

Transmission coupling mechanisms: cultural group selection

Robert Boyd^{1,*} and Peter J. Richerson²

¹*Department of Anthropology, University of California, Los Angeles, CA 90095, USA*

²*School of Environmental Science and Policy, University of California, Davis, CA 95616, USA*

The application of phylogenetic methods to cultural variation raises questions about how cultural adaptation works and how it is coupled to cultural transmission. Cultural group selection is of particular interest in this context because it depends on the same kinds of mechanisms that lead to tree-like patterns of cultural variation. Here, we review ideas about cultural group selection relevant to cultural phylogenetics. We discuss why group selection among multiple equilibria is not subject to the usual criticisms directed at group selection, why multiple equilibria are a common phenomena, and why selection among multiple equilibria is not likely to be an important force in genetic evolution. We also discuss three forms of group competition and the processes that cause populations to shift from one equilibrium to another and create a mutation-like process at the group level.

Keywords: cultural transmission; multi-level selection; cultural adaptation

1. INTRODUCTION

The application of phylogenetic methods to cultural variation has burgeoned over the past decade. As evidenced by the papers in this issue, this project has led to important new inferences about cultural history, human demography and human migrations. Much of this work has been done without reference to an explicit theory of cultural adaptation, simply applying phylogenetic statistical methods developed in biology to cultural data, and there are also important unanswered questions that might be illuminated by a better understanding of how cultural evolution works. For example, cultural phylogenies allow the application of new statistical methods developed in evolutionary biology (e.g. Huelsenbeck *et al.* 2001) to solve Galton's problem. This approach has proven very useful (e.g. Holden & Mace 2005), but it also raises important questions about the relationship between cultural phylogenies and cultural evolution. There is little gene flow between most biological species, and thus to a first approximation all of the genes within a species share a common history. This is definitely not true of cultural lineages—traits and trait complexes flow from one lineage to another, and thus different trait complexes may have different histories from each other and from the genes in the populations that carry them. As a consequence, sometimes it is not easy to know which phylogeny should be used to constrain phylogenetic inference. Linguistic phylogenies are often used, but, contrary to what is sometimes assumed, they are not histories of biological populations. Language phylogenies are appropriate for traits that are transmitted along with language, but are

not appropriate for traits that have different patterns of transmission. Thus, it is important to ask how different processes of cultural adaptation work, and in particular, focus on adaptive processes that cause different trait complexes to have similar cultural histories.

For many traits, cultural evolutionary processes can lead to many distinct steady-state outcomes, or 'multiple stable equilibria'. Which outcome is reached will then be determined by the accidents of initial conditions, and knowing the adaptive consequences of different traits does not allow us to predict the outcome. However, if such a population is subdivided into partially isolated subpopulations, adaptive processes can maintain different subpopulations near different equilibria. Then if subpopulations near one equilibrium have lower extinction rates or produce more migrants, the variants that characterize that equilibrium can spread to the population as a whole. This process is not subject to the usual criticisms directed at group selection for altruistic variants because adaptation within groups does not compete with selection among groups. It can work even if populations are very large, and migration rates are substantial. The main requirement is that rates of adaptation within groups are high when compared with rates of migration between them, and as a result this process is more likely to be important for cultural evolution than for genetic evolution. When these conditions are satisfied, group selection will lead to the spread of the most group-beneficial equilibrium.

In this paper, we review and discuss cultural group selection, distinguishing it from other group selection processes, discussing the processes that lead to multiple equilibria, the processes that select among equilibria and the random processes that give rise to new group-level variants. We believe and think that this process is especially relevant to cultural

* Author for correspondence (rboyd@anthro.ucla.edu).

One contribution of 14 to a Theme Issue 'Cultural and linguistic diversity: evolutionary approaches'.

phylogenetics because some of the same processes that lead to multiple equilibria tend to cause traits to be transmitted vertically within populations rather than horizontally among them.

2. GROUP SELECTION HAS MANY FACES

The modern group selection controversy began in the early 1960s when Wynne-Edwards (1962) proposed that a number of interesting bird behaviours evolved because they promoted group survival. Populations in which the behaviour was common survived and prospered, while those in which it was rare perished. While casual group functionalism was common in those days, Wynne-Edwards was much clearer than his contemporaries that it was selection among groups that gave rise to such group-level adaptations. The book generated a storm of controversy, and luminaries like Williams (1966) and Maynard Smith (1964) penned critiques explaining why this mechanism, then called group selection, was unlikely to be an important evolutionary process. Moreover, they also showed how such traits could evolve owing to individual and kin selection. The result was the beginning of an ongoing, and highly successful revolution in our understanding of the evolution of animal behaviour, a revolution that is rooted in carefully thinking about the individual and nepotistic function of behaviours.

In the early 1970s, Price (1970, 1972) developed a powerful new mathematical formalism that describes all natural selection as going on in a series of nested levels: among genes within an individual, among individuals within groups and among groups. While this ‘multi-level’ approach and the older gene-centred approaches are mathematically equivalent, both have proved useful in understanding many evolutionary problems. However, rise of the multi-level approach also led to confusion about what kinds of evolutionary processes should be called ‘group selection’. Some authors use group selection to mean the process that Wynne-Edwards envisioned—selection between sizable groups made up of mostly genealogically distantly related individuals, while others use group selection to refer to selection involving any kind of group in a multi-level selection analysis including even pairs of individuals interacting in, say, the hawk–dove game.

