A model for the evolutionary diversification of religions

Michael Doebeli, Iaroslav Ispolatov

Department of Zoology, University of British Columbia, Vancouver, BC, Canada V6T 1Z4

Department of Mathematics, University of British Columbia, Vancouver, BC, Canada V6T 1Z4

ARTICLE INFO

Article history:
Received 19 January 2010
Received in revised form 29 July 2010
Accepted 8 September 2010
Available online 18 September 2010

Keywords:
Cultural evolution
Religion
Diversity
Epidemiological modeling

ABSTRACT

We address the problem of cultural diversification by studying selection on cultural ideas that colonize human hosts and using diversification of religions as a conceptual example. In analogy to studying the evolution of pathogens or symbionts colonizing animal hosts, we use models for host–pathogen dynamics known from theoretical epidemiology. In these models, religious content colonizes individual humans. Rates of transmission of ideas between humans, i.e., transmission of cultural content, and rates of loss of ideas (loss of belief) are determined by the phenotype of the cultural content, and by interactions between hosts carrying different ideas. In particular, based on the notion that cultural non-conformism can be negative frequency-dependent (for example, religion can lead to oppression of lower classes and emergence of non-conformism and dissent once a religious belief has reached dominance), we assume that the rate of loss of belief increases as the number of humans colonized by a particular religious phenotype increases. This generates frequency-dependent selection on cultural content, and we use evolutionary theory to show that this frequency dependence can lead to the emergence of coexisting clusters of different cultural types. The different clusters correspond to different cultural traditions, and hence our model describes the emergence of distinct descendant cultures from a single ancestral culture in the absence of any geographical isolation.

1. Introduction

Evolution can occur whenever there are units of reproduction that produce other such units which inherit some characteristics of the parent units. If the units of reproduction vary in their reproductive output, there will be evolutionary change. “Intellectual content” can satisfy these simple requirements. An idea or a theory can be viewed as colonizing the brain of an individual human (or animal). It can develop or mutate within that brain, and it can be passed on to the brains of other individuals, thereby reproducing itself. Such reproduction typically occurs with modification and through mixing and amalgamation with other ideas present in the human population. For a multitude of potential reasons, both cognitive and selective in nature (Henrich, 2009; Henrich et al., 2008; Richerson and Boyd, 2005), some ideas and theories are more successful at such reproduction through transmission than others, hence there is typically differential reproductive success. As a consequence, there is cultural evolution of intellectual content such as ideas and theories.

Such a perspective of cultural evolution has been lucidly advocated by Cavalli-Sforza and Feldman (1981) and Boyd and Richerson (1985), and by now there is a rather large body of literature on cultural evolution (see e.g. Richerson and Boyd, 2005). In particular, population genetic models have been adopted early on to study cultural evolution (Boyd and Richerson, 1985; Cavalli-Sforza and Feldman, 1981). On the other hand, an epidemiological perspective of cultural content colonizing human brains has been advocated in various verbal narratives by a number of researchers (e.g. Dietz, 1967; Sperber, 1996). However, despite some early attempts (e.g. Cavalli-Sforza and Feldman, 1981), it seems that to date this epidemiological perspective has not been rigorously adopted as a basis for the mathematical modeling of cultural evolution. There is a large body of theoretical and mathematical literature on the modeling of the epidemiological dynamics of pathogens in humans and many other animals. While one naturally has to be careful in adopting such models for cultural evolution, we believe that such epidemiological models might be well suited to study cultural dynamics (Dietz, 1967).

Here we apply mathematical epidemiological models to address the problem of cultural diversification. It seems perhaps relatively easy to understand how cultural differentiation can develop between human populations that live in isolation from each other (e.g. on different continents), because different climatic and biological circumstances may lead different cultural content to thrive in different areas of the globe. Much thought has been given to understanding what happens when cultures come into contact after they have diversified in isolation. Such an
In a more recent example, Whitehouse (1995) has observed processes within the Catholic church may have created and generate and/or become susceptible to new religious ideas. Thus, the peer pressure to conform to the existing doctrine, such as the removal of non-conformist bishops or the sale of indulgences, which is an ongoing splitting off of minor sectarian movements from a mainstream religious organization in Papua New Guinea. In the sociological literature, such a new branch of religion that is formed from within and in coexistence with the old religion is often referred to as a “sect” (Stark, 1996), while a following generated by an original set of ideas, sufficiently unrelated to any existing religion (at all, or typical to a particular geographical area) is defined as a “cult”. In this language, our the models presented here would (loosely) refer to the formation of new sects rather than new cults.

In particular, we propose to model cultural diversification arising without spatial segregation. We use the emergence of new religions, or sects, as a schematic backdrop for informing our modeling assumptions, but conceptually, our approach can be used to describe general scenarios for the evolution of new sets of cultural content (e.g. languages or ideologies) from a single ancestral culture. Borrowing from the epidemiological literature, our models incorporate human individuals as hosts for religious content. Just as in traditional epidemiological models, religious content can be lost by a human host, and it can be transmitted between human hosts. We use established mathematical techniques from evolutionary theory to describe the dynamics of host populations that harbour a variable set of religious ideas that are, for simplicity, described by a continuous real variable. The trait value of a religious idea determines the idea’s propensity of being lost by their human hosts, as well as their success in colonizing susceptible hosts. The basic question we address concerns the distribution of religious values: when does a unimodal distribution of non-conformist behaviors seem to be present in almost any cultural setting, ranging from languages to politics, economics, and religion.

