Contents lists available at ScienceDirect

Animal Behaviour

The effect of innovation and sex-specific migration on neutral cultural differentiation

Samuel Yeaman*, Redouan Bshary, Laurent Lehmann

Department of Biology, University of Neuchâtel

Studies of behaviour are increasingly focusing on acquisition of traits through cultural inheritance. Comparison of patterns of spatial population structure (F_{ST}) between neutral genetic loci and behavioural or cultural traits can been used to test hypotheses about demography, life history, and the mechanisms of inheritance/transmission of these traits in humans, chimpanzees and other animals. Here, we develop analytical expectations to show how F_{ST} in cultural traits can differ strongly from that measured at neutral genetic markers if migration is largely restricted to one sex but social learning is predominantly modelled on the other (e.g. males migrate, females serve as models for cultural traits), if one individual is the learning model for many, or if rates of innovation (individual learning) are high or rates of social learning are low. We discuss how comparisons of F_{ST} between genetic loci and behavioural traits can be applied to evaluate the importance of innovation in shaping patterns of cultural differentiation, as even low rates of innovation can considerably reduce F_{ST}, relative to observed structure at neutral genetic loci. Our results also suggest that differentiation in neutral cultural traits should occur over much smaller scales in species with male migration and female enculturation (or the reverse).

© 2011 The Association for the Study of Animal Behaviour. Published by Elsevier Ltd. All rights reserved.

In their seminal papers on chimpanzee, *Pan troglodytes,* and orang-utan, *Pongo pygmaeus*, cultures, Whiten et al. (1999) and van Schaik et al. (2003) documented patterns of differentiation among groups in a range of behavioural traits. The authors argued that these patterns could not be explained by differences in ecology, and hence concluded that at least some of these traits are likely to be culturally inherited. Considerable controversy has surrounded the problem of cultural transmission in chimpanzees, as some have argued that observed patterns could also be consistent with a genetic basis for behaviours (Laland & Janik 2006), while additional evidence has been presented based on cladistic analysis suggesting the behaviours are more likely to be culturally inherited (Lycett et al. 2007, 2010). As other research on chimpanzees has failed to find significant differences between patterns of variation among groups for genetic loci versus behavioural traits (Langergraber et al. 2011), the extent to which these behaviours are culturally inherited remains unclear. Even assuming many of these traits are culturally inherited, there is little understanding of why they are present in some groups but not in others. What is the role of geographical barriers or the relative contributions of social

versus individual learning for the maintenance of variation of cultural traits between groups?

Several of the traits studied in the great apes seem unlikely to be closely tied to fitness, such as the kiss-squeak in orang-utans or the rain dance in chimpanzees, and might therefore be expected to evolve neutrally. For neutrally evolving genetic traits, measuring the ratio of variance in allele frequency partitioned within versus among groups (population structure, F_{ST}; Weir & Cockerham 1984) has provided a powerful means to study patterns of dispersal and make comparisons among species (Manel et al. 2003). Data on population structure in cultural variants have also been used to study the evolution of birdsong (Lynch & Baker 1994), and to test models of cooperation in humans (Bell et al. 2009). As studies continue to collect data showing variation in behavioural traits within and among groups, measures of population structure may provide a relevant approach to analysing these data and making inferences about the mechanism of transmission. Explicitly applying an approach using F_{ST} to study population structure, where frequencies of alleles (genetic markers) are replaced by abundance frequencies of cultural traits (cultural markers), could increase the power of studies that have relied on comparisons of patterns of variation among groups averaged over many traits (in the absence of data on variance among individuals within groups, e.g. Langergraber et al. 2011).

Even for neutral traits, however, many factors may affect F_{ST} for cultural traits differently than they would affect F_{ST} for genetic traits (Cavalli-Sforza & Feldman 1981; Boyd & Richerson 1985). For

* Correspondence: S. Yeaman, Department of Biology, University of Neuchâtel, 11 Rue Emile Argand, 2000, Neuchatel, Switzerland.
  E-mail address: samuel.deblasi@unine.ch (S. Yeaman).

instance, in addition to vertical transmission from parent to offspring, cultural transmission (social learning) can also occur obliquely between individuals from different generations, or horizontally between individuals of the same generation (Cavalli-Sforza & Feldman 1981). When such social learning is independent of the frequency of cultural variants in a group, it may occur according to the class or status of the exemplar individual (such as sex, age, wealth, prestige, etc.), so that one influential individual may act as the cultural parent of many others in a social network, thereby causing one-to-many transmission (Cavalli-Sforza & Feldman 1981). But social-learning rules may also be frequency dependent (or a more complex combination of various updating rules). For instance, individuals may express preferences for common (conformist transmission) or minority (contrariness transmission) variants, or may even express preferences for a partial consensus (Figure 2 in Lumsden & Wilson 1980).

Although theoretical research in cultural evolution has explicitly evaluated neutral F_{ST} and other measures of variance within versus among groups (Cavalli-Sforza & Feldman 1973, 1981; Boyd & Richerson 1982, 1985; Lehmann et al. 2008, 2011), little research has analysed how F_{ST} for neutral cultural traits varies in response to different demographic scenarios. By contrast, the sensitivity of F_{ST} to variation in demographic scenarios for neutrally evolving genetic traits has been analysed in a wide range of models (e.g. Wright 1931; Birky et al. 1983; Weir & Cockerham 1984; Whitlock & McCauley 1990; Whitlock 1992; Rousset 2004, to name but a few). More recently, there has been a renewed interest in ascertaining population homogeneity, subpopulation differentiation and other variance components for cultural traits (Bentley et al. 2004; Kandler & Laland 2009; Danchin & Wagner 2010), and the tools and insights already obtained in the population genetic literature provide a useful starting framework to approach problems in cultural evolution. Exploring how population structure is shaped by cultural transmission is therefore an important component of the rapidly expanding field of cultural evolution (Whiten et al. 2011).

Generally speaking, high rates of migration lead to a very genetically homogeneous population structure, with little differentiation among groups (low F_{ST}). For autosomal genes, if dispersal is restricted to only one sex, it makes little difference whether it is the male or the female sex that is migratory; average population structure will be the same regardless. For uniparentally inherited genes such as those found in chloroplasts or mitochondria or for genes on the sex chromosomes, however, the situation is very different. As mitochondrial genes are inherited from the mother, even complete mixing of males among groups will not affect population structure if females are completely philopatric (Birky et al. 1983; Chesser & Baker 1996; Prugnolle & de Meeus 2002). Empirical differences in maternally versus paternally inherited genetic markers have been used to infer the amount of sex-specific migration in humans (Seielstad et al. 1998; Wilder et al. 2004), chimpanzees (Langergraber et al. 2007) and other animals (Baker et al. 1998; Natoli et al. 2005; Eriksson et al. 2006), so similar questions could be addressed using F_{ST} for cultural traits. While innovation may be analogous to mutation, it is unclear how closely the various well-studied models of mutation would correspond to innovation when a trait can have one of many variants or be completely absent. Although some genetic markers may be scored as present–absent (e.g. AFLP), the absence of a given AFLP band is heritable, whereas the absence of a cultural trait can occur through the failure to learn socially by a particular individual in a group during its life span.

Here, we explore how sex-biased migration, sex-biased social learning (i.e. higher probability of cultural transmission from males or females of the parental generation), and innovation of new variants (individual learning) interact to shape patterns of population structure. How drastically do features particular to cultural transmission modify the standard expectations of population differentiation from population genetics? To start addressing these questions, we quantify levels of population differentiation under various demographic assumptions when individuals can learn traits during their life span through either individual or social learning (e.g. Rogers 1988; Wakano et al. 2004; Richerson & Boyd 2005).

The demographic features and the learning rules we follow are based largely on data collected on primates. With respect to modes of migration, primates are highly variable. Male philopatry is relatively rare, but it does occur in some species, such as chimpanzee and red colobus, Procolobus badius, but female philopatry is much more common (Wrangham 1980; van Schaik 1983; Dunbar 1988). Monkeys and apes typically live in stable groups, and otherwise regularly aggregate with conspecifics (e.g. orang-utans). Thus, inexperienced individuals may encounter a variety of potential models for social learning. However, few individuals seem actually to influence the learning of other group members (de Waal 2001). An important feature of primate life is that, with the exception of few New World monkeys and humans, parental care is provided primarily by females. Hence, social learning of cultural traits from the parental generation may occur primarily through females. In orang-utans, for instance, young individuals appear to learn what to eat from their mother, which leads to high diet overlaps between mothers and their offspring (Jaeggi et al. 2010). Furthermore, field experiments on vervet monkeys, Chlorocebus aethiops, provide evidence that group members are more likely to pay attention to and hence learn socially from members of the philopatric sex (van de Waal et al. 2010; unpublished data). Thus, modes of migration may also influence which sex acts as a cultural exemplar model and which sex is largely ignored.

To represent the effects of these modes on F_{ST}, as well as any more general sex-specific biases in migration and learning, we use a combination of analytical theory and individual-based simulations to study the population structure at equilibrium of the cultural dynamics as a function of the rates of sex-biased migration, and individuals and social learning. We assume that this approach will apply only to effectively neutral cultural traits with linear frequency-dependent updating (which occurs with random mating and no selection of any form on the trait of interest); different patterns are expected if selection or some form of frequency-biased transmission is acting. We also allow for any probability that two individuals in a group share the same cultural parent, which allows us to represent cases where a dominant individual is the model for all individuals in the group (e.g. one-to-many transmission, Cavalli-Sforza & Feldman 1981). To illustrate the effect of sex-biased learning and migration, we compare the population structure for cultural traits with that of biparentally inherited genetic traits under the assumption of random mating within each group and iteroparous reproduction with Poisson-distributed offspring number (i.e. the standard Wright–Fisher model; Ewens 2004; Hartl & Clark 2007). We model cultural traits in two ways: where the trait is either present or absent in each individual or where individuals carry one of k variants of a cultural trait. While we find many patterns that have qualitative similarities to those found for population structure in genetic loci, our results present quantitative descriptions of how some features of cultural transmission can result in diagnostic differences that could be used to study the basis of inheritance and learning in complex animal behaviours. We discuss some possible approaches that could be used in empirical applications, especially for testing for the importance of innovation through individual learning.