The real scientific question is always: does the population structure in question lead to selection that favours genetic variants of interest? In the case of the mechanism proposed by Wynne-Edwards, we want to know, can selection among large groups of distantly related individuals, sometimes labelled ‘interdemic group selection’, lead to the evolution of group-beneficial traits when it is opposed by individual selection? The answer to this question is fairly clear: only when groups are small or there is very little gene flow between them. To see why, it will be useful to introduce Price’s formalism. In a population structured into groups, the change in frequency of a gene undergoing selection, Δp , is given by

$$\Delta p \propto \underbrace{V_G \beta_G}_{\text{between groups}} + \underbrace{\overline{V_W \beta_W}}_{\text{within groups}}$$

The first term gives the change owing to selection between groups, and the second term gives the change in frequency owing to changes within groups. The β s give the effect of the behaviour on the fitness of groups (β_G) and individuals (β_W). A behaviour is beneficial to the group when it increases group fitness, or $\beta_G > 0$. If it is costly to the individual $\beta_W < 0$. The V s are the variance in gene frequency between groups (V_G) and within groups (V_W). Population genetics theory tells us that when groups are large, selection is weak, and there is even a modest amount of migration among them, the variance between individuals (V_W) will be much larger than the variance between groups (V_G , Rogers 1990). Thus, unless selection within groups is much weaker than selection among groups ($\beta_G \gg \beta_W$), group selection cannot overcome opposing individual selection.

3. INTERDEMIC GROUP SELECTION CAN BE IMPORTANT WHEN THERE ARE MULTIPLE STABLE EQUILIBRIA

This does not mean, however, that interdemic group selection is never important—it can play a crucial role in determining evolutionary outcomes when there are multiple stable equilibria. Interestingly, this idea dates to the early 1930s when the great population geneticist Sewall Wright (1931) first outlined his ‘shifting balance’ theory of evolution. Wright knew from his empirical work that interaction between genes often leads to evolutionary systems with multiple stable equilibria. The simplest case is underdominance at a single locus. Suppose there are two alleles, A and B , and that the fitnesses of the three genotypes are $W_{AA} = 1$, $W_{AB} = 1 - s$, and $W_{BB} = 1 + t$ (where $t, s > 0$). It is easy to see that populations in which either allele is common can resist invasion by the alternative allele. For example, if A is common, most of the A alleles will occur in AA homozygotes and thus have average fitness of one, while most B alleles are in heterozygotes and have fitness $1 - s$. When B is common, it has higher fitness for the same reason. This means that if A is initially common, individual selection will never lead to the spread of the B allele, even though it leads to higher fitness. However, Wright argued that group selection can lead to the spread of the B allele. Suppose that a large population is subdivided into a number of genetically well-mixed demes linked by low rates of gene flow. Selection is strong enough that in any given deme either A is common or B is common. Now, apply the Price equation to this population: since one or the other allele is common in each deme, V_W is small in all demes, and since selection *within* different demes pulls in opposite directions, the average value of β_W will also be small. Thus, the within-group component of the Price equation is close to zero—because each deme is near a stable equilibrium, selection within groups has little effect on the frequency of the two alleles. Now, consider the between-group term: because selection is much stronger than migration, there will be lots of variation among demes. Thus, if the fact that the B allele has higher average fitness translates into between-group selection, this process

will lead to the spread of that allele. This can happen in at least three different ways. Higher average fitness could lead to more out-migration, and this in turn can lead to the spread of the *B* allele through differential proliferation, the basis of the third phase of Wright's shifting balance model (Gavrilets 1995). Second, *B* will spread if groups with higher average fitness have lower extinction rates and new groups are formed by the fissioning of existing ones (Boyd & Richerson 1990). Finally, *B* will spread if higher average fitness attracts immigrants, and as a consequence the larger group spreads or splits (Boyd & Richerson 2009).

4. MANY PROCESSES LEAD TO MULTIPLE STABLE EQUILIBRIA

We believe that evolutionary systems with many equilibria are very common. Engineering experience suggests that even ordinary adaptive problems like the design of tools or shelters typically have many locally optimal solutions. The frequency dependence introduced by social interaction vastly multiplies the potential for multiple equilibria. For example, coordination systems resulting from communication, group movement and bargaining generate many equilibria. The possibilities for multiple equilibria are further increased by repeated interactions and contingent behaviour. Especially important are systems of moral norms enforced by reputation, retribution or reciprocity, which can stabilize a vast range of behaviours. Finally, a conformist bias in social learning can stabilize virtually any behaviour.