2. Model setup

Religions are sets of ideas, statements and prescriptions of whose validity and applicability individual humans can become convinced. Thus, individual minds are the hosts of religious ideas, and these ideas can exert considerable influence on the behavior of their hosts. In principle, understanding the dynamics of religion can be achieved by understanding the interaction between religious ideas and their hosts, i.e., by understanding how religion affects not only the behavior of individual hosts, but also the social structure of host populations, and how human behavior, cognitive and social structures in turn affect the transmission of religious ideas among host individuals. For example, the number of host individuals adopting a given religion grows, this social structure may give rise to unrest, particularly in the lower social ranks. As a consequence, individuals may be enticed to start or adopt alternative, “unspoiled” religions, which offer less repression, and in which they can attain improved social and spiritual status, both in terms of intellectual satisfaction and hierarchical position. For example, it has been suggested that social unrest led to the split of the Protestant church from the Catholic church in the 16th century (Tuchman, 1985). In that time, political developments led to ever increasing financial needs of the Catholic church, which burdened its followers through taxation and other means, e.g., the sale of indulgences. This in turn led to unrest and spiritual decay and contributed to the secession of a more democratic and less repressive religion. In other words, hosts of Catholicism tended to lose faith due to effects that Catholicism itself generated in their host society. Moreover, belief-losing hosts became susceptible to a similar, but distinct type of religion that promised to improve the conditions of these hosts, probably at least in part because the new religion was not very common, and hence did not have the same detrimental effects on its hosts as ancestral one. Other potential examples of such secessions include the diversification of Judaism into more traditional Orthodox, Reform, and Humanistic branches, as well as a very recent split of the Anglican church based on different positions towards homosexual bishops and marriages.

From a complimentary perspective, the loss of belief due to overcrowding of the ancestral religion can be seen as the deepening of a conflict between non-conformists and progressively less flexible mainstream ideologies. As the number of followers of a religion grows, the initially diverse set of ideas inevitably solidifies into a number of increasingly dogmatic principles that are easy to communicate to new adepts and thus hard to alter or even doubt. A vivid example of such dogmatization is the emergence of mainstream Christianity from a large number of early branches with an initially diverse vision and interpretation of what later became known as Biblical events. Another example is the consolidation of a very multifaceted early
Marxism into rigorous ideologies of Stalinist USSR and Maoist China with physical extermination of even the slightest dissent. At more advanced stages of the development of the ancestral religion, it then becomes more promising for non-conformists to cede and start an independent school of thought rather than to argue for changes in the mainstream ideology. Of course, many other forces impinge on the well-being of hosts of particular religious ideas. For example, common religions may offer protection and a well-established machinery of spiritual and economic support, and rare ones may suffer persecution, so that many dissident movements may not be successful. Thus, there are numerous aspects of religious life that are described in terms of positive frequency dependence, which compete with any trends that are due to negative frequency dependence. Nevertheless, it seems reasonable that some fraction of followers of a religion would tend to value freedom from dogmatism and repression higher than certain levels of protection and comfort, thus becoming susceptible to dissent. Here we propose that mechanisms such as the ones alluded to above could cause negative frequency-dependent selection on religious ideas, and, as a consequence, adaptive cultural diversification.

Under the perspective of host individuals being colonized by religious ideas, it is natural to attempt a formal analysis of the evolution of religion using epidemiological models. Such models are very well studied in the context of disease dynamics (Otto and Day, 2007). In the simplest case, there is only one type of religion present, and the corresponding model describes the dynamics of two variables, each associated with a subpopulation of the total host population: \( S \) denotes the density of susceptible individuals in the host population, i.e., individuals that are not yet hosts to the given religion, and \( C \) denotes the density of colonized hosts, i.e., host individuals whose minds have adopted the given religion.

Our analysis is based on the following “ecological” model for the dynamics of susceptible and colonized hosts:

\[
\begin{align*}
\frac{dS}{dt} &= r_SS \left( 1 - \frac{S + C}{K_S} \right) - \tau SC + IC \\
\frac{dC}{dt} &= r_C C \left( 1 - \frac{S + C}{K_C} \right) + \tau SC - IC
\end{align*}
\]

(1)

Here we have assumed that both susceptible and infected hosts grow logistically. Thus, in the absence of religious colonization, susceptible hosts have an intrinsic growth rate \( r_S \) and grow logistically to carrying capacity \( K_S \), and in the absence of susceptibles, hosts colonized by the religion have an intrinsic growth rate \( r_C \) and grow logistically to carrying capacity \( K_C \). For simplicity, we assume that offspring of religious hosts are also religious (in principle, part or all of these offspring could first join the susceptible class), and that offspring of susceptible hosts also belong to the susceptible class. The two death terms for susceptible and colonized hosts are coupled by assuming that growth depends on the sum of the two populations \( S \) and \( C \). In addition, susceptible hosts adopt religion, i.e., become colonized, at a per capita rate \( \tau C \) that is proportional to the number of religious hosts. However, religious people also lose their belief at a per capita rate \( l \) and become susceptible once again, leading to a decrease in \( C \) at a rate \( IC \) and a corresponding increase in \( S \).

To introduce variability in religion and thus to allow for religious diversification, we expand this model by making the very simplistic assumption that religion types are characterized by a one-dimensional trait \( x \) with distribution \( C(x) \). For example, it has been proposed by Whitehouse (1995) that religions can be characterized by either having a high frequency of low-arousal rituals, or having a low frequency of high-arousal rituals. One could then describe religions simply by the frequency of their rituals. Mathematically, \( C(x) \) is the population density of hosts colonized by religious ideas with values in the interval \( (x-x_0 dx) \).