Table 1
List of symbols

Symbol	Definition
ρ	Frequency of the focal trait in the population
m_f, m_m	Female and male migration rate
N, N_f, N_m	Total, female, and male group size
θ_{ij}	Probability that two randomly sampled distinct individuals from the same group (of sex i and j) both carry the focal trait or the same focal variant of the trait in the k-variant model
ϕ_{ij}	Probability that for two distinct individuals from the same group (of sex i and j), one carries the focal trait (focal variant for the k-variant model) and the other does not
F_{ij}	Index of cultural differentiation for two individuals of sex i and j
F_{ST}	Index of cultural differentiation averaged over all F_{ij}
p_s	Probability that two individuals in the same group have the same cultural parent
p_f, p_m	Probability that two individuals in the same group have the same female or male cultural parent
p_{ff}, p_{mm}	Probability that two individuals in the same group have different cultural parents that are both female or both male
n_d	Number of groups
k	Number of cultural variants
μ	Probability of gaining a trait by innovation
ν	Probability of switching among variants
β	Probability of gaining a trait by social learning
α	Probability that an offspring that is learning socially has a female cultural parent

MODEL ASSUMPTIONS

Measure of Cultural Population Differentiation

Consider a population consisting of an infinite number of groups, each with a constant number N_m of males and N_f of females (see Table 1 for a list of symbols). Random migration occurs between groups and we assume that individuals within groups are endowed with physiological mechanisms allowing them to learn nongenetically determined traits by two means. First, individuals may acquire traits by individual learning (trial-and-error, lucky accidents) and, second, by social learning (imitation, copying).

Our aim is to evaluate two quantities in this population. First, we evaluate the probability ρ that a randomly sampled adult individual from the population carries a focal cultural trait, whose dynamic is assumed to be independent from other traits the individual may possibly acquire during its life span. For simplicity, we assume that both sexes have the same learning parameters, which means that both males and females carry the trait with the same probability. Second, we aim to characterize the sex-specific differentiation between groups in the frequency of the trait that arises as a consequence of individuals and social learning. In complete analogy with the classical measure of population differentiation used in population genetics (e.g. Wright 1951; Weir & Cockerham 1984; Rousset 2004), we define the index of cultural differentiation as

$$F_{ij} = \frac{\theta_{ij} - \rho^2}{\rho(1 - \rho)}, \tag{1}$$

which is the correlation between the abundance frequency (zero or one) of the focal trait between individual of sex i and another of sex j. This index of differentiation depends on θ_{ij}, which is the probability that two randomly sampled distinct adult individuals from the same group, one of sex i and the other of sex j, both carry the focal cultural trait, and on ρ which is the probability that an individual of either sex carries the trait. Note that we use the symbol F_{ij} as a shorthand for F_{STij} and use F_{ST} to represent the average over all F_{ij}, namely

$$F_{ST} = \frac{N_f\left(N_f - 1\right)F_{ff} + 2N_mN_fF_{fm} + N_m(N_m - 1)F_{mm}}{N(N - 1)}, \tag{2}$$

where N is the total group size ($N = N_f + N_m$) and $N_f(N_f - 1)/N(N - 1)[2N_mN_f/N(N - 1)]$ is the ratio of the number of pairs of distinct females [pairs of males and females] in a group to the total number of pairs of individuals in a group.

When F_{ij} is positive, two individuals sampled in the same group are more likely to carry the focal trait than are two individuals sampled at random from the population. In this case, groups are culturally differentiated and the frequency of individuals carrying the focal cultural trait in different groups will be different. To compute F_{ij} explicitly under given life cycle assumptions, we need to evaluate θ_{ij}, which, in turn, requires that we evaluate the probability ϕ_{ij} that, among two randomly sampled individuals from the same group, one individual carries the focal trait and the other does not. The probability that two randomly sampled individuals from the same group do not carry the focal trait is then given by $1 - \theta_{ij} - \phi_{ij}$.

Life Cycle

The cultural state probabilities (ρ, θ_{ij}, ϕ_{ij}) will be evaluated under the following life cycle, which is basically Wright's (1931) infinite-island model with a round of social and individual learning. (1) Each of the N_f females in a group produces a large number of offspring. (2) Offspring grow and develop, a period during which they may acquire the focal cultural trait from a cultural parent, either by vertical transmission (mother or father) or by oblique transmission. We assume that during this social-learning stage each offspring acquires the focal trait from its cultural parent (which may or may not be its biological parent) with probability β if the cultural parent carries it. We also consider that with probability p_s two randomly sampled offspring in a group may have the same cultural parent, which could be either a female or a male. Hence, the coalescence probability p_s is a composite parameter that could be written as $p_s = p_f + p_m$, where $p_f (p_m)$ is the probability that two randomly sampled offspring in a group may have the same female (male) as their cultural parent. The probability that these two offspring have two different cultural parents that are two females is denoted p_{ff} and the probability that the two cultural parents of the offspring are males is denoted p_{mm}. With probability $1 - p_s - p_{ff} - p_{mm}$ the two offspring thus have two different cultural parents, one a male and the other a female. (3) Each offspring that has not acquired the focal cultural trait through social learning may invent it through individual learning with probability μ. (4) Each offspring either remains philopatric or disperses to another group; males disperse with probability m_m, while females disperse with probability m_f. (5) Individuals of the

parental generation die and density-dependent mortality occurs among juveniles, which brings the male and female population back to the census sizes N_m and N_f, respectively. Random mating occurs and the cycle starts again.

Because different individuals in a group may copy the focal trait from the same cultural parent ($p_s > 0$), there will be fluctuations of focal trait frequencies between groups, which may result in cultural differentiation between them. In other words, cultural kinship (Cavalli-Sforza & Feldman 1981; Allison 1991), $F_{ij} > 0$, may build up among group members as a result of the transmission process.

The description given above assumes that an individual either carries or does not carry a focal cultural trait, and that only one variant segregates in the population with imperfect transmission. But different variants of this trait may also segregate in the population. To take this case into account, we also evaluate F_{ij} under the situation where k cultural variants of the focal trait segregate in the population (as per Crow & Aoki 1984; Enquist et al. 2010), in which case the cultural state probabilities (ρ, θ_{ij}, ϕ_{ij}) are evaluated for a focal variant of the focal trait. Here, we assume that an individual inheriting the focal cultural trait during stage 2 of the life cycle actually inherits the variant of its cultural parent. It then keeps that variant with probability $1 - \nu$ and switches to another specific variant with probability $\nu/(k - 1)$. An individual not inheriting the focal trait from its cultural parent invents it with probability μ during stage 3 of the life cycle, as under the presence–absence model, and in doing so adopts one of the k variants of the trait.

RESULTS

We present the recurrence equations for the dynamics of the cultural state probabilities (ρ, θ_{ij}, ϕ_{ij}) under the presence–absence and the k-cultural variants model in the Appendix. The equilibrium values of these probabilities then allow us to evaluate the fixation index F_{ij}. We begin by separately considering analytical predictions under the presence–absence model ($k - 1$, $\beta < 1$) and the k-variant model assuming perfect cultural transmission ($\beta = 1$). We then use simulations to study a model in which individuals carry either one of the k variants of the focal trait or do not carry the trait at all ($\beta < 1$ and $k > 1$), as this case is analytically more complicated.

Presence–absence Model

Trait abundance frequency

At steady state of the cultural dynamics in the presence–absence model, we find that the probability that a single, randomly sampled individual from the population carries the focal trait is

$$\rho = \frac{\mu}{1 - \beta(1 - \mu)}, \tag{3}$$

which increases with increasing rates of both innovation, μ, and transmission, β (see equations (A1) and (A4) in the Appendix).

Equation (3) holds regardless of the other parameter values of the model. It thus does not depend on the mode of migration and the various probabilities (p_s, p_{mm}, p_{ff}) of copying cultural traits from the same or different parents. This simple result has been established previously for panmictic populations (equation 3 in Enquist et al. 2010; equation A20 in Lehmann et al. 2011), and thus holds more generally. By contrast, we were unable to find a simple expression for F_{ij} holding for all parameter values (see equation (A7)), so we examine special cases instead.

Differentiation under equal migration rate in both sexes

To gain intuition about the model, we first present the values of F_{ij} when the migration rate is the same in both sexes ($m_f = m_m$). In this case, the measure of cultural differentiation (equation (A7)) between groups takes the value

$$F_{ij} = \frac{\beta^2(1 - \mu)^2(1 - m)^2 p_s}{1 - \beta^2(1 - \mu)^2(1 - m)^2(1 - p_s)}, \tag{4}$$

where $p_s = p_f + p_m$, which is the same for all i and j; that is, differentiation between groups is the same regardless of the sexes of the pair of individuals under consideration.

Note that this fixation index behaves exactly as the standard measure of population differentiation. The cultural population structure increases with the probability of coalescence of cultural traits, p_s (the probability that two individuals have the same cultural parent) and the rate of social learning β, and decreases with the migration rate m and innovation rate μ. In this context, the rate of offspring growing to adulthood without learning socially ($1 - \beta$) has the same homogenizing effect on population structure as innovation or migration, but low β can result in very low trait abundance frequencies over the entire population (equation (3)). Thus, much like classical models in population genetics (Wright 1943; Maynard Smith 1970), population structure will be most strongly affected by changes in whichever of the parameters m, μ, or $1 - \beta$ has the largest value. Therefore, when $m \gg \mu$ and $m \gg 1 - \beta$, genetic and cultural population structure should be very similar, in the absence of other differences in demography/transmission. Furthermore, when copying is perfect ($\beta = 1$; i.e. all individuals have an equal chance of being cultural parents), equation (4) reduces to the standard F_{ST} value for haploid individuals (or diploids with same dispersal rate among the sexes) in the island model of dispersal under the infinite allele model of mutation and with the coalescence probability among pairs of genes being given by p_s (e.g. Wright 1931; Rousset 2004).

