.2010.1191)
68. Henrich J, Broesch J. 2011 On the nature of cultural transmission networks: evidence from Fijian villages for adaptive learning biases. *Phil. Trans. R. Soc. B* **366**, 1139–1148. (doi:10.1098/rstb.2010.0323)