(a) *Ordinary adaptive problems often have many solutions*

Textbook examples of evolution as an optimization process sometimes portray the adaptive problem as climbing a smooth hill with a single local maximum. However, there are good reasons to believe that real adaptive problems often have vast numbers of locally optimal solutions. Real world design problems have many dimensions that can interact in a complicated, nonlinear fashion. Even seemingly simple problems have much hidden complexity. Consider, for example, the design of bows. The overall length of the bow affects how strongly it must be bent to generate a given amount of force. Thus, the optimal construction depends on length. Shorter bows must sustain greater strains, and this affects the best cross-section, the kind of wood that is used, how the wood is cut from the tree, whether the bow is sinew backed, whether the handle is live or static and a host of other attributes. The best choice for any given attribute affects what is best for others. Once most people in a society have converged on a particular solution, trial-and-error often will not generate progress because small changes will make the design worse. However, different groups may come to different solutions, which then can compete either directly, say in warfare, or indirectly to attract imitators.

(b) *Coordination games*

In many kinds of social interactions, individuals can increase their payoffs if they can coordinate their

Table 1. Payoff matrix for a simple coordination game.

		younger son	
		partition	primogenitor
older son	partition	2,2	0,0
	primogenitor	0,0	5,1

choices. Game theorists refer to such interactions as coordination games. Bargaining interactions provide a good illustration of why coordination games lead to multiple stable equilibria. Suppose in a particular population there are two cultural variants governing beliefs about inheritance: equal partition among brothers, and primogenitor (only the oldest brother inherits). To keep things simple, let us suppose that all families have exactly two sons, and that the payoffs associated with each combination of beliefs within a family are given in table 1. When brothers agree, they have a higher payoff than when they disagree because disputes are costly. This means that once either variant becomes common, people with the common variant achieve a higher payoff on average, and if the cultural evolution is driven by payoffs (for example, because people imitate the successful), then both inheritance institutions will be evolutionarily stable. Also notice that coordination games may involve conflicts of interest. Younger sons prefer partition while older sons prefer primogenitor.

A wide range of social interactions give rise to coordination games. Classic examples are social conventions, drive on the right versus drive on the left, matrilineal versus patrilineal post-marital residence. Signalling systems may also have many equilibria and the behaviour signaller and receiver have to be coordinated. People could signal their health or wealth in many different ways. Some may be better than others, but once the whole population converges on a given channel, individuals who deviate will lose out.

(c) *Conformist social learning*

There are good reasons to believe that social learning is often subject to a conformist bias, meaning that individuals are disproportionately likely to imitate the most common variant they observe in their social environment. Conformist makes sense if we think of the psychology of social learning as having been designed to acquire adaptive information. Adaptive processes will tend to increase the frequency of the locally adaptive behaviour. Transmission and learning errors, and changing environments will reduce it, but under many circumstances the most adaptive behaviour will, on average, be the most commonly observed behaviour. Thus, preferentially imitating the most common behaviour, individuals will increase the chance of acquiring the most adaptive behaviour. This intuitive argument is supported by modelling work which indicates that selection favours a conformist psychology in variable environments (Henrich & Boyd 1998; McElreath *et al.* 2008) and when cultural

transmission is error-prone (Henrich & Boyd 2002). Laboratory experiments support these predictions (Efferson *et al.* 2008; McElreath *et al.* 2008).

Conformist social learning creates an evolutionary force that causes common variants to become more common and rare variants to become more rare. If this effect is strong compared with migration, then variation among groups can be maintained. To see why, think of a number of groups linked by migration. Assume that the two cultural variants affect religious beliefs: ‘believers’ are convinced that moral people are rewarded after death and the wicked suffer horrible punishment for eternity, while ‘heretics’ do not believe in any afterlife. Because they fear the consequences, believers behave better than heretics—more honestly, charitably and selflessly. As a result, groups in which believers are common are more successful than groups in which heretics are common. Moreover, it is plausible that people’s decision to adopt one cultural variant or the other might not be strongly affected by content bias. True, people seek comfort, pleasure and leisure and this can cause them to behave wickedly. However, a desire for comfort also causes people to worry about spending an eternity buried in a flaming tomb. Since people are uncertain about the existence of an afterlife, they might not be strongly biased in favour of one cultural variant or the others. As a result, they are strongly influenced by the cultural variant that is common in their society. People who grow up surrounded by believers, choose to believe, while those who grow up among worldly atheists do not.

(d) *Moralistic punishment*

Moralistic punishment can also stabilize a very wide range of behaviours. To see why, consider the following simple example. Imagine a population subdivided into a number of groups. Cultural practices spread between groups because either people migrate, or they sometimes adopt ideas from neighbouring groups. Two alternative culturally transmitted moral norms exist in the population, norms that are to be enforced by moralistic punishment. Let us call them norm x and norm y . These could be ‘must wear a business suit at work’ and ‘must wear a dashiki to work’, or ‘a person owes primary loyalty to their kin’ and ‘a person owes primary loyalty to their group’. In groups where one of the two norms is common, people who violate the norm are punished. Suppose that people’s innate psychology causes them to be biased in favour of norm y , and therefore y will tend to spread, all other things being equal. Nonetheless, when norm x is sufficiently common, the effects of punishment overcome this bias and people tend to adopt norm x . In such groups, new immigrants whose beliefs differ from the majority (or people who have adopted ‘foreign’ ideas) rapidly learn that their beliefs get them into trouble and adopt the prevailing norm. When more believers in norm y arrive, they find themselves to be in the minority, rapidly learn the local norms and maintain norm x despite the fact that it is not the norm that fits best with their evolved psychology.