To introduce frequency-dependent selection on religious content, we first introduce a measure of “overcrowding” by defining, for any given type \( x \), the function

\[
A(x) = \int_y (x-y)C(y) \, dy
\]

(2)

where \( A(x-y) \) is a unimodal function of the form

\[
z_0 \exp \left( -\frac{1}{2} \frac{(x-y)^2}{\sigma_x^2} \right)
\]

(3)

The exponent \( b_1 \) in \( A(x-y) \) is a positive real number. For example, if \( b_1 = 2 \), \( A(x-y) \) is a Gaussian function. Technically speaking, \( A(x) \) is a convolution of the density distribution \( C(y) \) with the “kernel” \( A(x-y) \). Such a convolution corresponds to a weighted sum over all densities \( C(y) \), with the weights \( z_0 A(x-y) \). Since \( A(x-y) \) has a maximum at \( x=y \) and decreases to 0 as the distance \( |x-y| \) increases, the densities \( C(y) \) of hosts colonized with religious types \( y \) that are very different from the focal type \( x \) have little weight, and hence matter little for calculating the quantity \( A(x) \), where as the density of hosts colonized by types that are more similar to the focal \( x \) have more weight in calculating the overcrowding \( A(x) \) at \( x \). In general, if the distribution \( C(x) \) is unimodal with a single maximum at \( x=x_0 \), then overcrowding \( A(x) \), Eq. (3), tends to be large for \( x \) close to \( x_0 \), i.e., for common \( x \), and, conversely, \( A(x) \) tends to be small for \( x \) very different from \( x_0 \), i.e., for rare \( x \).

We then assume that the per capita rate of loss of the religious idea, \( l \), is a function of overcrowding, \( l(A(x)) \), where \( l(z) \) increases monotonically with increasing \( z \). For simplicity, we assume \( l(z) = z \). This implies that the rate of loss is high for hosts carrying religious ideas \( x \) for which \( A(x) \) is large, whereas the rate of loss is small for hosts carrying religious ideas \( x \) for which \( A(x) \) is small. Thus, hosts are more likely to lose common religious content than rare religious content. As mentioned above, one rationale for this assumption is that once a religion becomes common, the social structure may change such that the benefits gained from adopting the religion decrease for the majority of hosts, so that, on average, hosts of such religion become more likely to lose belief.

With religious variability, the differential equation for \( C \) must be replaced by a partial differential equation describing the dynamics of the distribution \( C(x) \). To model this, we assume that offspring of hosts colonized by religious type \( x \) on average also carry type \( x \), but with a certain probability the religion carried by the offspring undergoes a small mutation (as e.g. when children adopt religious notions that are slightly different from those of their parents, see Henrich, 2009). Thus, the offspring of a parent with religion \( y \) has a probability \( N_{x, x_0} \) to lie in the interval \( (x+x_0 dx) \), where \( N_{x_\pm} \) is a normal distribution with mean the parental type \( y \) and mutational variance \( \sigma_m \). We are aware of the fact that this is an extremely simplified view of the cognitive process of learning religions (Henrich, 2009). In fact, it has been argued that cognitive biases are the main determinants of cultural evolution, but this argument has been largely defused (Henrich et al., 2008). Nevertheless, the simple models presented here do not address the important questions of cognitive bias and selective learning, and instead assume a very simple form of transmission of cultural content. It should be pointed out, though, that it is entirely possible to include cognitive biases and more complicated forms of transmission (e.g. learning from the population mean) in epidemiological models of the type studied here.

With this in mind, the epidemiological dynamics for religiously variable host populations becomes

\[
\frac{dS}{dt} = r_SS \left( 1 - \frac{S + \int x C(x) \, dx}{K_S} \right) - \tau SC + IC + \int A(x)C(x) \, dx
\]

(4)
\[
\frac{\partial C}{\partial t} = r_C \int N_y \sigma_C(y) dy - \frac{r_C C(x)}{K_C} \int \frac{f_x C(x) dx}{K_C} + \tau(x)SC(x) - A(x)C(x)
\] (5)

To include mutation at birth, we have separated the birth and death term of the logistic equation, with \(r_C \int N_y \sigma_C(y) dy\) describing all offspring that are born to parents with all possible religious types \(y\) and whose type mutated to \(x\). In the death term, \(\int A(x)C(x) dx\) is the total population size of religious hosts.

To describe recruitment of a new host to a religion, we assume that only a susceptible individual can be converted to a bear of a certain trait \(x\). This is in agreement with the observation made in Stark (1996) that both cults and sects recruit new members most successfully from the non-religious population rather than by converting strong believers. We have made the additional assumption that the transmission rate \(\tau(x)\) is a function of the religious trait \(x\). This function is assumed to reflect some intrinsic properties of religious ideas that determine their likelihood of transmission to susceptible hosts. For example, some ideas might entice their carriers to proselytize more than others, but such activities might come at certain costs. The function \(\tau(x)\) is assumed to reflect the balance of such costs and benefits. Specifically, we assume that this function is unimodal, so that there is a unique "optimal" religious type in terms of transmissibility. This introduces a stabilizing component of selection on the trait \(x\), restricting the proliferation of very extreme ideas characterized by large deviation of \(x\) from the optimum \(x_0\). The unimodality of \(\tau(x)\) is obviously a simplification, and in general the transmission function could have multiple local maxima (especially in multi-dimensional phenotype spaces). However, the unimodality assumption is made here for the sake of argument and to show that even the simplest assumptions can lead to interesting outcomes.