Differentiation under sex-specific migration rate

When migration is sex specific ($m_f \neq m_m$), the model becomes more complicated, but the case that is of most interest to us is when one sex remains completely philopatric while the other migrates. If females are the philopatric sex, we set $m_f = 0$ and $m_m = m$, which greatly simplifies the F_{ij} values (equation (A7)). If we further assume that offspring copy only individuals of the philopatric sex, then we can set $p_m = 0$, $p_{mm} = 0$ and $p_{ff} = 1 - p_f$, whereby the sex-specific fixation indexes simplify to

$$F_{ff} = \frac{\beta^2(1 - \mu)^2 p_f}{1 - \beta^2(1 - \mu)^2\left(1 - p_f\right)}$$

$$F_{fm} = \frac{\beta^2(1 - \mu)^2(1 - m)p_f}{1 - \beta^2(1 - \mu)^2\left(1 - p_f\right)} \tag{5}$$

$$F_{mm} = \frac{\beta^2(1 - \mu)^2(1 - m)^2 p_f}{1 - \beta^2(1 - \mu)^2\left(1 - p_f\right)}.$$

These equations show that the migration rate affects population structure most strongly among males, that differentiation among females does not depend on migration, and that as the coalescence probability increases, the amount of population structure always increases. These effects are illustrated in Fig. 1, which shows the population structure that occurs when learning is exclusively from either males or females. Figure 1 also illustrates that population structure in females is independent of migration when learning is female specific, as transmission is completely decoupled from migration (Fig. 1a, solid line). In general, when males migrate there
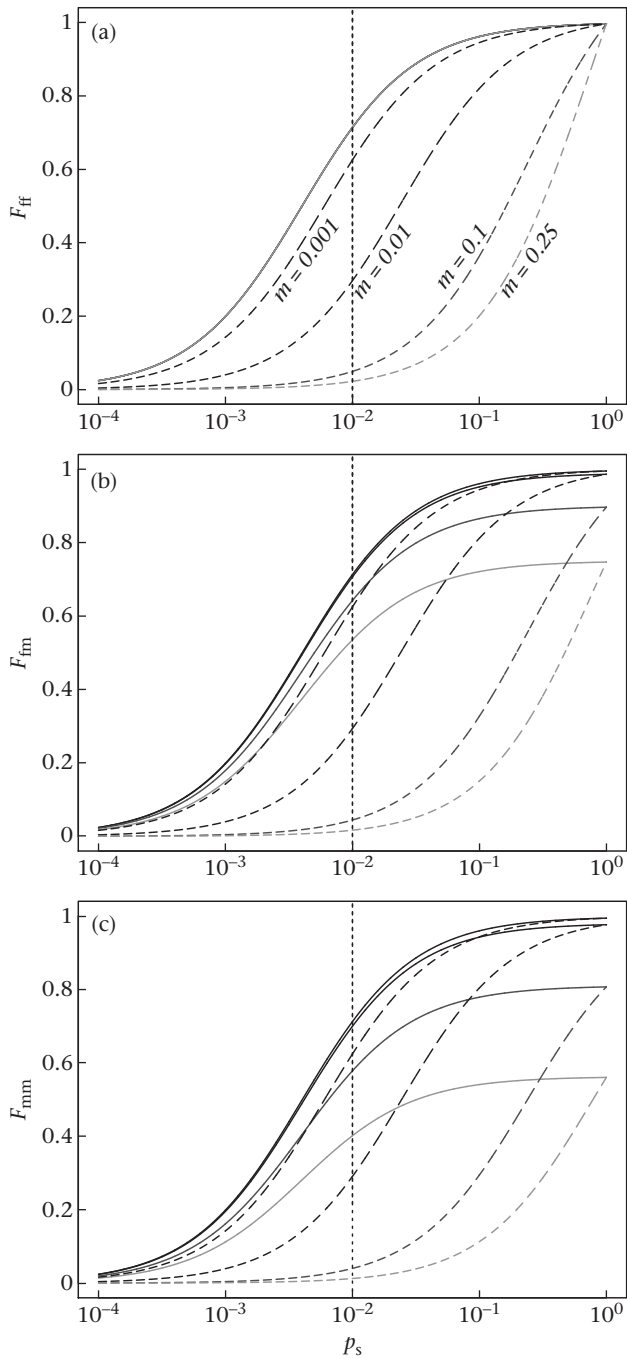
Figure 1. Increase in the sex-specific indexes of differentiation F_{ij} given by equation (A7) as a function of the probability p_s that two individuals have the same cultural parent: (a) for two females, (b) for a male and a female, and (c) for two males. The graphs depict cases in which only males migrate, and cultural parents are either exclusively female ($p_s = p_f$, $p_{ff} = 1 - p_f$, $p_m = 0$, $p_{mm} = 0$; solid lines, which depict equation (5)) or male ($p_s = p_m$, $p_{mm} = 1 - p_m$, $p_f = 0$, $p_{ff} = 0$; dashed lines), for a range of migration rates, as shown in (a). The dashed vertical line indicates the case where $N_f = N_m = 100$ and all individuals of the same sex have the same probability of being cultural parents. $\mu = 10^{-3}$, $\beta = 1 - \mu$.

is much less population structure when offspring learn from males (Fig. 1, dashed lines) than when they learn from females (Fig. 1, solid lines), and this is most evident at higher migration rates. The difference between population structures under male- versus female-biased learning is also most pronounced when the coalescence probability is significantly less than 1 (generally, when

$p_s \ll m$, see Fig. 1), as high probabilities of coalescence within groups result in structuring even under high m. The opposite patterns are expected if only females migrate.

Cultural inheritance from both sexes

The coalescence probability, p_s, can represent a range of social-learning rules, from purely random copying to highly sex-biased or one-to-many transmission (Cavalli-Sforza & Feldman 1981). To explore the sensitivity of the results of Fig. 1 to intermediate levels of sex-biased social learning when only males migrate, we set the probability that a random offspring learns from a female cultural parent as α, such that with complementary probability $1 - \alpha$ it learns from a male. With this, $p_f = \alpha^2/N_f$, $p_m = (1 - \alpha)^2/N_m$, $p_{ff} = \alpha^2(N_f - 1)/N_f$, and $p_{mm} = (1 - \alpha)^2(N_m - 1)/N_m$ (assuming all individuals of the same sex are equally likely to be cultural parents). Substituting these expressions into equation (A7) and then evaluating the average F_{ST} according to equation (2), we find that the high population structuring that occurs when males migrate and offspring learn from females alone (Fig. 1) decreases with decreasing values of α (Fig. 2a).

As a point of comparison, we also show levels of population structure in genetic traits under the same migration schemes but with random mating within groups and Poisson-distributed offspring number [i.e. the standard Wright−Fisher model of population genetics, Wright 1931; Hartl & Clark 2007, which results in coalescence probability of $1/(2N_i)$ in individuals of sex i]. Even in this scenario, under which the models are as similar as possible, there remain three factors that can cause differences in F_{ST}. When $\alpha \to 0.5$, the noncoalescence probabilities (p_{ff} and p_{mm}) are exactly the same as in the diploid genetic model (Gandon 1999), but p_s is still lower in diploids because there are twice as many gene copies as cultural trait copies segregating in the population. Also, when $\alpha = 0.5$, p_s would not be exactly twice as low in the genetic model because inbreeding occurs within individuals, which raises p_s. Finally, mutation in the genetic model involves transition probabilities that are slightly different from the transition probabilities due to innovation and social learning in the cultural model.

When there are opposite-sex biases in learning and migration (i.e. when $\alpha > 0.5$ and females tend to be philopatric), population structure in cultural traits is markedly increased relative to the genetic model (Fig. 2). When all learning is from females, $\alpha = 1$, there are pronounced differences between population structures in the cultural and genetic traits. Furthermore, under the presence−absence model, the rate of innovation (μ) plays a critical role in determining the amount of structuring seen in the cultural traits, much like mutation in genetic traits. When innovation rates are high relative to migration ($\mu \gg m$), there is very little population structure, both because a high rate of innovation reduces the similarity within groups, and because the same variant is commonly being invented de novo in each group (Fig. 2b). This shows that even when social learning is still relatively high ($\beta = 1 - \mu = 0.9$), individual learning can swamp the effect of migration on population structure. Results from individual-based simulations (points) agree well with analytical predictions across the range of parameter space explored in Fig. 2 (see Appendix 2 for simulation details and for a test of the sensitivity of the analytical results to the infinite-groups assumption).

k-cultural Variant Model with Perfect Transmission

Trait abundance frequency

We now consider the model with k variants and perfect transmission ($\beta = 1$). Here, individuals carry the focal trait with certainty
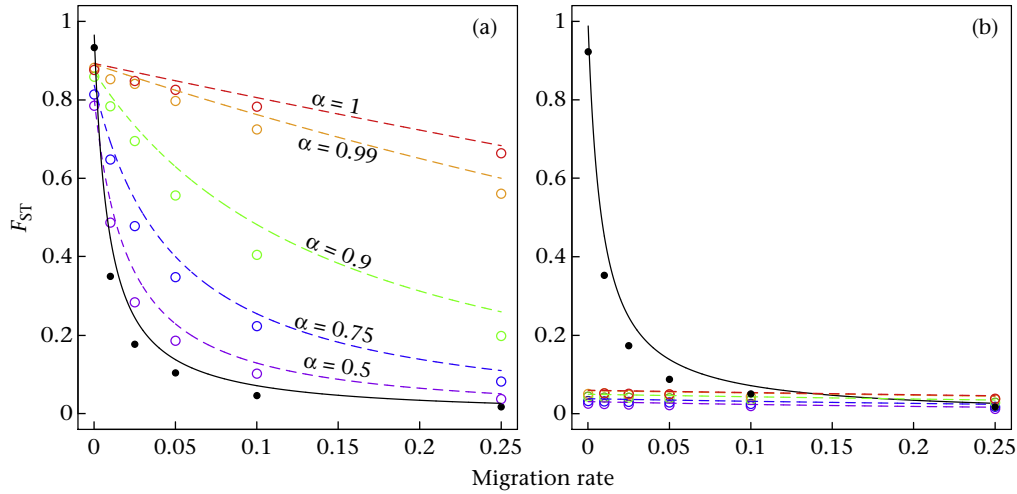
Figure 2. Population structure (F_{ST}) in cultural (dashed lines; open circles) and genetic traits (solid lines; filled circles) under low (a; $\mu = 0.001$) and high (b; $\mu = 0.1$) rates of cultural innovation with different probabilities of learning from females (α), when only males migrate ($m_f = 0$); that is, we used the average F_{ST} given by equation (2) over the F_{ij}'s given by equation (A7) for the cultural model and the same average of the sex-specific identities in a diploid genetic model (e.g. Gandon 1999). Lines show analytical results, points show simulation results for a population with a finite number n_d of groups. In all cases, the rate of social learning is set to $\beta = 1 - \mu$ so that the mean frequency of the cultural trait is $\rho \approx 0.5$. Parameters for individual-based simulations are: genetic mutation rate $v_G = 10^{-4}$, number of groups $n_d = 10$, $N_m = N_f = 30$, number of cultural and genetic loci $l_C = l_G = 5$. Genetic loci are diallelic, whereas the presence–absence model is a two-state model; we note that, while similar, the transition probabilities due to mutation/innovation in these two models are not exactly equivalent.