5. INTERDEMIC GROUP SELECTION IS PROBABLY MORE IMPORTANT IN CULTURAL THAN GENETIC EVOLUTION

This mechanism only works when the adaptation within groups is a much stronger evolutionary force than migration among groups, and thus is not likely to be an important force in genetic evolution. Evolutionary biologists normally think of selection as being weak, and, although there are many exceptions to this rule, it is a useful generalization. So, for example, if one genotype had a 5 per cent selection advantage over the alternative genotype, this would be thought to be an extremely strong selection. So, suppose that a novel, group-beneficial genotype has arisen, and that it has become common in one local group where it has a 5 per cent advantage over the genotype that predominates in the population as a whole. For group selection to be important, the novel type must remain common long enough to spread by group selection, and this will only be possible if the migration rate per generation is substantially less than 5 per cent. Otherwise, the effects of migration will swamp the effects of natural selection. But this is not very much migration. In most group-living primates, the members of one sex leave at sexual maturity, and there are about two generations present at any moment, and thus the migration rate between neighbouring such primate groups is of the order of 25 per cent per generation. While migration rates are notoriously difficult to measure, most likely migration rates are typically high among small local groups that suffer frequent extinction. Migration rates between larger subdivisions of a population are probably much lower, but so too will be the extinction rates. In contrast, we know that social learning processes are very rapid, and that they can maintain behavioural differences among neighbouring human groups despite substantial flows of people and ideas between them.

As a result, human groups are more like different species than populations of the same species, and this may be why phylogenetic methods work so well for cultural variation. If two human groups have different adaptations to the same ecological niche, the dynamics of their evolution often look more like competitive exclusion than conventional multi-level evolution in a metapopulation of the same species. The speed of cultural competitive exclusion is often enhanced because people moving from the losing group can be assimilated into the winning one, or because ideas diffuse from winning to losing groups. Barriers between human ‘species’ are often selective, limiting the effects of migration into the successful competitor but accelerating the flow of successful ideas into the less successful group.

Recently, Lehmann, Feldman and colleagues (Lehmann & Feldman 2008; Lehmann *et al.* 2008) have published several theoretical studies which they claim show that culture does not facilitate the evolution of cooperation by the mechanisms outlined above. In each paper, they present models that they claim (in one case ‘exactly’: Lehmann *et al.* 2008, p. 22) capture the processes discussed here, and then derive results showing that culture makes it harder for selection to favour cooperative behaviour. These

claims are mistaken. In all of these papers, Lehmann and colleagues assume that selection (or analogous adaptive cultural processes) is weak enough that it can be ignored when calculating the variation among groups (or alternatively, the relatedness within groups). This ‘quasi-equilibrium’ assumption means that neither the multiple adaptive equilibria, nor conformist social learning maintains variation among groups. Instead, their models assume that groups are small enough that there is a substantial probability that two individuals chosen randomly from within a group acquired their culture from the same model, and, as a consequence, common descent and limited migration can give rise to substantial variation among large groups. Because the way that variation is maintained is radically different in these models, they have no relevance to the processes discussed in this paper. Moreover, their explanation of between-group variation is not empirically plausible. In the modern world, there is substantial variation in beliefs and norms among ethnic groups and nation states that number millions of individuals (Bell *et al.* 2010). It is not plausible that four million Kamba (East African ethnic group) share language and many beliefs (Bell *et al.* 2010) because a substantial fraction of the Kamba acquired their beliefs by imitating the same person. Nor is this account believable for the small-scale societies that dominated most of human history because even in such societies, the scale of cultural variation is larger than the scale of everyday interaction. For example, Australian groups that shared a common language and culture typically numbered between 500 and 5000 (Keen 2004). If we assume that bands numbered between 10 and 100 people, and that everybody in a band imitates a single individual, then the formulae used by Lehman and colleagues predict that only a small fraction of cultural variation will be between ethnolinguistic units.

6. THREE TYPES OF INTERGROUP COMPETITION HAVE BEEN STUDIED

In the *Origin of Species*, Darwin (1859) famously argued that three conditions are necessary for adaptation by natural selection: first, there must be a ‘struggle for existence’ so that not all individuals survive and reproduce. Second, there must be variation so that some types are more likely to survive and reproduce than others, and finally, variation must be heritable so that the offspring of survivors resemble their parents. While Darwin usually focused on individuals, the same three postulates apply to any reproducing entity—molecules, genes and cultural groups. We have seen that rapid cultural adaptation in human societies combined with multiple equilibria give rise to stable, between-group differences that are heritable at the group level. Symbolic boundary markers act to limit the flow of ideas from one group to the other. Thus, there will be adaptation at the group level as long as groups compete in such a way that the cultural variants that characterize successful groups spread. We have been able to think of three different mechanisms of intergroup competition.