Specifically, we will use the form

\[
\tau(x) = \tau_0 \exp \left( -\frac{1}{2} \left[ \frac{x-x_0}{\sigma_x} \right]^b \right)
\] (6)

where the exponent \(b\) is a positive real number. Note that for \(b=2\), \(\tau(x)\) is a Gaussian function. The rate at which hosts colonized with type \(x\) convince susceptible individuals of their religion is \(\tau(x)C(x)\), so that the total per capita rate of transmission for susceptible hosts is \(\int \tau(x)C(x) dx\). Note again that this assumed process of "convincing" is extremely simplified. In reality, even if a person is successful in convincing another person of some religious ideas, it is likely that in this newly convinced person, the religion takes a slightly different form from that present in the person from which the transmission occurred. Thus, cultural mutations may occur during transmission from colonized to susceptible hosts, the magnitude of which in general may depend on the nature of religion and the power of rituals (Henrich, 2009). In addition, a newly susceptible person might adopt a religion only after listening to various models, so that the adopted religion would represent some average over already colonized hosts. In our models, such effects could for example be implemented by adding a diffusion term \(D \partial^2 C/\partial x^2\), describing random steps in \(x\) space during the transmission processes. However, our simulations show that as long as the diffusion constant \(D\) is not very large, adding such a diffusion term does not affect the dynamics of the model qualitatively, which is why we have not incorporated diffusion into the description given here.

For colonized hosts with religious type \(x\) the per capita rate of loss is \(A(x)\) as described above, so that the total rate of loss is \(\int A(x)C(x) dx\). For simplicity, we assume \(r_S=r_C=r\) and \(K_S=K=r\) in the sequel. We note that with the above assumptions, the religious trait \(x\) only affects rates of loss and transmission, but it does not affect the birth and death rates of colonized hosts, which are assumed non-evolving (constant) on the timescale relevant to the diversification of culture. Thus, selection on religion is not mediated by differential viability and/or reproductive success in the host population, but instead by differential loss and gain of ideas by colonized and susceptible hosts which in its turn depends on the density of these hosts.

3. Results

The dynamical system given by Eqs. (5) and (6) is in general analytically intractable but can always be solved numerically. Such simulations reveal two basic dynamic regimes. In the first one, all colonized host are concentrated in a narrow vicinity of the maximum of the transmission rate \(\tau(x)\). In this state, religious variation is controlled only by mutation, i.e. random deviations of the hosts from the optimal religion type, as illustrated in Fig. 1a. In the second regime, frequency-dependent selection on religious types leads to the maintenance of religious variation. Maintenance of variation in turn occurs in two different ways. At equilibrium, the distribution of colonized hosts, \(C(x)\), is either a unimodal function with a large positive variance (much larger than expected from mutation alone), as shown in Fig. 1b, or the equilibrium distribution is multimodal, as shown in Fig. 1c. Multimodal pattern formation as shown in Fig. 1c corresponds to the emergence of different religions, and hence to religious diversification.

In fact, even if religious diversity ultimately manifests itself in a unimodal distribution as in Fig. 1b, starting from homogenous populations essentially containing only one type of religion, this equilibrium distribution is reached through a series of "bifurcations" into distinct religious strains, as can be seen in Fig. 1b. Over time, different strains give rise to new strains, a process which eventually fills in the religious space and results in a unimodal equilibrium distribution. This can be seen more clearly using the individual-based models introduced below (see Fig. S1b).

It is worth noting that the case of Gaussian functions \(\tau(x)\) and \(\sigma(x-y)\) shown in Fig. 1b is special in the sense that for these functions, it is possible to find an analytical expression for the equilibrium distribution of colonized host. Specifically, it is not hard to see that the equilibrium distribution \(C(x)\) must satisfy the equation \(\int \sigma(x-y)C(y) dy = \tau(x)S\), and if both \(\tau(x)\) and \(\sigma(x)\) are Gaussian, this equation has a Gaussian solution \(C(x)\) with width \(\sigma = \sqrt{\sigma_x^2 - \sigma_y^2}\). This is the width of the equilibrium shown in Fig. 1b. However, the existence of such an equilibrium is a special property of the "Gaussian" case, and finding analytical expressions for equilibrium distributions in cases where the exponents \(b_1\) and \(b_2\) appearing in the functions \(\tau(x)\) and \(\sigma(x-y)\) are not equal to 2 is in general impossible. In particular, it is in general not true that such equilibrium distributions are unimodal, as the example in Fig. 1c shows.

Whether diversification occurs, and whether diversification manifests itself in unimodal or multimodal distributions, depends on the parameters of the model. First of all, diversification occurs when \(\sigma_x\) is small enough compared to \(\sigma_y\). This is revealed by numerical simulations, and in Appendix A we will use the framework of adaptive dynamics to provide some analytical justification for this threshold. Because \(\sigma_x\) is a measure for how fast religious types can gain an advantage by being different from common types, and \(\sigma_y\) measures how fast transmissibility decreases with increasing distance from the optimum \(x_0\), this can be roughly interpreted as diversification in religion occurring if the advantage gained from rarity outweighs the disadvantage due to having lower transmissibility.

Second, whether diversification, if it occurs, results in unimodal or multimodal equilibrium distributions depends on the exponents \(b_1\) and \(b_2\), i.e., on the nature of the functions \(\sigma(x-y)\) and \(\tau(x)\). Generally speaking, multimodal distributions, and hence
diversification into multiple distinct religious strains, require larger exponents in these functions. For example, in Fig. 1b showing unimodal diversification, these exponents were set to 2, i.e., both functions $\alpha(x-y)$ and $\tau(x)$ were of Gaussian form. As mentioned, in this case the dynamical system given by Eqs. (5) and (6) has an equilibrium distribution of colonized hosts that is itself Gaussian, and hence unimodal. However, increasing these exponents to $b_2 = 3$ and $b_1 = 3$, as in Fig. 1c, results in multimodal equilibrium distributions. Thus, “platykurtic” functions $\alpha(x-y)$ and $\tau(x)$ that fall off less sharply from their maximum tend to favour multimodal diversification. This is in agreement with results from models of adaptive diversification in organismic evolution (Doebeli et al., 2007; Leimar et al., 2008; Pigolotti, 2010).