but different individuals may express different variants of that trait. When the cultural dynamics are at a steady state, we find that the probability that a single randomly sampled individual from the population carries a focal variant of the cultural trait is

$$\rho = \frac{v}{k}, \tag{6}$$

which is the ratio of the variant innovation rate to the total number of cultural variants of the focal cultural trait (see equation (A9)). This result is analogous to neutral models in population genetics (i.e. the k-allele model, Ewens 2004).

Differentiation under equal migration rate in both sexes

As above, we start by presenting F_{ij} values when the migration rate is the same in both sexes ($m_f = m_m$). In this case, we find from equations (A7) and (A10) that the measure of cultural differentiation between groups in the k-cultural variant model is

$$F_{ij} = \frac{(1 - vk/\{k-1\})^2(1-m)^2 p_s}{1 - (1 - vk/\{k-1\})^2(1-m)^2(1-p_s)}. \tag{7}$$

This is very similar to equation (4) except that β is equal to one and the innovation rate in equation (4) has been replaced by $vk/\{k-1\}$. If p_s is interpreted as the coalescence probability of pairs of genes, this equation is equivalent to the standard F_{ST} value for haploid individuals in the island model of dispersal under the k allele model of mutation (e.g. Crow & Aoki 1984; Rousset 2004). If $v = \mu$ the population structure predicted by equation (7) becomes equivalent to that predicted by equation (4) with $\beta = 1$ as k becomes large ($k \to \infty$). This implies that population structure decreases slightly with decreasing k, as this causes the probability of identity in state to increase.

Differentiation under sex-specific migration rate

We now consider sex-specific migration ($m_f \neq m_m$). As above, we assume that females are philopatric, and use $m_f = 0$ and $m_m = m$. Furthermore, when we got $p_{mm} = 0$, $p_f = p_s$ and $p_{ff} = 1 - p_f$, the sex-specific fixation indexes simplify to

$$F_{ff} = \frac{(1 - vk/\{k-1\})^2 p_f}{1 - (1 - vk/\{k-1\})^2 \left(1 - p_f\right)}$$

$$F_{fm} = \frac{(1 - vk/\{k-1\})^2(1-m)p_f}{1 - (1 - vk/\{k-1\})^2 \left(1 - p_f\right)} \tag{8}$$

$$F_{mm} = \frac{(1 - vk/\{k-1\})^2(1-m)^2 p_f}{1 - (1 - vk/\{k-1\})^2 \left(1 - p_f\right)}.$$

These equations show that the effects of migration and coalescence probability in the k-variant model are similar to the presence–absence model in equation (5), with differences arising from the 'innovation term' (γ in equation (A7)). To illustrate the effects of μ, v and β in these models, in Fig. 3 we show the fixation indexes for the k-variants model as a function of the coalescence probability p_f along with results for the presence–absence model under a range of rates of social learning. Generally speaking, the k variants and presence–absence models show similar qualitative sensitivity to the coalescence probability, as long as rates of social learning are high. The presence–absence model closely matches the k-variants model when k is large and $\beta = 1$ and when $k = 2$ and $\beta = 1 - \mu$ (Fig. 3). Because the analytical k-variants model does not allow for the absence of a trait, these models can differ considerably when the rate of social learning is low relative to the rates of individual learning ($1 - \beta \gg \mu$) and migration ($1 - \beta \gg m$), as this is the region of parameter space where rate of social learning exerts a relatively large effect (as in equation (4)).

k-cultural Variant Model with Imperfect Transmission

We now consider simulations in which each individual from the population carries either one of the k variants of the focal trait or does not carry the trait at all ($\beta < 1$ and $k > 1$), which thus combines the two models investigated above. Although we were unable to find a simple analytical way to combine these two models, simulations suggest that the amount of population structure arising under this situation is very similar to the predictions of
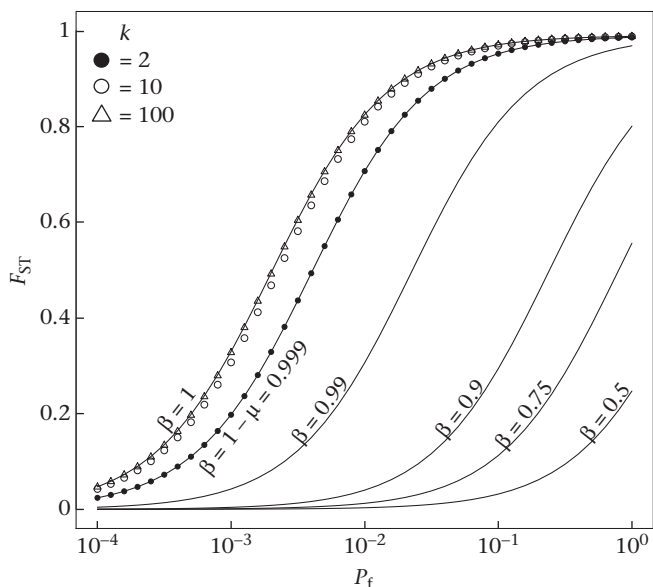
Figure 3. Analytical predictions for F_{ST} in cultural traits as a function of p_f with male migration, complete female philopatry and female-biased learning under the presence–absence model for six values of β (lines) and under the k-variants model for three values of k (points). The presence–absence model closely matches the k-variants model when β = 1 and k is large (triangles) and when β = 1 − μ and k = 2 (dots). When the rate of social learning is low, however, the presence–absence model predicts much less population structuring, with pronounced effects of social learning when $1 − β \gg μ$ and $1 − β \gg m$. Measures of F_{ST} are averaged over sex-specific structuring indexes from equation (5) for the presence–absence model and equation (8) for the k-variant model; m = 0.01 and μ = 0.001.

Figure 4. F_{ST} in cultural traits as a function of male migration under the analytical presence–absence model (solid lines without points), analytical k-variants model (points without lines) and in simulations combining presence–absence and k-variants models (points with dashed lines). For the presence–absence model and the simulations, the rates of social learning and innovation are scaled such that μ = ν = 1 − β, so that approximately half of the individuals carry cultural traits. For the analytical k-variants model, in all cases β = 1, but the values of μ and ν correspond to those used for the simulations. In all cases, only males migrate and social learning is modelled on females. Measures of F_{ST} are averaged over sex-specific structuring indexes from equation (5) for the presence–absence model and equation (8) for the k-variant model, with p_f = 1/N_f, N_m = N_f = 30. Other parameters are as in Fig. 2, but n_d = 50.
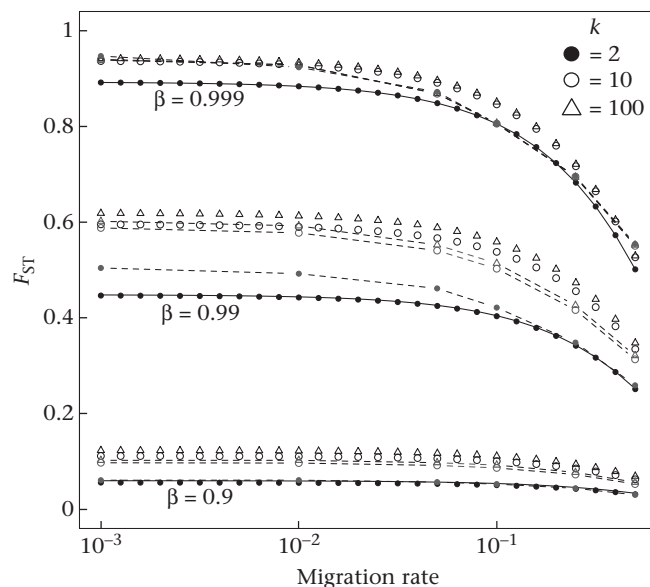
the other models (Fig. 4). As described above, there is slightly higher population structure under higher k, because of the reduced probability of identity in state due to innovation of the same variant. As $k → ∞$, the probability of identity in state due to innovation of the same variant tends to zero, but population structure can still be low if high rates of innovation ($ν \gg m$) reduce the similarity within groups.

DISCUSSION

Comparisons with Population Genetic Theory

In many respects, our results for neutral cultural traits under linear frequency-dependent updating are similar to the qualitative patterns of population structure expected from population genetic theory developed for neutral loci, but they also illustrate the importance of nuances particular to cultural transmission. For example, under the extreme case when learning is exclusively modelled on females and only males migrate (m_f = 0, p_m = 0, p_{mm} = 0), cultural F_{ST} (or cultural relatedness) behaves much the same as genetic F_{ST} for maternally inherited mitochondrial markers (Birky et al. 1983; Chesser & Baker 1996; Prugnolle & de Meeus 2002), as high levels of population structure can occur even under high male migration rates (Fig. 2a). Our approach also derives analytical predictions of population structure for intermediate levels of sex-biased inheritance (social learning), which is one scenario that is particular to cultural inheritance, with no clear parallel in genetics.