(a) *Variation in extinction rates*

The simplest mechanism is intergroup competition. The spread of the Nuer at the expense of the Dinka in the nineteenth century Sudan provides a good example. During the nineteenth century, each language group was divided into a number of independent polities. Cultural differences in norms between the two groups meant that the Nuer were able to organize larger war parties than the Dinka. The Nuer, who were driven by the desire for more grazing land, attacked and defeated their Dinka neighbours, occupied their territories and assimilated tens of thousands of Dinka into their communities. This example illustrates the requirements for cultural group selection by intergroup competition. Contrary to some critics (Palmer *et al.* 1997), there is no need for groups to be strongly bounded, individual-like entities. The only requirement is that there be persistent cultural differences between groups, and these differences must affect the group’s competitive ability (Boyd & Richerson 1990). Losing groups must be replaced by the winning groups. Interestingly, the losers do not have to be killed. The members of losing groups just have to disperse or to be assimilated into the victorious group. Losers will be socialized by conformity or punishment, so even very high rates of physical migration need not result in the erosion of cultural differences. This kind of group selection can be a potent force even if groups are usually very large.

Group competition is common in small-scale societies. The best data come from New Guinea, which provides the only large sample of simple societies studied by professional anthropologists before they experienced major changes owing to contact with Europeans. Joseph Soltis (Soltis *et al.* 1995) assembled data from the reports of early ethnographers in New Guinea. Many studies report appreciable intergroup conflict and about half mention cases of social extinction of local groups. Five studies contained enough information to estimate the rates of extinction of neighbouring groups (table 2). The typical pattern is for groups to be weakened over a period of time by conflict with neighbours and finally to suffer a sharp defeat. When enough members become convinced of the group’s vulnerability to further attack, members take shelter with friends and relatives in other groups, and the group becomes socially extinct. At these rates of group extinction, it would take between 20 and 40 generations, or 500–1000 years, for an innovation to spread from one group to most of the other local groups by cultural group selection.

These data suggest that cultural group selection is a fairly slow process. But then, so are the actual rates of increase in political and social sophistication we observe in the historical and archaeological records. Change in the cultural traditions that eventually led to large-scale social systems like those that we live in proceeded at a modest rate. The relatively slow rate of evolution of cultural group selection may explain the 5000-year lag between the beginnings of agriculture and the first primitive city-states, and the five millennia that passed between the origins of simple states and modern complex societies.

Table 2. Extinction rates for cultural groups from five regions in New Guinea. Adapted from Soltis *et al.* (1995).

region	number of groups	number of social extinctions	number of years	% groups extinct every 25 years
Mae Enga	14	5	50	17.9
Maring	13	1	25	7.7
Mendi	9	3	50	16.6
Fore/Usurufa	8–24	1	10	31.2–10.4
Tor	26	4	40	9.6

(b) Imitation of successful neighbours

A propensity to imitate the successful can also lead to the spread of group-beneficial variants. People often know about the norms that regulate behaviour in neighbouring groups. They know that we can marry our cousins here, but over there they cannot; or anyone is free to pick fruit here, while individuals own fruit trees there. Suppose different norms are common in neighbouring groups, and that one set of norms causes people to be more successful. Both theory and empirical evidence suggest that people have a strong tendency to imitate the successful (Henrich & Gil-White 2001; Richerson & Boyd 2005; McElreath *et al.* 2008). Consequently, behaviours can spread from groups at high payoff equilibria to neighbouring groups at lower payoff equilibria because people imitate their more successful neighbours. A mathematical model suggests that this process will spread group-beneficial beliefs from one group to another, resulting in a wave-like advance, and that this occurs over a wide range of conditions (Boyd & Richerson 2002). The model also suggests that such spread can be rapid. Roughly speaking, it takes about twice as long for a group-beneficial trait to spread from one group to another as it does for an individually beneficial trait to spread within a group.

This kind of group selection is also likely to be faster than that owing to differential extinction because it readily leads to the recombination of group-beneficial strategies that initially arise in different groups (Boyd & Richerson 2002). The exact combination of strategies necessary to support complex, adaptive social institutions would seem unlikely to arise through a single chance event. It is much more plausible that complex institutions are assembled in numerous small steps. Differential extinction models are analogous to the evolution of an asexual population in which they lack any mechanism that allows the recombination of beneficial strategies that arise in different populations, and thus require innovations to occur sequentially in the same lineage. In contrast, the spread of ideas from successful groups allows recombination of different strategies and thus more rapid cumulative change.

The rapid spread of Christianity in the Roman Empire may provide an example of this process. Between the death of Christ and the rule of Constantine, a period of about 260 years, the number of Christians increased from a only a handful to somewhere between 6 and 30 million people (depending on whose estimate you accept). This sounds like a huge increase, but it turns out that it is equivalent to a 3–4% annual rate of increase, about the same as