Some of the above arguments can be made more precise using the framework of adaptive dynamics (Dieckmann and Law, 1996; Geritz et al., 1998; Metz et al., 1996) to describe the cultural evolutionary dynamics resulting from the epidemiological assumptions described in the Model setup. Adaptive dynamics has proven to be a very useful tool for identifying various scenarios of evolutionary diversification and speciation in organismal biology (e.g. Dieckmann and Doebeli, 1999; Dieckmann et al., 2004), and Appendix A describes how this theory can be applied in the present context. In contrast to the often analytically intractable partial differential equation models used above, adaptive dynamics theory allows one to derive certain analytical results about the conditions that lead to cultural diversification. For example, one can easily prove that if the two exponents $b_2$ and $b_1$ occurring in the functions $\alpha(x,y)$ (Eq. (4)) and in the function $\tau(x)$ (Eq. (6)) are equal to 2, then diversification occurs whenever $\sigma_2 < \sigma_1$ (see Appendix A).

4. Discussion

The purpose of this paper is to translate verbal narratives of cultural epidemiology into quantitative epidemiological modeling of cultural evolution, and to illustrate the potential usefulness of such models for understanding processes of cultural diversification that are due to frequency-dependent selection on cultural content, rather than based on geographical isolation or other diversifying effects. Naturally, cultural evolution is an extremely complicated and multifaceted process, where diversification, shifts and mergers (Hochberg, 2004) between cultures are the product of cooperation and competition between a multitude of forces driving and inhibiting diversity. These forces can be very indirect, such as religion-induced xenophobia that reduces the spread of epidemics and thus reinforced isolation and diversification (Fincher and Thornhill, 2008). Our goal in this paper is not to investigate all possible reasons for the diversification of systems of ideas, schools of thought, and religions, but to suggest a logically appealing and observationally relevant description of a so far underexplored, but potentially important driving force behind such diversification.

Thus, as any complex phenomenon, in reality the diversity of systems of ideas cannot be explained by a single mechanism or process, and here we suggest one particular candidate for such a force and provide simple quantitative arguments for its feasibility as a mechanism for cultural diversification. The mechanism we consider is negative frequency-dependent selection, or “overcrowding”, reflecting an inherent tendency to dissent as ideologies become more dominant, rigid, and repressive.
New Guinea, where Whitehouse (1995) described the coexistence of different cultures. In Papua, but reunited very recently. The sequential branching and later speciation events outside of Russia, which were divided by the 1917 revolution. The Russian Orthodox Church and the Russian Orthodox Church peaked in the 19th century, may be a good example. Some of the sects that emerge from the fragmentation of the Protestant church, which underwent intense religious conflicts and schisms, can be seen as evolutionary branches. The initial branches often further diversify, leading to the proliferation of Protestant denominations in the United States. Thus, this type of diversification occurs in situ from a single ancestral religion. The historic record contains many examples of both types of diversification occurring in our models: emergence of partially overlapping sects that differ from each other, for example, in the details of the interpretation of holy texts (such as the proliferation of Protestant denominations in the United States), and major splits that lead to the emergence of separate religions, such as between the Catholic, Protestant, and Eastern Orthodox churches. The initial branches often further diversify, for which the fragmentation of the Protestant church, which peaked in the 19th century, may be a good example. Some of the branches may later merge again, such as in the reintegration of the Russian Orthodox Church and the Russian Orthodox Church. Outside of Russia, which were divided by the 1917 revolution but reunited very recently. The sequential branching and later reunification can be observed in the behaviour of our model (e.g. Fig. 1b). A more recent example of the type of cultural evolutionary branching modeled here may be occurring in Papua New Guinea, where Whitehouse (1995) described the coexistence of a mainstream religious cult with periodically emerging sectarian splinter groups. We think that it might be interesting and fruitful to study the driving force for the emergence of such "modes of religiosity" (Whitehouse, 1995) based on the perspectives of frequency-dependent selection and evolutionary branching in religious content.

Even though we have used religion as a guiding example for our modeling approach, it should be pointed out that our models may serve more generally as a conceptual basis for understanding general processes of cultural diversification that are due to frequency-dependent non-conformism, e.g. in politics or in the evolution of language, as well as in more mundane areas, such as the evolution of fashion trends. Thus we purposefully avoided any reference to supernatural concepts (Atran, 2002; Boyer, 2001), intrinsic to religions, but usually irrelevant in more general secular cultural contexts. Evidently, the inclusion of such religion-specific concepts could affect certain quantitative aspects of our model (such as transmission rates and the probability of belief loss, see Henrich, 2009 for an example), yet the underlying negative frequency dependence paradigm and qualitative conclusions would remain similar. Our models illustrate that diversification processes should be expected to operate whenever the likelihood of secession from a dominant culture increases with growth of dominance of the mainstream culture. Intuitively, it is not hard to imagine that the attractiveness of a culture diminishes as the culture becomes more dominant, dogmatic, and perhaps oppressive, and that the desire to stand out and be different increases in more conformist cultures.