As another example of similarities between genetic and cultural models of population structure, innovation (individual learning) is akin to mutation in genetic models, introducing novelty, increasing within-group variance, and thereby decreasing F_{ST}. But whereas loci with high mutation rates are sometimes considered as an

empirical nuisance in genetic studies (e.g. Whitlock 2011), high rates of innovation may be of primary empirical interest in the study of culture. Rates of social learning may also have an effect similar to that of mutation, but in the opposite way to individual learning, as F_{ST} decreases with decreasing rates of social learning (i.e. with increasing values of 1 − β). Even very low rates of individual learning (μ) or of growth to adulthood without social learning (1 − β) can greatly reduce population structure when either quantity is higher than the migration rate (m), which may be very common in some species. On the other hand, whenever $m \gg μ$ and $m \gg 1 − β$, there should be little difference between cultural and genetic F_{ST} due to these aspects of transmission. Again, while these patterns can be understood by analogy to the effects of mutation in population-genetic models, this formulation explicitly shows the considerable impact of learning rates on cultural population structure.

Cultural variation among individuals may be quantified in slightly different ways, whether traits are recorded as present–absent or as a range of variations on a theme (k-variants model). Under these two innovation schemes, the population structure has the same qualitative sensitivity to parameter values as long as the rate of social learning is high in the presence–absence model (Fig. 3). This is akin to standard results in population genetics, where the particular nature of the mutation scheme does not qualitatively affect F_{ST} values (Tachida 1985; Rousset 2004). As social-learning rates decrease in the presence–absence model, the similarity among individuals within groups decreases and the relative importance of the innovation rate increases. Because all innovating individuals invent the same variants in the presence–absence model, the identity in state due to innovation results in higher similarity among groups and lower measures of population structure (Figs 3, 4). Population structure is

slightly higher when there are many potential variants, as individuals are less likely to innovate the same variant as k increases (Fig. 4).

When there is a high probability of coalescence (owing to high variance in fecundity for genetic traits or one-to-many transmission for cultural traits), both genetic and cultural models predict high F_{ST} values, even under high migration rates. While we do not explicitly model the factors shaping coalescence probabilities, these could be added to incorporate features of transmission particular to culture. As there is considerable interest in the effect of social networks on cultural transmission (Newman et al. 2004), it would be interesting to derive coalescence probabilities for social networks nested within each group of a population. All else being equal, networks with scale-free architectures and one-to-many motifs would have much higher coalescence probabilities (and higher F_{ST}) than networks with more diffuse architectures.

Empirical Applications

Depending on the ways in which the demography of a species affects cultural versus genetic transmission, we may expect very similar or very different levels of population structure in genetic loci versus cultural traits. For example, we have shown here how male migration and female-biased learning can cause very different patterns of population structure in genetic loci versus cultural traits (Fig. 2a), but very similar patterns could also be expected if the rate of individual learning was on the same order as the rate of male migration (Fig. 2b). For any direct comparisons of F_{ST} between diploid genetic loci versus cultural traits, it is necessary to account for the effect of the two-fold difference in effective population size; as there are twice as many gene copies as cultural trait copies in the population, F_{ST} at diploid genetic loci can be up to two-fold lower, all else being equal. It is also important to stress that high variance in fecundity, as might occur under extreme polygamy, can also lead to very high levels of genetic F_{ST}, even under high migration rates. In this case, cultural F_{ST} could be lower than genetic F_{ST}, even if females are cultural role models and only males migrate. Our comparisons of genetic versus cultural population structure (Fig. 2)

assume a genetic model with random mating and Poisson-distributed offspring number, but in many social animals this may not be the case. In the absence of considerable additional evidence about patterns of migration, mating systems, parental care and possible biases in social learning (all of which may be difficult to estimate with accuracy), it may sometimes be difficult to use comparisons of F_{ST} in genetic loci versus behavioural traits to make inferences about unknown aspects of demography and the basis of inheritance for the behaviours of interest. We now discuss several potential applications of F_{ST} and their promises and pitfalls.

It seems that the most basic question, whether variation in a trait is genetically or culturally inherited, may actually be the most difficult to address using inference from F_{ST}. Similarity between F_{ST} for genetic loci versus behavioural traits is expected even if the trait is culturally inherited under many demographic scenarios (as suggested by Langergraber et al. 2011). On the other hand, even if pronounced differences in genetic versus putative cultural F_{ST} are found, this could be explained by natural selection affecting a genetically determined behavioural trait, either through spatially homogeneous selection (low F_{ST}) or heterogeneous selection (high F_{ST}), as discussed with respect to Q_{ST} in quantitative genetic traits (Spitze 1993; Whitlock 2008). Various forms of selection can operate on culturally inherited traits, either through effects on biological fitness (natural selection) or through differences between variants in how likely they are to be learned and transmitted among individuals (cultural selection). Both cultural and natural selection can either decrease or increase F_{ST} in the cultural traits, depending on whether the same or different variants are favoured in different groups. Coupling comparisons of population structure in neutral markers versus behavioural traits (e.g. Langergraber et al. 2011) with studies of covariance in kinship and behavioural variants (e.g. Krützen et al. 2005) could provide a more powerful method to differentiate between an innate genetic basis and social learning.

As an example of a pattern of population structure in vervet monkeys that suggests a role for cultural inheritance, very little genetic population structure is observed among neighbouring social groups, even at maternally inherited mitochondrial markers,

Table 2
Explanations consistent with observed patterns

Possible causal factors	Observed patterns in F_{ST}		
	$F_{ST}[trait] \gg F_{ST}[genetic]$	$F_{ST}[trait] \ll F_{ST}[genetic]$	No significant difference in $F_{ST}[trait]$ vs $F_{ST}[genetic]$
Difference in copy number	Two-fold higher copy number in diploid genetic loci vs behavioural traits can cause up to two-fold lower $F_{ST}[genetic]$	No clear prediction	No clear prediction
Biases in migration and learning	Sex-biased migration, opposite sex-biased learning	Higher cultural migration rates (oblique or horizontal learning from other groups without mating and gene flow)	No consistent opposite-sex biases in learning/migration
Differences in coalescence probability between behavioural traits and genetic loci	One-to-many cultural transmission coupled with low variance in fecundity among individuals	Polygamy or polyandry coupled with random learning or biased learning modelled on the less promiscuous parent	No significant differences between mating system and cultural transmission system
Innovation rate	No clear prediction	High cultural innovation rate relative to the migration rate $\mu \gg m$ or $v \gg m$	No clear prediction
Social-learning rate	No clear prediction	Low rate of social learning relative to migration rate $(1-\beta) \gg m$. Note that if $(1-\beta) \gg \mu$, the trait segregates at low frequency [equation (1)]	No clear prediction
Selection	Spatially heterogeneous natural or cultural selection on the focal trait (different trait variants favoured in each group)	Spatially homogeneous natural or cultural selection on the focal trait (the same trait variant favoured in each group)	No strong or consistent effects of selection on focal trait (neutral or nearly neutral)

Many different learning and demographic scenarios can lead to similar patterns in comparisons of F_{ST} for genetic loci ($F_{ST}[genetic]$) versus behavioural traits ($F_{ST}[trait]$). This table lists the various explanations that are consistent with a given observed pattern, assuming that $F_{ST}[genetic]$ is estimated using genetic markers that are effectively neutral, with mutation rates that are lower than the migration rate the innovation rate and $(1 - \beta)$.

but there are pronounced differences in behaviour between groups (van de Waal et al. 2010; unpublished data). Migration between social groups is mainly by males and most care of offspring is provided by mothers and females within the social group. In this case, patterns of population structure in the behavioural traits are consistent with sex-biased cultural inheritance and migration and a high rate of social learning, but it is not possible to rule out some effect of selection. By contrast, chimpanzees have female-biased dispersal and much lower levels of differentiation in behaviour among groups separated by much greater distances (Whiten et al. 1999; Langergraber et al. 2011). If the behaviours studied in chimpanzees and vervets are in fact neutrally evolving cultural traits (and this remains an open question), both of these observations are consistent with expectations based on the differences in sex-biased migration. However, natural selection on genes for behaviour could also explain observed patterns of differentiation, as spatially heterogeneous selection increases population structure, whereas globally homogeneous selection decreases it, relative to neutrality (Whitlock 2008). Such cases will thus often require other lines of evidence in addition to comparisons of F_{ST} to prove that a given behaviour is culturally inherited.

It seems likely that analysing F_{ST} for behavioural traits could provide a powerful approach to analysing the importance of innovation (a component of individual learning). As considerable theory has now investigated the conditions that should favour individual versus social learning and showed that very often these two processes are likely to be jointly selected for (e.g. Rogers 1988; Richerson & Boyd 2005; Wakano & Aoki 2006; Borenstein et al. 2008; Rendell et al. 2010), it would be helpful to have a tool that could be used to study the consequences of these processes on the distribution of cultural traits within and between groups. As described above, high rates of innovation should greatly reduce population structure in most cases, relative to F_{ST} at genetic loci, except when the cultural coalescence probability (p_s) is very high. If the rate of innovation is high for a trait with several possible variants, then comparisons of groups that are geographically distant should reveal high genetic F_{ST} but low cultural F_{ST}. Failure to observe such a pattern would indicate that the rate of innovation is either relatively low or is effectively opposed by spatially homogeneous natural selection (or a cultural analogue) or extreme one-to-many social transmission (and high p_s). Again, interpreting any observed patterns in F_{ST} will generally require additional information about the biological and social life of the study species. In humans, understanding how the processes of innovation and transmission shape variation in cultural traits is of particular interest, and has been studied in examples ranging from pottery production in Cameroon (Wallaert-Pêtre 2001) and South Africa (Fowler 2008) to Turkmen textiles (Tehrani & Collard 2002). Analysing patterns of F_{ST} in these and other empirical cases may yield novel insights about the importance of innovation versus social learning. Table 2 shows a summary of the various possible explanations described above that are consistent with a given observation of F_{ST} for genetic loci versus behavioural traits. We note that as several mechanisms can have opposite influences on F_{ST}, a given observed pattern will seldom be conclusive without considerable additional evidence to rule out possible competing explanations.