the growth rate of the Mormon Church over the past century. According to the sociologist Rodney Stark, many Romans converted to Christianity because they were attracted to what they saw as a better quality of life in the early Christian community. Pagan society had weak traditions of mutual aid, and the poor and sick often went without any help at all. In contrast, in the Christian community norms of charity and mutual aid created ‘a miniature welfare state in an empire which for the most part lacked social services’ (Johnson 1976, p. 75, quoted in Stark 1997). Such mutual aid was particularly important during the several severe epidemics that struck the Roman Empire during the late Imperial period. Unafflicted pagan Romans refused to help the sick or bury the dead. As a result, some cities devolved into anarchy. In Christian communities, strong norms of mutual aid produced solicitous care of the sick, and reduced mortality. Both Christian and Pagan commentators attribute many conversions to the appeal of such aid. For example, the emperor Julian (who detested Christians) wrote in a letter to one of his priests that Pagans needed to emulate the virtuous example of the Christians if they wanted to compete for their souls, citing ‘their moral character even if pretended’ and ‘their benevolence toward strangers’ (Stark 1997; pp. 83–84). Middle class women were particularly likely to convert to Christianity, probably because they had higher status and greater marital security within the Christian community. Roman norms allowed polygyny, and married men had great freedom to have extramarital affairs. In contrast, Christian norms required faithful monogamy. Pagan widows were required to remarry, and when they did, they lost control of all of their property. Christian widows could retain property, or, if poor, would be sustained by the church community. Demographic factors were also important in the growth of Christianity. Mutual aid led to substantially lower mortality rates during epidemics, and a norm against infanticide led to substantially higher fertility among Christians.

This form of group selection may also explain the spread of moral norms that stigmatize ‘victimless’ crimes, for example, drunkenness or prostitution. There is by now a large literature that indicates that people often have time-inconsistent preferences and as a result, they often make choices in the short run that they know are not in their long-run interest. It is plausible that social norms help people solve these problems by creating short-run incentives to do the right thing. I may not be able to resist a drink when the costs are all in the distant future, but make a different decision if I suffer immediate social disapproval. It is

also easy to see why such norms persist once they are established. If everyone agrees that self-control is proper behaviour and punish people who disagree, then the norm will persist. The problem is that the same mechanism can stabilize any norm. People could just as easily agree that excessive drinking is proper behaviour and punish teetotalers. If, however, groups in which drinkers are stigmatized achieve better outcomes, and if those outcomes are observable, the norm can spread from one group to another by differential imitation.

(c) *Selective migration*

Selective migration, the tendency of people to move from less desirable to more desirable societies, can also lead to the spread of some kinds of equilibria. We are very familiar with this process in the modern world where streams of migrants flow from societies that migrants perceive as offering them fewer opportunities toward ones that appear to offer them more (Martin 2005). The extensive literature on this topic (e.g. Alba & Nee 2003; Borjas 1994) supports two generalizations: (i) that migrants flow from societies where immigrants find their prospects poor to those where they perceive them to be better, and (ii) most immigrant populations assimilate to the host culture within a few generations. Ethnographic evidence suggests that selective immigration is not limited to industrialized nation states, and thus may be an ancient phenomenon (Knauff 1985; Cronk 2002). The spread of cultural institutions associated with ancient complex societies, such as China, Rome and India, supports the idea that this process is not new. Ancient imperial systems often expanded militarily but the durable ones, such as Rome, succeeded by assimilating conquered peoples and by inducing a flow of migrants across their boundaries. Although the Roman empire eventually faded, its most attractive institutions were adapted by successor polities and persist in modified form to this day. Rome, India, China and Islamic civilization stand in stark contrast to pure conquest empires like that of the Mongols, which expanded but did not assimilate.

The simple mathematical model of this process (Boyd & Richerson 2009) indicates that it has two qualitatively evolutionary outcomes. The model assumes that there are two possible evolutionary equilibria in an isolated population, and one equilibrium leads to higher average welfare than the other. The population is subdivided into two subpopulations linked by migration. There is more migration from low-payoff to high-payoff subpopulations than the reverse. When local adaptation is strong enough when compared with migration to maintain cultural variation among subpopulations, the population as a whole evolves towards a polymorphic equilibrium at which the variants that produce higher average welfare are more common, but the lower payoff variant also persists. Initial subpopulation size and the sizes of the basins of attraction play relatively minor roles. When migration is stronger, however, initial population sizes and sizes of the basins of attraction predominate. The variant that is common in the

larger of the two populations tends to spread and the other variant tends to disappear even it yields a higher payoff.

7. THREE PROCESSES CAN SHIFT GROUPS TO NOVEL EQUILIBRIA

Selection always requires a source of variation, and so it is with all of the group-selection mechanisms described above. If all of the groups are characterized by the same equilibrium behaviour, group selection can have no effect. There must be some process that causes groups to shift from one equilibrium to another, the analogue of a new mutation at the group level. Three different processes may have this effect. First, sampling variation affecting who happens to get copied and who happens to interact with whom will generate random changes in the frequencies of cultural variants analogous to genetic drift, and these will occasionally cause populations to shift from the neighbourhood of one stable equilibrium to a second equilibrium. In large populations, the waiting time until such shifts occur can be very long (e.g. Lande 1985). Environments that vary in time so that adaptive forces shift magnitude and direction can also create drift-like forces that lead to shifts from one peak to another, but these forces do not depend on population size (Gillespie 2000). Note that the environmental variation need not directly affect the trait in question if the transmission of different traits is linked, for example, because individuals tend to acquire a suite of traits from the same individual. Finally, the frequency of cultural traits is affected by learning, and chance variation in cues from the environment will lead to drift-like shifts in trait frequencies. Moreover, if the cues available to different individuals in the population are correlated, this could lead groups to shift rapidly from the basin of attraction of a second equilibrium. For example, according to Dower (1999), the experience of losing World War II led many Japanese people to adopt strongly pacifist beliefs. If things had gone differently at Midway, Japan might not have lost, and the Japanese population might have instead maintained their previously held, strongly militaristic beliefs.