In this paper we have deliberately avoided the term "meme" (Dawkins, 1976) for denoting a unit of cultural or religious content, because in the literature, this term is rather loaded and its use subject to considerable controversy. For example, the use of memes in cultural evolution is sometimes criticized because memes are supposedly discrete entities that replicate faithfully. In contrast, in the present context, memes of religious content would be continuously varying rather than discrete entities, with different memes given by different real numbers. Also, such quantitative memes are not necessarily replicated faithfully, since the copy y of meme x that is transmitted from one human host to another might have type y ≠ x. In fact, the assumptions for transmission in our model are very simplistic, and one could envision much more complicated transmission schemes, in which for example the meme colonizing a new host is an average of a sample of memes from already colonized hosts, and in which the acquisition of new memes by susceptible hosts is biased by cognitive processes. In traditional evolutionary terms, the issue of whether memes are categorical and discrete or quantitative is analogous to whether one studies evolutionary models with discrete or continuous trait values, and of course, depending on the problem studied, either approach is feasible. Similarly, the issue of the modes of cultural transmission would, in classical evolutionary biology, roughly correspond to questions about the modes of reproduction and inheritance, and in principle, evolutionary models are not restricted to certain classes of reproduction and inheritance. For example, if cultural reproduction, i.e., transmission, occurs through an averaging process, this would correspond to a process of blending inheritance (Cavalli-Sforza and Feldman, 1981). Such cultural mixing could in principle impede the emergence of new cultural strains (similar to when random mating impedes speciation in organismic evolution). In this situation, cognitive constraints, such as an inability to learn from averaging of distinctly different models, could lead to cultural assimilations during transmission and thereby promote diversification (similar to when assortative mating promotes speciation in organismic evolution). It would seem of interest to pursue such extensions in future work. Overall, we think that our models for cultural evolution could well be formulated in terms of a generalized theory of memes, in which memes are quantitative, and in which transmission of memes is not necessarily faithful and may involve cognitive as well as selective biases. This view seems to be in agreement with a recent paper by Henrich et al., 2008, who tried to clarify a number of misconceptions about the theory of cultural evolution.

We think that our paper makes two potentially novel contributions. First, it uses explicit epidemiological models to describe the dynamics of cultural evolution, and second it studies the process of cultural diversification from a single ancestral religion in well-mixed, i.e., geographically unstructured populations. Sets of cultural ideas, such as languages, fashions, and ideological categories, clearly exhibit reproduction and heredity through their transmission between human hosts. Of course, these ideas ultimately need their human bearers for survival and reproduction (for example, a book's content only comes "alive" once the book is read). Just as the survival and reproduction of symbionts and pathogens is tied
to their effects on their hosts, the evolutionary fate of cultural ideas is tied to their impact on human individuals. And just as viewing individual organisms as hosts of evolving symbionts or pathogens offers the appropriate perspective for studying the evolution of those symbionts and pathogens, viewing human individuals as hosts of evolving cultural content offers a useful perspective.

Formal epidemiological models for cultural evolution do not seem to have been used very much in the past, even though the feasibility of the approach was already pointed out in Cavalli-Sforza and Feldman (1981) and Dietz (1967). An interesting exception is Lafferty et al. (2008), who use explicit epidemiological models to analyze the spread of (terrorist) ideologies. Just like ours, their models are very general and conceptual, but they are also potentially fruitful, because framing the spread of ideology in an epidemiological context allows one to identify the conditions that are most favourable to such spread, and the measures that would allow one to control the spread of detrimental ideologies, e.g. by identifying the various factors that impinge on the basic reproductive number \( R_0 \) of an infectious ideology.

Epidemiological terminology and narratives have been used in cultural contexts by well-known scholars such as Richard Dawkins and Daniel Dennett. Indeed, the fact that such prominent figures use epidemiological metaphors for culture seems to call for a more quantitative investigation of cultural epidemiology. In our opinion, one advantage of using this modeling framework is that it shifts the perspective from considering the adaptive value of culture for human hosts to considering the adaptive value of cultural traits for the cultural meme itself. This perspective is common when studying the epidemiology of microbial pathogens or symbionts, where an evolutionary understanding of microbial traits is primarily based on their adaptive value for the microbes (which is often mediated through their effects on the human hosts). A similar perspective might be useful for studying cultural epidemiology. This is not to say that other approaches cannot be put to very good use when studying the adaptive value of cultural traits for the cultural meme itself. For example, Boyd and Richerson (1985) have already pointed out, using models adopted from population genetics, that cultural traits can increase in frequency even if they are maladaptive for their human hosts. Nevertheless, we think that there is potential in the formal epidemiological approach for making progress in understanding cultural evolution based on adaptation in cultural memes.

Regarding processes of cultural diversification, a more traditional approach consists of viewing different human populations as carrying different cultural ideas, and of investigating competition between such human populations. In the language of host–pathogen models, this would correspond to considering different host populations carrying different pathogens and asking which of the host populations can outcompete the other. Because in this perspective success is based on characteristics imparted or imposed by the pathogen on a group of hosts, this perspective is akin to group selection. In contrast, studying the evolution of pathogens or the evolution of cultural content in a single host population is based on individual selection on the pathogens or on the cultural content. In our model, this difference to traditional approaches is reflected in the fact that the religious trait \( x \) does not affect survival and reproduction in the host. Thus, selection on culture is not mediated by differential viability or reproductive success in the host population, but by the fact that different religious ideas have different rates of being transmitted to susceptible hosts, and different rates of being lost from colonized hosts (Henrich, 2009). Loss of belief is frequency-dependent, because the rate of loss depends on overcrowding, which is the driving force of diversification. Boyd and Richerson (1985) have already pointed out that frequency-dependent selection can maintain cultural polymorphisms. However, these earlier treatments did not describe the gradual emergence of a diversified set of coexisting cultural lineages out of a single ancestral lineage, as exemplified by cultural evolutionary branching.

Just as in host–pathogen or host–symbiont systems, coevolution between humans and culture may be very important, and one can easily envisage many extensions of the model presented here to more complicated scenarios, in which the effects of culture on individual hosts as well as on the demographics of entire host populations are described in more mechanistic detail (e.g. in terms of propensity of host reproduction, sacrifice, and susceptibility to diseases, Fincher and Thornhill, 2008), and in which genetic evolution in the host occurs as a response to constraints imposed by cultural content, which in turn changes cultural opportunities. We believe that the perspective of cultural ideas as the evolutionary units will be very useful for such studies. Cultural content is best viewed not as fixed and pre-existing, but as evolving due to its effect on human individuals, who ultimately decide whether to accept or reject such content and how vigorously to spread it upon acceptance.