Thus far, we have focused primarily on how demography and learning affect genetic versus cultural F_{ST}, rather than identifying specific statistical tests and methodological approaches. In addition to direct comparisons of F_{ST} in genetic loci versus cultural traits, it may be possible to adopt approaches used in genetics to prioritize further research. For instance, comparisons of F_{ST} among genetic loci have been used to identify candidate loci that might have adaptive effects (i.e. outliers with high F_{ST}; Lewontin & Krakauer 1973; Beaumont & Nichols 1996), providing a simple screen to prioritize

further study. While there are many factors that can cause high F_{ST} in cultural traits, a similar outlier approach could be used to identify cultural traits where interesting demographic features may be affecting their diversity and distribution. An important consideration for any empirical applications is how to test statistically for differences in cultural versus genetic F_{ST}. This problem has been addressed for comparisons of F_{ST} versus Q_{ST} (population structure in quantitative genetic traits; Whitlock 2008; Whitlock & Guillaume 2009), although we have not explored how such tests might need to be modified when applied to F_{ST} for cultural traits. In whatever way F_{ST} is applied to empirical examples, it is critical to consider how alternative demographic scenarios might yield similar patterns. Many different proximate behaviours may lead to biases in transmission that either increase or decrease variance within groups, affecting F_{ST} in potentially unexpected ways. While using F_{ST} to make inferences about demographics should therefore always be accompanied by other lines of evidence, it can be a powerful addition to the study of behavioural and cultural variation, as originally suggested by Cavalli-Sforza & Feldman (1981).

Acknowledgments

References

Allison, P. D. 1991. Cultural relatedness under oblique and horizontal transmission. *Ethology and Sociobiology*, **13**, 153–169.

Baker, C., Medrano-Gonzalez, L., Calambokidis, J., Perry, A., Pichler, F., Rosenbaum, H., Straley, J., Urban-Ramirez, J., Yamaguchi, M. & Von Ziegesar, O. 1998. Population structure of nuclear and mitochondrial DNA variation among humpback whales in the North Pacific. *Molecular Ecology*, **7**, 695–707.

Beaumont, M. & Nichols, R. 1996. Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society B*, **263**, 1619–1626.

Bell, A. V., Richerson, P. J. & McElreath, R. 2009. Culture rather than genes provides greater scope for the evolution of large-scale human prosociality. *Proceedings of the National Academy of Sciences, U.S.A.*, **106**, 17671–17674.

Bentley, R. A., Hahn, M. H. & Shennan, S. J. 2004. Random drift and culture change. *Proceedings of the Royal Society B*, **271**, 1443–1450.

Birky, C., Maruyama, T. & Fuerst, P. 1983. An approach to population and evolutionary genetic theory for genes in mitochondria and chloroplasts, and some results. *Genetics*, **103**, 513–527.

Borenstein, E., Aoki, K. & Feldman, M. W. 2008. Evolution of learning in fluctuating environments: when selection favors both social and exploratory individuals. *Evolution*, **62**, 586–602.

Boyd, R. & Richerson, P. J. 1982. Cultural transmission and the evolution of cooperative behavior. *Human Ecology*, **10**, 325–351.

Boyd, R. & Richerson, P. J. 1985. *Culture and the Evolutionary Process*. Chicago: University of Chicago Press.

Cavalli-Sforza, L. & Feldman, M. W. 1973. Models for cultural inheritance 1. Group mean and within group variation. *Theoretical Population Biology*, **4**, 42–55.

Cavalli-Sforza, L. & Feldman, M. W. 1981. *Cultural Transmission and Evolution*. Princeton, New Jersey: Princeton University Press.

Chesser, R. K. & Baker, R. J. 1996. Effective sizes and dynamics of uniparentally and diparentally inherited genes. *Genetics*, **144**, 1225–1235.

Cockerham, C. C. & Weir, B. S. 1987. Correlations, descent measures: drift with migration and mutation. *Proceedings of the National Academy of Sciences, U.S.A.*, **84**, 8512–8514.

Crow, J. F. & Aoki, K. 1984. Group selection for a polygenic behavioral trait: estimating the degree of population subdivision. *Proceedings of the National Academy of Sciences, U.S.A.*, **81**, 6073–6077.

Danchin, E. & Wagner, R. H. 2010. Inclusive heritability: combining genetic and non-genetic information to study animal behavior and culture. *Oikos*, **119**, 210–218.

Dunbar, R. I. M. 1988. *Primate Social Systems*. Ithaca, New York: Cornell University Press.

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

Eriksson, J., Siedel, H., Lukas, D., Kayser, M., Erler, A., Hashimoto, C., Hohmann, G., Boesch, C. & Vigilant, L. 2006. Y-chromosome analysis confirms highly sex-biased dispersal and suggests a low male effective population size in bonobos (*Pan paniscus*). *Molecular Ecology*, **15**, 939–949.

Ewens, W. J. 2004. *Mathematical Population Genetics*. New York: Springer-Verlag.

Fowler, K. 2008. Zulu pottery production in the Lower Thukela Basin, KwaZulu-Natal, South Africa. *South African Humanities*, **20**, 477–511.

Gandon, S. 1999. Kin competition, the cost of inbreeding and the evolution of dispersal. *Journal of Theoretical Biology*, **200**, 345–364.

Hartl, D. & Clark, A. G. 2007. *Principles of Population Genetics*. 4th edn. Cambridge, Massachusetts: Sinauer.

Jaeggi, A. V., Dunkel, L. P., Van Noordwijk, M. A., Wich, S. A., Sura, A. A. L. & Van Schaik, C. P. 2010. Social learning of diet and foraging skills by wild immature Bornean orangutans: implications for culture. *American Journal of Primatology*, **72**, 62–71.

Kandler, A. & Laland, K. N. 2009. An investigation of the relationship between innovation and cultural diversity. *Theoretical Population Biology*, **76**, 59–67.

Krützen, M., Mann, J., Heithaus, M. R., Connor, R. C., Bejder, L. & Sherwin, W. B. 2005. Cultural transmission of tool use in bottlenose dolphins. *Proceedings of the National Academy of Sciences, U.S.A.*, **102**, 8939–8943.

Laland, K. N. & Janik, V. M. 2006. The animal cultures debate. *Trends in Ecology & Evolution*, **21**, 542–547.

Langergraber, K. E., Siedel, H., Mitani, J. C., Wrangham, R. W., Reynolds, V., Hunt, K. & Vigilant, L. 2007. The genetic signature of sex-biased migration in patrilocal chimpanzees and humans. *PloS One*, **2**, e973, doi:10.1371/journal.pone.0000973.

Langergraber, K., Boesch, C., Inoue, E., Inoue-Maruyama, M., Mitani, J., Nishida, T., Pusey, A., Reynolds, V., Schubert, G. & Wrangham, R., et al. 2011. Genetic and 'cultural' similarity in wild chimpanzees. *Proceedings of the Royal Society B*, **278**, 408–416.

Lehmann, L., Feldman, M. W. & Foster, K. R. 2008. Cultural transmission can inhibit the evolution of altruistic helping. *American Naturalist*, **172**, 12–24.

Lehmann, L., Aoki, K. & Feldman, M. W. 2011. On the number of independent cultural traits carried by individuals and populations. *Philosophical Transactions of the Royal Society B*, **366**, 424–435.

Lewontin, R. C. & Krakauer, J. 1973. Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics*, **74**, 175–195.

Lumsden, C. J. & Wilson, E. O. 1980. Translation of epigenetic rules of individual behavior into ethnographic patterns. *Proceedings of the National Academy of Sciences, U.S.A.*, **77**, 4382–4386.

Lycett, S. J., Collard, M. & McGrew, W. C. 2007. Phylogenetic analyses of behavior support existence of culture among wild chimpanzees. *Proceedings of the National Academy of Sciences, U.S.A*, **104**, 17588–17592.

Lycett, S. J., Collard, M. & McGrew, W. C. 2010. Are behavioral differences among wild chimpanzee communities genetic or cultural? An assessment using tool-use data and phylogenetic methods. *American Journal of Physical Anthropology*, **142**, 461–467.

Lynch, A. & Baker, A. 1994. A population memetics approach to cultural evolution in chaffinch song: differentiation among populations. *Evolution*, **48**, 351–359.

Manel, S., Schwartz, M., Luikart, G. & Taberlet, P. 2003. Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology & Evolution*, **18**, 189–197.

Maynard Smith, J. 1970. Population size, polymorphism, and the rate of non-Darwinian evolution. *American Naturalist*, **104**, 231–237.

Natoli, A., Birkun, A., Aguilar, A., Lopez, A. & Hoelzel, A. 2005. Habitat structure and the dispersal of male and female bottlenose dolphins (*Tursiops truncatus*). *Proceedings of the Royal Society B*, **272**, 1217–1226.

Newman, M., Barabási, A.-L. & Watts, D. 2004. *The Structure and Dynamics of Networks*. Princeton, New Jersey: Princeton University Press.

Prugnolle, F. & de Meeus, T. 2002. Inferring sex-biased dispersal from population genetic tools: a review. *Heredity*, **88**, 161–165.

Rendell, L., Boyd, R., Cownden, D., Enquist, M., Eriksson, K., Feldman, M. W., Fogarty, L., Ghirlanda, S., Lillicrap, T. & Laland, K. N. 2010. Why copy others? insights from the social learning strategies tournament. *Science*, **328**, 208–213.

Richerson, P. J. & Boyd, R. 2005. *Not by Genes Alone*. Chicago: University of Chicago Press.

Rogers, A. R. 1988. Does biology constrain culture? *American Anthropologist*, **90**, 819–831.

Rousset, F. 1996. Equilibrium values of measures of population subdivision for stepwise mutation processes. *Genetics*, **142**, 1357–1362.

Rousset, F. 2004. *Genetic Structure and Selection in Subdivided Populations*. Princeton, New Jersey: Princeton University Press.

van Schaik, C. 1983. Why are diurnal primates living in groups. *Behaviour*, **87**, 120–144.

van Schaik, C., Ancrenaz, M., Borgen, G., Galdikas, B., Knott, C., Singleton, I., Suzuki, A., Utami, S. & Merrill, M. 2003. Orangutan cultures and the evolution of material culture. *Science*, **299**, 102–105.