8. GENE–CULTURE COEVOLUTION

Over longer time scales, social environments shaped by cultural group selection may have affected the genetic evolution of the human species. The archaeological record suggests that cumulative cultural evolution arose in the human lineage sometime between 250 and 500 thousand years ago. As a consequence, social environments shaped by cultural group selection may have generated novel selection pressures on genes influencing human social behaviour. For example, the existence of group-beneficial norms enforced by moralistic punishment might select for moral emotions like shame, and cognitive mechanisms like cheater detection because such genetically transmitted adaptations reduced the chance that their bearers would be punished (Richerson & Boyd 2005). It has also been suggested that cultural group

selection may explain the low levels of genetic variation within the human species (Premo & Hublin 2009). These authors argue that cultural variation between groups reduces the amount of gene flow among groups, and this in turn increases the fraction of human genetic variation among groups. Then, competition between culturally different groups led to group extinction, and thus reduced the genetic variation in the human species as a whole.

9. CONCLUSION: WHAT COMES NEXT?

The theory of cultural group selection is fairly well worked out, and there are a number of convincing examples of the process at work. We believe that three kinds of additional research will be especially valuable. First, there has been little systematic quantitative empirical work that allows an assessment of the relative importance of cultural group selection compared with other processes that shape cultural variation. The main exception is the work of Soltis *et al.* (1995) estimating group extinction rates, described above. Similar estimates for a wider range of societies would be useful, as would analogous work on group selection by differential imitation and differential migrations. Second, group selection predicts that societies should exhibit design at the group level, that we should be able to understand the structure and variation of norms in terms of how they enhance group welfare (Wilson 2002). Of course, there is a long tradition of functionalist explanation in the social sciences, but for the most part this work takes the form of group-level just so stories. What is needed are sharp, testable hypotheses about how group-functional behaviours, especially group-functional norms, should vary with ecology, group size and other measurable variables. The field of 'law and economics' is a rich source of such hypotheses (e.g. Posner 1980). Finally we believe that group selection should leave detectable patterns in the ethnographic and archaeological records. There is a rich body of techniques for detecting individual selection using correlations among traits and biogeographic patterns, and the analogous methods may be useful in detecting group selection.

We would like to thank Joe Henrich, Ruth Mace, Richard McElreath, Luke Premo, Stephen Shennan and James Steele for useful discussions of the ideas reported in this paper.