Acknowledgments

M.D. gratefully acknowledges the support of NSERC (CANADA) and of the Human Frontier Science Program.

Appendix A

To augment the analysis based on partial differential equation described in the main text, here we use the mathematical framework of adaptive dynamics (Dieckmann and Law, 1996; Geritz et al., 1998; Metz et al., 1996) to describe the evolutionary dynamics of culture based on the epidemiological assumptions described in the Model setup section of the main text. In the adaptive dynamics framework, one considers monomorphic resident populations consisting of a single religious type, and then investigates the fate of rare mutant types that appear in the resident population, e.g. because one of the hosts colonized by the resident religious type has slightly changed their belief and is now host to a slightly altered “mutant” type.

To do this, we first have to consider the dynamics of monomorphic resident populations. If the population is monomorphic for type \( x \), the distribution \( C(z) \) is a delta function with total weight \( C(x) = C(x) \) (Eq. (2)). Eqs. (5) and (6) then become a system of two ordinary differential equations:

\[
\frac{dS}{dt} = rs\left(1 - \frac{S + C(x)}{K}\right) - \tau(x)SC(x) + 3_0C(x)C(x) \tag{7}
\]

\[
\frac{dC(x)}{dt} = rc(x)\left(1 - \frac{S + C(x)}{K}\right) + \tau(x)SC(x) - 2_0C(x)C(x) \tag{8}
\]

It is easy to see that this system has a unique equilibrium

\[
(S^*, C^*) = \left(\frac{K\tau}{3_0 + \tau(x)}, \frac{K\tau}{3_0 + \tau(x)}\right) \tag{9}
\]

at which both \( S^* > 0 \) and \( C^* > 0 \). Moreover, the Jacobian matrix of system (7), (8) at the equilibrium \((S^*, C^*)\) has two negative eigenvalues, and the equilibrium is globally stable in the sense that the system will converge to this equilibrium from any initial condition with both densities \( > 0 \).

Let us assume that the host population is colonized by a single resident type \( x \), and that the resident dynamics given by Eqs. (7) and (8) has settled at its equilibrium \((S^*, C^*)\). This equilibrium constitutes the environment for a rare mutant type \( y \) that appears in the host population. If the mutant is rare, its logistic growth
term is determined by the total resident density $S^* + C^*$, its transmissibility is $\tau(y)$, and its rate of loss of belief is $A(y) = x_0(2y - x)x$. Therefore, the growth of the population of hosts colonized by the mutant religion type $y$ is

$$\frac{dC(y)}{dt} = \underbrace{rC(y)\left(1 - \frac{S^* + C^*}{K}\right)}_{\text{transmissibility}} + \underbrace{\tau(y)S^*C(y) - x(y - x)C^*C(y)}_{\text{death, transmission and loss of belief}}. \quad (10)$$

The invasion fitness $f(x,y)$ of a rare mutant $y$ in the resident $x$ lies at the basis of adaptive dynamics analyses and is defined as the per capita growth rate of $y$-types, i.e., by the right-hand side of Eq. (10) divided by $C(y)$:

$$f(x,y) = r\left(1 - \frac{S^* + C^*}{K}\right) + \tau(y)S^* - x(y - x)C^* \quad (11)$$

According to general theory (Dieckmann and Law, 1996), the adaptive dynamics of the religious trait $x$ is then given by the selection gradient

$$D(x) = \left. \frac{df(x,y)}{dy} \right|_{y=x} = \tau(x)S^* \quad (12)$$

More precisely, the adaptive dynamics of the trait $x$ is

$$\frac{dx}{dt} = \mu D(x) \quad (13)$$

where $\mu$ is a quantity describing the rate at which resident types give rise to mutational variants.

The analysis of the evolutionary dynamics given by (13) proceeds in two steps. First one finds stable equilibria of the dynamical system (13), and then one checks the evolutionary stability of these equilibria, as follows. Equilibria of (13) are points $x^*$ in trait space satisfying $D(x^*) = 0$. In the present case, i.e., with $D(x)$ given by (12), there is only one such point: the maximum of the function $\tau(x)$, $x^* = x_0$. Dynamic stability of this so-called singular point is determined by the derivative of $D(x)$ at the singular point, i.e., by $dD/dx(x_0)$. In the present case, this derivative is proportional to the second derivative of $\tau(x)$ at $x_0$, which is negative. Therefore, the singular point $x^* = x_0$ is a locally stable attractor for the dynamics (13), and it follows that starting from any initial resident value $x$, the religious trait will converge to the value $x_0$.

However, despite this convergence stability the singular point $x_0$ need not be evolutionarily stable. Evolutionary stability is determined by the shape of the invasion fitness function around the singular point. Note that by definition of the singular point as a solution of $D(x^*) = 0$, the first derivative of the invasion fitness function is necessarily 0 at a singular point. Thus, generically the invasion fitness function either has a maximum or a minimum at $x_0$. It is shown from adaptive dynamics theory (Geritz et al., 1998) that if $x_0$ is a fitness minimum, this generates the phenomenon of evolutionary branching. Once the resident is at $x_0$, every nearby mutant can invade. Moreover, two nearby mutants on either side of the singular value $x_0$ can coexist, leading to religion populations consisting of two coexisting strains. Finally, in each of these two strains selection favours trait values lying further away from the singular point, which means that the two strains will diverge evolutionarily. The phenomenon of convergence to a singular point that is a fitness minimum and subsequent emergence and divergence of coexisting strains is called evolutionary branching, and the singular point is called an evolutionary branching point. For example, if we assume that the two exponents $b_x$ and $b_z$ occurring in the functions $z(x,y)$ (Eq. (3)) and in the function $\tau(x)$ (Eq. (6)) are equal to 2, then one can show that the singular point $x_0$ is a fitness minimum if

$$\sigma_2 < \sigma_1$$

In particular, evolutionary branching occurs if $\sigma_2$ is small enough compared to $\sigma_1$.