Seielstad, M., Minch, E. & Cavalli-Sforza, L. 1998. Genetic evidence for a higher female migration rate in humans. *Nature Genetics*, **20**, 278–280.

Spitze, K. 1993. Population structure in daphnia obtusa: quantitative genetic and allozymic variation. *Genetics*, **135**, 367–374.

Tachida, H. 1985. Joint frequencies of alleles determined by separate formulations for the mating and mutation systems. *Genetics*, **111**, 963–974.

Takahata, N. 1983. Gene identity and genetic differentiation of populations in the finite island model. *Genetics*, **104**, 497–512.

Tehrani, J. & Collard, M. 2002. Investigating cultural evolution through biological phylogenetic analyses of Turkmen textiles. *Journal of Anthropological Archaeology*, **21**, 443–463.

de Waal, F. B. M. 2001. *The Ape and the Sushi Master: Cultural Reflections by a Primatologist*. New York: Basic Books.

van de Waal, E., Renevey, N., Favre, C. M. & Bshary, R. 2010. Selective attention to philopatric models causes directed social learning in wild vervet monkeys. *Proceedings of the Royal Society B*, **277**, 2105–2111.

Wakano, J. Y. & Aoki, K. 2006. A mixed strategy model for the emergence and intensification of social learning in a periodically changing natural environment. *Theoretical Population Biology*, **70**, 486–497.

Wakano, J. Y., Aoki, K. & Feldman, M. W. 2004. Evolution of social learning: a mathematical analysis. *Theoretical Population Biology*, **66**, 249–258.

Wallaert-Pêtre, H. 2001. Learning how to make the right pots: apprenticeship strategies and material culture, a case study in handmade pottery from Cameroon. *Journal of Anthropological Research*, **57**, 471–493.

Weir, B. S. & Cockerham, C. C. 1984. Estimating F-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

Whiten, A., Goodall, J., McGrew, W., Nishida, T., Reynolds, V., Sugiyama, Y., Tutin, C., Wrangham, R. & Boesch, C. 1999. Cultures in chimpanzees. *Nature*, **399**, 682–685.

Whiten, A., Hinde, R., Laland, K. L. & Stringer, C. 2011. Culture evolves. *Philosophical Transactions of the Royal Society B*, **366**, 938–948.

Whitlock, M. 1992. Temporal fluctuations in demographic parameters and the genetic variance among populations. *Evolution*, **46**, 608–615.

Whitlock, M. C. 2008. Evolutionary inference from Q(ST). *Molecular Ecology*, **17**, 1885–1896.

Whitlock, M. C. 2011. G'ST and d do not replace FST. *Molecular Ecology*, **20**, 1083–1091.

Whitlock, M. C. & Guillaume, F. 2009. Testing for spatially divergent selection: comparing Q(ST) to F-ST. *Genetics*, **183**, 1055–1063.

Whitlock, M. & McCauley, D. 1990. Some population genetic consequences of colony formation and extinction: genetic correlations within founding groups. *Evolution*, **44**, 1717–1724.

Wilder, J., Kingan, S., Mobasher, Z., Pilkington, M. & Hammer, M. 2004. Global patterns of human mitochondrial DNA and Y-chromosome structure are not influenced by higher migration rates of females versus males. *Nature Genetics*, **36**, 1122–1125.

Wolfram, S. 2003. *Mathematica*. 5th edn. Cambridge: Cambridge University Press.

Wrangham, R. 1980. An ecological model of female-bonded primate groups. *Behaviour*, **75**, 262–300.

Wright, S. 1931. Evolution in Mendelian populations. *Genetics*, **16**, 97–159.

Wright, S. 1943. Isolation by distance. *Genetics*, **28**, 114–138.

Wright, S. 1951. The genetical structure of populations. *Annals of Eugenics*, **15**, 323–354.

APPENDIX 1. ANALYTICS

Presence–absence Model

Recursions

To obtain recurrence equations describing the dynamics of the cultural state probabilities under the life cycle assumptions presented in the main text, we express these probabilities in a descendant generation at the adult stage (ρ', θ'_{ij}, ϕ'_{ij}) as a function of their values in the previous generation at that stage (ρ, θ_{ij}, ϕ'_{ij}) by decomposing the total change in the probabilities into the three different life cycle events that affect their dynamics that is, migration individual learning and social learning.

After migration and regulation, the cultural state probabilities are given by

$$\rho' = \rho^{**}$$
$$\theta'_{ij} = (1-m_i)(1-m_j)\theta^{**}_{ij} + \left[1 - (1-m_i)(1-m_j)\right](\rho^{**})^2$$
$$\phi'_{ij} = (1-m_i)(1-m_j)\phi^{**}_{ij} + \left[1 - (1-m_i)(1-m_j)\right]2\rho^{**}(1-\rho^{**}),$$
(A1)

where ρ^{**}, θ^{**}_{ij}, ϕ^{**}_{ij} are the cultural state probabilities before migration and right after individual learning. The first line of equation (A1). follows from the fact that migration does not affect the cultural state of a randomly sampled individual from the population. The second line of equation (A1) follows from the fact that if two randomly sampled individuals in a group are of philopatric origin, then the probability that they both carry the focal trait is θ^{**}_{ij}, while if at least one of the individuals is an immigrant, the two individuals are unlikely to have inherited the focal trait from the same cultural ancestor. Then, the ancestral lineages of the traits of the two individuals are independent and the two individuals carry the focal trait with probability $(\rho^{**})^2$. The third line of equation (A1) follows from similar considerations, but here if individuals carry independent trait lineages, the probability that one carries the trait and the other does not is $2\rho^{**}(1-\rho^{**})$.

Right after individual learning, the cultural state probabilities are given by

$$\begin{aligned}
\rho^{**} &= \rho^* + \mu(1 - \rho^*) \\
\theta_{ij}^{**} &= \theta_{ij}^* + \mu\phi_{ij}^* + \mu^2\left(1 - \theta_{ij}^* - \phi_{ij}^*\right) \\
\phi_{ij}^{**} &= (1 - \mu)\theta_{ij}^* + 2\mu(1 - \mu)\left(1 - \theta_{ij}^* - \phi_{ij}^*\right),
\end{aligned} \quad (A2)$$

where ρ^*, θ_{ij}^*, θ_{ij}^* are the cultural state probabilities right after social learning. The first line of equation (A2) can be understood by noting that if a focal individual carries the focal trait before individual learning, it keeps it with probability one, while if it does not carry that trait, it invents it with probability μ. The second line of equation (A2) follows from the fact that when, among two randomly sampled individuals, only one carries the focal trait before the stage of individual learning, they both carry it after individual learning with probability μ. On the other hand, if both individuals do not carry the trait before individual learning, they both carry it with probability μ^2 after individual learning. The third line of equation (A2) can be understood by noting that when, among two randomly chosen individuals, only one carries the focal trait before learning, this state is not changed if the individual lacking the trait did not invent it, while if both individuals do not carry the trait before individual learning, one carries after individual learning with probability $2\mu(1 - \mu)$.

We use the same notation as above to represent the probabilities that two individuals acquire traits from the same cultural parent (p_s, p_{ff} and p_{mm}), so that the cultural state probabilities right after social learning can be written as

$$\begin{aligned}
\rho^* &= \beta\rho \\
\theta_{ij}^* &= \beta^2\left[p_s\rho + p_{ff}\theta_{ff} + p_{mm}\theta_{mm} + \left(1 + p_s - p_{ff} - p_{mm}\right)\theta_{mf}\right] \\
\phi_{ij}^* &= p_s[2\beta(1 - \beta)\rho] + p_{ff}\left[2\beta(1 - \beta)\theta_{ff} + \beta\phi_{ff}\right] \\
&\quad + p_{mm}[2\beta(1 - \beta)\theta_{mm} + \beta\phi_{mm}] \\
&\quad + \left(1 - p_s - p_{ff} - p_{mm}\right)\left[2\beta(1 - \beta)\theta_{mf} + \beta\phi_{mf}\right],
\end{aligned} \quad (A3)$$

where ρ, θ_{ij}, ϕ_{ij} are the cultural state probabilities in the parental generations.

The first line of equation (A3) follows from the fact that a single focal individual carries the trait after social learning if its cultural parent carries that trait (probability ρ) and if the individual acquired it (probability β). The second line of equation (A3) follows from the fact that if two randomly sampled different offspring had the same cultural ancestor, the ancestor carried the trait with probability ρ, in which case both individuals acquire that trait with probability β^2. If the two offspring have different cultural ancestors, one of sex i and the other of sex j, then the probability that they both carry the trait is $\beta^2\theta_{ij}$. Finally, the third line of equation (A3) can be understood by noting that, among two offspring that have the same cultural parent that has the focal trait, only one offspring carries the focal trait after social learning with probability $2\beta(1 - \beta)$. If the two offspring have two different cultural ancestors, one of sex i and the other of sex j, then the probability that only one offspring carries the trait is $2\beta(1 - \beta)\theta_{ij} + \beta\theta_{ij}$.

Equilibrium Values

To obtain the equilibrium values of the cultural state probabilities, we substitute equation (A2) and (A3) into equation (A1), set $\rho' = \rho$, $\theta_{ij}' = \theta_{ij}$ and $\phi_{ij}' = \phi_{ij}$ and then solve for ρ, θ_{ij}, ϕ_{ij}. The resulting values then allow us to evaluate the sex-specific fixation index F_{ij} (equation (1)) explicitly. Regardless of the parameter values, we find that the probability that a single, randomly sampled individual from the population carries the focal trait is

$$\rho = \frac{\mu}{1 - \beta(1 - \mu)}. \quad (A4)$$

By contrast, the expressions for θ_{ij} and ϕ_{ij} are unwieldy when all model parameters can vary and so we investigated special cases (a Mathematica, Wolfram 2003; notebook with the full expression is available on request), which lead to the expressions of F_{ij} presented in the main text (equations (4) and (5)). However, as detailed in the next section, there is a simpler way to obtain directly the equilibrium values of F_{ij}.