REFERENCES

- Alba, R. & Nee, V. 2003 *Remaking the American mainstream: assimilation and the new immigration*. Cambridge, MA: Harvard University Press.
- Bell, A. V., Richerson, P. J. & McElreath, R. 2010 Culture rather than genes provides greater scope for the evolution of large-scale human prosociality. *Proc. Natl Acad. Sci. USA* **106**, 17 671–17 674. (doi:10.1073/pnas.0903232106)
- Borjas, G. J. 1994 The economics of immigration. *J. Econ. Lit.* **32**, 1667–1717.
- Boyd, R. & Richerson, P. J. 1990 Group selection among alternative evolutionarily stable strategies. *J. Theor. Biol.* **145**, 331–342. (doi:10.1016/S0022-5193(05)80113-4)
- Boyd, R. & Richerson, P. J. 1992 Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethol. Sociobiol.* **13**, 171–195. (doi:10.1016/0162-3095(92)90032-Y)
- Boyd, R. & Richerson, P. J. 2002 Group beneficial norms spread rapidly in a structured population. *J. Theor. Biol.* **215**, 287–296. (doi:10.1006/jtbi.2001.2515)
- Boyd, R. & Richerson, P. J. 2009 Voting with your feet: payoff biased migration and the evolution of group beneficial behavior. *J. Theor. Biol.* **257**, 331–339. (doi:10.1016/j.jtbi.2008.12.007)
- Cronk, L. 2002 From true Dorobo to Mukogodo Maasai: contested ethnicity in Kenya. *Ethnology* **41**, 27–49. (doi:10.2307/4153019)
- Darwin, C. 1859 *On the Origin of Species by Means of Natural Selection*, 1st edn. London, UK: John Murray.
- Dower, J. W. 1999 *Embracing defeat: Japan in the wake of World War II*. New York, NY: Blackstone, W. W. Norton Co.
- Efferson, C., Lalive, R., Richerson, P. J., McElreath, R. & Lubell, M. 2008 Conformists and mavericks: the empirics of frequency-dependent cultural transmission. *Evol. Human Behav.* **29**, 56–64. (doi:10.1016/j.evolhumbehav.2007.08.003)
- Gavrilets, S. 1995 On phase three of the shifting balance theory. *Evolution* **50**, 1034–1041. (doi:10.2307/2410644)
- Gillespie, J. H. 2000 Genetic drift in an infinite population: the pseudohitchhiking model. *Genetics* **155**, 909–919.
- Henrich, J. & Boyd, R. 1998 The evolution of conformist transmission and the emergence of between-group differences. *Evol. Hum. Behav.* **19**, 215–241. (doi:10.1016/S1090-5138(98)00018-X)
- Henrich, J. & Boyd, R. 2002 On modeling cognition and culture: why cultural evolution does not require replication of representations. *Cult. Cogn.* **2**, 87–112. (doi:10.1163/156853702320281836)
- Henrich, J. & Gil-White, F. J. 2001 The evolution of prestige—freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evol. Human Behav.* **22**, 165–196. (doi:10.1016/S1090-5138(00)00071-4)
- Holden, C. & Mace, R. 2005 The cow is the enemy of matriliney: using phylogenetic methods to study cultural evolution in Africa. In *The evolution of cultural diversity: a phylogenetic approach* (eds R. Mace, C. Holden & S. Shennan). London, UK: UCL Press.
- Huelsenbeck, J. P., Ronquist, F., Nielsen, R. & Bollback, J. P. 2001 Bayesian inference of phylogeny and its impact on evolutionary biology. *Science* **294**, 2310–2314. (doi:10.1126/science.1065889)
- Johnson, P. 1976 *A history of Christianity*. London, UK: Weidenfeld & Nicolson.
- Keen, I. 2004 *Aboriginal economy and society: Australia at the threshold of colonisation*. Oxford, UK: Oxford University Press.
- Knauff, B. M. 1985 Good company and violence: sorcery and social action in a lowland New Guinea society. In *Studies in Melanesian anthropology*. Berkeley, CA: University of California Press.
- Lande, R. 1985 Expected time for random genetic drift of a population between stable phenotypic states. *Proc. Natl Acad. Sci. USA* **82**, 7641–7645. (doi:10.1073/pnas.82.22.7641)
- Lehman, L. & Feldman, M. W. 2008 Cultural transmission can inhibit the evolution of altruistic helping. *Theor. Popul. Biol.* **73**, 506–515. (doi:10.1016/j.tpb.2008.02.004)
- Lehmann, L., Feldman, M. W. & Foster, K. 2008 Cultural transmission can inhibit the evolution of altruistic helping. *Am. Nat.* **173**, 12–24. (doi:10.1086/587851)
- Martin, P. 2005 Migrants in the global labor market. pp. 1–57. *Global Commission on International Migration* (see <http://www.gcim.org/en/>).

- Maynard Smith, J. 1964 Group selection and kin selection. *Nature* **201**, 1145–1147. (doi:10.1038/2011145a0)
- McElreath, R., Bell, A., Efferson, C., Lubell, M., Richerson, P. J. & Waring, T. 2008 Beyond existence and aiming outside the laboratory: estimating frequency-dependent and payoff-biased social learning strategies. *Phil. Trans. R. Soc. B* **363**, 3515–3528. (doi:10.1098/rstb.2008.0131)
- Palmer, C. T., Fredrickson, B. E. & Tilley, C. F. 1997 Categories and gatherings: group selection and the mythology of cultural anthropology. *Evol. Human Behav.* **18**, 291–308. (doi:10.1016/S1090-5138(97)00045-7)
- Posner, R. 1980 A theory of primitive society, with special reference to law. *J. Law Econ.* **23**, 1–53. (doi:10.1086/466951)
- Premo, L. L. & Hublin, J. J. 2009 Culture, population structure, and low genetic diversity among Pleistocene hominins. *Proc. Natl Acad. Sci. USA* **106**, 33–37. (doi:10.1073/pnas.0809194105)
- Price, G. R. 1970 Selection and covariance. *Nature* **227**, 520–521. (doi:10.1038/227520a0)
- Price, G. R. 1972 Extensions of covariance selection mathematics. *Ann. Human Genet.* **35**, 485–490. (doi:10.1111/j.1469-1809.1957.tb01874.x)
- Richerson, P. J. & Boyd, R. 2005 *Not by genes alone: how culture transformed human evolution*. Chicago, IL: University of Chicago Press.
- Rogers, A. R. 1990 Group selection by selective emigration: the effects of migration and kin structure. *Am. Nat.* **135**, 398–413. (doi:10.1086/285053)
- Soltis, J., Boyd, R. & Richerson, P. J. 1995 Can group-functional behaviors evolve by cultural group selection? An empirical test. *Curr. Anthropol.* **36**, 437–494. (doi:10.1086/204381)
- Stark, R. 1997 *The rise of Christianity: how the obscure, marginal Jesus movement became the dominant religious force in the Western world in a few centuries*. San Francisco, CA: Harper/Collins.
- Williams, G. C. 1966 *Adaptation and natural selection: a critique of some current evolutionary thought*. Princeton, NJ: Princeton University Press.
- Wilson, D. S. 2002 *Darwin's cathedral: evolution, religion, and the nature of society*. Chicago, IL: University of Chicago Press.
- Wright, S. 1931 Evolution in Mendelian populations. *Genetics* **16**, 97–159.
- Wynne-Edwards, V. C. 1962 *Animal dispersion in relation to social behavior*. Edinburgh, UK: Oliver and Boyd.