A single bout of evolutionary branching leads to coexistence of diverging strains, and it is in principle possible to analyze the (two-dimensional) adaptive dynamics of these coexisting strains using invasion fitness functions. In a typical scenario, the two coexisting strains evolve to a new equilibrium (i.e., a new singular point in two-dimensional trait space), and this singular point may or may not be a branching point. If it is, further bouts of evolutionary diversification occur, resulting in the coexistence of more than two strains. If it is not, evolution comes to a halt in a diversified population consisting of two distinct and coexisting strains. Adaptive dynamics after diversification into two coexisting strains can in principle be studied analytically (in a similar way as above, see e.g. Dieckmann and Law, 1996), but it is also illustrative to study the evolutionary dynamics in individual-based models. In such models, the various terms on the right hand side of Eqs. (4) and (5) are interpreted as rates at which birth, death, transmission and loss of belief occur, resulting in a stochastic model for the evolutionary dynamics. The detailed setup of these models is described in Appendix B. Fig. S1 shows different scenarios of evolutionary branching occurring in the individual-based model.

---

**Fig. S1.** Evolution of the religious distribution $C(x)$ obtained from individual-based models. Parameter values are the same as in the corresponding panels of Fig. 1. (a) No diversification. (b) Continuous sequential branching in the Gaussian case results in an essentially unimodal distribution of religious types, corresponding to the unimodal equilibrium distribution in Fig. 1b. (c) Branching stops after two bouts and results in the coexistence of three distinct religious clusters. Note that in contrast to the deterministic model in Fig. 1c, which exhibits five clusters, the individual-based model only shows three clusters due to finite population size.
Appendix B

To construct individual-based stochastic models that correspond to the deterministic model given by Eqs. (4) and (5), we have to distinguish the different types of events that can occur at the level of individuals: birth, death, loss of belief, and transmission of belief. Each of these events occur at certain rates. For example, all host individuals have a per capita birth rate \( r_B = r_B \), so that the total birth rate of susceptible hosts, \( B_S = r_BS \), and the total birth rate of colonized hosts, \( B_C = r_C \), where \( S \) and \( C \) are the number of susceptible and colonized hosts present in the population at any given time (note that in contrast to Eqs. (4) and (5), where \( S \) and \( C \) are population densities, and hence real numbers, in the individual-based models \( S \) and \( C \) are integers). For both susceptible and colonized host individuals, the per capita death rate is \( r_D(S+C)/K_D = r_C(S+C)/K_C \), and total death rates \( D_S \) and \( D_C \) for susceptible and colonized individuals are \( r_S(S+C)/K_S \) and \( r_C(S+C)/K_C \), respectively. For a host colonized by religious type \( x \), the per capita rate at which this type is transmitted to susceptible hosts is \( P(tx)S \), where \( t(x) \) is the transmission function (6). The total rate of transmission, \( T \), is therefore \( \sum t(x)S \), where the sum runs over all colonized hosts. Finally, the per capita rate of loss of religion of host individuals colonized by religious type \( x \) is given by \( c(A(x)) = A(x) \) (Eq. (2)), so that the total rate of loss, \( L \), is \( \sum c(A(x)) \).

The individual-based model is implemented as follows. At any given time \( t \), all individual rates as well as the total rates \( B_S, B_C, D_S \), \( D_C, T \) and \( L \) are calculated as described above. Then the type of event that occurs next, birth or death of a susceptible or a colonized host, transmission of religious content, or loss of religious content, is chosen with probabilities proportional to the total rates for these events (i.e., with probabilities \( B_S/E \), etc., where \( E = B_S + B_C + D_S + D_C + T + L \) is the total event rate). For the event chosen, the individual to perform this event is chosen with probabilities proportional to the individual rates for the chosen event. For example, if loss of belief is the chosen event, individual \( i \) is chosen to lose belief with probability \( A(x)/L \), where \( x \) is the religious type of individual \( i \). This individual is then removed from the population of colonized hosts, and the number of susceptible hosts is augmented by 1. Similarly, if the chosen event is transmission, individual \( i \) among the colonized hosts is chosen for transmission with probability \( t(x)S/T \), where \( x \) is again the religious type of individual \( i \). Individuals for birth and death events are chosen analogously. If an individual dies it is removed from the population (and the numbers \( S \) or \( C \) are updated accordingly). If a susceptible individual gives birth the number \( S \) is augmented by 1. If an individual colonized with type \( x \) gives birth, a new colonized host is added to the population carrying a type \( x \) that is chosen from a normal distribution with mean the parental type \( x \) and a certain (small) width \( \sigma_x \). This reflects "mutation" during transmission of religious content from parent to offspring.

Performing one individual event in the manner described above completes one computational step in the individual-based model, which advances the system from time \( t \) to time \( t + \Delta t \) in real time. To make the translation from discrete computational steps to continuous real time, \( \Delta t \) is drawn from an exponential probability distribution with mean \( 1/E \), where \( E \) is the total event rate. Thus, if the total event rate \( E \) is high, the time lapse \( \Delta t \) between one event and the next is small, and vice versa if the total event rate is low. Starting from some initial population containing \( S_0 \) susceptible hosts and \( C_0 \) hosts colonized by religious types \( x_1, \ldots, x_n \), at time 0, iteration of the computational steps described above generates the stochastic cultural evolutionary dynamics of a finite population in continuous time.

References