Transition equations for F_{ij}

In the previous sections we used the cultural state probabilities (ρ, θ_{ij}, ϕ_{ij}) to evaluate the correlations F_{ij} at steady state of the cultural dynamics. Owing to the fact that in the presence–absence model, the only force that creates positive F_{ij} is identity-by-descent of cultural traits (when $p_s > 0$), we can directly write down transition equations for the F_{ij} as is usually done in population genetics (e.g. Wright 1951; Cockerham & Weir 1987; Rousset 2004), without the need to consider explicitly the dynamics of the state probabilities ρ, θ_{ij}, ϕ_{ij}. We now present these calculations, which are simpler than the previous ones and will also allow us to evaluate the F_{ij}'s under the k-variant model more simply.

First note that, as is usually done in population genetics for the infinite-island model of migration, we can interpret F_{ij} as the probability that two individuals randomly sampled from the same group at the adult stage (after migration and density-dependent competition), one of sex i and the other of sex j, have acquired the focal cultural trait from the same common ancestor. After migration and regulation, we can write this probability as

$$F_{ij}' = (1 - m_i)(1 - m_j)F^J, \quad (A5)$$

where $(1 - m_i)(1 - m_j)$ is the probability that the two individuals are of philopatric origin and F^J is the probability that two individuals sampled before migration (regardless of their sexes) carry the cultural trait inherited from the same common ancestor.

Taking into account both the effect of individual and social learning on F^J we have:

$$F^J = (1 - \mu)^2\beta^2\left[p_s + p_{ff}F_{ff} + \left(1 - p_s - p_{ff} - p_{mm}\right)F_{mf}\right. \\
\left. + p_{mm}F_{mm}\right]. \quad (A6)$$

This equation can be understood by noting that two offspring in a group before dispersal carry the same cultural trait inherited from a common cultural ancestor if they both have not innovated the focal trait and copied it from their cultural parents (probability $(1 - \mu)^2\beta^2$). When this event occurs, they have inherited the focal trait from the same cultural parent living in the previous generation with probability p_s, in which case they have a common ancestor with probability one. With probability p_{ff} (p_{mm}), the two offspring copy the focal trait from two distinct females (males), in which case the probability that the two offspring carry a trait from a common cultural ancestor is given by the value of the correlation in the past generation that is, F_{ff} (F_{mm}). Finally, one offspring may have inherited the trait from a male and the other from a female, in which case they carry the trait from a common ancestor with probability F_{mf}.

To obtain the equilibrium values of the F_{ij} values, we substitute equation (A6) into equation (A5), set $F_{ij}' = F_{ij}$, and solve for the system of equations for F_{ff}, F_{mm} and F_{mf}, which produces

$$F_{\mathrm{ff}} = \frac{\gamma\left(1 - m_{\mathrm{f}}\right)^2 \left(p_{\mathrm{f}} + p_{\mathrm{m}}\right)}{1 - \gamma\left[\left(1 - m_{\mathrm{f}}\right)^2 p_{\mathrm{ff}} + \left(1 - m_{\mathrm{m}}\right)\left(1 - m_{\mathrm{f}}\right)\left(1 - p_{\mathrm{s}} - p_{\mathrm{ff}} - p_{\mathrm{mm}}\right) + \left(1 - m_{\mathrm{m}}\right)^2 p_{\mathrm{mm}}\right]}$$

$$F_{\mathrm{mf}} = \frac{\gamma(1 - m_m)\left(1 - m_{\mathrm{f}}\right)\left(p_f + p_m\right)}{1 - \gamma\left[\left(1 - m_{\mathrm{f}}\right)^2 p_{\mathrm{ff}} + \left(1 - m_{\mathrm{m}}\right)\left(1 - m_{\mathrm{f}}\right)\left(1 - p_{\mathrm{s}} - p_{\mathrm{ff}} - p_{\mathrm{mm}}\right) + \left(1 - m_{\mathrm{m}}\right)^2 p_{\mathrm{mm}}\right]} \tag{A7}$$

$$F_{\mathrm{mm}} = \frac{\gamma(1 - m_{\mathrm{m}})^2 \left(p_{\mathrm{f}} + p_{\mathrm{m}}\right)}{1 - \gamma\left[\left(1 - m_{\mathrm{f}}\right)^2 p_{\mathrm{ff}} + \left(1 - m_{\mathrm{m}}\right)\left(1 - m_{\mathrm{f}}\right)\left(1 - p_{\mathrm{s}} - p_{\mathrm{ff}} - p_{\mathrm{mm}}\right) + \left(1 - m_{\mathrm{m}}\right)^2 p_{\mathrm{mm}}\right]},$$

which holds for $0 < \rho < 1$, and where

$$\gamma = \beta^2 (1 - \mu)^2. \tag{A8}$$

Equation (A7) could also have been obtained by substituting equations (A3) and (A2) into equation (A1), solving for ρ, θ_{ij}, θ_{ij} and then substituting into equation (1) of the main text.

When $m_{\mathrm{f}} = m_{\mathrm{m}}$ equation (A7) gives equation (A4) of the main text. When $p_{\mathrm{m}} = 0$, $p_{\mathrm{mm}} = 0$ and $p_{\mathrm{ff}} = 1 - p_{\mathrm{f}}$, equation (A7) produces equation (A5) of the main text.

k-variant model

For the case where individuals carry the focal trait with certainty but there are k cultural variants of this trait segregating in the population and each individual can switch its variant at rate ν per generation, the transition equation for the probability that a randomly sampled individual from the population carries a focal variant is

$$\rho' = \rho(1 - \nu) + (1 - \rho)\frac{\nu}{k - 1}, \tag{A9}$$

which, at equilibrium, gives $\rho = \nu/k$.

Fortunately, the equilibrium values of the F_{ij} values are still given by equation (A7) and we need only to adjust the value of γ. This follows from results in population genetics that show that when the mating and mutation systems are independent, the effect of mutation under the k-allele model on the equilibrium values of F statistics is summarized by the parameter γ (e.g. Tachida 1985; Rousset 1996). Here, this corresponds to the independence between the individual and social-learning systems, and the relevant value of γ for the k-allele mode is

$$\gamma = \left(1 - \frac{\nu k}{k - 1}\right)^2 \tag{A10}$$

(equation 6 in Rousset 1996).

Substituting equation (A10) into equation (A7) and setting $p_{\mathrm{m}} = 0$, $p_{\mathrm{mm}} = 0$ and $p_{\mathrm{ff}} = 1 - p_{\mathrm{f}}$ produces equation (A8) of the main text. This equation could alternatively be obtained by evaluating the cultural state probabilities (ρ, θ_{ij}, θ_{ij}), as was done above, but such calculations are much more tedious (a Mathematica notebook, Wolfram 2003, with these calculations is available on request).

APPENDIX 2. SIMULATIONS

For the individual-based simulations, the life cycle assumptions are the same as those for the analytical model. In addition, each individual has l_{G} neutrally evolving, unlinked, diploid diallelic genetic loci and can learn up to l_{C} cultural traits, which are either present or absent in each individual. When individuals are born, they inherit their genes from two randomly chosen parents and then undergo a period of enculturation, where for each of the l_{C} traits, they have an overall probability of learning the trait of β. For each cultural trait that individuals may learn, they inherit the trait from a randomly chosen female in the group with probability α or a randomly chosen male

with probability $1 - \alpha$. Individuals may also innovate (individually learn) traits that they have not learned socially, with probability μ. For the k-variants version of the simulations, individuals that successfully learn from their cultural parents change their trait to one of the other $k - 1$ variants with probability ν. Mutation between the two allelic states at each genetic locus occurs in the gametes at a rate ν_G.

Simulations are initialized with 10 groups of 60 individuals (30 of each sex) with randomly drawn genotypes and cultural repertoires (equal probabilities of the two genetic alleles and of having/ not having memes or of having one of k variants) and then run for 1000 generations, with 25 replicates per parameter set. Given the population size, this was sufficient time to reach equilibrium. Measures of population structure at the genetic and cultural loci are calculated using equation (1), and averaged over the three sex-pair combinations to yield the F_{ST} values, as in equation (2).

Infinite Number of Groups Assumption

Our analytical approach assumes an infinite number of groups to simplify the analysis. To test the sensitivity of this assumption, we ran individual-based simulations with different numbers of groups, n_{d} (Fig. A1). As the analytical model is based on the assumption of an infinite number of groups, the quantitative agreement between the analytical and simulated results was best under a large number of groups ($n_{\mathrm{d}} = 100$), but the qualitative patterns did not change substantially, even under a very small number of groups ($n_{\mathrm{d}} = 3$), which is a standard result for the differentiation of neutral genetic loci (Takahata 1983; Rousset 2004).
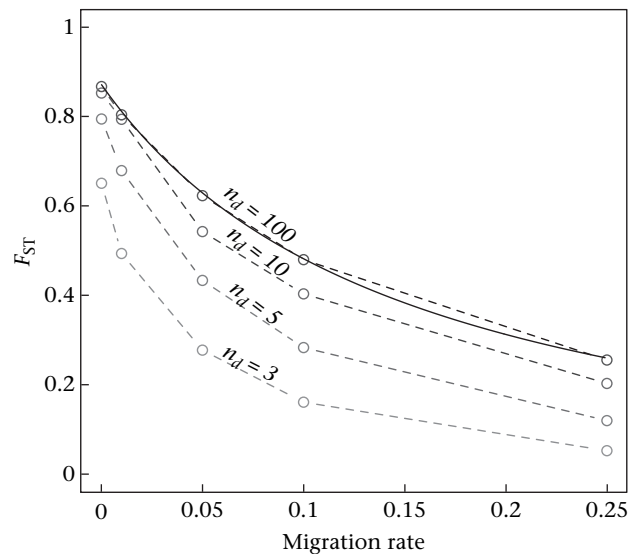
Figure A1. Population structure in cultural traits from individual-based simulations with different numbers of groups (n_{d}; dashed lines), compared to analytical predictions (solid line). The probability of social learning from a female cultural parent, $\alpha = 0.9$; all other statistics and parameters for simulations are as in Fig. 2.